МОСКОВСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ
имени М. В. ЛОМОНОСОВА

ФЕДЕРАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ ЦЕНТР
«ИНФОРМАТИКА И УПРАВЛЕНИЕ»
РОССИЙСКОЙ АКАДЕМИИ НАУК

МОСКОВСКИЙ ФИЗИКО-ТЕХНИЧЕСКИЙ ИНСТИТУТ

ИНСТИТУТ ПРОБЛЕМ УПРАВЛЕНИЯ РАН

РОССИЙСКОЕ ОБЩЕСТВО ИССЛЕДОВАНИЯ ОПЕРАЦИЙ

РОССИЙСКОЕ НАУЧНОЕ ОБЩЕСТВО
ИССЛЕДОВАНИЯ ОПЕРАЦИЙ

IX Московская международная конференция по исследованию операций (ORM2018)

Москва, 22–27 октября 2018

ТРУДЫ

ТОМ I

IX Moscow International Conference on Operations Research (ORM2018)

Moscow, October 22-27, 2018

PROCEEDINGS

VOLUME I

МОСКВА – 2018

УДК 519.8
ББК 22.18

Издание осуществлено при поддержке
Российского фонда фундаментальных исследований по проекту 18-01-20080.

Ответственные редакторы:
профессор **А. А. Васин**, профессор **А. Ф. Измаилов**

IX Московская международная конференция по исследованию операций (ORM2018): Москва, 22–27 октября 2018 г.: Труды. Том I / Отв. ред. А. А. Васин, А. Ф. Измаилов. — М.: МАКС Пресс, 2018. — 373 с.

ISBN **!!!**

В сборнике представлены труды IX Московской международной конференции по исследованию операций. Конференция проводится факультетом вычислительной математики и кибернетики МГУ им. М.В. Ломоносова, Федеральным исследовательским центром «Информатика и управление» РАН (ФИЦ ИУ РАН), Московским физико-техническим институтом (МФТИ), Институтом проблем управления РАН (ИПУ РАН), Российским обществом исследования операций (РосОИО) и Российским научным обществом исследования операций (РНОИО). На конференции обсуждаются математические вопросы исследования операций, последние достижения в этой области, модели исследования операций в экономике, экологии, социологии, биологии, медицине, политологии, а также численные методы исследования операций.

Ключевые слова: исследование операций; математическое моделирование; методы оптимизации.

УДК 519.8
ББК 22.18

IX Moscow International Conference on Operations Research (ORM2018): Moscow, October 22–27, 2018: Proceedings: Vol. I — M.: MAKS Press, 2018. — 373 p.

This issue collects the proceedings of IX Moscow International Conference on Operations Research. The conference is organized by Lomonosov Moscow State University (MSU), Federal Research Center “Computer Science and Control” of RAS (FRC CSC RAS), Moscow Institute of Physics and Technology (MIPT), Institute of Control Sciences (ICS RAS), and Russian Operational Research Society (RuORS), Its main topics include mathematical problems of operations research, latest achievements in this field, new models in economics, ecology, sociology, biology, medicine, political science, etc., as well as numerical methods of operations research.

Keywords: operations research; mathematical modeling; optimization methods.

ISBN !!!

IX Московская международная конференция по исследованию операций (ORM2018)

Москва, 22–27 октября 2018

Конференция проводится МГУ имени М.В. Ломоносова, ФИЦ ИУ РАН, МФТИ, ИПУ РАН, Российским обществом исследования операций (РосОИО) и Российским научным обществом исследования операций (РНОИО), и посвящена столетию со дня рождения профессора Ю. Б. Гермейера. На конференции обсуждаются теоретические аспекты и различные приложения исследования операций.

Сопредседатели оргкомитета

А. А. Васин (МГУ), Ю. А. Флеров (ФИЦ ИУ)

Председатель программного комитета

И. Г. Поспелов

Программный комитет

Ф. Т. Алескеров, В. Н. Бурков, А. А. Васин, Г.-В. Вебер, В. А. Горелик, В. А. Гурвич, Ю. Г. Евтушенко, А. Ф. Измаилов, В. В. Мазалов, Н. М. Новикова, Ю. Р. Павловский, Г. И. Савин, А. Фишер, А. А. Шананин, М. Ячимович

Оргкомитет

Ф. И. Ерешко (зам. председателя), Ф. Т. Алескеров, Л. Г. Афанасьева, А. А. Белолипецкий, Н. В. Белотелов, Е. В. Булинская, В. Н. Бурков, Н. К. Бурова, А. В. Вахранев, А. В. Гасников, Д. Ю. Голембиовский, В. А. Горелик, М. А. Горелов, Д. В. Денисов, А. Г. Дивцова, И. А. Зонн, А. Ф. Измаилов, Ф. В. Костюк, Н. С. Кукушкин, А. В. Лотов, Ю. Е. Малашенко, И. С. Меньшиков, Е. И. Моисеев, В. В. Морозов, Е. З. Мохонько, Н. М. Новикова, В. В. Подиновский, И. Г. Поспелов, И. И. Поспелова, И. А. Соколов, М. Г. Фуругян, В. В. Шевченко, Е. Б. Яровая

Секретари конференции

Ю.В. Гусева, З.В. Мешина

При поддержке РФФИ (проект 18-01-20080).

IX Moscow International Conference on Operations Research (ORM2018)

Moscow, October 22-27, 2018

The conference is organized by Lomonosov Moscow State University (MSU), Federal Research Center “Computer Science and Control” of RAS (FRC CSC RAS), Moscow Institute of Physics and Technology (MIPT), Institute of Control Sciences (ICS RAS), and Russian Operational Research Society (RuORS), and is dedicated to 100th anniversary of Professor Ju. B. Germeyer. The conference brings together scientists from all over the world to discuss theoretical aspects and various applications of operations research.

Co-chairs of Organizing Committee
Yu. A. Flerov (FRC CSC RAS) and A. A. Vasin (MSU)

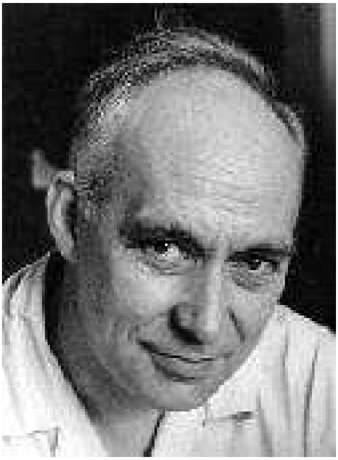
Chair of Program Committee
I. G. Pospelov

Program Committee
F. T. Aleskerov, V. N. Burkov, Yu. G. Evtushenko, A. Fischer, V. A. Gorelik, V. A. Gurvich, A. F. Izmailov, M. Jacimovic, V. V. Mazalov, N. M. Novikova, Yu. N. Pavlovsky, G. I. Savin, A. A. Shananin, I. A. Sokolov, A. A. Vasin, G.-W. Weber

Organizing Committee
F. I. Ereshko (vice-chair), F. T. Aleskerov, L. G. Afanasyeva, A. A. Belolipetskiy, N. V.Belotelov, V. N. Burkov, N. K. Burova, E. V. Bulinskaya, D. V. Denisov, A. G. Divtsova, M. G. Furugyan, A. V. Gasnikov, D. Yu. Golembiovskiy, V. A. Gorelik, M. A. Gorelov, F. V. Kostjuk, A. F. Izmailov, N. S. Kukushkin, A. V. Lotov, Yu. E. Malashenko, I. S. Menshikov, E. I. Moiseev, E. Z. Mokhonko, V. V. Morozov, N. M. Novikova, V. V. Podinovskiy, I. G. Pospelov, I. I. Pospelova, V. V. Shevchenko, I. A. Sokolov, A. V. Vakhranev, E. B. Yarovaya, I. A. Zonn

Conference Secretaries
Julija Guseva and Zlata Meshina

Supported by RFBR (project 18-01-20080).



Yuriy Borisovich Germeyer (18.07.1918–24.06.1975) was born in Atkarsk town, in Saratovskiy region of Russia. In July of 1941 he graduated from Faculty of Mechanics and Mathematics of Lomonosov Moscow State University (MSU), and was sent to the aviation plant 490 in Stalingrad for a position of an engineer-calculator. For a long time, he was concerned with developing of new military aircrafts and their armament. Till the end of the Great Patriotic War he worked at several plants of the Ministry of aviation industry. At the same time he completed the aspirant (PhD) program at MSU. He got his candidate degree in 1947, and doctor of sciences in physics and mathematics degree in 1963. His doctoral thesis examined some optimization problems and stochastic processes related to evaluation of the aircraft systems efficiency. Since 1966, and until the end of his life, Yu. B. Germeyer has been working in Computing Center of the Academy of Sciences of the USSR. In 1974 he organized a laboratory of Operations Research there. He was also a founder of Operations Research department at Faculty of Computational Mathematics and Cybernetics of MSU in 1970. Since that time, he has also been a head of this department.

Yu. B. Germeyer made an outstanding contribution to the development of Operations Research. He formulated the principle of the maximal guaranteed result for decision making under random and uncertain factors, introduced the concept of hierarchical games and proposed efficient methods for computation of their solutions. His books "Introduction to Operations Research" (1971) and "Non-antagonistic games"(1976) remain basic textbooks for students at Lomonosov MSU and at Moscow Institute of Physics and Technology.

Contents

Multiple objective decision making **103**

Improper optimization problems **131**

OR in economics 162

Markets and auctions analysis and design 173

OR in military and computer-aided design 190

OR in finance and banking 195

OR in insurance and risk-management **213**

OR in biology, medicine, physics and ecology **226**

Game-theoretic models **239**

The typical models for congested traffic **264**

Discrete optimization problems

Optimization of syscall sequences using minimal spanning trees search

N.N. Efanov and E.S. Shtypa
Moscow Institute of Physics and Technology,
Dolgoprudnyy, Moscow Region, Russian Federation

1. Introduction

Modern computer systems are much high loaded by complex tasks. Thereby, suspending and resuming the states of executional instances is strictly important. The approach which leads to system overhead minimization and reconstruction in a natural way is process-tree reconstruction by sequences of system calls operations [1–2]. Unfortunately, large combinatorial complexity of suitable tree generation does not allow to build naive direct syscall-based reconstructors [2]. Alternatively, the formal grammar-based solution via derivation of covering syscall sequences from some process-tree is proposed [1]. The main advantages of a formal language model for restoring the OS process-tree are:

- Unification and simplification of the process of developing systems of high-loaded and distributed computing, virtualization and checkpoint-restore
- Elimination of architectural flaws in checkpoint recovery systems, increasing their productivity
- Increasing the fault tolerance and security of checkpoint-restore systems and operating systems

- Unification of checkpoint-restore support tools
- Reduction of overhead and downtime for live migration

Nevertheless, formal grammar approach has a list of drawbacks. First of all, it is shortening: processes can terminate by exit() syscall, and initial process-tree reconstruction is possibly-feasible by reverse-reparenting with a set of heuristics, strict formal recovery is impossible in common case because of equivalence of such Chomsky type-0 grammar to universal Turing machine [3]. Another feature of existing syscall grammar [1–2] is ambiguous by design: overwhelming majority of process-trees have more than one derivation tree, so there are more than one syscall sequence to construct the certain state of some process. For example, a child process can be created by a single fork() call or by fork()-fork()-exit() sequence and so on. Different syscall sequences commits different overhead into system workflow because of different amount of context-switches and performing times. In this paper, the method of optimal syscall sequences obtaining is designed. Grammar and profiling approaches are also discussed, as the potential sources of syscall sequences, and graph-searching-based algorithm is proposed.

2. Problem Statement and Decomposition

Getting a certain process tree L and set of sources: derivation trees $\{D\}$ that derives L, profile traces $\{P\}$ of L construction and so on, construct build a directed acyclic graph (DAG) G, which describes all of available syscall sequences, then obtain acyclic oriented graph G', which contains the optimal sequences of syscalls for L construction, using graph G.

The problem reduction to some known and solved problems of graph theory is presented by the work in a set of statements:

Statement 1 *Graph G construction equals to tree into directed acyclic graph merge problem [4].*

Proof. Each instance from $\{D\}$ is a tree by design. Each instance from $\{P\}$ is received by profiling of some program which consists of single or multiple processes and can be represented by sequence of syscalls and states achieved by these syscalls, or set of such sequences, with some common subsequences from start state to the point where fork() syscall is executed, can be merged into P_tree on input in $O(N)$, where N is the number of process' states. Thus, $\{D\}$ and $\{P\}$ contains the trees which should be merged to construct G. ■

Statement 2 *The subproblem of G′ getting from G can be transformed into minimal spanning tree search problem.*

Proof. The method obtain syscall sequences with minimal overhead on system, so for each sequentially connected nodes $u, v \in G$ the most suitable branch from u to v should be chosen. If such branch is not unique, there are two ways to resolve ambiguity: a) strip the graph into a set of graphs, each of which contains only one branch from mentioned above, so each of those graphs is tree contains of minimal cost syscall sequences by design; b) introduce a rule of unique branch picking. For example, by alternative metric examining. Cases a) and b) determines minimal spanning trees by definition [6]. ■

Based on decomposition and statements given above, the proposed solution is reduced to well-known designed methods from graph theory and discrete mathematics. The details of such subproblems solution are given in Section 3.

3. Tree Merging Details and Minimal Spanning Trees Obtaining

The problem of trees merging is well-investigated in different applications of graph theory and computer science. The wide set of applications produces the set of different merge algorithms [4–5]. Most of those algorithms have $O(N \cdot M)$ time and space complexity, where N - the maximal number of nodes in any of M trees. Nevertheless, the authors suggest to use relatively simple and generalized approach, which also works in $O(N \cdot M)$: the attributes, which describes an each node position (number, label and so on) should be logged into a list by DFS[6], then the same labels should be connected by additional undirected edges E' in the procedure of duplicates checking, then all of trees should be united into the graph, with equal nodes deduplication by rule

$$\{(V_j, E_k)\}\Big|_{E_k(V_j)\to V_j} \longrightarrow \{(V_j)\}\,. \tag{1}$$

And redundant edges deduplication via

$$\{(V_j, E''_k)\}\Big|_{\left|E''_k(V_j)\to V_{j+1}\right|>1} \longrightarrow \{(V_j, E''_1)\}\,, \tag{2}$$

where E''_k is the set outbound edges with same label, e.g. syscall name for the given problem. All of operation requires at case, $4O(N \cdot M)$ – passes, so total complexity is $O(N \cdot M)$, and merged multi-tree structure is the output of algorithm if at least one node in each pair of trees is the same. This condition is accomplished by the source of problem and described in the Section 2. Thus, merging the trees of $\{D\}$ and $\{P\}$ sets leads to DAG G.

The subproblem of getting minimal spanning tree from G named minimal spanning arborescence requires to application of already designed methods of optimum directed branching search like Edmond's search [7]. It is proposed to use Tarjan modification of algorithm [8] because of relative sparsity of G: derivation trees and profiling traces not cover all of potential patches between the nodes by practical design [1], so the sparsity factor of G is near $\frac{1}{|V|}$, where $|V|$ is the power of nodes set. The Tarjan modification works in $O(|E|\log(|V|))$ on sparse DAGs with $|E|$ edges. Possible ambiguity in searching can be solved by branches picking by some strict rules provided in Section 2. The cost metrics, which are supported by proposed method, are edge-weight additive metrics presented in the Section 4. Thus, the total complexity of described in Section 2 problem is $O(N \cdot M + \log(N \cdot M)|E|)$.

4. Cost Metrics

Different syscalls impact different overhead into the program workflow because of various times of operations performing. Moreover, the

procedure of context switch from userspace to kernel code during syscall is relatively slow and potential security-unsafe itself [9]. Thus, cost metrics should consider two independent factors: whole number of context switches and average times of each syscall execution. To handle the factors above effectively, two cost metrics are presented: combined time-based (CT) and context switch metric (CS). The combined time-based metrics is defined as:

$$CT = \sum_{i=1}^{m} (n_i \cdot t_i), \tag{3}$$

where n_i is the number of syscalls type i from the set of types $1 \ldots m$, t_i is the average time of single syscall type i performing. The minimization in this metrics is intended to accounting for the full overhead from the system calls execution during runtime. Context switch metrics is defined as:

$$CS = \sum_{i=1}^{m} n_i, \tag{4}$$

where n_i is the number of syscalls type i from the set of types $1 \dots m$. The minimization in this metrics is intended to exclude redundant context-switches [9].

According to the optimization procedure and weights nature of CS, CT can be obtained from CS by setting all of times of kernel execution to 1.

5. Impact into the Formal Grammar of System Calls Analysis and Future Work

Optimal syscall sequences obtaining is constructed to improve each of execution environment technics: traces and profiling, heuristic-based methods and formal grammar derivation. Nevertheless, the impact to the grammar-based solution [2] is extremely significant, because of ambiguity of the designed grammar: it is potentially possible to implement improved grammar parser: such parser should dynamically exclude non-optimal branches from derivation tree during the main procedure of sequences and subtrees obtaining from an intermediate representation. According to the complexity estimations above, the parser still be polynomial, and the degree of polynomial is bigger in worst case on 2 than original. This complexity is still competitive with a set of grammars from
mild-context-sensitive formalisms, which also partially supports
shortening, like domain-based PMCFG [10] and so on. The further works of the author will be focused on the practical construction of such
conditional-analytical optimizing parser for syscalls formal grammar [1], simultaneously with the grammar expansion.

References

1. Efanov N.N., Emelyanov P.V. Constructing the formal grammar of system calls // In Proceedings of the 13th Central & Eastern European Software Engineering Conference in Russia (CEE-SECR'17). 2017. Article 12. 5 pages.
2. Efanov N.N., Emelyanov P.V. Postroenie formal'noj grammatiki sistemnyh vyzovov // Informacionnoe obespechenie matematicheskih modelej. 2017. P. 83–91.
3. Partee B.H., Ter Meulen A., Wall R.E. Turing Machines, Recursively Enumerable Languages and Type 0 Grammars // Mathematical Methods in Linguistics (Studies in Linguistics and Philosophy). 1993. V. 30. P. 505–525.
4. Mailund T. Merging Trees into a DAG, 2004.

5. Morozov D., Weber G. Distributed merge trees // In Proceedings of the 18th ACM SIGPLAN symposium on Principles and practice of parallel programming (PPoPP'13). 2013. P. 93–102.
6. Sedgewick R. Algorithms in C++ Part 5: Graph Algorithms, 3rd Edition // Pearson Education. 2002.
7. Edmonds J. Optimum Branchings // Journal of Research of the National Bureau of Standards - B. Mathematics and Mathematical Physics. 1967. V. 71B, no. 4. P. 233–240.
8. Tarjan R.E. Finding Optimum Branchings // Networks. 1977. V. 7. P. 25–35.
9. Hruby T., Crivat T., Bos H., Tanenbaum A. On sockets and system calls minimizing context switches for the socket API // In Proceedings of the 2014 International Conference on Timely Results in Operating Systems (TRIOS'14). 2014. P. 8.
10. Hiroyuki S., Nakanishi R., Kaji Y., Ando S., Kasami T. Parallel multiple context-free grammars, finite-state translation systems, and polynomial-time recognizable subclasses of lexical-functional grammars // In 31st Meeting of the Association for Computational Linguistics (ACL'93). 1993. P. 121–129.

Heavy-ball method in unconstrained minimization

F.V. Kostyuk

Dorodnicyn Computing Center of Federal Research Center Computer Science and Control of Russian Academy of Sciences, Moscow, Russia

Consider the unconstrained optimization problem

$$\Pi : \quad F(x) \to min, x \in R_k,$$

where $F(x)$ is assumed to be continuously differentiable on $R_k, i.e.F(x) \in C_1(R_k)$ and $\nabla F(\cdot) \in C_{Lip}(R_k, L)$ ($C_{Lip}(X, L)$ denotes the class of vector-valued functions which satisfy a Lipschitz condition on $X \in R_k$ with a constant$L > 0$). We also assume that the Lebesgue sets

$$X_L(f) = \{x \in R_k \mid F(x) \leq f\}, \quad f \in R_1,$$

are bounded for all $f \in R_1$. Obviously under these assumptions the optimal set

$$X_{opt} = \left\{x \in R_k \mid F(x) = min_{x' \in R_k} F(x') = f_{opt}\right\}$$

is not empty and compact.

The stationary set X_{stat} of the problem Π is defined by

$$X_{stat} = \{x \in R_k \mid \nabla F(x) = 0\},$$

clearly X_{stat} is the set of all points satisfying the first order necessary optimality condition for Π. Assume that X_{stat} is bounded.

Let us consider the following optimization algorithm which is known as "the heavy ball method"[1]:

$$HBM: \qquad x^{n+1} = x^n - \alpha\nabla F(x^n) + \beta(x^n - x^{n-1}),$$

$$n = 1, 2, \ldots, \qquad (x^0, x^1) \in R_k * R_k,$$

where $\alpha > 0, \beta \geq 0$ are parameters (step sizes); (x^0, x^1) is a pair of starting points.

The method owes its name to the following physical analogy. The motion of a body ("the heavy ball") in a potential field under the force of friction (or viscosity) is described by a second-order differential equation

$$\theta\frac{d^2x(t)}{dt^2} = -\nabla F(x(t)) - \kappa\frac{dx(t)}{dt} \qquad (1.1)$$

The body eventually reaches a local minimum point of the potential $F(\cdot)$ because of energy loss caused by viscosity. Thus, the heavy ball method "solves"the corresponding minimization problem. Considering the difference analogue of the equation (1.1), getting the iterative method HBM. Note, for $\beta = 0$, HBM turns into the simplest algorithm of the gradient descent method:

$$GrM: \qquad x^{n+1} = x^n - a\nabla F(x^n), \quad n = 1, 2, \ldots, x^1 \in R_k.$$

In the case $\beta < 0$ HBM can be considered as a gradient algorithm of the prox-method [2].

There are two main reasons why the significant attention is being drawn to the heavy ball technique:

Firstly, standard gradient descent technique possesses a wide range of fairly good properties, i.e. the global convergence to the optimal set or, in the multiextremal case, to the stationary set X_{stat} and robustness, i.e. computational errors do not essentially affect its convergence. It works under unrestricted requirements on the objective function's smoothness, unlike the conjugate gradient and the quasi-Newton techniques which need the twice differentiability of $F(\cdot)$ etc. But its convergence rate to ill-posed solutions of the problem Π is too low. Thus, the first reason for developing the heavy ball technique is the need in simply constructed algorithms which solve optimization problems with once differentiable data more effectively than the standard gradient descent technique. In particular, Nesterov [3] suggested an original algorithm which construction includes HBM's steps and proved that its convergence rate is $O(\frac{1}{n^2})$ on the class of convex problems with C^1-data, and what is important this convergence rate on the class of convex differentiable problems cay not be essentially improved be any iterative algorithm as was shown by Nemirovski and Yudin [4].

Secondly, HBM's trajectories get the property to cross neighborhoods of nondistinct slantish local minima like the physical analogy of the method thanks to kinetic energy of the moving ball. Thus, in multiextremal problems the attraction area of the global minima (i.e. the set of starting points of trajectories converging to the global minima) of HBM is wider than such area for the standard gradient descent technique. Furthermore, in the present paper it is shown that special algorithms of HBM possess the property to cross neighborhoods of nondeep (but may be sharp) local minima. These properties are important for using HBM within the multistart method for the search of a global minimum as was shown by Rinnooy Kan [5]. Thus, the second reason for developing the heavy ball method is attractive global optimization properties of this local descent technique.

Stability properties of the continuous process (1.1) were analyzed by Zirilli, Parisi and Alluffi-Pentini [6].

Local convergence of HBM at a neighborhood of a nonsingular solution x^* of Π was explored in [7]. It was assumed that $F(\cdot)$ is twice continuously differentiable on a neighborhood of $x^*, \nabla F(x^*) = 0$, and the eigenvalues $\lambda_1, \ldots, \lambda_k$ of $\nabla^2 F(x^*)$ are as follows:

$$0 < \lambda_1 \leq \cdots \leq \lambda_k$$

The condition μ of x^* is defined by

$$\mu := \frac{\lambda_k}{\lambda_1}$$

.

It was obtained that for any (a, b) satisfying

$$(a, b) \in \Omega_P = \left\{ (a', b') \mid 0 \leq b' < 1, 0 < a' < \frac{2(1 + b')}{L} \right\} \tag{1.2}$$

there exist real $\epsilon, q > 0, q = q(a,b) < 1$, such that every trajectory $\{x^n\}$ of HBM with $x^0, x^1) \in B_\epsilon(x^*) \times B_\epsilon(x^*)$ has the property

$$\lim_{n\to\infty} \|x^n - x^*\| = 0,$$

and, moreover,

$$\|x^n - x^*\| \leq C(q(a,b)+\delta)^n, \quad n = 1, 2, \ldots,$$

for any $\delta > 0$ and the corresponding $C = C(\epsilon, \delta) > 0$. Furthermore,

$$\min_{(a,b)\in\Omega_P} q(a,b) = \frac{\sqrt{\mu}-1}{\sqrt{\mu}+1} < q_* = \frac{\mu-1}{\mu+1}$$

where q_* is the best possible characteristic of the linear convergence rate to x^* of the gradient descent method GrM (i.e. HBM with $b = 0$).

Thus, when Π is ill-posed at x^* (i.e. $\mu \gg 1$) the convergence rate of HBM with an appropriate $(a_0, b_0) \in \Omega_P$ is much better than the convergence rate of the standard gradient descent technique. However, the condition (1.2) does not guarantee the global convergence of HBM.

In the case when $F(\dot{})$ is differentiable for the global convergence of HBM was proved that for any (a,b), satisfying

$$(a,b) \in \Omega_A = \left\{(a',b') \mid 0 \leq b' < \frac{1}{3}, 0 < a' < \frac{2(1-3b')}{L}\right\} \qquad (1.2)$$

every trajectory $\{x^n\}$ of HBM converges to X_{opt}. Note that $\Omega_A \subset \Omega_P$.

The present report concerns the stability analysis of HBM at multiextremal problems. We shall investigate global convergence properties of HBM and then construct global optimization procedures based on this aproach.

Consider an iterative process

$$x_{n+1} \in \Psi(n, x^n, x^{n-1}, x^{n-2}, \ldots, x^{n-s+1}), \quad (x^1, \ldots, x^s) \in X^1 \times \cdots X^s,$$

$$n = s, s+1, \ldots,$$

where $\Psi(\cdot,\cdot) : N \times \otimes_{i=1}^s R_k \to \Re(R_k)$ (...$\Re(X)$ denotes the set of all subsets of $X, N = \{1, 2, \ldots\}$), $X^i \subset R_k, i = 1, \ldots, s$. Denote by X... the set of all trajectories of the process.

The process is said to be Lagrange stable if every trajectory $\{x^n\} \in X$... is bounded. We say that a bounded trajectory $\{x^n\}$ converges to a set A if

$$\overline{lt}\{x^n\} \subset A, \qquad (1.3)$$

where $\overline{lt}\{x^n\}$ denotes the set of all limit points of $\{x^n\}$. For a Lagrange stable process the set $\Xi := \bigcup_{\{x^n\}\in X_{\dots}} \overline{lt}\{x^n\}$ is called the attractor set of the process. Clearly that the attractor set is defined as the smallest set A with the property that every trajectory of the process attracts to A, i.e. (1.3) holds for every $\{x^n\} \in X_{\dots}$.

As a rule the global stability analysis of nonlinear optimization methods at multiextremal problems consist in the following:

i) to establish the Lagrange stability of the method;

ii) to prove the convergence of the method to the stationary set X_{stat}, i.e. to obtain the attractor set of the method;

iii) to clarify what kind of perturbations does not essentially affect on the method's convergence properties (it is important for the further developing of the method). Explicitly, it has considered the situation when the method deals at the n-th iteration with, for instance, $\nabla F(x^n) + p_1(n, x)$ instead of the exact value $\nabla F(x^n), n = 1, 2, \dots$, where $p_1(\cdot, \cdot) : N \times R_k \to R_k$ is a perturbation of the objective function's gradient. Then the aim is to find condentions on $p_1(\cdot, \cdot)$ guaranteeing that the perturbed method still converges to the original attractor set of the method (i.e. to the attractor set of the unperturbed method) and, moreover, possibly preserves the original convergence rate. Certainly these conditions appear to be rather restricted and, in particular, imply that

$$\lim_{n\to\infty} \|p_1(n, x^n)\| = 0$$

on every trajectory $\{x^n\}$ of the perturbed method.

In the present paper all the problems i)-iii) of the global stability analysis are considered. It appears that this modification provides to HBM the property to smooth a surface under the moving ball which is very important for the application of the technique to global optimization.

References

1. Polyak B.T. Introduction to Optimization. New York: Optimization Software, Inc., Publication Division, 1987.
2. Polyak B.T., W.S. Petrov and L.M. Kravchukov. On the sufficient conditions in global optimization // Economics and Mathematical Methods. 1972. V. 8. P. 130–135.
3. Nesterov Yu.E., An 0(l/k)-rate of convergence method for smooth convex functions minimization // Dokl. Acad. Nauk SSSR. 1983. V. 269: P. 543–547, .

4. Nemirovsky A., Yudin D. Problem complexity and method efficiency in optimization. John Wiley and Sons, 1983
5. Boender, C.G.E., A.H.G. Rinnooy Kan; L. Strougie; G.T. Timmer. A stochastic method for global optimization // Mathematical Programming. 1982. V. 22. P. 125–140. doi:10.1007/BF01581033
6. Aluffi-Pentini F., Parisi V., Zirilli F. An inexact continous methid in complementarity problems //IFAC Control Applications of Nonlinear Programming and Optimization. Capri, Itali. 1985. P. 19–26
7. Polyak B.T. Some methods of speeding up the convergence of iteration methods // USSR Computational Mathematics and Mathematical Physics. 1964. V. 5, No 4, P. 1–17

Continuous optimization problems

Model of stabilization for inter-branch balance by Leontiev*

[1]A.S. Antipin, [2]E.V. Khoroshilova, [3]M. Jacimovic, and [3]N. Mijajlovic
[1]*Federal Research Center «Computer Science and Control» of RAS,*
[2]*Lomonosov Moscow State University, Moscow, Russia*
[3]*University of Montenegro, Podgorica, Montenegro*

Introduction. The inter-branch balance model is one of the most famous and popular models of economic and mathematical modeling. Thousands of pages in scientific texts are devoted to various versions of this model. There are static models, dynamic controlled models and dynamically uncontrolled models, as well as their various combinations. Models with controlled dynamics are especially popular in the theory of optimal control. The methodology of these models assumes that it is possible to compute in advance the program control and the corresponding program trajectory over the entire time interval, for example, for one year. In this case, the calculated trajectory is quite adequate to reflect the development in time of a real inter-branch balance. However, the practice of calculation shows that very often serious difficulties arise with an adequate description of the real economic process.

The paper considers another approach to mathematical modeling of this dynamic situation, namely, we assume that the control goal is a balanced inter-branch balance at the end of the planning period. In this

*This research is supported by the Russian Foundation for Basic Research (Project 18-01-00312)

case, the balanced state of the system at the end of the time interval will be considered the equilibrium state of the control object. However, often under the influence of various perturbing factors, the object loses its equilibrium and is far from the equilibrium point. In this case, the problem arises to select the control so that the object is returned to the equilibrium state. If, for example, program control is calculated in advance, then any such control can always be considered as a disturbed state of the control object. In this case, we get the stabilization problem to return the control object from the disturbed state again to the equilibrium state. The stabilization problem is one of the two main problems of the general control theory (another problem is the controllability problem).

Formulation of the problem. Consider a linear differential model of optimal control [1]–[3]. This model includes controlled dynamics and a boundary-value finite-dimensional problem of the inter-branch balance by Leontiev. When the control runs through the entire set of controls $u(\cdot) \in \mathbb{U}$, then the linear controlled dynamics of the problem generates trajectories $x(\cdot)$, $t \in [t_0, t_1]$, whose right-hand ends $x(t_1) = x_1$ describe the terminal set $X_1 = X(t_1) \subset \mathbb{R}^n$, called the reachability set. The control problem can be treated as a stabilization problem for the case when the control object needs to be transferred from an arbitrary initial state to an optimal terminal state:

$$x_1^* \in \text{Argmin}\left\{\frac{1}{2}\left|(A-E)x_1 - y\right|^2 \mid x_1 \in X_1 \subset \mathbb{R}^n\right\}, \tag{1}$$

$$\frac{d}{dt}x(t) = D(t)x(t) + B(t)u(t),\ x(t_0) = x_0,\ x(t_1) = x_1^*, \tag{2}$$

$$u(\cdot) \in \mathrm{U} \subset \mathbb{L}_2^r[t_0, t_1]. \tag{3}$$

The objective finite-dimensional function from (1) is generated by the interbranch balance model

$$x = Ax + y,$$

where $A = a_{ij},\ i = 1, ..., n;\ j = 1, ..., n$. They show the amount of the product of the i-th industry, which must be spent to produce a product unit of j-th industry. The i-th component of $x = (x_1, ..., x_n)$ is the gross output of j-th product; each component of $y = (y_1, ..., y_n)$ is the final consumption. This model has been the subject of intensive scientific research for almost 100 years [Lotov A.V. (1984)].

The trajectory of differential system (2) identically satisfies condition

$$x(t) = x(t_0) + \int_{t_0}^{t} (D(\tau)x(\tau) + B(\tau)u(\tau))d\tau. \tag{4}$$

for almost all $t \in [t_0, t_1]$ and is an absolutely continuous function. The existence of a solution of the dynamic problem (1)-(3) is well-known fact.

The formulated problem belongs to the class of optimal control problems, for which the maximum principle is the main tool for constructing iterative methods for solving problems. However, the maximum principle is a necessary condition for optimality. This condition does not guarantee that the limit point is the solution to the problem. In fact, this means that such an approach is not conclusive and justified, and all solutions obtained by this method require additional expert justification. But this is already a sphere of heuristics, and not a proof theory. To get out of this situation, it is necessary to use sufficient optimality conditions that ensure that the solution found is a true solution of the original problem.

In reality, there are several types of sufficient conditions for optimality. First of all, these are sufficient conditions based on the Hamilton-Jacobi inequalities [Krotov V.F. (1973), Dykhta V.A. (1995)], sufficient condi- tions generated by the field theory of extremals [Velichenko V.V. (1974)]. In our paper, we use sufficient conditions of the duality theory, that is, saddle-point sufficient conditions. On the basis of these conditions, we construct a saddle-point iterative methods of extragradient type. The duality theory assumes the convexity of the original problem. But this circumstance is not a rigid restriction, since any smooth problem can always be replaced by a sequence of convex problems.

Following the proposed strategy, we linearize the initial problem of terminal control (1)-(3) and replace it with the linear programming problem formulated in the functional space

$$x_1^* \in \text{Argmin}\{\langle \nabla\varphi(x_1^*), x_1 - x_1^* \rangle \mid \; x_1 \in X_1 \subset \mathbb{R}^n\}, \tag{5}$$

$$\frac{d}{dt}x(t) = D(t)x(t) + B(t)u(t), \; x(t_0) = x_0, \; x(t_1) = x_1^*, \tag{6}$$

$$u(t) \in \mathrm{U} \subset \mathbb{R}^n, u(\cdot) \in \mathbb{L}_2^r[t_0, t_1], \tag{7}$$

where $\varphi(x_1) = \frac{1}{2} \mid (A-E)x_1 - y \mid^2$, the gradient $\nabla\varphi(x_1) = (A-E)^T((A-E)x_1 - y)$, and the linearized objective function $\langle \nabla\varphi(x_1), x(t_1) - x_1 \rangle = \langle (A-E)^T \; ((A-E)x_1 - y), x(t_1) - x_1 \rangle$. Linearization is carried out at the point $x_1 = x_1^*$, which is the solution of the problem (1)-(3).

Lagrangians with respect to primal and dual variables. We write out the dual problem for system (1)-(3) in an explicit form. For the scalarized problem (5)-(7) we have the Lagrange function

$$L(\psi(\cdot); x_1, x(\cdot), u(\cdot)) =$$

$$= \langle \nabla\varphi(x_1^*), x_1 - x_1^* \rangle + \int_{t_0}^{t_1} \langle \psi(t), D(t)x(t) + B(t)u(t) - \frac{d}{dt}x(t) \rangle dt, \quad (8)$$

for all $(\psi(\cdot); x_1, x(\cdot), u(\cdot))$.

By definition, the saddle point $(\psi^*(\cdot); x^*(t_1), x^*(\cdot), u^*(\cdot))$ of the Lagrange function, formed by primal $(x^*(t_1), x^*(\cdot), u^*(\cdot))$ and dual $(\psi^*(\cdot))$ variables, satisfies saddle-point inequalities

$$\langle \nabla\varphi(x_1^*), x_1^* - x_1^* \rangle + \int_{t_0}^{t_1} \langle \psi(t), D(t)x^*(t) + B(t)u^*(t) - \frac{d}{dt}x^*(t) \rangle dt$$

$$\leq \langle \nabla\varphi(x_1^*), x_1^* - x_1^* \rangle + \int_{t_0}^{t_1} \langle \psi^*(t), D(t)x^*(t) + B(t)u^*(t) - \frac{d}{dt}x^*(t) \rangle dt$$

$$\leq \langle \nabla\varphi(x_1^*), x_1 - x_1^* \rangle + \int_{t_0}^{t_1} \langle \psi^*(t), D(t)x(t) + B(t)u(t) - \frac{d}{dt}x(t) \rangle dt \quad (9)$$

for all $(\psi(\cdot); x_1, x(\cdot), u(\cdot))$.

If the regularity conditions (Slater's condition) are satisfied, if the original problem (1)-(3) has a primal and dual solution, then they form a saddle point for the Lagrange function. The converse is also true: the components of the saddle point of (9) are primal and dual solutions to problem (1)-(3) and dual to it.

Using conjugate operators, we write out the conjugate Lagrangian

$$\mathcal{L}^T(x_1, \psi(\cdot), x(\cdot), u(\cdot)) = \langle \nabla\varphi(x_1^*) - \psi(t_1), x_1^* \rangle +$$

$$+ \int_{t_0}^{t_1} \langle D^T(t)\psi(t) + \frac{d}{dt}\psi(t), x(t) \rangle dt$$

$$+ \int_{t_0}^{t_1} \langle B^T(t)\psi(t), u(t) \rangle dt - \langle \psi(t_1), x(t_1) \rangle + \langle \psi(t_0), x(t_0) \rangle \quad (10)$$

for all $(\psi(\cdot); x_1, x(\cdot), u(\cdot))$. Both Lagrangians (primal and dual) have one and the same the saddle point $(\psi^*(\cdot); x_1^*, x^*(\cdot), u^*(\cdot))$, which satisfies the saddle-point dual system.

The saddle-point system generates mutually-dual problems, which in turn generate a differential system with respect to primal and dual variables:

$$\frac{d}{dt}x^*(t) = D(t)x^*(t) + B(t)u^*(t), \ \ x^*(t_0) = x_0,$$

$$D^T(t)\psi^*(t) + \frac{d}{dt}\psi^*(t) = 0, \ \ \nabla\varphi(x_1) - \psi^*(t_1) = 0,$$

$$\int_{t_0}^{t_1} \langle B^T(t)\psi^*(t), u^*(t) - u(t)\rangle dt \leq 0, \ u(\cdot) \in \mathrm{U}. \tag{11}$$

This system is a saddle-point sufficient optimality condition. This condition make possible to construct proving methods for solving very complex problems.

As a result, we obtain an iterative process with two half-steps on one iteration [1]–[4]:

1) predictive half-step

$$\frac{d}{dt}x^k(t) = D(t)x^k(t) + B(t)u^k(t), \ x^k(t_0) = x_0, \tag{12}$$

$$\frac{d}{dt}\psi^k(t) + D^T(t)\psi^k(t) = 0, \ \ \psi_1^k = \nabla\varphi(x_1^k), \tag{13}$$

$$\bar{u}^k(t) = \pi_U(u^k(t) - \alpha B^T(t)\psi^k(t)); \tag{14}$$

2) basic half-step

$$\frac{d}{dt}\bar{x}^k(t) = D(t)\bar{x}^k(t) + B(t)\bar{u}^k(t), \ \bar{x}^k(t_0) = x_0, \tag{15}$$

$$\frac{d}{dt}\bar{\psi}^k(t) + D^T(t)\bar{\psi}^k(t) = 0, \ \bar{\psi}_1^k = \nabla\varphi(\bar{x}_1^k), \tag{16}$$

$$u^{k+1}(t) = \pi_U(u^k(t) - \alpha B^T(t)\bar{\psi}^k(t)), \ k = 0, 1, 2, ... \tag{17}$$

Conclusion. The theorem on convergence of the proposed method (12)–(17) to solution of the problem was proved. In particular, it was shown that convergence in controls is weak, convergence in state and conjugate trajectories is strong.

References

1. Antipin A.S., Khoroshilova E.V. Controlled dynamic model with boundary-value problem of minimizing a sensitivity function // Optim. Lett. 2017. DO1 10.1007/s11590-017-1216-8. (Published online: 17 November 2017).

2. Antipin A.S., Khoroshilova E.V. Saddle point approach to solving problem of optimal control with fixed ends // J. of Global Optimiza- tion. 2016. V. 65, Issue 1, P. 3–17.
3. Antipin A.S., Jacimovic M. and Mijajlovic N. Extragradient method for solving quasivariational inequalities // Optimization. A Journal of Mathematical Programming and Operations Research. 2018. V. 67, Issue 1. P. 103–112.
4. Khoroshilova E.V. Extragradient method of optimal control with terminal constraints // Automation and Remote Control. 2012. V. 73. No 3. P. 517–531.

A refinement of the maximum principle for state constrained optimal control problems under a regularity condition*

A.V. Arutyunov, D.Yu. Karamzin, and F.L. Pereira

Peoples' Friendship University of Russia, Federal Research Center "Computer Science and Control" of the Russian Academy of Sciences, Moscow, and University of Porto, Portugal

In this short note, we refer to the classic monograph [1], see Chapter 6 therein, in which the state constrained optimal control problems have been investigated and corresponding optimality conditions derived. We notice that a certain refinement to this known result (originally obtained by R.V. Gamkrelidze in [2]) holds, namely, the fact that the measure Lagrange multiplier is continuous. Continuity of the multiplier follows in view of the regularity conditions imposed on the optimal trajectory w.r.t. the state constraints. Below we provide rigourous proof for this fact. This proof is carried out by virtue of the same idea and similar arguments to the ones suggested in [3], however, the hypothesis for smoothness of the data is reduced. In [3], even the Hölder continuity of the measure is established. However, the proof therein uses an extra smoothness w.r.t. the u-variable. Herein, we intend to use the same class of smoothness for the data, as in [1]. Moreover, unlike [3], geometrical constraints on control, given by an arbitrary feasible closed set U, are considered.

[*] This work was supported by the Russian Foundation for Basic Research, projects 16-31-60005, 17-51-12064, 18-01-00106, by the Ministry of Education and Science of the Russian Federation (Project no. 1.962.2017/4.6), and by FCT (Portugal): support of FCT R&D Unit SYSTEC – POCI-01-0145-FEDER-006933/SYSTEC funded by ERDF | COMPETE2020 | FCT/MEC | PT2020 extension to 2018, and Project STRIDE – NORTE-01-0145-FEDER-000033, by ERDF | NORTE 2020. The first author was supported by the Russian Science Foundation, project no. 17-11-01168.

The continuity of the measure-multiplier appears to be important for numerical implementation in the framework of indirect computational approach. Indeed, the fact that the multiplier jumps, thereby having singularities, creates obstacles on the way to apply the standard indirect approach and the shooting algorithm to resolve the co-state equation, as the "number of variables" is, in general, higher than the "number of equations". In what follows, we demonstrate the absence of singularities of this kind, if the above mentioned regularity conditions are satisfied. Therefore, the presented theoretical aspects may be useful in practical applications.

Consider the following optimal control problem with state constraints.

$$\begin{array}{ll} \text{Minimize} & \displaystyle\int_0^1 f_0(x,u,t)dt, \\ \text{subject to} & \dot{x} = f(x,u,t), \\ & x(0) = x_A, \;\; x(1) = x_B, \\ & u(t) \in U \;\text{ for a.a.}\, t \in [0,1], \\ & g(x(t),t) \le 0 \;\; \forall\, t \in [0,1]. \end{array} \tag{1}$$

Here, $\dot{x} := \frac{dx}{dt}$, $t \in [0,1]$ signifies the time variable, x is the state variable with values in $\mathbb{R}^n$. The mappings $f_0 : \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R} \to \mathbb{R}$, $f : \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R} \to \mathbb{R}^n$, and $g : \mathbb{R}^n \times \mathbb{R} \to \mathbb{R}^k$ satisfy a certain smoothness assumption specified below, the set U is a closed subset of $\mathbb{R}^m$. The vectors x_A and x_B in the state-space are the so-called starting and terminal positions. The measurable bounded function $u(\cdot)$ is termed control. The feasible trajectory $x(\cdot)$ is supposed to be absolutely continuous and to satisfy the differential constraints $\dot{x}(t) = f(x(t),u(t),t)$ for a.a. $t \in [0,1]$, and the inequalities $g^j(x(t),t) \le 0 \;\forall\, t \in [0,1]$, $j = 1,...,k$, which define the *state constraints*.

Let $u(\cdot)$ be a control function, and $x(\cdot)$ be a corresponding trajectory. The pair (x,u) is said to be *feasible process* provided that the endpoint and state constraints and the control constraints: $u(t) \in U$ for a.a. $t \in [0,1]$, are satisfied. A feasible process (x^*,u^*) is termed *optimal* if the value of the cost integral is the least possible over the set of all feasible processes. This is the well-known notion of the so-called global minimum.

The data will satisfy the following hypothesis.

Hypothesis 1 *Functions f, f_0 are C^1-, and g is C^2-smooth.*

The second-order derivative for the vector-valued function g, which defines state constraints, is required due to the fact that the Gamkrelidze set of necessary conditions is used in that which follows.

Recall this set of conditions, [1,4,5]. Consider the extended Hamilton-Pontryagin function

$$\bar{H}(x,u,\psi,\mu,\lambda,t) := \langle \psi, f(x,u,t)\rangle - \langle \mu, \Gamma(x,u,t)\rangle - \lambda f_0(x,u,t),$$

where ψ,μ,λ are variables in $\mathbb{R}^n$, $\mathbb{R}^k$, and $\mathbb{R}$ respectively, and function Γ is defined as follows

$$\Gamma(x,u,t) = g'_x(x,t)f(x,u,t) + g'_t(x,t).$$

Definition 1 *The control process (x^*,u^*) in Problem (1) is said to satisfy the nondegenerate maximum principle provided that there exist Lagrange multipliers: a number $\lambda \geq 0$, an absolutely continuous function $\psi : [0,1] \to \mathbb{R}^n$, and a bounded function $\mu : [0,1] \to \mathbb{R}^k$, such that $\mu \neq 0$ on $(0,1)$ if $\lambda = 0$ and $\psi = 0$, each component μ^j, $j = 1,..,k$, is decreasing, thereby, defining a Borel measure, and*

$$\begin{aligned}
&\dot{\psi}(t) = -\bar{H}'_x(x^*(t),u^*(t),\psi(t),\mu(t),\lambda,t) \ \textit{for a.a.}\, t \in [0,1],\\
&\max_{u\in U} \bar{H}(x^*(t),u,\psi(t),\mu(t),\lambda,t) =\\
&\qquad\qquad = \bar{H}(x^*(t),u^*(t),\psi(t),\mu(t),\lambda,t) \ \textit{for a.a.}\, t \in [0,1],\\
&\int_0^1 \langle g(x^*(t),t), d\mu\rangle = 0.
\end{aligned}$$

As is known, under natural regularity or controllability conditions w.r.t. the state constraints, any optimal process (x^*,u^*) satisfies the nondegenerate maximum principle, [6]. In that which follows we aim to provide a few of such regularity concepts. In this regularity framework, the maximum principle is not only nondegenerate, but also the measure-multiplier μ is continuous on $(0,1)$.

Consider a feasible process (x^*,u^*). Define $J(t) = \{j : g^j(x^*(t),t) = 0\}$, and let $\mathcal{U}(t)$ be the closure of $u^*(t)$ w.r.t. the Lebesgue measure (see in [3]). By convention, we set $\mathcal{U}^-(0) := \mathcal{U}^+(0)$, $\mathcal{U}^+(1) := \mathcal{U}^-(1)$. By $\mathcal{U}^+(t)$ and $\mathcal{U}^-(t)$ denote the right and the left closure w.r.t. measure respectively.

The definition that follows is an extension of the regularity condition imposed on the optimal trajectory in [1].

Definition 2 *The feasible process* (x^*, u^*) *is said to be left-regular at point* $t \in [0,1]$ *w.r.t. the state constraints, provided that for all* $u \in \mathcal{U}^-(t)$ *there exists a vector* $d = d(u,t) \in T_U(u) \cap N_U^*(u)$ *such that*

$$\Big\langle \frac{\partial \Gamma^j}{\partial u}(x^*(t), u, t), d \Big\rangle > 0 \ \ \forall j \in J(t). \tag{2}$$

Respectively, it is said to be right-regular at this point, provided that for all $u \in \mathcal{U}^+(t)$ *there exists a vector* $d = d(u,t) \in T_U(u) \cap N_U^*(u)$ *such that*

$$\Big\langle \frac{\partial \Gamma^j}{\partial u}(x^*(t), u, t), d \Big\rangle < 0 \ \ \forall j \in J(t). \tag{3}$$

Process (x^*, u^*) *is said to be regular if it is either left-, or right-regular in each point of the time interval.*

Here, $T_U(u)$ stands for the contingent tangent cone to the set U at point u, and $N_U^*(u)$ is the dual cone to the limiting normal cone $N_U(u)$ defined in [7].

It follows that, in the scalar case, that is, when $k = 1$ the regularity concept can be weakened. (In fact, it can also be weakened for $k > 1$, but then some extra rigid assumptions on the behavior of trajectory arise, [3].) Define sets

$$\begin{aligned} T_0 &:= \{t \in [0,1] : g(x^*(t), t) = 0\}, \\ Z(t) &:= \{u \in U : \Gamma(x^*(t), u, t) = 0\}. \end{aligned}$$

Definition 3 *The feasible process* (x^*, u^*) *is said to be regular w.r.t. the state constraints, provided that for all* $t \in T_0$ *there exist* $u \in \mathcal{U}(t)$ *and* $d \in T_U(u) \cap N_U^*(u)$ *such that*

$$\begin{aligned} &\big\langle \Gamma'_u(x^*(t), u, t), d \big\rangle > 0 \ \ \textit{if} \ \ u \in Z(t) \cap \mathcal{U}^-(t), \\ &\big\langle \Gamma'_u(x^*(t), u, t), d \big\rangle < 0 \ \ \textit{if} \ \ u \in Z(t) \cap \mathcal{U}^+(t). \end{aligned} \tag{4}$$

Note that (2) is automatically satisfied when $\mathcal{U}(t) \subsetneq Z(t)$.

This type of regularity is weaker than the one in Definition 2. Thus, we will also term these types as weakly regular, and strongly regular, respectively. Note that: *Any feasible process is weakly regular provided that for all* $x \in \mathbb{R}^n$, $u \in U$, *and* $t \in [0,1]$ *such that* $g(x,t) = 0, \Gamma(x,u,t) = 0$, *there exist vectors* $d_+, d_- \in T_U(u) \cap N_U^*(u)$ *such that*

$$\big\langle \Gamma'_u(x,u,t), d_+ \big\rangle > 0, \ \textit{while} \ \big\langle \Gamma'_u(x,u,t), d_- \big\rangle < 0. \tag{5}$$

It is possible to point out to various classes of control problems with state constraints for which this global and a priori verification condition works.

The following lemma is our main result.

Lemma 1 *Let (x^*, u^*) be an extremal process w.r.t. the nondegenerate maximum principle. Assume that at least one of the following conditions is satisfied:*

i) process (x^*, u^*) is strongly regular in the sense of Definition 2, while the set U satisfies the regularity property that, for any $u \in U$, $d \in T_U(u)$, there exists a function $o(\cdot) : \mathbb{R} \to \mathbb{R}^m$ such that $u + \varepsilon d + o(\varepsilon) \in U\ \forall \varepsilon > 0$, and $o(\varepsilon)/\varepsilon \to 0$ as $\varepsilon \to 0$;[*]*

ii) process (x^*, u^*) is weakly regular in the sense of Definition 3, and $k = 1$.

Then, for any set of the Lagrange multipliers (ψ, μ, λ) corresponding to (x^*, u^*) by virtue of Definition 1, the measure Lagrange multiplier $\mu(\cdot)$ is continuous on $(0, 1)$.

References

1. Pontryagin, L.S., Boltyanskii, V.G., Gamkrelidze, R.V., and Mishchenko, E.F.: The Mathematical Theory of Optimal Processes, Interscience, New York (1962)
2. Gamkrelidze, R.V.: Optimal control processes with restricted phase coordinates, Izv. Akad. Nauk SSSR Scr. Mat., 24, 315–356 (1960)
3. Arutyunov, A.V., Karamzin, D.Yu. On Some Continuity Properties of the Measure Lagrange Multiplier from the Maximum Principle for State Constrained Problems. SIAM Journal on Control and Optimization, 53, 4 (2015)
4. Neustadt, L.W.: An abstract variational theory with applications to a broad class of optimization problems. II: Applications, SIAM J. Control, 5 (1967)
5. Arutyunov, A.V., Karamzin, D.Yu., Pereira, F.L.: The Maximum Principle for Optimal Control Problems with State Constraints by R.V. Gamkrelidze: Revisited. J. Optim. Theory Appl., 149 (2011)

[*] This means that the contingent tangent cone coincides everywhere with the inner tangent cone.

6. Arutyunov, A.V.: Optimality conditions: Abnormal and Degenerate Problems. Mathematics and Its Application. Kluwer Academic Pub-lisher (2000)
7. Mordukhovich, B.S., Variational Analysis and Generalized Differen-tiation. Volume I. Basic Theory. Springer-Verlag, 2006, Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences], Berlin.

Projection onto a convex quadratic surface

V.A. Bereznev
Federal Research Center of Computer Science and Control,
Moscow, Russia

We consider the problem of computing the projection of the vector $x^0 \in \mathbb{R}_n$ on a convex quadratic surface bounding a certain convex set X,

$$\min_{x \in X} \|x - x^0\|, \quad X = \{x \in \mathbb{R}_n \mid g(x) = \frac{1}{2}\langle x, Hx\rangle - \langle c, x\rangle + b \leq 0\}, \quad (1)$$

where H is symmetric positive definite matrix $n \times n$, $c \in \mathbb{R}_n$, $b \in \mathbb{R}_1$. Without loss of generality, we can assume that $x^0 = 0$. Denote by $\hat{x} = H^{-1}c$ is a point of unconditional minimum of the function $g(x)$. In addition, we assume that $X \neq \oslash$ and satisfies the Slater's regularity condition , i.e. $g(\hat{x}) < 0$.

If the point $x^0 \in X$, then obviously, it is a solution of (1). Assume that $x^0 \notin X$, that is $g(x^0) = b > 0$. Since the set of X is convex, then the solution of the problem (1) is the projection of x^* of a point x^0 on X. Moreover, it is obvious that x^* belongs to the boundary of X, i.e. $g(x^*) = 0$. Thus, the (1) problem is equivalent to the

$$\min_{x \in X_0} \frac{1}{2}\|x\|^2, \quad X_0 = \{x \in \mathbb{R}_n \mid g(x) = 0\}. \quad (2)$$

Problem (2) can be solved using number of methods. One of the most effective the rate of convergence is the Newton-Lagrange method. It should be noted, however, that this method is iterative, and each iteration is associated with the solution of system of equations. Below we provide the method (call it *a two-surface method*) that provides a solution to (2) over one iteration, which undoubtedly should be considered an advantage in comparison with Newton-Lagrange method.

First of all, we note the situation in which the solution of (2) is easily guessed. Let $c \neq 0$ and

$$H^{-1}c - \mu c = 0, \tag{3}$$

for some $\mu > 0$, that is, the vector c is the eigenvector of the operator H^{-1}. Then, as is easy to see, the solution is $x^* = \lambda^* H^{-1}c$, where

$$\lambda^* = 1 - \sqrt{1 - \frac{2b}{\langle c, H^{-1}c\rangle}}.$$

Assume that (3) does not hold. Consider the set $M = \{x \in \mathbb{R}_n \mid \langle g'(x), x\rangle = 0\}$. It is easy to see that M is a quadratic surface in $\mathbb{R}_n$, namely,

$$M = \{x \in \mathbb{R}_n \mid \langle x, Hx\rangle - \langle c, x\rangle = 0\}.$$

Denote by S a subset of the set M that has a non-empty intersection with the set X. It is easy to make sure that $\hat{x} \in M \cap X$, which implies that $S = M \cap X \neq \varnothing$.

We introduce a function of one variable $\varphi(\xi)$ under $\xi \in [0, 1)$. The computation of the values of $\varphi(\xi)$ is reduced to several elementary steps. In particular, consider the ray $\tilde{x}(\xi) = \xi\hat{x}$ and note that at each point $\xi \in [0, 1)$ gradient $g'(\tilde{x}(\xi)) = \xi H\hat{x} - c = (\xi - 1)c$ is equally directed and it differs only in the length of the vector. Let $\tilde{x}(\xi)$ be an arbitrary point of the specified part of the beam. Let us take a step from this point in the direction of the antigradient of the function $g(x)$ before crossing with the set S. So we have

$$u(\alpha, \xi) = \tilde{x}(\xi) - \alpha g'(\tilde{x}(\xi)) = \xi H^{-1}c - \alpha(\xi - 1)c, \tag{4}$$

where the step $\alpha = \alpha(\xi) > 0$ is subject to the condition

$$\langle Hu(\alpha, \xi) - c, u(\alpha, \xi)\rangle = 0, \tag{5}$$

and the gradient of the function $g(x)$ at the point $u(\alpha, \xi)$ is equal to $Hu(\alpha, \xi) - c$.

We can show that for any fixed $\xi \in [0, 1)$ the equation (5) is solvable with respect to $\alpha = \alpha(\xi)$, and there is a positive root of this equations. In particular,

$$\alpha = \frac{(2\xi - 1)\|c\|^2 - \sqrt{(2\xi - 1)^2\|c\|^4 - 4\xi(\xi - 1)\langle c, Hc\rangle\langle c, H^{-1}c\rangle}}{2(\xi - 1)\langle c, Hc\rangle}. \tag{6}$$

Note also that for any $u \in s$ there exists such $\lambda \in (0,1)$ that the point $x_\lambda = \lambda u \in X_0$. It is enough to solve the square equation

$$\frac{\lambda^2}{2}\langle u, Hu\rangle - \lambda\langle c, u\rangle + b = 0. \tag{7}$$

Taking into account that for any $u \in M$ we have the equality $\langle u, Hu\rangle = \langle c, u\rangle > 0$ and given that $\lambda \in (0,1)$, we have

$$\lambda = 1 - \sqrt{1 - \frac{2b}{\langle c.u\rangle}}. \tag{8}$$

Thus, fixing $\xi \in [0,1)$, we can uniquely calculate $\alpha(\xi)$ by the formula (6), then $u(\xi) = u(\alpha, \xi)$ by the formula (4) and $\lambda(\xi)$ by the formula (8). Finally, we obtain $x(\xi) = \lambda(\xi)u(\xi)$. Then we have the following

Theorem. *The point $x^* = x(\xi^*)$ is the solution of* (2) *if and only if ξ^* is the solution of the equation*

$$\varphi(\xi) = \langle g'(x(\xi)), g'(u(\xi))\rangle = 0. \tag{9}$$

Proof. *Necessity.* Let $x^* = x(\xi^*)$ be the solution problems (2). Then by theorem of Karush-Kuhn-Tucker there is a number $\gamma > 0$, that $x^* = -\gamma g'(x^*)$. Substituting from this equality the expression for $g'(x^*)$ into the formula (9), we get

$$\varphi(\xi^*) = -\frac{1}{\gamma}\langle x^*, g'(u(\xi^*))\rangle = -\frac{\lambda}{\gamma}\langle u(\xi^*), g'(u(\xi^*))\rangle = 0,$$

since $u(\xi^*) \in M$.

Sufficiency. Let $\varphi(\xi^*) = 0$. The statement of the theorem will be proved if there is such a number $\gamma > 0$ that $x(\xi^*) = -\gamma g'(x(\xi^*))$, which is a sufficient condition for the optimality of $x^* = x(\xi^*)$ due to the same Karush-Kuhn-Tucker theorem.

Assume that such $\gamma > 0$ does not exist. Note that for any $\xi \in [0,1)$ the following equality hold true:

$$g'(u(\xi)) = \xi c - \alpha(\xi - 1)Hc - c = (\xi - 1)c - \alpha(\xi - 1)Hc,$$

$$g'(x(\xi)) = g'(\lambda u(\xi)) = (\lambda\xi - 1)c - \alpha\lambda(\xi - 1)Hc.$$

In other words, the vectors $g'(u(\xi))$ and $g'(x(\xi))$ belong to the plane $\pi = \{x \in \mathbb{R}_n \mid x = \mu_1 c + \mu_2 Hc, \quad \mu_1, \mu_2 \in \mathbb{R}_1\}$, stretched over vectors c and Hc. Therefore, this plane contains the straight line

$$l = \{x \in \mathbb{R}_n \mid x = u(\xi) + \tau g'(u(\xi)), \quad \tau \in \mathbb{R}_1\}.$$

Consider a point $\tilde{u}(\tilde{\alpha},\xi^*) = x(\xi^*) - \tilde{\alpha}g'(x(\xi^*) = \lambda u(\xi^*) - \tilde{\alpha}g'(x(\xi^*))$, where $\tilde{\alpha}$ is selected so that $\tilde{u}(\tilde{\alpha},\xi^*) \in m$. Multiply scalar this equality on $g'(u(\xi^*))$. So we have

$$\langle \tilde{u}(\tilde{\alpha},\xi^*), g'(u(\xi^*))\rangle = \lambda\langle u(\xi^*), g'(u(\xi^*))\rangle - \tilde{\alpha}\langle g'(x(\xi^*)), g'(u(\xi^*))\rangle = 0,$$

i.e. vectors $\tilde{u}(\tilde{\alpha},\xi^*)$ and $u(\xi^*)$ are orthogonal to $g'(u(\xi^*))$. It follows that $\tilde{u}(\tilde{\alpha},\xi^*) = u(\xi^*)$, because only one perpendicular can be omitted from the origin of the π plane on the line l. Thus, the vectors $u(\xi^*)$ and $g'(x(\xi^*)$ are collinear, taking into account them in the opposite direction have $u(\xi^*) = -\rho g'(x(\xi^*)), \quad \rho > 0$, or

$$x(\xi^*) = \lambda(\xi^*)u(\xi^*) = -\lambda(\xi^*)\rho g'(x(\xi^*)) = -\gamma g'(x(\xi^*)),$$

where $\gamma = \lambda(\xi^*)\rho > 0$, which proves the theorem. ■

So, to solve the problem (2) it is enough to solve the equation (9) with respect to the only variable ξ. We can use the iterative procedure to do this, proposed in [1] and characterized by very low computing in comparison with the Newton-Lagrange method, as shown by the results of the computational experiment. However, it is possible specify a more efficient procedure for calculating the projection on a convex quadratic surface. The validity of the following statement is easily verified.

Lemma. *For the function* $\varphi(\xi)$ *in the interval* $\xi \in [0,1)$ *the relations are valid:* $\varphi(0) \geq 0$; $\lim\limits_{\xi\to 1}\varphi(\xi) < 0$.

Based on the statement of lemma, we can describe the scheme of solving the problem.

Algorithm.

Step 1. Calculate H^{-1} and verify that the condition (3) is hold. If (3) is hold, then put $x^* = \lambda^* H^{-1} c$ and stop. Otherwise, go to step 2.

Step 2. Calculate $\varphi(0)$. If $\varphi(0) = 0$, then put $x^* = \alpha_0 \lambda_0 c$, where

$$\alpha_0 = \frac{\|c\|^2}{\langle c, Hc \rangle}, \quad \lambda_0 = 1 - \sqrt{1 - \frac{2b}{\alpha_0 \|c\|^2}},$$

and stop. Otherwise, put $\delta = 10^{-2}, \quad \xi_1 = 0$, and go to step 3.

Step 3. Calculate $\varphi(1-\delta)$ and go to step 4.

Step 4. If $\varphi(1-\delta) < 0$, then put $\xi_2 = 1-\delta$ and go to step 5. Otherwise, put $\delta = \delta/2$ and go to step 3.

Step 5. Solve the problem of finding zero of the function φ on the interval $[\xi_1, \xi_2]$ and stop.

Example. Let

$$H = \begin{pmatrix} 1 & 2 \\ 2 & 5 \end{pmatrix}, \quad c = \begin{pmatrix} 4 \\ 7 \end{pmatrix}, \quad b = 5.2.$$

To obtain a solution $x^* = (0.65325444; 0.71858815)$ with a given accuracy required five iterations of the Newton-Lagrange method. To obtain the same solution using the iterative procedure of the two-surface method [1] required forty iterations. Finally, the above algorithm also yields x^*.

Taking into account the result obtained in [2], there is every reason to believe, that on the basis of the given algorithm a new effective version of the method of sequential quadratically constrained quadratic programming can be developed.

References

1. Bereznev V.A. Proection of vector on quadratic surface // Theoretical and applied problems of nonlinear analysis. 2015, P.3-11.
2. Bereznev V.A. A polynomial algorithm for the quadratic programming problem // Russian Journal of Numerical Analysis and Mathematical Modelling, 2014, V.29, Issue 3, P.139-144.

An approach to solve special classes of multi-extremal problems*

I.A. Bykadorov
Sobolev Institute of Mathematics SB RAS,
Novosibirsk State University,
Novosibirsk State University of Economics and Management,
Novosibirsk, Russia

We suggest an approach to optimize the monotone combination (e.g., sum, product) of several functions. The main attention is paid to the minimization of the sum of two functions. We assume that effective algorithms are known to solve the problems $f_i(x) \to \min_{x\in X}, i = 1,2$, where $X \subset \mathbb{R}^n$ while f_i are functions defined on X. Consider the problem

$$(P) : f(x) = f_1(x) + f_2(x) \to \min_{x\in X}.$$

For problem (P), the effective algorithms are known only for the case when the set X is polyhedral and the functions f_i have a special form. Let us only mention some works: [1], [2], [3]. A feature of the mentioned works is that they take into account special kind of problem, so they cannot be transferred directly to the more general form of functions f_i.

Let us associate with each $\nu_i \in \mathbb{R}$ the following subsets of the set X: $X_i(\nu_i) = \{x \in X : f_i(x) \le \nu_i\}, i = 1,2$, and consider the problems

$$(P_1(\nu_2)) : f_1(x) \to \min_{x\in X_2(\nu_2)}, \qquad (P_2(\nu_1)) : f_2(x) \to \min_{x\in X_1(\nu_1)}.$$

Let $x_{P_1(\nu_2)}$ and $x_{P_2(\nu_1)}$ be the solutions of the problems $(P_1(\nu_2))$ and $(P_2(\nu_1))$, respectively. Of course, for arbitrary choice of ν_1 and ν_2, it is possible that problems $(P_1(\nu_2))$ and $(P_2(\nu_1))$ have no solutions. But if $\nu_i^0 = \min_{x\in X} f_i(x), i = 1,2$, then problems $\left(P_1\left(\nu_2^0\right)\right)$ and $\left(P_2\left(\nu_1^0\right)\right)$ are solvable and $f_1\left(x_{P_2\left(\nu_1^0\right)}\right) = \nu_1^0$, $f_2\left(x_{P_1\left(\nu_2^0\right)}\right) = \nu_2^0$. Let us denote $\nu_1^{00} = f_1\left(x_{P_1\left(\nu_2^0\right)}\right)$ and $\nu_2^{00} = f_2\left(x_{P_2\left(\nu_1^0\right)}\right)$. A pair (ν_1,ν_2) is said to be **attainable** if a point $x \in X$ exists such that $f_i(x) = \nu_i, i = 1,2.$. Pairs $\left(\nu_1^0,\nu_2^{00}\right)$ and $\left(\nu_1^{00},\nu_2^0\right)$ are attainable. Let x^* be the solution of Problem (P) and $f_i(x^*) = \nu_i^* \in \left[\nu_i^0,\nu_i^{00}\right], i = 1,2$. Consider the right isosceles triangle ABC with vertices $A = \left(\nu_1^0 + \nu_{1,2}^0,\nu_2^0\right)$, $B = \left(\nu_1^0,\nu_2^0\right)$, $C = \left(\nu_1^0,\nu_2^0 + \nu_{1,2}^0\right)$, where $\nu_{1,2}^0 = \min\left\{\nu_1^{00} - \nu_1^0,\nu_2^{00} - \nu_2^0\right\}$ (see Fig. 1).

*The work was supported by the program of fundamental scientific researches of the SB RAS No I.5.1., project No 0314-2016-0018. Supported in part by RFBR grants 16-01-00108, 16-06-00101 and 18-010-00728.

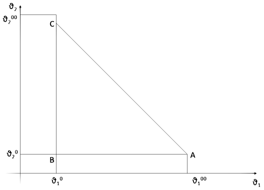


Fig. 1. Triangle ABC.

Now, problem (P) is equivalent to the following: *in triangle ABC, find a point such that its coordinates (ν_1, ν_2) form an attainable pair and moreover $\nu_1 + \nu_2 = \nu_1^* + \nu_2^*$.* Let us denote $T_i = \left[\nu_i^0, \nu_i^{00}\right], i = 1, 2$. In what follows, we assume that the following condition is fulfilled.

Condition (A). *For any pair $(\nu_1, \nu_2) \in T_1 \times T_2$, the points $x' \in X$ and $x'' \in X$ exist such that $f_1(x') = \nu_1,\ f_2(x') = f_2\left(x_{P_2(\nu_1)}\right),\ f_1(x'') = f_1\left(x_{P_1(\nu_2)}\right),\ f_2(x'') = \nu_2$.*

If functions f_1, f_2 are quasi-convex on X, then Condition (A) holds. Let $(\nu_1, \nu_2) \in T_1 \times T_2$. Under Condition (A), the pairs $\left(\nu_1, f_2\left(x_{P_2(\nu_1)}\right)\right), \left(f_1\left(x_{P_1(\nu_2)}\right), \nu_2\right)$ are attainable. Define decreasing function $G : T_1 \to \mathbb{R}$ as $G(\nu_1) = \min\{f_2(x) : x \in X_1(\nu_1), \nu_1 \in T_1\}$. Consider the set $Y = \{(\nu_1, G(\nu_1)) : \nu_1 \in T_1\}$. We associate with each pair $(\nu_1, \nu_2) \in Y$ the line $H(\nu_1, \nu_2)$ passing through it and parallel to the hypotenuse of triangle ABC (see Fig. 2). The pair (ν_1^*, ν_2^*) (the values of the functions f_1 and f_2 in the optimum) is characterized as follows: *for each pair $(\nu_1, \nu_2) \in Y$, the line $H(\nu_1, \nu_2)$ lies "above" the line $H(\nu_1^*, \nu_2^*)$.*

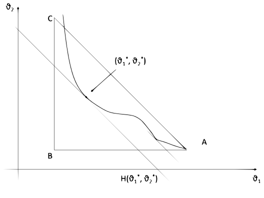


Fig. 2. Curve Y and lines $H(\nu_1, \nu_2)$.

The algorithm suggested is iterative and consists in the sequential refinement of the estimates of the values ν_1^*, ν_2^* and $\nu_1^* + \nu_2^*$, as well as the reduction the total area of the region containing the point (ν_1^*, ν_2^*).

Let $\nu_1 \in \left[\nu_1^0, \nu_1^0 + \nu_{1,2}^0\right]$. Let us set $\nu_2 = G(\nu_1)$. Note that $(\nu_1, \nu_2) \in Y$ by the definition of set Y. The following cases are possible:

1. (see Fig. 3) $\nu_1 + \nu_2 < \nu^0 \equiv \nu_1^0 + \nu_2^0 + \nu_{1,2}^0$;

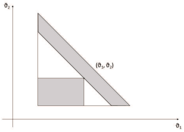


Fig. 3. Case 1: $\nu_1 + \nu_2 < \nu^0 \equiv \nu_1^0 + \nu_2^0 + \nu_{1,2}^0$.

2. (see Fig. 4) $\nu_1 + \nu_2 = \nu^0$;

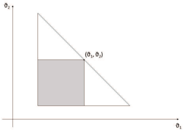


Fig. 4. Case 2: $\nu_1 + \nu_2 = \nu^0$.

3. (see Fig. 5) $\nu_1 + \nu_2 > \nu^0,\ \nu_2 < \nu_2^0 + \nu_{1,2}^0$;

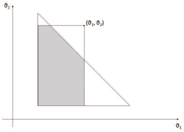


Fig. 5. Case 3: $\nu_1 + \nu_2 < \nu^0 \equiv \nu_1^0 + \nu_2^0 + \nu_{1,2}^0$.

4. (see Fig. 6) $\nu_1 + \nu_2 > \nu^0,\ \nu_2 \geq \nu_2^0 + \nu_{1,2}^0$.

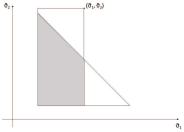


Fig. 6. Case 4: $\nu_1 + \nu_2 < \nu^0 \equiv \nu_1^0 + \nu_2^0 + \nu_{1,2}^0$.

In each of these cases we can exclude from further consideration the regions that do not contain the point of our interest (ν_1^*, ν_2^*). These areas correspond to the shaded parts of the triangle. This way, the bounds for value $\nu_1^* + \nu_2^*$ are refined, and the area of the region, which is guaranteed not containing the point (ν_1^*, ν_2^*), is increased. In the next steps, the procedure is applied to each of the obtained unshaded triangles, or to one of them (for example, the largest, "the most perspective").

Remark 1. As initial lower bounds for ν_1^*, ν_2^* and $\nu_1^* + \nu_2^*$, we can take, for example, the values $\underline{\nu}_1 = \nu_1^0$, $\underline{\nu}_2 = \nu_2^0$, $\underline{\nu} = \underline{\nu}_1 + \underline{\nu}_2$, while as the upper bounds, the values $\overline{\nu}_1 = f_1(x^1)$, $\overline{\nu}_2 = f_2(x^2)$, $\overline{\nu} = \min\left\{\underline{\nu}_1 + \overline{\nu}_2, \overline{\nu}_1 + \underline{\nu}_2\right\}$, where $x^i \in \{x \in X : f_i(x) = \underline{\nu}_i\}, \quad i = 1, 2$.

Remark 2. If $\nu_1 = (\underline{\nu}_1 + \overline{\nu}_1)/2$, $\nu_2 = f_2\left(x_{P_2(\nu_1)}\right)$, then we can delete from the triangle the region whose area is not less than half the area of the triangle, see Fig. 3 – Fig. 6. So, at each iteration, we can exclude from considerate triangle a part whose area is not less than half the area of the entire triangle. This allows us to tell about the effectiveness of the algorithm.

Remark 3. The disadvantage of the approach is the possible increase in the number of resulting triangles. In this case, one triangle can be chosen (for example, the largest one, i.e., "perspective") and temporarily "forget" about the others (see Fig. 7), thus obtaining new estimates of the quantities ν_1^*, ν_2^* and $\nu_1^* + \nu_2^*$ for this selected triangle. These new estimates may allow us to exclude some of the "forgotten" triangles from further consideration, since we remove all parts of the triangles lying "above" the corresponding hypotenuse (this part may coincide with the whole triangle, see figure Fig. 8). Thus, the number of considered triangles does not necessarily increase and, moreover, may even decrease.

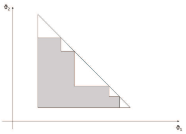


Fig. 7. Selection of the most "perspective" triangle.

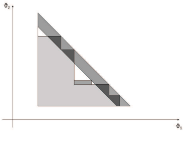


Fig. 8. The situation when the number of triangles decreases.

Remark 4. See details in [4]. The approach can be applied to the monopolistic competition models, see e.g. [5], [6], [7], [8].

References

1. Falk J.E., Palocsay S.M. Optimizing the sum of linear fractional functions //Recent Advances in Global Optimization. Princeton: Princeton University Press, 1992. P. 221–258.
2. Kuno T. A branch-and-bound algorithm for maximizing the sum of several linear ratios // Journal of Global Optimization. 2002. V. 22, No. 1-4. P. 155–174.
3. Gruzdeva T., Strekalovsky A. On a Solution of Fractional Programs via D.C. Optimization Theory // CEUR Workshop Proceeding. 2017. V. 1987. P. 246–252.
4. Bykadorov I. Solution of Special Classes of Multi-extremal Problems // CEUR Workshop Proceeding. 2017. V. 1987. P. 115–122.
5. Dixit A.K., Stiglitz J.E. Monopolistic Competition and Optimum Product Diversity // American Economic Review. 1977. V. 67, No. 3. P. 297–308.
6. Krugman P.R. Increasing returns, monopolistic competition and international trade // Journal of International Economics. 1979. V. 9, No. 4. P. 469–479.
7. Bykadorov I., Gorn A., Kokovin S., Zhelobodko E. Why are losses from trade unlikely? // Economics Letters. 2015. V. 129. P. 35–38.
8. Aizenberg N., Bykadorov I., Kokovin S. Beneficial welfare impact of bilateral tariffs under monopolistic competition // Abstracts of the Tenth International Conference Game Theory and Management. Saint Petersburg: Saint Petersburg State University, 2017. P. 5–7.

On the shape of power flow feasibility set and Jacobian singularity curves

I.P. Bogdanov
Keldysh Institute of Applied Mathematics of RAS, Moscow, Russia

Exploring the structure of power flow feasibility set for alternating current (AC) grid is a considerable part of research works, related to power market analysis, stability assessment and fuel cost optimization.

A simplified model of AC grid includes buses (characterized by complex voltage and power injection values: $U = |U| \cdot e^{i\delta}$ and $S = P + Q \cdot i$ respectively), connected by a set of transmission lines (characterized by impedance and transformation ratio). Power flow in an AC grid is described by a system of nonlinear equations, which formalize power balance in each bus and reflect relationships between components of nodal voltages and power injections. Vectors of nodal powers, guaranteeing the existence of solution for the regarded system of equations, form the power flow feasibility set.

The analysis of power flow feasibility boundary enables us to determine power reserves, obtain upper estimates for operational parameter values (e.g. in economic dispatch and stability assessment problems) and ignore obviously non-physical modes.

Consider the power network model, consisting of $n + 1$ buses. 0-th bus is the slack bus with fixed voltage magnitude value $|U_0|$ and zero phase angle value $\delta_0 = 0$. Buses from 1 to m are the PV-buses with specified active power (real part of complex power injection) values $P_1, \dots, P_m$ and specified voltage magnitude values $|U_1|, \dots, |U_m|$. Buses from $m + 1$ to n are the PQ-buses with specified active power values $P_{m+1}, \dots, P_n$ and specified reactive power (imaginary part of complex power injection) values $Q_{m+1}, \dots, Q_n$. Voltage magnitudes $|U_{m+1}|, \dots, |U_n|$ and phase angles $\delta_1, \dots, \delta_n$ are determined from the solution of power flow equations:

$$\begin{cases} P_k = |U_k| \sum\limits_{l=0}^{n} |U_l||Y_{kl}| \cos(\delta_k - \delta_l - \arg Y_{kl}), \ k = 1, \dots, n, \\ Q_j = |U_j| \sum\limits_{l=0}^{n} |U_l||Y_{jl}| \sin(\delta_j - \delta_l - \arg Y_{jl}), \ j = m+1, \dots, n, \end{cases} \tag{1}$$

where Y_{kl} (Y_{jl}) denote the elements of the admittance matrix [1]. Power flow feasibility set for the regarded grid is a $(2n - m)-$dimensional domain, comprising all vectors $(P_1, \ldots, P_n,\ Q_{m+1}, \ldots, Q_n)$, which provide the existence of a solution for (1).

One of the most common approaches for nodal power limits investigation implies identifying intersection of power flow feasibility boundary with linear trajectory:

$$\begin{cases} P_k = \lambda \cdot a_k^P + P_k^0,\ k = 1, \ldots, n, \\ Q_j = \lambda \cdot a_j^Q + Q_j^0,\ j = m+1, \ldots, n, \end{cases} \tag{2}$$

where λ is a scalar parameter, $P_1^0, \ldots, P_n^0,\ Q_{m+1}^0, \ldots, Q_n^0$ correspond to some basic mode and $a_1^P, \ldots, a_n^P,\ a_{m+1}^Q, \ldots, a_n^Q$ are given coefficients, determining the direction of regarded trajectory. Thus, we obtain the following nonlinear optimization problem:

$$\lambda \to \max_{\left(\substack{\lambda, \\ \delta_k, P_k,\ k=1,\ldots,n, \\ |U_j|, Q_j,\ j=m+1,\ldots,n}\right)},$$

subject to power flow equations (1) (where $P_1, \ldots, P_n,\ Q_{m+1}, \ldots, Q_n$ are treated as variables, while $|U_1|, \ldots, |U_m|$ are given and fixed) and constraints (2).

Various numerical techniques — direct methods, sequential quadratic programming, augmented Lagrangian function algorithm, continuation methods etc. — were proposed to solve this problem [1]. However, in general, power flow feasibility domain is non-convex [2] — for instance, it can be star-shaped or ring-shaped [3, 4]. Moreover, Jacobian matrix of power flow equations is singular on the outer boundary of power flow feasibility set and on complex-shaped curves inside this domain. In the vicinity of these singularity curves Jacobian matrix becomes ill-conditioned. Described peculiarities weaken the performance of numerical algorithms, mentioned above, inducing divergence or convergence to local extrema, corresponding to unstable mode.

The paper presents a set of power flow feasibility domains' and Jacobian singularity curves' images, constructed for a 3-bus test power system (Figure 1). Test system consists of one slack bus and two PV-buses, connected with three transmission lines (i.e. power flow is formalized by two equations, describing active power balance at PV-buses). Parameter values (line impedances and voltage magnitudes) are similar to those, regarded in [4]. Complex numbers beside transmission lines (shown in Figure 1) denote resistance (real part) and reactance (imaginary part) of the corresponding line.

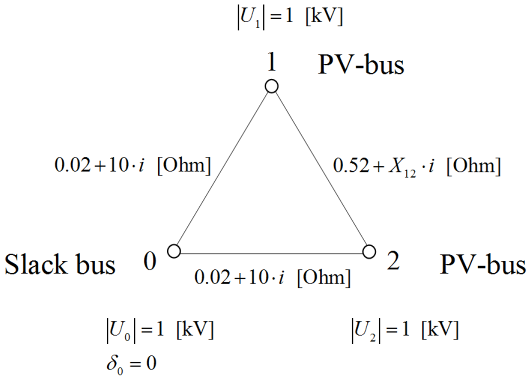


Fig. 1. 3-bus test network.

Plots, presented in the paper (Figures 2 and 3), demonstrate the evolution of the power flow feasibility domain's non-convexities and the shape of Jacobian singularity curves, according to the increasing value of reactance of the line, which connects PV-buses (X_{12} in Figure 1). Data for plotting were obtained via application of Maple tools for nonlinear equation solving (one of the possible alternative approaches implies the usage of Euler homotopy method [5]).

Further research efforts may include implementation of analytical results, concerning sufficient conditions of convexity or non-convexity [6], for the regarded and similar examples, developing algorithms for non-convex parts approximation and investigation of technical and security constraints' impact on the geometry of reduced feasibility set (e.g. in optimal power flow problems).

References

1. Bogdanov I.P. Mathematical methods for computing the extremal power values in the nodes of ac electric power systems // Computational Mathematics and Modeling. 2012. V. 23, N 2. P. 175–194.
2. Makarov Y.V., Dong Z.Y., Hill D.J. On convexity of power flow feasibility boundary // IEEE Transactions on Power Systems. 2008. V. 23, N 2. P. 811–813.
3. Tarasov V.I. Theoretical foundations for analysis of the load-flow problem in electric power systems. Novosibirsk: Nauka, 2002. (in

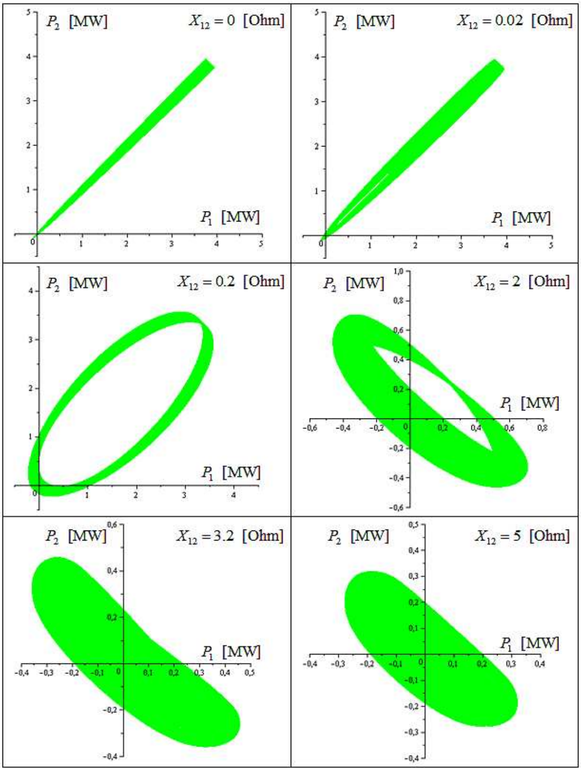


Fig. 2. Power flow feasibility sets.

Russian)

4. Vasin V.P. Region of existence of an electrical system steady-state regime. Moscow: Moscow Power Engineering Institute, 1982. (in Russian)
5. Hiskens I.A., Davy R.J. Exploring the power flow solution space boundary // IEEE Transactions on Power Systems. 2001. V. 16, N 3. P. 389–395.
6. Polyak B.T., Gryazina E.N. Convexity/nonconvexity certificates for power flow analysis // Advances in Energy System Optimization. Springer (Trends in Mathematics series), 2017. P. 221–230.

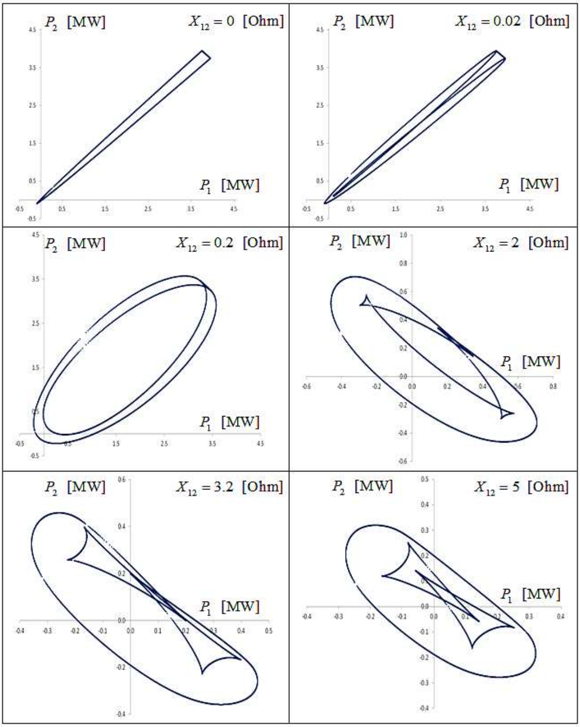


Fig. 3. Jacobian singularity curves.

On solving saddle point problems and monotone equations

O. Burdakov
Linköping University, Linköping, Sweden

We consider the problem of finding the saddle point $z^* = [x^*, y^*] \in R^n$ of function $f(x, y)$, which is assumed to be sufficiently smooth, strongly convex in $x \in R^{n_x}$ and strongly concave in $y \in R^{n_y}$, where $n_x + n_y = n$. By the definition, the saddle point z^* is required to satisfy the inequalities

$$f(x^*, y) \leq f(x^*, y^*) \leq f(x, y^*), \quad \forall x \in R^{n_x}, y \in R^{n_y}.$$

The assumptions concerning $f(x, y)$ guarantee that z^* exists, and it is unique.

In the case when $n_y = 0$, the variable y vanishes in f, and then the saddle point problem is reduced to minimization of $f(x)$ in $x \in R^{n_x}$. This allows us to view the approaches presented here as an extension of those available in the unconstrained optimization.

Denote $F(z) = E\nabla f(z)$, where

$$E = \begin{bmatrix} I_{n_x} & 0 \\ 0 & -I_{n_y} \end{bmatrix}.$$

Since the function $f(z)$ is strongly convex-concave, the mapping F is strongly monotone, that is there exists a scalar $c > 0$ such that

$$\langle F(u) - F(v), u - v \rangle \geq c\|u - v\|^2, \quad \forall u, v \in R^n.$$

Furthermore, the saddle point problem is equivalent to solving the system of nonlinear equations $F(z) = 0$.

In [1, 2], iterative processes of the form

$$z_{k+1} = z_k + \alpha_k p_k$$

were considered. Here $p_k \in R^n$ is a search direction, and the step length α_k is obtained by solving the equation

$$\langle F(z_k + \alpha_k p_k), p_k \rangle = 0.$$

Since the function $f(x, y)$ is strongly convex-concave, the solution to this equation exists and unique for any nonzero vector p_k.

This orthogonality-based principle of choosing α_k was introduced in [1, 2]. It is an extension of the exact line search used in the unconstrained optimization. In the saddle search problem, the resulting point z_{k+1} provides in the line $z_k + \alpha p_k$ a kind of a unique *trade-off* for x and y in the following sense.

We use the partitioning

$$p_k = [p_x, p_y] \quad \text{and} \quad \nabla_x f(z_{k+1}) = [\nabla_x f, \nabla_y f],$$

where $p_x, \nabla_x f \in R^{n_x}$ and $p_y, \nabla_y f \in R^{n_y}$. Suppose that $\langle \nabla_x f, p_x \rangle \neq 0$. Then $\langle \nabla_y f, p_y \rangle \neq 0$, because $E\nabla f(z_{k+1}) \perp p_k$. Given $\varepsilon > 0$, consider the two problems

$$f_x^* = \min_{t \in [-\varepsilon, \varepsilon]} f(x_{k+1} + tp_x, y_{k+1})$$

and

$$f_y^* = \min_{t \in [-\varepsilon, \varepsilon]} f(x_{k+1}, y_{k+1} + tp_y).$$

It can be easily seen that, for all sufficiently small $\varepsilon > 0$ the two optimal values of t are located in the opposed ends of the interval $[-\varepsilon, \varepsilon]$. Furthermore,

$$f_x^* = f(z_{k+1}) - \varepsilon |\langle p_x, \nabla_x f \rangle| + o(\varepsilon^2)$$

and

$$f_y^* = f(z_{k+1}) + \varepsilon |\langle p_y, \nabla_y f \rangle| + o(\varepsilon^2).$$

Thus, the gain in minimizing $f(x, y_{k+1})$ along p_x is equal to the gain in maximizing $f(x_{k+1}, y)$ along p_y to the first-order approximation. This means that the orthogonality-based line search provides in the resulting point z_{k+1} a kind of 'equal rights' for a local minimization over p_x and a local maximization over p_y.

The orthogonality-based line search works well for the Newton search direction $p_k = -(f''(z_k))^{-1} \nabla f(z_k)$, because it inherits the same quadratic rate of convergence to z^* (see [1, 2]) as the Newton method with the exact line search in the unconstrained optimization.

Consider the case when $f(z)$ is a quadratic function, or equivalently, when $F(z)$ is a linear mapping. Denote $A = Ef'' = F'(z)$. The matrix A is non-symmetric and positive definite. This feature was exploited in [1, 2]. Following [3], we call non-zero vectors $p_0, p_1, \ldots, p_m$ *A-pseudo-orthogonal* if

$$\langle Ap_j, p_i \rangle = 0, \quad \forall i, j, \ 0 \le i < j \le m.$$

They are linearly independent. The aforementioned iterative process with the orthogonality-based line search results in the orthogonality

$$\langle F(z_{m+1}), p_k \rangle = 0, \quad \forall k, \ 0 \le k \le m.$$

For any n A-pseudo-orthogonal directions, this property implies convergence to the saddle point z^* in at most n iterations. When $n_y = 0$, the exact line search is performed along conjugate directions.

In [1, 2], some conjugate direction methods, which are used for unconstrained minimization, were extended to finding saddle points of non-quadratic functions $f(x, y)$ and to solving systems of non-linear equations $F(z) = 0$, where F is a monotone mapping. It was proved in [1, 2] that the rate of convergence of these extensions is quadratic.

Here we survey the aforementioned approaches. Based on them, we developed a new limited memory conjugate-direction-type algorithm for solving large-scale linear systems originating from saddle point problems and monotone equations. The memory size is a parameter which can be tuned for a certain class of problems to be solved. Results of numerical experiments are presented. They show that the new algorithm is competing with the most efficient of the existing algorithms.

In the Barzilai-Borwein (BB) unconstrained minimization algorithm [4], a step is made along the steepest descent direction $-\nabla f(x_k)$ with the step size computed by a special formula. This algorithm is often used for solving large scale minimization problems. It competes with the conjugate gradient algorithm. Here we extend the BB algorithm to solving saddle point problems and nonlinear monotone equations. In the extended algorithm, a proper step is made along the direction $-F(z_k)$. We present results of numerical experiments for this algorithm in solving problems for linear and non-linear mapping F.

References

1. Burdakov O.P. Conjugate direction methods for solving systems of equations and finding saddle points. I // Engineering Cybernetics. 1982. V. 20, N 3. P. 13–19.
2. Burdakov O.P. Conjugate direction methods for solving systems of equations and finding saddle points. II // Engineering Cybernetics. 1982. V. 20, N 4. P. 23–32.
3. Voevodin V.V. On methods of conjugate direction // USSR Computational Mathematics and Mathematical Physics. 1979. V. 19, N 5. P. 228–233.
4. Barzilai J., Borwein J.M. Two-point step size gradient methods // IMA Journal of Numerical Analysis. 1988. V. 8, N 1. P. 141–148.

Variations of the v-change of time in problems with state constraints*

A.V. Dmitruk and N.P. Osmolovskii

Central Economics and Mathematics Institute of RAS, Moscow,
University of Technology and Humanities, Radom, Poland

We give a new proof of the maximum principle for optimal control problems with running state constraints. The proof uses the so-called method of $v-$change of the time variable introduced by Dubovitskii

*This research was supported by the RFBR grants 16-01-00585 and 17-01-00805.

and Milyutin. In this method, the time t is considered as a new state variable satisfying the equation $dt/d\tau = v$, where $v(\tau) \geqslant 0$ is a new control and τ a new time. Unlike the general $v-$change with an arbitrary $v(\tau)$, we use a piecewise constant $v(\tau)$. Every such $v-$change reduces the original problem to a problem in a finite dimensional space, with a continuum number of inequality constrains corresponding to the state constraints. The stationarity conditions in every new problem, being written in terms of the original time t, give a weak* compact set of normalized tuples of Lagrange multipliers. The family of these compacta is partially ordered by inclusion and possesses a maximal element. An arbitrary tuple of Lagrange multipliers belonging to the latter ensures the maximum principle.

The idea of v-change of the time variable, proposed in [1], is as follows. Consider the time t, varying in an interval $[t_0, t_1]$, as a new state variable $t = t(\tau)$ that depends on a new time τ, varying in an interval $[\tau_0, \tau_1]$. Let the function $t(\tau)$ satisfy the equation

$$\frac{dt(\tau)}{d\tau} = v(\tau), \quad t(\tau_0) = t_0, \quad v(\tau) \geqslant 0 \quad \text{a.e. in} \quad [\tau_0, \tau_1],$$

where the function $v(\tau)$ is a new control, measurable and essentially bounded. It follows that $t(\tau)$ is a nondecreasing function, mapping a time interval $[\tau_0, \tau_1]$ onto the original time interval $[t_0, t_1]$. The function $t(\tau)$ enables to match for any control $u(t)$ a new control $\tilde{u}(\tau) = u(t(\tau))$ on the set $M_+ := \{\tau \in [\tau_0, \tau_1] \,:\, v(\tau) > 0\}$, while on the set $M_0 := \{\tau \in [\tau_0, \tau_1] \,:\, v(\tau) = 0\}$ the new control can be defined arbitrarily, with the only condition that $u(\tau) \in U$. Assume e.g. that $v(\tau)$ vanishes just on an interval $[\tau', \tau''] \subset [\tau_0, \tau_1]$ of a positive measure. Then we can put $\tilde{u}(\tau) \equiv u_*$ on this interval, where $u_* \in U$ is an arbitrary fixed value, while on the complement to $[\tau', \tau'']$ we can use the formula $\tilde{u}(\tau) = u(t(\tau))$. Thus, we obtain a new control $\tilde{u}(\tau)$, while a new trajectory $\tilde{x}(\tau)$ is defined simply as $\tilde{x}(\tau) = x(t(\tau))$ for all $\tau \in [\tau_0, \tau_1]$. In this way, we can transform the initial optimal control problem with the independent variable t into a new problem, corresponding to this change, with the independent variable τ. Moreover, one can easily show that, under this transformation, any optimal process of the initial problem transforms into an optimal process of the new problem.

Furthermore, it is possible to return from the variable τ to the original variable t using the "inverse" (to be more precise, right inverse) change. Namely, for any $t \in [t_0, t_1]$, let $\tau(t)$ be the smallest root of the equation $t(\tau) = t$. Then, obviously, $t(\tau(t)) \equiv t$, $\tilde{u}(\tau(t)) = u(t)$, and $\tilde{x}(\tau(t)) = x(t)$

on $[t_0, t_1]$.

It might seem that the selected value u_* does not play any role in such transition, since the interval $[\tau', \tau'']$ maps into a single point $t_* \in [t_0, t_1]$. However, this is true only for the given control $v(\tau)$. But what happens if the function $v(\tau)$ is perturbed by a uniformly small variation $\bar{v}(\tau)$ such that still $v(\tau) + \bar{v}(\tau) \geqslant 0$ a.e. in $[\tau_0, \tau_1]$? Then the interval $[\tau', \tau'']$ maps onto a small interval $[t', t'']$ with $u(t) = u_*$ on it, so we obtain a needle shape variation of the original control! It is this fact that was used by Dubovitskii and Milyutin to obtain the MP by means of the v-change [1], and they systematically used this approach in other works. The v-change turned out to be a very powerful tool for obtaining necessary optimality conditions in the form of MP.

Note that though the v-variations are much alike the needle variations, they have some advantages. Whereas the usage of needle variations requires the assumption of piecewise continuity of the optimal control, for the usage of v-variations this assumption is not needed, and the optimal control can be an arbitrary measurable bounded function. Moreover, the v-variations generate a smooth control system, well defined for $v(\tau)$ of arbitrary sign, where the requirement $v(\tau) \geqslant 0$ can be regarded as a separate standard constraint, while the needle variations can be considered only for nonnegative widths of the needles, so one obtains functions defined just on the positive orthant in a finite-dimensional space, which is not convenient to differentiate. By these reasons, the v-change of time is a more preferable tool in research of optimality than the needle variations.

Note however, that for an arbitrary nonnegative $v(\tau)$, the v-change is rather complicated technically. To make the proof of MP more simple, A.A. Milyutin proposed in 2001 to use the v-change with a *piecewise constant* function $v(\tau)$ This idea was realized in [2, 3]. It gave a quite simple proof of MP for a general optimal control problem of the Pontryagin type, i.e. without state constraints. Such a primitive v-change allowed to pass to a family of smooth optimization problems in finite dimensional spaces (each of which corresponds to the parameters of the given v-change), i.e. to problems of mathematical programming, and then to use the well-known necessary optimality conditions in each problem. A proper arranging of the obtained family of optimality conditions made it possible to pass from them to one universal condition, which had the form of MP. We show that this approach works also for problems with state constraints [4], of course, with some additional technical results, especially in the case of several state constraints.

One of those results relates to the fact that the state constraints present an infinite number of inequality constraints on the parameters of the problem, so the latter is not a standard smooth problem of mathematical programming. Such problems are now called semi-infinite. However, they admit a simple version of a generalized Lagrange multipliers rule, involving Lebesgue–Stieltjes measures as multipliers at the state constraints [5]. Another result concerns the compactness of the tuples of Lagrange multipliers in each such problem, where we should consider those measures in the weak-* topology and use the Helly theorems on their convergence.

References

1. Dubovitskii A.Ya. and Milyutin A.A. Extremum problems in the presence of restrictions // USSR Comput. Math. and Math. Phys. 1965. v. 5, no. 3, p. 1–80.
2. Milyutin A.A., Dmitruk A.V., Osmolovskii N.P. Maximum principle in optimal control. Moscow: Mech-Math Faculty of Moscow State Univ., 2004 (in Russian).
3. Dmitruk A.V., Osmolovskii N.P. On the Proof of Pontryagin's Maximum Principle by Means of Needle Variations // J. of Math. Sciences. 2016. Vol. 218, no. 5, p. 581–598.
4. Dmitruk A.V., Osmolovskii N.P. Variations of the type of $v-$change of time in problems with state constraints // Proc. Inst. of Math. and Mech., the Ural Branch of RAS. 2018. Vol. 24, no. 1 (in Russian).
5. Dmitruk A.V., Osmolovskii N.P. A General Lagrange Multipliers Theorem // In: "Constructive Nonsmooth Analysis and Related Topics (CNSA-2017)". IEEE Xplore Dig. Lib., 2017. DOI: 10.1109/CNSA.2017.7973951

Critical solutions of nonlinear equations*

A.F. Izmailov, A.S. Kurennoy, and M.V. Solodov
Moscow State University, Moscow, Russia,
Tambov State University, Tambov, Russia,
IMPA, Rio de Janeiro, Brazil

1. Introduction and the key assumption

We consider a nonlinear equation

$$\Phi(u) = 0,$$

where $\Phi : \mathbb{R}^p \to \mathbb{R}^p$ is a given smooth mapping. Even though here we deal solely with the case when the number of equations is the same as the number of variables, it will never be assumed that the solution in question is isolated, and moreover, the case of nonisolated solutions will be of special interest. Observe that any nonisolated solution $\bar{u}$ is necessarily singular in a sense that $\Phi'(\bar{u})$ is a singular matrix.

In [7], it was shown that a solution $\bar{u}$ of the unconstrained equation "survives" perturbations in large classes if Φ is smooth enough, and there exists $\bar{v} \in \ker \Phi'(\bar{u})$ such that Φ is 2-regular at $\bar{u}$ in the direction $\bar{v}$, the latter meaning that

$$\operatorname{im} \Phi'(\bar{u}) + \Phi''(\bar{u})[\bar{v}, \ker \Phi'(\bar{u})] = \mathbb{R}^p.$$

Importantly, such $\bar{v}$ may exist even if $\bar{u}$ is a nonisolated solution.

Furthermore, as demonstrated in [7], if $\bar{u}$ is a singular solution, the needed $\bar{v}$ cannot belong to $T_{\Phi^{-1}(0)}(\bar{u})$, and hence it can never exist if $T_{\Phi^{-1}(0)}(\bar{u}) = \ker \Phi'(\bar{u})$. The latter is one of the two ingredients of the concept of noncriticality of solution $\bar{u}$, as introduced in [7] The second ingredient is Clarke regularity of $\Phi^{-1}(0)$ at $\bar{u}$, and as demonstrated in [7], under the appropriate smoothness assumptions, this combination of properties is equivalent to the local Lipschitzian error bound

$$\operatorname{dist}(u, \Phi^{-1}(0)) = O(\|\Phi(u)\|) \quad \text{as } u \in \mathbb{R}^p \text{ tends to } \bar{u},$$

which is known to be equivalent to the following upper Lipschitzian property:

$$\operatorname{dist}(u(w), \Phi^{-1}(0)) = O(\|w\|) \quad \text{as } w \in \mathbb{R}^p \text{ tends to } 0,$$

*Research of A.F. Izmailov was supported by the Russian Science Foundation Grant 17-11-01168 (Section 2). A.S. Kurennoy was supported by the Russian Foundation for Basic Research Grant 17-01-00125. M.V. Solodov was supported in part by CNPq Grant 303724/2015-3, and by FAPERJ Grant 203.052/2016.

where $u(w)$ is any solution of the perturbed equation

$$\Phi(u) = w,$$

close enough to $\bar{u}$. In addition, the results in [7] imply that singular noncritical solutions of the unconstrained equation can only be stable subject to very special perturbations. At the same time, critical solutions (i.e., those which are not noncritical), or, more precisely, those solutions for which $T_{\Phi^{-1}(0)}(\bar{u})$ is a proper subset of $\ker \Phi'(\bar{u})$, can naturally satisfy the 2-regularity condition with some $\bar{v} \in \ker \Phi'(\bar{u})$, and hence, be stable subject to wide classes of perturbations.

In this work, we demonstrate that 2-regularity in a direction $\bar{v} \in \ker \Phi'(\bar{u})$ (which is our key assumption, and which may never hold at noncritical singular solutions, as discussed above) makes $\bar{u}$ specially attractive for sequences generated by Newton-type methods. Apart form the basic Newton method (NM), we will consider some modifications of it, intended specially for tackling the case of nonisolated solution. Specifically, these are the Levenberg–Marquardt method (L-MM) and the LP-Newton method (LP-NM).

2. Perturbed Newton method

We define the perturbed Newton method (pNM) for equation in question as follows: for a current iterate $u^k \in \mathbb{R}^p$, the next iterate is $u^{k+1} = u^k + v^k$, with v^k computed as a solution of linear equation

$$\Phi(u^k) + (\Phi'(u^k) + \Omega(u^k))v = \omega(u^k),$$

where $\Omega : \mathbb{R}^p \to \mathbb{R}^{p \times p}$ and $\omega : \mathbb{R}^p \to \mathbb{R}^p$ characterizes perturbation.

The following can be regarded as an extension of [5, Lemma 5.1] from the basic NM to pNM.

Every $u \in \mathbb{R}^p$ is uniquely decomposed into the sum $u = u_1 + u_2$, $u_1 \in (\ker \Phi'(\bar{u}))^{\perp}$, $u_2 \in \ker \Phi'(\bar{u})$. Let Π be the orthogonal projector onto $(\operatorname{im} \Phi'(\bar{u}))^{\perp}$, and assume that the norm is Euclidian. Let $\mathbf{S}$ stand for the unit sphere in $\mathbb{R}^p$.

Theorem. *Let Φ be twice differentiable near $\bar{u} \in \mathbb{R}^p$, with its second derivative Lipschitz-continuous with respect to $\bar{u}$. Let $\bar{u}$ be a solution of the nonlinear equation in question, and assume that Φ is 2-regular at $\bar{u}$ in a direction $\bar{v} \in \ker \Phi'(\bar{u}) \cap \mathbf{S}$. Let $\Omega : \mathbb{R}^p \to \mathbb{R}^{p \times p}$ and $\omega : \mathbb{R}^p \to \mathbb{R}^p$ satisfy the estimates*

$$\Omega(u) = O(\|u - \bar{u}\|), \quad \Pi\Omega(u) = O(\|u_1 - \bar{u}_1\|) + O(\|u - \bar{u}\|^2),$$

$$\omega(u) = O(\|u-\bar{u}\|^2), \quad \Pi\omega(u) = O(\|u-\bar{u}\|\|u_1-\bar{u}_1\|) + O(\|u-\bar{u}\|^3).$$

Then there exist $\varepsilon = \varepsilon(\bar{v}) > 0$ *and* $\delta = \delta(\bar{v}) > 0$ *such that any starting point* $u^0 \in \mathbb{R}^p \setminus \{\bar{u}\}$ *satisfying*

$$\|u^0-\bar{u}\| \leqslant \varepsilon, \quad \left\|\frac{u^0-\bar{u}}{\|u^0-\bar{u}\|} - \bar{v}\right\| \leqslant \delta$$

uniquely defines the sequence $\{u^k\} \subset \mathbb{R}^p$ *of the pNM,* $u_2^k \neq \bar{u}_2$ *for all* k*, the sequence* $\{u^k\}$ *converges to* $\bar{u}$*, and*

$$\lim_{k\to\infty} \frac{\|u^{k+1}-\bar{u}\|}{\|u^k-\bar{u}\|} = \frac{1}{2}, \quad \|u_1^{k+1}-\bar{u}_1\| = O(\|u^k-\bar{u}\|^2).$$

Theorem above establishes the existence of a set with nonempty interior, which is star-like with respect to $\bar{u}$, and such that the pNM initialized at any point of this set converges linearly to $\bar{u}$. Moreover, if Φ is 2-regular at $\bar{u}$ in at least one direction $\bar{v} \in \ker \Phi'(\bar{u})$, then set of such $\bar{v}$ is open and dense in $\ker \Phi'(\bar{u}) \cap \mathbf{S}$: its complement is the null set of the nontrivial homogeneous polynomial $\det \Pi\Phi''(\bar{u})[\cdot]|_{\ker \Phi'(\bar{u})}$ considered on $\ker \Phi'(\bar{u}) \cap \mathbf{S}$. The union of convergence domains coming with all such $\bar{v}$ is also a star-like convergence domain with nonempty interior. In the case when $\Phi'(\bar{u}) = 0$ (full singularity) this domain is quite large. In particular, it is “asymptotically dense”: its complement is “asymptotically thin”, and the only excluded directions are those in which Φ is not 2-regular at $\bar{u}$, which is the null set of a nontrivial homogeneous polynomial.

The assumptions on perturbations in Theorem automatically hold if

$$\Omega(u) = O(\|\Phi(u)\|), \quad \omega(u) = O(\|u-\bar{u}\|\|\Phi(u)\|).$$

3. Levenberg–Marquardt method

The L-MM is a well-established tool for tackling possibly nonisolated solutions. The iteration subproblem of this method has the form

$$\text{minimize} \quad \frac{1}{2}\|\Phi(u^k) + \Phi'(u^k)v\|^2 + \frac{1}{2}\sigma(u^k)\|v\|^2, \quad v \in \mathbb{R}^p,$$

where $\sigma : \mathbb{R}^p \to \mathbb{R}_+$ defines the regularization parameter. In particular, from the results in [8] it follows that being initialized near a noncritical solution, the L-MM with $\sigma(u) = \|\Phi(u)\|^2$ generates a sequence which is quadratically convergent to a (nearby) solution.

The L-MM subproblem is equivalent to the linear system

$$(\Phi'(u^k))^{\mathrm{T}}\Phi(u^k) + ((\Phi'(u^k))^{\mathrm{T}}\Phi'(u^k) + \sigma(u^k)I)v = 0,$$

characterizing stationary points of that convex optimization problem.

From [5, Lemma 3.1] it can be seen that $\bar{v}$ in Theorem applied to *the basic NM* comes with a “conic neighborhood” such that for every u in it, $\Phi'(u)$ is invertible, and $(\Phi'(u))^{-1} = O(\|u - \bar{u}\|^{-1})$. Multiplying both sides of the iteration system by $(((\Phi'(u^k))^{\mathrm{T}})^{-1} = (((\Phi'(u^k))^{-1})^{\mathrm{T}}$, we now obtain

$$\Phi(u^k) + (\Phi'(u^k) + \sigma(u^k)(((\Phi'(u^k))^{-1})^{\mathrm{T}})v = 0,$$

which is the pNM iteration system with the perturbation terms

$$\Omega(u) = \sigma(u)(((\Phi'(u))^{-1})^{\mathrm{T}} = O(\|u - \bar{u}\|^{-1}\sigma(u)), \quad \omega \equiv 0$$

as $u \to \bar{u}$, and therefore, the needed requirements on Ω will hold, e.g., if $\sigma(u) = \|\Phi(u)\|^{\tau}$ with $\tau \geqslant 2$.

Moreover, in the case when $\Phi'(\bar{u}) = 0$ (full singularity) the appropriate values are all $\tau \geqslant 3/2$.

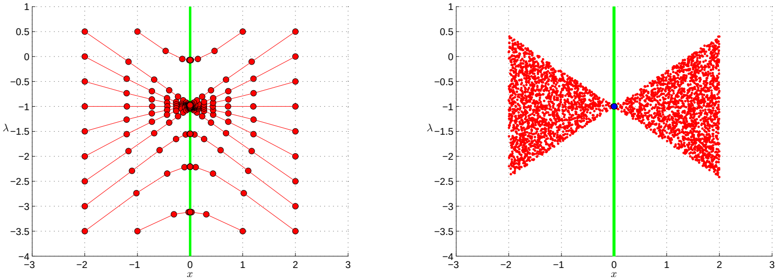


Fig. 1. Levenberg–Marquardt method with $\tau = 1$.

Example. Consider the equality-constrained optimization problem

$$\text{minimize} \quad x^2 \quad \text{subject to} \quad x^2 = 0.$$

Stationary points and associated Lagrange multipliers of this problem are characterized by the Lagrange optimality system which has the form

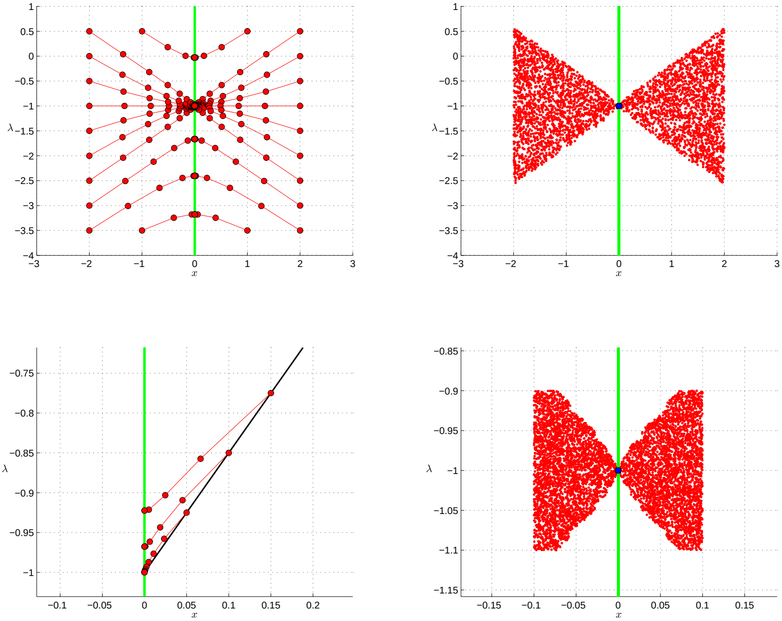


Fig. 2. Levenberg–Marquardt method with $\tau = 3/2$.

of a nonlinear equation with $\Phi : \mathbb{R}^2 \to \mathbb{R}^2$, $\Phi(u) = (2x(1+\lambda),\, x^2)$, where $u = (x,\, \lambda)$. The unique feasible point (hence, the unique solution, and the unique stationary point) of this problem is $\bar{x} = 0$, and the set of associated Lagrange multipliers is the entire $\mathbb{R}$. Therefore, the solution set of the Lagrange system (i.e., the primal-dual solution set) is $\{\bar{x}\} \times \mathbb{R}$. The unique critical solution is $\bar{u} = (\bar{x},\, \bar{\lambda})$ with $\bar{\lambda} = -1$, the one for which $\Phi'(\bar{u}) = 0$ (full singularity).

In Figures 1 and 2, the vertical gray line corresponds to the primal-dual solution set. These figures demonstrate some iterative sequences generated by the L-MM, and the domains from which convergence to the critical solution was detected.

4. LP-Newton method

A more recent approach, alternative to the L-MM, is the LP-NM proposed in [2]. The iteration subproblem of this method has the form

$$\begin{array}{ll}\text{minimize} & \gamma\\ \text{subject to} & \|\Phi(u^k)+\Phi'(u^k)v\| \leqslant \gamma\|\Phi(u^k)\|^2,\\ & \|v\| \leqslant \gamma\|\Phi(u^k)\|,\\ & (v,\,\gamma)\in\mathbb{R}^p\times\mathbb{R}.\end{array}$$

With l_∞ norm, this is a linear programming problem. As demonstrated in [2, 3], local convergence properties of the LP-NM (near noncritical solutions!) are the same as for L-MM.

The first constraint in the LP-NM subproblem can be interpreted as the pNM with $\Omega\equiv 0$ and some $\omega(\cdot)$, which will satisfy the assumptions in Theorem if the optimal value $\gamma(u)$ of the LP-NM subproblem with $u^k=u$ satisfies

$$\gamma(u)=O(\|\Phi(u)\|^{-1}\|u-\bar{u}\|)$$

as $u\to\bar{u}$.

In order to establish the needed estimate on $\gamma(\cdot)$, suppose again that u belongs to the "conic neighborhood" of $\bar{v}$ where *the basic NM* step $v(u)$ is uniquely defined, and $v(u)=O(\|u-\bar{u}\|)$. Then the point $(v,\,\gamma)=(v(u),\,\|v(u)\|/\|\Phi(u)\|)$ is feasible in the LP-N subproblem, and hence,

$$\gamma(u)\leqslant\gamma=\|\Phi(u)\|^{-1}\|v(u)\|=O(\|\Phi(u)\|^{-1}\|u-\bar{u}\|)$$

as $u\to\bar{u}$.

Figure 3 has the same meaning as Figure 2, but for LP-NM instead of L-MM, with the same conclusions.

A detailed exposition of these results can be found in [6]. The extensions of these and related results to constrained equations can be found in [1], [4].

References

1. A.V. Arutyunov and A.F. Izmailov Stability of possibly nonisolated solutions of constrained equations, with applications to complementarity and equilibrium problems. Set-Valued Var. Analys. 26 (2018), 327–352.
2. F. Facchinei, A. Fischer, and M. Herrich. An LP-Newton method: Nonsmooth equations, KKT systems, and nonisolated solutions. Math. Program. 146 (2014), 1–36.
3. F. Facchinei, A. Fischer, and M. Herrich. A family of Newton methods for nonsmooth constrained systems with nonisolated solutions. Math. Methods Oper. Res. 77 (2013), 433–443.

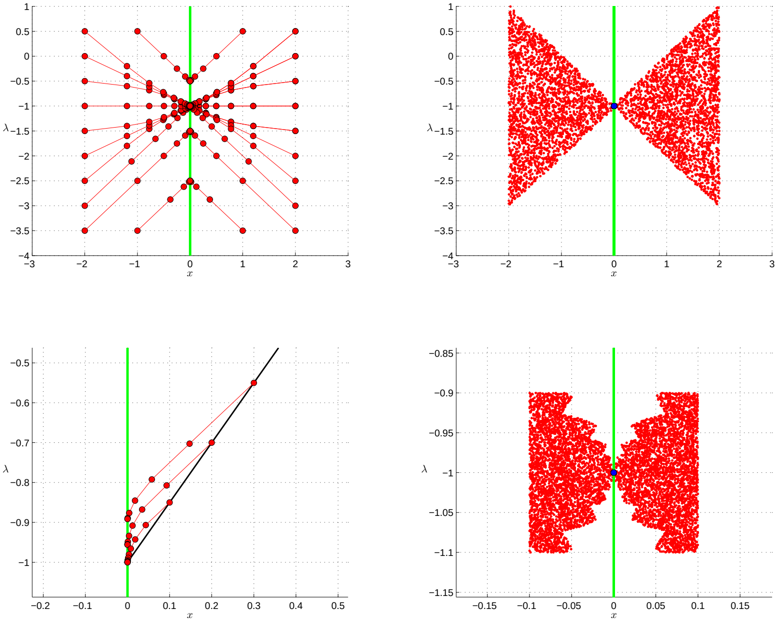


Fig. 3. LP-Newton method.

4. A. Fischer, A.F. Izmailov, and M.V. Solodov. Local attractors of Newton-type methods for constrained equations and complementarity problems with nonisolated solutions. J. Optim. Theory Appl. 2018. DOI 10.1007/s10957-018-1297-2.

5. A. Griewank. Starlike domains of convergence for Newton's method at singularities. Numer. Math. 35 (1980), 95–111.

6. A.F. Izmailov, A.S. Kurennoy, and M.V. Solodov. Critical solutions of nonlinear equations: local attraction for Newton-type methods. Math. Program. 167 (2018), 355–379.

7. A.F. Izmailov, A.S. Kurennoy, and M.V. Solodov. Critical solutions of nonlinear equations: stability issues. Math. Program. 168 (2018), 475–507.

8. N. Yamashita and M. Fukushima. On the rate of convergence of the Levenberg–Marquardt method. Computing. Suppl. 15 (2001), 237–249.

SPRG method for large-scale non-convex optimization problem with linear constraints

A. El Mouatasim[1] and J. Oudaani[2]

[1,2] *Laboratory LabSI, Ibn Zohr University - FPO,*
Ouarzazate 45800, Morocco

Introduction

In this work, the following nonconvex optimization problem with linear constraints (NCOPLC) is considered:

$$\begin{cases} \text{Minimize} & f(\mathbf{x}) \\ \text{subject to} & A\mathbf{x} = b, \\ & \mathbf{x} \geq 0, \end{cases} \tag{1}$$

where $f(.)$ is twice continuously differentiable on $I\!R^n$, A is an $m \times n$ matrix with full row rank and $b \in I\!R^m$.

There exist several application areas for NCOPLCs like pumping water operation, optimal control, machine learning see for instance [1,2].

Many algorithms have been proposed for solving NCOPLCs, such as the reduced gradient method [1] consisting of solving a sequence of subproblems in which the number of variables is implicitly reduced. These reduced problems are obtained by using the linear constraints to express certain variables, designated as 'basic', as functions of other variables.

We are mainly interested in the situation of NCOPLC where, on one hand, f is nonconvex and, on the other hand, the rank of matrix A can be less than or equal to m.

In this work, we propose an algorithm of stochastic perturbation of reduced gradient (SPRG) method for optimizing a large-scale non convex differentiable function subject to linear equality constraints and nonnegativity bounds on the variables. In particular, at each iteration, we compute a search direction by reduced gradient and optimal stepsize by bisection algorithm.

Reduced gradient method

We introduce basic and nonbasic variables according to

$$A = [B, N], \qquad \mathbf{x} = \begin{bmatrix} \mathbf{x}_B \\ \mathbf{x}_N \end{bmatrix}, \qquad \mathbf{x}_N \geq 0 \quad \text{and} \qquad \mathbf{x}_B > 0. \tag{2}$$

We denote by I_B and I_N the index sets of basic variables and nonbasic variables, respectively.

The reduced gradient method starts with a basis B and a feasible solution $\mathbf{x}^k = (\mathbf{x}_B^k, \mathbf{x}_N^k)$ such that $\mathbf{x}_B^k > 0$. The solution $\mathbf{x}$ is not necessarily a basic solution, i.e. $\mathbf{x}_N$ does not have to be identically zero. Such a solution can be obtained e.g. by the usual first phase procedure of simplex method. Using the basis B form $B\mathbf{x}_B + N\mathbf{x}_N = b$, we have

$$\mathbf{x}_B = B^{-1}b - B^{-1}N\mathbf{x}_N,$$

hence the basic variables $\mathbf{x}_B$ can be eliminated from the problem (1):

$$\begin{cases} \text{Minimize } \; f_N(\mathbf{x}_N) \\ \text{subject to } \; B^{-1}b - B^{-1}N\mathbf{x}_N > 0, \\ \qquad\qquad 0 \le \mathbf{x}_N, \end{cases} \tag{3}$$

where $f_N(\mathbf{x}_N) = f(\mathbf{x}) = f(B^{-1}b - B^{-1}N\mathbf{x}_N, \mathbf{x}_N)$. Using the notation

$$\nabla f(\mathbf{x})^t = \left[\nabla_B f(\mathbf{x})^t, \nabla_N f(\mathbf{x})^t\right],$$

the gradient of f_N, which is the so-called *reduced gradient*, can be expressed as

$$\nabla f_N(\mathbf{x})^t = -\big(\nabla_B f(\mathbf{x})^t B^{-1}N\big) + \big(\nabla_N f(\mathbf{x})^t\big).$$

Now let us assume that the basis is nondegenerate, i.e. only the nonnegative constraints $\mathbf{x}_N \ge 0$ might be active at the current iterate $\mathbf{x}^k$. Let the search direction be a vector $\mathbf{d}^t = (\mathbf{d}_B^t, d_N^t)$ in the null space of the matrix A defined as $\mathbf{d}_B = -B^{-1}N\mathbf{d}_N$ and $\mathbf{d}_N \ge 0$. If we define so, then the feasibility of $\mathbf{x}^k + \eta\mathbf{d}$ is guaranteed as long as $\mathbf{x}_B^k + \eta\mathbf{d}_B \ge 0$, i.e. as long as

$$\eta \le \eta_{max} = \min_{i \in B, d_i < 0} \{\frac{-x_i^k}{d_i}\}.$$

We still need to define $\mathbf{d}_N \ge 0$ such that it is a descent direction of f_N projected to the coordinate hyperplane active at the current point $\mathbf{x}_N^k$. So we have

$$d_j^k = \begin{cases} 0 & \text{if } \; x_j^k = 0 \text{ and } \dfrac{\partial f_N(\mathbf{x}_N^k)}{\partial x_j} \ge 0, \\ -\dfrac{\partial f_N(\mathbf{x}_N^k)}{\partial x_j} & \text{otherwise,} \end{cases} \qquad j \in N.$$

Using the bisection algorithm see for instance [1] we can determine the optimal stepsize as

$$\eta_k = \underset{0\leq\eta\leq\eta_{max}}{\operatorname{argmin}} f(\mathbf{x}^k + \eta\mathbf{d}^k).$$

Algorithm of reduced gradient

Step 0 (Initialization). Choose a feasible point $\mathbf{x}^0 \in \mathbf{R}^n$ and I_B^0, I_N^0 such that B^0 is nonsingular. Set the iteration counter $k = 0$.

Step 1 (Independent variables choice). If $k \neq 0$ choose the sets I_B^k and I_N^k.

Step 2 (Compute a search direction):

1. Let $(d_N)_j = \begin{cases} 0 & \text{if } (x_N)_j = 0 \quad \text{and } r_j \geq 0, \\ -r_j & \text{otherwise.} \end{cases}$
2. If d_N is equal to zero stop, the current point is a solution. Otherwise calculate $d_B = -B^{-1}Nd_N$.

Step 3 (Compute optimal stepsize).
Calculate $\eta_{\max}$ and η_k such that

$$\eta_{\max} = \min_{i\in B, d_i<0}\{\frac{-x_i^k}{d_i}\} \text{ and } \eta_k = \underset{0\leq\eta\leq\eta_{\max}}{\operatorname{argmin}} f(x^k + \eta d^k).$$

Step 4 (Compute the next point). Put $x^{k+1} = x^k + \eta_k d^k$.

Step 5 If $\eta_k < \eta_{\max}$ return to Step 2. Otherwise update B and N.

Step 6 (Basic variables choice). Choose I_B^{k+1} and I_N^{k+1}.
Let $k = k + 1$ and go to Step 1.

Stochastic perturbation of the reduced gradient method

Since we have the linear equality constraints, the sequence of reals variables $\{x_N^k\}_{k\geq 0}$ is replaced by a sequence of random variables $\{\mathbf{X}_N^k\}_{k\geq 0}$ involving a random perturbation $\mathcal{P}_k^N$ of the deterministic iteration:

- The nonbasic random perturbation $\mathcal{P}_k^N$ for nonbasic variables is defined as

$$\mathbf{X}_N^{k+1} = \mathbf{X}_N^k + \eta_k d_N^k + \mathcal{P}_k^N. \quad (4)$$

- Using the projection, the basic random perturbation $\mathcal{P}_k^B$ for basic variables has the form

$$\mathbf{X}_B^{k+1} = B^{-1}b - B^{-1}\mathbf{N}X_N^{k+1}. \qquad (5)$$

This procedure generates a sequence $U_k = f(X^k)$ and by construction this sequence is decreasing and there exist a constant l^* such that

$$\forall k \geq 0, \quad l^* \leq U_{k+1} \leq U_k. \qquad (6)$$

Then, there exist $U \geq l^*$ such that $U_k \to U \ for \ k \to +\infty$.

Numerical experiments

In the tables presented below:

- n is the number of variables.
- k_{sto} is the number of stochastic perturbations.
- f^* is the optimal value achieved by the algorithm.

Problem 1: The problem is defined as follows:

$$\begin{cases} \text{minimize} & \sum\limits_{i=1}^{n} \cos(2\pi x_i \sin(\frac{\pi}{20})) \\ \text{subject to} & x_i - x_{i+1} = 0.4, \quad i = 1, \dots, n-1. \end{cases}$$

Problem		Algorithm			
n	m	CPU time	Iter.	k_{sto}	f^*
100	99	0.1094	3	1	-3.7032
250	249	0.1250	3	1	-4.6087
500	499	0.1875	2	1	-4.0147
1000	999	1.578	3	1	-4.9829
2500	2499	13.25	2	1	-5.0117
5000	4999	107.43	2	1	-2.0491
6000	5999	165.84	2	1	-5.0372

Table 1. Results of SPRG algorithm for problem 1.

Problem 2: Consider Rastrigin's function

$$\begin{cases} \text{minimize} & \sum\limits_{i=1}^{n}(x_i^2 - 10\cos(2\pi x_i) + 10) \\ \text{subject to} & \sum\limits_{i=1}^{n} x_i = 0, \\ & -5.12 \leq x_i \leq 5.12, \quad i = 1, 2, \ldots, n. \end{cases}$$

Problem		Algorithm			
n	m	CPU time	Iter.	k_{sto}	f^*
100	1	5.3125	46	50	2.75e-9
250	1	3.5156	20	75	9.72e-5
500	1	10.6875	55	75	3.53e-4
750	1	12.5156	15	300	3.00e-4
1000	1	24.5313	16	500	3.10e-3
1500	1	22.0781	11	550	1.14e-4
2000	1	19.3438	7	700	9.80e-3

Table 2. Results of SPRG algorithm for problem 2.

The last numerical results to the large-scale problems (NCOPLCs) which are presented, show that this approach give a efficient results.

References

1. El Mouatasim A. Implementation of reduced gradient with bisection algorithms for non-convex optimization problem via stochastic perturbation. Journal of Numerical Algorithms. 2018. V. 78, N 1. P. 41–62.
2. El Mouatasim A. Mathematical programming and application. Scholars Press, 2018.

Adaptive conditional gradient algorithm

Z.R. Gabidullina
Kazan (Volga Region) Federal University, Kazan, Russia

Our goal is to study the following problem:

$$\min_{x \in D} f(x), \tag{1}$$

where $f(x)$ - is a continuously differentiable pseudoconvex function satisfying the so-called Condition A (introduced in [1]) on a convex closed and bounded subset D of Euclidean space $\mathbb{R}^n$. For solving this problem, we present a new efficient algorithm which has the estimates of the rate of its convergence and allows adaptive controlling both the parameter of an $\varepsilon-$normalization of a descent direction and the step length.

We start with some notations: $g(x)$ is the gradient of the function $f(x)$ at the point x, x_0 stands for the starting point of the iterative sequence constructed by minimizing the objective function, $f^* = \min\limits_{x \in D} f(x)$, $D^* = \{x \in D : f(x) = f^*\}$, $L = \{0, 1, \ldots\}$, and p_k^* corresponds to a projection of the iterative point x_k on the set D^*, $k \in L$.

To the best of our knowledge, the use of the extension of smooth convex functions, namely the continuously differentiable pseudoconvex ones, was pioneered by Mangasarian in [2].

Definition 1 *A function $f(x)$ given and continuously differentiable on an open convex set G from $\mathbb{R}^n$ is called a pseudoconvex one if for $\forall\, x, y \in G$ it holds the following implication:*

$$\langle g(x), y - x\rangle \geq 0 \Rightarrow f(x) \leq f(y),$$

or the equivalent implication:

$$f(y) < f(x) \Rightarrow \langle g(x), y - x\rangle < 0.$$

Definition 2 *We say that a continuous function $f(x)$ satisfies Condition A on the convex set $D \subseteq \mathbb{R}^n$ if there exist a positive constant μ and a nonnegative symmetric function $\tau(x, y)$ such that*

$$f(\alpha x + (1-\alpha)y) \geq \alpha f(x) + (1-\alpha) f(y) - \alpha(1-\alpha)\mu\tau(x, y),$$

$\forall x, y \in D, \alpha \in [0, 1]$.

For x, $y \in D \subseteq \mathbb{R}^n$, we call the function $\tau(x, y)$, the symmetric one if $\tau(x, y) = \tau(y, x)$, $\tau(x, x) = 0$. Condition A describes a sufficiently broad class of functions $A(\mu, \tau(x, y))$. It was shown in [1],[3-4] that the class $A(\mu, \|x - y\|^2)$, in particular, is wider than $C^{1,1}(D)$ - the well-known class of functions whose gradients satisfy the Lipschitz condition on the convex set $D \subset \mathbb{R}^n$. By the way, let us note that namely the Lipschitzian properties of gradients for this class of functions have been sought as favorable assumptions in justification of the theoretical estimates of the rate of convergence for the various modern differentiable optimization algorithms. In [4], there is given a variety of examples of functions that satisfy Condition A. For functions from $A(\mu, \tau(x, y))$, we also investigated their principle properties and criteria which allow to categorize some function as belonging to the treated class.

In particular, it was proved in [4] that for a continuously differentiable function satisfying Condition A on a convex set D the following extremely important inequality holds:

$$f(x) - f(y) \geq \langle g(x), x - y\rangle - \mu\tau(x, y).$$

Definition 3 *For functions from $A(\mu, \|x{-}y\|^v)$, $v \geq 2$, the vector $s \neq \mathbf{0}$ is called an $\varepsilon-$normalized descent direction ($\varepsilon > 0$) of the function f at the point $x \in D$ if the following inequality holds: $\langle g(x), s\rangle + \varepsilon\|s\|^v < 0$.*

If some descent direction s is not $\varepsilon-$normalized, then the vector constructed in such a way that $\hat{s} = t \cdot s/\varepsilon\|s\|^v$ satisfies the above definition when $0 < t \leq |\langle g(x), s\rangle|$.

Under the condition $v = 2$, we fix some point $x \in \mathbb{R}^n$, then it is not hard to see that all the points $z \in \mathbb{R}^n$, for which the vectors $z - x$ are $\varepsilon-$normalized directions of descent at the point x, belong to the $n-$dimensional ball of radius $R = \|g(x)\|/2\varepsilon$ with center at the point $u = x - g(x)/2\varepsilon$.

Let

$$\zeta = \begin{cases} (\varepsilon \cdot \mu^{-1})^{1/(v-1)}, & \text{if } \varepsilon < \mu, \\ 1, & \text{if } \varepsilon \geq \mu. \end{cases}$$

We further present the very useful new properties of the $\varepsilon-$normalized descent directions:

Theorem 1 *If s is an $\varepsilon-$normalized descent direction for the function f at the point x, then for $\forall\, \beta \in (0,\, 1)$ there exists a constant $\hat{\lambda} = \hat{\lambda}(\beta) > 0$ ($\hat{\lambda} = (1-\beta)^{1/(v-1)}\zeta$) such that for all $\lambda \in (0,\, \hat{\lambda}]$ it holds*

$$f(x) - f(x + \lambda s) \geq -\lambda\beta \cdot \langle g(x), s\rangle, \tag{2}$$

$$f(x) - f(x + \lambda s) \geq \lambda\beta \cdot \varepsilon\|s\|^v. \tag{3}$$

Theorem 2 *If s is an $\varepsilon-$normalized direction of descent for the function f at the point x, then there exists a constant $\hat{\lambda} > 0$ ($\hat{\lambda} = \zeta$) such that for all $\lambda \in (0,\, \hat{\lambda}]$ it is fulfilled*

$$f(x) - f(x + \lambda s) \geq -\lambda \cdot (\langle g(x), s\rangle + \varepsilon\|s\|^v) \geq 0. \tag{4}$$

According to Theorems 1–2, we can describe the rules of calculating the step-size satisfying (2)–(4). Let s be some $\varepsilon-$normalized direction of descent for f at the point x. Besides, let be fulfilled the following conditions: $\beta \in (0, 1)$, $\eta = (1-\beta)^{1/v-1}$, $\hat{\imath} = 1$, $J(\hat{\imath}) = \{\hat{\imath}, \hat{\imath}+1, \hat{\imath}+2, \ldots\}$.

We further determine i^* as the least index $i \in J(\hat{i})$ for which the following condition holds:

$$f(x) - f(x + \eta^i s) \geq -\eta^i \beta \cdot \langle g(x), s \rangle, \tag{5}$$

or the more weak condition:

$$f(x) - f(x + \eta^i s) \geq \eta^i \beta \cdot \varepsilon \|s\|^v. \tag{6}$$

Next, we set $\lambda = \eta^{i^*}$. The step-size calculated in accordance with these rules satisfies to (2) or (3), respectively. For the next rule of the choice of λ, we choose $\eta \in (0, 1)$ and set $\hat{i} = 0$. There should be found further i^* - the smallest index $i \in J(\hat{i})$ such that

$$f(x) - f(x + \eta^i s) \geq -\eta^i \cdot (\langle g(x), s \rangle + \varepsilon \|s\|^v), \tag{7}$$

and $\lambda = \eta^{i^*}$. In the case of finding λ from (5) or (6) under the assumption that $0 < \varepsilon < \mu$, we prove that $\lambda > (\varepsilon \mu^{-1} \cdot (1 - \beta)^2)^{1/v-1} > 0$. Consequently, it holds $\mu > \varepsilon \cdot (1 - \beta)^2 \lambda^{1-v}$. This implies that the procedure of diminishing the step length is finite.

Algorithm.

Step 0. Initialization. Select $x_0 \in D$, $\beta \in (0, 1)$, $\varepsilon_0 > 0$, $0 < \sigma_0 \leq 1$, $0 < \alpha \leq \alpha_0$. Set iteration counter $k = 0$.

Step 1. Choose the point y_k, $k = 0, 1, \ldots$ in such a way that under the conditions $0 < \sigma \leq \sigma_k \leq 1$, $0 < \alpha \leq \alpha_k$ it holds

$$\langle g(x_k), y_k - x_k \rangle \leq \max\{\sigma_k \min_{x \in D} \langle g(x_k), x - x_k \rangle, \, -\alpha_k\}.$$

If $\langle g(x_k), y_k - x_k \rangle = 0$, then we terminate the implementation of the algorithm (since x_k is a solution of problem (1)). In a different way, we set

$$s_k = \begin{cases} y_k - x_k, & \text{if } \langle g(x_k), y_k - x_k \rangle + \varepsilon_k \|y_k - x_k\|^v \leq 0, \\ \dfrac{t_k (y_k - x_k)}{\varepsilon_k \|y_k - x_k\|^v}, & \text{otherwise.} \end{cases}$$

Here $t_k = |\langle g(x_k), y_k - x_k \rangle|$.

Step 2. Let i_k be the least index $i \in J(\hat{i})$ for which there holds the condition of a choice of the iterative step-size for $x = x_k$, $s = s_k$, $\varepsilon = \varepsilon_k$. Then set $\lambda_k = \eta^{i_k}$.

Step 3. Compute the next iterate $x_{k+1} = x_k + \lambda_k s_k$.

Step 4. Update $\varepsilon_{k+1} = \zeta_k \varepsilon_k$. Here,

$$\zeta_k = \begin{cases} (1 - \beta)^{1 - i_k}, & \text{in the case of choice } \lambda_k \text{ from (5) or (6),} \\ \eta^{(1 - i_k)(v - 1)}, & \text{in the case of choice } \lambda_k \text{ from (7).} \end{cases}$$

Set $k = k + 1$ and go to Step 1.

The following theorem justifies a stopping criterion of the algorithm presented above.

Theorem 3 *Let $f(x)$ be a continuously differentiable pseudoconvex function on the convex set $D \subseteq \mathbb{R}^n$. Then for the function $f(x)$ to attain its minimum value on D at the point $x_k \in D$ it is necessary and sufficient to hold*

$$\langle g(x_k), y_k - x_k \rangle = 0.$$

Let $\{x_k\}$ be the sequence constructed by the algorithm. Furthermore, let it be fulfilled $x_k \notin D^*$, $\forall k = 0, 1, \ldots$ For the purpose of exploring the convergence of the numerical methods in the case of pseudoconvex functions, there is usually defined in the literature on optimization an auxiliary numeric sequence $\{\theta_k\}$ as follows.

$$\theta_k > 0,\, 0 < \theta_k \cdot (f(x_k) - f(x^*)) \leq \langle g(x_k), x_k - x^* \rangle,\ x^* \in D^*, k \in L. \quad (8)$$

From the definition of pseudoconvexity, it follows that for pseudoconvex functions such values θ_k must exist. In particular, if $f(x)$ is a smooth convex function, then $\theta_k = 1$, $k = 0, 1, \ldots$ The properties of elements of the sequence $\{\theta_k\}$ were investigated, for instance, in [1],[4].

We further consider the theorem establishing conditions for convergence of the sequence $\{x_k\}$ generated by the algorithm to a solution of problem (1).

Theorem 4 *1. If $f(x)$ is a continuously differentiable pseudoconvex function on the convex and closed set $D \subseteq \mathbb{R}^n$ satisfying Condition A with constant μ and function $\tau(x, y) = \|x - y\|^v$, $v \geq 2$,*
2. a numeric sequence $\{\theta_k\}$ defined by (8) satisfies the condition: $\exists \theta > 0$ such that $\theta_k \geq \theta, \forall k$,
3. there exists a constant $\gamma > 0$ such that $\|g(x)\| \leq \gamma < \infty$, $\forall x \in D$,
4. the Lebesgue set of the function $f(x)$ at the point $x_0 \in D$, denoted by $M_D(f, x_0) = \{x \in D : f(x) \leq f(x_0)\}$, is bounded,
5. $\{\alpha_k\}$, $\{\sigma_k\}$ are such that $\exists \eta > 0 : \|x_k - y_k\| \leq \eta, \forall k$,
6. the step-size λ_k, $\forall k \in L$ is chosen in accordance with one of the rules described in (5)–(7).
Then the sequence $\{x_k\}$, $k \in L$ is weakly convergent, i.e.

$$f(x_k) - f^* \sim O(1/k),$$

or equivalently, there exists the constant $C > 0$ such that it holds

$$f(x_k) - f^* \leq C \cdot k^{-1}.$$

For details of the proof of the preceding theorem, see [4]. Let us note that for estimating the rate of convergence in the case of convexity of the minimized function $f(x)$, the fourth condition of Theorem 4 can be changed to the claim of boundedness of D^*. Without evaluating the rate of convergence, there can be proved the convergence of $\{x_k\}$, $k \in L$ to a solution of problem (1) under the more weak conditions. Indeed, there is true the following theorem.

Theorem 5 *Let the 1, 4, 5 conditions of Theorem 4 be fulfilled, then for the sequence $\{x_k\}$, $k \in L$ generated by the algorithm it holds:*
1. $\lim\limits_{k\to\infty} \langle g(x_k), y_k - x_k\rangle + \varepsilon_k \|y_k - x_k\|^v = 0$,
2. Any limit point of $\{x_k\}$, $k \in L$ belongs to D^*, i.e. $\lim\limits_{k\to\infty} \|x_k - p_k^*\| = 0$.

Finally, we note that the presented algorithm as compared with the classical Frank-Wolfe algorithm has the advantage consisting in a possibility of inexact solving of the direction finding subproblem and handling the accuracy of its solution.

References

1. Gabidullina Z.R. Relaxation methods with step regulation for solving constrained optimization problems // Journal of Mathematical Sciences. 1995. V. 73, N 5. p. 538–543.
2. Mangasarian O.L. Pseudo-convex functions// J.Soc. Industr.and Appl. Math.- Ser.A Control. 1965. V. 3, p. 281–290.
3. Gabidullina Z.R. Convergence of the constrained gradient method for a class of nonconvex functions // Journal of Soviet Mathematics. 1990. V. 50, N 5. p. 1803-1809.
4. Gabidullina Z.R. Adaptive methods with step length regulation for solving pseudoconvex programming problems// Dissertation for the Degree of Candidate of Science in Physics and Mathematics. Kazan. 1994. 125 p.

On solution of bilevel problems with a matrix game at the lower level

T.V. Gruzdeva and A.V. Orlov
Matrosov Institute for System Dynamics and Control Theory of SB of RAS, Irkutsk, Russia

1. Introduction. This paper addresses one new approach to a special type of bilevel optimization problems [1] with an equilibrium at the lower level. Such problems arise in modeling of hierarchical systems, which are characterized by a disparate status of the participants, where one leader is connected with several followers. There are a lot of applications of the bilevel programming problems (BPPs) in control, economy, traffic, telecommunication networks, etc. [2].

As known, according to Pang [3] hierarchy and equilibrium are the promising paradigms in mathematical programming in recent years. Therefore, developing the efficient numerical methods even for the simplest classes of BPPs with an equilibrium is a challenge of modern Operations Research. In this work we consider bilevel problem with a parametric matrix game [4] at the lower level and with a linear goal function subject to linear constraints at the upper level.

In order to elaborate numerical methods for the solving of BPPs with a matrix game at the lower level we reformulate it as a single level optimization problem using a reduction theorem. This auxiliary problem turns out to be a global optimization problem with a nonconvex feasible set (see, e.g., [5–7]). Further for solving the single level problem obtained we apply a special Global Search Theory [7–10] developed by A.S. Strekalovsky for optimization problem with d.c. functions.

2. Problem formulation. Let us formulate the simplest case of BPPs with an equilibrium in the following way:

$$\langle c, x\rangle + \langle d_1, y\rangle + \langle d_2, z\rangle \uparrow \max_{x,y,z}, \quad x \in X, \ (y,z) \in C(\Gamma M(x)), \qquad (\mathcal{BP}_{\Gamma M})$$

where $X = \{x \in I\!R^m \mid Ax \leq a, \ x \geq 0, \ \langle b_1, x\rangle + \langle b_2, x\rangle = 1\}$,
$C(\Gamma(x))$ is the set of saddle points of the game

$$\left.\begin{array}{l} \langle y, Bz\rangle \uparrow \max\limits_{y}, y \in Y(x) = \{y \mid y \geq 0, \langle e_{n_1}, y\rangle = \langle b_1, x\rangle\}, \\ \langle y, Bz\rangle \downarrow \min\limits_{z}, z \in Z(x) = \{z \mid z \geq 0, \langle e_{n_2}, z\rangle = \langle b_2, x\rangle\}; \end{array}\right\} \quad (\Gamma M(x))$$

$c, b_1, b_2 \in I\!R^m$; $y, d_1 \in I\!R^{n_1}$; $z, d_2 \in I\!R^{n_2}$; $a \in I\!R^p$; $b_1 \geq 0, \ b_1 \neq 0$, $b_2 \geq 0, \ b_2 \neq 0$; A, B are matrices and $e_{n_1} = (1, ..., 1)$, $e_{n_2} = (1, ..., 1)$ are vectors of appropriate dimension.

The expression $\langle b_1, x\rangle + \langle b_2, x\rangle = 1$ can be interpreted as some resource, which should be distributed by the leader among the followers.

In order to elaborate numerical methods for the solving of bilevel problem $(\mathcal{BP}_{\Gamma M})$ we need to reformulate it as a single level problem.

Let us set $\xi_1 := \langle b_1, x\rangle$, $\xi_2 := \langle b_2, x\rangle$ (x is fixed) and formulate matrix game with parameters ξ_1, ξ_2:

$$\left.\begin{array}{l} \langle y, Bz\rangle \uparrow \max\limits_{y}, \quad y \in Y = \{y \mid y \geq 0, \langle e_{n_1}, y\rangle = \xi_1 > 0\}, \\ \langle y, Bz\rangle \downarrow \min\limits_{z}, \quad z \in Z = \{z \mid z \geq 0, \langle e_{n_2}, z\rangle = \xi_2 > 0\}. \end{array}\right\} \qquad (\Gamma M)$$

Further we formulate optimality conditions for generalized matrix game (ΓM). These conditions are a generalization of classical optimality conditions in a matrix game [4].

Theorem 1. The tuple $(y^*, z^*) \in C(\Gamma M)$ if and only if there exists a number v_* (an optimal value of the game (ΓM)) such that the following system is fulfilled:

$$\left.\begin{array}{lll} \xi_1(Bz^*) \leq v_* e_{n_1}, & z^* \geq 0, & \langle e_{n_2}, z\rangle = \xi_2; \\ \xi_2(y^* B) \geq v_* e_{n_2}, & y^* \geq 0, & \langle e_{n_1}, y\rangle = \xi_1. \end{array}\right\} \qquad (1)$$

Note, conditions (1) represent finite numbers of equalities and inequalities. Now we can replace a game at the lower level by its optimality conditions. Hence, for the bilevel problem $(\mathcal{BP}_{\Gamma M})$ it is possible to formulate the following equivalent single level problem:

$$\left.\begin{array}{c} f_0(x,y,z) \overset{\triangle}{=} \langle c, x\rangle + \langle d_1, y\rangle + \langle d_2, z\rangle \uparrow \max\limits_{x,y,z,v}, \\ Ax \leq a, \quad x \geq 0, \quad \langle b_1, x\rangle + \langle b_2, x\rangle = 1, \\ y \geq 0, \quad \langle e_{n_1}, y\rangle = \langle b_1, x\rangle, \quad z \geq 0, \quad \langle e_{n_2}, z\rangle = \langle b_2, x\rangle, \\ \langle b_1, x\rangle (Bz) \leq v e_{n_1}, \quad -\langle b_2, x\rangle (yB) \leq -v e_{n_2}. \end{array}\right\} \qquad (\mathcal{PM})$$

More precisely, the following theorem takes place.

Theorem 2. The triplet (x^*, y^*, z^*) is a global optimistic solution of the bilevel problem $(\mathcal{BP}_{\Gamma M})$, if and only if there exist a number v_* such that the 4-tuple (x^*, y^*, z^*, v_*) is a global solution of problem $(\mathcal{PM})$.

It can readily be seen, that problem $(\mathcal{PM})$ is a global optimization problem with a nonconvex feasible set (see, e.g., [5–7]). A nonconvexity in the problem $(\mathcal{PM})$ is generated by two vector constraints (two groups of $(n_1 + n_2)$ bilinear constraints in total). These constraints have arisen from optimality conditions for the generalized matrix game at the lower level of the bilevel problem $(\mathcal{BP}_{\Gamma M})$. It is known, that bilinear function is represented as a difference of two convex functions (i.e. bilinear function is d.c. function) [4]. Therefore, problem $(\mathcal{PM})$ belongs to the class of nonconvex optimization problems with d.c. constraints [7–10] and we can apply the Global Search Theory for solving this class of nonconvex problems.

3. D.C. decomposition. The first stage of the application of the Global Search Theory to the problem under scrutiny is a decomposition of nonconvex function as a difference of two convex functions. As noted above, $(n_1 + n_2)$ bilinear constraints generate the basic nonconvexity in the problem $(\mathcal{PM})$.

Therefore we should find an explicit decomposition of functions f_i by the difference of two convex functions. As an example, let us obtain an evident d.c. representation of the i-th scalar constraint of the first group:

$$f_i(x, z, v) = \langle b_1, x\rangle\langle (B)_i, z\rangle - v \le 0, \qquad i = 1, \dots, n_1, \tag{2}$$

where $(B)_i$ is an i-th row of the matrix B.

Introduce a denotation $Q_i^T = (b_1^{(1)}(B)_i;\ b_1^{(2)}(B)_i;\ \dots;\ b_1^{(m)}(B)_i)$, where $b_1^{(j)}$ is a j-th component of the vector b_1. Hence, we can reduce (2) to a standard bilinear form $f_i(x, z, v) = \langle x^T Q_i, z\rangle - v \le 0,\ i = 1, \dots, n_1$. And we can use here the known d.c. representation which is based on the property of a scalar product [4]:

$$f_i(x, z, v) = g_i(x, z, v) - h_i(x, z), \tag{3}$$

where $g_i(x, z, v) = \dfrac{1}{4}\|xQ_i + z\|^2 - v,\ \ h_i(x, z) = \dfrac{1}{4}\|xQ_i - z\|^2$. The remain part of constraints can be decomposed analogously.

Therefore we obtain the following problem with d.c. constraints:

$$\left.\begin{array}{c} f_0(x, y, z) \downarrow \min\limits_{x,y,z,v}, \quad (x, y, z) \in S, \\ f_i(x, z, v) := g_i(x, z, v) - h_i(x, z) \le 0,\ i = 1, \dots, n_1, \\ f_i(x, y, v) := g_i(x, y, v) - h_i(x, y) \le 0, i = n_1 + 1, \dots, n_1 + n_2, \end{array}\right\} \quad (\mathcal{P})$$

where the functions f_0 and $g_i, h_i,\ i \in I = \{1, ..., n_1 + n_2\}$, as well as the set $S = \{x, y, z \ge 0 \mid Ax \le a,\ \ \langle b_1, x\rangle + \langle b_2, x\rangle = 1,\ \ \langle e_{n_1}, y\rangle = \langle b_1, x\rangle,$ $\langle e_{n_2}, z\rangle = \langle b_2, x\rangle\}$, are convex.

4. Local and Global search. As mentioned above, for the purpose of solving the d.c. constraint problem $(\mathcal{P})$, we develop the Global Search Algorithm based on the Global Search Theory (GST) [7–9] using d.c. decomposition constructed above. According the GST, the algorithm for the solving of problem $(\mathcal{P})$ should consist of two principal stages: 1) a special local search method (LSM), which takes into account the structure of the problem under scrutiny [10]; 2) the procedure based on Global Optimality Conditions (GOCs) [7–9], that allow to improve the point provided LSM [10].

In order to find a local solution to problem $(\mathcal{P})$, we suppose that the feasible set $D := \{(x, y, z) \in S \mid f_i(\cdot) \leq 0, \;\; i \in I\}$ of problem $(\mathcal{P})$ is not empty and the optimal value $\mathcal{V}(\mathcal{P}) := \inf\{f_0(x, y, z) \mid (x, y, z, v) \in D\}$ of problem $(\mathcal{P})$ is finite: $\mathcal{V}(\mathcal{P}) > -\infty$.

Furthermore, let us denote $w := (x, y, z, v) \in I\!R^{m+n_1+n_2+1}$ and assume that a feasible starting point $w^0 \in D$ is given and, in addition, after several iterations it has derived the current iterate $w^s \in D$, $s \in \mathbb{Z}_+ = \{0, 1, 2, \ldots\}$.

In order to propose a LSM for problem $(\mathcal{P})$, apply a classical idea of linearization with respect to the basic nonconvexity of the problem (i.e. with respect to $h_i(\cdot), \;\; i \in I$) at the point w^s [10]. Thus, we obtain the following linearized problem:

$$\left.\begin{array}{c} f_0(x, y, z) \downarrow \min\limits_{x,y,z,v}, \quad (x, y, z) \in S, \\ \varphi_{is}(x, z, v) := g_i(x, z, v) - \langle \nabla h_i(x^s, z^s), (x, z) - (x^s, z^s)\rangle - \\ -h_i(x^s, z^s) \leq 0, \; i = 1, \ldots, n_1, \\ \varphi_{is}(x, y, v) := g_i(x, y, v) - \langle \nabla h_i(x^s, y^s), (x, y) - (x^s, y^s)\rangle - \\ -h_i(x^s, y^s) \leq 0, \; i = n_1 + 1, \ldots, n_1 + n_2. \end{array}\right\} \quad (\mathcal{PL}_s)$$

Suppose the point w^{s+1} is provided by solving problem $(\mathcal{PL}_s)$, so that

$$w^{s+1} \in D_s = \{(x, y, z) \in S \mid \varphi_{is}(\cdot) \leq 0, \quad i \in I\}$$

and inequality $f_0(x^{s+1}, y^{s+1}, z^{s+1}) \leq \mathcal{V}(\mathcal{PL}_s) + \delta_s$ holds. Here $\mathcal{V}(\mathcal{PL}_s)$ is the optimal value to Problem $(\mathcal{PL}_s)$ and the sequence $\{\delta_s\}$ satisfies the following condition functions $\sum\limits_{s=0}^{\infty} \delta_s < +\infty$.

Therefore, the LSM generates the sequence $\{w^s\}$, $w^s \in D_s$, $s \in \mathbb{Z}_+$, of solutions to problems $(\mathcal{PL}_s)$. As it was proven in [10], the cluster point $w_* \in D_*$ of the sequence $\{w^s\}$ is a solution to the linearized problem $(\mathcal{PL}_*)$ (which is problem $(\mathcal{PL}_s)$ with w_* instead of w^s), and w_* can be called the critical point with respect to the LSM. Thus, the algorithm constructed in this way provides critical points by employing suitable convex optimization methods for any given accuracy τ. The following inequality:

$$f_0(x^s, y^s, z^s) - f_0(x^{s+1}, y^{s+1}, z^{s+1}) \leq \frac{\tau}{2}, \quad \delta_s \leq \frac{\tau}{2},$$

can be chosen as a stopping criterion for the LSM [10].

Finally, we recall that the Global Search Procedure for finding a global solution to the problem $(\mathcal{P})$ are based on GOCs [7–9] and their so-called algorithmic (constructive) property. If there exist a point, which violates these conditions, then we can find a new point, which is better than the current point in problem $(\mathcal{P})$ (it works even for critical points and local solutions). In order to find such violative points we use the solution of linearized problems, local search, and a constructing of the level surface approximation of the convex function which generates the basic nonconvexity in the problem $(\mathcal{P})$.

References

1. Dempe S. Foundations of Bilevel Programming. Dordrecht: Kluwer Academic Publishers, 2002.
2. Colson B., Marcotte P., Savard G. An overview of bilevel optimization // Annals of operations research. 2007. V. 153, P. 235–256.
3. Pang J.-S. Three modeling paradigms in mathematical programming // Mathematical programming, Ser.B. 2010. V. 125, P. 297–323.
4. Strekalovsky A.S., Orlov A.V. Bimatrix games and bilinear programming. Moscow: FizMatLit, 2007 (in russian).
5. Horst R., Tuy H. Global Optimization. Deterministic Approaches. Berlin: Springer-Verlag, 1993.
6. Strongin R.G., Sergeyev Ya.D. Global Optimization with Non-Convex Constraints. Sequential and Parallel Algorithms. New York: Springer-Verlag, 2000.
7. Strekalovsky A.S. Elements of nonconvex optimization. Novosibirsk: Nauka, 2003 (in russian).
8. Strekalovsky A.S. Global Optimality Conditions in Nonconvex Optimization // Journal of Optimization Theory and Applications. 2017. V. 173, No. 3. P. 770–792.
9. Strekalovsky A.S. Global Optimality Conditions and Exact Penalization // Optimization Letters. DOI: 10.1007/s11590-017-1214-x (published online).
10. Strekalovsky A.S. On local search in d.c. optimization problems // Applied Mathematics and Computation. 2015. V. 255. P. 73–83

Antipodal theorems extended*

V.V. Kalashnikov[1,2], A.J.J. Talman[3], L. Alanís-López[4], and
N.I. Kalashnykova[4]

[1] *Central Economics and Mathematics Institute (CEMI), Moscow, Russian Federation*

[2] *Tecnológico de Monterrey (ITESM), Monterrey, Nuevo León, Mexico*

[3] *School of Economics and Management, Tilburg University, Tilburg, the Netherlands*

[4] *Universidad Autónoma de Nuevo León (UANL), San Nicolás de los Garza, Nuevo León, Mexico*

Since 1909 when Brouwer proved the first fixed-point theorem named after him, the fixed-point results in various settings play an important role in the optimization theory and applications. This technique has proven to be indispensable for the proofs of multiple results related to the existence of solutions to numerous problems in the areas of optimization and approximation theory, differential equations, variational inequalities, complementary problems, equilibrium theory, game theory, mathematical economics, etc. It is also worthwhile to mention that the majority of problems of finding solutions (zero-points) of functions (operators) can be easily reduced to that of discovering of fixed points of properly modified mappings.

Not only theoretical but also practical (algorithmic) developments are based on the fixed-point theory. For instance, the well-known simplicial (triangulation) algorithms help one to find the desired fixed points in a constructive way. That approach allows one to investigate the solvability of complicated problems arising in theory and applications.

In this thesis, making use of the triangulation technique, we extend some antipodal and fixed-point theorems to the case of non-convex, more exactly, star-shaped sets. Also, similar extensions are made for set-valued mappings defined over star-shaped sets. From now onward, we briefly explain the main points of our extensions.

The techniques using various fixed-point theorems have always been widely applied in Operations Research, Mathematical Programming, Game Theory, and other areas of optimization. Such techniques are

*The research activity of the first and the fourth authors was financially supported by the R&D Department “Optimization and Data Science” of the Tecnológico de Monterrey (ITESM), Campus Monterrey, as well as by the SEP-CONACYT projects CB-2013-01-221676 and FC-2016-01-1938 (Mexico)

appropriate in establishing the existence of solutions to mathematical programming problems, convex games, mathematical programs with equilibrium constraints (MPEC), to mention only few. Because of that, any extensions of the classical fixed-point theorems (like the Brouwer theorem for single-valued mappings and Kakutani theorem for multi-valued mappings) are interesting and important. In other words, topological tools are of a very high importance and use in the development of optimization theory and applications.

The Brouwer fixed point theorem and the Borsuk-Ulam theorem can be characterized as two most powerful topological tools of extremely similar structure. Most topology textbooks that include these theorems, such as, for example, Krasnosel'skii [1], Herings [2], Yang [3], do not mention that the two are tightly related (for instance, the Borsuk-Ulam theorem implies the Brouwer Fixed Point Theorem).

The majority of fixed-point theorems have been established for the mappings defined on convex domains. However, in many applications and real-life problems, the domains need not be convex; for example, the feasible sets of bilevel programming problems mostly lack this property even for linear bilevel problems (*cf.,* [4]). In this abstract, we outline how we extend the Borsuk-Ulam (antipodal) theorem to more general domains and to the case of multi-valued mappings as well. First, we recall the definition of star-shaped sets and some properties of projections onto them.

Definition 1 *A star-shaped region centered at x_0 in $\mathbb{R}^n$ is a set $D \subset \mathbb{R}^n$ such that for every $x \in D$ and $t \in [0,1]$ one has $\gamma_x(t) := x_0 + t(x - x_0) \in D$.*

Consider $\mathbb{R}^n$ with the maximum metric: $d_\infty(x,y) := \max\{|x_i - y_i|\}_{i=1}^n$, where $x, y \in \mathbb{R}^n$.

In order to work with the projections of the points inside and outside a star-shaped subset to its boundary, we need the following notation.

Definition 2 *Let D be a closed, bounded, and star-shaped region D centred at $x_0 = 0$. For an $r > 0$, such that $D \subset C^n(r) := \{x \in \mathbb{R}^n : \|x\|_\infty \leq r\}$, we define the following:*

- *For $x \in C^n(r) \setminus D$ define $\theta_x := max\{t \in [0,1] : \gamma_x(t) \in D\}$ and $\alpha(x) := \gamma_x(\theta_x)$.*
- *For $y \in D$ let $u_y := inf\{t \in [1,\infty) : \gamma_y(t) \notin D\}$ and $\lambda(y) := \gamma_y(u_y)$.*

- *For $z \in C^n(r)$ let $s_z := inf\{t \in [1, \infty) : \gamma_z(t) \notin C^n(r)\}$ and $\omega(z) := \gamma_z(s_z)$.*

The antipodal property is the key feature of the functions examined in the Brouwer and Borsuk-Ulam theorems.

Definition 3 *Let $D \subset C^n(r)$ be a closed star-shaped set and $\rho : D \longrightarrow \mathbb{R}^n$ a continuous function. The function is said to have the antipodal property, if for every $x \in \partial D$ it holds that $\rho(x) = -\rho(\lambda(-x))$ if $(-x) \in D$ and $\rho(x) = -\rho(\alpha(-x))$ when $(-x) \notin D$.*

The well-known extension of the antipodal property is the *nonparallel condition* defined below.

Definition 4 *Let $D \subset C^n(r)$ be a closed star-shaped set and $\rho : D \longrightarrow \mathbb{R}^n$ a continuous function. This function satisfies the nonparallel condition if for every $x \in \partial D$ it holds: for all $c \geq 0$, $\rho(x) \neq c\rho(\lambda(-x))$ if $(-x) \in D$ and $\rho(x) \neq c\rho(\alpha(-x))$ otherwise, that is, when $(-x) \notin D$.*
We say that ρ satisfies the weakly nonparallel condition when for all $c > 0$, $\rho(x) \neq c\rho(\lambda(-x))$ if $(-x) \in D$ and $\rho(x) \neq c\rho(\alpha(-x))$ otherwise, i.e., if $(-x) \notin D$.

Similar to the constructive proofs presented in [5–6], we deduce important existence results concerning zero points.

Theorem 1 *Let $D \subset C^n(r)$ a closed, bounded, and star-shaped set and $\rho : D \longrightarrow \mathbb{R}^n$ a continuous function which satisfies the weakly nonparallel condition. Then, there exists $x^* \in D$ such that $\rho(x^*) = 0$.*

Next, we extend the above results to multi-valued mappings. In order to do that, we refer to the distance between sets as introduced in [8–10].

Definition 5 *Let A, B be subsets of $\mathbb{R}^n$. The distance between A and B is defined as follows: $d(A, B) := \inf_{x \in A\ y \in B}\{\|x - y\|\}$.*

Now we can extend the antipodal and nonparallel conditions to the case of multi-valued mappings defined on star-shaped sets.

Definition 6 *Let $D \subset C^n(r)$ be a closed star-shaped set and $\rho : D \longrightarrow \mathcal{S}(\mathbb{R}^n)$ a multi-valued mapping. The mapping is said to have the antipodal property if for every $x \in \partial D$ one has $\rho(x) = -\rho(\lambda(-x))$ if $(-x) \in D$ and $\rho(x) = -\rho(\alpha(-x))$ if $(-x) \notin D$; here $-\rho(w) := \{-z : z \in \rho(w)\}$.*

Definition 7 *Let $D \subset C^n(r)$ be a closed, bounded, and star-shaped set and $\rho : D \longrightarrow \mathcal{S}(\mathbb{R}^n)$ a multi-valued mapping. The mapping satisfies the nonparallel condition if for every $x \in \partial D$ and $c \geq 0$ one has $y_1 \neq cy_2$ for all $y_1 \in \rho(x)$, $y_2 \in \rho(\lambda(-x))$ whenever $(-x) \in D$. Otherwise, if $(-x) \notin D$, then $y_1 \neq cy_2$ for all $y_1 \in \rho(x)$, $y_2 \in \rho(\alpha(-x))$. If the above inequalities are claimed for (only) $c > 0$ we say that ρ satisfies the weakly nonparallel condition.*

First, we demonstrate that the problem of finding zeros of a multi-valued mapping defined on a closed bounded star-shaped subset of $\mathbb{R}^n$ satisfying the nonparallel condition can be reduced to the problem of finding zeros of an extended multi-valued mapping boasting the antipodal property on some cube $C^n(r)$.

Lemma 1 *Let $D \subset C^n(r)$ be a closed, bounded, and star-shaped set and $\rho : D \longrightarrow \mathcal{S}(\mathbb{R}^n)$ a multi-valued mapping satisfying the weakly nonparallel condition. Then ρ can be extended to a mapping $\Psi : C^n(r) \longrightarrow \mathcal{S}(M)$ with the antipodal property and such that $x^* \in D$ whenever $(0, \ldots, 0) \in \Psi(x^*)$.*

Now the main result obtained for the multi-valued mappings on star-shaped set follows.

Theorem 2 *Let $\rho : D \subset C^n(r) \longrightarrow \mathcal{S}(\mathbb{R}^n)$ be an upper semicontinuous mapping defined on a star-shaped subset D satisfying the weakly nonparallel condition. Then there exists $x^* \in D$ such that $0 \in \rho(x^*)$.*

The abstract finishes with an extension of the nonparallel condition-related result for mappings defined on a star-shaped set times an interval $D \times [0,1]$ and having convex compact subsets of $\mathbb{R}^n$ as its values. Moreover, we will study a connected set of zero points intersecting both $D \times \{0\}$ and $D \times \{1\}$ in order to extend the Browder theorem.

Theorem 3 *Let $\rho : D \times [0,1] \longrightarrow \mathcal{H}(\mathbb{R}^n)$ be a multi-valued mapping from $D \times [0,1]$ to $\mathcal{H}(\mathbb{R}^n)$, where $\mathcal{H}(\mathbb{R}^n) = \{A \in \mathcal{S}(\mathbb{R}^n) : A \text{ is a convex set}\}$ and D is a compact star-shaped set in $\mathbb{R}^n$ with respect to the origin. Suppose that the following conditions hold for ρ:*

1. *ρ is upper semicontinuous.*
2. *For every $t \in [0,1]$, the restriction $\rho_t : D \longrightarrow \mathcal{C}(\mathbb{R}^n)$ defined by $\rho_t(x) := \rho(x,t)$ satisfies the nonparallel condition. Then*

there exists a connected subset Z in $D \times [0,1]$ such that $Z \cap D \times \{0\} \neq \emptyset$, $Z \cap D \times \{1\} \neq \emptyset$, and if $x \in Z$ then $0 \in \rho(x)$.

To conclude, we have presented the extensions of both the antipodal (Borsuk-Ulam) theorem and Browder theorem to the cases comprising both a star-shaped domain of the considered mapping and a multi-valued structure of that mapping. Moreover, an explicit algorithm constructing the desired connected path of the zero points of the mapping is developed. The properties of the latter algorithm are used in the proof of the extended Browder Theorem, similar to the existence proofs in the pioneer works by van der Laan and Talman [5–6].

In our future research, we are going to extend the above-mentioned results to other classes of non-convex domains of the examined multi-valued mappings.

References

1. Krasnosel'skii M.A. Topological Methods in the Theory of Nonlinear Integral Equations. Oxford: Pergamon Press, 1964.
2. Herings P.J.-J. Static and Dynamic Aspects of General Equilibrium Theory. Boston/London/Dordrecht: Kluwer Academic Publishers, 1996.
3. Yang Z.F. Computing Equilibria and Fixed Points. Boston/London /Dordrecht: Kluwer Academic Publishers, 1999.
4. Dempe S. Foundations of Bilevel Programming. Boston/London/ Dordrecht: Kluwer Academic Publishers, 2002.
5. van der Laan G., Talman A.J.J. A class of simplicial restart fixed point algorithms without an extra dimension // Mathematical Programming. 1981. V. 20, N. 1. P. 33–48.
6. van der Laan G. Existence and approximation of zeros // Mathematical Programming. 1984. V. 28, N. 1. P. 1–24.
7. Todd M.J. Improving the convergence of fixed point algorithms // Mathematical Programming Study. 1978. V. 7, N. 2. P. 151–179.
8. Munkres J.R. Topology. Upper Saddle River, NJ: Prentice Hall, 2000.
9. Searcóid M.Ó.. Metric Spaces. Springer Undergraduate Mathematics Series. 2007. V. 106.
10. Isac G., Bulavsky V.A., Kalashnikov V.V. Complementarity, Equilibrium, Efficiency and Economics. Boston/London/Dordrecht: Kluwer Academic Publishers, 2002.

Numerical comparison of the centered cutting plane methods for solving convex nondifferentiable optimization problems[*]

A.V. Kolosnitsyn
Melentiev Energy Systems Institute SB RAS, Irkutsk, Russia

We consider the convex optimization problem in the following form:

$$\begin{array}{ll} \text{minimize} & f(x) \\ \text{subject to} & Ax \leq b, \end{array}$$

where f is convex but not necessarily differentiable function, A is $m \times n$ matrix, $x \in \mathbb{R}^n$ and $b \in \mathbb{R}^m$. The feasible set $X = \{x \in \mathbb{R}^n : Ax \leq b\}$ is supposed to be nonempty and bounded.

We aim to compare different centered cutting plane methods with the following general scheme:

Step 0. Set $X^0 = X$, $k = 0$.

Step 1. Define x^k – the center of the set X^k.

Step 2. Find a vector α^k:

$$f(x) - f(x^k) \geq (\alpha^k)^T(x - x^k).$$

Step 3. Define the set X^{k+1}:

$$X^{k+1} = X^k \bigcap \{x : (\alpha^k)^T(x - x^k) \leq 0\}.$$

Step 4. Increment $k = k + 1$ and move to the Step 1.

Different techniques of determine the center of the set X from the general scheme lead to a bunch of centered cutting plane methods. We consider the following ways of finding the center of the set X^k.

Center of the inscribe ellipsoid with maximal volume. Solution of the problem

$$\begin{array}{ll} \text{maximize} & \sum\limits_{j=1}^{n} \ln \delta_j, \\ \text{subject to} & \sum\limits_{j=1}^{n} a_{ij}x_j + \sqrt{\sum\limits_{j=1}^{n} a_{ij}^2 \delta_{ij}^2} \leq b_i,\ i = 1, \ldots, m \end{array}$$

[*] This research is supported by RFBR, research project No. 18-07-01432.

is the pair $(x^\star, \delta^\star)$, where $x^\star$ is the center of maximal volume ellipsoid, $\delta^\star$ defines the length of ellipsoid semiaxes. Ellipsoid $E \subset X$ is defined as follows:

$$E = \left\{ x \in \mathbb{R}^n : \sum_{j=1}^{n} \frac{(x_j - x_j^*)^2}{(\delta_j^*)^2} \le 1 \right\}.$$

Center of the inscribe parallelepiped with maximal volume. Let us consider the following convex optimization problem

$$\begin{array}{ll} \text{maximize} & \sum\limits_{j=1}^{n} \ln \delta_j, \\ \text{subject to} & \sum\limits_{j=1}^{n} a_{ij} x_j + \sum\limits_{j=1}^{n} |a_{ij}| \delta_{ij} \le b_i, \ i = 1, \dots, m. \end{array}$$

Optimal solution $(x^\star, \delta^\star)$ has the following sense: $x^\star$ is the center of the maximal volume parallelepiped $\Pi \subset X$:

$$\Pi = \left\{ x \in \mathbb{R}^n : x_j^* - \delta_j^* \le x_j \le x_j^* + \delta_j^*, j = 1, \dots, n \right\},$$

and $2\delta^\star$ defines the length of the parallelepiped sides.

Center of the inscribe rhombus with maximal volume. In current case we solve the following convex optimization problem

$$\begin{array}{ll} \text{maximize} & \sum\limits_{j=1}^{n} \ln \delta_j, \\ \text{subject to} & \sum\limits_{j=1}^{n} a_{ij} x_j + |a_{ij}| \delta_{ij} \le b_i, \ i = 1, \dots, m, \ j = 1, \dots, n. \end{array} \tag{1}$$

Let the pair (x^*, δ^*) be a solution of the problem (1). Then x^* is the rhombus center and δ^* defines the length of its semiaxes. Rhombus $R \subset X$ has the form

$$R = \left\{ x \in \mathbb{R}^n : \sum_{j=1}^{n} \frac{|x_j - x_j^*|}{\delta_j^*} \le 1 \right\}.$$

Analytical center. An analytical center of the set X is the point $x^\star$:

$$x^* = \text{argmax} \left\{ \sum_{i=1}^{n} \ln(b_i - A^i x) : x \in \mathbb{R}^n \right\}.$$

Circumscribe sets. We consider two types of circumscribe localization sets: minimal volume ellipsoid [1, 2] with modifications developed in [3] and minimal volume simplex [4] with cutting plane shift modifications [5].

All the algorithms based on defined centers of the feasible set were tested on the nondifferentiable convex optimization problem that has the form

$$\begin{array}{ll} \text{minimize} & \max\limits_{1 \le i \le m} \{f_i(x)\}, \\ \text{subject to} & Ax \le b, \end{array}$$

where $f_i(x) = x^T Q^i x + p^i x + r_i$, Q^i are $n \times n$ positive definite matrixes, $p^i \in \mathbb{R}^n$, $i = 1, 2, ..., m$, $r \in \mathbb{R}^m$, $b \in \mathbb{R}^m$, A is $m \times n$ matrix. The results of the numerical experiment will be given in this work.

At last we notice one practical applications of the demonstrated centered cutting plane methods. Such approaches of solving convex nondifferentiable optimization problems are useful in two-stage stochastic linear programming problems [6]. In avoidance of huge amount of the variables it is reasonable to formulate the dual problem for the second stage. But in this case we obtain nondifferentiable optimization problem that can be solved by the suggested methods.

References

1. Shor N.Z. Minimization methods for non-differentiable functions. Springer-Verlag Berlin, Heidelberg, 1985. P. 164.
2. Nemirovsky A., Yudin D. Informational complexity and efficient methods for solution of convex extremal problems. J. Wiley & Sons, New York, 1983.
3. Ershova M., Khamisov O. A modification of the ellipsoid method // Studia Informatica Universalis. Hermann. 2012. V. 9, N 3. P. 43–63.
4. Antsiferov E.G., Bulatov V.P. An algorithm of simplex imbeddings in convex programming // U.S.S.R. Comput. Math. Math. Phys. 1987. V. 27, N 2. P. 36–41.
5. Kolosnitsyn A.V. Computational Efficiency of the Simplex Embedding Method in Convex Nondifferentiable Optimization // Computational Mathematics and Mathematical Physics. 2018. Vol. 58, N. 2. P. 215–222.
6. Yudin D.B. Mathematical Methods of Control under Incomplete Information. Tasks and Methods of stochastic programming. Moscow, Sovetskoye radio, 1974.

Some way to minimize a function with extra variable

V.V. Koulaguin
Institute of Problems in Mechanical Engineering, Russian Academy of Sciences, St. Petersburg, Russia

A mathematical programming problem for function of two variables, where minimum is taken over one of the variables, [1-5] is talked. Examples, existence conditions, and applications [6-13] are considered.

1. Introduction. The problem is some mathematical tool to make decision and control design under uncertainty. There are many ways to make decision under uncertainty. They all supposed to have an uncertainty parameter domain known and bounded (as minimax principle, for example). In our way to tackle uncertainty, the uncertainty parameter domain is a whole space of its values. As a result, we get a system that is effective (robust, reliable) for the maximal (in a certain sense) set of values of the uncertainty parameter. This set is the basic (attributive) property of the system we construct under that kind of uncertainty.

2. The minimization problem. There are a function $f(x,y)$, $x \in E_x$, $y \in E_y$; a set

$$A = \{(x,y) \mid x = \arg\min_x f_y(x)\},$$

where $f_y(x)$ – a cross-section of function $f(x,y)$ by a variable y fixed; and a set

$$Y(x) = \{y \mid (x,y) \in A\},$$

which characterizes the element $x \in E_x$ and is named an effectiveness set of element $x \in E_x$. The sets E_x, E_y are some finite-dimensional or infinite-dimensional spaces.

The problem is to find an element x^0 so that

$$Y(x^0) \supseteq Y(x), \ \forall x \in E_x. \qquad (*)$$

Here x^0 – a maximally effective element, $Y(x^0)$ – a maximal effectiveness set.

2.1. An example of the problem (*). Existence of a solution depends on a type of space E_x. Here is an example. First, let function $f(x,y)$ be $f(x,a,b,c) = ax^2 + bx + c$; where $x \in R_1$, $y = (a,b,c) \in R_3$. Effectiveness sets $Y(x)$ don't have common elements [3]; That means a solution of the problem (*) doesn't exist. Now let element x be a function of an argument (a, b, c), $x = x(a, b, c)$. In that case, (by properties of parabolas) a maximally effective element is $x^0 = -b/2a$, a maximal effectiveness set is R_3.

2.2. One more example of the problem (*). In [1], a control function was constructed that moves, in minimum time, a single mass point to the origin of the phase coordinates – for any initial motion data.

2.3. A trivial solution. A solution of the problem (*) always exists, when element $x \in E_x$ is a function $x(y)$. This solution cannot be acceptable for decision maker, because parameter y a priori isn't known. Usually, the decision maker knows a part of uncertainty parameter y, i.e. $x = x(\bar{y})$, $\bar{y} \in y$. The more of y he knows the more likely that the solution of problem (*) exists.

2.4. Some condition of existence. Let's consider a set

$$X(y) = \{x \mid (x, y) \in A\}$$

and a problem to find an element x^+, so that

$$x^+ \in \bigcap_y X(y). \tag{**}$$

Theorem [3]. An element x is a solution of problem (*), if and only if this element is a solution of problem (**).

3. Decision making problem. Let x be a decision, y be an uncertainty, a pair (x, y) be an act of decision making, which is estimated by functionals $g_i(x, y), h_j(x, y), J(x, y)$. Let A be a set of pairs (x, y), such that

$$g_i(x, y) \leqslant 0, \tag{3.1}$$

$$h_j(x, y) = 0, \tag{3.2}$$

$$J(x, y) = \min_x J_y(x, y), \tag{3.3}$$

where $J_y(x, y)$ – a cross-section of function $J(x, y)$ by a variable y fixed. It should be noted that for fixed y the relation (3.1-3.3) is an ordinary problem of conditional minimization.

To each decision x we associate the set

$$Y(x) = \{y \mid (x, y) \in A\},$$

which characterizes the element $x \in E_x$ and is named a robustness set of element $x \in E_x$.

The problem is to find an element x^0 so that either

$$Y(x^0) \supseteq Y(x), \ \ \forall x \in E_x, \tag{i}$$

or

$$f(Y(x^0)) = \max_x f(Y(x)), \tag{ii}$$

where $f(\cdot)$ – some function of the set $Y(x)$. Here x^0 – a maximally robust element, $Y(x^0)$ – a maximal robustness set.

4. Control design problem. In such applications, element $x \in E_x$ is supposed to be a function of time and/or a function of phase variables of the system controlled. If element x is a function of time $x(t)$, then a robustness sets $Y(x)$ are usually empty or contain only one element. On the contrary, if element x is a function of phase variables and time, then often a solution of problem (i) or problem (ii) might be obtained. In [2], one can find some technique how to construct function $f(\cdot)$. In particular, the problem (ii) can take the form

$$max_x root_y \rho(x, y), \tag{4.1}$$

where root denotes the root of equation $\rho(x, y) = 0,$ $\rho(x, y)$ is some distance from the ball of radius y inscribed in the set $Y(x)$ to the set $Y(x)$. Some problem (ii) in the form (4.1) is considered in [8]. The solution is obtained by using the relation

$$max_x root_y \rho(x, y) = root_y max_x \rho(x, y), \tag{4.2}$$

which is valid under certain assumptions.

In [6-9] some control design problems in the form (i) or (ii) are considered.

5. Engineering problem. In [10] the problem of reliability was formulated as a problem of robustness of machines and structures. In [11] the problem of maximal robustness shock absorber in form (ii) is considered.

6. Game theory. In [12-13] zero-sum game with constrained payoff is investigated. In this game, there is the equilibrium (4.2).

Conclusion. In this report some way to minimize a function that is burdened by an additional variable is discussed. The area of change of the additional variable is the entire space. Applying the proposed method, we obtain the system which is effective (robust, reliable) for maximal (in the sense (i), (ii)) set of values of the variable y.

References

1. Pontryagin L.S., Boltyansky V.G., Gamcrelidze R.V., Mishchenko E.F. Mathematical theory of optimal processes. – M.: Fizmatgiz, 1961. (In Russian)
2. Koulaguin V.V. Balance point method of design under some kind of uncertainty. // Report. International Conference Dedicated to the 90th Anniversary of L.S.Pontryagin. Optimal Control. Moscow State University, Moscow (1998).
3. Koulaguin V.V. Mathematical programming problem for function with extra variable. // Report. Conference "Mathematical programming and applications". Institute of Mathematics and Mechanics, Russian Academy of Sciences. Ekaterinburg (2003). (In Russian)
4. Koulaguin V.V. Robustness as an index of efficiency for a decision making under uncertainty. A maximal robustness problem. *In: Proceedings of VIII Moscow International Conference on Operations Research (ORM 2016).* Moscow State University. Moscow (2016). (In Russian)
5. Koulaguin V.V. On minimum of function with extra variable. // Report. International conference "Constructive Nonsmooth Analysis and Related Topics". Euler International Mathematical Institute. Saint-Petersburg (2017).
6. Koulaguin V.V., Slesar N.O. On optimal control design under uncertainty. // Report. International Conference "Differential Equations and Topology". Lomonosov Moscow State University. Moscow (2008). (In Russian)
7. Koulaguin V.V. Maximal admissible nonparametric uncertainty in a control design problem. *In: Proceedings of 11th IFAC international workshop "Control Applications of Optimization". St.-Petersburg, Russia. Published by Elsevier Science in PERGAMON for IFAC*, pp.193-198, (2001).
8. Koulaguin V.V. The problem of optimal control design under phase constraints and information deficit. *In the collection "Mathematical problems in the analysis of nonsmooth models". The Saint-Petersburg University Press*, pp 159-170 (1995). (In Russian)
9. Koulaguin V.V., Slesar N.O. Maximal robustness relocation of single mass point with a minimal velocity within the segment of real axis during the preassigned time. *Vtstnik of Nigniy Novgorod University*, No. 4(2), 189–190 (2011). (In Russian)

10. Koulaguin V.V. Reliability as robustness. Maximum robustness problem. // Report. IV International conference "Mathematical methods of reliability theory (MMR-2009)". Moscow (2009). (In Russian)
11. Koulaguin V.V., Prourzin V.A. Maximal robubustness shock absorber for an uncertain mass of object. *Mechanics of Solids, a Journal of Russian Academy of Sciences*, No. 1, (2005). (In Russian)
12. Koulaguin V.V. Two-person zero-sum game with a limited payment. *In: Proceedings of the International Conference in memory of V.I. Zubov "Stability and control processes". Saint-Petersburg University.* Saint-Petersburg (2005). (In Russian)
13. Koulaguin V.V. Game equilibrium as the equivalence of double problems for a function of two variables. *In: Proceedings of V International Conference on Operations Research (ORM 2007)*, pp 278-281. Moscow (2007). (In Russian)

A new two-stage non-Euclidean proximal method of solving the problem of equilibrium programming*

V.V. Semenov
Taras Shevchenko Kiev National University, Kiev, Ukraine

Consider the equilibrium problem for nonempty convex closed set $C \subseteq \mathbb{R}^d$ and bifunction $F : C \times C \to \mathbb{R}$:

$$\text{find} \;\; x \in C \;\; \text{such that} \quad F(x, y) \geq 0 \quad \forall\, y \in C, \tag{1}$$

where $F(y, y) = 0$ for all $y \in C$.

The equilibrium problem (1) (problem of equilibrium programming, Ky Fan inequality) is very general in the sense that it includes, as special cases, many applied mathematical models such as: variational inequalities, fixed point problems, optimization problems, saddle point problems, and Nash equilibrium point problems.

In this report, we propose and analyze a new iterative method for solving the equilibrium problem. Namely, using the Bregman distance (Bregman divergence) instead of the Euclidean we modifies the two-stage proximal algorithm from [1, 2]. The Bregman distance allows to take into account the geometry of an admissible set effectively in some

*This research is supported by the Ministry of Education and Science of Ukraine (project 0116U004777).

cases. We note that in the particular case of a variational inequality the obtained algorithm coincides with the version of the method of mirror descent recently studied by the author [3, 4].

We assume that the bifunction F satisfies the following conditions:

(A1) for all x, $y \in C$ from $F(x,y) \geq 0$ it follows that $F(y,x) \leq 0$ (pseudo-monotonicity);

(A2) $F : C \times C \to \mathbb{R}$ is lower semicontinuous on $C \times C$;

(A3) for all $x \in C$ the function $F(x, \cdot)$ in convex on C;

(A4) for all $y \in C$ the function $F(\cdot, y)$ is upper semicontinuous on C;

(A5) for all x, y, $z \in C$ the next inequality holds

$$F(x,y) \leq F(x,z) + F(z,y) + a\left\|x - z\right\|^2 + b\left\|z - y\right\|^2,$$

where a, b are positive constants (the Lipschitz-type property).

Let us consider the dual equilibrium problem:

$$\text{find} \;\; y \in C \;\; \text{such that} \quad F(x,y) \leq 0 \quad \forall x \in C. \tag{2}$$

The sets of solutions for problems (1) and (2) we denote as S and S^*. In the considered case the sets S and S^* are equal and are convex and closed. Further, we assume that the solution set S is nonempty.

Let's recall some facts about Bregman distance. Let $\varphi : \mathbb{R}^d \to \mathbb{R}$ be the continuous convex function on C, continuously differentiable on C. Assume that the function φ is strongly convex with the parameter $\sigma > 0$ in the norm $\|\cdot\|$, i. e.

$$\varphi(a) - \varphi(b) \geq (\nabla\varphi(b), a - b) + 2^{-1}\sigma\|a - b\|^2 \quad \forall a \in C, \; b \in C.$$

The Bregman divergence (generated by function φ) on the set C is defined by

$$D_\varphi(a,b) = \varphi(a) - \varphi(b) - (\nabla\varphi(b), a - b) \quad \forall a \in C, \; b \in C.$$

The useful three-point identity holds

$$D_\varphi(a,c) = D_\varphi(a,b) + D_\varphi(b,c) + (\nabla\varphi(b) - \nabla\varphi(c), a - b).$$

The minimization problem

$$F(a,y) + \lambda^{-1}D_\varphi(y,b) \to \min\nolimits_{y \in C} \quad (a, b \in C, \; \lambda > 0)$$

always has only one solution. Suppose, that we have an ability to solve efficiently this problem. For example, it is possible in the case of probability simplex, linearity F for the second argument and the Kullback-Liebler divergence.

Now, we introduce the following iterative algorithm for solving of the equilibrium problem (1).

Algorithm 1.

For x_1, $y_1 \in C$ generate the sequences of elements x_n, $y_n \in C$ with the iterative scheme

$$\begin{cases} x_{n+1} = \operatorname{argmin}_{y \in C} \left(F(y_n, y) + \frac{1}{\lambda} D_\varphi(y, x_n)\right), \\ y_{n+1} = \operatorname{argmin}_{y \in C} \left(F(y_n, y) + \frac{1}{\lambda} D_\varphi(y, x_{n+1})\right), \end{cases}$$

where $\lambda > 0$.

In Algorithm 1, at each iterative step, we must solve two optimization programs onto C with strongly convex functions.

If $\varphi(\cdot) = \frac{1}{2}\|\cdot\|_2^2$, then the Algorithm 1 takes the form

$$\begin{cases} x_{n+1} = \operatorname{prox}_{\lambda F(y_n, \cdot)} x_n, \\ y_{n+1} = \operatorname{prox}_{\lambda F(y_n, \cdot)} x_{n+1}, \end{cases} \tag{3}$$

where prox_g is the proximal operator, associated with convex lower semicontinuous proper function g

$$\mathbb{R}^d \ni x \mapsto \operatorname{prox}_g x = \operatorname{argmin}_{y \in \operatorname{dom} g} \left(g(y) + \frac{1}{2}\|y - x\|_2^2\right) \in \operatorname{dom} g.$$

Two-step proximal algorithm (3) was introduced in [1]. Also, it was reported at the VIII Moscow International Conference on Operations Research [2]. In the special case of variational inequality problem, i. e., if $F(x, y) = (Ax, y - x)$, it takes the form

$$\begin{cases} x_1 \in C, \ y_1 \in C, \\ x_{n+1} = P_C(x_n - \lambda A y_n), \\ y_{n+1} = P_C(x_{n+1} - \lambda A y_n), \end{cases}$$

where P_C is the operator of metric projection onto the set C.

For variational inequalities the Algorithm 1 has the form

$$\begin{cases} x_1 \in C, \ y_1 \in C, \\ x_{n+1} = \Pi_C \left((\nabla\varphi)^{-1}(\nabla\varphi(x_n) - \lambda A y_n)\right), \\ y_{n+1} = \Pi_C \left((\nabla\varphi)^{-1}(\nabla\varphi(x_{n+1}) - \lambda A y_n)\right), \end{cases}$$

where Π_C is the Bregman projection operator onto the set closed convex C defined by the rule

$$\Pi_C x = \operatorname{argmin}_{y \in C} D_\varphi(y, x).$$

This method was introduced in [3]. Its convergence was discussed at the conference CNSA–2017 [4].

We note first, that if for some number $n \in \mathbb{N}$ next equalities are satisfied

$$x_{n+1} = x_n = y_n \tag{4}$$

than $y_n \in S$ and the following stationarity condition holds

$$y_k = x_k = y_n \quad \forall\, k \geq n.$$

Further, we assume that for all numbers $n \in \mathbb{N}$ the condition (4) doesn't hold.

Lemma 1. Let sequences (x_n), (y_n) be generated by the Algorithm 1, and let $z \in S^*$. Then, we have

$$D_\varphi(z, x_{n+1}) \leq D_\varphi(z, x_n) - \left(1 - \frac{2\lambda b}{\sigma}\right) D_\varphi(x_{n+1}, y_n) -$$
$$- \left(1 - \frac{4\lambda a}{\sigma}\right) D_\varphi(y_n, x_n) + \frac{4\lambda a}{\sigma} D_\varphi(x_n, y_{n-1}).$$

Lemma 2. Let $\lambda \in \left(0, \frac{\sigma}{2(2a+b)}\right)$. Then all limit points of the sequence (x_n) belong to the set S.

Lemma 2 is valid without the assumption (A1) about the pseudomonotonicity of the bifunction F.

Lemma 3. Sequences (x_n), (y_n) generated by the Algorithm 1 converge to the solution $\bar{z} \in S$ of the problem (1).

Summing up, we formulate the main result.

Theorem 1. Let $C \subseteq \mathbb{R}^d$ be a nonempty convex closed set, for bifunction $F : C \times C \to \mathbb{R}$ the conditions (A1)–(A5) are satisfied and $S \neq \emptyset$. Assume that $\lambda \in \left(0, \frac{\sigma}{2(2a+b)}\right)$. Then sequences (x_n), (y_n) generated by the Algorithm 1 converge to the solution $\bar{z} \in S$ of the equilibrium problem (1).

References

1. Lyashko S.I., Semenov V.V. A New Two-Step Proximal Algorithm of Solving the Problem of Equilibrium Programming // Optimization and Its Applications in Control and Data Sciences. SOIA, vol. 115. Cham: Springer, 2016. P. 315–325.

2. Semenov V.V. A two-step proximal algorithm of solving the problem of equilibrium programming // VIII Moscow International Conference on Operations Research (ORM2016): Moscow, October 17–22, 2016: Proceedings: Vol. I. Moscow: MAKS Press, 2016. P. 61.
3. Semenov V.V. A Version of the Mirror descent Method to Solve Variational Inequalities // Cybernetics and Systems Analysis. 2017. V. 53. P. 234–243.
4. Semenov V.V. A variant of mirror descent method for solving variational inequalities // Constructive Nonsmooth Analysis and Related Topics (dedicated to the memory of V.F. Demyanov), 2017. doi: 10.1109/CNSA.2017.7974011

Accelerated subgradient method with Polyak's step[*]

P.I. Stetsyuk

V.M. Glushkov Institute of Cybernetics, Kiev, Ukraine

1. Formulation of the problem. Let $f(x)$ be a convex function, $x \in R^n$. Denote its minimal value as $f^* = f(x^*)$ and, without loss of generality, assume that the point x^* is the unique minimum point. A subgradient $\partial f(x)$ satisfies the following condition:

$$(x - x^*, \partial f(x)) \geq f(x) - f^*, \quad \forall x \in R^n. \tag{1}$$

Here (x, y) is the scalar product of vectors $x \in R^n$ and $y \in R^n$.

If f^* is known, then to find an approximation to the point $x^* \in R^n$ one can use the Polyak's subgradient method [1]:

$$x_{k+1} = x_k - h_k \frac{\partial f(x_k)}{\|\partial f(x_k)\|}, \quad h_k = \frac{f(x_k) - f^*}{\|\partial f(x_k)\|}, \quad k = 0, 1, 2, \ldots. \tag{2}$$

The step h_k is called the Polyak's step (or the Agmon-Motzkin-Schoenberg step). For the first time, Polyak's step was used for minimization of piecewise linear convex functions. In 1954 Agmon [2] and Motzkin, Schoenberg [3] used this step in relaxation method for finding at least one of the solutions of feasible system of linear inequalities. In 1965 Eremin [4] generalized this relaxation method for the systems of convex inequalities.

[*]This research is supported by Volkswagen Foundation, grant No 90 306.

The geometric sense of the method (2) is the following. The function $f(x)$ is approximated by a linear function $\tilde{f}(x) = f(x_k)+(\partial f(x_k), x-x_k)$ and the step is selected so that this approximation function becomes equal to f^* (i. e. $\tilde{f}(x_{k+1}) = f^*$). For convex function $f(x)$ the step h_k determines the value of the maximum shift in the direction of the normalized antisubgradient, which under condition (1) guarantees that angle between the antisubgradient and the direction from the point x_{k+1} to the minimum point will not be obtuse.

Consider the Polyak's subgradient method and its accelerated version for finding an approximation to the minimum point of ravine convex functions. As a stopping criterion, we use condition $f(x_k) - f^* \leqslant \varepsilon$; for an arbitrarily small $\varepsilon > 0$ it allows us to find the point $x^*_\varepsilon = x_k$ $f(x^*_\varepsilon) \leqslant f^* + \varepsilon$. We consider methods for a more general case of a convex function $f(x)$, when its subgradient $\partial f(x)$ satisfies the following condition:

$$(x - x^*, \partial f(x)) \geq m(f(x) - f^*), \quad \forall x \in R^n, \tag{3}$$

where parameter $m \geqslant 1$. Here the parameter m is introduced to take into account special classes of convex functions: for example, for a piecewise linear nonsmooth function $m = 1$, for a quadratic smooth function $m = 2$.

2. Method A (the Polyak's subgradient method). If the convex function $f(x)$ satisfies condition (3) and f^* is known, then to find the point $x^*_\varepsilon \in R^n$ such that $f(x^*_\varepsilon) \leqslant f^* + \varepsilon$, one can use the following iterative method.

Initialization. Let f^* and $m \geqslant 1$ are given. Let's select the point $x_0 \in R^n$ and the value $\varepsilon > 0$ and go to the next iteration with the value x_0.

Iterative process. Let the point $x_k \in R^n$ be found on the k-th iteration. To proceed to the $(k + 1)$-th iteration, we perform the following actions.

A1. Calculate $f(x_k)$ and $\partial f(x_k)$. If $f(x_k)-f^* \leqslant \varepsilon$, then STOP ($k^* = k$, $x^*_\varepsilon = x_k$).

A2. Calculate the next point

$$x_{k+1} = x_k - h_k \frac{\partial f(x_k)}{\|\partial f(x_k)\|}, \quad h_k = \frac{m(f(x_k) - f^*)}{\|\partial f(x_k)\|},$$

A3. Go to the $(k + 1)$-th iteration with x_{k+1}.

Theorem 1 *The sequence* $\{x_k\}_{k=0}^{k^*-1}$ *generated by method A satisfies the inequalities*

$$\|x_{k+1} - x^*\|^2 \leqslant \|x_k - x^*\|^2 - \frac{m^2(f(x_k) - f^*)^2}{\|\partial f(x_k)\|^2}, \quad k = 0, 1, 2, \ldots \tag{4}$$

Proof. From A2 for an arbitrary k $(0 \leqslant k \leqslant k^* - 1)$ we have

$$\|x_{k+1} - x^*\|^2 = \left\| x_k - x^* - h_k \frac{\partial f(x_k)}{\|\partial f(x_k)\|} \right\|^2 =$$
$$= \|x_k - x^*\|^2 - 2h_k \frac{(x_k - x^*, \partial f(x_k))}{\|\partial f(x_k)\|} + h_k^2.$$

Taking into account that from (3) it follows

$$\frac{\big(x_k - x^*, \partial f(x_k)\big)}{\|\partial f(x_k)\|} \geqslant \frac{m\big(f(x_k) - f^*\big)}{\|\partial f(x_k)\|} = h_k,$$

we have

$$\|x_{k+1} - x^*\|^2 \leqslant \|x_k - x^*\| - h_k^2 = \|x_k - x^*\|^2 - \left(\frac{m(f(x_k) - f^*)}{\|\partial f(x_k)\|}\right)^2,$$

that gives the inequalities (4). Theorem is proved.

Theorem 1 guarantees that in Polyak's subgradient method the distance to the minimum point decreases monotonically.

The disadvantage of method A is its slow convergence for ravine functions. For example, for nonsmooth function in two variables $f(x_1, x_2) = |x_1| + t\,|x_2|$, $t > 1$, the rate of convergence of method A is determined by the geometric progression with the common ratio

$$q = \sqrt{1 - 1/t^2} \tag{5}$$

and will be very slow for large values of t. An analogous situation holds also when minimizing the ravine quadratic function $f(x_1, x_2) = (x_1)^2 + t\,(x_2)^2$, $t \gg 1$. For example, if $t = 100$, $m = 2$ and $\varepsilon = 0.01$, then method A generates the sequence $x_0 = (1.00, 1.00)^T$, $x_1 = (0.99, -0.01)^T$, $\ldots$, $x_5 = (0.238, -0.002)^T$, $\ldots$, $x_8 = (0.058, 0.058)^T$, $x_9 = (0.057, -0.001)^T$.

Below we consider the modification of the subgradient method with Polyak's step, where the accelerated convergence with respect to method A can be provided by choosing the space transformation matrix.

3. Method B (subgradient method with Polyak's step in the transformed space of variables). Let's make the substitution of variables $x = By$, where B is a nonsingular $n \times n$-matrix (that is, there exists an inverse matrix $A = B^{-1}$). The subgradient $\partial\varphi(y)$ of the convex function $\varphi(y) = f(By)$ at the point $y = Ax$ of the transformed space of variables satisfies inequality

$$(y - y^*, \partial\varphi(y)) \geq m(\varphi(y) - \varphi^*), \quad \forall y \in R^n, \tag{6}$$

where $\partial\varphi(y) = B^T\partial f(x)$, $\varphi^* = \varphi(y^*) = f(By^*)$, $y^* = Ax^*$. Indeed, since $A = B^{-1}$ and $x = By$, inequality (3) can be rewritten in the form

$$(A(x - x^*), B^T\partial f(x)) \geq m(f(By) - f(By^*)), \quad \forall By \in R^n,$$

whence we obtain inequality (6).

To find the point x^*, the subgradient method with the Polyak's step in the transformed space (defined by the nonsingular matrix B) has the following form:

$$x_{k+1}=x_k-h_kB\frac{B^T\partial f(x_k)}{\|B^T\partial f(x_k)\|},\ h_k=\frac{m\,(f(x_k)-f^*)}{\|B^T\partial f(x_k)\|},\ k = 0, 1, 2, \ldots. \tag{7}$$

Here h_k is the Polyak's step (the Agmon-Motzkin-Schoenberg step), but in the transformed space of variables $y = Ax$. This follows from the fact that in the transformed space of variables method (7) is written as a subgradient process

$$y_{k+1} = y_k - h_k\frac{\partial\varphi(y_k)}{\|\partial\varphi(y_k)\|},\ h_k = \frac{m\big(\varphi(y_k) - \varphi^*\big)}{\|\partial\varphi(y_k)\|},\ k = 0, 1, 2, \ldots. \tag{8}$$

The Polyak's step in the transformed space of variables has the same properties as the Polyak's step in the original space. They are defined by an a priori knowledge of the minimum value of the function and the inequality (6) associated with it.

To find the point $x^*_\varepsilon \in R^n$ for which $f(x^*_\varepsilon) \leqslant f^* + \varepsilon$, the subgradient method with the Polyak's step in the transformed space is represented by the next iterative procedure.

Initialization. We have f^* and $m \geqslant 1$. We choose the point $x_0 \in R^n$, the nonsingular $n \times n$ matrix B, and the value $\varepsilon > 0$. We move to the next iteration with the value x_0.

Iterative process. Let $x_k \in R^n$ be found on the k-th iteration. To proceed to the $(k + 1)$-th iteration, we perform the following actions.

B1. Calculate $f(x_k)$ and $\partial f(x_k)$. If $f(x_k)-f^* \leqslant \varepsilon$, then STOP ($k^* = k$, $x^*_\varepsilon = x_k$).

B2. Calculate the next point

$$x_{k+1} = x_k - h_k B \frac{B^T \partial f(x_k)}{\|B^T \partial f(x_k)\|}, \quad h_k = \frac{m(f(x_k) - f^*)}{\|B^T \partial f(x_k)\|},$$

B3. Go to the $(k+1)$-th iteration with x_{k+1}.

Theorem 2 *The sequence $\{x_k\}_{k=0}^{k^*-1}$ generated by method B satisfies the inequalities*

$$\|A(x_{k+1} - x^*)\|^2 \leqslant \|A(x_k - x^*)\|^2 - \frac{m^2(f(x_k) - f^*)^2}{\|B^T \partial f(x_k)\|^2}, \quad k = 0, 1, \dots \quad (9)$$

Proof. From B2 for an arbitrary k ($0 \leqslant k \leqslant k^* - 1$) we have

$$\|A(x_{k+1} - x^*)\|^2 = \left\| A(x_k - x^*) - h_k \frac{B^T \partial f(x_k)}{\|B^T \partial f(x_k)\|} \right\|^2 =$$
$$= \|A(x_k - x^*)\|^2 - 2h_k \frac{(x_k - x^*, \partial f(x_k))}{\|B^T \partial f(x_k)\|} + h_k^2.$$

Taking into account that from (3) it follows the inequality

$$\frac{\big(x_k - x^*, \partial f(x_k)\big)}{\|B^T \partial f(x_k)\|} \geqslant \frac{m\big(f(x_k) - f^*\big)}{\|B^T \partial f(x_k)\|} = h_k,$$

we have

$$\|A(x_{k+1} - x^*)\|^2 \leqslant \|A(x_k - x^*)\| - h_k^2 =$$
$$= \|A(x_k - x^*)\|^2 - \left(\frac{m(f(x_k) - f^*)}{\|B^T \partial f(x_k)\|} \right)^2,$$

which gives the inequalities (9). Theorem is proved.

Theorem 2 guarantees that in subgradient method (8) with Polyak's step in the transformed space of variables the distance to the minimum point decreases monotonically in the transformed space.

If the matrix B is chosen such that in the transformed space of variables the level surfaces of ravine functions are less elongated than in the original space of variables, then method B will converge faster than method A [5, 6]. For example, if the matrix $B = \mathrm{diag}(1, 0.5)$, then for function $f(x_1, x_2) = |x_1| + t\,|x_2|$, $t > 1$, the rate of convergence of method B is determined by the geometric progression with the common ratio $q' = \sqrt{1 - 4/t^2}$, which is greater than $q = \sqrt{1 - 1/t^2}$ for method A (see the formula (5)). If $m = 2$, $B = \mathrm{diag}(1, 0.5)$ and $\varepsilon = 0.01$, then while minimizing function $f(x_1, x_2) = (x_1)^2 + 100\,(x_2)^2$ method B generates sequence $x_0 = (1.00, 1.00)^T$, $x_1 = (0.96, -0.01)^T$, $x_2 = (0.184, 0.184)^T$, $x_3 = (0.177, -0.002)^T$, $x_4 = (0.034, 0.034)^T$, $x_5 = (0.033, -0.000)^T$.

References

1. Polyak B.T. Minimization of unsmooth functionals // USSR Comput. Math. Math. Phys. 1969. V. 9, N 3. P. 14–29.
2. Agmon S. The relaxation method for linear inequalities // Canadien Journal of Mathematics. 1954. V. 6. P. 382–392.
3. Motzkin T., Schoenberg I.J. The relaxation method for linear inequalities // Canadien Journal of Mathematics. 1954. V. 6. P. 393–404.
4. Eremin I.I. Generalization of the Motzkin-Agmon relaxational method // Uespekhi Mat. Nauk. 1965. V. 20, N 2. P. 183–187.
5. Stetsyuk P.I. Acceleration of Polyak's subgradient method // Teoriia Optymalnykh Rishen. 2012. P. 151–160.
6. Stetsyuk P.I. Methods of Ellipsoids and r-Algorithms. Chisinau, Moldova: Eureka, 2014.

Multiplicatively Barrier Methods for Linear Cone Programming Problems[*]

V.G. Zhadan

Dorodnicyn Computing Centre,

FRC “Computer Science and Control” of RAS, Moscow, Russia

The linear cone programming problem is considered. This problem is the convex optimization problem in which a linear function is minimized over the intersection of the affine linear manifold with the convex closed cone [1]. Important cases of such cones are second order cone and the cone of positively semidefinite symmetric matrices [2]. Of particular interest are the cone programming problems with Cartesian product of various cones.

[*] This research was supported by the Russian Foundation for Basic Research (project no. 17-07-00510).

Numerical methods for solving cone programming problems are usually constructed by adjusting to these problems appropriate methods from linear programming. In this paper we give the generalizations of primal affine scaling algorithm and Newton's method proposed earlier for solving linear and nonlinear programming problems. Both these methods can be treated as special ways for solving optimality conditions for primal and dual cone problems. Variants of these methods for problems with second order cones were considered in [3,4]. The convergence of both methods is discussed.

1. Problem formulation and optimality conditions. Assume that in $\mathbb{R}^n$ there is the convex closed pointed cone K with nonempty interior $K_0 = \mathrm{int}K$, which induces in $\mathbb{R}^n$ a partial order: $x_1 \succeq_K x_2$, iff $x_1 - x_2 \in K$. The strong inequality $x_1 \succ_K x_2$ means that $x_1 - x_2 \in \mathrm{int}\, K$.

The linear cone programming problem is

$$\min \langle c, x\rangle, \quad \mathcal{A}x = b, \quad x \in K, \tag{1}$$

where $\mathcal{A}$ is a $m \times n$ matrix, $c = [c^1; \ldots; c^n] \in \mathbb{R}^n$ and $b = [b^1; \ldots; b^m] \in \mathbb{R}^m$ are nonzero vectors. Here and in what follows a semicolon in the enumeration of vectors or components of a vector indicates that one of them is placed under another. The dual problem to (1) is

$$\max \langle b, u\rangle, \quad v = c - \mathcal{A}^T u, \quad v \in K^*, \tag{2}$$

where $K^* = \{y \in K : \langle x, y\rangle \geq 0\}$ is a dual cone to K. We assume that both problems (1) and (2) have solutions and the rows of the matrix $\mathcal{A}$ are linear independent. We denote also the feasible set in problem (1) by X_P. Usually in the standard formulation of the cone programming problem set that $x = [x_1; \ldots; x_r]$ and $K = K_1 \times \cdots \times K_r$, where $x_i \in K_i$ and K_i is a convex closed cone in $\mathbb{R}^{n_i}$.

The necessarily and sufficient optimality conditions for problems (1) and (2) can be written as

$$\langle x, v\rangle = 0, \qquad \mathcal{A}x = b, \qquad v = c - \mathcal{A}^T u, \tag{3}$$

in which $x \in K$, $v \in K^*$.

The equality $\langle x, v\rangle = 0$ with regard for inclusions $x \in K$, $v \in K^*$ can be replace by n other inequalities. Assume that there is the differentiable mapping $x = \xi(y)$ such that the image $\xi\left(\mathbb{R}^n\right)$ of the space $\mathbb{R}^n$ coincides with K. Then omitting inclusion $x \in K$, we derive the problem

$$\min \tilde{f}(y), \quad \tilde{g}(y) = 0_m, \quad y \in \mathbb{R}^n, \tag{4}$$

where $\tilde{f}(y) = \langle c, \xi(y) \rangle$ and $\tilde{g}(y) = b - \mathcal{A}\xi(y)$. If y_* is a solution of problem (4), then $x_* = \xi(y_*)$ is a solution of (1).

Introduce the Lagrange function for problem (1) $\tilde{L}(y,u) = \tilde{f}(y) + \langle u, \tilde{g}(y) \rangle$, where $y \in \mathbb{R}^n$, $u \in \mathbb{R}^m$, and denote by $\tilde{J}(y) = \xi_y(y)$ the Jacobian of the mapping $\xi(y)$. It follows from Karush–Kuhn–Tucker conditions for problem (4) that

$$\tilde{L}_y(y,u) = \tilde{f}_y(y) + \tilde{g}_y^T(y)u = 0_n, \quad \tilde{L}_u(y,u) = \tilde{g}(y) = 0_m. \tag{5}$$

Let $f(x) = \langle c, x \rangle$ and $g(x) = b - \mathcal{A}x$. Derivatives of the functions $\tilde{f}(y)$, $\tilde{g}(y)$ and $f(x)$, $g(x)$ are connected between themselves by relations: $\tilde{f}_y(y) = \tilde{J}^T(y) f_x(\xi(y)$, $\tilde{g}_y^T(y) = \tilde{J}^T(y) g_x^T(\xi(y)$. Therefore, introducing into consideration the Lagrange function $L(x,u) = f(x) + \langle u, g(x) \rangle$, we get from (5) the equalities: $\tilde{J}^T(y) L_x(\xi(y), u) = 0_n$, $L_u(\xi(y), u) = 0_m$.

If y is a nonsingular point of the mapping $\xi(y)$, then in some neighborhood of $x = \xi(y)$ there is the inverse mapping $y = \xi^{-1}(x)$, and we obtain in x-space

$$J^T(x) L_x(x,u) = 0_n \quad L_u(x,u) = g(x) = 0_m. \tag{6}$$

Here $J(x) = \tilde{J}(\xi^{-1}(x))$. Hence, the vector $L_x(x,u) = c - \mathcal{A}^T u = v$ must belong to the null-space of the matrix $J^T(x)$.

Denote by $\mathcal{F}(x|K)$ the cone of feasible directions with respect to a cone K at the point $x \in K$, and denote by $\mathcal{F}^*(x|K)$ the dual cone to $\mathcal{F}(x|K)$. Moreover, let $\mathcal{N}(x|J^T)$ be the null-space of the matrix $J^T(x)$. We impose at the mapping $\xi(y)$ the following condition.

The compatibility condition. *Matrix $J(x)$ is defined in some domain containing the cone K, and at every point $x \in K$ the equality*

$$\mathcal{N}(x|J^T) = \text{lin } \mathcal{F}^*(x|K),$$

takes place, where lin $\mathcal{F}^*(x|K)$ *is a linear hull of the cone $\mathcal{F}^*(x|K)$.*

Proposition 1. *Let the differentiable mapping $\xi(y) : \mathbb{R}^n \to K$ be such that the compatibility condition holds for the matrix $J(x)$. Then for $x \in K$ and $v \in K^*$ the equality $\langle x, v \rangle = 0$ is valid iff*

$$G(x)v = G(x)\left(c - \mathcal{A}^T u\right) = 0_n, \tag{7}$$

where $G(x)$ is an arbitrary square matrix with the null-space, coinciding with the null-space of the matrix $J^T(x)$.

Remark. *If $x \in intK$, then the matrix $G(x)$ must be nonsingular. This matrix is singular iff x is a boundary point of K.*

In cone programming of particular important are convex cones which can be represent with the help of quadratic mappings (see [2]). Consider two examples of such cones K. They are self-dual, i.e. $K^* = K$.

1) $K = \mathbb{R}^n_+$. This cone is polyhedral. The problem (1) with the cone $\mathbb{R}^n_+$ is the ordinary linear programming problem. Introduce the Hadamard product $x \circ y$ between vectors $x, y \in \mathbb{R}^n$

$$x \circ y = [x_1 y_1;\ \ldots;\ x_n y_n], \tag{8}$$

and denote by $x^2 = x \circ x$. Then it is possible to define the quadratic mapping $x = \xi(y) = \frac{1}{2}y^2$. The cone $\mathbb{R}^n_+$ is the image of the space $\mathbb{R}^n$ under the mapping $\xi(y)$. The corresponding Jacobian $J(x)$ has the diagonal form $J(x) = \sqrt{2}D^{1/2}(x)$, where $D(x)$ is a diagonal matrix with a vector x at its diagonal. For $x, v \geq 0_n$, the first equality from (3) holds iff $x \circ v = 0_n$. Taking for example the matrix $G(x) = D(x)$, we have

$$x \circ v = G(x)\, v = G(v)\, x = G(x)\, G(v)\, \bar{e}. \tag{9}$$

Here $\bar{e}$ is n-dimensional vector of ones.

2) $K = K^n_2$, where K^n_2 is the second order (Lorentz) cone in $\mathbb{R}^n$:

$$K^n_2 = \left\{x = [x_0, \bar{x}] \in \mathbb{R} \times \mathbb{R}^{n-1} :\ x_0 \geq \|\bar{x}\|\right\}.$$

Here $\|\cdot\|$ is the standard Euclidean norm. Consider now instead of (8) the following product between vectors $x, y \in \mathbb{R}^n$ (see [2]):

$$x \circ y = \left[\begin{array}{c} x^T y \\ x_0 \bar{y} + y^0 \bar{x} \end{array}\right]. \tag{10}$$

Let $x^2 = x \circ x$. It can be inspected that the cone K^n_2 is the image of the space $\mathbb{R}^n$ under the quadratic mapping $x = \xi(y) = \frac{1}{2}y^2$. Again we obtain that for $x \in K^n_2$ and $v \in K^n_2$ the first equality from (3) holds iff $x \circ v = 0_n$.

We have for $x = \frac{1}{2}y^2$ with the product defined by (10):

$$x(y) = \frac{1}{2}\left[\begin{array}{c} y_0^2 + \|\bar{y}\|^2 \\ 2y_0\bar{y} \end{array}\right], \qquad \tilde{J}(y) = \mathrm{Arr}(y) = \left[\begin{array}{cc} y_0 & \bar{y}^T \\ \bar{y} & y_0 I_{n-1} \end{array}\right].$$

At nonzero boundary points K^n_2 the Jacobian has the form $J(x) = \sqrt{x_0}\mathrm{Arr}(x)$. Hence, it is reasonable to take $\mathrm{Arr}(x)$ as the matrix $G(x)$. For this $G(x)$ formula (9) is preserved.

The numerical algorithms. Using the proposition 1, we can rewrite the optimality conditions (3) in the following form

$$G(x)v = 0_n, \quad \mathcal{A}x = b, \quad v = c - \mathcal{A}^T u, \quad x \in K, \quad v \in K^*. \tag{11}$$

Substituting v into the first equality and multiplying it by the matrix $\mathcal{A}$, we obtain $\mathcal{A}G(x)\mathcal{A}^T u = \mathcal{A}G(x)c$. Add to this equality the second equality from (11) previously multiplied by a certain coefficient $\tau > 0$. As a result we get the equation with respect to the dual variable u:

$$\Gamma(x)u = \mathcal{A}G(x)c + \tau(b - \mathcal{A}x), \qquad \Gamma(x) = \mathcal{A}G(x)\mathcal{A}^T.$$

If the matrix $\Gamma(x)$ is nonsingular, then resolving this equation, we find

$$u = u(x) = \Gamma^{-1}(x)\left[\mathcal{A}G(x)c + \tau(b - \mathcal{A}x)\right].$$

Moreover, we have for weak dual variable $v = v(x) = c - \mathcal{A}^T u(x)$. Hence, the equality $G(x)v(x) = 0_n$ can be rewritten as

$$G(x)\left\{\left[I_n - \mathcal{A}^T\Gamma^{-1}(x)\mathcal{A}G(x)\right]c + \tau\mathcal{A}^T\Gamma^{-1}(x)(\mathcal{A}x - b)\right\} = 0_n. \tag{12}$$

This is the system of n nonlinear equations with respect to variable x.

Various numerical techniques for solving nonlinear equations can be applied for finding the solution of (12). In particular, using the simple iteration method, we come to the following iterative process:

$$x_{k+1} = x_k - \alpha_k G(x_k)\left\{\left[I_n - \mathcal{P}(x)G(x_k)\right]c + \tau\mathcal{A}^T\Gamma^{-1}(x_k)(\mathcal{A}x_k - b)\right\}, \tag{13}$$

where $x_0 \in K$ and $\mathcal{P}(x) = \mathcal{A}^T\Gamma^{-1}(x_k)\mathcal{A}$. The step length α_k is chosen by some manner, for example, it can be taken constant and sufficiently small.

Applying the Newton method for solving the system of equations (12), we obtain another iterative process

$$x_{k+1} = x_k - \Lambda^{-1}(x_k)G(x_k)v_k, \quad v_k = v(x_k), \tag{14}$$

where $\Lambda(x)$ is the Jacobian of $W(x) = G(x)v(x)$. We have

$$\Lambda(x) = \left[I_n - G(x)\mathcal{P}(x)\right]G(v(x)) + \tau G(x)P(x),$$

If the point $x \in X_P$ is non-degenerate, then matrices $\Gamma(x)$ and $\Lambda(x)$ are nonsingular.

Definition [5]. *Let $F_{\min}(x; K)$ be the minimal face of K, containing the point x, and let $F^*_{\min}(x; K)$ be the conjugate face to $F_{\min}(x; K)$. The pair $[x, v]$ is called strictly complementary, if $x \in \mathrm{ri} F_{\min}(x; K)$, $v \in \mathrm{ri} F^*_{\min}(x; K)$.*

Theorem. *Suppose that solutions x_* and $v_* = v(u_*)$ of primal and dual problems* (1) *and* (2) *satisfy the strict complementary condition. Then the iterative process* (13) *converges to x_* with a linear rate. The iterative process* (14) *converges to x_* at superlinear rate.*

References

1. Handbook of Semidefinite, Cone and Polynomial Optimizatin: Theory, Algorithms, Software and Applications, Anjos M.F., Lasserre J.B, eds. New York, USA: Springer. 2011.
2. Alizadeh F., Goldfarb D. Second-order cone programming // Math. Program., Ser. B. 2003. V. 95. P. 3–52.
3. Zhadan V.G. Primal Newton Method for the Linear Cone Programming Problem // Computational Mathematics and Mathematical Physics. 2018. V. 58, No 2. P. 207–214.
4. Zhadan V.G. Variant of Affine Scaling Method for the Cone Programming Problem with Second Order Cone // Proceedings of the Institute of Mathematics and Mechanics Ural Brunch of RAS. 2017. V. 23, No 3. P. 115–124. [in Russian]
5. Pataki G. Cone-LP's and Semidefinite Programs: Geometry and Simplex-type Method // Integer Programming and Combinatorial Optimization, Lecture Notes in Computer Science. Vol. 1084. Berlin: Springer. 1996. P. 162–174.

Multiple objective decision making

On connection between multiobjective optimization, polyhedral projection and automatic control

M.N. Demenkov
Institute of Control Sciences, Moscow, Russia

Multiobjective optimization [1,2] deals with simultaneous optimization, instead of one objective function, of multiple functions over the same constrained feasible set. Usually, the goal is to determine the Pareto, or non-dominated, set of solutions that cannot be improved in any one function without degrading of at least one another. Polyhedral projection [3] can be understood as computation of a "shadow" of convex polytope in lower dimensional space. To be precise, we consider general polyhedra — sets defined by systems of linear inequalities, and exclusion of certain variables from these inequalities so as to keep them valid for the remaining variables. The goal of automatic control [4,5] is to affect the behaviour of dynamical systems (represented by differential or difference equations and inclusions), which can describe some mechanical objects, like aircrafts or self-driving cars, or even financial systems. At first glance, all three fields seem quite disparate in nature. In fact, there are multiple connections between them, leading to new approaches in each field.

In control, operations on convex polytopes have been considered since 1980s to construct e.g. reachable and controllable sets for linear discrete systems. The precursor for this line of research was a little-known paper

[6], published in the former USSR 15 years before similar ideas appeared in [7]. According to [4,8], stabilizing regulators for differential inclusions could be constructed by projection. Finally, in the context of model predictive control (MPC, which is essentially an application of numerical optimization in real time) operations on polytopes are indispensable as well [5], including projection [9].

It is a common viewpoint that the projection of polyhedron is extremely hard to compute. It appears to be NP-hard for polyhedra defined by systems of linear inequalities [10] and the most known method for its solution, Fourier-Motzkin elimination, has double exponential complexity [11,12]. Recently, the equivalence has been established between linear multiobjective optimization and polyhedral projection [13]. If we calculate the complete exact Pareto front in the objective space (where coordinates represent values of different objective functions) of a specially constructed linear multiobjective problem, the result can be interpreted as a projection of some given polyhedron (and vice versa). The actual computations could be performed by e.g. Benson algorithm [14,15], for which software realization (www.bensolve.org) is available. It is an outer-approximation method (see e.g. [16]) with cutting planes (see [17] for its application to reachable sets). One can also try multicriteria simplex method (several variants have been proposed since the beginning of 1970s, see e.g. [2,19]). In model predictive control, the equivalence is shown between projection and multi-parametric linear programming [9], which could be solved as a linear complementarity problem [19].

In this note we want to "close the circle" and apply some tools, originally developed in automatic control, to the projection problem and therefore to the solution of linear multiobjective problem as well. For this, we study special kind of polytopes, called zonotopes. The method of zonotope construction, outlined below, has been proposed in the context of studying reachable sets [20,21] and control allocation problem [22,23].

Consider underdetermined system of linear equations $z = Gw$ (i.e. number of columns m higher than number of rows n). A zonotope is an image of the cube under an affine projection [3], that is, a polytope

$$Z = \{z \in \mathbb{R}^n : z = z_0 + \sum_{i=1}^{m} g_i w_i,\ -1 \leq w_i \leq 1\}$$

where $G = [g_1\ g_2\ ...\ g_m] \in \mathbb{R}^{n \times m}$, $m \geq n$. Column vectors $g_i \in \mathbb{R}^n$ are called "generators". For simplicity, let $z_0 = 0$. Note that for $h \in \mathbb{R}^n$

$$\max_{z \in Z} h^T z = \max_{||w||_\infty \leqslant 1} (h^T G w) = \sum_{i=1}^{m} h^T g_i \text{sign}(h^T g_i).$$

Any normal vector h_j of a zonotope facet pointing "outside" is orthogonal to some $n-1$ columns taken from the matrix G [3]:

$$h_j^T g_k = 0, k \in S_j, card(S_j) = n-1,$$

therefore we can can write the following H-representation of Z:

$$\pm h_j^T z \leq r_j, r_j = \sum_{i=1}^{m} h_j^T g_i \text{sign}(h_j^T g_i), \;\; j = 1, ..., K,$$

$$K = \left(\begin{array}{c} m \\ n-1 \end{array} \right) = \frac{m!}{(m-n-1)!(n-1)!}.$$

To compute h_j in case $n = 2$ one can swap two vector components and change the sign of one of them, for $n = 3$ the required vector is a cross product of two columns. In general case, we need to calculate one-dimensional null space.

Suppose that in addition to general linear inequalities we have interval constraint for each variable and we want to exclude variables in x_1:

$$Ax \leq b, \; Ax = A_1 x_1 + A_2 x_2, \; A = [A_1 \; A_2], \; x = [x_1^T \; x_2^T]^T, \; ||x||_\infty \leq 1.$$

The orthogonal projection onto the subspace of x_2

$$P_{x_2} = \{x_2 : \exists x_1, \; A_1 x_1 + A_2 x_2 \leq b, \; ||x_1||_\infty \leq 1, \; ||x_2||_\infty \leq 1\}.$$

Let us now introduce the vector of slack variables y (as in e.g. simplex method for LP):

$$Ax + y = b.$$

We denote as a_i the rows of $A \in \mathbb{R}^{q \times q}$ with components a_{ij} and the corresponding components of y and b as y_i and b_i. We can remove all inequalities $a_i^T x \leq b_i$, for which

$$\max_{||x||_\infty \leq 1} a_i^T x \leq b_i, \;\; \max_{||x||_\infty \leq 1} a_i^T x = \sum_{j=1}^{q} |a_{ij}|.$$

For other inequalities

$$a_i^T x + y_i = b_i, \;\; y_i \in [0, b_i - \min_{||x||_\infty \le 1} a_i^T x], \;\; \min_{||x||_\infty \le 1} a_i^T x = -\sum_{j=1}^{q} |a_{ij}|.$$

As a result, we have transformed our initial system into the form

$$z = A_1 x_1 + 0.5Dy^* = -A_2 x_2 + b - c,$$

$$w = [x_1^T \; y^{*T}]^T, \; ||w||_\infty \le 1,$$

where D is a diagonal matrix with components equal to widths of intervals for y_i, while c contains their centers. Transformation of this kind is the reverse of the projection operation and known in discrete optimization as “lift” or “extended formulation” [24].

Now, consider zonotope with $G = [A_1 \;\; 0.5D]$, $z_0 = 0$ and its H-representation

$$Hz \le r$$

where $z = -A_2 x_2 + b - c$. The polyhedron P_{x_2} is given by the following system of linear inequalities:

$$-HA_2 x_2 \le r - H(b - c), ||x_2||_\infty \le 1.$$

Let us demonstrate our approach on an elementary example:

$$x_1 = x_2, x_1 + x_2 \le 1, ||x||_\infty \le 1.$$

The projection onto x_2 is obviously given by $-1 \le x_2 \le 0.5$. In order to check if our algorithm works, introduce y so that

$$x_1 - x_2 = 0, x_1 + x_2 + y = 1, -1 \le x_{1,2} \le 1, 0 \le y \le 3$$

with $y^* = (y - 3/2)/(3/2)$, $x_1 + x_2 + 1.5y^* = -0.5$, $-1 \le y^* \le 1$ and

$$z = \begin{bmatrix} 1 & 0 \\ 1 & 1.5 \end{bmatrix} \begin{bmatrix} x_1 \\ y^* \end{bmatrix} = \begin{bmatrix} 1 \\ -1 \end{bmatrix} x_2 + \begin{bmatrix} 0 \\ -0.5 \end{bmatrix}.$$

Normals to facets are $h_1^T = \begin{bmatrix} -1 & 1 \end{bmatrix}$ and $h_2^T = \begin{bmatrix} 1 & 0 \end{bmatrix}$, $r_1 = 1.5$,$r_2 = 1$. It is easy to check that

$$(h_1^T z \le r_1) \rightarrow (x_2 \ge -1), \; (h_2^T z \le r_2) \rightarrow (x_2 \le 1),$$

$$(-h_1^T z \le r_1) \rightarrow (x_2 \le 0.5), \; (-h_2^T z) \le r_2) \rightarrow (x_2 \ge -1).$$

References

1. Lotov A.V., Bushenkov V.A., Kamenev G.K. Interactive decision maps. Approximation and visualization of Pareto frontier. Kluwer, Boston, 2004.
2. Ehrgott M. Multicriteria optimization. 2nd edition. Springer, 2005.
3. Ziegler G.M. Lectures on polytopes. Springer-Verlag, 1995.
4. Blanchini F., Miani S. Set-theoretic methods in control. Birkhäuser Boston, 2008.
5. Borrelli F., Bemporad A., Morari M. Predictive control for linear and hybrid systems. Cambridge University Press, 2017.
6. Lotov A.V. Numerical method of constructing attainability sets for a linear control system // USSR Computational Mathematics and Mathematical Physics. 1972. V. 12. No. 3. P. 279-283.
7. Keerthi S., Gilbert E. Computation of minimum-time feedback control laws for discrete-time systems with state-control constraints // IEEE Trans. on Autom. Control. 1987. V. 32. No. 5. P. 432-435.
8. Blanchini F. Non-quadratic Lyapunov functions for robust control // Automatica. 1995. V. 31. No. 3. P. 451-461.
9. Jones C.N., Kerrigan E.C., Maciejowski J.M. On polyhedral projection and parametric programming // Journal of Optimization Theory and Applications. 2008. V. 138, No 2. P. 207-220.
10. Tiwary H.R. Complexity of some polyhedral enumeration problems. Ph.D. Thesis. Saarland University, 2008.
11. Fukuda K. CDDLIB reference manual. Swiss Federal Institute of Technology, Zurich (available from https://www.inf.ethz.ch/personal/ fukudak/cdd_home/index.html).
12. Bastrakov S.I., Zolotykh Yu.N. Fast method for verifying Chernikov rules in Fourier-Motzkin elimination // Computational Mathematics and Mathematical Physics. 2015. V. 55. No. 1. P. 160-167.
13. Löhne A., Weißing B. Equivalence between polyhedral projection, multiple objective linear programming and vector linear programming // Math. Methods of Oper. Res. 2016. V. 84. No. 2. P. 411-426.
14. Benson H.P. An outer approximation algorithm for generating all efficient extreme points in the outcome set of a multiple objective

linear programming problem // J. of Glob. Optim. 1998. V. 13. P. 1-24.

15. Löhne A., Weißing B. The vector linear program solver Bensolve – notes on theoretical background // European Journal of Operational Research. 2017. V. 260. No. 3. P. 807-813.
16. Kamenev G.K. Duality theory of optimal adaptive methods for polyhedral approximation of convex bodies // Comp. Mathematics and Mathematical Physics. 2008. V. 48. No. 3. P. 376-394.
17. Shao L., Zhao F., Cong Y. Approximation of convex bodies by multiple objective optimization and an application in reachable sets // Optimization. 2018 (accepted for publication)
18. Rudloff B., Ulus F., Vanderbei R. A parametric simplex algorithm for linear vector optimization problems. Mathematical programming A. 2017. V. 163. P. 213-242.
19. Murty K.G. Linear complementarity, linear and nonlinear programming. Helderman-Verlag, 1988.
20. Girard A. Reachability of uncertain linear systems using zonotopes // Lec. Notes in Comp. Science, Springer. 2005. V. 3414. P. 29-305.
21. Althoff M., Stursberg O., Buss M. Computing reachable sets of hybrid systems using a combination of zonotopes and polytopes // Nonlinear Analysis: Hybrid Systems. 2010. V. 4. P. 233-249.
22. Durham W., Bordignon K.A., Beck R. Aircraft control allocation. John Wiley & Sons, 2017.
23. Demenkov M. Interval bisection method for control allocation // IFAC Proceedings Volumes. 2007. V. 40. No. 7. P. 183-188.
24. Kaibel V. Extended formulations in combinatorial optimization // Optima. 2011. V. 85. P. 2–7.

Increase of consistency index in paired comparisons[*]

A.E. Kurennykh, V.P. Osipov, and V.A. Sudakov
Moscow Aviation Institute (National Research University) and Keldysh Institute of Applied Mathematics (Russian Academy of Sciences), Moscow, Russia

The matrix of paired comparisons is the main tool used in the method of paired comparisons of criteria or alternatives, as well as in the Analytic hierarchy process (AHP). Both methods find wide application in

[*] This research is supported byRFBR, project 18-01-00382a.

decision support tasks, are well studied, simple and understandable for the expert using them. However, for all its clarity, the method of pairwise comparisons has one significant drawback: it exerts a high load on the expert, which is especially evident in problems of large dimension. It is well known that an ordinary human's brain can simultaneously operate with no more than seven objects, and it is quite obvious that almost all modern scientific, technical, research or management tasks are characterized by dozens of criteria that must be somehow compared and evaluated. The high load exerted on the expert leads, while filling the matrix of paired comparisons, to the following consequences: mistakes (inaccuracies) are made when determining the superiority of alternatives or criteria one on top of another. Such inaccuracies lead to a violation of the transitivity of judgments, which, in turn, leads to the use of an incorrectly composed matrix in the ranking process, and therefore there may be errors in the decisions made.

To assess the suitability of the matrix for its use in decision support, is used a special numerical index - the consistency index (CI), which was proposed by Thomas Saati [1]. It is considered that the matrix is sufficiently well matched if its CI does not exceed 0.1.The authors of this paper have extensive experience in the development of decision support systems (DSS) [2-4] and have already implemented development to ensure the reliability of the original data to support decision making [5, 6]. The problem solved in the framework of this research, again, is connected with ensuring the reliability of the initial data and consists in developing methodological and algorithmic support for adjusting the matrix of paired comparisons in such a way as to bring it to the proper degree of consistency.

The basis for the approach to adjusting the matrix are two principles that are inherent in ideal matrices:

- if the matrix is matched, its rows are collinear vectors;
- expert judgments must be transitive.

The first principle allows to identify elements that are poorly coordinated with the others. The second is to formulate a rule according to which numerical values are determined on which the elements of the original matrix must be varied in order to improve its consistency.

When searching for elements and their variation, two approaches are possible:

- provide the required value of the CI with the minimum number of changes introduced into the matrix of paired comparisons;

- provide the required value of the CI with a limitation on the total of variation (for example, the change may not exceed 30% of the original value).

These approaches are consistent with the requirements of experts and decision makers, and also provide the possibility of analyzing the resulting matrix due to the small number of changes introduced.

Problem of low coherence is well known, and the issue of increasing the consistency of judgments is extremely relevant. However, it is worth noting, many modern DSS, widely represented in the software market, do not have the functionality to improve consistency. Some of them can only calculate the value of CI, and the correction of the matrix falls on the expert. The authors of this paper suggest an algorithmic way to correct the matrix, which significantly reduces the work of the expert, requiring him to only analyze the proposed changes, leaving the person the right either to accept changes or reject them.

Some researchers from Russia suggested their methods of increasing the consistency of the matrix of paired comparisons. Comparison of the developed method with the existing ones allows us to find out the main advantage - the resulting matrix of paired comparisons is very close to the original one, there is only a small difference in the set of elements that have been changed to improve consistency. In the methods proposed, for example, in papers [7, 8], all elements of the matrix undergo changes, which is an obvious drawback, and substantially complicates the analysis of the resulting matrix.

References

1. Saaty T.L. Decision Making With Dependence and Feedback: The Analytic Network Process. New York: Rws Publications, 2001.
2. Osipov V.P., Sivakova T.V., Sudakov V.A., Trahtengerts E.A., Zagreev B.V. Methodological Base of Support Decision-making in the Planning of Sscientific and Applied Research and Experiments on the International Space Station (ISS) // Electrical and Data Processing Facilities and Systems. 2013. vol.9. N 3. P. 80–88.
3. Zagreev B.V., Osipov V.P., Sudakov V.A. A Decision Support System (DSS) for Developing Programs of Scientific and Applied Research and Experiments on the Russian Segment of the ISS //

5th European Conference for Aeronautics and Space Sciences (EUCASS 2013). Munich, Germany, 1-4 July 2013.

4. Eskin V.I., Sudakov V.A. Automated Decision Support Using a Hybrid Preference Function // Herald of the Bauman Moscow State Technical University. Series Instrument Engineering. 2014. N 3. P. 116–124.
5. Sudakov V.A., Nesterov V.A., Kurennykh A.E. Integration of Decision Support Systems „Kosmos“and WS-DSS with Computer Models // Management of Large-Scale System Development (MLSD), 2017 Tenth International Conference. Moscow, Russia, 2-4 Oct. 2017.
6. Kurennykh A.E., Sudakov V.A. Decision Support Based on Simulation. Scientific and Technical Journal of Information Technologies, Mechanics and Optics. 2017. vol. 17. N 2. P. 348–353 (in Russian). doi: 10.17586/2226-1494-2017-17-2-348-353.
7. Yarygin A.N., Yarygin O.N. Realtive Ranking Intellectual Competences by Interactive Paired Comparisions // Vektor Nauki of Togliatti State University. 2011. vol.16 N 2. P. 413–417.
8. Lomakin V.V., Lifirenko M.V. Algorithm to Enhance the Coherence of Paired Comparisons Matrix When Carrying out Expert Surveys // Technical sciences. 2013. N 11. P. 1798–1803.

Injection of optimal solutions into population of a genetic algorithm for Pareto frontier approximation*

A.V. Lotov and A.I. Ryabikov

Dorodnicyn Computing Centre of Federal Research Center "Computer Science and Control" of Russian Academy of Sciences, Moscow, Russia

The paper is devoted to a new technique for approximating the Edgeworth-Pareto Hull (EPH) of the feasible set in criteria space in complicated multi-criteria decision problems. The technique applies integration of optimization techniques and genetic algorithms. In the framework of it, the decision points that provide extrema of partial optimization criteria are periodically injected into the population of the genetic algorithm used in the study. Experiments show that such an approach is efficient in the case of complicated multi-criteria decision problems.

*This research was supported by the Russian Foundation for Basic Research, project no. 17-29-05108 ofi m

Approximating the Edgeworth-Pareto Hull is the first, most complicated step of the multi-criteria decision support method named the Interactive Decision Maps (IDM) technique [1, 2]. The next two steps of IDM are visualization of EPH that informs the decision makers about the Pareto frontier and subsequent identification of the preferred criterion point of the Pareto frontier by the decision makers. IDM proved to be an effective tool for solving the multi-criteria problems with three to nine criteria, especially in economic and environmental fields [1]. For non-convex non-linear problems, several methods were developed that proved to be able to approximate EPH in the case of relatively small number of local minima in possible functions of criteria [3, 4].

Mathematically, the problem of multi-criteria optimization looks as follows [5]. The set of feasible decisions (feasible set) X belongs to the decision space R^n. The vectors of decision criteria y belong to the criteria space R^m. For a decision vector x, the related criterion vector y is given by the function $F : R^n \to R^m$. Then, the set of feasible criterion vectors Y is given by $Y = F(X)$. For certainty, we assume that the decision makers are interested in minimization of the value of any partial criterion $y_j, j = 1, ..., m$, while other criterion values are constant. Such preferences of the decision makers are formalized by using the Pareto binary relation (Pareto domination). A criterion point (vector) y' is more preferable than the criterion point y (in other words, y' dominates y in Pareto sense) means that $y' \leq y$ and $y' \neq y$. Mathematical solution of the multi-criteria decision problem is provided by two sets: the set of non-dominated criterion vectors (Pareto frontier) given by $P(Y) = \{y \in Y : \{y' \in Y : y' \leq y, y' \neq y\} = \emptyset\}$, as well as the set $P(X)$ of Pareto efficient decision vectors, i.e. the set of such decisions $x \in X$ for which it holds $F(x) \in P(Y)$.

The Edgeworth-Pareto Hull (EPH) of the set Y is defined in multi-criteria minimization problems as $H(Y) = Y + \mathbf{R}^m_+$, where $\mathbf{R}^m_+$ is the non-negative orthant $\mathbf{R}^m$. The set $H(Y)$ is the maximal set that has the same Pareto frontier as Y.

In this paper a new technique for approximating EPH is proposed to be applied in complicated multi-criteria decision problems. Namely, the problem of constructing the release rules for a cascade of hydropower plants is considered taking multiple criteria into account. The problem is described by a non-linear dynamic multistep model (more than 2000 steps) [6]. The release rules are described by more than 300 decision variables (parameters of the release rules) and more than two dozen of criteria. The relations of the model are partially given by tables. Consequences of a decision can be found by using simulation if a variant of the release rule is provided as the input of the model. The classes of release rules developed by specialists in water management are described by non-differentiable dependence of the decision variables with an extremely high Lipschitz constant. The criteria are given by piecewise constant functions of violation of external economic, environmental, communal and other requirements to the cascade. Actually, any criterion is the sum of several thousands of Heaviside functions of the violations. To overcome complications related to Heaviside functions, auxiliary continuous functions were used for approximation. However, it turned out that, even in the case of simple functions of several criteria, the problem of minimization of such functions is characterized by an extremely large number of local minima. This is the reason why application of the methods [3, 4] in this case is not effective.

Another approach to approximating the Pareto frontier is the use of genetic methods [7]. Our experiments show, however, that one of the most known genetic algorithms of NSGA [7] is also not sufficient in the case of the model of hydroelectric power plants cascade. For this reason, a new technique for EPH approximation was proposed. It is based on genetic algorithms augmented by periodic injection of solutions of partial criteria optimization problems (Partial Optimal Solutions denoted by POS) into the population of genetic algorithms. Thus, the proposed method is named Partial Optimal Solutions Injection into Genetic Algorithm (POSIGA) technique. In addition, the new sopping rule is used. Instead of the traditional stopping rule (number of iteration), maximal deviation of the new criterion points (obtained at an iteration) from the current EPH approximation is used. If the deviation is small enough (less than a given positive number ε_{gm}), the approximation process is completed. The result is obtained in the form of a list of non-dominated criterion points (approximation base) T_{gm} that is the basis of the EPH approximation obtained in the form $H(T_{gm})$. The set of efficient decisions Q_{gm}, for which $T_{gm} = F(Q_{gm})$, is provided as well.

Let us consider the description of the POSIGA method, in the framework of which the version [8] of the algorithm NSGA is used as the genetic algorithm. The desirable number of points of the output set N_{pop}, the parameter of the stopping rule ε_{gm} and the period K of POS injection must be given in advance.

Step 1. The initial set (initial population) is prepared.

First of all, the set POS is constructed, i.e. m global minimization problems for the partial criteria are solved. The problems must be solved fairly precisely by using the multistart method of global optimization (see, for example [9]). The set of initial population of a genetic method denoted by Q_0 includes points of POS as well as any convenient number of random points of the set X.

Step 2. Genetic algorithm NSGA complemented with the new stopping rule and the injection of POS computes the EPH approximation.

Iteration number k (substeps 2.1-2.4).

It is assumed that population of points Q_{k-1} from X has already been constructed at the previous iteration.

2.1 If $k = nK$ where $n = 1, 2, 3, ...$, then the points of POS are additionally included into Q_{k-1}.

2.2 A subset Q^*_{k-1} of the set Q_{k-1} is selected on the basis of some rules (see [8]). The set of new points (descendants) C_{k-1} is generated by application of cross-over and mutation operators [8] to Q^*_{k-1}. Let $Q^{**}_{k-1} = Q^*_{k-1} \cup C_{k-1}$.

2.3 If the number of points of the set Q^{**}_{k-1} is less than N_{pop}, then $Q_k = Q^{**}_{k-1}$. In the opposite case, the linear order between points of the set Q^{**}_{k-1} is established on the basis of Pareto domination between the related criterion vectors and the deviation of the related criterion vectors from the rest of criterion vectors (see [8]). Then, the first N_{pop} points of the set Q^{**}_{k-1} constitute the set Q_k.

2.4 Maximal deviation ε_k of the criterion points $F(Q_k)$ from $H(F(Q_{k-1}))$ is computed. If $\varepsilon_k > \varepsilon_{gm}$, then start the next iteration.

Compute $Q_{gm} = P(Q_k \cup Q_0)$ and $T_{gm} = F(Q_{gm})$.

Step 3. At this step, a refinement of the EPH approximation $H(T_{gm})$ is performed in the neighborhood of several criterion points, the most important from the practical point of view. Usually, the Plastering algorithm [5] is carried out. As the result of its application, the final approximation of EPH and the related set of efficient decisions are obtained.

The influence of the POS injection into the initial population was studied. It means that the original NSGA algorithm was compared with the simplified POSIGA method while the intermediate and final injections are not used. The initial population of the NSGA method consisted of N_{pop} random points of X, and the initial population of the POSIGA method consisted of points of POS complemented with random points of X to get the same number of points in the set Q_0. Both methods constructed EPH approximation while the number of computing the criterion function value was the same. We used $N_{pop} = 10000$ and $\varepsilon_{gm} = 0.01$. As the result, the POSIGA method stopped after 545-th iteration. Computing of POS required about 10 million calculations of the criterion function, and 545 iterations of step 2 required about 5.5 million calculations of it. This number of its calculations was sufficient for 1545 iterations of the NSGA algorithm.

Several methods were used for comparison of EPH approximations. First, the inclusion functions method [10] was used. Let T_{NS} be the approximation base obtained by the NSGA algorithm and T_{POS} – by the POSIGA algorithm. Let $\nu_{NS}(\varepsilon)$ be the share of points of T_{NS}, which belong to ε vicinity of $H(T_{POS})$ and $\nu_{POS}(\varepsilon)$ be the share of points of T_{POS}, which belong to ε-vicinity of $H(T_{NS})$. It has turned out that $\nu_{NS}(\varepsilon) \leq \nu_{POS}(\varepsilon)$. In particular, it has turned out that the set $H(T_{POS})$ contains 8% of points of T_{NS}, its ε-vicinity at $\varepsilon = 0.016$ contains already more than 95% of points of T_{NS} and the deviation of the most distanced point is 0.026. It means that points of T_{NS} are concentrated in the small neighborhood of the set $H(T_{POS})$. On the contrary, the set $H(T_{NS})$ does not contain points of T_{POS}. In addition, at $\varepsilon = 0.016$, the ε-vicinity of the set $H(T_{NS})$ contains less than 20% of points of T_{POS} and the deviation of the most distanced point is 0.074. It means that points of T_{POS} are fairly distant from the set $H(T_{NS})$. Thus, the EPH approximation given by T_{POS} is obviously much better.

Another method of approximation comparison is based on deviations of a known criterion point that for sure belongs to the set Y and is pretty close to its Pareto frontier. It has turned out that its deviation from the approximation constructed by the NSGA algorithm was 1.5 times greater than the deviation from the approximation constructed by the POSIGA method. This evidence supports the conclusion obtained with the help of inclusion functions method.

References

1. Lotov A.V., Bushenkov V.A., Kamenev G.K. Interactive Decision Maps. Approximation and Visualization of Pareto Frontier. Boston: Kluwer Academic Publishers, 2004.
2. Lotov A.V., and Miettinen K. Visualizing the Pareto Frontier // Multiobjective Optimization. Interactive and Evolutionary Approaches, Lecture Notes in Computer Science, V. 5252, Springer, Berlin-Heidelberg, 2008, p.213-244.
3. Berezkin V.E., Kamenev G.K., and Lotov A.V. Hybrid Adaptive Methods for Approximating a Nonconvex Multidimensional Pareto Frontier // Computational Mathematics and Mathematical Physics. 2006. V. 46, no. 11. P. 1918-1931.
4. Berezkin V.E., Lotov A.V., and Lotova E.A. Study of Hybrid

Methods for Approximating the Edgeworth-Pareto Hull in Nonlinear Multicriteria Optimization Problems // Computational Mathematics and Mathematical Physics. 2014. V. 54, no. 6. P. 919-930.

5. Miettinen K. Nonlinear multiobjective optimization. Boston: Kluwer, 1999.
6. Lotov A.V., and Riabikov A.I. Multiobjective feedback control and its application to the construction of control rules for a cascade of hydroelectric power stations // Trudy of Institute for Mathematics and Mechanics. 2014. V. 20, no. 4. P. 187-203 (in Russian).
7. Deb K. Multi-objective optimization using evolutionary algorithms. Chichester, UK: Wiley, 2001. 515 p.
8. Deb K., Pratap A., Agarwal S., and Meyarivan T. A Fast and Elitist Multiobjective Genetic Algorithm : NSGA-II // IEEE Transactions on Evolutionary Computation. 2002. V. 6, no. 2. P. 182-197.
9. Horst R., Pardalos P.M. Handbook on global optimization. Dordrecht, NL: Kluwer, 1995.
10. Berezkin V.E., and Lotov A.V. Comparison of Two Pareto Frontier Approximations // Computational Mathematics and Mathematical Physics. 2014. V. 54, no. 9. P. 1402-1410.

Aggregation of preferences in attitude scales

V.N. Nefyodov, S.O. Smerchinskaya, and N.P. Yashina
Moscow Aviation Institute, Moscow, Russia

The problem of multi-criteria choice with nonuniform scales of criteria is considered.

A set of alternatives $A = \{a_1, a_2, ..., a_n\}$ and a set of criteria $K = \{K_1, K_2, ..., K_m\}$ are given. Criteria are of equal importance, and for each criterion K_t $(t = 1, ..., m)$ numerical estimates of alternatives or information on how many times one alternative is preferable to another one can be given. It is required to construct an aggregated preference relation ρ which allows to compare alternatives on all criteria simultaneously.

Let the estimates of the alternatives $a_1, a_2, ..., a_n$ by the criterion K_t $(t = 1, ..., m)$ be given by a vector $x^t =< x_1^t, x_2^t, ..., x_n^t >$ with positive real components: x_i^t – an estimation of alternative a_i by criterion K_t.

Definition 1. A matrix of preferences $R^t = \|r_{ij}^t\|$ by criterion K_t is a square matrix of order n (n is the number of alternatives) with elements

$$r_{ij}^t = \begin{cases} \dfrac{x_i^t}{x_i^t + x_j^t}, & \text{if the values of the estimates on the scale } K_t \\ & \text{are maximized,} \\ \dfrac{x_j^t}{x_i^t + x_j^t}, & \text{if the values of the estimates on the scale } K_t \\ & \text{are minimized,} \end{cases}$$

$t = 1, ...m$.

Notice that for the elements of the preference matrix R^t constructed for the criterion K_t, the following holds:

$$1) \quad \frac{r_{ij}^t}{r_{ji}^t} = \begin{cases} \dfrac{x_i^t}{x_j^t}, & \text{if the values of the estimates on the scale } K_t \\ & \text{are maximized,} \\ \dfrac{x_j^t}{x_i^t}, & \text{if the values of the estimates on the scale } K_t \\ & \text{are minimized.} \end{cases}$$

Information is stored about how many times the alternative a_i is preferable to the alternative a_j.

2) $r_{ij}^t + r_{ji}^t = 1$ $(i, j = 1, ..., n)$, this actually replaces the procedure of bringing the scales of criteria to uniform.

If information is specified by the K_t criterion that the alternative a_i is preferable to the alternative a_j in α_{ij} times (the information should not contain contradictions), then the elements of the preference matrix R^t are calculated by the formulas:

$$r_{ji}^t = \frac{\alpha_{ij}}{1 + \alpha_{ij}}, \quad r_{ij}^t = \frac{1}{1 + \alpha_{ij}}$$

$(t \in \{1, ..., m\}, i, j = 1, ..., n)$.

Algorithm of aggregation of preferences in attitude scales

1. Formation of matrices of preferences $R^1, R^2, ..., R^m$ by the criteria $K_1, K_2, ..., K_m$.

2. The construction of the matrix of total preferences $P = \|p_{ij}\|$ is a square matrix of order n (number of alternatives) with elements

$$p_{ij} = \sum_{t=1}^{m} r_{ij}^t.$$

3. The construction of the preference matrix $R = \|r_{ij}\|$ of the aggregated relation ρ based on the matrix of total preferences P:

$$r_{ij} = \begin{cases} 1, & \text{if } p_{ij} > p_{ji} \\ \frac{1}{2}, & \text{if } p_{ij} = p_{ji} \\ 0, & \text{if } p_{ij} < p_{ji}, \end{cases}$$

where p_{ij} are the elements of the matrix P.

We note that the matrix of total preferences can be constructed taking into account the different importance of the criteria. In the presence of weighting coefficients of the criteria importance $k_1, k_2, ..., k_m$, the elements of the matrix $P = \|p_{ij}\|$ are calculated by the formula

$$p_{ij} = \sum_{t=1}^{m} k_t r_{ij}^t.$$

The aggregated relation constructed by this algorithm can be nontransitive, and the corresponding digraph $G =< A, \rho >$ contains contradictory cycles. The procedure for destroying contradictory circuits and constructing a transitive aggregate relation was proposed in [2].

Consider the case where alternatives are evaluated according to two quality criteria.

Proposition 1. For alternatives evaluated according to two quality criteria, estimates on which are positive and maximized, an alternative with a large value of the product of the components of the vector estimate is preferred.

Proposition 2. For alternatives evaluated according to two quality criteria, estimates on which are positive and minimized, an alternative with a smaller product of the components of the vector estimate is preferred.

Proposition 3. Let estimates of two alternatives, evaluated according to the quality criteria K_1 and K_2, be positive and criterion K_1 maximized, and K_2 minimized. Then an alternative with a large value of the ratio of the K_1 evaluation to the K_2 evaluation is preferable.

Theorem 1 follows from Propositions 1–3.

Theorem 1. An aggregated preference relation constructed for two quality criteria with positive scales is transitory.

The requirement of such a natural for the decision-maker, the conditions, as the transitivity of the aggregated preference relation, is one of the most important in decision theory. The fulfillment of this condition actually ensures consistency of the results obtained.

We will perform a comparative analysis of the proposed algorithm for aggregating preferences in attitude scales and the algorithm for constructing additive convolution for two criteria. Let the criteria on which alternatives are evaluated have positive scales and equal importance. We will consider vector estimates $< x_1, y_1 >$ and $< x_2, y_2 >$, as points of a plane. The coordinates of points that are equivalent to a fixed $< x_2, y_2 >$, belong to the positive branch of the hyperbola $y = \frac{c}{x}$, where $c = x_2 y_2$.

Consider the following example. Let us take the interval $(0; 10]$ as a scale and fix the point $< 5; 5 >$. Points equivalent to $< 5; 5 >$ belong to the hyperbola $y = \frac{25}{x}$. If the values on the criterion scales K_1 and K_2 are maximized, then vector estimates, less preferable estimates $< 5; 5 >$, lie under the branch of the hyperbola, including the Pareto relation $(Pareto^{-})$. Vector estimates, more preferable than $< 5; 5 >$ lie above the branch of the hyperbola. In the case of additive convolution, vector estimates equivalent to $< 5; 5 >$, belong to the straight line $y = 10 - x$. Vector estimates, preferably $< 5; 5 >$ lie above the line. In Fig. 1 compares the results obtained by aggregation and additive convolution. The area allocated between the line and the hyperbole illustrates the differences of the results of the aggregation method in attitude scales and additive convolution.

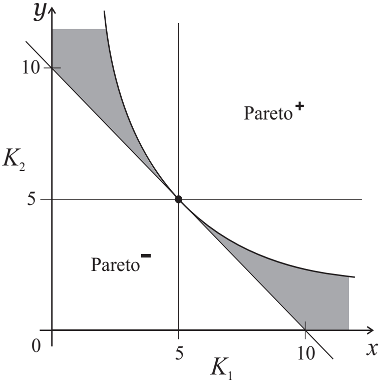


Fig. 1.

Let's make a comparative analysis of aggregation methods in attitude scales and additive convolution for m criteria.

Theorem 2. Let the alternatives be evaluated according to m quality criteria, the estimates for which are positive and maximized. Among all the alternatives with vector estimates $< y_1, y_2, ..., y_m >$ for which $y_1 + y_2 + \cdots + y_m = c$ $(c \in R^+)$ is fulfilled, the most preferable by the aggregation method is the alternative with the estimate $< \frac{c}{m}, ..., \frac{c}{m} >$.

It follows from Theorem 2 that from among all alternatives that are equivalent by the convolution method to the aggregation method, the most preferable is an alternative with equal components of the vector estimate. For example, alternatives with vector estimates $< 5, 5, 5, 5, 5 >$ and $< 10, 10, 2, 2, 1 >$ are equivalent by the method of additive convolution, since Have an equal sum of components. It is natural to assume that for a decision-maker, an alternative with average values for all criteria is preferable to an alternative with two distinct and three very bad estimates.

The obtained results indicate that the choice of the solution method depends not only on the type of initial information, but also on the preferences of the decision-maker. For this purpose, vector estimates equivalent to the decision-maker are in the conversational mode, and then they are approximated by curves. For the method of convolution, the least remote from the alternatives equivalent by two criteria should be a straight line, for the aggregation method in attitude scales – the hyperbola.

The aggregated preference relation can be found from the condition of minimization of the total distance from the preference matrix of this relation to the preferences matrices by criteria. In this case, the aggregated relation depends on the method of choosing the formula for the distance between the matrices.

Define the distance between the matrices R^k and R^t by formula

$$d(R^k,\, R^t) = \sum_{i=1}^{n} \sum_{j=1}^{n} |r_{ij}^k - r_{ij}^t|.$$

The following theorem holds.

Theorem 3. The total distance $D(Q) = \sum_{t=1}^{m} d(Q,\, R^t)$ is minimal for an odd number of experts if all the elements q_{ij} of the preference matrix Q are equal the median of the corresponding elements of the preferences matrices $r_{ij}^1, ..., r_{ij}^m$.

If the distance between matrices of preferences is given by formula

$$d(R^k,\, R^t) = \sum_{i=1}^{n} \sum_{j=1}^{n} (r_{ij}^k - r_{ij}^t)^2,$$

then Theorem 4 holds.

Theorem 4. The total distance $D(Q)$ for the introduced distance is minimal if all the elements q_{ij} of the preference matrix Q are equal the arithmetic mean of the corresponding elements of the preferences matrices $r_{ij}^1, ..., r_{ij}^m$.

Since $Q = \frac{1}{m}P$, the relation corresponding to the preference matrix equal to Q will coincide with the aggregated relation ρ constructed earlier.

In the case when weighting coefficients the importance of the criteria $k_1, k_2, ..., k_m$, are given, the elements q_{ij} of the matrix Q will be equal the median or the arithmetic mean of the elements, respectively $k_1 r_{ij}^1, k_2 r_{ij}^2$, ..., $k_m r_{ij}^m$.

References

1. Smerchinskaya S.O., Yashina N.P. On an algorithm for pairwise comparison of alternatives in multi-criteria problems // International Journal of Modeling, Simulation, and Scientific Computing. 2018. Vol. 9, N 1. P. 71–85. DOI: 10.1142/S179396231850006X.
2. Nefyodov V.N., Osipova V.A., Smerchinskaya S.O., Yashina N.P. Non-Contradictory Aggregation of Strict Order Relations // Russian Mathematics. 2018. Vol. 62, N 5. P. 61–73.

Adaptive evolution algorithm for multi-objective optimization of innovation projects[*]

S.V. Pronichkin

Federal Research Center "Computer Science and Control" of Russian Academy of Sciences, Moscow, Russia

The innovation project (IP) optimization problems are highly nonlinear with large problem dimension and a large number of equality and inequality constraints. Due to the complexity of the underlying problem, a penalty function based approach for constraint enforcement in evolutionary algorithms (EA) was deemed impractical. We propose different representation schemas that make monotonicity constraint satisfaction inherent to the optimization process. In addition, the problem dimension

[*] The reported study was funded by RFBR according to the research projects 16-06-00224 and 17-06-00237.

is reduced to by transforming the equality constraints into an inherent property of the representation methods. The optimization problem thus obtained has only boundary constraints so that an EA generates only feasible solutions at all times. This approach of automatic resolution of constraints through use of a suitable representation schema is key to efficient EA-based optimization. We also consider the utilization of domain knowledge to facilitate the optimization. In this paper, we propose an alternative approach to the discrete cohort approach, which poses the computing and smoothing of one-year IPs as a constrained optimization problem. The objective function to be minimized is an error function that calculates the discrepancy between the predicted transition matrices and the empirical data over the required time horizon. All the required structural properties of the one-year IP are captured in the form of constraints. The problem however is complex due to the following reasons: - The objective function is highly nonlinear due to the nonlinear nature of the error function and the matrix exponential operation involved in calculating the later-year transition probability matrices. - The problem dimension is very high as the number of variables in the IP is of the order of a few hundreds considering that there may be a large number of innovation ratings. - The optimized one-year IP is expected to satisfy structural and default properties. Due to the above difficulties, a traditional nonlinear programming method is not efficient to find the global optimal solution. This motivated a consideration of a set of population-based evolutionary algorithms, as they have shown success in many real world applications [1–4], and their requirement of computing resources is reasonably acceptable for offline applications. The industry standard for estimating discrete time project transition probability matrices is the innovation approach [5]. This approach applies to discrete innovation migration data and employs two key assumptions: (a) Future rating transitions are independent of past ratings (Markovian assumption). (b) The transition probabilities are solely a function of the distance between dates and are independent of the calendar dates (time-homogeneity assumption). In this approach, we first calculate the discrete cohort IP for one-year. Then, a later-year IP is obtained by raising the one-year IP to a power. For example, assume the one-year IP as M; then a t-year TPM is given by M^t. In practice, we would like these projected IPs to be as close as possible to the empirical IPs, M^t, obtained from empirical data. However, this may not be true for the discrete cohort one-year empirical IP. Due to the natural tendency of obligors to maintain their status quo, the diagonal value for any particular rating is expected to be greater than off-diagonal values (except the value that corresponds to the default probability) in the same way of the one-year IP. Also, it is expected that a particular obligor has a greater probability of migrating to a nearer innovation rating than a farther one. Therefore, the IP matrix elements should be monotonically decreasing on either side of the diagonal. However, this property may not be satisfied by the empirical IPs. Other required properties of the IPs are default constraints: - The default probability increases over time for each rating category. - In each year, a higher rating has a lower probability of default than a lower rating. It is clear that we prefer a one-year IP that satisfies the structural stability constraints and default constraints, and when pushed through time minimizes the deviation from the multiyear IPs. Assume that the last row and the last column of an n x n IP are associated with the state of default, while other rows and columns correspond to normal innovation ratings. We can represent the calculation of a smoothing one-year IP M in the form of a constrained error minimization problem as follows: $\min\limits_{M} \sum\limits_{t=1}^{T} w_t f(M^t, \bar{M}_t, w_{ij})$, (1) subject to $0 \leqslant m_{t,ij} \leqslant 1$, $\sum\limits_{j=1}^{N} m_{t,ij} = 1$ for $t = 1$, (2) $m_{t,nn} = 1$, and $m_{t,ni} = 0$ if $i < n$ for $t = 1$, (3) and $m_{t,ij} \leqslant m_{t,ik}$, if $j < k \leqslant i$ or $i < k < j < n$, (4) $m_{t,in} \leqslant m_{t,jn}$, if $i < j$, (5) for $\forall t \in \{1, 2, \ldots, T\}$ and $\forall i, j \in \{1, 2, \ldots, n\}$, where T is the number of years of interest. In (1), f is an error function that measures the dependency between M^t and $\bar{M}_t$, and w_t is the weight for the dependency at the t-th year. The i-th row and j-th column elements, $m_{t,ij}$ and $\bar{m}_{t,ij}$, of M^t and $\bar{M}_t$ represent the predicted and empirical transition probabilities from the i-th to the j-th innovation rating over a t-year period, respectively. w_{ij} is the weight for the dependency between $m_{t,ij}$ and $\bar{m}_{t,ij}$. By varying the weights, the optimization can be customized to emphasize defaults, transitions, or specific rating categories, according to business needs. The equality constraints in (2) ensure that each row of a transition probability matrix sum up to 1. Eq. (3) implies that the rating of default is an absorbing state, i.e., once an obligor reaches the default state, it is assumed to remain there indefinitely (in practice, an obligor that emerges from default is treated as a new obligor). Thus, the problem in (1) is actually to optimize over $n\,(n-1)$ variables. Note that the constraints (2) and (3) are automatically satisfied for $t > 1$, as long as they hold for $t = 1$. Eq. (4) formulates the structural constraints such that the elements of a IP are monotonically decreasing on either side of the diagonal. This acts as a mechanism to smooth the probability surface. Eq. (5) states that a higher rating has a lower probability of default than a lower rating. Different from constraints (2) and (3), constraints (4) and (5) do not necessarily hold for multi-year IP (i.e., $t > 1$) even when they are satisfied by the one-year IP. However, our experiments on a wide variety of test data show that these constraints are implicitly satisfied by the optimized IP for later years once they hold for $t = 1$. For simplicity, we explicitly impose these constraints for $t = 1$ in this paper, while assuming that the error minimization procedure implicitly enforces them for later-year IPs. Evolutionary algorithms are stochastic, population-based search methods that mimic the process of natural biological evolution. They generally operate on a population of potential solutions applying the principle of survival of the fittest to produce better and better approximations to a solution. As shown by successes in various fields such as engineering, finance, biology [5], evolutionary algorithms consistently perform well in searching optimal solutions to various types of problems. Differential evolution follows the basic procedure of an evolutionary algorithm. The initial population is randomly generated according to a uniform distribution between the lower and upper bounds defined for each component of an individual vector. After the initialization, algorithms enters a loop (called a generation in the EA literature) of evolutionary operations: mutation, crossover and selection. In addition, in an adaptive algorithm, control parameters are adapted at the end of each generation. Mutation: At each generation g, this operation creates mutant vectors $\mathrm{v_{i,g}}$, based on the current parent population $\{\mathrm{x_{i,g}}\ | i = 1, 2, ..., NP\}$, where NP is the population size. The mutation vectors are generated as follows: $\mathrm{v_{i,g}} = \mathrm{x_{r3,g}} + F\left(\mathrm{x_{r1,g}} - \mathrm{x_{r2,g}}\right)$, (6) where the indices $r1$, $r2$ and $r3$ are distinct integers uniformly chosen from the set $\{1,\ 2,\ \ldots,\mathrm{NP}\} \setminus \{\mathrm{i}\}$, $\mathrm{x_{r1,g}} - \mathrm{x_{r2,g}}$ is a difference vector to mutate the parent, and $F \in (0, 1]$ is the mutation factor that is fixed throughout the optimization process. Different from (6), we adopt a relatively greedy mutation strategy: $\mathrm{v_{i,g}} = \mathrm{x_{i,g}} + F_{\mathrm{i}}\left(\mathrm{x}^{p}_{best,g} - \mathrm{x_{i,g}}\right) + F\left(\mathrm{x_{r1,g}} - \mathrm{x_{r2,g}}\right)$, (7) where $\mathrm{x}^{p}_{best,g}$ is randomly chosen as one of the top 100p% individuals in the current population, and $F_{\mathrm{i}} \in (0, 1]$ is the mutation factor associated with each individual $\mathrm{x_{i,g}}$ and is randomly generated by the parameter self-adaptation . Crossover: After mutation, a binary crossover operation forms the final trial vector $\mathrm{u_{i,g}} = (\mathrm{u_{1,i,g}}, \mathrm{u_{2,i,g}}, \ldots, \mathrm{u_{D,i,g}})$: $\mathrm{u_{1,i,g}} = \begin{cases} \mathrm{v_{j,i,g}} \text{ if } rand_{\mathrm{j}}(0,1) \leqslant \mathrm{CR_i} \ or\ j = \mathrm{j_{rand}}, \\ \mathrm{x_{j,i,g}}\ otherwise, \end{cases}$ (8) where $rand_{\mathrm{j}}(0,1)$ is a uniform random number on the interval (a,b] and newly generated for each $j, \mathrm{j_{rand}} = \ randint_{\mathrm{i}}(1, D)$ is an integer randomly chosen from

1 to D and newly generated for each i, and the crossover probability, $\mathrm{CR_i} \in (0, 1]$, roughly corresponds to the average fraction of vector components that are inherited from the mutant vector. The crossover probabilities are newly generated by the parameter self-adaptation at each generation. Selection: The selection operation selects the better one from the parent vector $\mathrm{x_{i,g}}$ and the trial vector $\mathrm{u_{i,g}}$ according to their fitness values f. For example, since we consider a minimization problem, the selected vector is given by

$$\mathrm{x_{i,g+1}} = \begin{cases} \mathrm{u_{i,g}} \text{ if } f\left(\mathrm{u_{i,g}}\right) < f\left(\mathrm{x_{i,g}}\right), \\ \mathrm{x_{i,g}} \; otherwise, \end{cases} \quad (9)$$

and used as a parent vector in the next generation. If the trial vector $\mathrm{u_{i,g}}$ succeeds, the selection is considered as a successful update and the corresponding control parameters F_{i} and $\mathrm{CR_i}$ are called a successful mutation factor and successful crossover probability, respectively. The two involved control parameters, F and CR, are usually problem dependent and need to be tuned by trial and error by a self-adaptation mechanism that is based on a simple principle. Better values of control parameters tend to generate individuals that are more likely to survive and thus these values should be propagated. To be specific, F_{i} and $\mathrm{CR_i}$ are generated by two random processes: $\mathrm{CR_i} = randn_{\mathrm{i}}\left(\mu_{CR}, 0.1\right)$, (10) $\mathrm{F_i} = randc_{\mathrm{i}}\left(\mu_F, 0.1\right)$, (11) The mean μ_{CR} and location parameter μ_F are updated in a self-adaptive manner. The proposing algorithm (PA) works best in terms of both convergence rate and robustness for set of IP optimization problems. PA generally obtains near-optimal values in 500 generations, compared to the values achieved after 20000 generations. As a comparison, existing algorithm usually approaches the optimal value after 2000 generations, also has difficulty to solve the IP problems due to premature convergence, although its convergence rate is fastest during the early generations. In view of the high dimension of innovation project optimization problems, a promising improvement over the current method could be based on the co-evolutionary strategies. Indeed, although the optimization problem is non-separable, the correlation among different probability terms of the IP matrix is not uniform. The probabilities associated with the same innovation rating are strongly correlated, while the interactions among probability terms of different ratings are relatively weak. Thus, it might be beneficial to optimize the probability terms related to each single rating by a separable sub-population and control the interaction among all probability terms by a standard co-evolutionary strategy. In this paper, we have considered the problem of computationally smoothing a one-year transition probability matrix by minimizing the discrepancy between predicted later-year IPs and empirical data over the time horizon of interest. The minimization problem is very complex not only due to its non-convex non-separable properties but also due to the large number of variables and constraints (desired properties) involved. A novel self-adaptive algorithm is adopted to calculate the optimal solution. Simulation results show that the proposed methodology, perform significantly better than other methods in terms of both the convergence speed of the optimization and quality of the final solution obtained.

References

1. Fanti M., Iacobellis G., Ukovich W. A simulation based Decision Support System for logistics management // Journal of Computational Science. 2015. V. 10, P. 86–96.
2. Haupt R., Haupt S. Practical genetic algorithms. Hoboken: John Wiley & Sons, 2004.
3. Hussain M., Al-Sultan K. A Hybrid Genetic Algorithm for Nonconvex Function Minimization // Journal of Global Optimization. 1997. V. 11. P. 313–324.
4. Haslinger J., Jedelsky D., Kozubek T., Tvrdik J. Genetic and Random Search Methods in Optimal Shape Design Problems // Journal of Global Optimization. 2000. V. 16, P. 109–131.
5. Gottschlich J., Hinz O. A decision support system for stock investment recommendations using collective wisdom // Decision Support Systems. 2014. V. 59, P. 52–62.

A fuzzy case of the Pareto set reduction in bi-criteria discrete problems*

A.O. Zakharov and Yu.V. Kovalenko
Saint Petersburg State University, St. Petersburg, Russia,
Sobolev Institute of Mathematics, Novosibirsk, Russia,
Novosibirsk State University, Novosibirsk, Russia

A discrete multicriteria problem includes the following components: a finite set of feasible solutions X and vector criterion (vector-valued function) $f = (f_1, f_2, \dots, f_m)$ defined on set X. Optimal solution to the problem is usually supposed to be the Pareto set [1], which is rather wide in real-world problems and rises difficulties in choosing a final solution.

*This research is supported by RFBR grant 17-07-00371 (A. Zakharov) and by the Ministry of Science and Education of the Russian Federation under the 5-100 Excellence Programme (Yu. Kovalenko).

For that reason numerous state-of-the-art methods are developed [2]: multiattribute utility theory, outranking approaches, verbal decision analysis, various iterative procedures with man-machine interface, etc.

In this paper we investigate discrete bi-criteria problems and apply to them the axiomatic approach of the Pareto set reduction proposed by V. Noghin [3]. The set of pareto-optimal solutions is defined as set $P_f(X) = \{x \in X \mid \nexists x^* \in X : f(x^*) \ \geq \ f(x) \ \}$, and the Pareto set $P(Y) = f(P_f(X))$.

V. Noghin considered such preference relation of the decision maker (DM), that reflects not only preferences but also a confidence degree of wishes (fuzzy case). The full description of apparatus of fuzzy set and fuzzy binary relation could be found in [4]. Further, we formulate the basic concepts and results of the axiomatic approach exactly for discrete bi-criteria problems. According to [3] we have the following fuzzy bi-criteria choice problem $< X, f, \mu >$:

- a finite set of feasible solutions X;
- a vector criterion $f = (f_1, f_2)$ defined on set X;
- an asymmetric fuzzy preference relation of the DM μ defined on set Y, $Y = f(X)$.

A fuzzy preference relation μ is defined by its membership function $\mu\colon Y \times Y \to [0, 1]$ as follows. If for vectors $y', y'' \in Y$ the equality $\mu(y', y'') = \mu^*$ holds, then the DM prefers the solution y' to the solution y'' with degree of confidence μ^* showing assurance in the choice.

The fuzzy relation μ satisfies the axioms of “fuzzy reasonable” choice [3], and it is irreflexive, transitive, invariant with respect to a linear positive transformation and compatible with each criteria f_1, f_2. The compatibility means that the DM is interested in increasing value of each criterion when value of other criterion is constant with degree of confidence one. In [3], the author established the Edgeworth–Pareto principle, according which under axioms of “fuzzy reasonable” choice any fuzzy set of selected outcomes $C(Y)$ belongs to the Pareto set $P(Y)$ (crisp set): $\lambda^C(y) \leqslant \lambda^P(y)$ for all $y \in Y$. Here the fuzzy set of selected outcomes is defined by its membership function $\lambda^C(\cdot)$ and interpreted as some abstract set corresponded to the set of outcomes, that satisfy all hypothetic fuzzy preferences of the DM. Membership function $\lambda^P(\cdot)$ assigns the Pareto set. So, the optimal “fuzzy” choice should be done only within the Pareto set if preference relation μ fulfills the axioms of “fuzzy reasonable” choice.

In order to narrow "fuzzy" upper bound on the fuzzy set of selected outcomes (in the Edgeworth–Pareto principle) V. Noghin introduced a specific information on the DM's fuzzy preference relation μ. The following definition we will give in terms of two criteria. Later on, we suppose that $i, j \in \{1, 2\}$, $i \neq j$.

Definition 1. Say that there exists a "fuzzy quantum of information" with degree of confidence $\mu^* \in [0, 1]$ if vector $y' \in R^2$ with components $y'_i = w_i > 0$, $y'_j = -w_j < 0$ satisfies the expression $\mu(y', 0_2) = \mu^*$. Value μ^* shows how the DM is sure that the i-th criterion is more important then the j-th one.

The quantity of relative looseness is set by the so-called coefficient of relative importance $\theta = w_j/(w_i + w_j)$, therefore $\theta \in (0,\ 1)$. It is easy to check that Definition 1 is equivalent to the existence of such vector $y'' \in R^2$ with components $y''_i = 1 - \theta$, $y''_j = -\theta$, that the relation $\mu(y'', 0_2) = \mu^*$ holds.

According to [3] the use of "fuzzy quantum of information" consists in solving two crisp multicriteria problems, that means finding corresponding Pareto sets. This process yields the fuzzy set, which is an upper bound on fuzzy set of the selected outcomes. By $\lambda^M(\cdot)$ we further denote the membership function of this set.

Let the i-th criterion is more important then the j-th one with parameters θ and μ^*. Firstly, we solve the problem $< X, f >$ without any additional information, i.e. find the Pareto set $P(Y)$. Then we put $\lambda^M(y) = 1$ for all vectors $y \in P(Y)$, and $\lambda^M(y) = 0$ for all vectors $y \in Y \setminus P(Y)$.

Secondly, we consider the problem $< X, \hat{f} >$, where vector criterion $\hat{f}$ has the components $\hat{f}_i = f_i$, $\hat{f}_j = \theta f_i + (1-\theta) f_j$. Then we set $\lambda^M(y) = 1 - \mu^*$ for all vectors $y \in P(Y) \setminus \hat{P}(Y)$, where $\hat{P}(Y) = f(P_{\hat{f}}(X))$. At the same time vectors y from set $\hat{P}(Y)$ still have degree of confidence, which is equal to 1. Thus, we get the membership function $\lambda^M(\cdot)$ defining fuzzy set M that the inequality $\lambda^C(y) \leqslant \lambda^M(y) \leqslant \lambda^P(y)$ holds for all $y \in Y$. Fuzzy set M forms a narrower "fuzzy" upper bound on the fuzzy set of selected outcomes, than the Pareto set $P(Y)$.

We investigate the degree of the Pareto set reduction with respect to values of coefficient of relative importance θ and degree of confidence μ^*. In any multicriteria discrete problem there exists such non-decreasing sequence of coefficients of relative importance $0 < \hat{\theta}_1 \leqslant \hat{\theta}_2 \leqslant \ldots \leqslant \hat{\theta}_k < 1$ that on each interval $(0,\ \hat{\theta}_1)$, …, $[\hat{\theta}_i,\ \hat{\theta}_{i+1})$, …, $[\hat{\theta}_k,\ 1)$ the set of vectors y having $\lambda^M(y) = 1 - \mu^*$ (in other words, $y \in P(Y) \setminus \hat{P}(Y)$) will be the same (invariant). We note that any θ from an interval $[\hat{\theta}_i, \hat{\theta}_{i+1})$ (and also $(0,\ \hat{\theta}_1)$, $[\hat{\theta}_k,\ 1))$, $i = 1, \ldots, k-1$, gives the same reduction of the Pareto set. In the paper we identify the following classes of the bi-criteria discrete problem upon its Pareto set structure: *"cascade"* and *"stairs"*.

We say that the Pareto set has "cascade" structure if its elements lay on p parallel lines so that $P(Y) = \bigcup_{i=1}^{p}\{(y_1^{(i)}, y_2^{(i)}) : y_2^{(i)} = a^{(i)} - ky_1^{(i)}, y_1^{(i)} \in Y_1^{(i)}\}$. Here, $Y_1^{(i)} = \{\sum_{j=0}^{i-1} l_j, \sum_{j=0}^{i-1} l_j + 1, \ldots, \sum_{j=0}^{i} l_j - 1\}$, l_0 is the value of the 1-st coordinate of the point having the maximum value on the 2-nd criterion (it lays on upper line), l_i is the number of points on the i-th line, $i \in \{1, 2, \ldots, p\}$. Besides, $a^{(1)} > a^{(2)} > \ldots > a^{(p)}$ and $k > 1$. Let $\hat{n} = \sum_{i=2}^{p} l_i$.

Theorem 1. *Let the Pareto set $P(Y)$ has "cascade" structure with p lines.*

1) Suppose the 1-st criterion f_1 is more important than the 2-nd one f_2 with coefficient of relative importance θ and the degree of confidence μ^*. Then if $\theta \in (0,\ k/(k+1))$ the reduced Pareto set $\hat{P}(Y)$ coincides with the Pareto set $P(Y)$, in the case of $\theta \in [k/(k+1),\ 1)$ the set $\hat{P}(Y)$ includes at most p elements.

2) Suppose the 2-nd criterion f_2 is more important than the 1-st one f_1 with coefficient of relative importance θ and the degree of confidence μ^*. Then if $\theta \in (0,\ 1/(k+1))$ the reduced Pareto set $\hat{P}(Y)$ includes at most $\hat{n}$ elements less then the Pareto set $P(Y)$. In the case of $\theta \in [1/(k+1),\ 1)$ the reduced Pareto set $\hat{P}(Y)$ consists of one element.

We say that the Pareto set has "stairs" structure if its elements lay on p parallel lines so that $P(Y) = \bigcup_{i=1}^{p}\{(y_1^{(i)}, y_2^{(i)}) : y_2^{(i)} = a^{(i)} - ky_1^{(i)}, y_1^{(i)} \in \tilde{Y}_1^{(i)}\}$. Here, $\tilde{Y}_1^{(i)} = \{l_0 + i - 1, l_0 + p + i - 1, \ldots, l_0 + (n-1)p + i - 1\}$, l_0 is the value of the 1-st coordinate of the point having the maximum value on the 2-nd criterion, n is the number of points on each line, $i \in \{1, 2, \ldots, p\}$. Besides, $a^{(1)} < a^{(2)} < \ldots < a^{(p)}$ and $k > 1$.

Theorem 2. *Let the Pareto set $P(Y)$ has "stairs" structure with p lines.*

1) Suppose the 1-st criterion f_1 is more important than the 2-nd one f_2 with coefficient of relative importance θ and the degree of confidence μ^*. Then if $\theta \in (0,\ k/(k+1))$ the reduced Pareto set $\hat{P}(Y)$ includes at least n elements, in the case of $\theta \in [k/(k+1),\ 1)$ the set $\hat{P}(Y)$ consists of one element.

2) Suppose the 2-nd criterion f_2 is more important than the 1-st one f_1 with coefficient of relative importance θ and the degree of confidence μ^*. Then if $\theta \in (0,\ 1/(k+1))$ the reduced Pareto set $\hat{P}(Y)$ includes at *least $p+n-1$ elements. In the case of $\theta \in [1/(k+1),\ 1)$ the set $\hat{P}(Y)$ contains at most p elements.*

The membership function $\lambda^M(\cdot)$ is calculated in theorems as mentioned before, and value of confidence degree does not influence on number of elements in set $\hat{P}(Y)$. We also identify boundary values of the coefficient of relative importance in intervals: $[k/(k+1),\ 1)$ and $(0,\ 1/(k+1))$ for "cascade" case; $(0,\ k/(k+1))$, $(0,\ 1/(k+1))$, and $[1/(k+1),\ 1)$ for "stairs" case. This values divide corresponding intervals on subintervals such that the number of vectors y having $\lambda^M(y) = 1 - \mu^*$ is identical for all θ in a subinterval.

We construct instances of the well-known bi-criteria Knapsack problem [5] with "cascade" and "stairs" structures of the Pareto set.

References

1. Podinovskiy V.V., Noghin V.D. Pareto-optimal'nye resheniya mnogokriterial'nyh zadach (Pareto-optimal solutions of multicriteria problems). Moscow: Fizmatlit, 2007 (In Russian).
2. Figueira J.L., Greco S., Ehrgott M. Multiple criteria decision analysis: state of the art surveys. New York: Springer-Verlag, 2005.
3. Noghin V.D. Reduction of the Pareto Set: An Axiomatic Approach. Springer International Publishing, 2018.
4. Zadeh L.A. Fuzzy sets // Information and control. 1965. V. 8, N 3. P. 338–353.
5. Ehrgott M. Multicriteria Optimization. Springer-Verlag, 2005.

Improper optimization problems

Tikhonov's solution of approximate and improper LP problems

V.I. Erokhin[1], A.V. Razumov[1], and A.S. Krasnikov[2],
[1]*Mozhaisky Military Space Academy, St. Petersburg, Russia,*
[2]*Financial University under the Government of the Russian Federation, Moscow, Russia*

This work relies on the results in [1], and is an extension of the development in [2]. The idea of the work is to extend Tikhonov's approximate system of linear algebraic equations solution approach to to find a stable solution of approximate linear programming problem.

Let

$$L(A, b, c) : Ax = b,\ x \geqslant 0,\ c^\top x \to \max,$$

$$L^*(A, b, c) : u^\top A \geqslant c^\top,\ b^\top u \to \min$$

– be a pair of mutually dual problems of the linear programming (LP), $c, x \in \mathbb{R}^n$, $b, u \in \mathbb{R}^m$, $A \in \mathcal{M}^{m\times n}$, $\mathcal{M}^{m\times n}$ – is a set of real matrices of size $m \times n$, $\mathcal{X}(A, b) \triangleq \{x|\, Ax = b, x \geqslant 0\}$, $\mathcal{U}(A, c) \triangleq \{u|\, u^\top A \geqslant c^\top\}$ – admissible solution set of the specified problems, furthermore the existance of the solution for the $L(A, b, c)$ and $L^*(A, b, c)$ problems is not specified.

Let there be a matrix $A_0 \in \mathcal{M}^{m\times n}$ and vectors $b_0 \in \mathbb{R}^m$, $c_0 \in \mathbb{R}^n$ such that problems $L(A_0, b_0, c_0)$ and $L^*(A_0, b_0, c_0)$ are proper problems. Let's define the matrix A_0 and vectors b_0, c_0 as *precise*, the matrix A and vectors b, c as *approximate*, and the corresponding LP problems as problems with the precise and approximate data. Let's assume that the following conditions are satisfied $\|A - A_0\| \leqslant \mu$, $\|b - b_0\| \leqslant \delta_b$, $\|c - c_0\| \leqslant \delta_c$, where $\mu, \delta_b, \delta_c > 0$, – some constants known a priori and the symbol $\|\cdot\|$ denotes Euclidean matrix norm (the symbol will also be used to denote an Euclidean vector norm through the paper).

Similar to the flow of the works [3, 4], let's consider the following problem.

Problem $Z_{A,b,c}$: *Find* $A_1 \in \mathcal{M}^{m\times n}$, $b_1 \in \mathbb{R}^m$ $c_1, \in \mathbb{R}^n$, $x_1 \in \mathbb{R}^n$, $u_1 \in \mathbb{R}^m$ *such as* $\|A - A_1\| \leqslant \mu$, $\|b - b_1\| \leqslant \delta_b$, $\|c - c_1\| \leqslant \delta_c$, $x_1 \in \mathcal{X}(A_1, b_1)$, $u_1 \in \mathcal{U}(A_1, c_1)$, $c_1^\top x_1 = b_1^\top u_1$, $\|x_1\|^2 + \|u_1\|^2 \to \min$.

Note that earlier, also similar to the flow of the works [3, 4], the following problem with the conditions $\delta_b = \delta_c = 0$ [2] was considered.

Problem Z_A : *Find* $A_1 \in \mathcal{M}^{m\times n}$, $x_1 \in \mathbb{R}^n$, $u_1 \in \mathbb{R}^m$ *such as* $\|A - A_1\| \leqslant \mu$, $x_1 \in \mathcal{X}(A_1, b)$, $u_1 \in \mathcal{U}(A_1, c)$, $c^\top x_1 = b^\top u_1$, $\|x_1\|^2 + \|u_1\|^2 \to \min$.

Formulation of the problem $Z_{A,b,c}$ allows the next interpretation: the pair of problems $L(A, b, c)$ and $L^*(A, b, c)$ (may be improper) with the approximate matrix A and vectors b, c is mapped to the corresponding solvable problems $L(A_1, b_1, c_1)$ and $L^*(A_1, b_1, c_1)$. Vectors x_1 and u_1 – are the solutions of the specified problems and the sum of squares of their Euclidean norms is minimal.

It can be shown that if $\mu, \delta_b, \delta_c \to 0$ the following holds $A_1 \to A_0$, $b_1 \to b_0$, $c_1 \to c_0$, $x_1 \to x_0$, $u_1 \to u_0$, where x_0 and u_0 – are the solutions of the problems $L(A_0, b_0, c_0)$ and $L^*(A_0, b_0, c_0)$ respectively, moreover the sum of the squares of the Euclidean norms of the vectors x_0 and u_0 is minimal. That is the vectors x_1 and u_1 are the stable approximation of the normal solution of the pair of mutually dual LP problems with the precise data.

The following theorems and auxiliary problems are important tool in finding the solution of the problem $Z_{A,b,c}$:

Theorem 1 [5]. *The system* $Ax = b, u^\top A = v$ *is solvable for the matrix* $A \in \mathcal{M}^{m\times n}$ *with the vectors* $x, v \in \mathbb{R}^n$, $u, b \in \mathbb{R}^m$, $x \neq 0$, $u \neq 0$, *if and only if the following holds* $v^\top x = u^\top b = \alpha$. *The unique solution with the minimal Euclidean norm* $\hat{A}$ *is defined with the following formula*

$$\hat{A} = \frac{bx^\top}{x^\top x} + \frac{uv^\top}{u^\top u} - \alpha\frac{ux^\top}{x^\top x \cdot u^\top u},$$

$$\|\hat{A}\|^2 = \frac{\|b\|^2}{\|x\|^2} + \frac{\|v\|^2}{\|u\|^2} - \frac{\alpha^2}{\|x\|^2 \cdot \|u\|^2}.$$

The family of matrices that are solutions of the system $Ax = b$, $u^\top A = v$ *has the form* $A = \hat{A} + \Delta A$, *where* $\Delta A \in \mathbb{R}^{m\times n}$ *is an arbitrary matrix* *such as* $\Delta Ax = 0$, $u^\top \Delta A = 0$ *and condition* $\|A\|^2 = \|\hat{A}\|^2 + \|\Delta A\|^2$ *is then satisfied.*

Problem $C:$ *Given* $A \in \mathcal{M}^{m\times n}$, $b \in \mathbb{R}^m$, $c \in \mathbb{R}^n$, $x \in \mathbb{R}^n$, $u \in \mathbb{R}^m$. *Find* $H \in \mathcal{M}^{m\times n}$, $h_b \in \mathbb{R}^m$, $h_c \in \mathbb{R}^n$ *such as* $x \in \mathcal{X}(A+H, b+h_b)$, $u \in \mathcal{U}(A+H, c+h_c)$, $(c+h_c)^\top x = (b+h_b)^\top u$, $\|H\|^2 + \|h_b\|^2 + \|h_c\|^2 \to$ min.

Theorem 2 [1]. *The solution* $\hat{H}$, $\hat{h}_b$, $\hat{h}_c$ *of the problem* C *exists and is unique for any* A, b, c, x, u,

$$\begin{bmatrix} \hat{H} & -\hat{h}_b \\ -\hat{h}_c^\top & 0 \end{bmatrix} = \frac{\begin{bmatrix} b - Ax \\ \upsilon \end{bmatrix} \begin{bmatrix} x^\top & 1 \end{bmatrix}}{x^\top x + 1} + \frac{\begin{bmatrix} u \\ 1 \end{bmatrix} \begin{bmatrix} g^\top & \omega \end{bmatrix}}{u^\top u + 1} - \\ - \alpha \frac{\begin{bmatrix} u \\ 1 \end{bmatrix} \begin{bmatrix} x^\top & 1 \end{bmatrix}}{(x^\top x + 1)(u^\top u + 1)},$$

$$\left\| \begin{bmatrix} \hat{H} & -\hat{h}_b \\ -\hat{h}_c^\top & 0 \end{bmatrix} \right\|^2 = \frac{\|b - Ax\|^2 + \upsilon^2}{\|x\|^2 + 1} + \frac{\|g\|^2 + \omega^2}{\|u\|^2 + 1} - \\ - \frac{\alpha^2}{\left(\|x\|^2 + 1\right) \cdot \left(\|u\|^2 + 1\right)},$$

$$g = (g_j) \in \mathbb{R}^n,\ g_j = \begin{cases} 0, \text{ if } \left(c - A^\top u\right)_j \leqslant 0 \text{ and } x_j = 0, \\ \qquad \left(c - A^\top u\right)_j \text{ otherwise,} \end{cases} \tag{1}$$

υ, ω, α *are the solution of SLAE*

$$\begin{bmatrix} 1 & 0 & -1 \\ 0 & 1 & -1 \\ (x^\top x+1)^{-1} & (u^\top u+1)^{-1} & -(x^\top x+1)^{-1} \cdot (u^\top u+1)^{-1} \end{bmatrix} \cdot \begin{bmatrix} \upsilon \\ \omega \\ \alpha \end{bmatrix} = \begin{bmatrix} u^\top Ax - u^\top b \\ u^\top Ax - c^\top x \\ 0 \end{bmatrix}.$$

Remark. By virtue of duality theory for LP problems, vector x is the solution of the $L(A + \hat{H}, b + \hat{h}_b, c + \hat{h}_c)$ problem, vector u is the solution of the $L^*(A + \hat{H}, b + \hat{h}_b, c + \hat{h}_c)$ problem.

Theorem 3. *Let* $A \in \mathcal{M}^{m\times n}$, $b \in \mathbb{R}^m$, $c \in \mathbb{R}^n$, $x \in \mathbb{R}^n$, $u \in \mathbb{R}^m$, $w_b, w_c \in \mathbb{R}$, $x \geqslant 0$, $w_b, w_c > 0$, *vector* g *is constructed by the formula* (1),

$$\begin{bmatrix} \tilde{H} & -\tilde{h}_b \\ -\tilde{h}_c^\top & 0 \end{bmatrix} \triangleq \frac{\begin{bmatrix} b - Ax \\ v \end{bmatrix} \begin{bmatrix} x^\top & w_b^2 \end{bmatrix}}{x^\top x + w_b^2} + \frac{\begin{bmatrix} u \\ w_c^2 \end{bmatrix} \begin{bmatrix} g^\top & \omega \end{bmatrix}}{u^\top u + w_c^2} - \\ - \alpha \frac{\begin{bmatrix} u \\ w_c^2 \end{bmatrix} \begin{bmatrix} x^\top & w_b^2 \end{bmatrix}}{(x^\top x + w_b^2)\,(u^\top u + w_c^2)}, \quad (2)$$

where v, ω, α are the solution of SLAE

$$\begin{bmatrix} w_c & 0 & -1 \\ 0 & w_b & -1 \\ \frac{w_b}{x^\top x + w_b^2} & \frac{w_c}{u^\top u + w_c^2} & \frac{-w_b \cdot w_c}{(x^\top x + w_b^2) \cdot (u^\top u + w_c^2)} \end{bmatrix} \cdot \begin{bmatrix} v \\ \omega \\ \alpha \end{bmatrix} = \begin{bmatrix} u^\top Ax - u^\top b \\ u^\top Ax - c^\top x \\ 0 \end{bmatrix}. \quad (3)$$

Then vector x is the solution of $L(A + \tilde{H}, b + \tilde{h}_b, c + \tilde{h}_c)$ problem, u is the solution of $L^*(A + \tilde{H}, b + \tilde{h}_b, c + \tilde{h}_c)$ problem.

Problem R : *Given $A \in \mathcal{M}^{m\times n}$, $b \in \mathbb{R}^m$, $c \in \mathbb{R}^n$, $\mu, \delta_b, \delta_c \in \mathbb{R}$, $\mu, \delta_b, \delta_c > 0$. Find $x \in \mathbb{R}^n$, $u \in \mathbb{R}^m$, $w_b, w_c \in \mathbb{R}$, $x \geqslant 0$, $w_b, w_c > 0$ such that $\|\tilde{H}\| \leqslant \mu$, $\|\tilde{h}_b\| \leqslant \delta_b$, $\|\tilde{h}_c\| \leqslant \delta_c$, objects $\tilde{H}$, $\tilde{h}_b$, $\tilde{h}_c$ are constructed by formulas* (1)-(3), $\|x\|^2 + \|u\|^2 \to \min$.

Let x^*, u^*, w_b^*, w_c^* and also $\tilde{H}^*$, $\tilde{h}_b^*$, $\tilde{h}_c^*$ be a solution of the problem R. The following theorem holds.

Theorem 4. *Problem $Z_{A,b,c}$ has a solution (for sufficiently small μ, δ_b, δ_c is unique), which has the form: $A_1 = A + \tilde{H}^*$, $b_1 = b + \tilde{h}_b^*$, $c_1 = c + \tilde{h}_c^*$, $x_1 = x^*$, $u_1 = u^*$.*

Numerical example (taken from [6], the vector c_0 is changed). Let the exact data for the problems L, L^* have the form

$$A_0 = \begin{bmatrix} 2 & -2 & 1 & 1 & 1/2 \\ 1 & 1 & 1 & 0 & 1/2 \\ 3 & -1 & 2 & 1 & 1 \end{bmatrix}, b_0 = \begin{bmatrix} 2 \\ 1 \\ 3 \end{bmatrix}, c_0 = \begin{bmatrix} -1 & -1 & 1 & 1 & 1 \end{bmatrix}^\top,$$

$$x_0 = \begin{bmatrix} 0 & 1/14 & 0 & 17/14 & 13/7 \end{bmatrix}^\top, u_0 = \begin{bmatrix} 1/3 & 1/3 & 2/3 \end{bmatrix}^\top.$$

Controlled approximate data have an appearance $A = A_0 + \dfrac{\mu}{\|\Delta A\|} \cdot \Delta A$,

$b = b_0 + \dfrac{\delta_b}{\|\Delta b\|} \cdot \Delta b$, $c = c_0 + \dfrac{\delta_c}{\|\Delta c\|} \cdot \Delta c$, where

$$\Delta A = \begin{bmatrix} 0.017 & 0.051 & -0.044 & -0.304 & 0.001 \\ -0.074 & -0.028 & 0.483 & 0 & -0.473 \\ 0.449 & 0.347 & 0.239 & 0.339 & 0.073 \end{bmatrix}, \Delta b = \begin{bmatrix} 0.031 \\ 0.343 \\ 0.158 \end{bmatrix},$$

$$\Delta c = \begin{bmatrix} 0.342 & -0.391 & -0.186 & -0.214 & -0.363 \end{bmatrix}^\top,$$

$$\varepsilon = (\mu = \delta_b = \delta_c) = 0.5, 0.1, 0.05, 0.01, ..., 0.000005.$$

The computations showed that the problem $L(A, b, c)$ turned out to be an improper problem of LP of the 1st kind, and the problem $L^*(A, b, c)$ turned out to be an improper problem of LP of the 2nd kind [7]. The results of the numerical solution of the $Z_{A,b,c}$ problem obtained in the Matlab® environment using the `fmincon` solver are shown in Table 1 and in Figure 1.

Table 1: The results of solving two $Z_{A,b,c}$ problems

ε	0.5	0.1
$x_1(\varepsilon)$	0.000000000058872 0.000000000000000 0.000000000000000 1.103916802786946 1.293518106519200	−0.000000000265282 0.034983813620750 0.000000002065893 1.153832523852359 1.743212162368541
$u_1(\varepsilon)$	0.304413631462186 −0.111818587775316 0.235593220635580	0.366646929000393 0.303657129746492 0.595586937489473
$w_b^*(\varepsilon)$	1.357109365670370	1.357107524155774
$w_c^*(\varepsilon)$	1.208336885898787	1.004875440528429
$\|x_0 - x_1(\varepsilon)\|$	0.578754003836975	0.134026160942559
$\|u_0 - u_1(\varepsilon)\|$	0.620339340922493	0.083921395457292

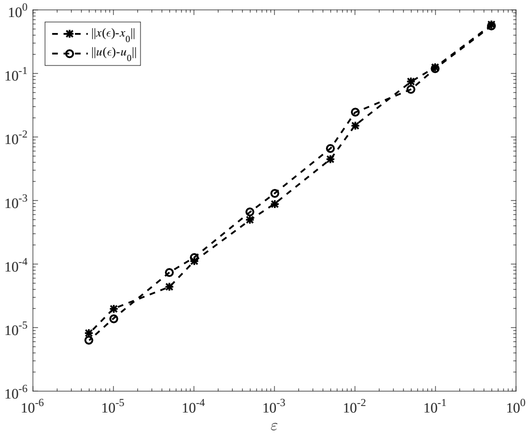


Fig. 1. The results of solving a series of $Z_{A,b,c}$ problems.

References

1. Volkov V. V. et al. Minimum-Euclidean-norm matrix correction for a pair of dual linear programming problems //Comput. Math. Math. Phys. 2017. V. 57, N. 11. P. 1757–1770.
2. Erokhin V. I. A stable solution of linear programming problems with the approximate matrix of coefficients // Proc. CNSA (dedicated to the memory of V.F. Demianov), St. Petersburg, Russia, May 22–27, 2017. P. 88–90.
3. Tikhonov A. N. Approximate systems of linear algebraic equations // USSR Comput. Math. Math. Phys. 1980. V. 20, N 6, P. 10–22.
4. Tikhonov A. N., Arsenin V.Ja. Methods for Solving ill-posed Problems. Moscow: Nauka, 1986 (in Russian).
5. Erokhin V. I. Matrix correction of a Dual Pair of Improper Linear Programming Problems // Comput. Math. Math. Phys. 2007. V. 47, N 4. P. 564–578.
6. Agayan G.M., Ryutin A.A., Tikhonov A.N. The problem of linear programming with approximate data // USSR Comput. Math. Math. Phys. 1984. V. 24, N 5, P. 14–19.
7. Eremin I. I., Mazurov V. D. and Astaf'ev N. N. Improper Problems of Linear and Convex Programming. Moscow: Nauka, 1983 (in Russian).

Solution of the approximate system of linear algebraic equations minimal with respect to the ℓ_1 norm*

V.I. Erokhin[1] and V.V. Volkov[2]
[1]*Mozhaisky Military Space Academy, St. Petersburg, Russia,*
[2]*Borisoglebsk branch of Voronezh State University, Borisoglebsk, Russia*

The paper is devoted to the problem of finding solutions of approximate systems of linear algebraic equations (SLAE). The generalization of the A. N. Tikhonov's method to non-Euclidean vector norms and considered. As a criterion for optimality of the solution in the resulting mathematical programming problems are used minimum of the Holder norm with exponent $p = 1$.

Approximate SLAE can arise in a large number of applied problems of science and technology (see, for example, [1–2]).

Sometimes, singularities in the initial data can lead that methods using the Euclidean norm as a criterion for the solution (for example, the least squares method) will give a worse result than solutions in other norms, in particular, in ℓ_1-norm (the Holder norm with exponent $p = 1$). The probability basis of using the polyhedral ℓ_1 norm is given, in particular, in [3].

We consider the exact joint system of linear algebraic equations $A_0 x = b_0$, where $A_0 \in \mathbb{R}^{m\times n}$, $b_0 \in \mathbb{R}^m$, $b_0 \neq 0$, the relationship between the dimensions of the matrix A_0 and its rank is not specified $x_0 \in \mathbb{R}^n$ is the solution of this system with minimal Holder norm with exponent $p = 1$. Numerical values of A_0, b_0 and x_0 are unknown, instead of them, approximate matrix $A \in \mathbb{R}^{m\times n}$ and vector $b \in \mathbb{R}^m$, $b \neq 0$ are given, such that $\|A_0 - A\|_{1,\psi} \leqslant \mu$, $\psi(b_0 - b) \leqslant \delta < \psi(b)$, where $\mu \geqslant 0$ and $\delta \geqslant 0$ are known parameters, $\psi(\cdot)$ is arbitrary vector norm, $\|\cdot\|_{1,\psi}$ is matrix norm, such that $\|A\|_{1,\psi} := \max\limits_{x\neq 0} \frac{\psi(Ax)}{\|x\|_1}$. The matrix A is not obliged to have a full rank and the compatibility of the system $Ax = b$ are not assumed in the general case.

To find matrix $A_1 \in \mathbb{R}^{m\times n}$ and vectors $b_1 \in \mathbb{R}^m$, $x_1 \in \mathbb{R}^n$ such that $\|A - A_1\|_{1,\psi} \leqslant \mu, \psi(b - b_1) \leqslant \delta, A_1 x_1 = b_1, \|x_1\|_1 \to \min$.

We shall denote this problem by $Z_{1,\psi}(\mu, \delta)$.

*The reported study was funded by RFBR according to the research project 18-31-00083

The problem $Z_{1,\psi}\,(\mu,\delta)$ is the modification (generalization) of the problem considered by A. N. Tikhonov [4] and called by him "regularized least squares method" (RLS). The problem in the original formulation (using Euclidean matrix and vector norms) are explored, for example, in papers [5, 6].

We note that in [7] the problem of finding a normal solution of regularized problems of SLAU is also solved, but with the use of another technique: pairs of mutually ambiguous problems of conditional optimization are used.

The article is based on the results presented in [8].

Let us proceed from problem $Z_{1,\psi}\,(\mu,\delta)$ to the equivalent problem $R_{1,\psi}(\mu,\delta) : \|x\|_1 \to \min\limits_{\psi(b-Ax)=\mu\cdot\|x\|_1+\delta} (=: \chi_{1,\psi})$.

When in the problem $Z_{1,\psi}\,(\mu,\delta)$ the Holder norm with exponent $p = 1,\infty$ is chosen as $\psi(\cdot)$, we obtain the following generalized problems of RLS and their reductions to the mathematical programming problems[8] (the symbol "$\mapsto$" means "is reduced to"):

$$Z_{1,1}\,(\mu,\delta) \mapsto R_{1,1}(\mu,\delta) : \|b-Ax\|_1 \leqslant \mu\cdot\|x\|_1+\delta,\ \|x\|_1 \to \min, \quad (1)$$

$$Z_{1,\infty}\,(\mu,\delta) \mapsto R_{1,\infty}(\mu,\delta) : \|b-Ax\|_\infty \leqslant \mu\cdot\|x\|_1+\delta,\ \|x\|_1 \to \min\,. \quad (2)$$

The problems (1)–(2), despite the seeming simplicity of the formulations, do not have obvious methods and algorithms for solving in the general case. Therefore, for now, we consider a special case of solving these problems. In subsequent calculations it is assumed that all elements of the vector x are non-negative (or the signs of all elements of this vector are known to us). In this case, $\|x\|_1 = \sum_{i=1}^n x_i$ and the problems (1)–(2) can be reduced to linear programming (LP) problems.

Consider the problem $Z_{1,1}\,(\mu,\delta)$.

Let the matrix $A \in \mathbb{R}^{m\times n}$ and the vector $b \in \mathbb{R}^m$, $b \neq 0$ are known.

It is required to find $A_1 \in \mathbb{R}^{m\times n}$, $x_1 \in \mathbb{R}^n$, $b_1 \in \mathbb{R}^m$ such that $\|A-A_1\|_1 \leqslant \mu$, $\|b-b_1\|_1 \leqslant \delta$, $A_1x_1 = b_1$, $\|x_1\|_1 \to \min$, where $\mu,\delta \geqslant 0$ are known a priori, simultaneously non-zero constants.

Theorem 1. Problem $Z_{1,1}\,(\mu,\delta)$ have a solution if and only if mathematical programming problem $R_{1,1}(\mu,\delta)$ have a solution.

Consider linear programming problem:

$$\begin{gathered}
-d \leqslant b-Ax \leqslant d,\\
1_n^\top x = \chi,\\
1_m^\top d \leqslant \mu\chi+\delta,\\
d \geqslant 0,\quad x \geqslant 0,\quad \chi \geqslant 0,\quad \chi \to \min\,.
\end{gathered} \quad (3)$$

Theorem 2. If problem $R_{1,1}(\mu, \delta)$ have a solution, it can be found as x^*, where $\chi^* = \|x^*\|_1$ is a solution of problem (3), $1_n \in \mathbb{R}^n$ and $1_m \in \mathbb{R}^m$ are unit vectors, $d \in \mathbb{R}^m$, $\chi \in \mathbb{R}$, $x \in \mathbb{R}^n \geqslant 0$.

Reasoning similarly, consider the problem $Z_{1,\infty}(\mu, \delta)$.

Theorem 3. Problem $Z_{1,\infty}(\mu, \delta)$ have a solution if and only if mathematical programming problem $R_{1,\infty}(\mu, \delta)$ have a solution.

Consider linear programming problem:

$$\begin{gathered} -\beta \cdot 1_m \leqslant b - Ax \leqslant \beta \cdot 1_m, \\ 1_n^\top x = \chi, \\ \beta \leqslant \mu\chi + \delta, \\ \beta \geqslant 0, \quad \chi \geqslant 0, \quad x \geqslant 0, \quad \chi \to \min . \end{gathered} \tag{4}$$

Theorem 4. If problem $R_{1,\infty}(\mu, \delta)$ have a solution, it can be found as x^*, where $\chi^* = \|x^*\|_1$ is a solution of problem (4), $1_n \in \mathbb{R}^n$ and $1_m \in \mathbb{R}^m$ are unit vectors, $\pi, \theta \in \mathbb{R}$, $x \in \mathbb{R}^n \geqslant 0$.

Computational experiments

There are the results of a numerical solution of model systems for problems (1) and (2). Calculations are performed using Matlab. The corresponding auxiliary LP problems were solved by the simplex method (using the linprog solver with the 'simplex' option). $x_{RLN(1,1)}$ is the solution of the problem (1) (RLN — Regularized Least Norm, regularized solution using the norm ℓ_1, $x_{RLN(1,\infty)}$ is the solution of the problem (2) (regularized solution using the norms ℓ_1 and ℓ_∞). The symbol $|| \cdot ||_E$ is used to denote the Euclidean (ℓ_2) matrix norm.

A series of problems of the form (1) (Fig. 1) and the form (2) (Fig. 2) with a decreasing error was considered. For each of these problems, solutions were found by three methods. The results of the computational experiments are presented below dependences of the errors of the solution $\varepsilon_{RLN(1,1)} = \|x_{RLN(1,1)} - x_0\|$ (Fig. 1 only), $\varepsilon_{RLN(1,\infty)} = \|x_{RLN(1,\infty)} - x_0\|$ (Fig. 2 only), $\varepsilon_{RLS} = \|x_{RLS} - x_0\|$ and $\varepsilon_{LS} = \|x_{LS} - x_0\|$, on the error parameter e.

The graphs shown in Figs. 1 and 2 show that there are such approximate SLAE for which the solution $x_{RLN(1,1)}$ and $x_{RLN(1,\infty)}$ is closer (according to the Euclidean norm) to the solution of the “exact” SLAE than the solutions obtained by LS and RLS methods. In addition, the graphs show that with a decrease in the amount of error imposed on the original matrix, the error of all the considered methods decreases.

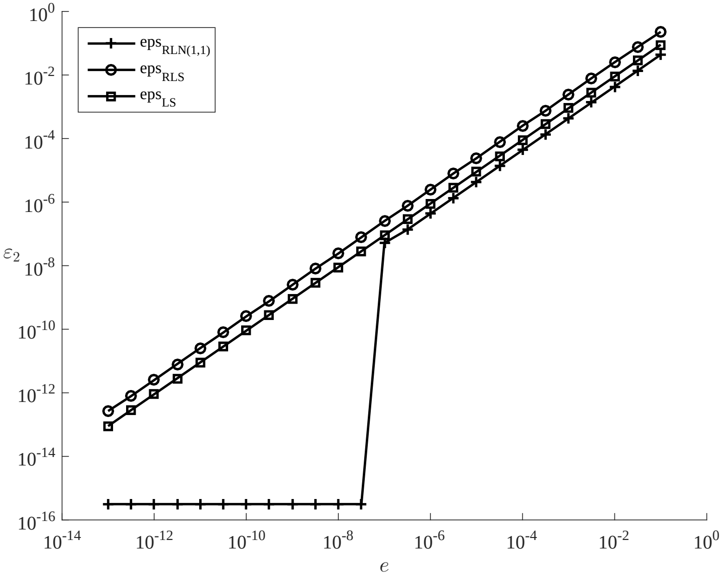


Fig. 1. The result of the first computational experiment (problem $Z_{1,1}(\mu,\delta)$).

References

1. Erokhin V.I. et al. Using negative regularization parameter in Tikhonov's regularized least squares method // Izvestija Sankt Peterburgskogo Gosudarstvennogo Techno-logicheskogo Instituta (Technicheskogo Universiteta), 2014. N. 24(50). P. 86–92. (In Russian)
2. Erokhin V.I., Volkov V.V. Recovering images, registered by device with inexact point-spread function, using tikhonov's regularized least squares method // Int. Journal of Artificial Intelligence. 2015. V. 13, N. 1. 12 p. Available at: http://www.ceser.in/ceserp/index.php/ijai/article/view/3531. (accessed: 21.03.2018)
3. Gorelik V.A., Trembacheva (Barkalova) O. S. Solution of the linear regression problem using matrix correction methods in the l_1 metric // Comput. Math. Math. Phys. 2016. V. 56, N. 2. P. 200–205.
4. Tikhonov A.N. Approximate systems of linear algebraic equations

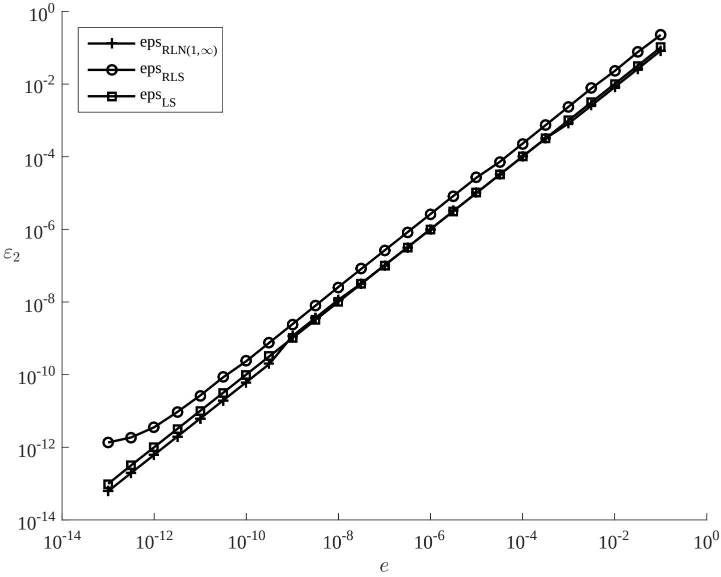


Fig. 2. The result of the second computational experiment (problem $Z_{1,\infty}(\mu,\delta)$).

// USSR Comput. Math. Math. Phys. 1980. V. 20, N. 6, P. 10–22.

5. Volkov V.V., Erokhin V.I. Tikhonov solutions of approximately given systems of linear algebraic equations under finite perturbations of their matrices // Comput. Math. Math. Phys. 2010. V. 50, N. 4. P. 589–605.
6. Erokhin V.I., Volkov V.V. About A. N. Tikhonov's regularized least squares method // Vestnik Sankt-Peterburgskogo Universiteta, Prikladnaya Matematika, Informatika, Protsessy Upravleniya, 2017. Issue 1. P. 4–16. (In Russian)
7. Golikov A.I., Evtushenko Y.G. Regularization and normal solutions of systems of linear equations and inequations // Proc. of the Steklov Institute of Mathematics. 2015. V. 289, N. 1. P. 102–110.
8. Volkov V.V. et al. Generalizations of Tikhonov's regularized method of least squares to non-Euclidean vector norms // Comput. Math. Math. Phys. 2017. V. 57, N. 9. P. 1416–1426.

Parametric correction of inconsistent systems of linear equations and improper linear programming problems*

V.A. Gorelik and T.V. Zolotova
Dorodnicyn Computing Centre, FRC CSC RAS,
Moscow, Russia,
Financial University under the Government of the Russian Federation,
Moscow, Russia

This paper is devoted to a new class of parameters correction problems for improper optimization problems. The inaccuracy of the model's initial data, the stringent conditions imposed on the variables, the contradictory nature of the requirements imposed on the model of the system, can lead to the fact that the constraints are inconsistent, reflecting the absence of homeostasis in the system. The corresponding optimization problems have no solution and are called improper. For such problems, procedures for minimum data correction are proposed [1], as a result of which approximating problems already have a solution ensuring achievement of the homeostasis region of the system.

It should be taken into account that the parameters of the system models are, as a rule, interrelated (for example, the technological parameters of the ecological and economic systems influence the level of environmental pollution). In this connection, a class of parametric correction problems is introduced in this paper, in which the entries of the constraint matrix cannot be corrected directly, but change due to the correction of the other matrix entries. Such mathematical statements can arise, for example, in the modeling of production problems, when the technology of the enterprise depends on the level of technology development in another field creating this technology, or when there are restrictions on the level of pollution, while the matrix of pollution coefficients depends on the technology of production.

Let us consider a class of linear programming problems (LP) in which the matrix of left-hand side constraints is formed by a linear relation:

$$\max_x \{\langle c, x\rangle \mid D(A)x = b,\ x \geqslant 0\}, \tag{1}$$

$$D(A) = D_0 + D_1 A, \tag{2}$$

*The work was carried out in the framework of Project No. 1.8535.2017 of the Ministry of Education and Science of the Russian Federation

where $x = (x_1, ..., x_n)$ is the vector of variables, $c = (c_1, ..., c_n)$ is the vector of the coefficients of the objective function, $b = (b_1, ..., b_m)$ is the right-hand side constraint vector, A is a matrix of a size $m \times n$, D_0 is a matrix of a size $m \times n$, D_1 is a non-degenerate matrix of a size $m \times m$ (D_1^{-1} exists).

Suppose that the system of constraints in the problem (1) is inconsistent. We formulate such statements of data correction problems in which a direct correction of the entries of the matrix A is possible, and the entries of the matrix D undergo a change according to (2). Such problems will be called parametric correction.

We will use here the matrix norm l_1 as a minimized criterion for the correction value (the sum of the modules of its entries). We denote the correction matrix of the entries of the matrix A by H. This matrix must satisfy the condition

$$D_0x + D_1(A - H)x = b$$

or

$$Hx = D_1^{-1}D_0x + Ax - D_1^{-1}b. \tag{3}$$

The minimum correction matrix H according to the norm l_1 for a fixed x is determined by the formulas

$$|h_{ij_0}| = \frac{|(D_1^{-1}D_0x + Ax - D_1^{-1}b)_i|}{\max\limits_j x_j}, \; h_{ij} = 0, \; j \neq j_0, i \; = \; 1, \ldots, m, \tag{4}$$

where $(\cdots)_i$ is the i-th component of the vector in parentheses, and the index j_0 is determined from the condition $x_{j_0} = \max\limits_j x_j$.

Indeed, the i-th component of the vector Hx of the left-hand side of the formula (3) is a linear function of the components of the i-th row of the matrix H, and the coefficients are non-negative components of the vector x. This linear function is equal to some constant, and if this constant is positive, the minimum of the sum of components of the i-th row of the matrix H is achieved when they all are equal to zero except for the component with the largest coefficient and it is equal to the constant divided by this coefficient. If this constant is negative, then everything is the same, only for a minimum of the sum of the modules. Thus, for fixed x, the correction matrix H, which is minimal in the norm l_1 and satisfies (3), has one nonzero column and modules of its component are computed by the formulas (4). The minimal correction problem is reduced to minimizing the sum of the modules of these entries of the matrix H with respect to x, where the vector x must satisfy the condition $x \geqslant 0$.

We introduce the variable $y = (\max_j x_j)^{-1}$ and the vector $z = yx$. The components of the vector z must satisfy the conditions $0 \leq z_j \leq 1$, and there exists an index j_0 such that $z_{j_0} = 1$. We also introduce the variables u_i, which are not less than the expressions on the right-hand side of the formulas (4), and m-dimensional vectors $u = (u_1, ..., u_m)$, $e = (1, ..., 1)$. In the new variables we get the problem of mathematical programming

$$\begin{gathered}\min_{u,z,y}\{\langle e, u\rangle | u \geqslant \pm(D_1^{-1} D_0 z + Az - D_1^{-1} by), \\ u \geqslant 0,\ y \geqslant 0,\ z \geqslant 0,\ z_j \leqslant 1,\ \exists j : z_j = 1\}.\end{gathered} \tag{5}$$

This is a problem of linear, partially integer programming. It can be reduced to solving n ordinary LP problems by setting $z_j = 1$ for j =1,. . . ,n, and choosing from these problems the one that gives the least value of the criterion $\langle e, u\rangle$. By the obtained from this problem y and z we find the vector $x = z/y$ and by the formulas (4) the entries of the correction matrix (the signs of entries are determined by the signs of the expression inside the module).

Example 1. $A = \begin{pmatrix} 10 & 17 \\ 18 & 19 \\ 17 & 16 \end{pmatrix}$, $D_0 = \begin{pmatrix} 1 & 3 \\ 4 & 6 \\ 1 & 2 \end{pmatrix}$, $D_1 = \begin{pmatrix} 1 & 2 & 6 \\ 3 & 5 & 4 \\ 7 & 5 & 1 \end{pmatrix}$, $b = \begin{pmatrix} 19 \\ 15 \\ 20 \end{pmatrix}$, $c = \begin{pmatrix} 7 \\ 8 \end{pmatrix}$.

The system of constraints $D(A)x = b$, $x \geqslant 0$ in the problem (1) is inconsistent.

Solving the correction problem (5) for $z_1 = 1$, we obtain $z_2 = 0$, $u = \begin{pmatrix} 0 \\ 23.656 \\ 8.902 \end{pmatrix}$, y=2.541, $\langle e, u\rangle = 32.557$, $x = \frac{z}{y} = \begin{pmatrix} 0.394 \\ 0 \end{pmatrix}$, $\langle c, x\rangle = 2.755$. Solving the correction problem (5) for $z_2 = 1$, we obtain $z_1 = 0$, $u = \begin{pmatrix} 0 \\ 28.059 \\ 2.131 \end{pmatrix}$, y=4.479, $\langle e, u\rangle = 30.19$, $x = \frac{z}{y} = \begin{pmatrix} 0 \\ 0.223 \end{pmatrix}$, $\langle c, x\rangle = 1.786$.

The second case gives a smaller value of the objective function $\langle e, u\rangle = 30.19$ of the problem (5), therefore $x = \begin{pmatrix} 0 \\ 0.223 \end{pmatrix}$ is the solution of the problem (1) with a correction matrix $H = \begin{pmatrix} 0 & 0 \\ 0 & 28.059 \\ 0 & 2.131 \end{pmatrix}$.

So far we have considered the correction of the system of constraints for the LP problem without taking into account its initial criterion. For an inconsistent system of equations this is natural, but for the LP problem under correction there is essentially a two-criteria problem (maximization of the initial criterion and minimization of the correction matrix norm). As in the above example, they are usually contradictory: the smaller the correction matrix, the smaller the value of the original criterion (given that it is subject to maximization).

In [2], an approach was proposed to this two-criteria problem, which consists in transfer the initial criterion into a constraint (by the way, even with a consistent constraint system, the requirement of achieving a certain threshold value of the criterion may lead to the improper problem).

We apply this approach to the problem (1), namely, we introduce a threshold value c_0 and the requirement $\langle c, x\rangle \geqslant c_0$. It gives an additional condition $\langle c, z\rangle \geqslant c_0 y$ in the variables y, z.

We obtain the correction problem

$$\begin{gathered}\min_{u,z,y}\{\langle e, u\rangle | u \geqslant \pm(D_1^{-1}D_0 z + Az - D_1^{-1}by),\ \langle c,z\rangle \geqslant c_0 y, \\ u \geqslant 0,\ y \geqslant 0,\ z \geqslant 0,\ z_j \leqslant 1,\ \exists j : z_j = 1\}.\end{gathered} \tag{6}$$

Example 2. We take the values of the matrices and vectors A, D_0, D_1, b, c from Example 1 and find the solution of the problem (6) with $c_0 = 3$.

Solving the correction problem (6) for $z_1 = 1$, we obtain $z_2 = 0$, $u = \begin{pmatrix} 0.745 \\ 23.314 \\ 9.549 \end{pmatrix}$, y=2.333, $\langle e, u\rangle = 33.608$, $x = \frac{z}{y} = \begin{pmatrix} 0.429 \\ 0 \end{pmatrix}$, $\langle c, x\rangle = 3$.

Solving the correction problem (6) for $z_2 = 1$, we obtain $z_1 = 0$, $u = \begin{pmatrix} 6.502 \\ 25.075 \\ 7.78 \end{pmatrix}$, y=2.667, $\langle e, u\rangle = 39.357$, $x = \frac{z}{y} = \begin{pmatrix} 0 \\ 0.375 \end{pmatrix}$, $\langle c, x\rangle = 3$.

The first case gives a smaller value of the objective function $\langle e, u\rangle = 33.608$ of the problem (6), therefore, $x = \begin{pmatrix} 0.429 \\ 0 \end{pmatrix}$ is the solution of problem (1) with the correction matrix $H = \begin{pmatrix} 0.745 & 0 \\ 23.314 & 0 \\ 9.549 & 0 \end{pmatrix}$ and the value of the target function $\langle c, x\rangle = 3$.

References

1. Eremin I.I., Mazurov V.D., Astafiev N.N. The improper problems of linear and convex programming. Moscow: Fizmatlit, 1983.
2. Gorelik V.A. Matrix correction of a linear programming problem with inconsistent constraints // Computational Mathematics and Mathematical Physics. 2001. V. 41, No. 11. P. 1697–1705.

Matrix correction of a dual pair of improper linear programming problems with a minimum weighted Euclidean norm*

M.N. Khvostov[1], V.I. Erokhin[2], and S.V. Sotnikov[2]

[1]Borisoglebsk Branch of Voronezh State University, Borisoglebsk, Russia, [2]Mozhaisky Military Space Academy, St. Petersburg, Russia

Linear optimization models are widely used in various fields of science and economic. The problems of linear programming occupy an important place among these models, for example

$$L(A, b, c): \; Ax = b, \; x \geqslant 0, \; c^\top x \to \max \tag{1}$$

where $A \in \mathbb{R}^{m\times n}$, $c, x \in \mathbb{R}^n$, $b, u \in \mathbb{R}^m$ $\mathcal{X}(A, b) \triangleq \{x \,|\, Ax = b, \; x \geqslant 0\}$ is the feasible set. The linear programming problem (1) is given in the canonical form.

However, the problems of linear programming are often insoluble in practice. Thus, the constraint system of problem (1) can be inconsistent, i.e. $\mathcal{X}(A, b) = \varnothing$. There is a need to clarify, change such problems. As a result, we must obtain a solvable linear programming problem that is similar to the original problem of linear programming in some sense. Thus, there is a need for matrix correction of linear programming problems. Matrix correction of linear programming problems is a change (correction) of any coefficients of both left and right parts of equations and

*The reported study was funded by RFBR according to the research project N 18-31-00083

inequalities of constraints of linear programming problems, arbitrary sets of these coefficients and/or coefficients of objective functions in order to ensure the consistency of the indicated constraints [1]:

$$P\left(A,b\right):\quad \begin{cases} \|[H \quad -h]\| \to \min, \\ \mathcal{X}\left(A+H,b+h\right) \neq \varnothing, \end{cases} \tag{2}$$

where $\|\cdot\|$ is the norm for the matrix, for example, the Euclidean norm.

Often there is a need to include in the model (2) additional data on the laboriousness of changing each of its parameters. One of the most well-known methods is the application of the weighted Euclidean norm. One of the most common options for applying the weighted norm is to calculate the Euclidean matrix norm after multiplying the correction matrix by nondegenerate matrices on the left and right [2]:

$$Z^{LR}\left(A,b\right):\quad \begin{cases} \|L\cdot[H \quad -h]\cdot R\| \to \min, \\ \mathcal{X}\left(A+H,b+h\right) \neq \varnothing, \end{cases} \tag{3}$$

where $L \in \mathbb{R}^{m\times m}$, $R \in \mathbb{R}^{(n+1)\times(n+1)}$, L, R is nondegenerate.

Another way to add additional data to the linear optimization model is to apply the weighted Euclidean norm obtained by using the Hadamard multiplication for a correction matrix by a matrix with positive coefficients [3]:

$$Z^{W}\left(A,b\right):\quad \begin{cases} \|W\circ[H \quad -h]\| \to \min, \\ \mathcal{X}\left(A+H,b+h\right) \neq \varnothing, \end{cases}$$

where $W \in \mathbb{R}^{m\times(n+1)}$, $W_{ij} > 0$, $i \in 1,\ldots,m$, $j \in 1,\ldots,n+1$.

Correction of the constraint system of the linear programming problem makes the feasible set non-empty. But this does not guarantee that the new linear programming problem will be solvable. Correction methods for improper linear programming problems that guarantee the property of corrected problems require the use of duality theory and correction of the dual linear programming problem [4]:

$$L^{*}(A,b,c):\ u^{\top}A \geqslant c^{\top},\ b^{\top}u \to \min, \tag{4}$$

where $\mathcal{U}(A,c) \triangleq \left\{u \left| u^{\top}A \geqslant c^{\top}\right.\right\}$ is the feasible set of problem (3). Henceforth, the problem (1) will be called the primal linear program problem.

The above methods make the feasible sets of the direct problem of linear programming and the dual linear programming problem nonempty.

That allows to correct any improper problem of linear programming

$$D\left(A,b,c\right):\quad\left\{\begin{array}{l}\left\|\left[H\quad -h\right]\right\|\to\min,\\ \left(A+H\right)x=\left(b+h\right),\ x\geqslant 0,\ c^{\top}x\to\max,\\ u^{\top}\left(A+H\right)x\geqslant c^{\top},\ \left(b+h\right)^{\top}u\to\min.\end{array}\right.\tag{5}$$

Thus, we try to combine the advantages of problems (3) and (5). We obtain matrix correction of the double pair of incorrect linear programming problems with a minimal correction matrix by the weighted Euclidean norm

$$W\left(A,b,c\right):\quad\left\{\begin{array}{l}\left\|L\cdot\left[H\quad -h\right]\cdot R\right\|\to\min,\\ \left(A+H\right)x=\left(b+h\right),\ x\geqslant 0,\ c^{\top}x\to\max,\\ u^{\top}\left(A+H\right)x\geqslant c^{\top},\ \left(b+h\right)^{\top}u\to\min.\end{array}\right.\tag{6}$$

The following theorems are modifications of the corresponding theorems for matrix correction of the double pair of incorrect linear programming problems (5).These theorems contain the condition for the existence of a solution and its form for problem (6) for a given nonzero solution.

Theorem 1. Let $b\in\mathbb{R}^m$ and $c\in\mathbb{R}^n$ be given. A is the set of matrices for which the vector $\bar{x}\in\mathbb{R}^n$, $\bar{x}\geqslant 0$ is included in the set of solutions of the linear programming problem $L(A,b,c)$, the vector $\bar{u}\in\mathbb{R}^m$, $\bar{u}\neq 0$ is included in the set of solutions of the linear programming problem $L^{*}(A,b,c)$. A exists if and only if

$$c^{\top}\bar{x}=b^{\top}\bar{u}=\alpha.$$

A has the form

$$A=\widehat{A}+\triangle A,$$

$$\widehat{A}=\frac{b\bar{x}^{\top}\left(RR^{\top}\right)^{-1}}{\bar{x}^{\top}\left(RR^{\top}\right)^{-1}\bar{x}}+\frac{\left(L^{\top}L\right)^{-1}\bar{u}d^{\top}}{\bar{u}^{\top}\left(L^{\top}L\right)^{-1}\bar{u}}-\alpha\frac{\left(L^{\top}L\right)^{-1}\bar{u}\bar{x}^{\top}\left(RR^{\top}\right)^{-1}}{\bar{u}^{\top}\left(L^{\top}L\right)^{-1}\bar{u}\bar{x}^{\top}\left(RR^{\top}\right)^{-1}\bar{x}},$$

where $\widehat{A}$ is matrix with a minimal weighted Euclidean norm,

$$d=\left[d_1,\ldots,d_n\right]^{\top},\quad d_i=\left\{\begin{array}{l}0,\ \text{if } c_i\leqslant 0 \text{ and } x_i=0,\\ c_i,\ \text{otherwise},\end{array}\right.$$

$\triangle A$ is any matrix such that

$$\bar{u}^{\top}\triangle A=0,\quad\triangle A\bar{x}=0.$$

Also

$$\left\|L\widehat{A}R\right\|^2 = \frac{\|Lb\|^2}{\|R^{-1}\bar{x}\|^2} + \frac{\left\|d^\top R\right\|^2}{\|L^{-1}\bar{u}\|^2} - \frac{\alpha^2}{\|R^{-1}\bar{x}\|^2\,\|L^{-1}\bar{u}\|^2},$$

where $\|\cdot\|$ is the Euclidean norm, $L \in \mathbb{R}^{m\times m}$, $R \in \mathbb{R}^{n\times n}$.

Theorem 2. Let (1) and (4) be improper linear programming problems. $[H \quad -h]$ is the set of matrices for which the vector $\bar{x} \in \mathbb{R}^n$, $\bar{x} \geqslant 0$ is included in the set of solutions of the linear programming problem $L(A+H, b+h, c)$, the vector $\bar{u} \in \mathbb{R}^m$, $\bar{u} \neq 0$ is included in the set of solutions of the linear programming problem $L^*(A+H, b+h, c)$. The matrix $[H \quad -h]$ has the form

$$[H \quad -h] = \left[\widehat{H} \quad -\widehat{h}\right] + [\triangle H \quad -\triangle h]$$

where $\left[\widehat{H} \quad -\widehat{h}\right]$ is matrix with a minimal weighted Euclidean norm,

$$\begin{aligned}\left[\widehat{H} \quad -\widehat{h}\right] = \frac{(b - A\bar{x})\left[\bar{x}^\top \;\; 1\right]\left(RR^\top\right)^{-1}}{\left[\bar{x}^\top \;\; 1\right]\left(RR^\top\right)^{-1}\left[\bar{x}^\top \;\; 1\right]^\top} + {} \\ + \frac{\left(L^\top L\right)^{-1}\bar{u}\left[d^\top \;\; b^\top\bar{u} - c^\top\bar{x}\right]}{\bar{u}^\top\left(L^\top L\right)^{-1}\bar{u}} - {} \\ - \alpha\frac{\left(L^\top L\right)^{-1}\bar{u}\left[\bar{x}^\top \;\; 1\right]\left(RR^\top\right)^{-1}}{\left[\bar{x}^\top \;\; 1\right]\left(RR^\top\right)^{-1}\left[\bar{x}^\top \;\; 1\right]^\top\bar{u}^\top\left(L^\top L\right)^{-1}\bar{u}},\end{aligned} \tag{7}$$

$$\alpha = b^\top\bar{u} - \bar{u}^\top A\bar{x},$$

$$d = [d_1, \ldots, d_n]^\top, \quad d_i = \begin{cases} 0, \text{ if } \left(c - A^\top\bar{u}\right)_i \leqslant 0 \text{ and } x_i = 0, \\ \left(c - A^\top\bar{u}\right)_i, \text{ otherwise,}\end{cases}$$

$[\triangle H \quad -\triangle h]$ is any matrix such that

$$\bar{u}^\top[\triangle H \quad -\triangle h] = 0, \quad [\triangle H \quad -\triangle h]\left[\bar{x}^\top \;\; 1\right]^\top = 0.$$

Also

$$\begin{aligned}\left\|L\left[\widehat{H} \quad -\widehat{h}\right]R\right\|^2 = \frac{\|L(b - A\bar{x})\|^2}{\left\|R^{-1}\left[\bar{x}^\top \;\; 1\right]^\top\right\|^2} + \frac{\left\|\left[d^\top \;\; b^\top\bar{u} - c^\top\bar{x}\right]R\right\|^2}{\|L^{-1}\bar{u}\|} - {} \\ - \frac{\alpha^2}{\|R^{-1}\bar{x}\|\,\|L^{-1}\bar{u}\|},\end{aligned} \tag{8}$$

where $H, \widehat{H}, \triangle H \in \mathbb{R}^{m\times n}$, $h, \widehat{h}, \triangle h \in \mathbb{R}^m$, $L \in \mathbb{R}^{m\times m}$, $R \in \mathbb{R}^{(n+1)\times(n+1)}$.

Example. We find the correction matrix that is minimal in the weighted Euclidean norm, which ensures the existence of given nonzero solutions. The problem has parameters

$$A = \begin{bmatrix} -3 & 2 & 1 & 3 & -2 \\ 2 & -3 & 4 & 1 & 0 \\ 5 & 3 & 1 & 2 & -3 \\ 1 & 0 & -1 & 1 & 0 \end{bmatrix}, b = \begin{bmatrix} 1 \\ 2 \\ 2 \\ -3 \end{bmatrix}, c = \begin{bmatrix} -1 \\ 1 \\ 0 \\ 3 \\ 2 \end{bmatrix},$$

$$L = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 3 & 0 \\ 0 & 0 & 0 & 4 \end{bmatrix}, R = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 & 0 & 0 \\ 0 & 0 & 3 & 0 & 0 & 0 \\ 0 & 0 & 0 & 4 & 0 & 0 \\ 0 & 0 & 0 & 0 & 5 & 0 \\ 0 & 0 & 0 & 0 & 0 & 6 \end{bmatrix},$$

$$\bar{x}^\top = \begin{bmatrix} 2 & 3 & 2 & 1 & 5 \end{bmatrix}, \bar{u}^\top = \begin{bmatrix} 2 & -1 & -3 & 0 \end{bmatrix}.$$

Using (7), we obtain

$$\widehat{H} = \begin{bmatrix} -1.4851 & 0.2712 & 1.6465 & 1.4512 & -1.3753 \\ -0.8922 & -0.2988 & -0.2843 & -0.2035 & 0.1013 \\ -2.0459 & -0.7196 & -0.4742 & -0.2981 & 0.0494 \\ -1.0277 & -0.3854 & -0.1142 & -0.0321 & -0.1028 \end{bmatrix},$$

$$\widehat{h}^\top = \begin{bmatrix} 7.6513 & -0.9466 & -1.2502 & 0.0143 \end{bmatrix},$$

and using (8) , we obtain $\left\| L \left[\widehat{H} \quad -\widehat{h} \right] R \right\|^2 = 2982.8568.$

Calculations confirm the fulfillment of the conditions

$$\left(A + \widehat{H}\right)\bar{x} = b + \widehat{h}, \quad \bar{u}^\top \left(A + \widehat{H}\right) \geqslant c^\top.$$

Comment. The choice of the forms of linear programming problems does not affect the generality of reasoning. Since the theorems proved for some forms of representation can be used for other forms.

References

1. Gorelik V.A., Erokhin V.I., Pechenkin R.V. Minimax matrix correction of inconsistent systems of linear algebraic equations with block matrices of coefficients // Journal of Computer and Systems Sciences International. 2006. Vol. 45. No 5. P. 727–737.

2. Erokhin V.I. Matrix correction of the improper linear programming problems on the minimum of Euclidean norm with the arbitrary weights and the fixed elements // Mathematical programming: Proceedings of XIII Baikal International School-seminar ”Optimization methods and their applications”, July , 2 - 8, Irkutsk, Baikal, 2005. Vol. 1. Irkutsk: Melentiev Energy Systems Institute SB RAS. 2005. P. 105–110.
3. Khvostov M.N. On sufficient conditions for the solvability of improper linear programming one of the first kind with minimal weighted Euclidean norm for structural matrix correction of the feasible region // Proceedings of Voronezh State University. Series: Physics. Mathematics. 2015. No 2. P 150–167.
4. Eremin I.I., Vatolin A.A. Duality in improper mathematical programming problems under uncertainty // Stochastic optimization. Proc. Int. Conf. Lect. Notes Control Inf. Sci. 81, 1986. P. 326–333.

Minimax matrix correction of inconsistent systems of linear inequalities and improper linear programming problems[*]

O.V. Murav’eva

Moscow Pedagogical State University, Moscow, Russian Federation

Introduction

Consider a linear programming problem (LPP) in standart form with inconsistent constraints

$$(c,x) \to \max, \; Ax \leqslant b, \; x \geqslant 0,$$

where $A \in \mathbb{R}^{m\times n}$, $x, c \in \mathbb{R}^n$, $b \in \mathbb{R}^m$, $X = \{Ax \leqslant b, \; x \geqslant 0\} = \varnothing$.

We formulate the problem of minimal matrix correction after which the LPP has a feasible solution as follows:

$$\inf_{x,H}\{\|H\|_\infty \colon (A+H)x \leqslant b, x \geqslant 0\}, \qquad (1)$$

where $\|H\|_\infty = \max_{i,j} |h_{ji}|$.

In [1-7], it was proved for the matrix norm $\|H\|_\infty$ that many matrix correction problem can be reduced to LPP.

[*] This research is supported by Ministry of Education and Science, project 1.8524.2017/6.7

In what follows, we consider the matrix correction of incompatible systems of linear inequalities with nonnegativity condition and infeasible LPP in standart form.

Matrix correction of inconsistent systems of linear inequalities

In [3], it was proved that solving the problem of matrix correction of an inconsistent system of linear equations with the nonnegativity condition is reduced to a linear programming problem.

Using slack variables, we can convert a system of linear inequalities $Ax \leqslant b$, $x \geqslant 0$ into a system of linear equalities with $n+m$ variables and m constraints. Then the correction problem [1] is reducible to a LPP.

If the LPP

$$\min_{u,e_0,e}\left\{u\colon \sum_{j=1}^{n} e_j = 1,\ Ae - e_0 b + y \leqslant \mathbf{1}^m u,\ -Ae + e_0 b - y \leqslant \mathbf{1}^m u,\right.$$
$$\left. e, e_0, y \geqslant 0\right\} \quad (2)$$

has the optimal solution $u^* \in \mathbb{R}$, $e^* \in \mathbb{R}^{n+1}$, $y^* \in \mathbb{R}^m$, then

$$\inf_{x,H}\{\|H\|_\infty\colon (A+H)x \leqslant b, x \geqslant 0\} = u^*.$$

If $e_0^* \neq 0$, when

$$x^* = \frac{1}{e_0^*}e^*,\ H^* = (b - Ax^* - \frac{1}{e_0^*}y^*)(x^*)^+,$$

where $(x^*)^+$ is dual to the vector x^* relative to vector norm $\|\cdot\|_1$.

The associated [2] dual linear program is given by

$$\max_{v,s,t}\left\{v\colon \sum_{i=1}^{m} s_i + \sum_{i=1}^{m} t_i \leqslant 1,\ s - t \leqslant 0,\ -(b,s) + (b,t) \leqslant 0,\right.$$
$$\left. -A^T s + A^T t \geqslant v\mathbf{1}^n,\ s,t \geqslant 0\right\}. \quad (3)$$

The LPP [3] can be reduced to the LPP with $m+1$ variables and $n+2$ constraints.

As a result, we obtain the following representation for problem [1].

Let the system of linear inequalities $Ax \leqslant b$, $x \geqslant 0$ be inconsistent.

1. If the primal and dual LPP

$$\max_{z_0,z}\left\{\sum_{j=1}^{n} z_j\colon Az - z_0 b \leqslant \mathbf{1}^m,\ z, z_0 \geqslant 0\right\}, \tag{4}$$

$$\min_{y}\left\{\sum_{i=1}^{m} y_i\colon (b,y) \leqslant 0,\ A^T y \geqslant \mathbf{1}^n,\ y \geqslant 0\right\} \tag{5}$$

   have optimal solutions $z_0^*, z^* = (z_1^*, \ldots, z_n^*)$ and $y^* = (y_1^*, \ldots, y_m^*)$, respectively, then

$$v^* = \inf_{x,H}\{\|H\|_\infty\colon (A+H)x \leqslant b, x \geqslant 0\} = \frac{1}{\sum\limits_{j=1}^{n} z_j^*} = \frac{1}{\sum\limits_{i=1}^{m} y_i^*}.$$

2. If $z_0^* > 0$, then the correction problem also has an optimal solution

$$x^* = \frac{1}{z_0^*} z^*,\ H^* = \triangle b (x^*)^+, \triangle b_i = \begin{cases} 0, & \text{if } (b - Ax^*)_i \geqslant 0, \\ v^*, & \text{otherwise.} \end{cases}$$

   If for all optimal solutions of [4] $z_0^* = 0$, the correction problem [1] has no optimal solution.

3. If the LPP [4] and [5] are improper (the primal problem [4] is unbounded, and the dual problem [5] is infeasible), then the objective value of the correction problem [1] is 0, and the correction problem has no optimal solution.

Matrix correction of infeasible LPP in standart form

For an improper LPP with incompatible constraints

$$(c,x) \to \max,\ Ax \leqslant b,\ x \geqslant 0,$$

we consider the problem of minimal correction of the constraint matrix under the restriction from below to the value of the objective function $(c,x) \geqslant c_0$.

There are two cases: the objective function coefficients are adjusted or fixed.

Let us denote $\tilde{A} = \begin{pmatrix} A \\ -c \end{pmatrix}$, $\tilde{b} = \begin{pmatrix} b \\ -c_0 \end{pmatrix}$, $\tilde{H} = \begin{pmatrix} H \\ -h \end{pmatrix}$. We obtain the following formalization of the problems of minimal matrix correction of the infeasible LPP in standart form:

$$\inf_{x,\tilde{H}} \{\|\tilde{H}\|_\infty \colon (A+H)x \leqslant b,\ (c+h, x) \geqslant c_0,\ x \geqslant 0\}. \tag{6}$$

$$\inf_{x,H} \{\|H\|_\infty \colon (A+H)x \leqslant b,\ (c, x) \geqslant c_0,\ x \geqslant 0\}. \tag{7}$$

It can be proved that both problems are reduced to LPP.

1. If the primal and dual LPP

$$\max_{z_0,z} \left\{ \sum_{j=1}^{n} z_j \colon \tilde{A}z - z_0\tilde{b} \leqslant \mathbf{1}^{m+1},\ z, z_0 \geqslant 0 \right\},$$

$$\min_{y} \left\{ \sum_{i=1}^{m+1} y_i \colon (\tilde{b}, y) \leqslant 0,\ \tilde{A}^T y \geqslant \mathbf{1}^n,\ y \geqslant 0 \right\}$$

have optimal solutions $z_0^*, z^* = (z_1^*, \ldots, z_n^*)$ and $y^* = (y_1^*, \ldots, y_{m+1}^*)$, respectively, then

$$\inf_{x,\tilde{H}} \{\|\tilde{H}\|_\infty \colon (A+H)x \leqslant b,\ (c+h, x) \geqslant c_0,\ x \geqslant 0\} =$$

$$= \frac{1}{\sum\limits_{j=1}^{n} z_j^*} = \frac{1}{\sum\limits_{i=1}^{m+1} y_i^*}.$$

If $z_0^* > 0$, then the correction problem [6] also has an optimal solution

$$x^* = \frac{1}{z_0^*} z^*,\ \tilde{H}^* = \triangle\, b(x^*)^+, \triangle\, b_i = \begin{cases} 0, & \text{if } (\tilde{b} - \tilde{A}x^*)_i \geqslant 0, \\ \frac{1}{\sum\limits_{j=1}^{n} z_j^*}, & \text{otherwise.} \end{cases}$$

2. If the primal and dual LPP

$$\max_{z_0,z} \left\{ \sum_{j=1}^{n} z_j \colon \tilde{A}z - z_0\tilde{b} \leqslant \begin{pmatrix} \mathbf{1}^m \\ 0 \end{pmatrix},\ z, z_0 \geqslant 0 \right\},$$

$$\min_{y}\left\{\sum_{i=1}^{m} y_i\colon (\tilde{b}, y) \leqslant 0,\ \tilde{A}^T y \geqslant \mathbf{1}^n,\ y \geqslant 0\right\}$$

have optimal solutions $z_0^*, z^* = (z_1^*, \ldots, z_n^*)$ and $y^* = (y_1^*, \ldots, y_{m+1}^*)$, respectively, then

$$\inf_{x,H}\{\|H\|_\infty\colon (A+H)x \leqslant b,\ (c,x) \geqslant c_0,\ x \geqslant 0\} = \frac{1}{\sum\limits_{j=1}^{n} z_j^*} = \frac{1}{\sum\limits_{i=1}^{m} y_i^*}.$$

If $z_0^* > 0$, then the correction problem [7] also has an optimal solution

$$x^* = \frac{1}{z_0^*} z^*,\ H^* = \triangle b (x^*)^+, \triangle b_i = \begin{cases} 0, & \text{if } (b - x^*)_i \geqslant 0, \\ \frac{1}{\sum\limits_{j=1}^{n} z_j^*}, & \text{otherwise.} \end{cases}$$

References

1. Vatolin A.A. Correction of the Augmented Matrix of an Inconsistent System of Linear Inequalities and Equations // Mathematical Methods of Optimization in Economical-Mathematic Modeling. Moscow: Nauka, 1991. P. 240–249 [in Russian].
2. Gorelik V. A. Ibatullin R. R. Correction of a System of Constraints of a Linear Programming Problem with a Minimax Constraint // Modeling, Decomposition, and Optimization of Complex Dynamic Processes. Dorodnitsyn Computing Center of the Russian Academy of Sciences, Moscow, 2001. P.89–107 [in Russian].
3. Gorelik V. A., Erokhin V. I., Pechenkin R. V. Numerical Methods for the Correction of Improper Linear Programming Problems and Structured Systems of Equations // Dorodnitsyn Computing Center of the Russian Academy of Sciences, Moscow, 2006 [in Russian].
4. Gorelik V. A., Erokhin V. I., Pechenkin R. V. Minimax matrix correction of inconsisent systems of linear algebraic equations with block matrices of coefficients // Journal of Computer and Systems Sciences International. 2006. V. 45, N. 5. P. 727–737.
5. Murav'eva O. V. Robustness and correction of linear models // Automation and Remote Control. 2011. V. 72, N. 3. P. 556–569.

6. Barkalova O. S. Correction of improper linear programming promblem in canonical form by applying the minimax criterion // Computational Mathematics and Mathematical Physics. 2012 V. 52, N. 12, P. 1624–1634.
7. Murav'eva O. V. Studying the stability of solution to system of linear inequalities and constructing separating hyperplanes // Journal of Applied and Industrial Mathematics. 2014 V. 8, N. 3, P. 349–356.

Solution of the approximate system of linear algebraic equations minimal with respect to the ℓ_∞ norm*

V.V. Volkov[1], V.I. Erokhin[2], A.Yu. Onufrei[2], V.V. Kakaev[2], and A.P. Kadochnikov[2]

[1]Borisoglebsk branch of Voronezh State University, Borisoglebsk, Russia, [2]Mozhaisky Military Space Academy, St. Petersburg, Russia

The paper is focused on the problem of finding solutions of approximate systems of linear algebraic equations (SLAE). The generalization of the A. N. Tikhonov's method to non-Euclidean vector norms and considered. As a criterion for optimality of the solution in the resulting mathematical programming problems are used minimum of the Holder norm with exponent $p = \infty$.

Approximate SLAE can arise in a large number of applied problems of science and technology (see, for example, [1–3]).

We consider the exact solvable system of linear algebraic equations $A_0x = b_0$, where $A_0 \in \mathbb{R}^{m\times n}$, $b_0 \in \mathbb{R}^m$, $b_0 \neq 0$, the relationship between the dimensions of the matrix A_0 and its rank is not specified, $x_0 \in \mathbb{R}^n$ is the solution of this system with minimal Holder norm with exponent $p = \infty$. Numerical values of A_0, b_0 and x_0 are unknown, instead of them, approximate matrix $A \in \mathbb{R}^{m\times n}$ and vector $b \in \mathbb{R}^m$, $b \neq 0$, are given, such that $\|A_0 - A\|_{\infty,\psi} \leqslant \mu$, $\psi(b_0 - b) \leqslant \delta < \psi(b)$, where $\mu \geqslant 0$ and $\delta \geqslant 0$ are known parameters, $\psi(\cdot)$ s arbitrary vector norm, $\|\cdot\|_{\infty,\psi}$ is matrix norm, such that $\|A\|_{\infty,\psi} := \max\limits_{x\neq 0} \frac{\psi(Ax)}{\|x\|_\infty}$. The matrix A is not obliged to have a full rank and the compatibility of the system $Ax = b$ are not assumed in the general case.

*The reported study was funded by RFBR according to the research project 18-31-00083

To find matrix $A_1 \in \mathbb{R}^{m\times n}$ and vectors $b_1 \in \mathbb{R}^m$, $x_1 \in \mathbb{R}^n$ such that $\|A - A_1\|_{\infty,\psi} \leqslant \mu$, $\psi(b - b_1) \leqslant \delta$, $A_1x_1 = b_1$, $\|x_1\|_\infty \to \min$.

We shall denote this problem by $Z_{\infty,\psi}\,(\mu,\delta)$.

The problem $Z_{1,\psi}\,(\mu,\delta)$ is the modification (generalization) of the problem considered by A. N. Tikhonov [4] and called by him "regularized least squares method" (RLS). The problem in the original formulation (using Euclidean matrix and vector norms) are explored, for example, in papers [5, 6].

We note that in [7] the problem of finding a normal solution of regularized problems of SLAU is also solved, but with the use of another technique: pairs of mutually dual problems of conditional optimization are used.

This article is based on the results presented in [8].

Let us proceed from problem $Z_{\infty,\psi}\,(\mu,\delta)$ to the equivalent problem $R_{\infty,\psi}(\mu,\delta) : \|x\|_\infty \to \min\limits_{\psi(b-Ax)=\mu\cdot\|x\|_\infty+\delta} (=: \chi_{\infty,\psi})$.

When in the problem $Z_{1,\psi}\,(\mu,\delta)$ the Holder norm with exponent $p = 1,\infty$ is chosen as $\psi(\cdot)$, we obtain the following generalized problems of RLS and their reductions to the mathematical programming problems [8] (the symbol "$\mapsto$" means "is reduced to"):

$$Z_{\infty,1}\,(\mu,\delta) \mapsto R_{\infty,1}(\mu,\delta) : \|b - Ax\|_1 \leqslant \mu \cdot \|x\|_\infty + \delta, \|x\|_\infty \to \min, \quad (1)$$

$$Z_{\infty,\infty}\,(\mu,\delta) \mapsto R_{\infty,\infty}(\mu,\delta) : \|b - Ax\|_\infty \leqslant \mu\cdot\|x\|_\infty+\delta, \|x\|_\infty\to \min\,. \quad (2)$$

The problems (1)–(2) can be reduced to the set of $2n$ linear programming (LP) problems

Consider the problem $Z_{\infty,1}\,(\mu,\delta)$:

Let the matrix $A \in \mathbb{R}^{m\times n}$ and the vector $b \in \mathbb{R}^m$, $b \neq 0$ are known. It is required to find $A_1 \in \mathbb{R}^{m\times n}$, $x_1 \in \mathbb{R}^n$, $b_1 \in \mathbb{R}^m$ such that $\|A - A_1\|_1 \leqslant \mu$, $\|b - b_1\|_1 \leqslant \delta$, $A_1x_1 = b_1$, $\|x_1\|_\infty \to \min$, where $\mu, \delta \geqslant 0$ are known a priori, simultaneously non-zero constants.

Theorem 1. Problem $Z_{\infty,1}\,(\mu,\delta)$ have a solution if and only if mathematical programming problem $R_{\infty,1}(\mu,\delta)$ have a solution.

Consider set of $2n$ linear programming problems, generated by the enumeration of two parameters: index $j = 1, 2, ..., n$ (external level) and scalar $z = -1, 1$ (internal level):

$$\begin{gathered} -p \leqslant b - Ax \leqslant p, \\ -\theta \cdot 1_n \leqslant x \leqslant \theta \cdot 1_n, \\ x_j = z \cdot \theta, \\ 1_m^\top p \leqslant \mu\theta + \delta, \\ p \geqslant 0, \quad \theta \geqslant 0, \quad \theta \to \min\,. \end{gathered} \quad (3)$$

Theorem 2. If problem $R_{\infty,1}(\mu,\delta)$ have a solution, it can be found as $x^* \in \operatorname{Argmin}\left\{\left\|x^1\right\|_\infty, ..., \left\|x^k\right\|_\infty, ..., \left\|x^K\right\|_\infty\right\}$, where $K \leqslant 2n$ is the number of solvable LP problems of the form (3), x^k is the solution of solvable LP problem of the form (3) with index k, $1_n \in \mathbb{R}^n$ and $1_m \in \mathbb{R}^m$ are unit vectors, $p \in \mathbb{R}^m$, $\theta \in \mathbb{R}$.

Reasoning similarly, consider the problem $Z_{\infty,\infty}(\mu,\delta)$.

Theorem 3. Problem $Z_{\infty,\infty}(\mu,\delta)$ have a solution if and only if mathematical programming problem $R_{\infty,\infty}(\mu,\delta)$ have a solution.

Consider set of $2n$ linear programming problems, generated by the enumeration of two parameters: index $j = 1, 2, ..., n$ (external level) and scalar $z = -1, 1$ (internal level):

$$\begin{gathered} -\pi \cdot 1_m \leqslant b - Ax \leqslant \pi \cdot 1_m, \\ -\theta \cdot 1_n \leqslant x \leqslant \theta \cdot 1_n, \\ x_j = z \cdot \theta, \\ \pi \leqslant \mu\theta + \delta, \\ \pi \geqslant 0, \quad \theta \geqslant 0, \quad \theta \to \min. \end{gathered} \tag{4}$$

Theorem 4. If problem $R_{\infty,\infty}(\mu,\delta)$ have a solution, it can be found as $x^* \in \operatorname{Argmin}\left\{\left\|x^1\right\|_\infty, ..., \left\|x^k\right\|_\infty, ..., \left\|x^K\right\|_\infty\right\}$, where $K \leqslant 2n$ is the number of solvable LP problems of the form (4), x^k is the solution of solvable LP problem of the form (4) with index k, $1_n \in \mathbb{R}^n$ and $1_m \in \mathbb{R}^m$ are unit vectors, $\pi, \theta \in \mathbb{R}$.

Computational experiments

There are the results of a numerical solution of model systems for problems (1) and (2). Calculations are performed using Matlab®. The corresponding auxiliary LP problems were solved by the simplex method.

A series of problems of the form (1) (Fig. 1) and the form (2) (Fig. 2) with a decreasing error was considered. For each of these problems, solutions were found by three methods. The results of the computational experiments are presented below by dependences of the errors (estimated on Euclidean norm) of the solution $\varepsilon_{RLN(\infty,1)_i} = \|x_{RLN(\infty,1)_i} - x_0\|$ (Fig. 1 only), $\varepsilon_{RLN(\infty,\infty)_i} = \|x_{RLN(\infty,\infty)_i} - x_0\|$ (Fig. 2 only), $\varepsilon_{RLS_i} = \|x_{RLS_i} - x_0\|$ and $\varepsilon_{LS_i} = \|x_{LS_i} - x_0\|$, on the error parameter e_i. Here $x_{RLN(\infty,1)}$ is the solution of the problem (1) (RLN – Regularized Least Norm, regularized solution using the norms ℓ_∞ and ℓ_1), $x_{RLN(\infty,\infty)}$ is the solution of the problem (2) (regularized solution using the norm ℓ_∞), x_{LS} is the solution by least squares method, and x_{RLS} is the solution by regularized least squares method.

The graphs shown in Figs. 1 and 2 show that there are such approximate SLAE $(A + \Delta A_i)x = b + \Delta b_i$ (where $\Delta A_i = e_i \Delta A$, $\Delta b_i = e_i \Delta b$,

$e_i = 10^{-1-0.5\cdot i}$) for which the solution $x_{RLN(\infty,1)}$ and $x_{RLN(\infty,\infty)}$ is closer (according to the Euclidean norm) to the solution of the "exact" SLAE than the solutions obtained by LS and RLS methods for any level of error.

In addition, the graphs show that all the considered methods are stable: with a decrease in the amount of error imposed on the original matrix, the error decreases.

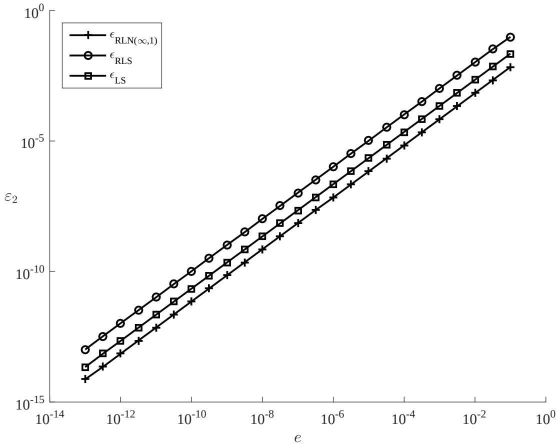


Fig. 1. The result of the first computational experiment (problem $Z_{\infty,1}(\mu,\delta)$).

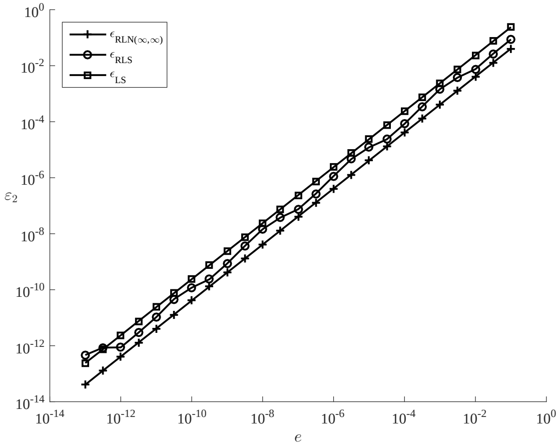


Fig. 2. he result of the second computational experiment (problem $Z_{\infty,\infty}\,(\mu,\delta)$).

References

1. Erokhin V.I., Volkov V.V. Methods and models of recovering linear dependencies from inaccurate information // Izvestija Sankt Peterburgskogo Gosudarstvennogo Technologicheskogo Instituta (Technicheskogo Universiteta)/ 2011. N 10. P. 52–57. (In Russian)
2. Erokhin V.I. et al. Using negative regularization parameter in Tikhonov's regularized least squares method // Izvestija Sankt Peterburgskogo Gosudarstvennogo Technologicheskogo Instituta (Technicheskogo Universiteta), 2014. N 24(50). P. 86–92. (In Russian)
3. Erokhin V.I., Volkov V.V. Recovering images, registered by device with inexact point-spread function, using tikhonov's regularized least squares method // Int. Journal of Artificial Intelligence. 2015. V. 13, N 1. 12 p. Available at: `http://www.ceser.in/ceserp/index.php/ijai/article/view/3531`. (accessed: 21.03.2018)
4. Tikhonov A.N. Approximate systems of linear algebraic equations // USSR Comput. Math. Math. Phys. 1980. V. 20, N 6, P. 10–22.
5. Volkov V.V., Erokhin V.I. Tikhonov solutions of approximately given systems of linear algebraic equations under finite perturbations of their matrices // Comput. Math. Math. Phys. 2010. V. 50,

N 4. P. 589–605.

6. Erokhin V.I., Volkov V.V. About A. N. Tikhonov's regularized least squares method // Vestnik Sankt-Peterburgskogo Universiteta, Prikladnaya Matematika, Informatika, Protsessy Upravleniya, 2017. Issue 1. P. 4–16. (In Russian)
7. Golikov A.I., Evtushenko Y.G. Regularization and normal solutions of systems of linear equations and inequations // Proc. of the Steklov Institute of Mathematics. 2015. V. 289, N 1. P. 102–110.
8. Volkov V.V. et al. Generalizations of Tikhonov's regularized method of least squares to non-Euclidean vector norms // Comput. Math. Math. Phys. 2017. V. 57, N 9. P. 1416–1426.

OR in economics

The model of privatization of a unitary enterprise[*]

V.I. Arkin and A.D. Slastnikov

Central Economics and Mathematics Institute, Moscow, Russia

Privatization in Russia, which began in the 1990s, generated a number of problems concerning its effectiveness, forms and methods of conducting (see, for example, [1]). There is no common opinion, what kind of an ownership is more effective: state (public) or private. Many researches support the proposition that privately owned firms are more efficient and more profitable than state-owned ones. On the other hand, privatization can also lead to certain problems, in particular, resulting in an increase in prices of goods and services, corruption etc. In this paper we don't discuss all pros and cons of privatization, but focus on its optimization (from the state's perspective), when a decision on privatization has already been made.

The methodology of this research is based on real options theory (see, e.g. [2]). The situation of partial privatization was studied in [3], where authors using real options approach derived the optimal strategy for private investor to enter a public sector and the optimal degree of privatization.

1. *The model.* Let us consider a unitary enterprise, i.e. a state-owned enterprise with indivisible assets, which may not be distributed among the agents in any way.

Due to budgetary limitations, the state wants to sell this enterprise to the private investor on the certain conditions. The state assigns a

[*] This research is supported by Russian Foundation for Basic Researches (project No. 18-010-00666)

certain price for the privatization transaction (sale price) and burdens the potential buyer with the obligation to upgrade the enterprise and make it more efficient (that requires additional costs from the buyer). The aim of the privatization process is to bring and optimize additional revenues to the state budget.

Let π_t^1, $t \geqslant 0$ be the cash-flow from state-owned enterprise at time t, and π_t^2 be the cash-flow from this enterprise after privatization. We consider π_t^1 and π_t^2 as a stochastic processes, defined at some probability space with filtration $(\Omega, \mathcal{F}, \{\mathcal{F}_t, t \geqslant 0\}, \mathrm{P})$.

Assume that private investor buys the enterprise at the time τ. Then his expected net present value (NPV) from this transaction is :

$$N(\tau) = \mathrm{E}\left(\int_\tau^\infty (1-\gamma)\pi_t^2 e^{-\rho t} dt - (P+M)e^{-\rho\tau}\right), \tag{1}$$

where P is the privatization price, M stands for the enterprise upgrading cost, γ means the tax burden rate (i.e. a part of total taxes in cash-flow), and ρ is the discount rate.

The benefits from the privatization (at the time τ) for the state are evaluated with the expected discounted budgetary effect $B(\tau, P)$ that consists of:

1) the payments into the budget from state-owned enterprise that are the assigned proportion θ, $0 < \theta \leqslant 1$ of the cash-flow, i.e. $\theta\pi_t^1$ (before the time τ);

2) the taxes from the private enterprise $\gamma\pi_t^2$ (after τ); and

3) the privatization price P (at the time τ).

More precisely,

$$B(\tau, P) = \mathrm{E}\left(\int_0^\tau \theta\pi_t^1 e^{-\rho t} dt + \int_\tau^\infty \gamma\pi_t^2 e^{-\rho t} dt + P e^{-\rho\tau}\right). \tag{2}$$

2. *The problem.* It is assumed that private investor has an opportunity to choose the privatization time τ and he follows the principle of maximal NPV :

$$N(\tau) \to \max_\tau, \tag{3}$$

where maximum is taken over all stopping times (Markov moments) τ.

The state wants to optimize his benefits from selling the enterprise, and puts the privatization price P that maximizes the budgetary effect (2) under the optimal behavior of private investor:

$$B(\tau^*, P) \to \max_P, \tag{4}$$

where maximum is taken over all acceptable prices P, and τ^* is a solution to the investor problem (3) (τ^*, of course, depends on P).

The problem (3)–(4) can be considered as Stackelberg equilibrium in the "privatization game" between the state and private investor.

3. *Mathematical assumptions.* The cash-flows from enterprise before and after the privatization (π_t^1 and π_t^2) are modeled as geometric Brownian motions with parameters (α_1, σ_1) and (α_2, σ_2) resp. The starting point for the cash-flow π_t^2 at the time of privatization τ is connected with the final point of cash-flow before privatization by the relation: $\pi_\tau^2 = k\pi_\tau^1$, where k represents the given "upgrading coefficient".

4. *Optimal behavior of the private investor.* For a given privatization price P the optimal time for privatization (i.e. a solution to the problem (3)) is the following:

$$\tau^* = \inf\left\{t \geqslant 0 : \pi_t^2 \geqslant \pi^* = \frac{\beta}{\beta-1}\cdot\frac{\rho-\alpha_2}{1-\gamma}(P+M)\right\},$$

where β is the positive root of the equation $\frac{1}{2}\sigma_2^2\beta(\beta-1)+\alpha_2\beta-\rho=0$.

5. *Optimal privatization price.* Let us denote

$$d = \frac{1}{1-\gamma}\left(1+\frac{\gamma}{\beta-1}-\frac{\beta}{\beta-1}\cdot\frac{\rho-\alpha_2}{\rho-\alpha_1}\cdot\frac{\theta}{k}\right).$$

If $d \leqslant 0$, then there is no optimal privatization price, and the budgetary effect $B(\tau^*, P)$ for any price $P \geqslant 0$ will be less than the budgetary effect from the enterprise *without* privatization.

In the case when $d > 0$ the optimal privatization price (solving the problem (4) over all $P \geqslant 0$) exists and is represented by the following formula:

$$P^* = \max\left\{0, \left(\frac{\beta}{\beta-1}\cdot\frac{1}{d}-1\right)M\right\}.$$

References

1. Polterovich V.M. Privatization and the rational ownership structure. Part 1. Privatization: the problem of efficiency // Economic science of modern Russia. 2012. N 4(59). P. 7–23 (in Russian).
2. Dixit A.K. and Pindyck R.S. Investment under Uncertainty. Princeton: Princeton University Press, 1994.

3. Chavanasporn W. and Ewald C.-O. Privatization of businesses and flexible investment: a real option approach // Decisions in Economics and Finance. 2012. V. 35. N 1. P. 75–89.

Dynamic model of the control process of the balance of the joint–distribution pension system

P.V. Kalashnikov
Far East Federal University, Vladivostok, Russia

Ensuring a decent standard of living for the older generation is a priority objective of the social and economic policies of most modern countries. At present, pension reform is a subject of wide discussion. Its tremendous social significance is obvious and is the foundation for the effective progressive development of society and the active involvement of citizens in the economic life of the country throughout the entire period of work.

The object of the study is the pension system of the Russian Federation, considered in the context of modeling the balance of incoming insurance contributions and payments of old-age labor pension on general grounds, which is the most massive type of pension benefits.

The subject of the study is an assessment of the main indicators of the budget of the Pension Fund of the Russian Federation in terms of the amount of contributions and liabilities used to pay the insurance part of the old-age labor pension on general grounds.

The goals of the study are: to build a model for the formation of contributions to the Pension Fund of the Russian Federation (PFR), as well as to estimate the size of the PFR deficit with the existing structure of the population and the legislative framework.

The state of the pension system is currently heavily influenced by the unfavorable demographic situation, which is reflected in the increase in the number of older people and in the growth of their relative share in the total population. According to the Federal State Statistics Service [1] at present, 35 555 thousand people receive a retirement pension, which is about 24 % of the total popula tion, and the total number of pensioners receiving various types of pensions exceeds 42 million people. This fact causes a steady increase in the level of demographic burden on the able-bodied population by the elderly [2]

The task of actuarial evaluation of the joint-distribution pension system involves the analysis of demographic, socio-economic, and institutional parameters (pension legislation) [3] .Each of these groups of parameters, in turn, is divided into a wide range of subtasks, the state of the pension system as a whole, and also to develop mechanisms for its effective functioning in the short and long term.

The model of the Russian pension system constructed in the course of the study includes the following basic elements: the demographic component, the unit for calculating the amount of contributions and obligations of the Pension Fund for the payment of the insurance part of the old-age labor pension on general grounds, the unit for analyzing the effect on the results of calculations of control actions on the parameters of the actuarial basis, as well as a block of study of the stability of the model to change the values of the basic calculated quantities and the presence of an error in the initial data.

The forecast of the level of balance of the budget of the PFR is presented in Figure 1. When analyzing the possible control actions aimed at reducing the budget deficit of the PFR in the long term, the following are considered: an increase in the birth rate per 100,000 people per year, an increase in the migration increase to 500,000 people per year and the number of legally working temporary migrants in the territory up to 2.5 million a person annually, raising the retirement age for the male and female population to 65, as well as the complete abolition of the funded part of the old-age pension for the population under the age of 1967. birth and contribution only on the insurance component of the pension benefit in question. Such a change in the parameters of the actuarial

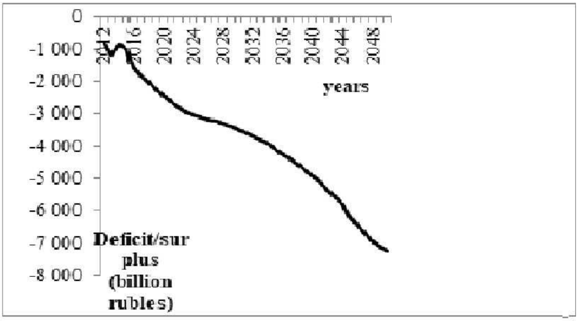


Fig. 1. Forecast of the level of balance of insurance contributions and payments to the Pension Fund of the old-age Pension Fund on general grounds in the long-term period

basis is based on an analysis of the dynamics of the variables under consideration in the period preceding the base year of calculations, the study of long-term forecasts of the socio-economic development of the Russian Federation prepared by the Federal Service for State Statistics, as well as studying the world experience in reforming the pension system at the state level.

The calculation of the level of balance of the budget of the PFR for each of the scenarios for changing the parameters of the actuarial basis is presented in Figure 2.

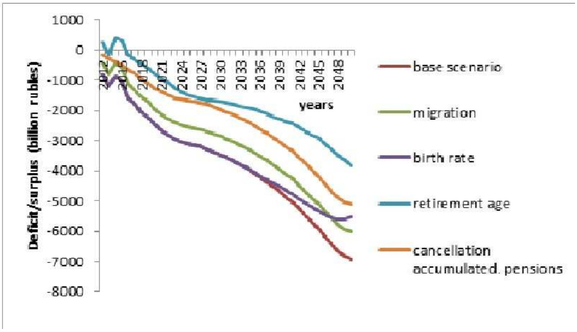


Fig. 2. The level of balance of the budget of the PFR for various options for socio-economic policy

References

1. The number of pensioners and the average size of designated pensions by types of pensions and categories of pensioners [Electronic resource]: data Federal Service of State Statistics of the Russian Federation. URL: `http://www.gks.ru/free_doc/new_site/population/urov/urov_p2.htm`
2. Smirnov I.V. Demography: A Training Manual.Kaluga: a branch of North-West Academy of Public Administration in Kaluga, 2004.
3. Simonenko V.N. Scenario of the pension system modeling in the context of the current state of the economy // Bulletin of the Pacific State University. 2010. V. 4, N 19. P. 145–152.

Example calculation of exotic options in incomplete $\{1, S\}$–market

E.A. Shelemekh
CEMI RAS, Moscow, Russia

1. The report presents new example calculations of exotic options in incomplete market generated by Markov chain.

Option is a financial contract between the Seller and the Buyer. It establishes the Buyer's right to receive payment in the future (amount of payment depends on risk asset's price in agreed moment in the future) or his/her right to buy/exchange risk assets on conditions specified in the contract. It is the Seller's obligation to make the payment or to sell/exchange risk asset according to the contract. To purchase above stated right the Buyer pays the Seller some fee also called value of the option or the option's premium.

There are a few approaches to option's fair value conception. There in the report it is assumed that the market is a set of fisk-free and risk assets. For each asset dynamics of it's price is specified by a random sequence. The set of martingale measures equivalent to distribution of risk asset's prise is supposed to contain more then one element, i.e. the market is incomplete. Under this approach supremum in equivalent martingale measures of expected value for the Seller's obligation at the moment of execution is stated to be the option's fair value [1]. Note that in most cases it is quite a problem to find the supremum.

In the report we provide solution examples for binary, barrier options and European option on maximum prise of risk asset in incomplete market generated by Markov chain. Note that by now examples for exotic options are known only for the cases of complete market (see [2]–[3]).

2. Statement of the problem. Suppose that on filtered stochastic basis $(\Omega, \mathcal{F}, (\mathcal{F}_n)_{n\geq 0}, \mathsf{P})$ a random sequence $(S_n, \mathcal{F}_n)_{n\geq 0}$ is set. It specifies dynamics of the risk asset prise. Let $\mathcal{F}_n \triangleq \sigma\{S_1, ..., S_n\}$, $n \geq 1$.

Condition 1. *Random sequence $\{S_n\}_{n\geq 0}$ satisfies recurrent relation $S_n = S_{n-1}\lambda^{\varepsilon_n}$, $S_n|_{n=0} = \lambda^{\varepsilon_0}$, where ε_0 is an integer, $\{\varepsilon_n\}_{n\geq 1}$ is a sequence of independent identically distributed random variables taking values in $\{-1, 0, 1\}$ with positive probabilities, $\lambda > 1$ is a given constant.*

It is well known that random sequence $\{S_n\}_{n\geq 0}$ under condition 1 specifies an incomplete $\{1, S\}$–market [1]. In spite of it's simplicity the market allows stability of risk asset's price as well as increase and decrease. So our model seems to be relevant.

Definitions and designations (according to [1]). Let:

1) $S_0^n \triangleq (S_0, ..., S_n)$;

2) $\mathfrak{R}$ be the set of equivalent martingale measures;

3) τ be a stopping moment for random sequence $\{S_n\}_{n\geq 0}$, taking values in $\{0, ..., N\}$, where N is a positive integer;

4) $1_{\{\tau=n\}}$ is an indicator of random event $\{\tau = n\}$, f_n is a bounded $\mathcal{F}_n$–measurable random variable, thus $\left\{1_{\{\tau=n\}} f_n\right\}_{0\leq n\leq N}$ is an exotic option's dynamic payoff;

5) $(\beta_n, \mathcal{F}_n)_{n\geq 0}$ and $(\gamma_n, \mathcal{F}_n)_{n\geq 0}$ are predictable sequences specifying number of risk-free and risk assets in dynamics, respectively. The set $\pi \triangleq (\beta, \gamma)$ is called portfolio, while $X_n^\pi \triangleq \beta_n + \gamma_n S_n$ is a capital of portfolio π. Portfolio π is said to be self-financing if for any n probability $\mathsf{P}(\triangle\beta_n = -\triangle\gamma_n S_{n-1}) = 1$. A non-decreasing sequence $(C_n, \mathcal{F}_n)_{n\geq 0}$, $C_n|_{n=0} = 0$, is called consumption, a pair (π, C) is a portfolio with consumption. Capital of portfolio with consumption (π, C) is defined by equality $X_n^{(\pi,C)} \triangleq X_n^\pi - C_n$, $0 \leq n \leq N$.

Definition. They say, that self-financing portfolio with consumption (π^*, C^*) is a superhedging portfolio of exotic option $\left\{1_{\{\tau=n\}} f_n\right\}_{0\leq n\leq N}$, if $\mathsf{P}\left(X_\tau^{(\pi^*,C^*)} \geq f_\tau\right) = 1$.

Definition. Portfolio with consumption (π^*, C^*) is a minimal one for given exotic option, if for any other superhedging portfolio with consumption (π, C) we have $\mathsf{P}\left(X_\tau^{(\pi^*,C^*)} \leq X_\tau^{(\pi,C)}\right) = 1$.

Under the approach used in the report to calculate exotic option means to find minimal superhedging portfolio with consumption (π^*, C^*) and fair value of the option, i.e. $X_0^{(\pi^*,C^*)}$.

3. General solution of the problem.

Theorem 1. *Suppose condition 1 is satisfied, there is exotic option with dynamic payoff* $\left\{1_{\{\tau=n\}} f_n\right\}_{0\leq n\leq N}$ *and* (π^*, C^*) *is a portfolio with consumption such, that for any* $n \in \{1, ..., N\}$*:*

1) *portfolio's capital* $\left\{X_n^{(\pi^*,C^*)}\right\}_{0\leq n\leq N}$ *satisfies recurrent relation*

$$\begin{aligned}
&X_n^{(\pi^*,C^*)}(S_0^n) = \\
&1_{\{\tau=n\}} f_n + 1_{\{\tau>n\}} \sup_{\mathsf{Q}\in\mathfrak{R}} \mathsf{E}^{\mathsf{Q}} X_{n+1}^{(\pi^*,C^*)}(S_0^n, S_n\lambda^{\varepsilon_{n+1}}) = \\
&= 1_{\{\tau=n\}} f_n + 1_{\{\tau>n\}} \max\Big\{X_{n+1}^{(\pi^*,C^*)}(S_0^n, S_n)\,; \\
&\tfrac{\lambda}{1+\lambda} X_{n+1}^{(\pi^*,C^*)}\left(S_0^n, S_n\lambda^{-1}\right) + \tfrac{1}{1+\lambda} X_{n+1}^{(\pi^*,C^*)}(S_0^n, S_n\lambda)\Big\}, \\
&X_n^{(\pi^*,C^*)}(S_0^n)\,|_{n=N} = 1_{\{\tau=N\}} f_N;
\end{aligned} \tag{5}$$

2) $\gamma_0^* = 0$, $\gamma_n^*\left(S_0^{n-1}\right) =$

$$= \frac{\lambda}{(\lambda^2-1)S_{n-1}}\left[X_n^{(\pi^*,C^*)}\left(S_0^{n-1}, S_{n-1}\lambda\right) - X_n^{(\pi^*,C^*)}\left(S_0^{n-1}, S_{n-1}\lambda^{-1}\right)\right]; \tag{6}$$

3) $\triangle\beta_n^* = -\triangle\gamma_n^* S_{n-1}$, $\beta_0^* = X_0^{(\pi^*,C^*)}(S_0)$;

4) $\triangle C_n^* = \gamma_n^*\triangle S_n - \triangle X_n^{(\pi^*,C^*)}\left(S_0^{n-1}, \lambda^{\varepsilon_n}S_{n-1}\right)$, $C_0^* = 0$.

Then portfolio (π^*, C^*) is a minimal superhedging portfolio of above stated exotic option in incomplete $\{1, S\}$–market.

4. Example calculation of exotic options. From Theorem 1 it follows, that in incomplete $\{1, S\}$–market generated by condition 1 solution of exotic option's problem is to be achieved by solution of equation (5). In the report the author presents solutions of (5) for for binary, barrier options and European option on maximum prise of risk asset.

4.1. *Binary option Up And Out* is a contract for the Seller to get from the Buyer one dollar at the moment, when risk asset's price exceeds some given value λ^m, if this takes place before the moment N inclusively (m is an integer). So the dynamic payoff of binary option Up And Out takes form $\{1_{\{\tau=n\}}\}_{0\le n\le N}$, where $\tau = \min\{0 \le n \le N : S_n \ge \lambda^m\}$ [3].

Let us denote:

1) $Bi(n,k,p) \triangleq \sum\limits_{i=0}^{k} \frac{n!}{k!(n-k)!}p^i(1-p)^{n-i}$, where $p \in [0,1]$ and $n \in \{0,...,N\}$, $k \in \{0,...,n\}$ is the Binomial distribution;

2) $[\cdot]$ is an integer part of a real number;

3) $A(t) \triangleq [0,5(N-n+\log_\lambda S_n+t)]$, $a(t) \triangleq [0,5(N-n-\log_\lambda S_n+t)]$.

Theorem 2. *Suppose condition 1 is satisfied and dynamic payoff of exotic option is $\{1_{\{\tau=n\}}\}_{0\le n\le N}$, where $\tau = \min\{0 \le n \le N : S_n \ge \lambda^m\}$. Then for any $n \in \{0,...,N\}$ capital of minimal superhedging portfolio may be submitted in the form:*

$$X_n^{(\pi^*,C^*)} = 1_{\{S_n\ge\lambda^m\}} + 1_{\{S_n<\lambda^m\}}\Big[Bi\left(N-n, A(-m), \tfrac{\lambda}{1+\lambda}\right) + \\ +S_n\lambda^{-m}\left\{1 - Bi\left(N-n, a(m), \tfrac{\lambda}{1+\lambda}\right)\right\}\Big],$$

$$X_n^{(\pi^*,C^*)}|_{n=N} = 1_{\{S_N\ge\lambda^m\}}.$$

4.2. *Barrier option* is a combination of binary and vanilla European option, namely: if the event fixed in the contract has occurred, then at the moment N the Buyer receives payment equal to paymant according to corresponding vanilla option. So for the barrier option Up and In, $\tau = (N+1) \wedge \min\{n \geq 0 : S_n \geq \lambda^m\}$, $f_n = (S_n - \lambda^k)^+$, where $0 < k < m$ are constants [2].

Theorem 3. *Suppose condition 1 is satisfied and exotic option is specified by* $\tau = (N+1) \wedge \min\{n \geq 0 : S_n \geq \lambda^m\}$, $f_n = (S_n - \lambda^k)^+$, $0 < k < m$ *are constants. Then for any* $n \in \{0, ..., N\}$ *capital of minimal superhedging portfolio may be submitted in the form:*

$$\begin{aligned}
&X_n^{(\pi^*,C^*)} = \\
&= 1_{\{S_n \geq \lambda^m\}} \Big\{ S_n Bi\left(N-n, A(-k), \tfrac{1}{1+\lambda}\right) - \\
&- \lambda^k Bi\left(N-n, A(-k), \tfrac{\lambda}{1+\lambda}\right)\Big\} + \\
&+ 1_{\{S_n < \lambda^m\}} \Big\{ S_n Bi\left(N-n, A(-m), \tfrac{1}{1+\lambda}\right) = \\
&- \lambda^k Bi\left(N-n, A(-m), \tfrac{\lambda}{1+\lambda}\right) + \\
&+ \lambda^m \left[Bi\left(N-n, a(2m-k), \tfrac{1}{1+\lambda}\right) - Bi\left(N-n, a(m), \tfrac{1}{1+\lambda}\right)\right] - \\
&- S_n \lambda^{k-m} \left[Bi\left(N-n, a(2m-k), \tfrac{\lambda}{1+\lambda}\right) - Bi\left(N-n, a(m), \tfrac{\lambda}{1+\lambda}\right)\right], \\
&X_n^{(\pi^*,C^*)}|_{n=N} = 1_{\{S_N \geq \lambda^m\}} (S_N - \lambda^k)^+.
\end{aligned}$$

An interesting fact follows from Theorems 2–3. For binary and barrier options for any $n \in \{0, ..., \tau\}$ number of risk assets γ_n^* in minimal superhedging portfolio calculated according to (6) is non-negative, i.e. minimal superhedging portfolio does not imply borrowing.

4.3. *European option on maximum prise of risk asset* is specified by $\tau = N$ and $f_n = \left(\max\limits_{0 \leq k \leq N} S_k - \lambda^m S_N\right)^+$, where $m \geq 0$ is a constant [2].

Theorem 4. *Suppose condition 1 is satisfied and exotic option is specified by* $\tau = N$, $f_n = \alpha^n \left(\max\limits_{0 \leq k \leq n} S_k - a S_n\right)^+$, *where constants* $a \geq 0$, $0 < \alpha \leq 1$. *Then for any* $n \in \{0, ..., N\}$ *capital of minimal superhedging portfolio may be submitted in the form:*

$$\begin{aligned}
&X_n^{(\pi^*,C^*)} = -\lambda^m S_n + \\
&+ \max_{0 \leq i \leq n} S_i \times \sum_{k=0}^{N-n} \sum_{j=0}^{(N-n-m-k-1)/2} \tfrac{N-n-2j}{N-n+1} \tfrac{(N-n)!}{j!(N-n-j)!} \left(\tfrac{\lambda}{1+\lambda}\right)^{N-n-j-1}.
\end{aligned}$$

5. Conclusion. An incomplete market generated by Markov chain was in consideration. In this market general form of solution for exotic option problem has been set. This solution includes recurrent equation of simplified form. The explicit solutions of the recurrent equation are presented for binary, barrier options and European option on maximum prise of risk asset. Most of above stated results and appropriate proofs were also presented in [4].

References

1. Shiryaev A.N. Probability. Moscow: MTsNMO, 2004 (in Russian).
2. Hull J.C. Options, Futures and other Derivatives. Upper Saddle River: Pearson Prentice Hall, 2009.
3. Föllmer H., Schied A. Stochastic Finance. An Introduction in Discrete Time. Berlin: Walter de Gruyter, 2004.
4. Shelemekh E.A. Calculationof exotic options in incomplete markets // Economica i matematicheskie metodi. 2017. 53:3. P. 78–92 (in Russian).

Markets and auctions analysis and design

The development of the electric power system, taking into account the balance reliability[*]

N. Aizenberg and S. Perzhabinskii
Melentiev Energy Systems Institute of SB RAS , Irkutsk, Russia

The level of reliability of electricity system depends on the amount of excess capacity that the generating companies agree to service, and on the network capacity. This determines the additional costs that affect the price of electricity. The consumer agree to pay these costs if he is interested in reliable electricity supply.

The problem of adequacy optimization of electric power systems consists in finding of optimal structure of generators and transmission lines for meeting electricity demand with random variations of load and failures of equipment.

In market conditions the problem should be solved taking into account different interests of economic agents such as generation and network companies and consumers [1]. This fact determines of changing of optimi- zation criteria. Instead of minimization of system costs we should maximize payoff function of each agent [2, 3]. Decision making of economic agents is based on results of adequacy estimation of electric power system. Effective method for such problem is developed earlier [4]. We modified the method for adequacy analysis of electric power system in market conditions. The method consists of four main blocks.

[*] This research is supported by grant RFBR 16-06-00071.

1. Determination of the equilibrium point (possible price and volume) of the electricity market based on the aggregate demand. This is a Cournot-type model with network restrictions.

2. Modeling of random states of electric power system. Each state is characterized of the set of random values such as load and available generator capacity in nodes, power line capacity (the initial approximation is defined in step 1).

3. The model of power shortage estimation of electric power system. Power shortage is estimated for each random state of electric power system. The model of power shortage estimation with quadratic power losses in power lines [4] is used in this stage. Taking account of quadratic losses guarantees uniqueness of deficit distribution over system nodes.

4. Computation of reliability indexes. Reliability indices are probability of no shortage system work, mathematical expectation of power shortage and electricity sacrifice of consumers, coefficient of availability of power. Optimal set of equipment in electric power system is defined as a result of comparison of variants of electric power systems development. That's way we are adding the next stage of computation.

5. Calculation of the generating company's profits, taking into account possible penalties for the electricity shortage. The calculation of prices as on base the equilibrium price obtained at stage 1 as on base the elasticity of demand for services quality depending on balance reliability indicators.

6. Comparison of reliability indices with companies' profit according to different configurations of electric power system. On this stage it is choosing of optimal variant of system development which balances reliability level necessary for consumer and profit margin of generation and network companies. Maintaining the integrity of the specifications.

The model of power shortage estimation

Let's consider the scheme of electric power system. The electric power systems scheme consists of n nodes and set of links between nodes. According to method for estimating of electric power systems adequacy it is necessary to simulate random states of electric power systems a many times. The simulations are occurred on a base of Monte Carlo method.

Let N is the given number of electric power systems states. Each state is characterized by set of means of random values such as generating capacity $\bar{x}_i^k$, load value $\bar{y}_i^k$ in the node i, line capacity $\overline{z}_{ij}^k$ between nodes i and j, $i = 1, \ldots, n$, $j = 1, ..., n$, $i \neq j$, $k = 1, ..., N$.

We use following problem for power shortage estimation of random states of electric power systems. The powerx_i and the load y_i in node i, power flow z_{ij} from node i to node j, $i = 1, ..., n$, $j = 1, ..., n$ are variables of the problem. The considered problem for some k, $k = 1, ..., N$, is

$$\sum_{i=1}^{n} y_i \to \max, \tag{1}$$

subject to the constraints

$$x_i - y_i + \sum_{j=1}^{n}(1 - a_{ji}z_{ji})z_{ji} - \sum_{j=1}^{n} z_{ij} = 0, i = 1, \ldots, n, i \neq j, \tag{2}$$

$$0 \leqslant y_i \leqslant \bar{y}_i^k, i = 1, \ldots, n, \tag{3}$$

$$0 \leqslant x_i \leqslant \bar{x}_i^k, i = 1, \ldots, n, \tag{4}$$

$$0 \leqslant z_{ij} \leqslant \bar{z}_{ij}^k, i = 1, \ldots, n, j = 1, ..., n, i \neq j, \tag{5}$$

where positive coefficients of power losses a_{ij} are given.

As a rule, the adequacy analysis of electric power system is realized for year. Every hour of work of electric power system is modeled. The failures of generators and power lines are used as random parameters. The repair time of equipment and fluctuations of load in the year are taking account in modeling. The rules of simulations of random values such as available capacity of generator and power line or load value are discussed in [3].

Let set of $\hat{x}_i^k$, $\hat{y}_i^k$, $\hat{z}_{ij}^k$ is optimal solution of the problem (1) – (5), $k = 1, ..., N$, $i = 1, \ldots, n$, $j = 1, ..., n$, $i \neq j$. The optimal value of power shortage in node i, $i = 1, \ldots, n$, is defined by the formula

$$d_i^k = \bar{y}_i^k - \hat{y}_i^k, k = 1, ..., N.$$

The state of electric power systems is deficit if the value

$$d^k = \sum_{i=1}^{n} d_i^k$$

is not equal to zero. This index is corresponded to index LOLP.

Mathematical expectation of power shortage in nodes of energy system is computed by next rule

$$MD_i = \sum_{j=1}^{H} \frac{d_i^j}{N}, i = 1, \ldots, n.$$

That's way mathematical expectation of power shortage in electric power systems is calculated by

$$MD = \sum_{i=1}^{n} MD_i.$$

Modeling the demand for electricity

Demand specification is accounting for different types of consumers. In this case, the price function is determined for each type of consumer, or, for each network node. This requires determination of the reliability indices r_i for the individual nodes $i = 1, \ldots, n$ or individual groups of nodes. Then the price that the consumer is willing to pay at node i can be represented in the form:

$$g_i(r_i, \tilde{y}_i) = (r_i)^m p_i(\tilde{y}_i). \tag{6}$$

where $p_i(\tilde{y}_i)$ is the maximum price that the consumer is willing to pay in the node i, $i = 1, \ldots, n$, without of power failures; m is the parameter that determine show much we take into account there liability factor, $m \in [0,\ 1]$ (the higher$m$, the more important for the consumer the qualitative power supply). The function $g(r, \tilde{y}_i)$ in the form (6) is concave, increasing with respect to r.

Modeling the behavior of generating companies

When looking for the Cournot-Nash equilibrium, it is necessary to solve the problem of maximizing profit for each generating company on residual demand. We take into account the conditions of positive and limited volumes of production. If the aggregate demand function is linear and the cost functions for all generating companies are quadratic, then the equilibrium exists and unique [5]. It determines the price of interaction without taking into account the reliability parameter. Based on this price and the reliability parameters obtained from the solution of the problem, we form the price for estimating the company's profit for the selected configuration of electricity system.

$$\pi_s = g_i(r_i, \tilde{y}_i)\sum_{i=1}^{n} x_{si} - C_s\left(\sum_{i=1}^{n} x_{si}\right) \to \max_{x_s}, \qquad s = 1, ..., S, \tag{7}$$

$$\sum_{i=1}^{n} y_i = \sum_{s=1}^{S}\sum_{i=1}^{n} x_{si}; \tag{8}$$

$$\sum_{i=1}^{n} x_{si} \leqslant \bar{x}_s, \quad x_{si} \geqslant 0, \qquad i = 1, ..., n, \qquad s = 1, ..., S. \tag{9}$$

where $C_s = \sum_{i=1}^{n} x_{si}$ is costs of the generating company s, $s = 1, ..., S$ depending on the reliability index r_i in each node i, $i = 1, ..., n$, power generation and capital costs of electricity generation. The cost function is increasing and convex. We form cost function on the basis of simulation modeling, comparing costs and balances reliability.

Problem (7)-(9) is solved by method of find by feeling to Cournot [5].

The reliability indexes are computed for each variant of development of electric power system. Profit of generating and network companies depends on meanings of reliability indexes. After analyzing all development variants companies will choose the reliability level which is optimal for them. They can provide this reliability level by inputting new generating and network equipment in the energy system. If all companies choose the same development variant then this is Nash equilibrium which may be not the best solution for the system. If the companies choose the different development variants then we accept a decision about an effective variant. Each company should have a positive profit and the reliability index should not be below some given level. The experimental research of the developed method is suggested on test scheme of electric power system which is constructed on the base of real technical data.

References

1. Aizenberg N., Perzhabinskii S. Method of adequacy optimization of electric power systems under market conditions // E3S Web of Conferences. 2017. V.25. P.02004.
2. Chuang A.S., Varaiya P. A game-theoretic model for generation expansion planning: problem formulation and numerical comparisons // IEEE Transactions on Power Systems, 2001. V.16 (4). P. 885-891.

3. Krupenev D., Perzhabinsky S. Algorithm for the Adequacy Discrete Optimization by Using Dual Estimates When Planning the Develop- ment of Electric Power Systems // Proceeding of The 17th Interna- tional Scientific Conference "'Electric power engineering"'. 2016. P.1-5.
4. Zorkaltsev V.I., Lebedeva L.M., Perzhabinsky S.M. Model for estimating power shortage in electric power systems with quadratic losses of power in transmission lines // Numerical Analysis and Applications. 2010. V. 3(3). P. 231-240.
5. Moulin E. Theory of games with examples from mathematical economics. Moscow: Mir, 1985.

Monopolistic competition with investments in *R&D*[*]

I.A. Bykadorov
Sobolev Institute of Mathematics SB RAS,
Novosibirsk State University,
Novosibirsk State University of Economics and Management,
Novosibirsk, Russia

We consider a monopolistic competition model with additive separable consumer's utility and the endogenous choice of technology. We study the impact of technological innovation on the equilibrium and socially optimal variables. We obtain the comparative statics of the equilibrium and socially optimal solutions with respect to the technological innovation parameter and utility level parameter. More precisely, we study a monopolistic competition model with endogenous choice of technology in the closed economy case. We consider "technological innovation" non-negative parameter α that influences on costs. Moreover, we consider "consumer utility level" non-negative parameter β that influences on utility. The aim is to make comparative statistics of equilibrium and social optimal solutions with respect to parameters α and β.

[*] The work was supported by the program of fundamental scientific researches of the SB RAS No I.5.1., project No 0314-2016-0018. Supported in part by RFBR grants 16-01-00108, 16-06-00101 and 18-010-00728. The author would like to thank Sergey Kokovin, Richard Ericson, Peter Neary, Kristian Behrens and Federico Etro for valuable comments, Jacques-Francois Thisse for guiding and advising, Evgeny Zhelobodko (25.09.1973-27.03.2013), Irina Antoshchenkova and Ekaterina Shelkova for participation in the early stages of this research.

Due to [1], the main assumptions of Monopolistic Competition are: consumers are identical, each endowed with one unit of labor; labor is the only production factor; consumption, output, prices etc. are measured in labor; firms are identical, but produce “varieties” (“almost the same”) of good; each firm produces one variety as a price-maker, but its demand is influenced by other varieties; each variety is produced by one firm that produces a single variety; each demand function results from additive utility function; number of firms is big enough to ignore firm’s influence on the whole industry/economy; free entry drives all profits to zero; labor supply/demand is balanced.

Our key findings are the following.

When parameter α increases,

– individual consumption x and individual investments f in *R&D* both increase;

– the behavior of the equilibrium and socially optimal variables does not depend on the properties of the costs as a function of investments f in *R&D*;

– the behavior of the equilibrium variables depends on the elasticity of demand only;

– the behavior of the socially optimal variables depends on the elasticity of sub-utility only;

– the equilibrium variables depend on the elasticity of demand and the socially optimal variables depend on the elasticity of utility in the identical way.

When parameter β increases,

– the behavior of the equilibrium individual investments f in *R&D*, individual consumption x, and mass of firms N depend on the behavior of the demand elasticity;

– the behavior of the social optimal individual investments f in *R&D*, individual consumption x, and mass of firms N depend on the behavior of the utility elasticity;

– the behavior of the equilibrium total investments Nf in *R&D* depends on the behavior of the elasticities of both demand and marginal costs;

– the behavior of the social optimal total investments Nf in *R&D* depends on the behavior of the elasticities of both sub-utility and marginal costs.

We study the impact of technological innovation on the equilibrium and socially optimal variables, namely, consumption, costs, the mass of firms and prices (in the equilibrium case). We obtain the comparative

Table 1. Equilibrium: Comparative statics w.r.t. α.

	$r'_u < 0$	$r'_u = 0$	$r'_u > 0$
$E_{x^*/\alpha}$	> 0	> 0	> 0
$E_{f^*/\alpha}$	> 0	> 0	> 0
$E_{N^*/\alpha}$	< 0	< 0	?
$E_{N^*f^*/\alpha}$	< 0	$= 0$	> 0
$E_{p^*/\alpha}$	< 0	< 0	< 0

Table 2. Social Optimality: Comparative statics w.r.t. α.

	$\varepsilon'_u > 0$	$\varepsilon'_u = 0$	$\varepsilon'_u < 0$
$E_{x^{opt}/\alpha}$	> 0	> 0	> 0
$E_{f^{opt}/\alpha}$	> 0	> 0	> 0
$E_{N^{opt}/\alpha}$	< 0	< 0	?
$E_{N^{opt}f^{opt}/\alpha}$	< 0	$= 0$	> 0

statics of the equilibrium and socially optimal solutions with respect to parameters α and β.

More precisely, we study the elasticities

$$E_{x/\alpha} = \frac{dx}{d\alpha} \cdot \frac{\alpha}{x} \,, \qquad E_{f/\alpha} = \frac{df}{d\alpha} \cdot \frac{\alpha}{f} \,, \qquad E_{N/\alpha} = \frac{dN}{d\alpha} \cdot \frac{\alpha}{N} \,,$$

$$E_{Nf/\alpha} = \frac{d(Nf)}{d\alpha} \cdot \frac{\alpha}{Nf} \,, \qquad E_{p/\alpha} = \frac{dp}{d\alpha} \cdot \frac{\alpha}{p}$$

with respect to the parameter α. In equilibrium, the signs of the elasticities can be found in Table 1, where the symbol "?" means that the sign of corresponding elasticity is not uniquely determined. In Social Optimality, the signs of the elasticities can be found in Table 2. As to comparative statics with respect to β, we study the elasticities

$$E_{x/\beta} = \frac{dx}{d\beta} \cdot \frac{\beta}{x} \,, \qquad E_{f/\beta} = \frac{df}{d\beta} \cdot \frac{\beta}{f} \,, \qquad E_{N/\beta} = \frac{dN}{d\beta} \cdot \frac{\beta}{N} \,,$$

$$E_{Nf/\beta} = \frac{d(Nf)}{d\beta} \cdot \frac{\beta}{Nf} \,, \qquad E_{p/\beta} = \frac{dp}{d\beta} \cdot \frac{\beta}{p} \,.$$

Table 3. Equilibrium: Comparative statics w.r.t. β.

	$\frac{\partial r_u}{\partial \beta} < 0$			$\frac{\partial r_u}{\partial \beta} > 0$		
	$\varepsilon_c' > 0$	$\varepsilon_c' = 0$	$\varepsilon_c' < 0$	$\varepsilon_c' > 0$	$\varepsilon_c' = 0$	$\varepsilon_c' < 0$
$E_{x^*/\beta}$	> 0	> 0	> 0	< 0	< 0	< 0
$E_{f^*/\beta}$	> 0	> 0	> 0	< 0	< 0	< 0
$E_{N^*/\beta}$	< 0	< 0	< 0	> 0	> 0	> 0
$E_{N^* f^*/\beta}$	< 0	$= 0$	> 0	> 0	$= 0$	< 0
$E_{p^*/\beta}$	< 0	< 0	< 0	> 0	> 0	> 0

Table 4. Social Optimality: Comparative statics w.r.t. β.

	$\frac{\partial \varepsilon_u}{\partial \beta} > 0$			$\frac{\partial \varepsilon_u}{\partial \beta} < 0$		
	$\varepsilon_c' > 0$	$\varepsilon_c' = 0$	$\varepsilon_c' < 0$	$\varepsilon_c' > 0$	$\varepsilon_c' = 0$	$\varepsilon_c' < 0$
$E_{x^{opt}/\beta}$	> 0	> 0	> 0	< 0	< 0	< 0
$E_{f^{opt}/\beta}$	> 0	> 0	> 0	< 0	< 0	< 0
$E_{N^{opt}/\beta}$	< 0	< 0	< 0	> 0	> 0	> 0
$E_{N^{opt} f^{opt}/\beta}$	< 0	$= 0$	> 0	> 0	$= 0$	< 0

In equilibrium, the signs of the elasticities can be found in Table 3. In Social Optimality, the signs of the elasticities can be found in Table 4.

The analysis shows that the equilibrium variables depend on the elasticity of demand in a similar way as the socially optimal variables depend on the elasticity of utility.

The paper concerns with [2], [3], [4] and [5]. Our research technique uses [6].

The results can be generalized to another monopolistic competition models: retailing [7], market distortion [8], international trade [9], and to the marketing models: optimization of communication expenditure [10] and the effectiveness of advertising [11], pricing [12].

References

1. Dixit A.K., Stiglitz J.E. Monopolistic Competition and Optimum Product Diversity // American Economic Review. 1977. V. 67, No. 3. P. 297–308.
2. Antoshchenkova I.V., Bykadorov I.A. Monopolistic competition model: The impact of technological innovation on equilibrium and social optimality // Automation and Remote Control. 2017. V. 78, No. 3. P. 537–556.
3. Bykadorov I. Monopolistic Competition Model with Different Technological Innovation and Consumer Utility Levels // CEUR Workshop Proceeding. 2017. V. 1987. P. 108–114.
4. Aizenberg N., Bykadorov I., Kokovin S. Beneficial welfare impact of bilateral tariffs under monopolistic competition // Abstracts of the Tenth International Conference Game Theory and Management. Saint Petersburg: Saint Petersburg State University. 2017. P. 5–7.
5. Bykadorov I., Kokovin S. Can a larger market foster R&D under monopolistic competition with variable mark-ups? // Research in Economics. 2017. V. 71, No. 4. P. 663–674.
6. Zhelobodko E., Kokovin S., Parenti M., Thisse J.-F. Monopolistic competition in general equilibrium: Beyond the Constant Elasticity of Substitution // Econometrica. 2012. V. 80, No. 6. P. 2765–2784.
7. Bykadorov I.A., Kokovin S.G., Zhelobodko E.V. Product Diversity in a Vertical Distribution Channel under Monopolistic Competition // Automation and Remote Control. 2014. V. 75, No. 8. P. 1503–1524.
8. Bykadorov I., Ellero A., Funari S., Kokovin S., Pudova M. Chain Store Against Manufacturers: Regulation Can Mitigate Market

Distortion // Lecture Notes in Computer Sciences. 2016. V. 9869. P. 480–493.

9. Bykadorov I., Gorn A., Kokovin S., Zhelobodko E. Why are losses from trade unlikely? // Economics Letters. 2015. V. 129. P. 35–38.
10. Bykadorov I., Ellero A., Moretti E. Minimization of communication expenditure for seasonal products // RAIRO Operations Research. 2002. V. 36, No. 2. P. 109–127.
11. Bykadorov I., Ellero A., Funari S., Moretti E. Dinkelbach Approach to Solving a Class of Fractional Optimal Control Problems // Journal of Optimization Theory and Applications. 2009. V. 142, No. 1. P. 55–66.
12. Bykadorov I., Ellero A., Moretti E., Vianello S. The role of retailer's performance in optimal wholesale price discount policies // European Journal of Operational Research. 2009. V. 194, No. 2. P. 538–550.

Energy markets: optimization of transportation system

A.A. Vasin, O.M. Grigoryeva, and N.I. Tsyganov
Lomonosov Moscow State University, Moscow, Russia

Markets of energy resources (natural gas, oil, etc.) play an important role in economies of many countries. Every such market includes its own transmission network system. The present paper provides a method for computation of the optimal transmission system with respect to the total social welfare. A formal model generalizes two well-known optimization problems. The first one is the transportation problem. The second related problem is the social welfare optimization for a market with several goods under perfect competition [1]. The difficulty of the problem is that an expansion of any line requires valuable fixed costs. The problem is in general NP-hard since the transportation problem with non-convex transmission costs is NP-hard [2]. For a market with a tree-type network, we propose a method of the supply-demand balances transfer to the root node. The method originates from the known Welfare Theorem and relies on a solution of the auxiliary problem of convex optimization with zero fixed costs of the lines expansion. Complexity of the method is quadratic with respect to the number of nodes. We also modify the method in order to obtain an approximate solution of the original problem and estimate the welfare loss for such solution.

Consider a market of a homogeneous commodity consisting of several local markets and a transportation network system. Let N denote the set of nodes and $L \subseteq N \times N$ be the set of edges. Each node $i \in N$ corresponds to a local perfectly competitive market. Demand function $D_i(p_i)$ and supply function $S_i(p_i)$ depend on price p_i at node i; they characterize consumers and producers in the market, respectively, and meet standard conditions. The demand function is non-increasing and equal to 0 for sufficiently large prices p_i. It relates to the consumption utility function of consumers at the node i: $U_i(\widehat{v_i}) = \int\limits_0^{\widehat{v_i}} D_i^{-1}(v)dv$, where $\widehat{v_i}$ is a consumption volume of node i. So the inverse demand function $D_i^{-1}(v)$ shows the marginal utility of consumption. The supply function $S_i(p)$ determines the optimal (profit-maximizing) production volume at the node i, i.e, $S_i(p) = Arg \max\limits_v (pv - c_i(v))$, where monotonically increasing convex function $c_i(v)$ shows the minimal production cost of volume v at node i.

For any $(i, j) \in L$, the transmission line is characterized by the initial transmission capacity Q_{ij}^0, the unit transmission cost e_t^{ij}, the cost of the transmission capacity increment, including fixed cost E_f^{ij} and variable cost $E_v^{ij}(Q_{ij} - Q_{ij}^0)$. Let q_{ij} be the flow from the market i to market j, $q_{ji} = -q_{ij}$. The total transmission costs for edge (i, j) are:

$$E^{ij}(q_{ij}) = \begin{cases} E_f^{ij} + E_v^{ij}(|q_{ij}| - Q_{ij}^0) + e_t^{ij} \cdot |q_{ij}|, & if |q_{ij}| > Q_{ij}^0, \\ e_t^{ij} \cdot |q_{ij}|, & if |q_{ij}| \leq Q_{ij}^0, \end{cases} \quad (1)$$

where the variable term E_v^{ij} is a monotonically increasing convex function of increment $(Q_{ij} - Q_{ij}^0)$; $E_v^{ij}(0) = 0$. For given flows $\overrightarrow{q} = (q_{ij}, (i, j) \in L)$ and production volumes $\overrightarrow{v} = (v_i,\ i \in N)$, the total social welfare is

$$W(\overrightarrow{q}, \overrightarrow{v}) = \sum_{i \in N} [U_i(v_i + \sum_{l \in N(i)} q_{li}) - c_i(v_i)] - \sum_{(i,j) \in L,\ i<j} E^{ij}(q_{ij}). \quad (2)$$

The problem under consideration is

$$\max_{\overrightarrow{q} \geq 0, \overrightarrow{v} \geq 0} W(\overrightarrow{q}, \overrightarrow{v}). \quad (3)$$

Consider an auxiliary problem of the social welfare optimization under a fixed set $\overline{L} \subseteq L$ of expanded lines:

$$\max_{\overrightarrow{q} \geq 0, \overrightarrow{v} \geq 0} W(\overrightarrow{q}, \overrightarrow{v}, \overline{L}), \quad (4)$$

where $W(\overrightarrow{q},\overrightarrow{v},\overline{L}) = \sum\limits_{i\in N}[U_i(v_i+\sum\limits_{l\in N(i)} q_{li})-c_i(v_i)] - \sum\limits_{(i,j)\in L,\ i<j} |q_{ij}|e_t^{ij} - \sum\limits_{(i,j)\in\overline{L},\ i<j} (E_f^{ij}+E_v^{ij}(|q_{ij}|-Q_{ij}^0))$, $|q_{ij}| \le Q_{ij}^0$ for all $(i,j)\in L\setminus\overline{L}$. Let $V(\overline{L})$ denote the maximal welfare in this problem. Then problem (3) reduces to finding $L^* = Arg\max\limits_{\overline{L}\subseteq L} V(\overline{L})$. The triple consisting of the price vector $p=(p_i, i\in N)$, the output vector $v=(v_i, i\in N)$, and the flow vector $q=(q_{ij},(i,j)\in L)$ is called a competitive equilibrium of the market if it satisfies the following conditions: $v_i = S_i(p_i)$, $i\in N$; $\Delta S_i(p_i)=\sum_{j\in N(i)} q_{ij}$ for any $i\in N$; for any $(i,j)\in L$ $\quad |p_i-p_j|<e_t^{ij} \Rightarrow q_{ij}=0$, $|p_i-p_j|=e_t^{ij}\Rightarrow |q_{ij}|\le Q_{ij}^0$, $\forall (i,j)\notin\overline{L}$ $\quad |p_i-p_j|>e_t^{ij}\Rightarrow |q_{ij}|=Q_{ij}^0$, $\forall (i,j)\in\overline{L}$ $\quad q_{ij}>Q_{ij}^0\Rightarrow p_j-p_i=e_t^{ij}+e_v^{ij}(q_{ij})$. Our paper [3] establishes that problem (4) is convex, and its solution $(\overrightarrow{q},\overrightarrow{v})(\overline{L})$ meets the first-order conditions (FOCs) which determine the competitive equilibrium of the corresponding network market.

Consider an efficient method for solving problem (4) in case of a market with a tree-type network. The idea is to transfer the S-D balances from all nodes to the root of the tree. Let N_1 denote the set of final nodes, $N_k=\{i\notin\bigcup\limits_{j=1}^{k-1}N_j \mid |Z(i)\setminus\bigcup\limits_{j=1}^{k-1}N_j|=1\}$ be the set of k-level nodes, $k=1,..,r$, $N_r=\{0\}$, where 0 is the root node of the tree.

Stage 1 Transfer of S-D balance to the root. Let $\Delta\overline{S}_j(p_j)$ denote the S-D balance at node j taking into account the transfer from all following nodes.

Substage 1.1 For every final node of the tree, let $\Delta\overline{S}_i(p_i)=\Delta S_i(p_i)$, $i\in N_1$.

Substage 1.l, $l=2,..,h$ For every node $j\in N_l$, we set

$$\Delta\overline{S}_j(p_j)=\Delta S_j(p_j)+\sum_{i\in\sigma^{-1}(j)}\Delta S_{ij}(p_j), \tag{5}$$

where the transfer from i to j for $(i,j) \in \overline{L}$ is

$$\Delta S_{ij}\,(p_j) = \begin{cases} 0,\ \overline{\tilde{p}}_i - e_t^{ij} < p_j < \overline{\tilde{p}}_i + e_t^{ij}, \\ \Delta\overline{S}_i(p_j - e_t^{ij}),\ \overline{\tilde{p}}_i + e_t^{ij} \le p_j < (\Delta\overline{S}_i)^{-1}(Q_{ij}^0) + e_t^{ij}, \\ Q_{ij}^0,\ (\Delta\overline{S}_i)^{-1}(Q_{ij}^0) + e_t^{ij} \le p_j \le (\Delta\overline{S}_i)^{-1}(Q_{ij}^0) + \\ \qquad\qquad + e_t^{ij} + e_v^{ij}(0), \\ \{q_{ij} | q_{ij} = \Delta\overline{S}_i(p_j - e_t^{ij} - e_v^{ij}(q_{ij}))\},\ p_j > (\Delta\overline{S}_i)^{-1}(Q_{ij}^0) + \\ \qquad\qquad + e_t^{ij} + e_v^{ij}(0), \end{cases} \tag{6}$$

$\overline{\tilde{p}}_i$ is a solution of $\Delta\overline{S}_i(p_i) = 0$. If $(i,j) \notin \overline{L}$, then

$$\Delta S_{ij}\,(p_j) = \begin{cases} 0,\ \overline{\tilde{p}}_i - e_t^{ij} < p_j < \overline{\tilde{p}}_i + e_t^{ij}, \\ \Delta\overline{S}_i(p_j - e_t^{ij}),\ \overline{\tilde{p}}_i + e_t^{ij} \le p_j < (\Delta\overline{S}_i)^{-1}(Q_{ij}^0) + e_t^{ij}, \\ Q_{ij}^0,\ p_j \ge (\Delta\overline{S}_i)^{-1}(Q_{ij}^0) + e_t^{ij}. \end{cases} \tag{7}$$

For $p_j \le \overline{\tilde{p}}_i - e_t^{ij}$ the value $\Delta S_{ij}(p_j) < 0$ is determined in a similar way. As a result of stage 1, we obtain $\Delta\overline{S}_0(p_0)$.

Stage 2 Determination of the equilibrium prices and the optimal strategy

Substage 2.1 We determine $\tilde{p}_0$ from the equation $\Delta\overline{S}_0(\tilde{p}_0) = 0$, and set $v_0^* = S_0(\tilde{p}_0)$.

Substage 2.l, $l = 2, .., h$ Consider a node $i \in N_{(h-l+1)\bigcap\sigma^{-1}(j)}$, where $j \in N_{h-l+2}$. From (5)-(7), we find

$$q_{ij}^* = \Delta S_{ij}(\tilde{p_j}),\ \tilde{p}_i = (\Delta\overline{S_i})^{-1}(\Delta S_{ij}(\tilde{p_j})). \tag{8}$$

Then we set $v_i^* = S_i(\tilde{p}_i)$. Finally we obtain q_{ij}^*, $\tilde{p}_i$ and v_i^* for every node $i \in N$ and $j = \sigma(i)$.

Theorem 4 *The given algorithm determines a solution of problem (4). Its complexity with respect to the number of the nodes is* $O(|N|^2)$.

An algorithm for approximate solution of the general problem is a modification of the given algorithm for problem (4).

Substage 1.2m Consider $i \in N_1$ and function $\Delta S_{ij}(p_j)$ defined by (5,6). Inverse function $c_{ij}(q) = (\Delta S_{ij})^{-1}(q)$ determines the marginal cost of the transfer from i to j. Let $a_{ij} = Q_{ij}^0$, $a_{ji} = -Q_{ji}^0$, b_{ij} and b_{ij} denote solutions of equations

$$\begin{array}{l} (b_{ij} - a_{ij}) \cdot c_{ij}(b_{ij}) - \int_{a_{ij}}^{b_{ij}} c_{ij}(q)dq = E_f^{ij}, \\ \int\limits_{b_{ji}}^{a_{ji}} c_{ij}(q)dq - (a_{ji} - b_{ji}) \cdot c_{ij}(b_{ji}) = E_f^{ij}. \end{array} \tag{9}$$

We introduce

$$c^m_{ij}(q) = \begin{cases} c_{ij}(q), & q \in (-Q^0_{ji}, Q^0_{ij}), \quad q > b_{ij} \quad or \quad q < b_{ji}, \\ c_{ij}(b_{ij}), & q \in (a_{ij}, b_{ij}), \\ c_{ij}(b_{ji}), & q \in (b_{ji}, a_{ji}), \end{cases} \tag{10}$$

define $\Delta S^m_{ij}(p) = (c^m_{ij})^{-1}(p)$ and call (a_{ij}, b_{ij}) and (b_{ji}, a_{ji}) connected intervals of the transfer ΔS^m_{ij}. In these intervals we increase and equalize the marginal costs in order to cover the fixed cost of the transfer. We define $\Delta \overline{S}^m_j(p_j)$ for $j \in N_2$ according to (5) with one change: we substitute ΔS^m_{ij} for ΔS_{ij}. The set of connected intervals for this function is:

$$Int(j) = \{(\overline{a}_{ij}, \overline{b}_{ij}), (\overline{b}_{ji}, \overline{a}_{ji}), i \in \sigma^{-1}(j)\}, \text{where } \overline{a}_{ij} = a_{ij} + \Delta_{ij}(b_{ij}),$$

$$\overline{b}_{ij} = b_{ij} + \Delta_{ij}(b_{ij}),\ \overline{a}_{ji} = a_{ji} + \Delta_{ij}(b_{ji}),\ \overline{b}_{ji} = b_{ji} + \Delta_{ij}(b_{ji}),$$

$$\Delta_{ij}(v) = \sum_{r \in \sigma^{-1}(j) \setminus i} \Delta S^m_{rj}(p_{ij}(b_{ij})) + \Delta S_j(p_{ij}(b_{ij})),\ p_{ij}(v) = \Delta S^{-1}_{ij}(v).$$

These formulas relate to a typical case where ΔS_j and ΔS^m_{rj} have no jumps at $p_{ij}(b_{ij})$ and $p_{ij}(b_{ji})$, $r \in \sigma^{-1}(j) \setminus i$.

Substage 1.lm, $l = 3, .., h$. Now we consider $j \in N_l$ and $i \in \sigma^{-1}(j)$. The function $\Delta \overline{S}^m_i(p_i)$ and the set $Int(i)$ are determined at the previous substage. Let $Int(i) = ((c^i_1, d^i_1), (c^i_2, d^i_2), ..., (c^i_{n(i)}, d^i_{n(i)}))$, $n(i) \leq 2 \cdot (l-2)$. We define function $\Delta S^m_{ij}(p_j)$ according to (7), the only change is that we employ $\Delta \overline{S}^m_i$ instead of $\Delta \overline{S}_i$. Next, we set $a_{ij} = Q^0_{ij}$, determine b_{ij} as a minimal solution of (9), exclude from $Int(i)$ all intervals that intersect with (a_{ij}, b_{ij}) and add their unification with (a_{ij}, b_{ij}) to the set $Int(i)$. In a symmetric way we determine a_{ji} and b_{ji} and change the set $Int(i)$. Thus we obtain the set $Int(i,j) = ((c^{ij}_1, d^{ij}_1), ..., (c^{ij}_{n(i,j)}, d^{ij}_{n(i,j)}))$. The prices $p^{ij}_{kL} \leq p^{ij}_{kH}$ corresponding to the ends of interval k proceed from relation $S^m_{ij}(p^{ij}_{kL}) \ni c^{ij}_k$, $S^m_{ij}(p^{ij}_{kH}) \ni d^{ij}_L$. For every $r \in \sigma^{-1}(j)$, we define the set of connected price stretches $P^{rj} = \{[p^{rj}_{1L}, p^{rj}_{1H}], ..., [p^{rj}_{n(r,j)L}, p^{rj}_{n(r,j)H}]\}$ and examine the unifications $\bigcup\limits_{r \in \sigma^{-1}(j)} P^{rj}$. We consider unifications of all intersecting stretches from this set and order them by increase. Thus, we obtain the set of connected price stretches $P^j = \{[p^j_{1L}, p^j_{1H}], ..., [p^j_{n(j)L}, p^j_{n(j)H}]\}$ for the function $\Delta \overline{S}^m_j(p_j)$. For every $k = 1, .., n(j)$, we determine the corresponding connected interval (c_k^j, d_k^j) of production volumes for this function. There exists of least one $i \in \sigma^{-1}(j)$ such that $p_{k'L}^{ij} = p_{kL}^j$ for some $k' \leq n(i,j)$. In the regular case, where functions $\Delta S_{rj}(p), r \neq i$, and $\Delta S_j(p)$ have no jumps at this price, we set $c_k^j = c_{k'}^{ij} + \sum\limits_{r\in\sigma^{-1}(j)\setminus i} \Delta S_{rj}(p_{kL}^j) + \Delta S_j(p_{kL}^j)$. Similarly, there exists $i' \in \sigma^{-1}(j)$ such that $p_{k'H}^{i'j} = p_{kH}^j$ for some $k' \leq n(i',j)$, and in the regular case we set $d_k^j = d_{k'}^{i'j} + \sum\limits_{r\in\sigma^{-1}(j)\setminus i} \Delta S_{rj}(p_{kH}^j) + \Delta S_j(p_{kH}^j)$. As a final result of stage 1, we obtain the function $\Delta\overline{S}_0^m(p_0)$ and the set $Int(0)$.

Stage 2m Equilibrium prices $\overline{\overline{p}}_i^m$ and strategy $(\overrightarrow{v}^m, \overrightarrow{q}^m)$ are determined according to stage 2 of the basic algorithm for problem (4) where at every substage we employ functions $\Delta\overline{S}_i^m$ and c_{ij}^m instead of $\Delta\overline{S}_i$ and c_{ij}. Modification of the marginal costs c_{ij}^m is equivalent to the following change of the transmission costs:

$$e_m^{ij}(q_{ij}) = \begin{cases} E^{ij\prime}(q_{ij}), & if\ q_{ij} \notin (a_{ij}, b_{ij}) \cup (b_{ji}, a_{ji}), \\ e_M^{ij} - (\Delta\overline{S}_i^m)^{-1}(q_{ij}), & if\ q_{ij} \in (a_{ij}, b_{ij}), \\ e_M^{ji} - (\Delta\overline{S}_i^m)^{-1}(q_{ij}), & if\ q_{ij} \in (b_{ji}, a_{ji}), \end{cases} \tag{11}$$

where $a_{ij} = Q_{ij}^0$, $a_{ji} = -Q_{ji}^0$, b_{ij} and b_{ji} are determined as minimal solutions of (9), $e_M^{ij} = (\Delta S_{ij})^{-1}(b_{ij})$, $e_M^{ji} = (\Delta S_{ij})^{-1}(b_{ji})$, $E_f^{ij} = 0$. Denote $W^m(\overrightarrow{q}, \overrightarrow{v})$ as the total social welfare function with modified transmission costs:

$$W^m(\overrightarrow{q}, \overrightarrow{v}) = \sum_{i\in N} [U_i(v_i + \sum_{l\in N(i)} q_{li}) - c_i(v_i)] - \sum_{(i,j)\in L,\ i<j} E_m^{ij}(q_{ij}).$$

Theorem 5 *The strategy $(\overrightarrow{v}^m, \overrightarrow{q}^m)$ determined by the modified algorithm is a solution of the welfare optimization problem with perturbed transmission cost functions $E_m^{ij}(q_{ij})$. If $0 \notin \bigcup\limits_{k=1,\ldots,n(0)} (c_k^0, d_k^0)$ then $(\overrightarrow{v}^m, \overrightarrow{q}^m)$ is a solution of general problem (3).*

Consider a particular case, where consumers and producers are separated by the root node 0 in the following sense: some branches of the tree are producing, and the flows in these branches go from final nodes to the root, and the rest branches are primarily consuming, and the flows there go from the root to the final nodes. In this case, we can simplify our method as follows. For every supply branch, we determine the net supply transfer to the root node and do not care about the negative part of the supply-demand balance. For every consuming branch, we determine the net demand transfer to the root node in a symmetric way. Thus we obtain the aggregated supply $\overline{S}_0(p_0)$, the aggregated demand $\overline{D}_0(p_0)$, the connected intervals for the both functions, the equilibrium volume $\tilde{v}$ and the equilibrium price for the root node. For the case, where $\tilde{v}$ belongs to some connected interval, we propose the following approximate solutions. Denote $\{(c^k_{0\overline{S}}, d^k_{0\overline{S}}), k = 1,..,n(\overline{S}_0)\}$ as the set of connected intervals for the agregated supply, $\{(c^k_{0\overline{D}}, d^k_{0\overline{D}}), k = 1,..,n(\overline{D}_0)\}$ — as a similar set for the agregated demand. Let $v_L \overset{\text{def}}{=} max\{v \in [0, \widetilde{v}) | v \notin \cup_k (c^k_{0\overline{S}}, d^k_{0\overline{S}}) \vee \cup_k (c^k_{0\overline{D}}, d^k_{0\overline{D}})\}$, $v_u \overset{\text{def}}{=} min\{v \geq \widetilde{v} | v \notin \cup_k (c^k_{0\overline{S}}, d^k_{0\overline{S}}) \vee \cup_k (c^k_{0\overline{D}}, d^k_{0\overline{D}})\}$.

We determine the prices $p^S_L = \overline{S}_0^{-1}(v_L)$ and $p^D_L = \overline{D}_0^{-1}(v_L)$. Then we set the production volumes v^L_i and the flows q^L_{i0} in every producing node $i \in \sigma^{-1}(0)$. We employ relations (8), but use p^S_L instead of $\tilde{p}_0$. We continue to determine $q^L_{i\sigma(i)}$ and v^L_i according to (8) in all following nodes in these branches.

Then, proceeding from the price p^D_L, we successively determine $q^L_{i\sigma(i)}$ and v^L_i for all following nodes i in the consuming branches.

Theorem 6 *The welfare losses for the given approximate solutions meet the following estimates:*

$$W^* - W(\overrightarrow{v}^L, \overrightarrow{q}^L) \leq \int\limits_{v^L}^{\tilde{v}} (\overline{D}_0^{-1}(v) - \overline{S}_0^{-1}(v))dv,$$

$$W^* - W(\overrightarrow{v}^U, \overrightarrow{q}^U) \leq \int\limits_{\tilde{v}}^{v^U} (\overline{S}_0^{-1}(v) - \overline{D}_0^{-1}(v))dv.$$

References

1. Arrow K.J., Debreu G., Existence of an Equilibrium for a Competitive Economy, Econometrica 22, 265-290 (1954).
2. Guisewite G.M., Pardalos P.M., Minimum concave-cost network flow problems: Applications, complexity, and algorithms. Annals of Operations Research, 25 (1); 1990. p. 75–99.
3. Vasin A.A., Grigoryeva O.M., Tsyganov N.I., Optimization of an energy market transportation system, Doklady Mathematics,2017, Vol. 96, No.1, pp. 1-4.

OR in military and computer-aided design

Model-oriented programming: CAD methods in the programs design

Yu.I. Brodsky
Federal Research Center "Computer Science and Control" of the Russian Academy of Sciences, Moscow, Russia

In this work, the new programming paradigm is offered, with higher level of encapsulation, than in the object-oriented approach. Its key features - an exclusion of imperative programming, and focusing on the distributed and high-performance calculations. The approach proposed is applicable for rather wide class of tasks including creation of simulation models of complex multi-component systems.

Model synthesis and the model-oriented programming as methods of the description, synthesis and program realization of simulation models of complex multicomponent systems, were developed in the department of Simulation systems of Computing Centre of Academy of Sciences of the USSR and further Russian Academy of Sciences, since the end of the 80th. The concept of model-component - the universal modelling agent - is the base of model synthesis. The model-component is similar to the object of the object analysis but supplied with not only characteristics and methods, capable to do something useful if they are caused, but a certain analogy of system services of an operating system, always functioning and always ready to give standard answers to the standard requests.

Let us formalize, basing on the closeness hypothesis and its consequences, the family of simulation models of complex systems, by the family of species of structures [2] in N. Bourbaki's sense. Base sets of the family representatives are the sets of characteristics of the model, methods (what the model is able to do) and events (on what the model has to be able to react). The family of species of structures model-component possesses two important properties:

1. The organization of calculations is same for all representatives of the family. Besides, the considerable part of these calculations can be executed in parallel. It means possibility of creation of the universal program focused on high-performance or distributed computing capable to execute any simulation model, if that is the mathematical object supplied with the species of structure of the model-component family.

2. The family of species of structures model component is closed, under the operation of uniting components into the model-complex. The complex received by association of models-components belongs to the family of species of structures model-component, and, therefore, can be included in new complexes, and the organization of simulation calculations of any model-complex can be carried out by the same universal program.

The properties of the model component family of species of structures allow offering the new model-oriented programming method for program realization of simulation models of complex systems.

At this approach the program complex seems as a complex of models-components, whose behaviour is no need to arrange (for example, by calling any methods) - all components always behave as they can. Programming consists in the description of the component's arrangement and behaviour (in fact - in the description of the corresponding species of structure) and in the description of creation of complexes from the components.

Historically, the changings of programming paradigms was accompanied by the aggregation of the base instruments of the programmer's activity. It all started with the machine instruction, then, with the advent of high-level languages – such a tool became an operator, which implements some action, possibly with a few machine instructions.

The victory of structured programming ideas replaced individual operators and variables by standard constructions such as loop, branching, subroutines-functions and data structures. With the advent of object analysis, the object became the main unit of the design. It unites some kind of data structure with a set of methods necessary for the data processing. In addition, through the inheritance mechanism, you can build a hierarchy of object classes, developing, implementing and embodying basic ideas of the root classes of this hierarchy. This programming paradigm is currently the dominant and its basic concepts, such as class, object, data typing, inheritance, encapsulation, polymorphism is implemented with some nuances in modern imperative programming languages, such as C++, Java, C#, Delphi and others.

Inheritance relationship for the set of classes of object-oriented programming language is a partial order. Classes that have no ancestors, but have descendants are called to them root or base. Classes that do not have descendants are called leaf.

Designing of a large software system within the object paradigm is laying the basic concepts and ideas of this system into the base classes of objects and then building a hierarchy of classes, developing, specifying and embody these ideas in a variety of leaf classes, with the help of which the target software system will be built. Even if the object-oriented design has built the greatest hierarchy of classes using inheritance, still all the organization of the computational process lies on the developer of the system: for the system operation – the developer is to organize calling of appropriate methods in the desired sequence.

To describe complex systems in the object paradigm, the unified modelling language UML was proposed in the late 90s [1]. The creators of UML have opted for a sharp increase in the number of initial concepts and ideas. They say about the language: “UML is subject to the rule of 80/20, i.e., 80% of most problems can be solved using 20% of the UML” [1]. Apparently, any system can be described with the help of the UML, and even from several points of view. The question is what to do next with such descriptions – there is no unity in opinions. Some specialists believe that the main value of UML is just in the application as a mean of recording and sharing formalized descriptions of the stages of the sketch and design of complex program systems. However, there are a number of tools, which allows compiling the UML-descriptions into the billet classes of universal programming languages, and in this case we can speak about the mode of using UML as a programming language, though a hard problem of compilation quality arises. Here we will describe a different approach to the programming – a model-oriented one, which is based on the model synthesis [2, 3], on the concept of the model-component. A model-component of the model synthesis is more complex and aggregate structure, than the object of the object analysis. Its main difference from the object – the possession of its own behavior, in the sense in which, for example, a computer with the operating system loaded, has behavior – the ability to respond on a set of standard requests in predetermined manner, known in advance.

In this paper we will describe a different approach to the implementation of complex systems models - a model-oriented one, which is based on the model synthesis, on the concept of the model-component, very close to that in mathematical modelling and especially in simulation. A model-component of the model synthesis is more complex and aggregate structure, than the object of the object analysis. Its main difference from the object - the possession of its own behaviour, in the sense in which, for example, a computer with the operating system loaded, has behaviour - the ability to respond on a set of standard external and internal requests in predetermined manner, known in advance. It turns out that the way of the model's behaviour arrangement (organization of simulation computations) can also be standard - the same for any model, no matter how large and complex it may be.

The proposed idea of model synthesis is minimalistic in a set of basic concepts: it has the only basic concept – a model-component and an auxiliary concept – a model-complex, which after all can also be treated as a model-component, hence itself may be included in the next level model-complexes as their component, etc. Any model may be run according the same standard rules, i.e., can be performed by once written and debugged program of models' execution. In addition, the program of models' execution is such that most of its calculations allow parallelization. The process of the programming breaks down firstly, into series of declarative descriptions of models-components and models-complexes combining models-components; and secondly, on the programming of certain functional dependencies, which are functions in the mathematical – not programist sense (i.e., unambiguous and have no states and side effects), and therefore can be programmed in a functional paradigm. This leads to the fact that the software implementation of even the most complicated fractal arranged system excludes the imperative programming – the most difficult at the stage of debugging.

The fee for this is a higher level of encapsulation. In contrast to the object analysis it is impossible (and unnecessary) to call the methods of the model-component "manually" with some "foreign" parameters. They are always called automatically, and only with the subset of characteristics of the model-component as parameters, in accordance with the described behavior of the model.

For the descriptions of models-components and models-complexes, the LCCD declarative language (language of the description of complexes and components) is elaborated. LCCD descriptions are compiled into database tables. Therefore, the question of compilation quality is not too actual - correctness of the compilation is important. The quality of computations lies in the universal program of models' execution, which may be optimized once and forever.

The proposed concept has been realized in series of simulation models implemented under the influence of the model-oriented programming paradigm. For example, some episodes of Reagan's SDI (Strategic Defence Initiative) functioning and the model of interaction of several countries were simulated. In addition, the system for model-oriented programming was incorporated into the simulation system MISS [4].

The concepts of model synthesis and model-oriented programming are applicable primarily for the description, design and software implementation of simulation models of complex multi-component systems. However, it is hoped that a similar approach can be used for the development of complex software systems, with the organization that fits the closeness hypothesis, including the software systems focused on high-performance computing.

References

1. Booch G., Rumbaugh J., Jacobson I. UML. User Guide, 2-ed., Addison-Wesley, 2005.
2. Brodsky Yu.I. Bourbaki's Structure Theory in the Problem of Complex Systems Simulation Models Synthesis and Model-Oriented Programming // Computational Mathematics and Mathematical Physics Vol. 55 No. 1, 2015. P. 148-159.
3. Brodsky Yu.I. Model synthesis and model-oriented programming – the technology of design and implementation of simulation models of complex multicomponent systems // In the World of Scientific Discoveries, Series B, 2014, Vol. 2, No 1, P. 12-31.
4. Brodsky Yu.I., Lebedev V.Yu. Instrumental'naya sistema imitatsionnogo modelirovaniya MISS [Instrumental Simulation System MISS] Moscow: CC AS of the USSR, 1991, 180 p. (in Russian)

OR in finance and banking

On low bounds for American call option[*]

V.V. Morozov and O.A. Migacheva
Lomonosov Moscow State University, Moscow, Russia

A low bound for American call based on threshold decision rule was introduced by Broady and Detemple in [1]. In [2] it was improved by Chung, Hung and Wang using threshold function $L\exp(a(T-t))$.

1. We construct two bounds generalizing just mentioned in the following way. The asset price $S(t)=S(0)\exp(\tilde{\alpha}t+\sigma z(t)),\ 0\leqslant t\leqslant T$, is geometric Brown motion process where T – expiring time, $z(t)$ – standard Wiener process, σ – volatility, $\tilde{\alpha}=\alpha-\sigma^2/2$ and α – mean asset return. Take $T_1\in(0,T)$, $L_1\leqslant L_2$, $a_1\geqslant 0, a_2\geqslant 0$ and define functions

$$f_1(L_1,L_2,T_1,t)=\begin{cases} L_2, & 0\leqslant t<T_1,\\ L_1, & T_1\leqslant t\leqslant T,\end{cases}$$

$$f_2(L,a_1,a_2,T_1,t)=\begin{cases} L\exp(a_1(T-T_1)+a_2(T_1-t)), & 0\leqslant t<T_1,\\ L\exp(a_1(T-t)), & T_1\leqslant t\leqslant T.\end{cases}$$

Let τ_i $(i=1,2)$ be stopping moments when the process $S(t)$ first time hits the graph of the function f_i. (If it hits the graph of f_1 at time T_1 then $S(t)\in[L_1,L_2]$.) Define functions

$$V_1(S,L_1,L_2,T_1)=\mathrm{E}[\exp(-r\tau_1)(S(\tau_1)-K)^+|S(0)=S],$$

[*]The reported study was funded by RFBR according to the research project 16-01-00353 a.

$$V_2(S,L,a_1,a_2,T_1) = \mathrm{E}[\exp(-r\tau_2)(S(\tau_2)-K)^+|S(0)=S],$$

where K is a strike and r is a constant bank rate. Denote

$$\beta_{1,2} = \frac{-\tilde{\alpha} \pm \xi}{\sigma^2},\ \xi = \sqrt{\tilde{\alpha}^2 + 2r\sigma^2},\ S^* = \frac{\beta_1 K}{\beta_1 - 1}.$$

Low bounds for call option price at time 0 are

$$V_1^*(S) = \max_{(L_1,L_2,T_1)\in D_1} V_1(S,L_1,L_2,T_1),$$

$$D_1 = \{(L_1,L_2,T_1)|\max(S,K) \leqslant L_1 \leqslant L_2 \leqslant S^*,\ 0 \leqslant T_1 \leqslant T\},$$

$$V_2^*(S) = \max_{(L,a_1,a_2,T_1)\in D_2} V_2(S,L,a_1,a_2,T_1),$$

$$D_2 = \{(L,a_1,L_2,T_1)|\max(S,K) \leqslant L \leqslant S^*,\ a_1,\ a_2 \geqslant 0,\ 0 \leqslant T_1 \leqslant T\}.$$

2. In this section we remind the low bounds for call option from [1,2]. Take $a \geqslant 0,\ S = S(0) < Le^{aT}, K < L$. Define a Brown motion process $x_1(t,a) = (\tilde{\alpha}+a)t + \sigma z(t),\ z(0)=0$, and consider the threshold rule

$$\tau_a = \min\{t \mid S(t) = Le^{a(T-t)}\} = \min\{t \mid x_1(t,a) = x_a \overset{def}{=} \ln(Le^{aT}/S)\}$$

when the process $S(t)$ (process $x_1(t,a)$) for a first time reaches the threshold function $Le^{a(T-t)}$ (threshold x_a). Cumulative and density functions for the random variable τ_0 are

$$G_1(x_0,t,\tilde{\alpha}) = \mathbb{P}(\tau_0 \leqslant t) = \Phi\Big(\frac{-x_0+\tilde{\alpha}t}{\sigma\sqrt{t}}\Big) + \exp\Big(\frac{2\tilde{\alpha}x}{\sigma^2}\Big)\Phi\Big(\frac{-x_0-\tilde{\alpha}t}{\sigma\sqrt{t}}\Big),$$

$$g_1(x_0,t,\tilde{\alpha}) = G'_{1t}(x_0,t,\tilde{\alpha}) = \frac{x_0}{\sigma t^{3/2}}\phi\Big(\frac{-x_0+\tilde{\alpha}t}{\sigma\sqrt{t}}\Big),\ t>0,$$

where Φ and ϕ – cumulative and density functions of standard normal distribution $\mathcal{N}(0,1)$. Write also functions

$$P_1(x_0,y,T,\tilde{\alpha}) = \mathbb{P}(x_1(T,0) \leqslant y,\ \tau_0 \geqslant T),$$

$$p_1(x_0,y,T,\tilde{\alpha}) = P'_{1y}(x_0,y,T,\tilde{\alpha})$$

for the random variable $x_1(T,0)$ when the process $x_1(t,0)$ doesn't reach the level x_0 before time T :

$$P_1(x_0,y,T,\tilde{\alpha}) = \Phi\Big(\frac{y-\tilde{\alpha}T}{\sigma\sqrt{T}}\Big) - \exp\Big(\frac{2\tilde{\alpha}x_0}{\sigma^2}\Big)\Phi\Big(\frac{y-2x_0-\tilde{\alpha}T}{\sigma\sqrt{T}}\Big),$$

$$p_1(x_0,y,T,\tilde{\alpha}) = \frac{1}{\sigma\sqrt{T}}\Big[\phi\Big(\frac{y-\tilde{\alpha}T}{\sigma\sqrt{T}}\Big) - \exp\Big(\frac{2\tilde{\alpha}x_0}{\sigma^2}\Big)\phi\Big(\frac{y-2x_0-\tilde{\alpha}T}{\sigma\sqrt{T}}\Big)\Big].$$

Similar distributions for stopping moment τ_a are $G_1(x_a,t,\tilde{\alpha}+a)$,

$$g_1(x_a,t,\tilde{\alpha}+a), P_1(x_a,y+aT,T,\tilde{\alpha}+a) = \mathbb{P}(x(T,0)\leqslant y,\ \tau_a \geqslant T) =$$

$$= \mathbb{P}(x(T,a)\leqslant y+aT,\ \tau_a\geqslant T),\ p_1(x_a,y+aT,T,\tilde{\alpha}+a) =$$

$$= \frac{1}{\sigma\sqrt{T}}\Big[\phi\Big(\frac{y-\tilde{\alpha}T}{\sigma\sqrt{T}}\Big) - \exp\Big(\frac{2(\tilde{\alpha}+a)x_a}{\sigma^2}\Big)\phi\Big(\frac{y-2x_a-\tilde{\alpha}T}{\sigma\sqrt{T}}\Big)\Big].$$

Define notations:

$$\beta_{1,2}(a) = \frac{-(\tilde{\alpha}+a)\pm\xi_a}{\sigma^2},\ \xi_a = \sqrt{\tilde{\alpha}+a)^2+2(r+a)\sigma^2},$$

$$\hat{\beta}_{1,2}(a) = \frac{-(\tilde{\alpha}+a)\pm\hat{\xi}_a}{\sigma^2},\ \hat{\xi}_a = \sqrt{\tilde{\alpha}+a)^2+2r\sigma^2},$$

$$d_1(M,b) = \frac{\ln(M)+(b+\sigma^2)T}{\sigma\sqrt{T}},\ d_2(M,b) = \frac{\ln(M)+bT}{\sigma\sqrt{T}}.$$

The following function W bounds the price of call option from below:

$$W(S,L,T,a) = V(S,L,T,a)+U(S,L,T,a),$$

$$V(S,L,T,a) = \mathrm{E}[e^{-r\tau_a}(Le^{a(T-\tau_a)}-K)1_{\tau_a\leqslant T}] =$$

$$= \int\limits_0^T e^{-rt}(Le^{a(T-t)}-K)g_1(x_a,t,\tilde{\alpha}+a)dt =$$

$$= Le^{aT}e^{-\beta_1(a)x_a}G_1(x_a,T,\xi_a) - Ke^{-\hat{\beta}_1(a)x_a}G_1(x_a,T,\hat{\xi}_a),$$

$$U(S,L,T,a) = e^{-rT}\mathrm{E}[(e^{x_1(T,0)}-K)1_{\tau_a>T}] =$$

$$= e^{-rT}\int\limits_{\ln(K/S)}^{x_0}(Se^y-K)p_1(x_a,y+aT,T,\tilde{\alpha}+a)dy.$$

$$= U_1(S,L,T) - U_2(S,L,T,a),$$

$$U_1(S,L,T) = e^{-rT}\int\limits_{\ln(K/S)}^{x_0}(Se^y-K)\frac{1}{\sigma\sqrt{T}}\phi\Big(\frac{y-\tilde{\alpha}T}{\sigma\sqrt{T}}\Big)dy =$$

$$= e^{-\delta T} S\Big[\Phi(d_1(S/K,\tilde{\alpha})) - \Phi(d_1(S/L,\tilde{\alpha}))\Big] -$$

$$-e^{-rT} K\Big[\Phi(d_2(S/K,\tilde{\alpha})) - \Phi(d_2(S/L,\tilde{\alpha}))\Big],$$

$$U_2(S,L,T,a) = Se^{-\delta T}\exp\Big(\frac{2(\tilde{\alpha}+a+\sigma^2)x_a}{\sigma^2}\Big)\cdot$$

$$\cdot\Big[\Phi\Big(d_1\Big(\frac{L^2e^{2aT}}{SK},\tilde{\alpha}\Big)\Big) - \Phi\Big(d_1\Big(\frac{Le^{2aT}}{S},\tilde{\alpha}\Big)\Big)\Big] -$$

$$-Ke^{-rT}\exp\Big(\frac{2(\tilde{\alpha}+a)x_a}{\sigma^2}\Big)\Big[\Phi\Big(d_2\Big(\frac{L^2e^{2aT}}{SK},\tilde{\alpha}\Big)\Big) - \Phi\Big(d_2\Big(\frac{Le^{2aT}}{S},\tilde{\alpha}\Big)\Big)\Big]$$

The low bounds of call option (see [1,2]) are

$$B(S) = \max_{L\geqslant\max(K,S)} W(S,L,T,0),\ H(S) = \max_{L\geqslant\max(K,S),\ a\geqslant 0} W(S,L,T,a).$$

3. In this section we represent formulae for functions V_1, V_2. Denote $x_i = \ln(L_i/S),\ i = 1,2,\ L(T_1) = L\exp(a_2T_1 + a_1(T-T_1))$. Define functions

$$W_1(S,L_1,L_2,T_1) = e^{-rT_1}\int\limits_{-\infty}^{x_1} W(Se^y, L_1, T-T_1, 0)p_1(x_2,y,T_1,\tilde{\alpha})dy,$$

$$\tilde{U}_1(S,L_1,L_2,T_1) = U_1(S,L_2,T_1) - U_1(S,L_1,T_1).$$

$$W_2(S,L,a_1,a_2,T_1) =$$

$$= e^{-rT_1}\int\limits_{-\infty}^{x_{a_1}-a_1T_1} W(Se^y,L,T-T_1,a_1)p_1(L(T_1),y+a_2T_1,T_1,\tilde{\alpha}+a_2)dy,$$

So, $V_1(S,L_1,L_2,T_1) = V(S,L_2,T_1,0)+$

$$+\tilde{U}_1(S,L_1,L_2,T_1) - \Big(\frac{S}{L_2}\Big)^{-2\tilde{\alpha}/\sigma^2}\tilde{U}_1\Big(\frac{L_2^2}{S},L_1,L_2,T_1\Big) + W_1(S,L_1,L_2,T_1),$$

$$V_2(S,L_1,L_2,T_1) = V(S,L\exp(a_1(T-T_1),T_1,a_2) + W_2(S,L,a_1,a_2,T_1).$$

4. Examples. In tables $K = 100,\ r = 0.03,\ \delta = 0.07, \sigma = 0.4$. The 2th and 4th columns contain Broady-Detemple and Chung-Hung-Wang low bounds for call options, the 6th column contains true option value.

Table 1. $T = 0.5$.

S	BD	V_1^*	CHW	V_2^*	TV
90	5.6942	5.7106	5.7186	5.7211	5.7221
100	10.1901	10.2191	10.2329	10.2371	10.2387
110	16.1101	16.1531	16.1731	16.1796	16.1812

Table 2. $T = 3$.

S	BD	V_1^*	CHW	V_2^*	TV
90	15.6088	15.6773	15.7023	15.7184	15.722
100	20.6562	20.7395	20.7693	20.7892	20.7933
110	26.3365	26.4327	26.4678	26.4898	26.4944

References

1. Broadie M., Detemple J. American option valuation: new bounds, approximations and comparison with existing methods// Review of Financial Studies. 1996. V. 9. N 4. P. 1211−1250.
2. Chung S.L., Hung M.W., Wang Jr. Y. Tight bounds of American option prices// Journal of banking & finance. 2010. V. 34. N 1. P. 77-89.

Modeling of exchange order flows

V.S. Nimak and D.Y. Golembiovsky
Lomonosov Moscow State University, Moscow, Russia

Modeling the dynamics of stock market can provide information for creating some investment strategies. Consider the basic concepts of exchange trade.

- Bid — price that a buyer of financial instrument is willing to pay.
- Ask — price at which a seller is willing to give financial instrument.
- Spread — difference between the best ask and the best bid.
- Limit order — an order to trade a certain amount of asset at specified price or better.
- Market order — an order that is executed immediately after entering the exchange at the best price in order book.

- Modification order — an order to reduce or cancel specified limit order.

- Order book — a table of limit orders for buying and selling securities on a stock market. Table 1 shows an example. First column contains volumes of limit buy orders, second holds prices corresponding to the volumes, and third has volumes of limit sell orders.

Market agents can send three types of buy and sell orders. Limit orders fall into electronic trading system, where orders with the same price are summed up by volume, forming a single position for demand or supply. The minimum selling price in order book is called the best ask price, the maximum buying price is called the best bid price. When a market order arrives at the exchange, it matches the best price of order book, and the transaction takes place. For example, if a market buy order has arrived at the exchange, i.e. one of agents wants to buy a certain amount of some asset, then the specified volume of the asset is removed from the best ask. If the best ask volume is not sufficient to satisfy market order, than the best ask price changes and remaining volume meets new best ask. Agent also may withdraw his limit order, sending modification order. Each of the orders changes the state of the order book, which affects the behavior of market agents.

Bid Side	Price	Ask Side
	\$73.86	1800
	\$73.85	54
	\$73.84	245
	\$73.83	100
	\$73.82	330
	\$73.81	
280	\$73.80	
563	\$73.79	
400	\$73.78	
10	\$73.77	
35	\$73.76	

Table 1. Example of order book.

In the article [1] Kont introduces stochastic model for limit order book's behavior, describes estimation procedure of order's parameters, based on statistical data.

Order book events are modeled as independent Poisson processes:

- Limit orders arrive at a distance of i price ticks from the opposite best quote at independent exponential times with rate $\lambda(i)$.

- Modification orders arrive at a distance of i price ticks from the opposite best quote at independent exponential times with rate $\theta(i)$.

- Market orders arrive at independent exponential times with rate μ.

For limit orders, modification orders and market orders the following formulas are proposed to take corresponding rate estimates:

$$\hat{\lambda}(i) = \frac{N_l(i)}{T_*}, \tag{1}$$

$$\hat{\theta}(i) = \frac{N_c(i)}{T_*}, \tag{2}$$

$$\hat{\mu} = \frac{N_m}{T_*}, \tag{3}$$

where $N_l(i)$ is the total number of limit orders arrived at a distance of i from the opposite best quote, $N_c(i)$ is the total number of modification orders arrived at a distance of i from the opposite best quote, N_m is the total number of market orders arrived, T_* is the trading period.

As shown in [2], such estimates correspond to maximum likelihood estimates. Let us find the maximum likelihood estimate for Poisson process of market buy orders arrival.

The density function for exponential distribution is $f(t, \mu) = \mu e^{-\mu t}$. Let $t_i, i = 1, \ldots, N(T)$ be the moments of market buy orders arrival. Then the likelihood function:

$$\mathcal{L}(\mu) = \prod_{i=2}^{N(T)} \mu e^{-\mu(t_i - t_{i-1})} = (\mu)^{N(T)} e^{-\mu H(T)}, \tag{4}$$

where $N(T)$ is the number of market buy orders arrived by the time T, $H(T)$ is the arrival time of last market buy order by the time T.

The log-likelihood function:

$$ln\mathcal{L}(\mu) = N(T)ln(\mu) - \mu H(T). \tag{5}$$

Eventually we can find the maximum likelihood estimate:

$$\frac{\partial ln\mathcal{L}(\mu)}{\partial\mu} = \frac{N(T)}{\mu} - H(T) = 0, \tag{6}$$

$$\hat{\mu} = \frac{N(T)}{H(T)}. \tag{7}$$

Now let us consider some other distributions, that may also describe order flows behavior.

Any data set with finite moments can be fitted by a member of the Johnson families such as S_B, S_U, S_L [3]. For continuous random variable X we can apply three normalizing transformations, having the general form:

$$Z = \gamma + \delta h(X, \xi, \lambda), \tag{8}$$

$$-\infty < \gamma < +\infty,\ \delta > 0,\ -\infty < \xi < +\infty,\ \lambda > 0, \tag{9}$$

where $h(X, \xi, \lambda)$ is a transformation function, Z is a standard normal random variable, γ, δ are shape parameters, λ, ξ are scale and location parameters.

Transformation functions for corresponding families are:

$$h_{S_L} = ln(\frac{X-\xi}{\lambda}),\ X > \xi, \tag{10}$$

$$h_{S_B} = ln(\frac{X-\xi}{\xi+\lambda-X}),\ \xi < X < \xi+\lambda, \tag{11}$$

$$h_{S_U} = Arsh(\frac{X-\xi}{\lambda}),\ -\infty < X < +\infty. \tag{12}$$

We focus on type S_B, that shows the best results for our data among Johnson families. Let $Y = \frac{X-\xi}{\lambda}$. Than it's probability density function (PDF):

$$f_{S_B}(y) = \frac{\delta}{\sqrt{2\pi}}\frac{1}{y(1-y)}e^{-\frac{1}{2}(\gamma+\delta ln(\frac{y}{1-y}))^2},\ \xi < X < \xi+\lambda. \tag{13}$$

Consider gamma distribution too. The probability density function:

$$f(t) = \frac{1}{\Gamma(k)\theta^k}t^{k-1}e^{-\frac{t}{\theta}},\ k, \theta > 0, \tag{14}$$

where $\Gamma(k) = \int_0^\infty x^{k-1}e^{-x}dx$ is an Euler gamma function, k is a shape parameter, θ is a scale parameter.

We will also look at exponential distribution with PDF:

$$f(t) = \lambda e^{-\lambda t},\ \lambda > 0. \tag{15}$$

To compare different models we will use Bayesian information criterion (BIC). This criterion is defined as:

$$BIC = ln(n)k - 2ln(L), \tag{16}$$

where $L = f(x|\hat{\theta})$ is the maximized value of the likelihood function of the model, $\hat{\theta}$ is the parameter values that maximize the likelihood function, x is the observed data, n is the sample size, k is the number of parameters in the model. Such criteria is used only for comparison purposes and do not have any interpretation for it's absolute value. It also penalizes the number of parameters in the model. The lower BIC of the model, the better it fits data.

Let us compare Kont's model and maximum likelihood fitting for different distributions: exponential, gamma, and Jonson's S_B. We will use NASDAQ data for such companies, as Facebook, Intel, Microsoft, Cisco, Vodafone, Liberty Ventures and Liberty Global PLC.

Fig. 1 describes the results for FB (Facebook), limit buy order at the distance of 1 price tick from the best ask. Histogram shows the empirical distribution for arrival data with bin size 0.1 sec. for time period 3 hours (10800 sec.), 11:00 — 14:00, 2014-11-03. As data is always nonnegative, first bar corresponds to zero values. The data has some peculiarity: it contains lots of zero arrival times and times close to zero. Notice, that exponential distribution is a particular case of gamma distribution for the shape parameter $k = 1$. It can not provide condition of going to infinity at $t = 0$ unlike gamma distribution, which make it possible with the shape parameter $0 < k < 1$.

	Kont's model	Exponential	Gamma	Johnson
Limit-Bid-1	99751	-18442	-461372	6020
Limit-Ask-1	112698	-25973	-528153	-2145
Modify-Bid-1	85375	-10265	-292100	5571
Modify-Ask-1	86390	-10835	-305393	11893
Market-Bid	21621	21605	-184024	25062
Market-Ask	21837	21530	-163060	27021

Table 2. BIC values for FB (Facebook).

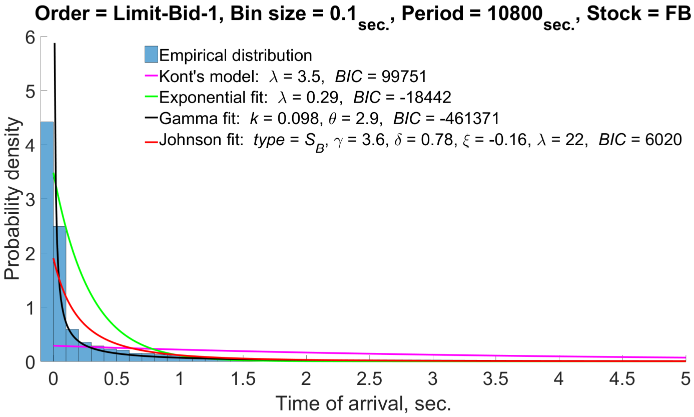


Fig. 1. Distributions for limit order arrival to the bid side at the distance of 1 price tick from the best ask.

For the same data Table 2 displays the example of BIC values for different orders. The results show that gamma distribution provides the best fitting and it suits data mush better than Kont's model.

References

1. Cont R., Stoikov S., Talreja R. Analysis of stochastic dual dynamic programming method // European Journal of Operational Research. 2010. V. 58, N 3. P. 63–72.
2. Rubisov A.D. Statistical Arbitrage Using Limit Order Book Imbalance. Toronto: University of Toronto, 2015.
3. Ramachandran K. .M. Estimation of Parameters of Johnson's System of Distributions // Journal of Modern Applied Statistical Methods. 2011. V. 10, N 2. P. 494–504.

On probability of default in local volatility models

S.G. Shorokhov and A.E. Buuruldai
Nikol'skii Mathematical Institute of RUDN University, Moscow, Russia

Structural credit risk models [1] are based on constant volatility Black-Scholes pricing model [2], where the dynamics for the value of the firm V_t in risk-neutral framework is described by stochastic differential equation (SDE)

$$dV_t = r\,V_t\,dt + \sigma\,V_t\,dW_t,\; V\left(0\right) = V_0 > 0 \tag{1}$$

with risk-free interest rate $r > 0$ and constant volatility $\sigma > 0$, W_t – standard Wiener process. In Black-Scholes model (1) the value of the firm V_t is distributed lognormally and

$$V_t = V_0\,e^{\left(r-\frac{\sigma^2}{2}\right)\,t+\sigma\,W_t}. \tag{2}$$

According to Merton default model [3] with debt face value K and maturity date T the default event D is determined as $D = \{V_T < K\}$ and probability of default PD is equal to

$$PD = \mathbb{P}\left[V_T < K\right] = \Phi\left(-\frac{1}{\sigma\sqrt{T}}\left[\ln\frac{V_0}{K} + \left(r-\frac{\sigma^2}{2}\right)\,T\right]\right), \tag{3}$$

where $\mathbb{P}$ is a risk-neutral probability measure.

But an assumption of constant volatility in Black-Scholes model fails to hold in practice because of the so-called ”volatility smiles” and fat tails of financial data distributions. One of the most natural approaches to these issues is the transition to general model of risk-neutral dynamics

$$dV_t = r\,V_t\,dt + \sigma\left(V_t,\, t\right)\,V_t\,dW_t,\; V\left(0\right) = V_0 > 0 \tag{4}$$

with volatility σ being a function of firm value V_t and time t. Although the advantages of local volatility models with SDE (4) are well recognized, there exist a limited number of volatility functions $\sigma\left(V_t,\, t\right)$ which admit closed form solutions of (4) for firm value V_t.

We are interested in models when the firm value V_t can be represented as a function of the standard Wiener process W_t and time t $V_t = \Psi\left(W_t,\, t\right)$, examined in [4]. This case is realized in shifted lognormal model [5]

$$dV_t = r\,V_t\,dt + \sigma\,\left(V_t - \alpha\,e^{r\,t}\right)\,dW_t,\; V\left(0\right) = V_0 > 0, \tag{5}$$

where the firm value V_t is distributed lognormally and

$$V_t = \alpha\,e^{r\,t} + \left(V_0 - \alpha\right)\,e^{\left(r-\frac{\sigma^2}{2}\right)\,t+\sigma\cdot W_t}. \tag{6}$$

In normal model, introduced by J. Cox and S. Ross [6],

$$dV_t = r\,V_t\,dt + \sigma\,dW_t,\; V\left(0\right) = V_0 > 0, \tag{7}$$

the firm value V_t is distributed normally and

$$V_t = V_0\, e^{r\,t} + \sigma\sqrt{\frac{e^{2r\,t}-1}{2\,r\,t}}\, W_t. \tag{8}$$

In hyperbolic-sine model

$$dV_t = r\,V_t\,dt + \sqrt{\sigma^2 + 2\,r\,V_t^2}\,dW_t,\; V\left(0\right) = V_0 > 0, \tag{9}$$

the firm value V_t can be represented as a function of W_t only

$$V_t = \frac{\sigma}{\sqrt{2\,r}}\sinh\left(\operatorname{arsinh}\left(\frac{\sqrt{2\,r}}{\sigma}V_0\right) + \sqrt{2\,r}\,W_t\right). \tag{10}$$

When the firm value V_T is a function of W_T and T, the probability of default PD can be easily calculated as

$$\begin{aligned} PD &= \mathbb{P}\left[V_T < K\right] = \mathbb{P}\left[\Psi\left(W_T,\, T\right) < K\right] = \\ &= \mathbb{P}\left[W_T < \Psi^{-1}\left(K,\, T\right)\right] = \Phi\left(\frac{1}{\sqrt{T}}\,\Psi^{-1}\left(K,\, T\right)\right) \end{aligned} \tag{11}$$

Then for shifted lognormal model (6) PD is equal to

$$PD = \Phi\left(-\frac{1}{\sigma\sqrt{T}}\left(\ln\frac{V_0-\alpha}{K-\alpha\, e^{r\,T}} + \left(r - \frac{\sigma^2}{2}\right)T\right)\right), \tag{12}$$

for normal (Cox-Ross) model (8) PD is equal to

$$PD = \Phi\left(\frac{K - V_0\, e^{r\,T}}{\sigma}\sqrt{\frac{2\,r}{e^{2r\,T}-1}}\right), \tag{13}$$

and for hyperbolic-sine model (10) PD is equal to

$$PD = \Phi\left(\frac{1}{\sqrt{2\,r\,T}}\left[\operatorname{arsinh}\left(\frac{\sqrt{2\,r}}{\sigma}K\right) - \operatorname{arsinh}\left(\frac{\sqrt{2\,r}}{\sigma}V_0\right)\right]\right). \tag{14}$$

Let us plot probability of default curves for the following common parameters

$$V_0 = 100,\; K \in \left[50,\, 170\right],\; T = 1,\; r = 7\%,\; \sigma = 30\%,\; \alpha = 45.$$

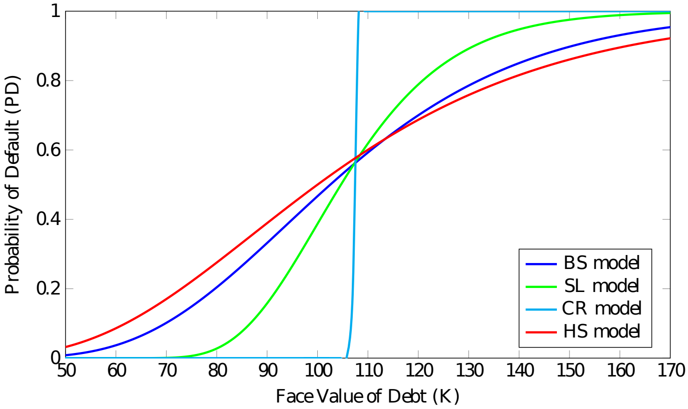


Fig. 1. Comparison of probabilities of default.

From fig. 1 it follows that PD for local volatility models differs from PD in Black-Scholes model, which gives us an opportunity to select most appropriate model of firm value dynamics for credit risk modelling.

References

1. Shorokhov S.G. Introduction into quantitative models of credit risk valuation. Moscow: RUDN, 2018. [In Russian]
2. Black F., Scholes M. The Pricing of Options and Corporate Liabilities // Journal of Political Economy. 1973. V. 81, N 3. P. 637–654.
3. Merton R. On the Pricing of Corporate Debt: The Risk Structure of Interest Rates // Journal of Finance. 1974. V. 29, N 2. P. 449–470.
4. Carr P., Tari M., Zariphopoulou T. Closed form option valuation with smiles // Preprint. NationsBanc Montgomery Securities, 1999.
5. Brigo D., Mercurio F. Fitting volatility skews and smiles with analytical stock-price models, Seminar Paper at Institute of Finance, University of Lugano, 2000.
6. Cox J.C., Ross S.A. The valuation of options for alternative stochastic processes // Journal of Financial Economics. 1976. V. 3, N 1-2. P. 145–166.

Application of mean field games approximation to economic processes modeling

N.V. Trusov

Lomonosov Moscow State University, Moscow, Russia

This work relies on the results in [1], and is its extension.

We present a trading executional model of professional trader and retail traders on financial market. On financial market we have the main investors that hold the asset shares with a view to a further growth of the asset share price, also known as long-term investors, and traders that are trying to maximize the revenue resulting from buying and selling blocks of asset shares. These traders are divided into two types of traders: a big trader, also known as professional trader, that can predict the further dynamic of asset share by analyzing the considered financial market, and a multitude of retail traders. Retail traders are non-professional traders that are acting similarly, trying to avoid abrupt changes.

In this abstract we consider the following problem. Professional trader has an initial budget $P_0 > 0$ at time $t_0 = 0$. By time $T > 0$ he wants to maximize his budget by entering the financial market and operating the blocks of asset shares. The behaviour of professional trader is described by Hamilton–Jacobi–Bellman equation evolving backward in time, and the behaviour of the retail traders is described by Fokker–Planck equation evolving forward in time. Coupling these equations we receive the mean field game problem.

We assume that the behaviour of the retail traders on the given time interval $[0, T]$ can be described by the utility function, presented in [1]. According to it, we present a value function:

$$u(t,x) = \max_{\alpha \in \mathcal{A}_1} \mathbb{E}\left(\int\limits_t^T \left(\ln m(\tau, x(\tau)) - \frac{1}{2}\alpha^2(\tau) - \lambda x^2(\tau) \right) d\tau \,\middle|\, \begin{smallmatrix} x(t)=x \\ \alpha(t)=\frac{dx}{dt} \end{smallmatrix} \right) \tag{1}$$

where $m(t,x)$ is a probability density function of the retail traders, $x(t) \in \mathbb{R}$ is the amount of asset shares held by retail traders at time $t \in [0,T]$, $x \in \mathbb{R}, t \in [0,T]$, $\alpha \in \mathcal{A}_1$ is a trading rate of the retail traders (real stochastic process $\gamma(t) \in \mathcal{A}_1$, if and only if $\mathbb{E}\left(\int\limits_0^T \gamma^2(t) dt \right) < \infty$), $\alpha(t,x) = \dfrac{\partial u}{\partial x}$, $\lambda > 0$. The coupled system of PDEs is:

$$\begin{cases} \dfrac{\partial u}{\partial t}(t,x) + \dfrac{\sigma^2}{2}\dfrac{\partial^2 u}{\partial x^2}(t,x) + \dfrac{1}{2}\left(\dfrac{\partial u}{\partial x}(t,x)\right)^2 - \lambda x^2 = -\ln m(t,x), \\ \dfrac{\partial m}{\partial t}(t,x) + \dfrac{\partial}{\partial x}\left(\dfrac{\partial u}{\partial x}(t,x) m(t,x)\right) - \dfrac{\sigma^2}{2}\dfrac{\partial^2 m}{\partial x^2}(t,x) = 0, \\ u(T,x) = -\theta(x-a)^2, \\ m(0,x) = m_0(x), \end{cases} \tag{2}$$

where $x \in \mathbb{R}, t \in [0,T]$, $m(t,x)$ is a probability density function of the retail traders by the amount of asset shares, $u(t,x)$ is defined in (1), $\theta \geqslant 0$, $a \in \mathbb{R}$ are given parameters.

By the assumptions that the density function of the retail traders has a Gaussian probability distribution of mean μ_0 and dispersion δ_0^2, the PDEs system (2) can be reduced to ODEs system

$$\begin{cases} u(t,x) = C_0(t) + C_1(t)x + C_2(t)x^2, \\ m(t,x) = \exp\left[D_0(t) + D_1(t)x + D_2(t)x^2\right], \\ \alpha(t,x) = C_1(t) + 2C_2(t)x, \end{cases} \tag{3}$$

where $x \in \mathbb{R}, t \in [0,T]$, and the functions $C_0(t)$, $C_1(t)$, $C_2(t)$, $D_0(t)$, $D_1(t)$ and $D_2(t)$ are the solutions of the six Riccati equations:

$$\begin{cases} \dot{D}_0 = \dfrac{\sigma^2}{2}D_1^2 - C_1D_1 + \sigma^2 D_2 - 2C_2, \\ \dot{D}_1 = 2\sigma^2 D_1D_2 - 2C_2D_1 - 2C_1D_2, \\ \dot{D}_2 = 2\sigma^2 D_2^2 - 4C_2D_2, \\ \dot{C}_0 = -D_0 - \dfrac{C_1^2}{2} - \sigma^2 C_2, \\ \dot{C}_1 = -D_1 - 2C_1C_2, \\ \dot{C}_2 = -D_2 - 2C_2^2 + \lambda, \end{cases} \tag{4}$$

with initial conditions $D_0(0) = -\dfrac{\mu_0^2}{2\delta_0^2} - \dfrac{1}{2}\ln\left(2\pi\delta_0^2\right)$, $D_1(0) = \dfrac{\mu_0}{\delta_0^2}$, $D_2(0) = -\dfrac{1}{2\delta_0^2}$, $C_0(T) = -a^2\theta$, $C_1(T) = 2a\theta$ and $C_2(T) = -\theta$; μ_0 is the amount of asset shares of retail traders at $t = 0$, $\mu_0 \in \mathbb{R}$, a is the amount of asset shares the retail traders want to achieve by time T, $a \in \mathbb{R}$ $\delta_0^2 > 0$, $\sigma^2 > 0$, $\lambda > 0$, $\theta \geqslant 0$ are given parameters. The function

$D_2(t)$ is negative at time $t \in [0,T]$, so it guarantees that the density function of the retail traders has a Gaussian probability distribution.

We assume that the asset share price satisfies the following dynamic equation:

$$S(t) = S^0(t) + \eta_1 M(t) + \eta_2 \beta(t), \tag{5}$$

where $t \in (0,T]$, $M(t) = \int\limits_{\mathbb{R}} \alpha(t,x)m(t,x)dx$, $\eta_1 > 0$, $\eta_2 > 0$, $S^0(t)$ represent the asset share price in absence of trading. We consider a step function for $S^0(t)$ that has a jump on h units at time $\tau \in (0,T)$.

It is obvious to assume that at time $t = 0$ professional trader has no asset shares, and by time T he wants to have a zero asset shares as well. We get the following optimal control problem for the optimal strategy of professional trader:

$$\begin{cases} \dot{y}(t) = \beta(t), \\ \dot{P}(t) = -\left(S^0(t) + \eta_1 \left(C_1(t) - C_2(t) \cdot \dfrac{D_1(t)}{D_2(t)}\right) + \eta_2\beta(t)\right)\beta(t), \end{cases} \tag{6}$$

where $t \in [0,T]$, $y(t)$ represents the amount of asset shares of professional trader at time $t \in [0,T]$, $\beta(t)$ is a trading rate of professional trader at time $t \in [0,T]$ ($y(0) = y(T) = 0$), $P(t)$ in a function that represents the amount of budget left by professional trader at time $t \in [0,T]$ (we assume that $P(t) > 0$, $t \in [0,T]$, $P(0) = P_0 > 0$). The system (6) includes the functions $C_1(t)$, $C_2(t)$, $D_1(t)$, $D_2(t)$ defined in (4) (see (3)). The goal is to maximize the budget of a retail trader at time $T > 0$, so we can write the functional to the system (6) as

$$J = P(T) \to \max. \tag{7}$$

Applying Pontryagin's maximum principle [2], we conclude that the optimal control β^* of professional trader is defined by:

$$\beta^* = \frac{1}{2\eta_2}\left(C - S^0(t) - \eta_1 C_1(t) + \eta_1 C_2(t)\frac{D_1(t)}{D_2(t)}\right), \tag{8}$$

where C is a constant unique defined after substitution the optimal control (8) into system (6).

Thus, the dynamic equation of asset share price (5) becomes:

$$S(t) = \frac{1}{2}\left(S^0(t) + C + \eta_1 C_1(t) - \eta_1 C_2(t)\frac{D_1(t)}{D_2(t)}\right). \tag{9}$$

Note that the obtained equation (9) does not depend on η_2, so the activity of professional trader does not impact on the asset share price.

Let $K(t), t \in [0,T]$ be a function that represents the budget of retail traders, $K(0) = K_0 > 0$, and defined by $\dot{K} = -S(t) \int\limits_{\mathbb{R}} \alpha(t,x)m(t,x)dx$. To analyze whether retail traders' revenue is positive or not, it is sufficient to analyze the sign of F, introduced in the following equation:

$$F = K(T) + x(T) \cdot S(T) - (K(0) + x(0) \cdot S(0)) \,. \tag{10}$$

Let $\tilde{P}$ represent the amount of invested money of professional trader. Introducing the ratio $\dfrac{P(T) - P(0)}{\tilde{P}}$, and dividing it by the ratio that shows the profit according to invested money, i.e. $\frac{\max\limits_{t\in[0,T]} S(t) - \min\limits_{t\in[0,T]} S(t)}{\min\limits_{t\in[0,T]} S(t)}$, we can define the profit coefficient of professional trader:

$$PC_{BT} = \frac{(P(T) - P(0)) \cdot \min\limits_{t\in[0,T]} S(t)}{\tilde{P}\left(\max\limits_{t\in[0,T]} S(t) - \min\limits_{t\in[0,T]} S(t)\right)}. \tag{11}$$

As an example, we consider the following parameters: $t_0 = 0$; $T = 6$; $\eta_1 = 4$; $\eta_2 = 0.5$; $\lambda = 1.6$; $\sigma = 3$; $P_0 = 1000$; $K_0 = 100$; $\mu_0 = 2$; $\delta_0 = 0.8$; $a = 3$; $\theta = 2.2$; $S_0 = 20$; $h = 10$; $\tau = 1$. The profit coefficient of professional trader: 0.47. The parameter F, defined in (10) is 16.97. Therefore, the retail traders make a profit due to the right forecast of the asset share price. The profit of professional trader is 170.29. In the following graphs we can observe the dependence of a budget of professional trader and his amount of asset shares:

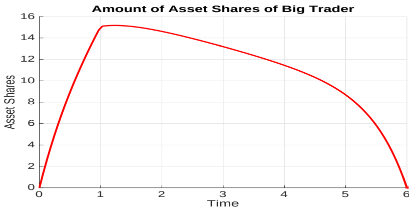


Fig. 1. Asset shares dependence of professional trader on time

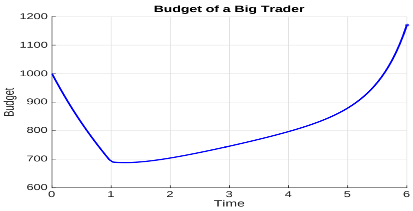


Fig. 2. Budgetary dependence of professional trader on time

If we fix the introduced parameters, and will vary only the parameters $\mu_0 \in [-2, 2]$ and $a \in [-2, 2]$, we can observe how the profit coefficient of a big trader and the profit of retail traders depend on the initial amount of asset shares μ_0 of the retail traders and the amount of asset shares the retail traders want to achieve by time T:

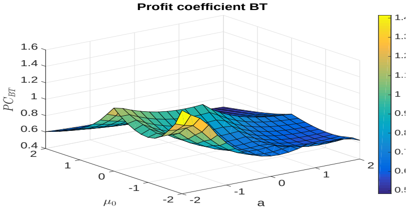


Fig. 3. Profit coefficient dependence of professional trader on (μ_0, a)

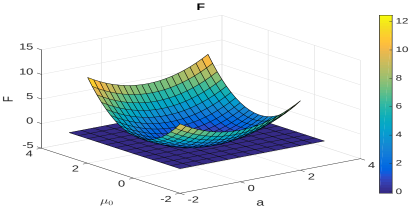


Fig. 4. Profit dependence of retail traders on (μ_0, a)

There is also a markable situation, when professional trader enters the short positions twice. Lets consider the following parameters: $t_0 = 0$; $T = 5$; $\eta_1 = 5$; $\eta_2 = 0.2$; $\lambda = 0.5$; $\sigma = 2$; $P_0 = 100$; $K_0 = 10$; $\mu_0 = -1$; $\delta_0 = 0.8$; $a = 1$; $\theta = 1$; $S_0 = 10$; $h = 3$; $\tau = 0.5$. In the following graphs we can observe the amount of asset shares of professional trader and the asset share price:

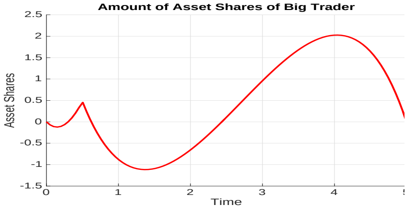


Fig. 5. Asset shares dependence of professional trader on time

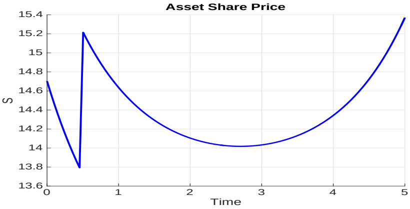


Fig. 6. Asset share price dependence on time

References

1. Fatone, L., Mariani, F., Recchioni, M.C. and Zirilli, F. (2014) A Trading Execution Model Based on Mean Field Games and Optimal Control. Applied Mathematics, 5, 3091-3116.
2. Pontryagin L.S., Boltyansky V.G., Gamkrelidze R.V., Mishchenko E.F. Mathematical theory of optimal processes (in Russian). Moscow: Nauka, 1983.

OR in insurance and risk-management

On sufficient conditions for survival probability in the life annuity insurance model with risk-free investment income

T.A. Belkina[1] and N.B. Konyukhova[2]
[1]*Central Economics and Mathematics Institute of RAS,*
[2]*Dorodnicyn Computing Center FRC CSC of RAS,*
Moscow, Russia

We study the life annuity insurance model with random revenues (see, e.g., [1]) when the whole surplus of the insurer is invested continuously in a risk-free asset. For the survival probability (SP) as a function of the initial surplus (IS) in this model, some associated singular problem for linear integro-differential equation (IDE) is formulated and studied in [2]*. In particular, for the case of exponential distribution of revenue sizes, the existence and uniqueness of the solution to this problem are stated. Here we prove the sufficiency theorem asserting that the solution of associated singular problem for IDE determines the SP in the original insurance model. For the classical collective risk model with investments, a similar approach was used in [4]; the corresponding existence results see, e.g., in [5].

The life annuity insurance model may be considered as dual to the classical non-life collective risk model and is called also the dual risk model (see, e.g., [6]). The surplus or equity of a company (in the absence

*In [2] the case of risky investments is also considered; for this case see [3] as well.

of investments) in the dual risk model is of the form

$$R_t = u - ct + \sum_{k=1}^{N(t)} Z_k, \quad t \geqslant 0. \tag{1}$$

Here R_t is the surplus of a company at time $t \geq 0$; u is the IS, $c > 0$ is the rate of expenses (total pension payments per unit time) assumed to be deterministic and fixed; $N(t)$ is a homogeneous Poisson process with intensity $\lambda > 0$ that, for any $t > 0$, determines the number of random revenues up to the time t; Z_k $(k = 1, 2, ...)$ are independent identically random variables with a distribution function $F(z)$ $(F(0) = 0$, $\mathbf{E}Z_1 = m < \infty, m > 0)$ that determine the revenue sizes and are assumed to be independent of $N(t)$; these revenues arise at the moments of the death of policyholders.

Let now the whole surplus be invested in a risk-free asset which evolves as $dB_t = rB_t\,dt$, $t \geqslant 0$, where $r > 0$ is the interest rate.

Then the resulting surplus process X_t is governed by the equation

$$dX_t = r\,X_t\,dt + dR_t, \quad t \geqslant 0. \tag{2}$$

with the initial condition $X_0 = u$, where R_t is defined by (1).

Denote $\varphi(u) = P\,(X_t \geq 0, \;\; t \geqslant 0)$ the SP (i.e., the probability that bankruptcy will never happen). It is easy to see that for the SP of process (2) the following relations are fulfilled:

$$\varphi(0) = 0,$$

and

$$\varphi(u) = 1, \;\; u \geqslant c/r,$$

i.e., a ruin will never occur if IS $u \geqslant c/r$.

Definition 1. Let $\mathcal{L}$ be the class of functions $\varphi(u)$ defined on $[0, \infty)$, continuously differentiable on $(0, c/r)$ and satisfying conditions

$$\varphi(0) = 0, \qquad \lim_{u \to c/r - 0} \varphi(u) = 1, \qquad \varphi(u) = 1, \;\; u \geqslant c/r. \tag{3}$$

The infinitesimal generator $\mathcal{A}$ of the process X_t has the form

$$(\mathcal{A}f)(u) = f'(u)[ru - c] - \lambda f(u) + \lambda \int_0^\infty f(u+z)\,dF(z), \tag{4}$$

for any function f from a class $\mathcal{L}$ [2].

Lemma 1 [2]. *Let all the parameters in* (4) *be fixed numbers, where* $c > 0,\ \lambda > 0,\ r > 0$. *Let the function* $\varphi \in \mathcal{L}$ *be satisfying IDE*

$$(\mathcal{A}\varphi)(u) = 0 \tag{5}$$

for all $u > 0$ *(perhaps, with exception of the point* $u = c/r$*). Then:* 1) *this solution is unique in* $\mathcal{L}$*;* 2) *the solution* $\varphi(u)$ *satisfies restrictions* $0 \leqslant \varphi(u) \leqslant 1,\ u \in \mathbb{R}_+$.

Theorem 1. *Let all the conditions of Lemma 1 be fulfilled. Then, for arbitrary* $u \in \mathbb{R}_+$*, the value* $\varphi(u)$ *of the function, defined in Lemma 1, is SP for the process* (2) *with initial state* $X_0 = u$.

Proof. For any $\epsilon > 0$ choose a positive constant $\delta < c/2r$, such that $\varphi(u) \leq \epsilon$ for $u \in [0, \delta]$, $\varphi(u) \geq 1 - \epsilon$ for $u \in [c/r - \delta, c/r]$. Then change $\varphi(u)$ to $\varphi_\delta(u)$ on $[0, \infty)$ and extend it on $(-\infty, 0)$, such that $\varphi_\delta(u) = 0$ for $u \in (-\infty, -\delta)$, $\varphi_\delta(u) = 1$ for $u \in [c/r + \delta, \infty)$, $\varphi_\delta(u) = \varphi(u)$ for $u \in [\delta, c/r - \delta]$, $\varphi_\delta(u)$ is continuously differentiable on $\mathbb{R}$ and $0 \leq \varphi_\delta(u) \leq \epsilon$ for $u \in [-\delta, \delta]$, $1 - \epsilon \leq \varphi_\delta(u) \leq 1$ for $u \in [c/r - \delta, c/r + \delta]$. By Lemma 1 and in accordance with the construction of the function $\varphi_\delta(u)$, it satisfies restrictions

$$0 \leqslant \varphi_\delta(u) \leqslant 1,\ u \in \mathbb{R}_+. \tag{6}$$

For the process X_t with initial state $u \in (\delta, c/r - \delta)$, define exit time $\tau_\delta = inf\{t \geq 0 : X_t \notin (\delta, c/r - \delta)\}$. Then, for $u \in (\delta, c/r - \delta)$, by Ito's Lemma (see, e.g., [7]) we have

$$\varphi_\delta(X_{t\wedge\tau_\delta}) = \varphi(u) + \int_0^{t\wedge\tau_\delta} (\mathcal{A}\varphi)(X_s)ds + M_t, \tag{7}$$

where

$$\begin{aligned} M_t \quad &= \quad \sum_{i=1}^{N(t\wedge\tau_\delta)} [\varphi_\delta(X_{\theta_i}) - \varphi(X_{\theta_i -})] \\ &\quad + \lambda \int_0^{t\wedge\tau_\delta} [\varphi(X_s) - \mathbf{E}(\varphi(X_s + Z))]ds, \end{aligned}$$

$\theta_i,\ i = 1, 2, ...,$ are the times of revenue arrivals and the expectation symbol corresponds to distribution of revenue size Z.

Since φ satisfies the equation (5), $u > 0$, we have in (7) $(\mathcal{A}\varphi)(X_s) = 0$. Further, M_t is martingale, hence, taking expectation in both sides of (7), we obtain

$$\mathbf{E}(\varphi_\delta(X_{t\wedge\tau_\delta})) = \varphi(u), \quad u \in (\delta, c/r - \delta).$$

Since $\varphi_\delta(x)$ is bounded, by dominated convergence, we obtain, letting $t \to \infty$:

$$\mathbf{E}(\varphi_\delta(X_{\tau_\delta})) = \varphi(u), \quad u \in (\delta, c/r - \delta). \tag{8}$$

We write $X_t(u)$ for the process (2) with IS u and $\tau(u)$ for the ruin time of this process, i.e., $\tau(u) = \inf\{t \geq 0 : X_t(u) < 0\}$. Note that, for $\varepsilon > 0$, on the set $\{\tau(u) = \infty\}$

$$X_t(u + \varepsilon) \geq \varepsilon,\ t \geq 0, \quad \text{and} \quad X_t(u + \varepsilon) \to \infty,\ t \to \infty. \tag{9}$$

Indeed, for the two processes defined in (2) with initial states u and $u+\varepsilon$, it is easy to see that $X_t(u + \varepsilon) - X_t(u) = \varepsilon e^{rt}$, whence follows (9).

Now fix $\varepsilon > 0$, such that $u + \varepsilon \in (\delta, c/r - \delta)$. In view of (8) and the non-negativity of φ_δ, we have

$$\varphi(u+\varepsilon) = \mathbf{E}[\varphi_\delta(X_{\tau_\delta(u+\varepsilon)}(u+\varepsilon))] \geq \mathbf{E}[\varphi_\delta(X_{\tau_\delta(u+\varepsilon)}(u+\varepsilon))\mathbf{I}\{\tau(u) = \infty\}].$$

Taking into account that on the set $\{\tau(u) = \infty\}$ the relations (9) are satisfied, we conclude that, for sufficiently small δ, the exit time $\tau_\delta(u+\varepsilon)$ coincides on this set with the crossing moment of the level $c/r-\delta$. Hence, $X_{\tau_\delta(u+\varepsilon)}(u+\varepsilon) \geq c/r-\delta$ and $\varphi_\delta(X_{\tau_\delta(u+\varepsilon)}(u+\varepsilon)) \geq 1-\epsilon$. Consequently, $\varphi(u + \varepsilon) \geq (1 - \epsilon)\mathbf{P}(\tau(u) = \infty)$. Then, letting $\epsilon \to 0$, $\varepsilon \to 0$, we obtain that

$$\varphi(u) \geq \mathbf{P}(\tau(u) = \infty), \quad u \in (0, c/r). \tag{10}$$

On the other hand, for $u \in (\delta, c/r - \delta)$, we have

$$\begin{aligned}\varphi(u) = \mathbf{E}[\varphi_\delta(X_{\tau_\delta(u)}(u))\mathbf{I}\{X_{\tau_\delta(u)}(u) > \delta\}]+\\ +\mathbf{E}[\varphi_\delta(X_{\tau_\delta(u)}(u))\mathbf{I}\{X_{\tau_\delta(u)}(u) \leq \delta\}] \leq\\ \leq \mathbf{E}[\varphi_\delta(X_{\tau_\delta(u)}(u))\mathbf{I}\{X_{\tau_\delta(u)}(u) \geq c/r - \delta\}]+\\ +\epsilon\mathbf{P}(X_{\tau_\delta(u)}(u) \leq \delta).\end{aligned}$$

Letting $\epsilon \to 0$ (then $\delta \to 0$), we obtain, in view of non-negativity of jumps of the process X_t and the validity of (6), that

$$\varphi(u) \leq \mathbf{P}(X_{\tau_0(u)}(u) \geq c/r) \leq \mathbf{P}(\tau(u) = \infty), \quad u \in (0, c/r). \tag{11}$$

Finally, (10) and (11) imply the equality $\varphi(u) = \mathbf{P}(\tau(u) = \infty)$ for $u \in (0, c/r)$. For the other values of u, this equality is obvious. Thus, for $u \geq 0$, $\varphi(u)$ is SP for the process (2) with initial state $X_0 = u$. Theorem 1 is proved.

As an example of the application of Theorem 1 we consider the case of an exponential revenue sizes. In [2] the following result is established.

Theorem 2. *Let* $F(z) = 1 - \exp(-z/m)$*, all the parameters* r*,* m*,* c*,* λ *be fixed positive constants. Then the following assertions hold:*

(I) *in the class* $\mathcal{L}$ *there exists a solution to IDE* (5) *(it satisfies* (5)*, perhaps, with exception of the point* $u = c/r$*);*

(II) *this solution is unique in* $\mathcal{L}$ *and, on the interval* $[0, c/r)$*, it has the form*

$$\varphi(u) = 1 - \int\limits_u^{c/r} \psi(s)\, ds. \tag{12}$$

where $\psi(u)$ *is defined by the formula*

$$\psi(u) = \left[\int\limits_0^{c/r} (c/r - u)^{\lambda/r - 1} \exp(u/m)\, du\right]^{-1} (c/r - u)^{\lambda/r-1} \exp(u/m). \tag{13}$$

In accordance with Theorem 1 the value $\varphi(u)$ of the function, defined in Theorem 2, is the SP of the process (2) with initial state $X_0 = u$.

We use the approach based on so called sufficiency theorem for SP and the existence theorem for the corresponding singular problem for IDE. This unified approach eliminates need to proof regularity of the survival probability (in details for other models see [4]).

Note that for $\lambda \leqslant r$ the solution of IDE (5), defined in Theorem 2, is *non-smooth* function. This function may be considered as *viscosity solutions* of IDE (5). The uniqueness theorem for a viscosity solution (it is formulated in [8] for more general model) in application to the problem considered here is an alternative tool to prove the fact that the function defined above determines the corresponding survival probability.

References

1. Grandell J. Aspects of Risk Theory. Berlin: Springer, 1991.
2. Belkina T.A., Konyukhova N.B., and Slavko B.V. Analytic-numerical investigations of singular problems for survival probability in the dual risk model with simple investment strategies // Analytical and Computational Methods in Probability Theory and Its Applications (V.V.Rykov et al. (Eds.): ACMPT 2017. Springer International Publishing AG 2017 / Series: Lecture Notes in Computer Science (LNCS), 2017. V. 10684. P. 236–250.

3. Kabanov Yu., Pergamenshchikov S. In the insurance business risky investments are dangerous: the case of negative risk sums // Finance Stochast. 2016. V. 20, N 2. P. 355–379.
4. Belkina T. Risky investment for insurers and sufficiency theorems for the survival probability // Markov Processes Relat. Fields. 2014. V. 20, N 3. P. 505–525.
5. Belkina T.A., Konyukhova N.B., and Kurochkin S.V. Dynamical insurance models with investment: Constrained singular problems for integrodifferential equations // Comput. Math. Math. Phys. 2016. V. 56, N 1. P. 43—92.
6. Albrecher H., Badescu A., and Landriault D. On the dual risk model with tax payments // Insurance Math. Econom. 2008. V. 42. P. 1086–1094.
7. Cont R., Tankov P. Financial Modelling with Jump Processes. Boca Raton: Chapman & Hall/CRC, 2004.
8. Belkina T., Kabanov Yu. Viscosity solutions of integro-differential equations for nonruin probabilities // Theory Probab. Appl. 2016. V. 60, N 4. P. 671–679.

Estimating the probability of company default based on system dynamics model

D.S. Kurennoy and D.Yu. Golembiovskiy
Lomonosov Moscow State University, Moscow, Russia

Nowadays a considerable number of mathematical methods have been developed to estimate the probability of borrower default. They are based on the analysis of quantitative and qualitative enterprise indicators values, [5]. However, most of these methods do not take into account the company structure, its dynamics when external factors are changed and presuppose a large sample of similar enterprises.

This work demonstrates a possibility of using system dynamic model [4, 7] to assess the probability of company default, which allows avoiding described shortcomings. In the system dynamic paradigm, an explored enterprise is represented in the form of continuously interacting elements and external factors. The links between the elements are described by functional dependencies and differential equations that determine compa- ny dynamics and its stability relative to various macroeconomic scenarios. The behavior of random macroeconomic variables in this research is described by ARIMA-GARCH and ARIMAX-GARCH models [1, 3], which are used in econometrics to predict non-stationary time series. The probability of company default is determined as a result of experiments over the obtained system dynamics model using the Monte Carlo simula- tion. The probability of default is a share of macroeconomic scenarios leading to the ruin of enterprise. The obtained numerical results are compared with estimations of the rating agencies Moody's and Fitch.

The system dynamic model of Bashneft [6], which has been producing since 1932 and is developing about 170 deposits, was built on the basis of the financial statements and information from other open sources for time period 2007-2015 years. The key element of this system is the "Ruble cash" stock. Its equality to zero means enterprise default. Prices of oil and oil products traded by the company, dollar-to-ruble rate, rate of attracted and redeemed loans, the basic rate of mineral extraction tax, unit costs of extraction, processing and general economic expenses were considered as external parameters that affect state of the model.

These external parameters are described using ARIMA-GARCH and ARIMAX-GARCH models. The basic concept of such models is to relate the current time series values with their previous values, taking into account random innovations and correlation between variables. Gene- ral view of ARIMAX(p, d, q) - GARCH(s, r) models is given by the equations:

$$(1-L)^d Y_t = c + (\sum_{i=1}^{p} a_i L^i)(1-L)^d Y_t + (\sum_{i=1}^{q} b_i L^i) e_t + w X_t,$$

$$e_t = \sigma_t z_t,$$

$$\sigma_t^2 = c_0 + \sum_{i=1}^{s} \gamma_i \sigma_{t-i}^2 + \sum_{i=1}^{r} \beta_i e_{t-i}^2,$$

where Y_t — considered time series; X_t — exogenous factor; $L : LY_t = Y_{t-1}$ — lagging operator; $a_i, b_i, \gamma_i, \beta_i, w$ — real numbers that are coeffici- ents of the model; p, q, d, s, r — natural numbers that determine the model order; $\{z_t\}$ — random processes of independent identically distribu- ted random variables; c, c_0 — constants. In this case, the ARIMA(p, d, q) - GARCH(s, r) model is obtained from the described one if we set w=0.

In order to assess the probability of default for oil producing and refining enterprise, company dynamics model was simulated from the second quarter of 2014 taking into account various scenarios of external macroeconomic parameters realized by the described ARIMA-GARCH and ARIMAX-GARCH models. The total number of experiments on the

Bernoulli scheme with two outcomes was 10,000. In each of them, fact of default or not default of the enterprise during different periods was fixed: one, two, three, four, five and ten years. As a result, the probability of default for each time interval was calculated by formula:

$$p = \frac{k}{n} \ ,$$

where k is number of experiments in which company has defaulted, n means total number of model runs. Then confidence intervals were constructed with confidence levels of 95% and 99% using the standard Clopper-Pearson method [2] for the binomial distribution. The results are presented in Table. 1. In this case, the probability of default indicates the percentage of cases that correspond to the enterprise default (it is defined as p 100).

In 2015, Bashneft was assigned a Ba1 rating by Moody's and BB+ by Fitch. Table 1 shows the average percentage of ruined companies with such rating for various periods (1-5 and 10 years), established on the basis of the data period 1983-1966 for Moody's and 1981-2015 for Fitch. Note that the average for the period 1983-2016 percentage of companies that went default during the year is below the same level for 2015. At the same time, the rating agencies' data differ by 1-1,5 points, which makes it possible to consider such a difference when comparing the results of modeling with the estimations of rating agencies as acceptable.

A comparative analysis of the obtained results and data from Moody's and Fitch demonstrates closeness of the simulated probability of company default and corresponding rating agencies estimates, which allows to conclude that the described approach is acceptable for estimating the probability of borrower default. In most cases, the model probability of default lies between the rating agencies estimates, and the confidence intervals found cover the level of the average percentage of ruined compa- nies with an error of 0.1-3 points. The most accurate are the results obtained in assessing the probability of default within one year, two, three and four years. Discrepancies that occur when considering forecas- ting periods of 5 and 10 years can be explained by the inadequate accuracy of ARIMA-GARCH for such large time intervals. They are incapable of taking into account the changes that have occurred in the market during this time. In addition, the information used in the construction of the Bashneft system-dynamic model is not exhaustively complete, since it is based only on the analysis of open sources. Note that banks have the ability to receive any data from their borrowers and thereby refine its system dynamic model.

T	L_{low}^{95}	L_{up}^{95}	L_{low}^{99}	L_{up}^{99}	p(%)	Moody's (1983-2016)	Fitch (1981-2015)
1 year	0.79	1.18	0.74	1.25	0.97	0.47	0.77
2 years	1.99	2.58	1.90	2.68	2.27	1.54	2.51
3 years	3.39	4.14	3.28	4.27	3.75	2.85	4.04
4 years	4.67	5.54	4.54	5.68	5.09	4.15	5.58
5 years	8.52	9.66	8.36	9.85	9.08	5.47	6.83
10 years	10.95	12.21	10.76	12.42	11.56	10.36	9.92

L_{low}^{95} , L_{up}^{95} – the lower and upper limits of the confidence interval with a confidence level of 95%.
L_{low}^{99} , L_{up}^{99} – the lower and upper limits of the confidence interval with a confidence level of 99%.
Moody's (1983-2016) - the average percentage of ruined companies for a different period (1-5 years and 10 years) and rated at Ba1.
Fitch (1981-2015) - the average percentage of ruined companies for a different period (1-5 years and 10 years) and rated at BB+.
Moody's (2015) - the percentage of companies ruined in 2015 with a rating of Ba1 is 0.905.

Table. 1. Estimation the probability of oil company default.

References

1. Bollerslev, T. Generalized autoregressive conditional heteroskedasti- city // Journal of Econometrics. - 1986. - N 31. - P. 309-328.
2. Clopper C. J. The use of confidence or fiducial limits illustrated in the case of the binomial / C. J. Clopper, E. S. Pearson // Biometrika. - 1934. - N 26. - P. 404-413.
3. Engle, R.F. Autoregressive conditional heteroskedasticity with esti- mates of the UK inflation // Econometrica. - 1982. - N 50. - P. 987-1008.
4. Forrester J. W. Urban Dynamics / Pegasus Communications. - 1969.
5. Gurny P., Gurny M. Comparison of credit scoring models on probability of default estimation for us banks // Prague economic papers. - 2013.

6. Kurennoy D.S., Golembiovskiy D.Yu. System dynamics credit risk model of an oil company // Issues of Risk Analysis. 2017. - N 01. - P. 6-22.
7. Sterman J.D. Business Dynamics: Systems Thinking and Modeling for a Complex World. Boston: McGraw-Hill Companies. - 2000.

Parameter estimation of ARIMA-GARCH model with variance-gamma distribution in financial series: expectation-maximization algorithm

D.S. Ogneva and D.Yu. Golembiovskiy
Lomonosov Moscow State University, Moscow, Russia

The financial time series of fixed income securities, stock indices and shares are characterized by the autocorrelation of their squared returns [1, 2]. In econometrics this feature is described by the heteroscedastic model ARCH [3, 4]. The ability to vary the returns distribution [5, 6], the model of conditional mean [7], and the ARCH form [8] allows to describe the real market price dynamics quite accurately.

This work consider ARIMA-GARCH model with variance-gamma distribution:

$$\begin{aligned}
\Delta^d y_t = m_t + \sqrt{h_t}\epsilon_t, &\quad \epsilon_t \sim VG(\theta,\mu,\sigma,\tau)\\
m_t = \mathbb{E}[\Delta^d y_t|\mathcal{F}_{t-1}] &= \phi_0 + \phi_1\Delta^d y_{t-1} + \phi_2\Delta^d y_{t-2} + \cdots + \phi_{\check{p}}\Delta^d y_{t-\check{p}} +\\
&+ \theta_1\sqrt{h_{t-1}}\epsilon_{t-1} + \theta_2\sqrt{h}_{t-2}\epsilon_{t-2} + \cdots + \theta_Q\sqrt{h}_{t-\check{q}}\epsilon_{t-\check{q}},\\
h_t = \mathbb{D}[\Delta^d y_t|\mathcal{F}_{t-1}] &= \alpha_0 + \alpha_1 h_{t-1} + \alpha_2 h_{t-2} + \cdots + \alpha_p h_{t-p} +\\
&+ \beta_1 h_{t-1}\epsilon_{t-1}^2 + \beta_2 h_{t-2}\epsilon_{t-2}^2 + \cdots + \beta_q h_{t-q}\epsilon_{t-q}^2,
\end{aligned}$$

and normalization restrictions:

$$\mathbb{E}[\epsilon_t] = \theta + \mu\tau = 0,\ \mathbb{D}[\epsilon_t] = \tau(\mu^2+\sigma^2) = 1,$$

where $y_t = s_t/s_{t-1}$ — considered returns of financial time series s_t; $\Delta x_t = x_t - x_{t-1}$ — difference operator, d — its order such that $\Delta^d y_t$ is the first second-order stationary time series; m_t and h_t — conditional mean and variance respectively; ϵ_t — independent standard variance-gamma random values, (θ,μ,σ,τ) — variance-gamma distribution parameters; $\phi_0, ..., \phi_{\check{p}}$ and $\theta_1, ..., \theta\hat{q}$ — ARIMA parameters, $\alpha_0, ..., \alpha_p$ and $\beta_1, ...\beta_q$ — GARCH parameters, $(\hat{p}, d, \hat{q})$ and (p, q) — ARIMA and

GARCH orders respectively. That is, $(\hat{p}+\hat{q}+1)$ parameters for ARIMA, $(p+q+1)$ parameters for GARCH and 2 parameters for variance-gamma distribution (because of two restrictions); denote, Θ — all of them.

Due to skewness and kurtosis variance-gamma distribution [9, 10] closely approximates financial series. However, its density has not got a simple form formula. So, classical approach to tuning the model coefficients solving the likelihood maximization problem for variance-gamma densities product is inconvenient to use. This implies numerical issues for parameters estimation.

Variance-gamma distribution can be defined as a normal variance-mean mixture where the mixing density is a gamma distribution [11]. By discretizing the second one [12], the first one is represented as a mixture of normal, so an expectation-maximization algorithm can be applied [13]. Expectation step computes expected value of complete log-likelihood function with respect to mixing variable w_k (gamma discretization) with densities π_k and missing values z_t (mixing components number) with their densities:

$$g_{tk} = p(z_{tk} = 1|y_t, \Theta^{old}) = \frac{\pi_k(\tau^{old}) \cdot \phi\big(y_t|\Theta^{old}, w_k\big)}{\sum\limits_{k=1}^{K} \pi_k(\tau^{old}) \cdot \phi\big(y_t|\Theta^{old}, w_k\big)},$$

$$z_t \in \{0,1\}^K : \sum_{k=1}^{K} z_{tk} = 1; \quad \pi_k(\tau) = \frac{w_k^{\tau-1} e^{-w_k}}{\sum\limits_{i=1}^{K} w_i^{\tau-1} e^{-w_i}},$$

$$\log \phi(y_t|\Theta, w_k) = -\frac{(y_t - (m_t + \theta\sqrt{h_t} + \mu\sqrt{h_t} w_k))^2}{2\sigma^2 h_t w_k} - \log(\sigma\sqrt{h_t w_k}),$$

where index *old* means last iteration parameters values. Maximization step maximizes the quantity $Q(\Theta, \Theta^{old})$ obtained in expectation step with respect to Θ:

$$\begin{aligned}\Theta^{new} &= \arg\max_{\Theta} Q(\Theta, \Theta^{old}) = \\ &= \arg\max_{\Theta} \left[\sum_{t=1}^{T} \sum_{k=1}^{K} g_{tk} \Bigg(\log \pi_k(\tau) + \log \phi(y_t|\Theta, w_k) \Bigg) \right]\end{aligned}$$

To verify the correctness of the obtained algorithm, experiments were performed with daily data on the SP500 index from 01.01.2016 to 01.10.2017. The efficiency of the constructed model was considered in comparison with normal (norm), Student-t (std), normal inverse gaussian (nig), generalized hyperbolic (ghyp) and Johnson's SU (jsu) distributions implemented in "rugarch" package in language R. The results are presented in Table 1. All the information criteria (consistent Akaike, or cAIC; Bayesian or BIC; Hannan-Quinn, or HQIC) showed that the quality of the variance-gamma (vg) model is comparable to the quality of normal inverse gaussian and much better than normal and Student-t distributions, but still slightly worse than generalized hyperbolic. Two mentioned models are generalized hyperbolic sub-families, its limiting forms, therefore, having more attractive distribution functions, they lose in accuracy of parameters selection.

Table 1. Information criteria for the compared models

	vg	norm	std	nig	ghyp	jsu
cAIC	3218.1	3261.2	3231.4	3219.0	3212.7	3226.0
BIC	3230.1	3270.2	3241.4	3230.0	3224.7	3237.0
HQIC	3258.5	3292.4	3266.1	3257.2	3254.4	3264.2

References

1. Heston S. A closed-form solution for options with stochastic volatility with applications to bond and currency options // The Review of Financial Studies. 1993. V. 6, N 2. P. 327–343.
2. Heston S., Nandi S. A closed-form GARCH option pricing model // The Review of Financial Studies. 2000. V. 13, N 3. P. 585–626.
3. Engle R. Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation // Econometrica. 1982. V. 50. N 4. P. 987–1008.
4. Bollerslev T. Generalized autoregressive conditional heteroskedasticity // Journal of Econometrics. 1986. V. 31, N 3. P 307–327.
5. Rossi E. Univariate GARCH models: a survey // Quantile. 2010. V. 8, N 8. P. 1–67.
6. Akgiray V. Conditional heteroscedasticity in time series of stock return: evidence and forecasts // The Journal of Business. 1989. V. 62, N 1. P. 55–80.
7. Xi Y. Comparison of option pricing between ARMA-GARCH and GARCH-M models // University of Western Ontario: Electronic Thesis and Dissertation Repository, 2013.

8. Bollerslev T. Glossary to ARCH (GARCH). Chapter in Volatility and Time Series Econometrics: Essays in Honour of Robert F. Engle (T. Bollerslev, J. R. Russell, M. Watson) // Oxford University Press, 2009.
9. Madan D., Seneta E. The variance gamma (V.G.) model for share market returns // The Journal of Business. 1990. V. 63, N 4. P. 511–524.
10. Madan D., Carr P., Chang E The variance gamma process and option pricing // Review of Finance. 1998. V. 2, N 1. P 79–105.
11. Kotz S., Kozubowski T.,Podgórski K. The laplace distribution and generalizations Boston: Birkhäuser, 2001.
12. Loregian A., Mercuri L., Rroji E. Approximation of the variance gamma model with a finite mixture of normals // Statistics & Probability Letters. 2012. V. 82, N 2. P. 217–224.
13. Dempster A., Laird N., Rubin D. Maximum likelihood from incomplete data via the EM algorithm // Journal of the Royal Statistical Society. Series B, N 39. 1–38.

OR in biology, medicine, physics and ecology

Opinion convergence in the Krasnoshchekov model

I.V. Kozitsin[1,3] and A.A. Belolipetskii[2]
[1]*Institute of Control Sciences of Russian Academy of Sciences, Moscow, Russia*
[2]*Dorodnicyn Computing Centre, FRC CSC RAS, Moscow, Russia*
[3] *Moscow Institute of Physics and Technology, Dolgoprudny, Russia*

Mathematical models describing opinion formation process has received considerable attention in recent years. A critical point in this field is to find whether individuals' opinions converge. Another important question is to find the settings which allow opinions to converge to the one common view. This outcome is usually called consensus [1]. Scholars have rigorously examined these issues for classic French-DeGroot and Friedkin-Johnsen models. These models has been built upon the *convex* opinion readjustment mechanism [2]. According to this rule, an individuals change their opinions to the *convex combinations* of their own and others' opinions.

There is a gap in the scientific knowledge with respect to the Krasnoshchekov model, which is less-known beyond the Russian scientific community [3–5]. This model has the very similar opinion formation rule. In this paper, we make an attempt to shed some light on this problem. Besides, we discuss some issues related to the time confusion puzzle in the Krasnoshchekov dynamics. To reconcile this problem, we introduce two evident explanations taking their roots in asynchronous dynamics and in game theory.

Consider an isolated system: a group of N agents. We suppose that the contingent does not change across the time which is assumed to be discrete. In the seminal paper [3], opinion dynamics of the agent i obeys the following equation:

$$y_i(t+1) = x_i y_i(t) + (1 - x_i) \sum_{j=1}^{N} c_{ij} y_j(t+1), \tag{1}$$

where $y_i \in [0, 1]$ represents agent's i opinion - *cognitive orientation* towards one fixed issue: for example, whether to visit museum on this weekend ($y_i(t) = 1$) or not to do it ($y_i(t) = 0$) [2]. In this example value $y_i(t) = 0.5$ corresponds to the maximum of uncertainty. Obviously, one may consider more complex opinion space by introducing opinions over the set of independent or even interdependent issues. Nonetheless, this work is confined to the scalar opinions standing for one fixed topic under discussion. Parameter x_i represents the agent's i resistance to the external interpersonal influence. Its conventional notation is *self-weight* or *self-appraisal.* Finally, $c_{ij} \in [0, 1]$ describes the influence agent j has on agent i. This set of coefficients is bounded by: $\sum\limits_{j=1}^{N} c_{ij} = 1$ and $c_{ii} = 0$. Using the notations $\vec{y}(t) = (y_1(t), ..., y_N(t))^T$, $X = diag(x_1, ..., x_N)$ and $C = (c_{ij})_{i,j \in \{1,...,N\}}$, one can rewrite equation (1):

$$\vec{y}(t+1) = X\vec{y}(t) + (I - X)C\vec{y}(t+1), \tag{2}$$

where I is the identity matrix. Matrix C is also called *relative interaction matrix* [2].

As it was mentioned above, opinion process (2) is very similar to ones from French-DeGroot and Friedkin-Johnsen models. Indeed, using the notations introduced above French-DeGroot model can be represented as

$$\vec{y}(t+1) = X\vec{y}(t) + (I - X)C\vec{y}(t). \tag{3}$$

Similarly, one can obtain Friedkin-Johnsen dynamics:

$$\vec{y}(t+1) = X\vec{y} + (I - X)C\vec{y}(t), \tag{4}$$

where the vector $\vec{y}$ stands for agents' prejudices. In some cases one can suppose that they are agents' initial opinions: $\vec{y} = \vec{y}(0)$. One striking difference between (4) and (2) is that in Friedkin-Johnsen model the principal diagonal of the matrix C may not be zero.

Actually, protocol (2) is controversial in some sense: for given agents i and j, agent's i opinion at the time moment $t+1$ relies on the agent's j opinion at the time moment t (in case $c_{ij} > 0$ and $x_i < 1$). Hence, the agent i readjust its opinion after the agent j has made the same. However, similar reflection can be implemented with respect to the agent j. Hence, we have time contradiction.

In this paper, we introduce two explanations of the Krasnoshchekov dynamics. We suppose this treatments to be able to reconcile the above puzzle.

Treatment 1. *In contrast to protocols* (3),(4) *which imply synchronous interaction settings, Krasnoshchekov model takes into account the fact that agents interact and change its opinions not simultaneously but rather asynchronously [6]. The prominent example is public debates, where people communicate in a chaos way. To consider such systems, one can introduce time confusion.*

The second explanation has its roots in the game theory.

Treatment 2. *Let us introduce a cost function. This function stands for the desire of an individual to minimize the* ***current*** *cognitive dissonance with surrounding people. We assume this cognitive dissonance effect to have complex structure: it should depend on the interpersonal weights which are established among individuals. Namely, this cost function can be represented by*

$$F_i^* = [y_i(t+1) - x_i y_i(t) - (1-x_i)\sum_{j=1}^{N} c_{ij} y_j(t+1)]^2. \tag{5}$$

The purpose of the agent i is to minimize (5) *by choosing such strategy $y_i(t+1)$ at the time moment $t+1$, which derives the balance between agent's own opinion and its neighborhood views:*

$$[y - x_i y_i(t) - (1-x_i)\sum_{j=1}^{N} c_{ij} y_j(t+1)]^2 \longrightarrow \min_{y\in[0,1]} \tag{6}$$

The striking signature of the (5) is that agents try to "look forward". In other words, to solve the problem (6), one should evaluate others' optimal strategies. This task requires full awareness: agents have to possess all information concerning the matrices X, C and $\vec{y}(t)$. Obviously, in this case the protocol (1) derives the optimal solution of the problem (6). However, this conditions can be satisfied only for relatively small systems, where all know all the information about each others.

The matrix C describes the structure of social interactions and it can be, in turn, represented by the directed unweighted graph $G[C]$. It is assumed that the nodes of the $G[C]$ correspond to the agents and the arch (j, i) exists if and only if $c_{ij} > 0$. This definition is conventional in social network analysis. In [5], the following statement was obtained.

Statement 1. *Equation* (2) *has an unique solution if and only if* $G[C]$ *doesn't contain any closed strong component (see [5]), whose nodes entirely correspond to the agents-conformists (open-minded agents with* $x_i = 0$*). In case the solution is unique, it can be calculated by*

$$\vec{y}(t+1) = W\vec{y}(t), \tag{7}$$

where $W = (E - (E - X)C)^{-1}X$ *is a row-stochastic matrix, which is also called influence matrix.*

Actually, the matrix W is very similar to the matrix binding initial and terminal opinions in Friedkin - Johnsen model in case the latter exists. The difference is that in the Krasnoshchekov model the matrix W stands only for one iteration, instead of the Friedkin-Johnsen model. One can explain it by specifying time-scale: Krasnoshchekov dynamics implies one iteration, which processes during the durable time span.

Additionally, building upon this fact, one can infer the following statement.

Statement 2. *Using the protocol* (4)*, agents have to reach the optimal solution of the problem* (6) *not through one iteration but during the huge number of time steps. The crucial point here is that protocol* (4) ***does not*** *require full awareness: an agent has to know only its own parameters to obey this dynamics.*

To set up about the convergence problem let us consider a directed unweighted graph $G[W]$. We constitute this graph similarly to $G[C]$. The well-known fact is that the existence of the limit $\lim\limits_{t\to\infty} \vec{y}(t) = \vec{y}(\infty)$ and the settings under which it is consensus (all components of $y(\vec{\infty})$ are the same) can be found from the structural properties of the influence network $G[W]$. The matrices W and X, C are bounded by the nontrivial relation $W = (E - (E - X)C)^{-1}X$. Nonetheless, in [5], the simple rule was explored, spanning these entities.

Lemma 1. *Let suppose that* X *and* C *satisfy the conditions of the statement 1. Then, for* $i \neq j$ *the inequality* $w_{ij} > 0$ *is true if and only if there exists a walk in graph* $G[C]$ *from node* j *to node* i *such that (i) agent* j *is not a conformist; (ii) all nodes pertaining the walk excepting* j *are not correspond to the "stubborn" agents with self-appraisal* $x = 1$.

Besides, $w_{ii} > 0$ *if and only if* $x_i > 0$.

This statement corroborate our assumption about durable time interval, since Krasnoshchekov dynamics encourage all possible opinion percolation directions, which can occur only through long time span.

Using lemma 1 the convergence properties of the Krasnoshchekov model can be precisely examined.

Theorem 1. *For any initial condition* $\vec{y}(0)$ *the process* (7) *converges:*

$$\lim_{t \to \infty} \vec{y}(t) = \vec{y}(\infty) \tag{8}$$

Thus, in contrast to French-DeGroot or Friedkin-Johnsen processes opinion evolution obeyed (7) is free of periodic dynamics and converges eventually. The question is whether it reaches consensus?

Theorem 2. *In theorem 1 the vector* $\vec{y}(\infty)$ *is a consensus for any initial condition* $\vec{y}(0)$ *if and only if (i) the number of closed strong communities doesn't exceed* 1; *(ii) number of "stubborn" agents in the closed strong community doesn't exceed* 1.

Theorem 2 does not differ from the corresponding statement in French-DeGroot model. The analysis we introduced is limited: we do not consider the case when the matrix W does not exists. In this situation equation (2) has the unlimited set of solutions which have the prominent signature: they are all alike in such a way that on each time step conformists cooperate to choose the same random opinion [5]. This behavior is close to the conduction electrons interaction in superconductivity state. More rigorous analysis of such opinion phenomenon can be found in [5].

The critical peculiarity of our society is the persistent disagreement [2]. Correspondingly, one should introduce such opinion dynamics model which capable to capture this situation. This statement stands for the famous Abelson problem [2]. The critical points here are that if the influence network is strongly connected and agents' resistances are less than one (which are in fact the natural assumptions), then the humankind should reach consensus with respect to all possible topics. In other words, Krasnoshchekov dynamics can not simulate persistent disagreement without making some exotic assumptions. To reconcile this problem one should introduce e.g. bounded confidence settings or repulsive dynamics. Nonetheless, we consider the Krasnoshchekov model to be advanced from the perspective of high-intensity systems (public debates, mass meetings) description.

References

1. Castellano C., Fortunato S., Loreto V. Statistical physics of social dynamics // Reviews of modern physics. 2009. V. 81, N 2. P. 591.
2. Proskurnikov A.V., Tempo R. A tutorial on modeling and analysis of dynamic social networks. Part I // Annual Reviews in Control. 2017. V. 43, P. 65–79.
3. Krasnoshchekov P.S. The simplest mathematical model of behavior. Psychology of conformism // Matematicheskoe modelirovanie. 1998. V. 10, N 7. P. 76–92.
4. Belolipetskii A.A., Kozitsin I.V. Dynamic variant of mathematical model of collective behavior // Journal of Computer and Systems Sciences International. 2017. V. 56, N 3. P. 385–396.
5. Kozitsin I.V. Modified Krasnoshchekov model in case of the reducible matrix of social interactions // Matematicheskoe Modelirovanie. 2017. V. 29, N 12. P. 3–15.
6. Ding Z. et al. Asynchronous opinion dynamics with online and offline interactions in bounded confidence model. 2017.

Control of dusty void characteristics in complex plasma*

O.V. Kravchenko[1] and O.A. Azarova[2]

[1] *Scientific and Technological Center of Unique Instrumentation of RAS, Moscow, Russia*

[2] *Dorodnicyn Computing Centre, Federal Research Center "Computer Science and Control" of RAS% , Moscow, Russia*

Introduction

A number of investigations of dynamics of complex plasma with dusty component were published in recent years [1, 2]. The first dusty plasma structure containing empty domains (voids) was discovered in the course of the experiments on the board of the International Space Station [2]. Also, in laboratory conditions the voids have been found later [3–5]. At the same time, it is not so much studies in which numerical simulations of the structure generation in dusty plasma are performed. Evolution of the dynamics of a single symmetric void from equilibrium was described by electro–hydrodynamic model [6, 7] taking into account the effect of an ion attraction force as a nonlinear function of the speed of ions. The

*This research is partially supported by RFBR Grant 18-37-00059.

algorithm for simulations of appearance of dusty plasma void using the model [7] has been presented in [8] for the case of cylindrical geometry of electrical field. Generation of a single symmetric void and concentric symmetrical voids in unmoving flows of low–temperature dusty plasma has been obtained in [9]. A study of generation of voids in moving flows together with the research of the voids dynamics in unmoving media was presented in [10, 11].

In this paper the results on voids dynamics are presented in dependence on the value of initial electrical field for different initial density of dust component. The simulations are based on the model [7] of formation of a void in a dusty plasma flow. The Lax–Friedrichs scheme with re–calculation and the complex conservative difference scheme [12] have been used for the hydrodynamic part of the model.

Physical model

Simulations are based on the model [7]. Here the model is considered in dimensionless form, believing that the normalizing parameters from [7] have been applied:

$$\frac{\partial v_d}{\partial t} + v_d \frac{\partial v_d}{\partial x} = F_d - E - \alpha_0 v_d - \frac{\tau_d}{n_d}\frac{\partial n_d}{\partial x}, \tag{1}$$

$$\frac{\partial n_d}{\partial t} = -\frac{\partial (n_d v_d)}{\partial x} + D_0 \frac{\partial^2 n_d}{\partial x^2}, \tag{2}$$

$$\frac{dn_e}{dx} = -\frac{n_e E}{\tau_i}, \tag{3}$$

$$\frac{dE}{dx} = 1 - n_e - n_d, \tag{4}$$

$$F_d = \frac{aE}{b + |v_i|^3}, \ v_i = \mu E. \tag{5}$$

Here n_d, n_e are concentrations of dust and electrons components, v_d, v_i are velocities of dust and ions components, E is electric field currency, F_d is ion–drag force, D_0 is diffusion coefficient, a, b are fit parameters of ion–drag force approximation, μ is coefficient of ions mobility, α_0 is friction coefficient, τ_d, τ_i are coefficients of normalized temperature for dust and ion species, n_{e0} is the value of electrons density in initial time moment. In fact the governing equations in this model contain dust momentum (1) and dust continuity equations (2), balance equation of electrons neglecting the electron inertia (3), Poisson law which completes the nonlinear system of equations, and the expression for ion–drag force are presented in (4), (5). Equations (1) and (2) constitute the hydrodynamic part of the model and the equations (3)–(5) are the electrostatic part of it. Parameters of the simulations are collected in Tab. 1.

Table 1: Parameters of the calculations

Parameters	Values
τ_i	0.125
τ_d	0.001
a	7.5
b	1.6, 0.4
α_0	2.0
μ	1.5
n_{e0}	0.999

Numerical procedure

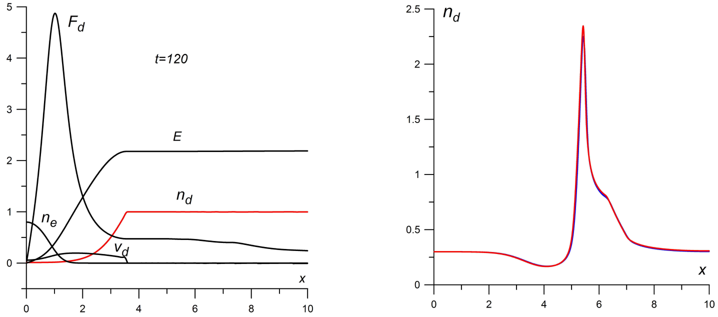
For the first order calculations of the hydrodynamic part of the model the Lax–Friedrichs scheme with re–calculation is used (with $D_0 = 0$) for the divergent form of the equations (1), (2):

$$\frac{\partial \mathbf{u}}{\partial t} + \frac{\partial \mathbf{F}}{\partial x} = \mathbf{f},$$
$$\mathbf{u} = \begin{pmatrix} n_d \\ v_d \end{pmatrix}, \ \mathbf{F} = \begin{pmatrix} n_d v_d \\ \tau_d \ln(n_d) + 0.5 v_d^2 \end{pmatrix}, \ \mathbf{f} = \begin{pmatrix} D_0 \dfrac{\partial^2 n_d}{\partial x^2} \\ F_d - E - \alpha_0 v_d \end{pmatrix}. \tag{6}$$

The second order scheme for the algorithm [8, 9] was obtained using the approach of [12] for increasing the difference scheme order. By this way the systems of divergent equations (6) are used for unknown functions together with the system for their space derivatives

$$\frac{\partial \mathbf{u}_x}{\partial t} + \frac{\partial \mathbf{F}_x}{\partial x} = \mathbf{0},$$
$$\mathbf{u}_x = \begin{pmatrix} n_{dx} \\ v_{dx} \end{pmatrix}, \ \mathbf{F}_x = \begin{pmatrix} (n_d v_d - D_0 n_{dx})_x \\ \left(\tau_d \ln(n_d) + 0.5 v_d^2\right)_x + E - F_d + \alpha_0 v_d \end{pmatrix}. \tag{7}$$

The space derivatives of the right parts are included into the flux functions in the systems of the derivatives (7). It provides the absence of the numerical (parasite) sources and runoffs. Equations (3), (4) are approximated with the second approximation order using the central form for the space derivatives and the method of Runge–Kutta of the fourth order of approximation is used, too. The next restrictions were imposed in the numerical simulations: if $n_d > 1$, then n_d was applied by 1.0 and v_d was applied by v_{d0} (index "0" refers to the initial values of the parameters).



a) b)

Fig. 1. Comparison of dynamics of dusty component n_d during the void formation calculated with the use of first (*blue curve*, $D_0 = 0$) and second (*red curve*, $D_0 = 0$) order schemes for the hydrodynamic part of the model, $n_{d0} = 0.3, v_{d0} = 0.1$: a) beginning stage (without restrictions on n_d and v_d); b) subsequent stage (with restrictions on n_d and v_d), *black curve* — $D_0 = 0.1$.

Comparison of the simulations using the difference schemes of the first and second approximation order is presented in Fig. 1. It can be seen that the obtained dynamics are quite the same (Fig. 1a). It should be noted that the restriction on v_d is introduced for "cutting" the high frequency numerical oscillations in the dynamics of v_d (when n_d becomes more than 1). These oscillations are seen in the profile of n_d calculated with the use of the second order scheme (Fig. 1b, *red curve*). Introduction into the scheme the term with the diffusion of concentration of the dusty component (which plays the role of "viscosity" in (2)) allowed smoothing these oscillations (Fig. 1b, *black curve* $D_0 = 0.1$). Note that the diffusion term is introduced into the difference approximations via the flux function in (7).

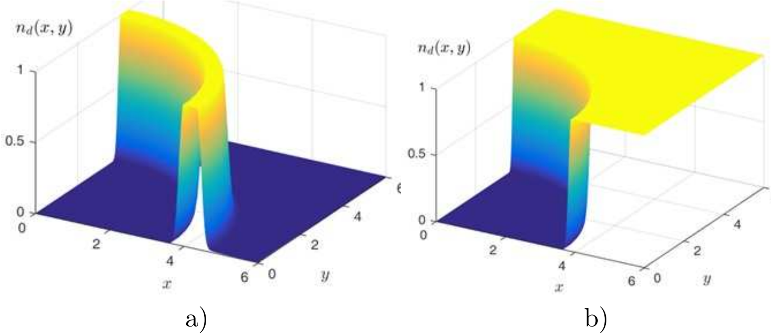


Fig. 2. Dynamics of concentration of dusty component n_d during the void formation, $E_0 = 4 \cdot 10^{-4}$ (first order scheme):
a) — non–dimensional time $t = 70$ (unsteady ring structure);
b) — $t = 200$ steady void.

Dynamics of a void generation

The initial and final stages of void generation in the unmoving flow ($v_{d0} = 0$) of dusty particles are shown in Fig. 2. Two–dimensional figures for n_d obtained by the rotation technique are presented. The mechanism of voids generation is connected with the superposition of the electrical field and the ion attraction force action. Void grows in time uder the constant parameters via the initial circular stage (Fig. 2a) and becomes saturated at time moment $t = 200$ (Fig. 2b).

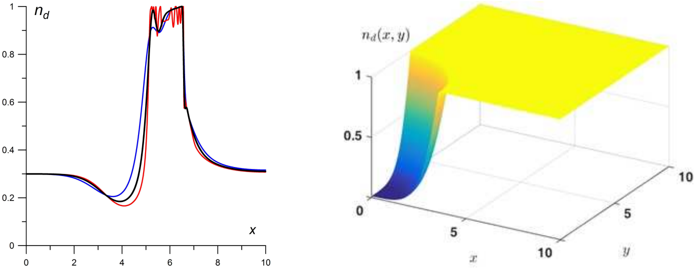


Fig. 3. Simulations with the use of second order scheme,
$n_{d0} = 0.2, b = 0.4, v_{d0} = 0, E_0 = 4 \cdot 10^{-5}, n_i = 2000,\ D_0 = 0.1:$
a) – profiles of the defining flow parameters on the stage when the steady void is formed, $t = 120$; b) – steady void obtained (n_d), $t = 120$.

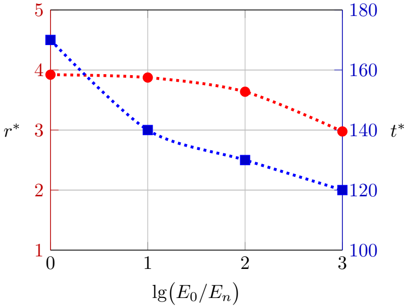


Fig. 4. Boundary of void r^* (*red curve*) and time of steady void establishing t^* (*blue curve*) *vs.* $\lg(E_0/E_n)$, $E_n = 4 \cdot 10^{-6}$.

Dynamics of a void generation with the use of the second order difference scheme for hydrodynamic part of the model [7] is shown in Fig. 3 (here $D_0 = 0.1$, $b = 0.4$). Figure 3a demonstrates the profiles of all the defining parameters for the steady flow containing the formed void. Two–dimensional figure for n_d obtained by the rotation technique is presented in Fig. 3b. Presence of diffusion of the concentration increases the values of n_d at the void's center but by decreasing the constant b in the expression of the ion–drag force F_d (5) the value of residual dusty component concentration can be decreased into the void's centre. In addition, the decrease in n_{d0} has been obtained to provide some icrease in the void's radius.

Dependences of the void's boundary r^* and time of steady state void establishing t^* on the initial relative electric field value E_0/E_n, $E_n = 4 \cdot 10^{-6}$ are presented in Fig. 4. It is seen that when increasing the initial electic field the void radius decreases toghether with the time value of steady void establishing.

Conclusion

Generation of steady structures with empty regions (voids) has been modelled numerically in complex plasma. The simulations are based on the known model of Avinash, Bhattacharjee and Hu of formation of a void in a field of dust particles. The model has been reduced to the divergent form and two algorithms for calculations have been suggested. The Lax–Firedrichs scheme with re–calculation and the complex conservative difference scheme have been used for the hydrodynamic part of the model. Results on the dynamics of voids have been obtained at the stage of circular ring structure generation and at the final stage of a steady round void formation. Dependences of the defining flow parameters on the initial value of the electric field have been obtained for different initial values of dusty component up to the steady voids formation.

References

1. Molotkov V., Thomas H., Lipaev A., Naumkin V., Ivlev A, and Khrapak S. Complex (dusty) plasma research under microgravity conditions: PK–3 Plus Laboratory on the International Space Station // International Journal of Microgravity Science and Application. 2015. V. 35, N 3. P. 1–8.
2. Fortov V.E., Morgill G.E. Complex and dusty plasmas: from laboratory to Space. CRC Press, 2009.
3. Feng H., Mao–Fu Y., Long W., and Nan J. Voids in experimental dusty plasma // Chin. Phys. Lett. 2004. V. 21, N 1. P. 121–124.
4. Sarkar S., Mondal M., Bose M., and Mukherjee S. Observation of external control and formation of a void in cogenerated dusty plasma // Plasma Sources Sci. Technol. 2015. V. 24, N 3. P. 1–7.
5. Feng H., Maofu Y., and Long W. Pattern phenomena in an RF discharge dusty plasma system // Science in China Series G: Physics, Mechanics & Astronomy. 2006. V. 49, N 5. P. 588–596.
6. Ng C.S., Bhattacharjee A., Hu S., Ma Z.W., and Avinash K. Generalizations of a nonlinear fluid model for void formation in dusty plasmas // Plasma Phys. Control. Fusion. 2007. V. 49. P. 1583–1597.
7. Avinash K., Bhattacharjee A., and Hu S. Nonlinear theory of void formation in colloidal plasmas // Phys. Rev. Lett. 2003. V. 90, N 7. P. 1–4.
8. Kravchenko O.V., and Pustovoit V.I. Numerical simulation of dynamics of concentric dusty–plasma structures // In proceedings of 15th International Workshop on Magneto–Plasma Aerodynamics. 2016. P. 1–8.
9. Kravchenko O.V. Simulation of spatially localized dusty plasma structures in complex plasma // Nanostructures. Mathematical physics and modeling. 2016. V. 15, N 2. P. 51–61 (in Russian).
10. Kravchenko O.V., Azarova O.A., Lapushkina T.A. Dusty plasma void dynamics in unmoving and moving flows // Proc. 7th European Conference for Aeronautics and Space Sciences (EUCASS2017), 2017. P. 1–12.

11. Kravchenko O.V., Azarova O.A. Generation of voids in unmoving and moving dusty plasma // Nanostructures. Mathematical physics and modeling. 2017. V. 17, N 1. P. 5–16 (in Russian).
12. Azarova O.A. Complex conservative difference schemes for computing supersonic flows past simple aerodynamic forms // J. Comp. Math. Math. Phys. 2015. V. 55, N 12. P. 2067–2092.

Game-theoretic models

Equilibria in games with common local utilities

Nikolai S. Kukushkin
Dorodnicyn Computing Centre, FRC CSC RAS, Moscow, Russia

The first ever examples of the games considered here were discovered in the 1970s by Rosenthal [1] and Germeier and Vatel' [2]. The former model played an important role in Monderer and Shapley's theory of potential games [3]. The latter model was generalized in [4,5] and especially in [6]. Here, we introduce a much wider class of games, which includes all those models. The theorems have been published in a series of articles [7–12]. Somewhat similar, but weaker, results can be found in [13–16].

Model

A *game with common local utilities* (a *CLU game*) may have an arbitrary (finite) set of players N and arbitrary strategy sets X_i, whereas the utility functions are generated by the following construction. There is a finite set A of *facilities*; we denote $\mathcal{B}$ the set of all nonempty subsets of A and $\mathcal{N}$ the set of all nonempty subsets of N. For each $i \in N$, there is a mapping $\mathrm{B}_i \colon X_i \to \mathcal{B}$ describing what facilities player i uses having chosen x_i.

For every $\alpha \in \mathrm{A}$, we denote $I_\alpha^- := \{i \in N \mid \forall x_i \in X_i\, [\alpha \in \mathrm{B}_i(x_i)]\}$ and $I_\alpha^+ := \{i \in N \mid \exists x_i \in X_i[\alpha \in \mathrm{B}_i(x_i)]\}$. For each $i \in I_\alpha^+$, we denote $X_i^\alpha := \{x_i \in X_i \mid \alpha \in \mathrm{B}_i(x_i)\}$; if $i \in I_\alpha^-$, then $X_i^\alpha = X_i$. Then we set $\mathcal{I}_\alpha := \{I \in \mathcal{N} \mid I_\alpha^- \subseteq I \subseteq I_\alpha^+\}$ and $\Xi_\alpha := \{\langle I, x_I\rangle \mid I \in \mathcal{I}_\alpha \;\&\; x_I \in X_I^\alpha\}$. Without restricting generality, we may assume $I_\alpha^+ \neq \emptyset$ – if nobody can use a facility, there would be no point in including it in the description of the game – and hence $\Xi_\alpha \neq \emptyset$ too. The *local utility function* at $\alpha \in \mathrm{A}$ is $\varphi_\alpha \colon \Xi_\alpha \to \mathbb{R}$. For every $\alpha \in \mathrm{A}$ and $x_N \in X_N$, we denote $N(\alpha, x_N) := \{i \in N \mid \alpha \in \mathrm{B}_i(x_i)\}$: the set of players using α under x_N; obviously, $I_\alpha^- \subseteq N(\alpha, x_N) \subseteq I_\alpha^+$. For every $i \in N$ and $x_i \in X_i$, there is an *aggregation rule*, a mapping $U_i^{x_i} \colon \mathbb{R}^{\mathrm{B}_i(x_i)} \to \mathbb{R}$. The *total utility function* of each player i is

$$u_i(x_N) := U_i^{x_i}\bigl(\langle \varphi_\alpha(N(\alpha, x_N), x_{N(\alpha,x_N)})\rangle_{\alpha \in \mathrm{B}_i(x_i)}\bigr)$$

for all $x_N \in X_N$.

Generally, CLU games do not have any remarkable properties. Interesting results emerge under certain assumptions about the aggregation rules and/or local utility functions.

Player i has a *negative impact* on facility α if, whenever $i \notin I \in \mathcal{I}_\alpha$, $I \cup \{i\} \in \mathcal{I}_\alpha$, $x_i \in X_i^\alpha$, and $x_I^\alpha \in X_I^\alpha$, there holds

$$\varphi_\alpha(I, x_I^\alpha) \geqslant \varphi_\alpha(I \cup \{i\}, \langle x_I^\alpha, x_i\rangle).$$

A definition of *positive impacts* is obtained by reversing the inequality sign. Γ is a *game with negative/positive impacts* if the appropriate condition holds for all $i \in N$ and $\alpha \in \mathrm{A}$.

Games with structured utilities are defined by the condition $I_\alpha^- = I_\alpha^+ =: N(\alpha)$ for all $\alpha \in \mathrm{A}$ (hence A must be finite); in simple words, the players cannot choose facilities. The games of [4] are distinguished by the minimum aggregation of local utilities. Every game with structured utilities exhibits both negative and positive impacts (by default); actually, such games are characterized by the combination of *strictly* negative and *strictly* positive impacts.

In a *generalized congestion game*, A is finite, $X_i \subseteq \mathcal{B}$ for each $i \in N$, and each B_i is an identity mapping (i.e., each player chooses just a set of facilities); besides, φ_α only depends on $\#I$. Proper congestion games of [1] are distinguished by additive aggregation of local utilities (i.e., the players just sum them up).

A facility $\alpha \in \mathrm{A}$ is *trim* if there is a real-valued function $\psi_\alpha(m)$ defined for integer m between $\min_{I \in \mathcal{I}_\alpha} \#I = \max\{1, \#I_\alpha^-\}$ and $\#I_\alpha^+ - 1$ such that

$$\varphi_\alpha(I, x_I) = \psi_\alpha(\#I) \tag{T}$$

whenever $I \in \mathcal{I}_\alpha$, $I \neq I_\alpha^+$, and $x_I \in X_I^\alpha$. In other words: whenever a trim facility is not used by all potential users, neither the identities of the users, nor their strategies matter, only the number of users. A CLU game is *trim* if so is every facility.

Both generalized congestion games and games with structured utilities are trim. In the first case, (T) holds for all $I \in \mathcal{I}_\alpha$, even for $I = I_\alpha^+$. In the second case, conversely, $I_\alpha^- = I_\alpha^+$ for each facility α; hence $\mathcal{I}_\alpha = \{I_\alpha^+\}$, and hence (T) is not required at all.

Results

In Theorems 1, 1*, and 4, appropriate topological assumptions should be imposed on sets X_i and functions φ_α. An *aggregation rule* in Theorems 2, 3, 5 and 6 is a mapping $U\colon \mathbb{R}^{\Sigma(U)} \to \mathbb{R}$, where $\Sigma(U)$ is a finite set (of potential names for local utilities). An aggregation rule is (*strictly*) *admissible* if it is continuous and strictly increases w.r.t. strong (weak) Pareto order on $\mathbb{R}^{\Sigma(U)}$.

Theorem 1. Let Γ be a CLU game with negative impacts where A is finite and each player uses the minimum aggregation. Then coalition improvements in Γ are acyclic and hence Γ possesses a strong Nash equilibrium.

Theorem 1*. Let Γ be a CLU game with positive impacts where each player uses the maximum aggregation. Then coalition improvements in Γ are acyclic and hence Γ possesses a strong Nash equilibrium.

Theorem 2. If a set $\mathfrak{U}$ of admissible aggregation rules ensures the existence of a weakly Pareto optimal Nash equilibrium in every generalized congestion game with negative impacts, then for every $U \in \mathfrak{U}$, there is a continuous and strictly increasing mapping $\lambda^U : \mathbb{R} \to \mathbb{R}$ such that

$$\forall U \in \mathfrak{U}\ \forall v_{\Sigma(U)} \in \mathbb{R}^{\Sigma(U)}\ \bigl[U(v_{\Sigma(U)}) = \lambda^U(\min_{s\in\Sigma(U)} v_s)\bigr];$$

and

$$\forall U', U \in \mathfrak{U}\ \bigl[\lambda^{U'} = \lambda^U \text{ or } \lambda^{U'}(\mathbb{R}) \cap \lambda^U(\mathbb{R}) = \emptyset \\ \text{or } \#\Sigma(U) = 1 = \#\Sigma(U')\bigr].$$

Theorem 3. If a set $\mathfrak{U}$ of admissible aggregation rules ensures the existence of a weakly Pareto optimal Nash equilibrium in every finite game with structured utilities, then for every $U \in \mathfrak{U}$, there is a continuous and strictly increasing mapping $\lambda^U : \mathbb{R} \to \mathbb{R}$ such that either

$$\forall U \in \mathfrak{U}\ \forall v_{\Sigma(U)} \in \mathbb{R}^{\Sigma(U)}\ \bigl[U(v_{\Sigma(U)}) = \lambda^U(\min_{s\in\Sigma(U)} v_s)\bigr],$$

or

$$\forall U \in \mathfrak{U}\ \forall v_{\Sigma(U)} \in \mathbb{R}^{\Sigma(U)}\ \bigl[U(v_{\Sigma(U)}) = \lambda^U(\max_{s\in\Sigma(U)} v_s)\bigr];$$

besides,

$$\forall U', U \in \mathfrak{U}\, \bigl[\lambda^{U'} = \lambda^{U} \text{ or } \#\Sigma(U) \neq \#\Sigma(U') \text{ or } \lambda^{U'}(\mathbb{R}) \cap \lambda^{U}(\mathbb{R}) = \emptyset\bigr].$$

Theorem 4. Let Γ be a trim CLU game where each player uses additive aggregation. Then individual improvements in Γ are acyclic and hence Γ possesses a Nash equilibrium.

Theorem 5. If $\#N \geqslant 2$ and there are sets $\mathfrak{U}_i$ $(i \in N)$ of strictly admissible aggregation rules ensuring the existence of a Nash equilibrium in every generalized congestion game, then:
(1) there are a continuous and strictly increasing mapping $\nu : \mathbb{R} \to \mathbb{R}$, and a continuous and strictly increasing mapping $\lambda^U : \nu(\mathbb{R})^{\Sigma(U)} \to \mathbb{R}$ for every $i \in N$ and $U \in \mathfrak{U}_i$ such that

$$\forall i \in N\ \forall U \in \mathfrak{U}_i\ \forall v_{\Sigma(U)} \in \mathbb{R}^{\Sigma(U)} \Bigl[U(v_{\Sigma(U)}) = \lambda^U\Bigl(\sum_{s \in \Sigma(U)} \nu(v_s)\Bigr)\Bigr];$$

(2) for every $i \in N$ and $U, U' \in \mathfrak{U}_i$ such that $\#\Sigma(U) > 1 < \#\Sigma(U')$, there is a constant $\bar{u}^{UU'} \in \mathbb{R} \cup \{-\infty, +\infty\}$ such that

$$\operatorname{sign}\bigl(\lambda^{U'}(u') - \lambda^{U}(u)\bigr) = \operatorname{sign}(u' - u - \bar{u}^{UU'})$$

for all $u' \in (\#\Sigma(U')) \cdot \nu(\mathbb{R})$ and $u \in (\#\Sigma(U)) \cdot \nu(\mathbb{R})$.

Theorem 6. If $\#N \geqslant 2$ and there are sets $\mathfrak{U}_i$ $(i \in N)$ of strictly admissible aggregation rules ensuring the existence of a Nash equilibrium in every finite game with structured utilities, then there are a continuous and strictly increasing mapping $\nu : \mathbb{R} \to \mathbb{R}$ and a continuous and strictly increasing mapping $\lambda^U : \nu(\mathbb{R})^{\Sigma(U)} \to \mathbb{R}$ for every $i \in N$ and $U \in \mathfrak{U}_i$ such that

$$\forall i \in N\ \forall U \in \mathfrak{U}_i\ \forall v_{\Sigma(U)} \in \mathbb{R}^{\Sigma(U)} \Bigl[U(v_{\Sigma(U)}) = \lambda^U\Bigl(\sum_{s \in \Sigma(U)} \nu(v_s)\Bigr)\Bigr].$$

Theorem 7. Every potential game is isomorphic to a game with structured utilities where each player just sums up the relevant local utilities.

Theorem 8. A set of aggregation rules ensures the existence of an exact potential in every trim CLU game if and only if every rule prescribes just summing up the relevant local utilities.

References

1. Rosenthal, R.W. A class of games possessing pure-strategy Nash equilibria // International Journal of Game Theory. 1973. V. 2, N 1. P. 65–67.

2. Germeier Yu.B., Vatel' I.A. On games with a hierarchical vector of interests // Izvestiya Akademii Nauk SSSR, Tekhnicheskaya Kibernetika. 1974. N 3. P. 54–69 [in Russian; English translation in Engineering Cybernetics. V. 12, N 3. P. 25–40].
3. Monderer, D., Shapley, L.S. Potential games // Games and Economic Behavior. 1996. V. 14, N 1. P. 124–143.
4. Vatel' I.A. The core of a game with public and private objectives // Avtomatika i Telemekhanika. 1980. No. 1. P. 91–96. [in Russian; English translation in Automation and Remote Control. V. 41, N 1. P. 73–77].
5. Men'shikov I.S., Men'shikova O.R. Strong equilibria and nucleolus in games with a hierarchical vector of interests // Computational Mathematics and Mathematical Physics. V. 25, N 5. P. 14–20.
6. Kukushkin N.S., Men'shikov I.S., Men'shikova O.R., Moiseev N.N. Stable compromises in games with structured payoffs // Zhurnal Vychislitel'noi Matematiki i Matematicheskoi Fiziki. 1985. V. 25, N 12. P. 1761–1776 [in Russian; English translation in USSR Computational Mathematics and Mathematical Physics. V. 25, N 6. P. 108–116].
7. Kukushkin, N.S. On existence of stable and efficient outcomes in games with public and private objectives // International Journal of Game Theory 1992. V. 20, N 3. P. 295–303.
8. Kukushkin, N.S. A condition for the existence of a Nash equilibrium in games with public and private objectives // Games and Economic Behavior 1994. V. 7, N 2. P. 177–192.
9. Kukushkin, N.S. Congestion games revisited // International Journal of Game Theory. 2007. V. 36, N 1. P. 57–83.
10. Kukushkin, N.S. Strong Nash equilibrium in games with common and complementary local utilities // Journal of Mathematical Economics. 2017. V. 68, N 1. P. 1–12.
11. Kukushkin, N.S. Inseparables: exact potentials and addition // Economics Bulletin. 2017. V. 37, N 2. P. 1176–1181.
12. Kukushkin, N.S. A universal construction generating potential games // Games and Economic Behavior, accepted in 2017.
13. Le Breton, M., Weber, S. Games of social interactions with local and global externalities // Economics Letters. 2011. V. 111, N 1. P. 88–90.
14. Harks, T., Klimm, M., Möhring, R.H. Characterizing the existence of potential functions in weighted congestion games // Theory of Computing Systems. 2011. V. 49, N 1. P. 46–70.

15. Harks, T., Klimm, M., Möhring, R.H. Strong equilibria in games with the lexicographical improvement property // International Journal of Game Theory. 2013. V. 42, N 2. P. 461–482.

16. Harks, T., Klimm, M. Equilibria in a class of aggregative location games // Journal of Mathematical Economics. 2015. V. 61. P. 211–220.

A cooperative game in international electric power integration*

I.M. Minarchenko

Melentiev Energy Systems Institute of Siberian Branch of the Russian Academy of Sciences, Irkutsk, Russia

The present study is concerned with the special model of the long-term development and functioning of electrical power systems. This model is called the ORIRES model [1]. The ORIRES is a Russian abbreviation that means the Optimization of Development and Operating Conditions of Electric Power Systems. One of its main properties is that it takes the development of the power system into consideration. In that way the model can provide not only the amount of generated electricity but also the capacities of power generators and electrical transmission lines installed in the system.

In this paper we focus on the problem of the long-term analysis of international electric power integration using the ORIRES model. This statement describes a cooperation of countries in the field of electric power industry. Such a cooperation may lower the costs of countries, which take part in the coalition. It is assumed that, if lack of the energy is the case, the transmission of electricity from the existing station of another country is less expensive than the expansion of own capacities. Thus we obtain the problem of dividing the surplus of the coalition between the countries in it. This problem can be solved by using cooperative game theory [2].

The ORIRES represents a static multi-node linear model. It allows for the four seasons, whereas every season consists of working days and holidays, every day is separated into 24 equal periods of time (hours). The variables in the model are capacities and operating powers of generating stations, capacities of transmission lines, and the amounts of elec-

*This work is a part of the project III.17.3.1 of the basic research program of Siberian Branch of the Russian Academy of Sciences, no. AAAA-A17-117030310444-2.

tricity transmitted between nodes via the lines. The total annual discounted cost of the system plays the role of the objective function. The costs consist of the following components: the power generation costs, as well as capital and maintenance costs of both power generators and transmission lines. A node represents a local energy system, which contains possibly various types of electrical power stations, for example, thermal, hydroelectric or nuclear. Every node belongs to a certain country presented in the system. Our investigation is based on the real data on the six countries of Asian region: Russia, Mongolia, China, North Korea, South Korea and Japan.

In order to define a cooperative game, we need to specify a characteristic function. Let N be the set of all countries under consideration. Every non-empty subset $K \subseteq N$ is referred to as a coalition. The set N is called the grand coalition. A characteristic function $v\colon \{K \mid K \subseteq N\} \to \mathbb{R}$ assigns to every coalition $K \subseteq N$ an attainable payoff $v(K)$ [2]. The value $v(K)$ represents the maximal payoff that may be guaranteed for K. As we mentioned before, building up capacities is significantly more expensive than electricity generating and transmitting. Hence the guaranteed payoff of a coalition is defined by this coalition's minimal costs in the case of its isolated functioning. In that way one need to solve $(2^{|N|} - 1)$ large-scale linear programming problems in order to specify $v(K)$ for every coalition $K \subseteq N$.

We investigate two solution concepts for the derived cooperative game: the Core and the Shapley value [2]. The Core is the set of undominated imputations which is defined by linear constraints. The Core in our problem is non-empty. Let us give an example of imputation that belongs to the Core. Table 1 describes Chebyshev center of the Core. The column “Isolated” represents costs for every country when it does not cooperate with any of other countries. The last two columns reflect the decreasing of costs, when countries cooperate in the grand coalition with the given imputation, as compared to isolated functioning.

The Shapley value is a unique imputation that always exists. Moreover, for our problem, the Shapley value turns out to be an element of the Core. Information on the Shapley value is gathered in Table 2.

References

1. Belyaev L.S., Podkovalnikov S.V., Savelyev V.A., Chudinova L.Yu. Effectiveness of International Electrical Communications (in Russian). Novosibirsk: Nauka, 2008.
2. Gilles R.P. The Cooperative Game Theory of Networks and Hie-

Table 1. Chebyshev center of the Core.

	Imputation (costs)		Isolated	Integration effect	
	mill. doll.	%	mill. doll.	%	mill. doll.
Russia	7353	2.21	7591	−3.14	−239
Mongolia	670	0.20	909	−26.26	−239
China	135301	40.63	147757	−8.43	−12455
North Korea	0	0	5047	−100.00	−5047
South Korea	54239	16.29	57153	−5.10	−2915
Japan	135438	40.67	138441	−2.17	−3004
Total	333001	100.00	356899	−6.70	−23898

Table 2. The Shapley value.

	Imputation (costs)		Isolated	Integration effect	
	mill. doll.	%	mill. doll.	%	mill. doll.
Russia	3867	1.16	7591	−49.07	−3725
Mongolia	350	0.11	909	−61.46	−559
China	141328	42.44	147757	−4.35	−6429
North Korea	527	0.16	5047	−89.56	−4520
South Korea	52578	15.79	57153	−8.01	−4575
Japan	134351	40.35	138441	−2.95	−4090
Total	333001	100.00	356899	−6.70	−23898

rarchies. Berlin: Springer-Verlag, 2010.

On an equilibrium in pure strategies for final-offer arbitration[*]

V.V. Morozov

Lomonosov Moscow State University, Moscow, Russia

A mechanism of final-offer arbitration (FOA) was introduced by Stevens [1] , in which the arbitrator has to choose between the offers submitted by the parties. In [2] the following zero-sum game was conceded. The first player-employee and the second player-employer offer a salary $x \in \mathbb{R}$ and $y \in \mathbb{R}$. If $x \leqslant y$ then the settled salary is $(x+y)/2$. Otherwise the arbitrator takes value z of random variable Z distributed with continuous cumulative probability function $F(z)$ and chooses the offer from the set $\{x, y\}$ closest to z. The first player payoff function is a mean salary

$$H(x,y) = yF\Big(\frac{x+y}{2}\Big) + x\Big(1 - F\Big(\frac{x+y}{2}\Big)\Big).$$

[*] The reported study was funded by RFBR according to the research project 16-01-00353 a.

In [2] a sufficient conditions for pure equilibrium existence in the game $\Gamma = \langle \mathbb{R}, \mathbb{R}, H\rangle$ were derived. In given work these conditions are refined. Furthermore, another conditions are formulated for a constrained game $\Gamma_+ = \langle \mathbb{R}_+, \mathbb{R}_+, H\rangle$.

Suppose that the probability density function $f(z) = F'(z)$ has at least one-sided derivative for each z. Let Mo and Me designate a mode and a median of the distribution. As shown in [2] under $f(\mathrm{Me}) > 0$ an equilibrium using pure strategies in the game Γ is necessary

$$x_0 = \mathrm{Me} + \frac{1}{2f(\mathrm{Me})},\ y_0 = \mathrm{Me} - \frac{1}{2f(\mathrm{Me})}.$$

The sufficient condition for local equilibrium (see [2]) is

$$f^2(\mathrm{Me}) > \frac{1}{4}|f'(\mathrm{Me})|. \tag{1}$$

For (x_0, y_0) to be a global equilibrium one can add to (1) the following sufficient conditions (see [2]):

$$f(z) \leqslant f(\mathrm{Me}) + 4f^2(\mathrm{Me})|\mathrm{Me} - z|, \forall\, z : |Me - z| \leqslant 1/(4f(Me)), \tag{2}$$

and for some constants c_1, c_2 such that $-\infty \leqslant c_1 \leqslant \mathrm{Me} \leqslant c_2 \leqslant +\infty$

$$f(z) \geqslant f(\mathrm{Me}) \exp(-2f(\mathrm{Me})|\mathrm{Me} - z|) \text{ iff } z \in [c_1, c_2]. \tag{3}$$

Instead of (2) and (3) let's formulate another conditions. The definition of the equilibrium

$$H(x, y_0) \leqslant H(x_0, y_0) \leqslant H(x_0, y),\ \forall\, x, y,$$

is equivalent to inequalities

$$G(x) \overset{def}{=} F\Big(\frac{x + y_0}{2}\Big) - 1 + \frac{1}{2f(\mathrm{Me})(x - y_0)} \geqslant 0,\ \forall\, x \geqslant \mathrm{Me}, \tag{4}$$

$$D(y) \overset{def}{=} \frac{1}{2f(\mathrm{Me})(x^0 - y)} - F\Big(\frac{x^0 + y}{2}\Big) \geqslant 0,\ \forall\, y \leqslant \mathrm{Me}. \tag{5}$$

Define two functions

$$g(x) = \frac{1}{4}f'\Big(\frac{x + y_0}{2}\Big) + \frac{1}{x - y_0} f\Big(\frac{x + y_0}{2}\Big),\ x \geqslant \mathrm{Me},$$

$$d(y) = -\frac{1}{4}f'\Big(\frac{x_0 + y}{2}\Big) + \frac{1}{x_0 - y} f\Big(\frac{x_0 + y}{2}\Big),\ y \leqslant \mathrm{Me}.$$

Theorem 1. *Let* (1) *be true and* $f(z) > 0$ *for all* z. *Suppose that function* g (*function* d) *on each interval* (Me, x_0) *and* (x_0, ∞) $((-\infty, y_0)$ *and* $(y_0, \mathrm{Me}))$ *has no more than one zero. Besides that at the zero of the function* g (*function* d) *it changes sign. Then* (x_0, y_0) *is the equilibrium in the game* Γ.

Note that for Cauchy distribution with $f(z) = 1/(1+\pi^2 z^2)$ conditions of the theorem 1 are fulfilled but (3) doesn't hold (see [2]).

Theorem 2. *Let* (1) *be true, variable* Z *be distributed on* $\mathbb{R}_+$ *and* $f(z) > 0$ *for all* $z > 0$. *Suppose that* $\mathrm{Me}+y_0 \geqslant 0$ *and function* g (*function* d) *on each interval* (Me, x_0) *and* (x_0, ∞) $((-x_0, y_0)$ *and* $(y_0, \mathrm{Me}))$ *has no more than one zero. Besides that at the zero of the function* g (*function* d) *it changes sign. Then* (x_0, y_0) *is the equilibrium in the game* Γ. *If* $\mathrm{Me} + y_0 < 0$ *then the equilibrium doesn't exist.*

In the game $\Gamma_+ = \langle \mathbb{R}_+, \mathbb{R}_+, H \rangle$ the random variable Z is distributed on $\mathbb{R}_+$. If $y_0 < 0$ then an equilibrium in the game Γ_+ is necessary $(x^*, 0)$ where $x^* = 2z^*$, $z^* = \arg\max\limits_{z \geqslant 0} z(1 - F(z))$. Note that z^* exists if

$$\lim_{z \to \infty} z(1 - F(z)) = 0. \tag{6}$$

Theorem 3 *Let* f *be unimodal function and* $\mathrm{Mo} \leqslant \mathrm{Me} < z^*$. *Then* $(x^*, 0)$ *is the equilibrium in the game* Γ_+.

Examples.

1. Gamma distribution with

$$f(z) = \frac{1}{\Gamma(a)} z^{a-1} e^{-z}, \ z \geqslant 0, \ a > 0.$$

It is known [3] that $\mathrm{Mo} = a - 1 < \mathrm{Me} < a$. Calculations show that (1) is true if $a > \hat{a} \approx 0.4148$ and $\mathrm{Me} + y_0 \geqslant 0$ is true if $a > \tilde{a} \approx 0.615$. In last case (x_0, y_0) is the equilibrium in the game Γ. One can prove that (1) holds if $a \geqslant 1$. Also calculations show that $z^* = x^*/2 > \mathrm{Me}$ if $a < a^* \approx 1.8318$. So, in the last case by theorem 3 $(x^*, 0)$ is the equilibrium in the game Γ_+. Under $a \geqslant a^*$ (x_0, y_0) is the equilibrium in the game Γ_+.

2. Weibullah-Gnedenko distribution with

$$F(x) = 1 - \exp(-z^c), \ z \geqslant 0, \ c > 0.$$

Here

$$\mathrm{Me} = (\ln 2)^{1/c}, \mathrm{Mo} = \begin{cases} ((c-1)/c)^{1/c}, & c \geqslant 1; \\ 0, & 0 < c < 1. \end{cases}$$

The inequality (1) is true iff $c > (1+\ln 2)^{-1}$ and $\mathrm{Me}+y_0 \geqslant 0$ iff $c \geqslant (2\ln 2)^{-1}$. In last case (x_0, y_0) is the equilibrium in the game Γ. Further if $0 < c < (\ln 2)^{-1}$ $(c \geqslant (\ln 2)^{-1})$ then $(x^*, 0)$ $((x_0, y_0))$ is the equilibrium in the game Γ_+.

3. Lognormal distribution with

$$f(z) = \frac{1}{\sigma z\sqrt{2\pi}} \exp\Big(-\frac{1}{2}\Big(\frac{\ln z - a}{\sigma}\Big)^2\Big),\ z > 0,\ \sigma > 0.$$

Here $\mathrm{Me} = \exp(a)$, $\mathrm{Mo} = \exp(a-\sigma^2)$. The pair (x_0, y_0) is the equilibrium in the game Γ iff $0 < c < 2\sqrt{2/\pi}$. Further if $\sigma > \sqrt{2/\pi}$ $(\sigma \leqslant \sqrt{2/\pi})$ then $(x^*, 0)$ $((x_0, y_0))$ is the equilibrium in the game Γ_+.

References

1. Stevens C.M. Is compulsory arbitration compatible with bargaining? // Industrial Relations. 1966. V. 5, N 2. P. 38–52.
2. Brams S.J., Merrill S. Equilibrium strategies for final-offer arbitration: there is no median convergence // Management Science. 1983. V. 29, N 8. P. 927–941.
3. Choi K.P. On the medians of gamma distributions and an equation of Ramanujan // Proceeding of American Mathematical Society. 1994. V. 121, N 1. P. 245–251.

Survey of statistical games of parameter estimation with linear minimax rules*

V.V. Morozov and V.A. Gribov

Lomonosov Moscow State University, Moscow, Russia

Consider density function $f(x|\theta)$ with unknown parameter θ. Statistician observes sequence of independent identical distributed random variables $X = (X_1, ..., X_n)$, $X_i \sim f(x|\theta)$, and estimates value $h(\theta)$ by decision rule $\delta(X)$. Let $\delta_0(X)$ designates unbiased estimate that is $\mathrm{E}[\delta_0(X)|\theta] = h(\theta)$. We restrict ourselves with games where linear minimax rule $\delta(X) = c_1\delta_0(X) + c_2$ exists among all class of decision rules Δ. We use risk function $R(\theta, \delta)$ $=\mathrm{E}[w(\theta)(h(\theta) - \delta(X))^2|\theta]$ with weight function $w(\theta)$. For a distribution $\xi(\theta)$ of random variable Θ define a mean risk function $R(\xi, \delta) = \mathrm{E}_\xi[R(\Theta, \delta)]$. In the statistical game $\Gamma = \langle\Lambda, \Delta, R(\xi, \delta)\rangle$ the nature (first player) chooses a prior distribution

*The reported study was funded by RFBR according to the research project 16-01-00353 a.

$\xi(\theta) \in \Lambda$ and the statistician (the second player) chooses a minimax estimation δ^* minimizing the worst value of risk function $\max\limits_{\theta} R(\theta, \delta)$.

For given distribution $\xi(\theta) \in \Lambda$ an estimation δ_ξ is said Bayesian if it minimizes $R(\xi, \delta)$ on $\delta \in \Delta$. Let $g(\theta|X)$ be a posterior distribution of the variable Θ. It is known (see for example [1]) that

$$\delta_\xi(X) = \frac{\mathrm{E}[w(\Theta)h(\Theta)|X]}{\mathrm{E}[w(\Theta)|X]}.$$

An estimation δ is said equalizing if $R(\theta, \delta)$ doesn't depend on θ. If an equalizing estimate is the limit of Bayesian ones then it is the minimax estimate.

1. The Poisson distribution $f(x|\theta) = \theta^x e^{-\theta}/x!,\ x \in \mathbb{Z}_+,\ \theta > 0$. Here $\mathrm{E}[X_i|\theta] = \mathrm{Var}[X_i|\theta] = \theta,\ i = 1, ..., n$. Let be $h(\theta) = \theta,\ w(\theta) = 1/\theta$. Then the unbiased estimate $\delta_0(X) = \bar{X} = (X_1 + ... + X_n)/n$ is equalizing with risk value $v = 1/n$. It's the limit of Bayesian estimates

$$\delta_g(X) = \frac{1}{\mathrm{E}[1/\theta|X]} = \frac{a - 1 + n\bar{X}}{b + n}$$

corresponding to the gamma distribution $g(\theta) = b^a \theta^{a-1} e^{-b\theta}/\Gamma(a)$ when $a \downarrow 1, b \downarrow 0$. So, $\bar{X}$ is the minimax estimate.

Let be $w(\theta) = 1$. In this case the risk function $R(\theta, \bar{X})$ is unrestricted from above. That's why instead of Λ the statistician may suppose that distribution $\xi(\theta) \in \Lambda(M) = \{\xi(\theta)|\mathrm{E}_\xi[\theta] \leqslant M\}$ where M is a known constant. Here $\bar{X}$ is also minimax estimate with risk value $v = M/n$.

2. Exponential distribution $f(x|\theta) = \theta e^{-\theta x},\ x \geqslant 0,\ \theta > 0$. Here $\mathrm{E}[X_i|\theta] = 1/\theta$, $\mathrm{Var}[X_i|\theta] = 1/\theta^2$. Let be $h(\theta) = \theta,\ w(\theta) = 1/\theta^2$.

Note that a random variable $Y = X_1 + ... + X_n$ has Erlang distribution with density function $p_n(y|\theta) = \theta^n y^{n-1} e^{-\theta y}/(n-1)!,\ y \geqslant 0$. Under $n > 2$ $\mathrm{E}[1/\bar{X}|\theta] = n\theta/(n-1)$, $\mathrm{E}[1/\bar{X}^2|\theta] = n^2\theta^2/((n-1)(n-2))$. For any distribution $\xi(\theta) \in \Lambda$ and an estimate $\delta(X) = k/\bar{X}$ we have $R(\xi, \delta) = 1 - 2kn/(n-1) + k^2n^2/((n-1)(n-2))$. The value $v = 1/(n-1)$ is a minimal value of the last expression which it reaches on $k^* = (n-2)/n$. The equalizing estimate $\delta^*(X) = (n-2)/(n\bar{X})$ is the limit of Bayesian estimates

$$\delta_g(X) = \frac{\mathrm{E}[1/\theta|X]}{\mathrm{E}[1/\theta^2|X]} = \frac{a + n - 2}{b + n\bar{X}}$$

corresponding to the gamma distribution $G(a, b)$ when $a \downarrow 0, b \downarrow 0$. So, $\delta^*(X)$ is the minimax estimate.

If we want to estimate a mean of the distribution $f(x|\theta)$ then we take $h(\theta) = 1/\theta,\ w(\theta) = \theta^2$. Here the minimax estimate is $\delta^*(X) = n\bar{X}/(n+1),\ n \geqslant 1$. The value of the game is $v = 1/(n+1)$.

3. Normal distribution $f(x|\theta) = \sqrt{\theta}e^{-\theta x^2/2}/\sqrt{2\pi}$, where θ is a measure of exactness ($1/\theta$ is a variance). Let be $h(\theta) = 1/\theta,\ w(\theta) = \theta^2$. Denote $S = X_1^2 + ... + X_n^2$. We have $\mathrm{E}[S/n|\theta] = 1/\theta,\ \mathrm{E}[S^2|\theta] = (3n + n(n-1))/\theta^2$.

For any distribution $\xi(\theta) \in \Lambda$ and an estimate $\delta(X) = kS/n$ we have $R(\xi, \delta) = 1-2k+k^2(n+2)/n$. The value $v = 2/(n+2)$ is a minimal value of the last expression which it reaches on $k^* = n/(n+2)$. The equalizing estimate $\delta^*(X) = k^*S/n = S/(n+2)$ is the limit of Bayesian estimates $\delta_g(X) =\mathrm{E}[\theta|X]/\mathrm{E}[\theta^2|X] = (b + S/2)/(b + n/2 + 1)$ corresponding to the gamma distribution $G(a, b)$ when $a \downarrow 0, b \downarrow 0$. So, $\delta^*(X)$ is the minimax estimate.

Let be $h(\theta) = 1/\sqrt{\theta},\ w(\theta) = \theta$. A random variable $S' = \sqrt{\theta S}$ has χ^2_n distribution. For any distribution $\xi(\theta) \in \Lambda$ and an estimate $\delta(X) = k\sqrt{S}$ we have $R(\xi, \delta) =\mathrm{E}[\theta(1/\sqrt{\theta} - \delta(X))^2] = 1 - 2k\mathrm{E}[S'] + k^2\mathrm{E}[(S')^2]$. A value $k^* =\mathrm{E}[S']/\mathrm{E}[(S')^2] = (\sqrt{2}/n)\Gamma((n + 1)/2)/\Gamma(n/2)$ minimizes the last expression. The equalizing estimate $\delta^*(X) = k^*\sqrt{S}$ is a limit of Bayesian estimates

$$\delta_g(X) = \frac{\mathrm{E}[\sqrt{\theta}|S]}{\mathrm{E}[\theta|S]} = \frac{(b + S/2)^{1/2}\Gamma(a + (n + 1)/2)}{(a + n/2)\Gamma(a + n/2)}$$

corresponding to the gamma distribution $G(a, b)$ when $a \downarrow 0, b \downarrow 0$. So, $\delta^*(X)$ is the minimax estimate.

The estimate $\bar{X}$ is minimax for the estimation of the mean when the variance is known [2].

3. Uniform distribution $f(x|\theta) = I_{[0,\theta]}(x)/\theta$, Denote $Z = \max\limits_{i=1,\dots,n} X_i$, $\mathrm{E}[Z|\theta] = n\theta/(n+1)$, $\mathrm{E}[Z^2|\theta] = n\theta^2/(n+2)$. Let be $h(\theta) = \theta,\ w(\theta) = 1/\theta^2,\ \delta(X) = kZ$. Then for any distribution $\xi(\theta) \in \Lambda$

$$R(\xi, \delta) = \mathrm{E}\Big[\frac{(\theta - \delta(X))^2}{\theta^2}\Big] = 1 - \frac{2nk}{n+1} + \frac{nk^2}{n+2}.$$

A value $k^* = (n + 2)/(n + 1)$ minimizes the last expression. Take the prior Pareto distribution with density function

$$g(\theta) = \begin{cases} \alpha\theta_0^\alpha/\theta^{\alpha+1}, & \theta > \theta_0, \\ 0, & \theta \leqslant \theta_0. \end{cases}$$

A posterior distribution is

$$g(\theta|X)=\begin{cases}(\alpha+n)(Z')^{\alpha+n}/\theta^{\alpha+n+1}, & \theta>Z',\\ 0, & \theta\leqslant Z'\end{cases}$$

where $Z'=\max(\theta_0,Z)$ (see [4]). A Bayesian estimate

$$\delta_g(X)=\frac{\mathrm{E}[1/\theta|X]}{\mathrm{E}[1/\theta^2|X]}=\frac{(\alpha+n+2)}{\alpha+n+1}Z'$$

converges to the equalizing estimate $\delta^*(X)=(n+2)Z/(n+1)$ while $\theta_0\downarrow 0,\ \alpha\downarrow 0$. So, $\delta^*(X)$ is the minimax estimate.

4. Binomial distribution $f(t|\theta)=C_n^t\theta^t(1-\theta)^{n-t},\ t=0,1,...,n$. Here θ is a probability of success in Bernoulli trials. Denote $T=X_1+...+X_n$ where $X_i=1$ if it's a success in ith trial and $X_i=0$ otherwise. T/n is an unbiased estimate of θ. Let be $h(\theta)=\theta,\ w(\theta)=1$. Then in [3] it is shown that an estimate

$$\delta^*(T)=\frac{T}{n+\sqrt{n}}+\frac{1}{2(\sqrt{n}+1)}=\delta_g(T)=\mathrm{E}[\theta|T]=\frac{a+T}{a+b+n}$$

is equalizing and Bayesian corresponding to beta distribution $g(\theta)=\theta^{a-1}(1-\theta)^{b-1}/B(a,b)$ where

$$B(a,b)=\int_0^1\theta^{a-1}(1-\theta)^{b-1}d\theta,\ a=b=\frac{\sqrt{n}}{2}.$$

So, $\delta^*(T)$ is the minimax estimate.

Let be $h(\theta)=\theta(1-\theta),\ w(\theta)=1$. An estimate $\delta^*(T)=c_1^*\delta_0(T)+c_2^*$ where

$$\delta_0(T)=\frac{T(n-T)}{n(n-1)},\ c_1^*=\frac{1}{1+\sqrt{\omega}},\ c_2^*=\frac{1}{2n(\omega+\sqrt{\omega})},\ \omega=\frac{4n-6}{n(n-1)}$$

is equalizing and Bayesian for some discrete distribution on $[0,1]$ for $n=3,5,6,...,13$. So, $\delta^*(T)=c_1^*\delta_0(T)+c_2^*$ is the minimax estimate [4]. If $w(\theta)=1/h(\theta)$ then an estimate

$$\delta^*(T)=k^*\delta_0(T),\ k^*=\sqrt{\frac{n(n-1)}{n(n-1)+2(2n-3)}}$$

is minimax for $n=2,4,...,11$ [5].

5. Negative binomial distribution $f(t|\theta) = C^t_{r+t-1}\theta^r(1-\theta)^t,\ t \in \mathbb{Z}_+,\ r > 0$. Let be

$$h(\theta) = \mathrm{E}[T|\theta] = \frac{r(1-\theta)}{\theta},\ w(\theta) = \frac{r}{\mathrm{Var}[T|\theta]} = \frac{\theta^2}{1-\theta}.$$

Then the estimate $\delta^*(T) = rT/(r+1)$ is minimax [6].

References

1. Gren J. Statistical Games and their Applications (translated from the Polish), Moscow: Statistika, 1975.
2. Wolfowitz Minimax estimates of the mean of a normal distribution with known variance // The Annals of Mathematical Statistics. 1950. V. 21, N 2. P. 218–230.
3. Hodges J.L., Lehmanm E.L. Some problems in minimax point estimation // Annals of Mathematical Statistics. 1950. V. 21, N 2. P. 182–192.
4. DeGroot M.H. Optimal statistical decisions (translated from English). Moscow: Mir, 1974.
5. Ferguson T.S, Kuo L. Minimax estimation of a variance// Annals of the Institute of Statistical Mathematics. 1994. V. 46, N 2. P. 295–308.
6. Ferguson T.S. Mathematical statistics. New-York and London: Academic Press, 1975.

Dynamic multicriteria games with finite horizon[*]

A.N. Rettieva
Institute of Applied Mathematical Research,
Karelian Research Centre of RAS,
Saint Petersburg State University,
Petrozavodsk, Saint Petersburg, Russia

Mathematical models involving more than one objective seem more adherent to real problems. Players can have more that one goal which are often not comparable. These situations are typical for game-theoretic models in economics and ecology. Traditionally, equilibrium analysis in

[*] This research is supported by the Russian Science Foundation, project no. 17-11-01079.

multicriteria problems is based on the static variant. Some concepts have been suggested to solve multicriteria games (the ideal Nash equilibrium [5], the E-equilibrium concept [1]). However the notion of Pareto equilibrium [4] is the most studied concept in multicriteria game theory.

In [3] a new approach to construct noncooperative equilibrium in dynamic multicriteria games is presented. A multicriteria Nash equilibrium is constructed adopting the bargaining concept (via Nash products) with the guaranteed payoffs playing the role of the status quo points.

This work is dedicated to linking multicriteria games with cooperative dynamic games. A new approach to obtain the cooperative equilibrium in dynamic games with many objectives is proposed. To construct cooperative behavior we adopt the approach presented for the game-theoretic models with asymmetric players [4]. Namely, we obtain cooperative strategies and payoffs in multicriteria dynamic game via Nash bargaining solution. The multicriteria Nash equilibrium payoffs play the role of the status quo points.

Consider a multicteria dynamic game with two participants in discrete time. The players exploit a common resource and both wish to optimize m different criteria. The state dynamics is in the form

$$x_{t+1} = f(x_t, u_{1t}, u_{2t})\,, \ \ x_0 = x\,, \tag{1}$$

where $x_t \geqslant 0$ is the resource size at time $t \geqslant 0$, $f(x_t, u_{1t}, u_{2t})$ gives the natural growth function, and $u_{it} \in U_i$ denotes the strategy of player i at time $t \geqslant 0$, $i = 1, 2$.

The payoff functions of the players over the finite time horizon are defined by

$$J_1 = \begin{pmatrix} J_1^1 = \sum\limits_{t=0}^{n} \delta^t g_1^1(u_{1t}, u_{2t}) \\ \cdots \\ J_1^m = \sum\limits_{t=0}^{n} \delta^t g_1^m(u_{1t}, u_{2t}) \end{pmatrix}, \ J_2 = \begin{pmatrix} J_2^1 = \sum\limits_{t=0}^{n} \delta^t g_2^1(u_{1t}, u_{2t}) \\ \cdots \\ J_2^m = \sum\limits_{t=0}^{n} \delta^t g_2^m(u_{1t}, u_{2t}) \end{pmatrix}, \tag{2}$$

where $g_i^j(u_{1t}, u_{2t}) \geqslant 0$ gives the instantaneous utility, $i = 1, 2$, $j = 1, \ldots, m$, and $\delta \in (0, 1)$ denotes a common discount factor.

We design the equilibrium in dynamic multicriteria game applying the Nash bargaining products [3]. Therefore, we begin with the construction of guaranteed payoffs which play the role of status quo points.

There are three possible concepts to determine the guaranteed payoffs. In the first one four guaranteed payoff points are obtained as the solutions of zero-sum games. In particular, the first guaranteed payoff point is a solution of zero-sum game where player 1 wishes to maximize her first criterion and player 2 wants to minimize it. Other points are obtained by analogy. Namely,
G_1^j is the solution of zero-sum game $\langle I, II, U_1, U_2, J_1^j \rangle$, $j = 1, \ldots, m$,
G_2^j is the solution of zero-sum game $\langle I, II, U_1, U_2, J_2^j \rangle$, $j = 1, \ldots, m$.

The second approach can be applied when the players' objectives are comparable. Consequently, the guaranteed payoff points for player 1 (G_1^1,...,G_1^m) are obtained as the solution of a zero-sum game where she wants to maximize the sum of her criteria and player 2 wishes to minimize it. And, by analogy, for player 2. Namely,
$G_1^1, \ldots, G_1^m$ are the solution of zero-sum game $\langle I, II, U_1, U_2, J_1^1 + \ldots + J_1^m \rangle$,
$G_2^1, \ldots, G_2^m$ are the solution of zero-sum game $\langle I, II, U_1, U_2, J_2^1 + \ldots + J_2^m \rangle$.

In the third approach the guaranteed payoff points are constructed as the Nash equilibrium with the appropriate criteria of both players, respectively. Namely,
G_1^1 and G_2^1 is the Nash equilibrium in the game $\langle I, II, U_1, U_2, J_1^1, J_2^1 \rangle$,
...
G_1^m and G_2^m is the Nash equilibrium in the game $\langle I, II, U_1, U_2, J_1^m, J_2^m \rangle$.

To construct multicriteria payoff functions we adopt the Nash products. The role of the status quo points belongs to the guaranteed payoffs of the players:

$$H_1(u_{1t}, u_{2t}) = (J_1^1(u_{1t}, u_{2t}) - G_1^1) \cdot \ldots \cdot (J_1^m(u_{1t}, u_{2t}) - G_1^m)\,, \qquad (3)$$
$$H_2(u_{1t}, u_{2t}) = (J_2^1(u_{1t}, u_{2t}) - G_2^1) \cdot \ldots \cdot (J_2^m(u_{1t}, u_{2t}) - G_2^m)\,. \qquad (4)$$

Definition 1. *A strategy profile (u_{1t}^N, u_{2t}^N) is called a multicriteria Nash equilibrium of the problem (1),(2) if*

$$H_1(u_{1t}^N, u_{2t}^N) \geqslant H_1(u_{1t}, u_{2t}^N) \;\; \forall u_{1t} \in U_1\,, \qquad (5)$$
$$H_2(u_{1t}^N, u_{2t}^N) \geqslant H_2(u_{1t}^N, u_{2t}) \;\; \forall u_{2t} \in U_2\,. \qquad (6)$$

Under the presented equilibrium concept players maximize the product of the differences between the optimal and guaranteed payoffs (3),(4).

To construct cooperative behavior we adopt the approach presented for the game-theoretic models with asymmetric players [4]. Namely, the multicriteria cooperative equilibrium is obtained as a solution of a Nash bargaining scheme with the multicriteria Nash equilibrium playing the role of status quo points.

First we have to determine noncooperative payoffs as players' gains when they apply multicriteria Nash strategies (u_{1t}^N, u_{2t}^N):

$$J_1^N = \begin{pmatrix} J_1^{1N} = \sum\limits_{t=0}^{n} \delta^t g_1^1(u_{1t}^N, u_{2t}^N) \\ \ldots \\ J_1^{mN} = \sum\limits_{t=0}^{n} \delta^t g_1^m(u_{1t}^N, u_{2t}^N) \end{pmatrix}, J_2^N = \begin{pmatrix} J_2^{1N} = \sum\limits_{t=0}^{n} \delta^t g_2^1(u_{1t}^N, u_{2t}^N) \\ \ldots \\ J_2^{mN} = \sum\limits_{t=0}^{n} \delta^t g_2^m(u_{1t}^N, u_{2t}^N) \end{pmatrix}. \qquad (7)$$

Then we construct Nash product where the sum of players' noncooperative payoffs plays a role a status quo points. To construct the cooperative behavior we adopt Nash bargaining solution, so it is required to solve the next problem

$$(V_1^{1c} + V_2^{1c} - J_1^{1N} - J_2^{1N}) \cdot \ldots \cdot (V_1^{mc} + V_2^{mc} - J_1^{mN} - J_2^{mN}) =$$
$$= (\sum_{t=0}^{n} \delta^t (g_1^1(u_{1t}^c, u_{2t}^c) + g_2^1(u_{1t}^c, u_{2t}^c)) - J_1^{1N} - J_2^{1N}) \cdot \ldots$$
$$\cdot (\sum_{t=0}^{n} \delta^t (g_1^m(u_{1t}^c, u_{2t}^c) + g_2^m(u_{1t}^c, u_{2t}^c)) - J_1^{mN} - J_2^{mN}) \to \max_{u_{1t}^c, u_{2t}^c}\,, \qquad (8)$$

where J_i^{jN} are the noncooperative gains determined in (7), $i = 1, 2$, $j = 1, \ldots, m$.

The next definition presents the suggested solution concept.

Definition 2. *A strategy profile* (u_{1t}^c, u_{2t}^c) *is called a multicriteria cooperative equilibrium of the problem (1),(2) if it solves the problem (8).*

A dynamic multicriteria model related with the bioresource management problem (harvesting) is investigated to show how the suggested concepts work.

References

1. Pusillo L., Tijs S. E-equilibria for multicriteria games // Annals of the International Society of Dynamic Games. 2013. V. 12. P. 217–228.
2. Rettieva A.N. A discrete-time bioresource management problem with asymmetric players // Automation and Remote Control. 2014. V. 75, N 9. P. 1665–1676.
3. Rettieva A.N. Equilibria in dynamic multicriteria games // International Game Theory Review. 2017. V. 19, N 1. P. 1750002.
4. Shapley L.S., Rigby F.D. Equilibrium points in games with vector payoffs // Naval Research Logistic Quarterly. 1959. V. 6, N 1. P. 57–61.
5. Voorneveld M., Grahn S. and Dufwenberg M. Ideal equilibria in noncooperative multicriteria games // Mathematical Methods of Operations Research. 2000. V. 52, N 1. P. 65–77.

Axiomatic approach to conspiracy theory

M.A. Savchenko
Moscow State University, Moscow, Russia

Lets consider n–player normal–form games with incomplete information. Axiomatic foundation for analysis of such games was developed by Robert J. Aumann in [1]. His correlated equilibrium is powerful abstraction that allows modeling of very different conflict situations with informational asymmetry. We are focusing on one particular kind of such situations — games were groups of players can secretly use shared strategies to improve their individual rewards. For example, constant sum game of three players with following payouts:

$\frac{1}{3}, \frac{1}{3}, \frac{1}{3}$	$\frac{1}{2}, 0, \frac{1}{2}$
$0, \frac{1}{2}, \frac{1}{2}$	$\frac{1}{2}, \frac{1}{2}, 0$

$\frac{1}{2}, \frac{1}{2}, 0$	$0, \frac{1}{2}, \frac{1}{2}$
$\frac{1}{2}, 0, \frac{1}{2}$	$\frac{1}{3}, \frac{1}{3}, \frac{1}{3}$

Here first player picks the row, second the column and third the matrix. All classical, non–correlated Nash equilibria in this game are predicting payouts of $(\frac{1}{3}, \frac{1}{3}, \frac{1}{3})$. On the other side, if players are allowed to communicate privately, we can easily imagine different viable outcomes. For example, first and second player can secretly arrange synchronous choice between top–left and bottom–right cells, while third player won't be able to adjust his choice of matrix accordingly, giving payout of $(\frac{1}{2}, \frac{1}{2}, 0)$ when he guesses wrong. Thanks to Aumann, this payout can be predicted by Nash equilibrium in correlated strategies with world consisting of one event with two equally probable outcomes (fair coin toss), which can be distinguished by first and second player, but not by the third, leading to payout of $(\frac{5}{12}, \frac{5}{12}, \frac{1}{6})$.

Correlated equilibrium as model of private communication between players seems natural and intuitive in simple cases like above, but, when one looks closer, noticeable gap in interpretation appears. Our choice of parameters for Aumann's solution was pretty much intuitive — no formal theory prescribes usage of this particular world of fair coin toss to be model of such collusion. Moreover, there is infinite amount of parameter sets, which lead to same payout prediction — for example, we can replace coin toss with roulette here. Unfortunately it means that in more complex cases (larger number of players, sophisticated communication structure, etc) we can't be sure that intuitively designed correlated extension actually represents situation in question.

To address this difficulties we propose new formal model using notion of probability spaces [2] and two core abstractions derived from Aumann's model. First one is *correlation space* $\Phi = \langle A, \Omega, \mathfrak{I}^a, \mathbb{P}, a \in A \rangle$ that represents set of parameters for correlated extension of any normal–form game with A as player set. Here Ω is sample space of the world with probability measure $\mathbb{P}$. For any player a σ–field $\mathfrak{I}^a$ represents set of events regarding which a is informed, making triplet $\langle \Omega, \mathfrak{I}^a, \mathbb{P} \rangle$ into his *personal correlation subspace*. Basically, this abstraction relates to Aumann's correlated extension without its subjectivity aspect.

Seconly, for each group of players A_* we take notice of its *shared* *secret* $\langle \Omega, \kappa(A_*), \mathbb{P} \rangle$ — correlation subspace, where

$$\kappa(A_*) = \{X \in \bigcap_{a \in A_*} \mathfrak{I}^a \mid \mathbb{P}(X \cap Y) = \mathbb{P}(X)\mathbb{P}(Y), \forall Y \in \sigma(\bigcup_{a \in A \setminus A_*} \mathfrak{I}^a)\}$$

is set of events, regarding which all members of A_* are informed, while all the others can get no hint, even by pooling their knowledge. For convenience we also define

$$\eta(\mathfrak{X}) = \sigma(\bigcup_{X \in \mathfrak{X}} \kappa(X)), \forall \mathfrak{X} \subseteq 2^A$$

that is σ–field, generated by joining σ–fields in shared secrets belonging to arbitrary set of player groups.

On this basis we introduce *conspiracy* of *structure* $\mathfrak{A} \subseteq 2^A$ — any correlation space with following properties:

1. For each non–empty subset of players A_*, its shared secret has either measure consisting of singular atom ($A_* \notin \mathfrak{A}$, interpreted as group without means of private communication), or diffuse measure ($A_* \in \mathfrak{A}$, for groups capable of secretly sharing any amount of information);

2. $\mathfrak{I}^a = \eta(\mathfrak{A}^a), \forall a \in A$, where $\mathfrak{A}^a = \{X \in \mathfrak{A} \mid a \in X\}$ — simply stated, every player's personal correlation subspace is generated by joining shared secrets of all groups he belongs to.

Let's return to our example game to use it as showcase for new abstractions. What would be the conspiracy structure in situation where 1st and 2nd player are colluding against 3rd? Firstly, conspiracy structure includes all singleton player groups $\{1\}$, $\{2\}$ and $\{3\}$, because all players can individually use uncorrelated mixed strategies. Secondly, it includes group $\{1, 2\}$, as 1st and 2nd player should be able to share information privately. Finally, remaining groups $\{1, 3\}$, $\{2, 3\}$ and $\{1, 2, 3\}$ are not included in conspiracy structure, as problem statement says nothing about 3rd player ability to communicate with anyone else. This gives us conspiracy structure $\{\{1\}, \{2\}, \{3\}, \{1, 2\}\}$.

What easily constructible correlation space can serve as example of conspiracy with this structure? Most natural answer is world of 4 independent roulette wheels, one per group in conspiracy structure. With that, personal correlation subspaces are defined in such a way that 1st (2nd and 3rd respectively) player is informed regarding state of 1st (2nd and 3rd respectively) roulette, plus 1st and 2nd players are informed regarding 4th roulette. Original game extended with such correlation space can be solved in terms of ordinary correlated Nash equilibrium, successfully predicting payout of $(\frac{5}{12}, \frac{5}{12}, \frac{1}{6})$.

Why should we give some special attention to solution with this set of parameters? Importance of conspiracies subclass among all correlation spaces arises from following

Theorem 1 *For any normal–form game with finite strategy sets all conspiracies of same structure produce identical sets of payouts in correlated Nash equilibria.*

This theorem allows us to abstract away from specific probability spaces, highlighting structure of conspiracy as its only meaningful characteristic. Game solved for any correlation space, which is conspiracy of given structure, becomes automatically solved for whole class of correlation spaces, that are conspiracies of same structure. World of one independent roulette per each group in conspiracy structure emerges as canonic recipe for prediction of important outcomes in any game, which interpretation is consistent with definition of conspiracy.

Also, model of conspiracy gives opportunity to ask further questions:

- if we have a game with player set A, can there be correlated Nash equilibrium with payout not achievable in conspiracy of structure 2^A?
- can this model be extended for games with more complex player sets (population games, for example)?
- what about different solution concepts, such as strong or coalition-proof Nash equilibrium?

References

1. Aumann, Robert J. Subjectivity and correlation in randomized strategies // Journal of Mathematical Economics. 1974. V. 1, N 1. P. 67–96.
2. Kolmogorov A.N. Foundations of the Theory of Probability (2nd ed.). New York: Chelsea, 1956.

The repeated game modeling an agreement on protection of the environment*

A.A. Vasin and A.G. Divtsova
Moscow State University, Moscow, Russia

We consider repeated games with sliding planning horizons. The initial game includes two stages: the first stage is a game that generalizes the known model "The Tragedy of the Commons" [1] . At the second stage the players redistribute the payoffs by means of side payments. Our purpose is to find conditions for existance of a subgame perfect equilibrium (SPE) realizing in the repeated game some Pareto-optimal outcome, that is, an outcome that maximizes the total payoff in the one-short game. The problem is of interest in context of the study of international ageements on limitation of environmental pollution (see [2,3]). Existance of the SPE means the possibility for a stable and efficient agreement of such sort. This concept takes into account a possibility of unexpected breaking the agreement by some country and assumes that only endogeneous economic mechanisms in frame of the agreement prevent such breaking. Side payments reflect possibilities for redistribution of the welfare among the players. Note that in the one-short game there is a "bad" Nash equilibrium in dominant strategies corresponding to a high pollution level. We examine two types of SPE realizing some Pareto-optimal outcome: 1) after any deviation, all players start playing dominant strategies; 2) if one player deviates, the rest continue cooperation maximizing their total payoff under the dominant strategy of the disturber; after the second deviation everybody plays his dominant strategy. For each type, we determine the set of Pareto-optimal SPE outcomes.

We consider the game $\Gamma = \langle\ I,\ Z_i\ ,\ F_i(z),\ i\ \in\ I\ \rangle$ with the set of players $I = \{1, ..., n\}$. Strategy $z_i \in Z_i = [0, z_{max_i}]$ of player i shows the pollution amount induced by his production activity. His payoff function $F_i(z) = v_i(z_i) - H_i(\sum\limits_{j\in I} \pi_{ji} z_j)$, $z = (z_1, ..., z_n)$, aggregates his utility function $v_i(z_i)$ that relates to production activity, is monotonous and concave, and loss function $H_i(\hat{z}_i)$ that is monotonous and convex and dependes on the total pollution level $\hat{z}_i = \sum_{j\in I} \pi_{ji} z_j$ for player i. Here $\pi_{ji} \in [0;1], j, i \in I$, is the share of the pollution amount induced by player j and transferred to player i.

For player i, z_i^* is called a dominat strategy if, for any z, $F_i(z|z_i^*) \geqslant F_i(z)$. We assume that for each player there exists a dominating strategy

*This research is supported by RFBR 16-01-00353/16

in the game. Then z^* is Nash equilibrium. In particular, such equilibrium exists if $H_i(\hat{z}_i)$ is a linear function or $\pi_{ii} = 0$. Profile $\bar{z}$ is Pareto optimal if it realizes $\max_{\bar{z}} \sum_{i=1}^{n} F_i(\vec{z})$.

In reality the interaction repeats many times. We assume that players can make side payments in order to compensate utility losses to those who reduce the production activity and the corresponding pollution amount. Each player chooses his strategy at some period t is concerned with his total payoff till period $t + T_i$, where T_i is his planning horizon.

The interaction model is a repeated game with complete information on prehistory and sliding planning horizons $T_i, i = 1, ..., n$. Each period $t = 1, 2, ...$ of the interaction includes two stages: at stage $t1$ each player determines the amount of pollution z_i^t, $i \in I$ and gets payoff $F_i^t = F_i(z^t)$, $i \in I$. At stage $t2$ players get or pay side payments $y_i^t, i \in I, \sum\limits_i y_i^t = 0$. They determine final payoffs $F_i^t(z^t, y^t) = F_i(z^t) + y_i^t, i \in I$ for this period t. At every stage t each player choses his action proceeding from the history of the interaction up to this time: $h^{t-1} = (z^\tau, y^\tau)_{\tau=1}^{t-1}$ for stage $t1$. At stage $t2$ the realized situation z^t is also known to all the players. A strategy of player i is formally given by functions $z_i^t = \sigma_i^1(h^{t-1}), y_i^t = \sigma_i^2(h^{t-1}, z^t)$. Strategy profile σ^* is called an SPE of the repeated game if, for any i, t and h,

$$\sigma_i^* = arg \max_{\sigma_i} \sum_{\tau=t}^{t+T_i} F_i(z^\tau(\sigma^* || \sigma_i), y^\tau(\sigma^* || \sigma_i)),$$

Below we aim to find such SPE that realize Pareto optimal profile $\bar{z}$ in every period. We examine two types of such SPE. The strategies for the simpler type a order to play profile $\bar{z}$ and pay agreed side payments until some player deviates from this. Then the rest players stop side payments and switch to realization of equilibrium z^* in the following periods (see [2]). For an SPE of the type b, after the first deviation, the rest players continue cooperation maximizing their total payoff under the dominant strategy of the disturber. After the second deviation everybody plays his dominant strategy.

For any vector-function $(f_i(z), i \in I)$ introduce the following notation: $f_i^* := f_i(z^*), \bar{f}_i := f_i(\bar{z}), f_\Sigma(z) := \sum\limits_{i \in I} f_i(z), f_{\Sigma(I \setminus i)}(z) := \sum\limits_{j \in I \setminus i} f_j(z)$.

Theorem 1. An SPE of the type a exists if and only if

$$\sum_{i \in I} ((H_i^* - H_i(\sum_{j \in I \setminus i} \pi_{ji} \bar{z}_j + \pi_{ii} z_i^*))/(1 + T_i)) \leqslant \bar{F}_\Sigma - F_\Sigma^*.$$

Under this condition, side payments may be set as

$$y_i = F_i^* - \bar{F}_i - (H_i^* - H_i(\sum_{j\in I\backslash i} \pi_{ji}\bar{z}_j + \pi_{ii}z_i^*))/(1+\lambda T_i),$$

where $\lambda \leqslant 1$ is a solution of equation

$$\sum_{i\in I}((H_i^* - H_i(\sum_{j\in I\backslash i} \pi_{ji}\bar{z}_j + \pi_{ii}z_i^*))/(1+\lambda T_i)) = \bar{F}_\Sigma - F_\Sigma^*.$$

Let $\bar{\bar{z}}_{I\backslash i}(i) = arg \max\limits_{(z_j, j\in I\backslash i)} \sum\limits_{j\in I\backslash i} F_j(z_{I\backslash i}, z_i{}^*)$, $\bar{\bar{z}}(i) = (\bar{\bar{z}}_{I\backslash i}, z_i^*), \bar{\bar{F}}_j(i) = F_j(\bar{\bar{z}}(i))\ \forall i, j \in I.$

Theorem 2. An SPE of the type b exists if and only if:
1)$\bar{\bar{F}}_\Sigma \leqslant \bar{F}_\Sigma$, 2)$\bar{\bar{F}}_{\Sigma(I\backslash i)} \geqslant F^*_{\Sigma(I\backslash i)}, \forall i \in I$,
3) $\sum_{i\in I} \frac{H_i(\sum_{j\in I} \pi_{ji}\bar{\bar{z}}_j(i)) - H_i(\sum_{j\in I\backslash i} \pi_{ji}\bar{z}_j + \pi_{ii}z_i^*)}{1+T_i} \leqslant \bar{F}_\Sigma - \bar{\bar{F}}_\Sigma$,
4) $\sum_{j\in I\backslash i} \frac{H_j(\sum_{k\in I} \pi_{kj}z_k^*) - H_j(\sum_{k\in I\backslash\{i,j\}} \pi_{kj}\bar{\bar{z}}_k + \pi_{ij}z_i^* + \pi_{jj}z_j^*)}{1+T_j} \leqslant \bar{\bar{F}}_{\Sigma(I\backslash i)}(i) - F^*_{\Sigma(I\backslash i)}, i \in I.$

Note 1: Theorem 1 and Theorem 2 permit the following generalization for a model where planning horizons $T_i(t)$ depend on the time. The SPE specified in these propositions exist also in this case if the conditions hold for $T_i = \min\limits_t T_i(t)$. However, the conditions are not necessary for the existence in general. Another generalization relates to repeated games with time-dependent discounting coefficients. In such game a player i at every period t aims to maximize the present value of his future payoffs $\sum_{\tau=t}^{\infty} d^i_{\tau t}W_i(\tau)$, where $W_i(\tau)$ is his payoff at time τ, $d^i_{\tau t}$ is the discounting coefficient, $d^i_{tt} = 1$. The theorems hold as sufficient conditions for the SPE existence if we set $T_i(t) = \sum_{\tau=t+1}^{\infty} d^i_{\tau t}$.

References

1. Hardin G. The Tragedy of the Commons // Science, New Series. 1968. V. 162, N 3859. P. 1243–1248.
2. Vasin A.A., Divtsova A.G. Game-teoretic model of agreement on limitation of transboundary atmospheric polution // Matematicheskaya teoriya igr i prilozhenie. 2017. V. 9, N 1. P. 27–44. (in Russian).
3. Sacco A., Zaccour G. Impact of Social Externalities on the Formation of an International Environmental Agreement: An Exploratory Analysis. Working paper. 2016.

The typical models for congested traffic

The estimation of the capacity of the highway intersected by a crosswalk without traffic lights[*]

L. Afanaseva, M. Gridnev, and S. Grishunina
Department of Probability, Faculty of Mathematics and Mechanics, Lomonosov Moscow State University, Moscow, Russia

The study of traffic flows has a long history (see, e.g. [1–5] and references therein). Various methods such as cellular automate [6], statistical mechanics and mathematical physics [7–11] or queueing theory [12–20] were used.

The purpose of the proposed study is an estimation of the carrying capacity of the automobile road, crossed by a crosswalk. Under the capacity we mean the upper limit of the intensity of the flow of cars, when the queue of cars does not tend to infinity. This means that the stability condition for the process determining the number of these cars is satisfied. Our analysis will be based on the results obtained in [12, 21, 22].

Let's move to the description of models.

We consider a road with two directions of traffic and m traffic lanes in each. The flow of cars X_i in $i-th$ direction is a regenerative flow with intensity $\lambda_X^{(i)} (i = 1, 2)$ [17]. The road is intersected by a two-directions pedestrian crossing (pic.1). We denote that pedestrians following from

[*] This research is supported by Russian Foundation for Basic Research grant 17-01-00468

right side (A) to the left side (B) have the first type, and from B to A - the second type. The flow of pedestrians of $i - th$ type is a Poisson flow with intensity $\lambda_i (i = 1, 2)$. Pedestrians cross the crosswalk independently of each other with a random (but constant during the entire time of being at the crosswalk) speed.

We assume that there are no traffic lights at the crossing and pedestrians have an absolute priority over cars. At this case the number of pedestrians at the crosswalk is a number of customers for an infinite-channel queueing system of the $M|G|\infty$ type.

Let's assume that $2b$ is an average time of crossing the road by a pedestrian. Then the probability P_0 that there is no pedestrian at the crosswalk in a stationary regime is defined by an expression

$$P_0 = e^{-2(\lambda_1+\lambda_2)b} \tag{1}$$

(see [23], for example).

First assume that a car can cross a pedestrian crossing only if there are no pedestrians at it (Model 1). Let's number the traffic lanes in the direction from A to B so that cars at lanes $1, 2, ..., m$ are going at one direction and cars at lanes $m + 1, ..., 2m$ - at another. We will consider the process $Q_1(t)$ - the number of cars at the lanes $1, 2, ..., m$ at time t (the consideration of lanes $m + 1, m + 2, ..., 2m$ is analogous).Denote $H_j(t)$ as the expected value of the number of cars that pass through the crosswalk at the lane j during time t under the condition that there are always cars at this lane and the crosswalk is free. Also denote that $H(t) = \sum_{j=1}^{m} H_j(t)$. In relation to the process $Q_1(t)$, we have a single-channel queueing system with an unreliable server. The operating time has an exponential distribution with the parameter $\lambda_1 + \lambda_2$, and time u_2 is the period of the system $M|G|\infty$ being busy.

Since $P_0 = \frac{Eu_1}{Eu_1+Eu_2}$, then $a = Eu_1 + Eu_2 = (\lambda_1 + \lambda_2)^{-1} e^{2(\lambda_1+\lambda_2)b}$.

Then basing on results from [22] it is not difficult to show that the traffic rate ρ is determined by an expression

$$\rho_1 = \frac{\lambda_X^{(1)} e^{2(\lambda_1+\lambda_2)b}}{(\lambda_1 + \lambda_2)h(\lambda_1 + \lambda_2)}, \tag{2}$$

where $h(\lambda) = \lambda \int_0^\infty e^{-\lambda y} H(y)dy$.

The necessary and sufficient condition for the stability of the process $Q_1(t)$ means that the inequality $\rho_1 < 1$ is fulfilled, and the capacity of the road $\bar{\lambda}_X^{(1)}$ is defined as

$$\bar{\lambda}_X^{(1)} = (\lambda_1 + \lambda_2)h(\lambda_1 + \lambda_2)e^{-2(\lambda_1+\lambda_2)b}.$$

If, for example $H(t) = m\nu t$, which corresponds to the assumption that each car crosses the pedestrian crossing for an exponentially distributed time with a parameter ν, then

$$\bar{\lambda}_X^{(1)} = m\nu e^{-2(\lambda_1+\lambda_2)b}.$$

When the real intensity $\bar{\lambda}_X^{(1)}$ is less then, but close to $\bar{\lambda}_X^{(1)}$, large queues accumulate before the crosswalk.
Their asymptotic analysis, as well as expressions for the characteristics of the process $Q_1(t)$ in a stationary regime, when $\rho^{(1)} < 1$ can be found at the article [12].

Now we will consider Model 2, in which the rules for crossing the crosswalk by a car are weakened. Namely, we assume that the car can move along the lane j ($j = 1, 2, ..m$) if there are no pedestrians of the first type (going from A to B) on the lanes $1, 2, ...j$, and on the lanes with numbers $j, j+1, ..., 2m$ there are no pedestrians of the second type. Denote $P_0(j)$ as the probability of this event in a steady state.

Since the number of pedestrians of the first type on the lanes $(1, 2, ...j)$ - is the number of customers at the system $M|G|\infty$ with the intensity λ_2 and with an average queueing time $(2m - j + 1)\frac{b}{m}$, then

$$P_0(j) = e^{-j\lambda_1\frac{b}{m}-(2m-j+1)\lambda_2\frac{b}{m}}. \tag{3}$$

So we have a queueing system with m unreliable servers. All the servers break when a pedestrian of the first (second) type appears on the lane 1 ($2m - th$).This means that the avalible period $\tau_j^{(1)}$ of $j - th$ server is exponentially distributed with parameter $\lambda_1 + \lambda_2$. Let $\tau_j^{(2)}$ be a block period of $j - th$ server and $a_j = E\tau_j^{(1)} + E\tau_j^{(2)}$. Then $P_0(j) = \frac{E\tau_j^{(1)}}{a_j}$, so

$$a_j = (\lambda_1 + \lambda_2)^{(-1)} e^{j\lambda_1\frac{b}{m}+(2m-j+1)\lambda_2\frac{b}{m}}.$$

Assuming that $h_j(\lambda) = \lambda\int_0^\infty e^{-\lambda t}H_j(t)dt$ and using results from [22], we can find the traffic rate ρ_2 for Model 2

$$\rho_2 = \lambda_X^{(1)}[(\lambda_1 + \lambda_2)\sum_{j=1}^{m} P_0(j)h_j(\lambda_1 + \lambda_2)]^{-1}. \tag{4}$$

If $h_j(\lambda) = \frac{1}{m}h(\lambda), j = 1,\bar{m}$, then (4) can be written as

$$\rho_2(m) = \begin{cases} \frac{m\lambda_X^{(1)}}{(\lambda_1+\lambda_2)h(\lambda_1+\lambda_2)} \frac{e^{\lambda_2 \frac{b}{m}} - e^{\lambda_1 \frac{b}{m}}}{e^{-(\lambda_1+\lambda_2)b} - e^{-2\lambda_2 b}}, when\lambda_1 \neq \lambda_2 \\ \\ \frac{\lambda_X^{(1)} e^{2\lambda b + \frac{\lambda b}{m}}}{2\lambda h(2\lambda)}, when\lambda_1 = \lambda_2. \end{cases}$$

When $m = 1$ we get

$$\rho_2(1) = \frac{\lambda_X^{(1)} e^{2\lambda_2 b + \lambda_1 b}}{(\lambda_1 + \lambda_2)h(\lambda_1 + \lambda_2)} < \frac{\lambda_X^{(1)} e^{2(\lambda_2 + \lambda_1)b}}{(\lambda_1 + \lambda_2)h(\lambda_1 + \lambda_2)} = \rho_1.$$

It is easy to show that for all $m \geqslant 1$ the inequality $\rho_2(m) < \rho_1$ holds. Weakening the rules of crossing the crosswalk increases the capacity of the road. To estimate this effect, we consider the ratio

$$\frac{\rho_2(m)}{\rho_1} = \begin{cases} \frac{m(e^{\lambda_2 \frac{b}{m}} - e^{\lambda_1 \frac{b}{m}})}{e^{(\lambda_1+\lambda_2)b} - e^{2\lambda_1 b}}, when\lambda_1 \neq \lambda_2 \\ \\ e^{-2\lambda b + \frac{\lambda b}{m}}, when\lambda_1 = \lambda_2 = 1. \end{cases}$$

Assume $x = e^{\lambda_1 b} \geqslant 1$, $\lambda_2 = \alpha\lambda_1$, we have

$$\phi(x) = \frac{\rho_2(m)}{\rho_1} = \begin{cases} \frac{m(x^{\frac{\alpha}{m}} - x^{\frac{1}{m}})}{x^{1+\alpha} - x^2}, \ when\ \alpha \neq 1 \\ \\ x^{-2+\frac{1}{m}}, \ when\ \alpha = 1. \end{cases}$$

We have that the effect of the weakened rule (Model 2) in comparison with the standard rule (Model 1) becomes stronger, when the number of lanes and the intensity of the flow of pedestrians increase.

References

1. Gideon R., Pyke R. Markov renewal modelling of Poisson traffic at intersections having separate turn lanes// Janssen J., Limnios N. (eds) Semi-markov models and applications. Springer, New York. 1999. P. 285–310.
2. Greenberg H. An analysis of traffic flows// Oper. Res. 1959. V. 7. P.79–85.
3. Greenshields B.D. A study of highway capacity// Proc Highway Res. 1935. V. 14. P. 448–477.

4. Inose H., Hamada T. Road traffic control// University of Tokyo Press, Tokyo. 1975.
5. May A.D. Traffic flow fundamentals// Pretice Hall, Englewood Cliffs, NJ. 1990.
6. Maerivoet S., De Moor B. Cellular automata models of road traffic// Phys. Rep. 2005. V. 419. P. 1–64.
7. Blank M. Ergodic proprerties of a simple determenistic traffic flow model// J. Stat. Phys. 2003. V. 111. P. 903-930.
8. Choudhury D. Vehicular traffic: a system of interacting particles driven far from equilibrium. arXiv:arXiv:cond-mat/9910173 v1 [cond-mat-stat-mech] 12 Oct 1999.
9. Fuks H., Boccara N. Convergence to equilibrium in a class of interacting particle system envolving in a discrete time// Phys. Rev. 2001. E 64:016117.
10. Helbing D. Traffic and related self-driven many-particle systems// Rev. Mod. Phys. 2001. V. 73. P. 1067–1141.
11. Schadschneider A. Statistical physics of traffic flow. arXiv:arXiv:cond-mat/0007418 v1 [cond-mat-stat-mech] 26 July 2000.
12. Afanasyeva L.G., Bulinskaya E.V. Some problems for the flows of interacting particles// Modern problems of mathematics and mechanics. 2009. V. 2. P. 55–86.
13. Afanasyeva L.G., Bulinskaya E.V. Mathematical models of transport systems based on queueing methods// Proceeding of Moscow Institute of physics and technology. 2010. V. 2, 4. P. 6-21.
14. Afanasyeva L.G., Bulinskaya E.V. Stochastic models of transport flows// Commun. Stat. Theory Methods. 2011a. V. 40, 16. P. 2830–2846.
15. Afanasyeva L.G., Bulinskaya E.V. Estimation of transport System Capacity// The nineth international conference on traffic and granular flow. Abstract book. Moscow, Russia. 2011. P. 138–139.
16. Afanasyeva L.G., Bulinskaya E.V. Asymptotic analysis of the traffic performance under heavy traffic assumption. Methodology and Computing in Applied Probability. 2013. V. 45, 4. P. 935–950.
17. Afanasyeva L.G. and Rudenko I.V. $GI|G|\infty$ queueing systems and their applications to the analysis of traffic models// Theory Probability Applications. 2013. V. 57, 3. P. 1–21.
18. Baykal-Gursoy M., Xiao W. Stochastic decomposition in $M|M|\infty$ queues with Markov-modulated service rates// Queueing systems. 2004. V. 48. P. 75–88.

19. Baykal-Gursoy M., Xiao W., Ozbay K. Modeling traffic flow interrupted by incidents// Eur. J. Op. Res. 2009. V. 195. P. 127–138.
20. Caceres F.C., Ferrari P.A., Pechersky E. (2007) A slow to start traffic model related to $M|M|1$ queue// J. Stat. Mech. 2007. PO7008, arXiv:arXiv:cond-mat/0703709 v2 [cond-mat-stat-mech] 31 May 2007.
21. Afanasyeva L.G. Queueing systems with a regenerative input flow// Modern problems of mathematics and mechanics. 2015. V. X, 3. P. 23-26.
22. Afanasyeva L.G., Tkachenko A. Stability Analysis of Multiserver Queueing System with a Regenerative Interuption Process// Book of Abstracts, ASMDA, De Morgan House, London, UK. 2017. P. 12–13
23. Saaty T.L. Elements of queueing theory, with applications// Dover, New York. 1983.

Non-ergodicity and non-Markovianity of some simple traffic flow models

M.L. Blank
Russian Academy of Sci. Inst. for Information Transmission Problems, and National Research University Higher School of Economics, Moscow, Russia

The celebrated Nagel-Schreckenberg model allows to study a reasonably large class of one-dimensional traffic flow models with parallel updates. Unfortunately further generalizations of this construction turn out to be not especially fruitful. Namely, numerical simulations of such generalizations did not demonstrate stable behavior qualitatively different from the original model, and more to the point their mathematical treatment is still not available even up to nowadays. I'll discuss several features of these models which partially explain the failure of both numerical and mathematical attempts to study generalizations of the Nagel-Schreckenberg model.

First, let us recall the original Nagel-Schreckenberg model belonging to the class of the so called cellular automatons. The road with cars is represented either by be-infinite binary sequences of type $x := \dots 0011100 \dots$ (where by ones we mark positions of cars and by zeros their absence), or a finite binary sequence in the case of a finite cyclic road. In both cases the dynamics is defined as follows each car moves (with probability p) to the next position to the right if and only if this position is not occupied (i.e. the next to the right element in x is equal to 0). Since all the movements proceed independently and simultaneously, this type of models is called models with parallel updates. If the probability p is chosen to be equal to 1 we get a pure deterministic version of the process. In mathematical terms models of this sort are known as discrete time TASEP (totally asymmetric exclusion process). In this setting the model is well understood (see physics arguments in [1,2] and the complete mathematical analysis in [3]). Among other things this model demonstrates the classical gas-liquid phase transitions which in terms of traffic flows is related to free (when all cars are moving at full speed) and congested (when traffic jams are present all the time) types of flows.

The first generalization that we consider is the case with several lanes of cars. Again each lane is represented by a be-infinite binary sequence (periodic if the road is cyclic). To take into account the presence of neighboring lanes we assume that if a car is blocked (i.e. locally it is represented by 11), it can change the lane with the probability q under the condition that this change does not affect the movement of cars at that lane. This simple generalization of the one-lane model is indeed very natural and can be easily implemented in numerical simulations. So, why these simulations do not demonstrate stable results (in distinction to the one-lane case)? Our analysis shows that already in the deterministic setting ($p = q = 1$) the system is becoming highly non-ergodic. The latter means that two initial configurations with the same density of cars may lead to very different types of behavior. For example, we demonstrate the situations, when the motion of cars on some lanes is free, while all other lanes are congested. Moreover, the density of cars on different lanes eventually will be different from each other.

Another possible generalization is to take into account that normally the cars in the traffic flow are rather different: they might have different local velocities and might have some preferences during overtaking (when a car with larger velocity overtakes other cars, or the car with even lower velocity overtakes other cars staying in a jam). Examples are fire or police cars, ambulance cars, etc. We shall discuss some very naturally looking models of these processes based again on the cellular automaton representation. Unfortunately as we shall show these models are non-ergodic in the manner similar to the one discussed above.

To be precise let us define the simplest model of this sort (which we shall call a Multi-species TASEP) explicitly. Each site of the integer lattice $\mathbb{Z}$ is occupied by a single particle (representing a car) of one of $r+1$ types. The particle's type plays the role of its priority under dynamics and the 0-th type corresponds to the "holes" (non-occupied sites). Thus a configuration of particles is described by the sequence $x := \{x_i\}_{i\in\mathbb{Z}}$ from the alphabet $\{0, 1 \ldots, r\}$.

Our model is a discrete time version with parallel updates of the model introduced in [4], in which to each bond between the sites one associates a random Poisson "alarm clock" and when it rings for the i-th particle (the probability of simultaneous ringing of several alarm clocks is zero) the adjacent particles swap places if the one on the left is of larger type. The discrete time parallel update means that all the particles in the configuration are trying to move simultaneously which not only makes the analysis more difficult, but demands to adjust the construction of the local dynamics since several pairs of particles with intersecting members may compete for the swapping simultaneously.

To overcome this difficulty we define the dynamics through the map T acting on particle configurations, described by the relation

$$(Tx)_i := \begin{cases} x_{i-1} & \text{if } x_{i-1} > x_i,\ x_{i-2} \leqslant x_{i-1} \\ x_{i+1} & \text{if } x_{i-1} \leqslant x_i,\ x_i > x_{i+1} \\ x_i & \text{otherwise.} \end{cases}$$

In other words, a particle moves forward if and only if the type of the preceding particle is not higher and the type of the succeeding particle is lower comparing to the type of the particle under consideration. Thus the particles of the highest type move completely independently on others, when the moving of all other particles is subordinated to the particles of higher types.

Mathematical arguments in the analysis of the Nagel-Schreckenberg model is very different in the deterministic ($p = 1$) and pure random ($0 < p < 1$) cases. In the later case one unavoidably needs to find the stationary probability of the process under study, and all the further analysis is based on the properties of this stationary probability. An important feature of the Nagel-Schreckenberg model is that this stationary probability turns out to be Markovian in space. Roughly speaking if we fix a position of one car at time t, then the positions of cars to the left and to the right from this position are independent from each other. This feature helps a lot in the exact calculation of average velocities as functions of car densities. From the first sight it seems that this feature should hold for the two generalizations that we have considered above. Moreover, numerical simulations indicated that this should be the case. Unfortunately, for none of these models the Markovian structure of the stationary distributions was not proven.

References

1. Gray L., Griffeath D. The ergodic theory of traffic jams // J. Stat. Phys., 105:3/4 (2001), 413-452.
2. Nagel K., Schreckenberg M. A cellular automaton model for freeway traffic // J. Physique I, 2 (1992), 2221-2229.
3. Blank M. Stochastic stability of traffic maps // Nonlinearity, 25:12 (2012) 3389-3408.
4. Ferrari P., Martin J. Stationary distributions of multi-type totally asymmetric exclusion process // Annals of Probability, 35:3 (2009), 807-832

Decentralized fast gradient method for computation of Wasserstein barycenter

D.M. Dvinskikh
Moscow Institute of Physics and Technology,
Skolkovo Institute of Science and Technology, Moscow, Russia

Wasserstein distance allows to find the distance between probabilistic measures on a certain metric space and it is defined as a solution of transport optimization problem which is a task of linear programming. Instead of Wasserstein distance, regularized Wasserstein distance was being considered which is the solution of an entropy-regularized optimal transport problem.

$$\mathcal{W}_\gamma(p,q) \triangleq \min_{X \in U(p,q)} \left\{ \langle M, X \rangle + \gamma \sum_{i,j=1}^{n} X_{ij} \log X_{ij} \right\},$$

where $U(p,q) \triangleq \left\{ X \in \mathbb{R}_+^{n\times n} \mid X\mathbf{1} = p, X^T\mathbf{1} = q \right\}$.

A wide-range of modern problems are based on calculating the average of a set of objects. These problems arise in image processing, computer graphics, statistics and clusterization. For example, calculation of the mean of objects is a base of K-means algorithm, well-known in machine learning, for finding cluster centers. Also, optimal transport approach gives good results for the classification of images, for example, recognition of handwritten digits from the MNIST dataset; for comparing texts, applying Wasserstein distance to the distance between words; and for many other tasks. The mean of probability measures is the solution of the problem minimizing the sum of Wasserstein distances to each element in the set.

$$\min_{p \in S_1(n)} \sum_{i=1}^{m} \lambda_i \mathcal{W}_{\gamma,q_i}(p). \tag{1}$$

This mean is known as the Wasserstein barycenter of discrete probability measures. Notation $\mathcal{W}_{\gamma,q}(p)$ means Wasserstein distance of any point p to a fixed histogram q on a probability simplex $S_1(n)$.

Large amount of data and their large-dimensional feature space motivates us to seek for distributed methods for computation of Wasserstein barycenter. For instance, the dimension of the image is the number of its pixels and for large number of images with good resolution we are faced with a problem of processing a huge number of features. The idea of distributed optimization is to distribute the calculations between computational agents that form a connected network. Also distributed optimization is based on the fact that the computing capabilities of agents are higher than the speed of information sharing with neighboring nodes. To present the problem (1) in a distributed manner, we introduce the connected graph $\mathcal{G} = (V, E)$ represented the network of agents, Laplacian matrix W(communication matrix) defining this graph, stacked column vectors $\mathsf{p} = [p_1^T, \cdots, p_m^T]^T$ and $\mathsf{q} = [q_1^T, \cdots, q_m^T]^T$ for each agent $i \in V$, $p_i, q_i \in S_1(n)$. After some transformations the problem (1) can be rewritten in the equivalent form as follows

$$\min_{\mathsf{y}} \mathcal{W}^*_{\gamma,\mathsf{q}}(\sqrt{W}\mathsf{y}) = \sum_{i=1}^{m} \mathcal{W}^*_{\gamma,q_i}([\sqrt{W}\mathsf{y}]_i), \tag{2}$$

where $\mathcal{W}^*_{\gamma,q_i}([\sqrt{W}\mathsf{y}]_i) = \max_{p_i \in S_1(n)} \left\{ \left\langle [\sqrt{W}\mathsf{y}]_i, p_i \right\rangle - \mathcal{W}_{\gamma,q_i}(p_i) \right\}$ is conjugate function (Fenchel-Legendre transform).

Assuming that every agent in a graph $\mathcal{G}$ has its probabilistic distribution, we prove that the agents are able to reach the consensus (barycenter) interacting with neighboring nodes. Our method is based on dual approach and uses accelerated gradient descent from [1] executed in a distributed manner [2].

Algorithm

Input Each agent $i \in V$ is assigned its distribution q_i.

1. All agents set $\tilde{w}_0^i = \tilde{y}_0^i = \tilde{z}_0^i = \mathbf{0} \in \mathbb{R}^n$ and N

```
Set $K = \exp(-M/\gamma)$
For each agent $i \in V$: $k = 0, 1, 2, \cdots, N-1$
    $\tau_k = \frac{2}{k+2}$ and $\alpha_{k+1} = \frac{k+2}{2}\frac{1}{L}$
    $\tilde{y}^i_{k+1} = \tau_k \tilde{z}^i_k + (1-\tau_k)\tilde{w}^i_k$
    Share $y^i_{k+1}$ with $\{j \mid (i,j) \in E\}$
    Calculate $p^*_j(\tilde{y}^j_{k+1}) = \exp(\tilde{y}^j_{k+1}/\gamma) \circ \left(K \cdot \frac{q_i}{K \exp(\tilde{y}^j_{k+1}/\gamma)}\right)$ by agent i for $\{j \mid (i,j) \in E\}$
    $\tilde{w}^i_{k+1} = \tilde{y}^i_{k+1} - \frac{1}{L}\sum_{j=1}^m W_{ij} p^*_j(\tilde{y}^j_{k+1})$
    $\tilde{z}^i_{k+1} = \tilde{z}^i_k - \alpha_{k+1}\sum_{j=1}^m W_{ij} p^*_j(\tilde{y}^j_{k+1})$
end
Set $(y^*_N)_i = \tilde{w}^i_N$, $\forall i \in V$
Set $(p^*_N)_i = \sum_{k=0}^{N-1} \frac{(k+2)}{N(N+3)} p^*_i(\tilde{y}^i_{k+1})$, $\forall i \in V$
```

We demonstrate the optimal estimates of the complexity of a distributed algorithm for calculating the Wasserstein barycenter of discrete probabilistic measures.

Assuming that $\|\nabla \mathcal{W}^*_{\gamma,\mathtt{q}}(\tilde{\mathtt{y}})\|_2 \leq G$ on a ball $B_R(0)$ and set $\gamma = \varepsilon/(4m\log n)$, we prove [3] that after

$$N \geq \sqrt{\frac{128G^2 m \log n}{\varepsilon^2}\chi(W)}$$

iterations the outputs of Algorithm, i.e. $\mathtt{p}^*_N = [(p^*_N)^T_1, \cdots, (p^*_N)^T_m]^T$ and $\mathtt{y}^*_N = [(y^*_N)^T_1, \cdots, (y^*_N)^T_m]^T$ have the following properties

$$\mathcal{W}_{0,\mathtt{q}}(\mathtt{p}^*_N) - \mathcal{W}_{0,\mathtt{q}}(\mathtt{p}^*) \leq \varepsilon \quad \text{and} \quad \|\sqrt{W}\mathtt{p}^*_N\|_2 \leq \varepsilon/(2R).$$

Thus, distributing the calculation between multi-agent network system, the agents compute the barycenter of probabilistic distribution having only its distribution and exchanging the information with neighboring agents.

References

1. Allen-Zhu Z., Orecchia L. Linear Coupling: An Ultimate Unification of Gradient and Mirror Descent. // arXiv preprint arXiv:1407.1537, 2014.

2. Uribe C.A, Lee S., Gasnikov A., Nedic A. Optimal Algorithms for Distributed Optimization. // arXiv preprint arXiv:1712.00232, 2017.
3. Uribe C.A., Dvinskikh D., Dvurechensky P., Gasnikov A, Nedic A. Distributed Computation of Wasserstein Barycenters over Networks // arXiv preprint arXiv:1803.02933, 2018.

On the complexity of optimal transport problem

P. Dvurechensky
Weierstrass Institute for Applied Analysis and Stochastics, Berlin, Germany

Optimal transport (OT) distances between probability measures, including the Monge-Kantorovich or Wasserstein distance, play an increasing role in different machine learning tasks, such as unsupervised learning, semi-supervised learning, clustering, text classification, as long as in image retrieval, clustering and classification, statistics, and other applications. Our focus in this work is on the computational aspects of OT distances for the case of two discrete probability measures with support of equal* size n. The state-of-the-art approach [1] for this setting is to apply Sinkhorn's algorithm (also known as balancing or RAS algorithm) to the entropy-regularized OT optimization problem. As it was recently shown in [2], this approach allows to find an ε-approximation for an OT distance in $\widetilde{O}\left(\frac{n^2}{\varepsilon^3}\right)$ arithmetic operations. In terms of the dependence on n, this result improves on the complexity $\widetilde{O}(n^3)$ achieved by the network simplex method or interior point methods, applied directly to the OT optimization problem, which is a linear program. Nevertheless, the cubic dependence on ε prevents approximating OT distances with good accuracy.

Approximating the OT distance amounts to solving the *OT problem* proposed by L. Kantorovich:

$$\min_{X \in \mathcal{U}(r,c)} \langle C, X \rangle,$$

$$\mathcal{U}(r,c) := \{X \in \mathbb{R}_{+}^{n \times n} : X\mathbf{1} = r, \, X^T\mathbf{1} = c\}, \tag{1}$$

*This is done for simplicity and all the results easily generalize to the case of measures with different support size.

where X is *transportation plan*, $C \in \mathbb{R}_+^{n\times n}$ is a given ground cost matrix, $r, c \in \mathbb{R}^n$ are given vectors from the probability simplex Δ^n, $\mathbf{1}$ is the vector of all ones. The *regularized OT problem* is

$$\min_{X\in\mathcal{U}(r,c)} \langle C, X\rangle + \gamma\mathcal{R}(X), \tag{2}$$

where $\gamma > 0$ is the *regularization parameter* and $\mathcal{R}(X)$ is a strongly convex *regularizer*, e.g. negative entropy or squared Euclidean norm. Our goal is to find $\widehat{X} \in \mathcal{U}(r,c)$ such that

$$\langle C, \widehat{X}\rangle \le \min_{X\in\mathcal{U}(r,c)} \langle C, X\rangle + \varepsilon. \tag{3}$$

In our work we choose a different approach. We construct the dual problem to the problem of entropy-regularized optimal transport and solve it by accelerated gradient descent from [3]. This algorithm allows to reconstruct also the primal solution of the problem and has accelerated convergence rate both for the primal objective residual and constraints feasibility. Then we apply the rounding procedure from [2] to obtain a feasible point from the transport polytop. For our approach we prove the following theorem.

Theorem 1 *Our algorithm outputs $\widehat{X} \in \mathcal{U}(r,c)$ satisfying* (3) *in*

$$O\left(\min\left\{\frac{n^{9/4}\sqrt{R\|C\|_\infty \ln n}}{\varepsilon}, \frac{n^2 R\|C\|_\infty \ln n}{\varepsilon^2}\right\}\right) \tag{4}$$

arithmetic operations. Here R is the norm of the solution to the dual problem with minimum norm.

We also perform numerical experiments to compare our result with the result obtained by Sinkhorn's algorithm.

References

1. Cuturi, M. Sinkhorn distances: Lightspeed computation of optimal transport. In Burges, C. J. C., Bottou, L., Welling, M., Ghahramani, Z., and Weinberger, K. Q. (eds.), Advances in Neural Information Processing Systems 26, pp. 2292–2300. 2013.
2. Altschuler, J., Weed, J., and Rigollet, P. Near-linear time approxfimation algorithms for optimal transport via sinkhorn iteration. In Guyon, I., Luxburg, U. V., Bengio, S., Wallach, H., Fergus, R., Vishwanathan, S., and Garnett, R. (eds.), Advances in Neural Information Processing Systems 30, pp. 1961–1971. 2017.

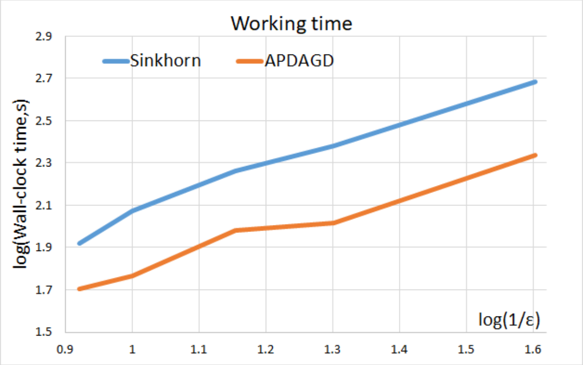


Figure 1: Comparison of working time of the Sinkhorn's algorithm and our algorithm APDAGD.

3. Dvurechensky, P., Gasnikov, A., Omelchenko, S., Tiurin, A. Adaptive Similar Triangles Method: a Stable Alternative to Sinkhorn's Algorithm for Regularized Optimal Transport, arXiv preprint 1706.07622

Formulation, algorithms for solution synthesis and computational complexity of minimax bi-assignment problem*

Yu.S. Fedosenko[1], D.I. Kogan[2], and D.A. Khandurin[2]
[1] *Volga State University of Water Transport, Nizhny Novgorod,*
[2] *Moscow Technological University, Moscow,*
Russia

1. The problem of the use of discrete resources arising in various applications is being investigated - the optimization of the distribution between agents of the pairs of non-mutually replaceable tasks. As an example of such an application, let us mention a logistic system of the Kamsky cargo area type, in which a dedicated group of multi-type cargo ships (multi-section ship convoys) is used to transport non-metallic construction materials (NCM) loaded in a single technological cycle with float-

*The Russian Foundation for Fundamental Research supported this research under grant 15-07-03141.

ing hydro-mechanized mining complexes (HMC) on the landfill sites. At the end of the session of the development of the next operational plan for the operation of the logistic system of the type in question, the dispatching service must unequivocally determine: a) to which HMC from the number of channeled deposits located at the landfill site, each individual ship of the selected group should be loaded for loading the NCM; b) what destination for discharge should be assigned to each particular vessel after its loading. For the development of operational plans for the functioning of a logistics system that are effective in the conditions of the developing operational environment, it is actual to develop and use a specialized digital management support system that includes both mathematical modeling procedures for the water transport logistics system and means for resolving suitably assigned extreme ship distribution tasks HMC for loading of the NCM and distribution of vessels at the unloading points of the NCM.

2. In discrete idealization, a mathematical model for the optimization of the distribution between agents of discrete resources of the type under consideration leads to the general bi-assignment problem with the minmax criterion formulated and studied below. There is set of agents $I = \{1, 2, \dots, n\}$ and two sets of tasks $\boldsymbol{P}= \{p_1, p_2, \dots, p_n\}$ and $\boldsymbol{Q}= \{q_1, q_2, \dots, q_n\}$. Each agent must be assigned to one of the tasks of the set $\boldsymbol{P}$ and to one of the tasks of the set $\boldsymbol{Q}$. Each of the tasks must be performed in full by exactly one agent. They are given by the given $(n \times n)$-matrices of the numerical estimates $\boldsymbol{A}= \{a_{ij}\}$ and $\boldsymbol{B}= \{b_{ij}\}$, where a_{ij} is the estimation of the performance of the task p_j by the agent i, and b_{ij} is the performance of the same agent q_j, $i = \overline{1, n}, j = \overline{1, n}$. We introduce the following notation: π_1 - the assignment of the agents to the tasks from the set $\boldsymbol{P}$, π_2 - the assignment of the agents to the tasks from the set $\boldsymbol{Q}$. Each assignment is a one-to-one mapping of the set $1, 2, \dots, n$ into itself. If $pi_1(i) = j$, then agent i should assign to task p_j. Similarly, the equality $\pi_1(i) = j$ means that agent i must also perform the task q_j.

A bi-assignment is called a pair of the form $< \pi_1(i), \pi_2(i) >$. It is considered that when implementing the bi-assignment $< \pi_1(i), \pi_2(i) >$, each agent i starts with the task number $\pi_1(i)$ starting from time 0, and immediately starts the task with the number $\pi_2(i)$.

In general form, the general problem of bi-assignment with a minmax criterion - problem 1 is written as follows:

$$\min_{\pi_1, \pi_2}(\max_{\alpha}[a_{\alpha\pi_1(\alpha)} + b_{\alpha\pi_2(\alpha)}]). \tag{1}$$

If the matrices $\boldsymbol{A}$ and $\boldsymbol{B}$ determine the duration of the task performed by the agents, then solving (1) we find a bi-assignment ensuring the minimum total length of the whole task package $\{p_1, p_2, \ldots, p_n, q_1, q_2, \ldots, q_n\}$.

It is easy to see that (1) is the generalization of the classical minmax assignment problem [1].

3. To construct the solving (1) algorithm, we use the concept of dynamic programming. Let i be a natural constant not exceeding n, and W_1, W_2 be arbitrary i-element subsets of $1, 2, \ldots, n$.

We denote by $Z(i, W_1, W_2)$ the subproblem of problem 1 in which the authors of the set $1, 2, \ldots, i$ should distribute the tasks with the lower indices (numbers) from the subsets W_1 and W_2; each agent must receive exactly one task of the set $\boldsymbol{P}$ (with the number entering the subset W_1) and exactly one task of the set $\boldsymbol{Q}$ (with the number entering the subset W_2). The set of tasks defined by subsets W_1, W_2 should be performed in the minimum time. The optimal value of the criterion in the problem $Z(i, W_1, W_2)$ is denoted by $B(i, W_1, W_2)$. It is obvious that $B(i, W_1, W_2)$ is the Bellman function for (1), and

$$B(1, \{j\}, \{k\}) = a_{1j} + b_{1k}, j, k \in \{1, 2, \ldots, n\}. \tag{2}$$

According to the Bellman principle, we have the following relation:

$$B(i, W_1, W_2) = \min(\max_{\alpha,\beta}[(a_{i\alpha} + b_{i\beta}), B(i-1, \{W_1 \backslash \alpha\}, \{W_2 \backslash \beta\})]). \tag{3}$$

where (α, β) are arbitrary pairs of indices from the set $W_1 \times W_2$.

Formulas (2), (3) are recurrence relations of dynamic programming for solving the problem (1).

The execution of the computational algorithm realizing these relations begins with the definition of the quantities $B(1, \{j\}, \{k\})$ for all singleton sets W_1 and W_2.

Then, in order of increasing i ($i = 2, 3, \ldots, n$), for all possible sets W_1 and W_2, the values of the function $B(i, W_1, W_2)$ are determined by formula (3); the value of $B(n, 1, 2, \ldots, n, 1, 2, \ldots, n)$ of the Bellman function with the extreme set of argument values is the optimal criterion value in problem 1.

In the process of performing the described computational procedure, for each triplet (i, W_1, W_2) of argument values, it is necessary to fix the pair (α, β) on which the minimum of the right-hand side of relation (3) is realized. This will allow, after finding the optimal value in task 1 of the criterion value, to write down the corresponding bi-assignment.

The complexity of the constructed decision algorithm (1) is determined by the number of calculated values of the Bellman function and, as is obvious, is determined by the quantity $O(4^n)$.

4. Let us introduce the application-specific laboriousness (1) problem 2, in which each of the available 2n tasks is characterized by its laboriousness: the task p_j has the laboriousness $t(p_j)$, the task q_j, has the laboriousness $t(q_j))$. We also assume that each agent i is characterized by its performance w_i, and the elements of the matrices $\boldsymbol{A}$ and $\boldsymbol{B}$ are calculated by the relations

$$a_{ij} = t(p_j)/w_i, b_{ij} = \{t(q_j)/w_i\}, i = \overline{1,n}, j = \overline{1,n}.$$

Under the conditions of problem 1 and problem 2, we select the corresponding recognition problems - problem 3 and problem 4, respectively.

In problem 3, with the initial data of problem 1 and the additionally indicated constant T, it is asked whether there is a bi-assignment, in the implementation of which the entire set of available tasks can be performed no later than the moment of time T (the directive deadline for completing the prescribed task package). In problem 4 an identical question is posed for the initial data of problem 2 and additionally indicated constant T. Obviously, the computational complexity of problem 1 is not lower than the computational complexity of problems 2 and 3, and the computational complexity of both problems and problems 3 is no less than the computational complexity of problem 4.

It is easy to show that problem 4 is polynomially equivalent to the NP-complete problem of "Combination with restrictions on weight" [2].

Thus, it is established that all the problems studied in this article are difficult to solve and, according to the natural scientific hypothesis "$P \neq NP$", algorithms of polynomial computational complexity for them can not be constructed.

5. Taking into account the applied significance of the discrete resources distribution problem under consideration, it is expedient to design an iterative solving algorithm for the general bi-assignment problem that realizes the concept of the branch and boundary method. To calculate the upper estimate of the value of the optimization criterion, the classical assignment problem with a minmax criterion, determined by the matrix A, is solved at the root of the variants of problem 1. The resulting assignment is denoted by π_1^*. Next, we construct the matrix $\boldsymbol{B*}$, each element of which is found by the formula $b_{ij}^* = b_{ij} + a_{i\pi^*(i)}$. The function

π_2^* is obtained analogously to the synthesis of π_1^* - as a result of solving the classical assignment problem defined by the matrix $\boldsymbol{B^*}$ with a minmax criterion.

The resulting bi-assignment $\pi^{**} =< \pi_1^*(i), \pi_2^*(i) >$ provides an upper bound in the root of the tree for the solution of the problem being solved. It is easy to see that, as the lower bound in the root of the variants tree, we can take the value of θ, calculated by the relation

$$\theta = \max_{\alpha}[\min_{\beta} a_{\alpha\beta} + \min_{\beta} b_{\alpha\beta}].$$

The above methods of obtaining the upper and lower estimates of the value of the optimized criterion in the root of the tree of variants induce in an obvious way the algorithms for finding them in subsequent intermediate vertices of this tree.

At the same time, the smallest of the upper bounds obtained in the process of the described construction of a variant of the variants tree sufficient for solving the problem is called the current record, and its value in general decreases in this process. The algorithm described above solves the problem as soon as a set of promising open vertices for future branching is empty.

The root of the variants tree is considered to be the vertex of the first rank; vertices of the k-th rank generate, on branching, a vertex of rank k+1.

The procedure of branching at an arbitrary vertex of the k-th rank consists in constructing from it branches, each of which corresponds to fixing for the k-th agent of some pair of free tasks from the set $\boldsymbol{P} \times \boldsymbol{Q}$; In this case, branches of the variants tree must be built only for those pairs (p_α, q_β) of free tasks for which the sum $a_{k\alpha} + b_{k\beta}$ is less than the value of the current record. At the same time, note that the number of branches of the tree of variants emerging from an arbitrary vertex of rank k can, in general, reach the value $(n - k) \times (n - k)$.

The final value of the current record is the optimal value of the criterion, and the path from the root to the top of the variants tree in which it is reached uniquely determines the solution of the bi-assignment problem.

In addition, we note that the presence of some pre-built even a small initial fragment of the variants tree can significantly shorten the duration of the solution of the problem by the method of dynamic programming. In fact, let the number U be the final (minimum) value of the current record when constructing such a fragment of the tree of variants. Then at each stage of calculations using formula (3) it is enough to consider only such pairs of indices (α, β) from $W_1 \times W_2$ for which the sum $a_{k\alpha} + b_{k\beta}$ does not exceed U.

The procedures for obtaining upper and lower bounds for the solution of problem 1 according to the scheme of BaB can be variously modified. We give the simplest example of such a modification.

References

1. Pentico D. Assignment problems: A golden anniversary survey // European Journal of Operational Research. 2007, N 176. P. 774–793.
2. Garey M.R. and Johnson D.S. Computers and Intractability: A Guide to the Theory of NP-Completeness, San Francisco: Freeman, 1979. Translated under the title Vychislitel'nye mashiny i trudnoreshaemye zadachi, Moscow: Mir, 1982.
3. Fedosenko Yu.S., Kogan D.I., Khandurin D.K. Non-standard types assignment problems: research and algorithms // Information Control and Technologies: VI International Scientific-Practical Conference. Odessa: ONMU, 2017. P. 300–302.

Dynamic network loading problem by means of link transmission approach

V.V. Kurtc and A.V. Prokhorov

A+S, Peter the Great St. Petersburg Polytechnic University, St. Petersburg, Russia

Introduction

Nowadays the problem of reliable forecast of traffic flows has a high degree of relevance. Realistic prediction of traffic flow propagation should be done both for urban roads and motorways out of the city. Moreover, it is necessary that the calculation be performed for large road networks in real or scalable time.These points lead to the dynamic traffic assignment problem [1–3] and to the need to use the Dynamic Network Loading models [4, 5]. These models allow to determine the link flows which correspond to given transport demand and route choices. The Iterative Link Transmission Model (ILTM) [6] provides realistic results according to first order kinematic wave theory [7] and allows traffic flow simulation in practical large scale networks in a reasonable time. In this paper we propose some modifications (improvements) of the ILTM. The contributions of this research are situated on three issues: the node model, calculation of turning portion matrix with respect to destination node and algorithm of distributing changes towards neighboring nodes.

The ILTM exploites the triangular shape of fundamental diagram with only two kinematic wave speeds (forward and backward). The ILTM procedure consists of three conseqtive steps going over all nodes: calculation of sending and receiving flows, running node model to find turning flows and updating cumulative value numbers at the link ends for all connected links. For more details the reader is referred to [6].

Modifications of the Iterative Link Transmission Mode

In this paper we suggest the node model as a solution of optimization problem that minimizes linear function, what corresponds to the optimal control within the intersection

$$\max_{TF_{ab}} \sum_{a \in BS_n} \sum_{b \in FS_n} TF_{ab}, \tag{1}$$

with linear constraints

$$\begin{gathered} TF_{ab} \geq 0, \forall a \in BS_n, \forall b \in FS_n, \\ v_a = \sum_{b \in FS_n} TF_{ab} \leq SF_a, \forall a \in BS_n, \\ u_b = \sum_{a \in BS_n} TF_{ab} \leq RF_b, \forall b \in FS_n, \\ TF_{ab} = \theta_{ab} \cdot v_a, \theta_{ab} = \frac{TP_{ab}}{\sum_{b \in FS_n} TP_{ab}}, \forall a \in BS_n, \forall b \in FS_n. \end{gathered} \tag{2}$$

Here TF_{ab} is a turning flow from link a to link b, BS_n and FS_n are incoming and outgoing links of the node n respectively, θ_{ab} is a turning fraction such that $\sum_{b \in FS_n} \theta_{ab} = 1, \forall a \in BS_n$. One can reformulate the problem (1), (2) in order to reduce the number of variables from $M * N$ to M (here $M = |BS_n|$, $N = |FS_n|$)

$$\begin{gathered} \max_{v_a} \sum_{a \in BS_n} v_a, \\ v_a \leq SF_a, \forall a \in BS_n, \\ \sum_{a \in BS_n} \theta_{ab} \cdot v_a \leq RF_b, \forall b \in FS_n, \end{gathered} \tag{3}$$

The problem (3) is a linear programming problem which can be solved by means of simplex algorithms. Afterwards the turning flows are calculated according to the following relation

$$TF_{ab} = \theta_{ab} \cdot v_a. \tag{4}$$

The second contribution introduced in this paper is the algorithm for calculation of turning flows in a node taking into account destination node. Let us consider node n and $TP_{ab}(i)$ is a number of vehicles transfering via node n from incoming link $a \in BS_n$ to outgoing link $b \in FS_n$ during $[t_i, t_{i+1}]$. Moreover the dynamic OD-matrix splitted according all routes is provided as one of the model input, that is $vol_p(i)$ is a number of vehicles departing on a route p during time period $[t_i, t_{i+1}]$. Firstly, let us define all routes going through the node n from link a to link b as P_{ab}. Next, considering P_{ab} one defines all routes P_{ab}^d, which end at the destination node d. Finally, turning portion from link a to link b towards destination d is as follows

$$TP_{ab}^d(i) = TP_{ab}(i) \frac{\sum_{p \in P_{ab}^d} vol_p(i)}{\sum_{p \in P_{ab}} vol_p(i)} \tag{5}$$

The third contribution of this research is the algorithm of distributing changes towards neighboring nodes. In [6] flows in a node are distributed according to sending flow values, but it is not taken into account that several routes with the same destination can use different outgoing links. Here we present another way for distributing changes towards neighbouring nodes considering above-mentioned issue. The formulas for computing upstream $U_{bd}(i)$ and downstream $V_{ad}(i)$ cumulative values are as follows

$$U_{bd}(i) = U_{bd}(i-1) + \sum_{a \in BS_n} TF_{ab} \frac{TP_{ab}^d(i)}{\sum_d TP_{ab}^d(i)}, \forall a \in BS_n, \forall d \in D \tag{6}$$

$$V_{ad}(i) = V_{ad}(i-1) + \sum_{b \in FS_n} TF_{ab} \frac{TP_{ab}^d(i)}{\sum_d TP_{ab}^d(i)}, \forall b \in FS_n, \forall d \in D \tag{7}$$

Results and Conclusion

We consider two networks to illustrate the performance of the ILTM with presented contributions. The first case study is from [5] (Fig. 1, left) whereas the second one is a practical large scale network – the part of Varshavskoe highway in Moscow which is more than three kilometers in length (Fig. 1, right). The second network includes 51 nodes, 106 links and 17 routes between different origin-destination pairs. The simulation time step $\Delta t = 10$ s and the total simulation time 150 min and 180 min respectively. Time-dependent OD-matrices spiltted according to all routes are imported every 180 seconds for the first and every 60 seconds for the second case study.

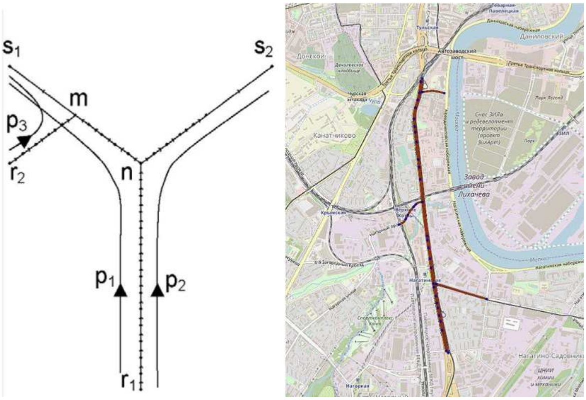


Fig. 1. Simple diverge netwok (left) and Varshavskoe highway network (right)

Figure 2 illustrates queue propagation according to ILTM for the simple diverge network. The height of the block represents the flow, whereas the color represents the speed. One can observe that merge node m acts as a bottleneck what leads to queue appearance on the link r_2m gradually growing and moving backwards.

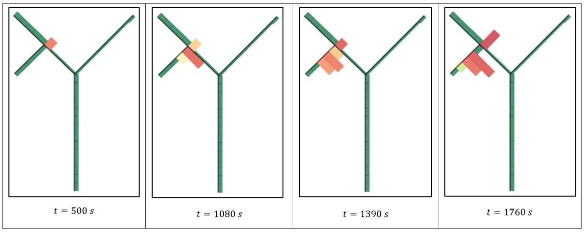


Fig. 2. Queue propagation over time. Simple diverge netwok

Simulation results for the second case study are presented in Figure 3. Initially network is empty. Later on densities on links increase and become stationary. At 1510 s flow from the north to the south disappears.

No queues and no bottlenecks are observed throughout all simulation period.

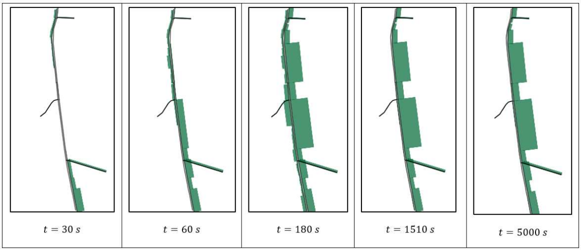


Fig. 3. Queue propagation over time. Varshavskoe highway, Moscow

The first case study demonstrates that the ILTM with presented in this paper contributions are able to reproduce basic phenomena of traffic flow such as bottlenecks activation, queue formation and its propagation backwards. The second case study shows the applicability of this model to practical networks. For further research, we plan to consider another practical and large scale networks and analyze the reliability of the simulation results. It will be done by comparison to real data or to results obtained with some commercial software.

References

1. Chiu Y.C. et. al. Dynamic Traffic Assignment. Transportation Network Modeling Committe. 2011. P. 1–39.
2. Szeto W., Wong S. Dynamic Traffic Assignment: model classifications and recent advances in travel choice principles // Open Engineering. 2012. Vol. 2, P. 1–18.
3. Viti F., Tampere C.M.J. Dynamic Traffic Assignment: Recent Advances and New Theories Towards Real Time Applications and Realistic Travel Behaviour. Edward Elgar, 2010.
4. Gentile G. The General Link Transmission Model for dynamic network loading and a comparison with the DUE algorithm // Proceedings of the Second International Symposium on Dynamic Traffic Assignment. Leuven, Belgium, 2008.
5. Yperman I. The Link Transmission Model for Dynamics Network Loading. June, 2007.
6. Himpe W. Integrated algorithms for repeated dynamic traffic assignments. The iterative link transmission model with equilibrium assignment procedure. March, 2016.

7. Jin W.L. Development and validation of kinematic wave traffic flow models for road networks // University of California Transportation Center Dissertation Grant, University of California. Davis, CA, 2002.

Fair energy flow redistribution after damage

Yu.E. Malashenko, I.A. Nazarova, and N.M. Novikova
Dorodnicyn Computing Centre, Federal Research Center "Computer Science and Control" of Russian Academy of Sciences, Moscow, Russia

A dynamic network flow model (MEF model) [1, 2] is constructed for the analysis of power supply processes, including the aggregated extraction, transportation, transformation, and redistribution of the main types of fuel and energy resources. Suppose that consumers and producers of energy resources are spatially distributed and integrated into a common energy complex (EC). In the MEF model, individual network micromodels correspond to each of the EC functional subsystems. Also, the common EC infrastructure is described by a set of network submodels. In general, the EC operation is reduced to the single-commodity flow optimization problem for a specially constructed multilayer graph whose layers correspond to different time intervals of EC analysis. Within the MEF model, energy supply options in spatially distributed systems after destructive impacts are studied. In this report, the strategies for controlling the flows are determined based on a posteriori information on the change in the capacity of the arcs of the model network. The control objective is the maximum of possible fulfillment of user requirements taking into account the regulatory constraints on the admissible levels of the accident/load on the network subsystems. This problem is considered as a multicriterial one [3].

Let $\mathcal{L}^+$ be the set of all MEF model's arcs (i,j):

$$i \in \mathcal{N} \cup \{\mathcal{U}\},\ \ j \in \mathcal{N} \cup \{\mathcal{S}\},\ \ i \neq j,\ \ l_{ij} \in \mathcal{L}, \text{ if }\ i = \mathcal{U}, \text{ then }\ j \neq \mathcal{S},$$

where $\mathcal{N}$ is the set of indices of model nodes, $\mathcal{U}$ is fictitious common source of infinite power, $\mathcal{S}$ is fictitious common sink, and $\mathcal{L}$ is the set of model arcs, except for those that connect the nodes to an fictitious source $\mathcal{U}$ and fictitious sink $\mathcal{S}$. The flow in the network is determined by the vector $\mathbf{x} = \langle x_{\mathcal{U}j}, ..., x_{ij}, ..., x_{i\mathcal{S}} \rangle$. Each component of $\mathbf{x}$ corresponds to the flow along the arc $l_{ij} \in \mathcal{L}^+$. Let us introduce the set $\mathbf{X}$ of flow vectors $\mathbf{x}$ that satisfy the conditions of flow conservation in transit nodes, constraints on the capacity of the corresponding arcs, and constraints on the given initial volumes of resources [1].The set $\mathbf{X}$ is called the set of feasible flows.

Let $\mathcal{K}$ be the set of indices of all consumer nodes of the MEF model; K be the total number of consumer nodes, $|\mathcal{K}| = K$; $v_m^k(t)$ be the sink vertex of the micromodel of the k-th consumer node for the m-th resource type, $k = \overline{1, K}$, $m \in \{1, ..., 4\}$ (for electricity, oil, gas, and coal respectively), the flow through which is considered on the interval $\tau_t \in [0, T]$; and $x_m^k(t)$ be the flow along one of the sink arcs of the micromodel of the kth consumer node into a single fictitious sink $v_{\mathcal{S}}$ equal to the amount of the mth resource fully utilized by the kth consumer for the time interval $\tau_t \in [0, T]$, $m \in \{1, ..., 4\}$, $k = \overline{1, K}$. Let us renumber all the sink vertices of all consumer nodes for all types of resources over all time intervals with natural numbers from 1 to $\overline{L}$, according to some rule and introduce the set $\overline{\mathcal{L}}$ of all sink arcs of the MEF model; i.e., let us establish a one-to-one correspondence

$$\overline{\mathcal{L}} = \{l_{i\mathcal{S}} \in \mathcal{L}^+ \mid (i, \mathcal{S}) = (v_m^k(t), v_{\mathcal{S}}),\ i = \overline{1, \overline{L}},\ k = \overline{1, K},\ m \in \{1, ..., 4\}\}.$$

It is assumed that each consumer node during the time interval τ_t has a specific request (requirement) for the supply of some quantity of the required types of energy resources. Let us define the kth consumer requirement for the mth resource type or the requirement at the vertex $v_m^k(t)$, for the time interval τ_t, $k = \overline{1, K}$, $m \in \{1, ..., 4\}$, by $f_m(k)(t)$, and let us proceed to the new notation: $f_{i\mathcal{S}}$ is the requirement on the sink arc $l_{i\mathcal{S}}$,

$$f_{i\mathcal{S}} = f_m(k)(t),\ l_{i\mathcal{S}} \in \overline{\mathcal{L}},\ i = \overline{1, \overline{L}},\ m \in \{1, ..., 4\},\ k = \overline{1, K}, \tau_t \in [0, T].$$

Vector $\mathbf{f} = \langle f_{1j}, ..., f_{ij}, ..., f_{\overline{L}j} \rangle$, $l_{ij} \in \overline{\mathcal{L}}$, $i = \overline{1, \overline{L}}$, $j = \mathcal{S}$, describes all the requirements for all types of energy resources for all consumer nodes for the entire considered period $[0, T]$ using sink arcs. Formally, assume the capacity d_{ij} of a sink arc $l_{ij} \in \overline{\mathcal{L}}$ to be equal to the request of the consumer node on this sink arc, i.e., $d_{ij} = f_{ij}$, $l_{ij} \in \overline{\mathcal{L}}$, $i = \overline{1, \overline{L}}$, $j = \mathcal{S}$. For the purposes of the present report, it is convenient to take the values obtained by consumers in the EC before the network damage (for example, according to the statistics, see [1]) as the model volume of the requirements. The measure for requirement satisfaction of the kth consumer node by the mth type of resource on the interval τ_t is equal to the flow running along the corresponding sink arc divided to the requirement at the sink vertex $v_m^k(t)$. Using the original indices, we denote this quantity by $\eta_m^k(t)$, i.e.,

$$\eta_m^k(t) = \frac{x_m^k(t)}{f_m(k)(t)}, \quad m \in \{1, ..., 4\}, \quad k = \overline{1, K}, \quad \tau_t \in [0, T],$$

or by proceeding to the notation given by the numbers of the sink arcs,

$$\eta_{ij} = \frac{x_{ij}}{f_{ij}} = \frac{x_{ij}}{d_{ij}}, \quad i = \overline{1, \overline{L}}, \quad j = \mathcal{S}.$$

The measure of the requirement satisfaction of the kth consumer node for the mth type of resource on the interval τ_t shows how much of the requested mth energy resource can actually be delivered to the kth consumer and is one of the most important characteristics of the EC after damage. Although in the initial network before the damage all $\eta_{ij} = 1$, and the network flow value corresponded to the sum of all requirements, under the conditions of large-scale damage, the flow value decreases. However, the measure of the requirement satisfaction of different users can change disproportionately. Therefore, in addition to the general problem of the minimum cost flow described in [1], let us consider in details the more particular problem: how to determine a flow that provides the most fair requirement satisfaction.

Suppose that the problem of controlling flows after damage is to reach the initial requirements of the users as much as possible. In this section, we a Assume that all consumers have equal rights and at any time none of them is given preference for any kind of energy resource. What is the guaranteed measure to fulfill all requirements? The answer to this question can be obtained by solving the problem of maximizing the minimal measure that fulfills the requirements given by the users.

Problem A. Let us find $\max\limits_{\mathbf{x} \in \mathbf{X}} \min\limits_{l_{ij} \in \overline{\mathcal{L}}} \eta_{ij}$

subject to

$$\eta_{ij} = \frac{x_{ij}}{d_{ij}}, \ l_{ij} \in \overline{\mathcal{L}}. \qquad (1)$$

The solution of problem A is equivalent to the solution of the following linear programming problem.

Problem B_1. Let us find $\max\limits_{\mathbf{x} \in \mathbf{X},\ \theta} \theta$

subject to

$$\theta \leqslant \frac{x_{ij}}{d_{ij}}, \ l_{ij} \in \overline{\mathcal{L}}. \qquad (2)$$

Let the optimal value of the parameter θ obtained as a result of solving problem A or B_1 subject to (1): be

$$\theta^* = \max_{\mathbf{x}\in\mathbf{X}} \min_{l_{ij}\in\overline{\mathcal{L}}} \eta_{ij} = \max_{\mathbf{x}\in\mathbf{X}} \min_{l_{ij}\in\overline{\mathcal{L}}} \frac{x_{ij}}{d_{ij}} = \max_{(\mathbf{x}\in\mathbf{X},\ \theta)\in(2)} \theta.$$

Taking into account constraints imposed on the flows (including those for sink arcs in which the capacity is equal to the requirements per flow) the value of θ cannot be greater than 1. If $\theta^* = 1$ then all user requirements can be fully satisfied. If $0 < \theta^* < 1$ then the requirements of some users at a time point can be fulfilled only at the level of $\theta^* \times 100\%$ of the requirements. If $\theta^* = 0$ then there is a consumer in the network who at least at one point in a time period cannot be provided with any kind of energy resource at all. It means that the damage in the energy resource transportation network cut off at least one consumer node from being supplied with at least one type of resource.

Suppose that while solving the original problem B_1 the best guaranteed measure to fulfill the requirements of consumers $\theta_1^* = \theta^*$ is found. In this case, the distribution of the flows along the sink arcs corresponding to the level of fulfillment θ_1^* is called the competitive distribution of flows. If $\theta_1^* < 1$, then there is a nonempty subset of sink arcs $\overline{\mathcal{L}}_1^*$, $\overline{\mathcal{L}}_1^* \subseteq \overline{\mathcal{L}}$, for which the measure of fulfillment of the requirements is exactly equal to θ_1^*, and it cannot be increased without violating the relation $\eta_{ij} \geqslant \theta_1^*$ for at least one sink arc $(i,j) \in \overline{\mathcal{L}}$:

$$\overline{\mathcal{L}}_1^* = \{l_{ij} \in \overline{\mathcal{L}} \mid \frac{x_{ij}^1}{d_{ij}} = \theta_1^* \ \ \forall\, x^1 :\ (x^1, \theta_1^*) \in \operatorname{Arg} \max_{(\mathbf{x}\in\mathbf{X},\ \theta)\in(2)} \theta\}.$$

Let us call the arcs in the set $\overline{\mathcal{L}}_1^*$ by the arcs of the first level of fulfillment of the requirements. The set $\overline{\mathcal{L}}_1^*$ can be constructed, for example, using the problem B_1 by the standard network flow programming methods. If $\overline{\mathcal{L}}_1^*$ coincides with $\overline{\mathcal{L}}$ then the problem of maximizing the minimum measure of fulfilling requirements for all sink arcs is solved. Otherwise, we will continue the maximization in order to find more fair competitive distribution of flows.

Just fix the flows along the arcs of the set $\overline{\mathcal{L}}_1^*$ with the achieved level θ_1^* and solve the following problem of maximizing the minimum measure of requirement satisfaction for the sink arcs from the set $\overline{\mathcal{L}}_1 = \overline{\mathcal{L}} \backslash \overline{\mathcal{L}}_1^*$.

Problem B_2. Let us find $\max\limits_{\mathbf{x}\in\mathbf{X},\ \theta} \theta$

subject to

$$\begin{cases} \theta \leqslant \dfrac{x_{ij}}{d_{ij}}, \;\; l_{ij} \in \overline{\mathcal{L}}_1, \\ \dfrac{x_{ij}}{d_{ij}} = \theta_1^*, \; l_{ij} \in \overline{\mathcal{L}}_1^*. \end{cases} \qquad (3)$$

The process of constructing Pareto-optimal flow vector will stop at the Sth step, as soon as it turns out that for all the sink arcs, whose flows are not yet fixed, the guaranteed measure of their requirement fulfillment is exactly equal to the optimal value of the parameter θ_S^*. The process will also be completed in the case when the guaranteed measure of fulfillment of the requirements of the considered sink arcs becomes unity.

Thus, we solve a finite lexicographic sequence of problems of type A (or B_1) and construct the flow that fairly meets the requirements. This means that it is impossible to increase the measure of the requirement satisfaction for any sink arc without simultaneously reducing the satisfaction requirement measure for another sink arc which has the same or worse conditions. Suppose that the lexicographic maximin problem is solved and $\theta_S^* < 1$. Thus it is impossible to fulfill the requirements of any consumer for either type of energy resources. Note that the process of solving the sequence of maximin problems can be interrupted at any step based on substantive considerations. In this case, $\theta_1^* < ... < \theta_s^* < 1$.

In order to identify network arcs, the insufficient capacity of which hinders increasing the target function, it is useful to carry out a postoptimal analysis of the obtained solution. In particular, if the flow along the arc is less than its capacity, then the latter is not required. In the case of equality, it is necessary to determine the values of the dual variables for the corresponding flow constraints along the arc. The large values of the dual variables indicate that the increase in the target function will be significant when this constraint is mitigated.

Within the MEF model, it is convenient to study the possibilities of supplying the network by one resource or by their group, estimate the measure of fulfillment of the requirements for one user or for a group of users. These features modify the type of functionals and constraints of problems but can be described within the propoused methodology for using the MEF model.

References

1. Kozlov M.V., Malashenko Yu.E., Nazarova I.A, et al. Fuel and energy system control at large-scale faults. 1. Network model and

software implementation // J. Comput. Syst. Sci. Int. 2017 V. 56, No 6. P. 945–968.

2. Malashenko Yu.E., Nazarova I.A, and Novikova N.M. Fuel and energy system control at large-scale faults. 2. Optimization problems // J. Comput. Syst. Sci. Int. 2018 V. 57, No 2. P. 208–221.
3. Germeier Yu.B. Introduction to Operations Research Theory. Moscow: Nauka, 1971. [in Russian].

Travel demand forecasting model based on dynamic traffic assignment

D.S. Mazurin

Institute for Systems Analysis, Federal Research Centre “Computer Science and Control” of Russian Academy of Sciences, Moscow, Russia

We consider the problem of travel demand forecasting in a large city. The common approach to solve this problem is the four-step model [1], which involves (i) trip generation, (ii) trip distribution, (iii) modal split and (iv) static traffic assignment. Static assignment models rely on simple link travel time functions and can’t reproduce observed network conditions, which raises questions about the accuracy of the forecasts. This work presents an alternative modelling framework for long-term transportation planning, transport policy evaluations and scenarios comparison based on dynamic traffic assignment (DTA) [2]. The general scheme is shown in Figure 1.

In our framework we take into account trip chaining behaviour: trips combine into tours starting and ending at the same place, usually at home [3]. In this paper we limit ourselves to the simplest two-leg tours *Home → Object → Home* with a single destination. The trip generation aims at predicting the total daily number of trips generated at and attracted to each zone for each demand stratum. The trip distribution model describes the activity location choice for each demand stratum. We use here the combined gravity model with two person groups with and without car availability:

$$Q_{ij}^k = a_i^k O_i^k b_j D_j f(c_{ij}^k),$$

$$\sum_j Q_{ij}^k = O_i^k, \quad \sum_i \sum_k Q_{ij}^k = D_j.$$

Here Q_{ij}^k is the total daily number of tours from i to j and back for person group k, O_i^k is the total daily generation of zone i for person

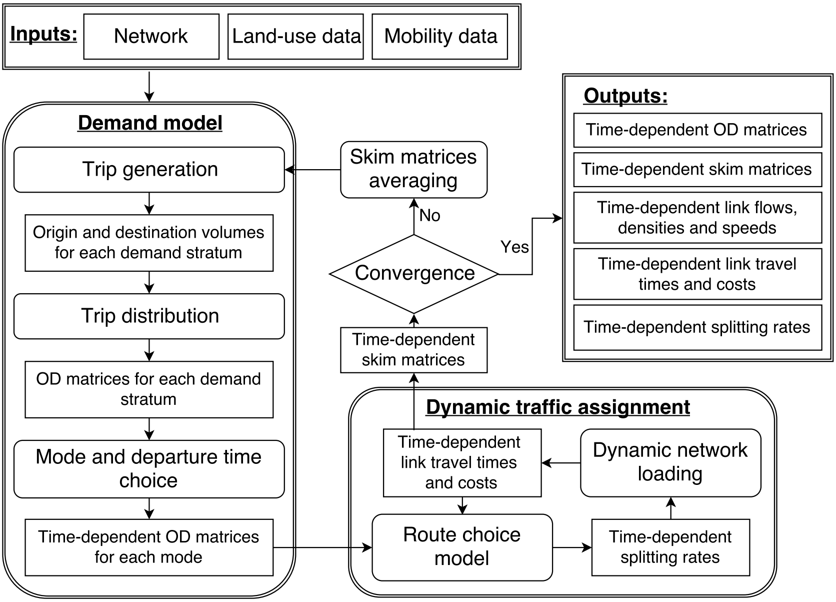

Inputs:
Network
Land-use data
Mobility data
Demand model
Trip generation
Origin and destination volumes for each demand stratum
Trip distribution
OD matrices for each demand stratum
Mode and departure time choice
Time-dependent OD matrices for each mode
Skim matrices averaging
No
Convergence
Yes
Time-dependent skim matrices
Outputs:
Time-dependent OD matrices
Time-dependent skim matrices
Time-dependent link flows, densities and speeds
Time-dependent link travel times and costs
Time-dependent splitting rates
Dynamic traffic assignment
Time-dependent link travel times and costs
Dynamic network loading
Route choice model
Time-dependent splitting rates

Fig. 1. The general scheme of the proposed modelling framework.

group k, D_j is the total daily attraction of zone j, c_{ij}^k is the aggregated generalized travel cost of choosing zone j from zone i for person group k, $f(c)$ is decreasing gravity function, a_i^k and b_j are balancing factors. The generalized travel cost usually includes travel time, monetary costs and some other components.

The next block, mode and departure time choice model, describes two more travel behaviour dimensions. It takes as input OD matrices for each demand stratum and splits them by mode (private car, public transport, etc.) and by departure time choice intervals (with time step $\sim$ 15 min) on the basis of discrete choice theory [4]. We consider two types of tours with respect to departure time choice: mandatory tours with fixed schedule (most of commuters, school trips, etc.) and discretionary tours with flexible schedule (shopping, leisure, some of commuters, etc.).

For the first tour type (e.g., *Home* → *Work* → *Home*) we assume preferred arrival (to work) time window $[t_1^a; t_2^a]$ and preferred departure (from work) time window $[t_1^d; t_2^d]$ given. For any deviation from preferred arrival/departure time so-called schedule delay penalty is included in generalized travel costs. Then the utility of the whole tour from i to j at time interval t_1 and back at time interval t_2 by mode m is the following:

$$\begin{gathered}U^m_{ijt_1t_2} = \alpha c^m_{ijt_1} + \beta \max(t^a_1 - \tau^m_{ijt_1} - t_1, 0) + \gamma \max(t_1 + \tau^m_{ijt_1} - t^a_2, 0) + \\ + \alpha c^m_{ijt_2} + \beta \max(t_2 - t^d_2, 0) + \gamma \max(t^d_1 - t_2, 0).\end{gathered}$$

Here τ^m_{ijt} and c^m_{ijt} are travel time and travel cost from i to j at time t accordingly, α, β and γ are travel behaviour parameters.

For the second tour type (e.g., *Home* → *Shopping* → *Home*) we assume preferred arrival (to shop) time window $[t^a_1; t^a_2]$ and activity duration Δ given. Then the utility of the whole tour is the following:

$$\begin{gathered}U^m_{ijt_1t_2} = \alpha c^m_{ijt_1} + \beta \max(t^a_1 - \tau^m_{ijt_1} - t_1, 0) + \gamma \max(t_1 + \tau^m_{ijt_1} - t^a_2, 0) + \\ + \alpha c^m_{ijt_2} + \beta \max(t_2 - t_1 - \tau^m_{ijt_1} - \Delta, 0) + \gamma \max(t_1 + \tau^m_{ijt_1} + \Delta - t_2, 0).\end{gathered}$$

Given utility values for each triple (t_1, t_2, m) we can calculate choice probabilities $s^{km}_{ijt_1t_2}$ and aggregated utility U^k_{ij} for the people group k. For example, Multinomial Logit model leads to:

$$s^{km}_{ijt_1t_2} = \frac{\exp(U^m_{ijt_1t_2})}{\sum_{(t_1,t_2)} \sum_{m \in M_k} \exp(U^m_{ijt_1t_2})},$$

$$U^k_{ij} = \ln\left(\sum_{(t_1,t_2)} \sum_{m \in M_k} \exp\left(U^m_{ijt_1t_2}\right) \right)$$

where M_k is the set of modes available for the people group k.

The dynamic traffic assignment model consists of two fundamental components: a route choice model and a dynamic network loading (DNL) model. The DNL model describes traffic flow propagation on road networks and reproduces traffic flow characteristics such as density, flow and average speed. It takes as input time-dependent path flows and outputs time-dependent link flows, densities and travel times. In this paper we use a macroscopic first-order traffic flow model, Link Transmission Model (LTM) [5]. This model captures spatial and temporal congestion dynamics (queue build-up, spillback, and dissipation) in terms of cumulative inflows and outflows and requires a minimum number of input data (capacity, free flow speed and jam density for each link).

The DNL model consists of a link model and a node model. The link model describes the traffic flow dynamics on homogeneous road sections without intersections (given the boundary conditions) and can be considered as a mapping $\mathbf{\Lambda}$ that determines the maximum possible inflows

$R_a(t)$ and outflows $S_a(t)$ for link a at time t given cumulative inflows $U_a(t')$ and outflows $V_a(t')$ at time $t' < t$:

$$(S_a(t), R_a(t)) = \mathbf{\Lambda}\left(U_a(t'), V_a(t')\right),\ t' < t.$$

The node model provides the connection between links and captures flow interactions at junctions [6]. For each node n it can be seen as a mapping $\mathbf{\Upsilon}$ from the maximum outflows $S_a(t)$ of incoming links $a \in BS(n)$ and the maximum inflows $R_b(t)$ of outgoing links $b \in FS(n)$ at time t to actual instantaneous outflows $v_a(t)$ and inflows $u_b(t)$ for these links at time t:

$$(v_a(t), u_b(t)) = \mathbf{\Upsilon}(S_a(t), R_b(t), \phi_{ab}(t)), a \in BS(n), b \in FS(n).$$

Here $\phi_{ab}(t)$ are the splitting rates determined by the route choice model (conditional probabilities of choosing outgoing link b for vehicles arrived to node n at time t from link a). In our model we use multi-commodity DNL procedure, keeping track of individual flow components towards different destinations, so all loading and routing variables are disaggregated by destination d. Given cumulative inflow and outflow temporal profiles $U_a(t)$ and $V_a(t)$ we can easily calculate link travel times $\tau_a(t)$ and traffic flow variables (flow, density and speed) at any intermediate point.

The route choice model describes drivers' route choice behaviour. It takes as input time-dependent link travel times and generalized travel costs and outputs time-dependent path flows. Each driver is assumed to follow the dynamic user equilibrium (DUE) route choice principle, which is the simplest dynamic extension of Wardrop's first principle [2]. It states that for each origin-destination pair, any routes used by travelers departing at the same time must have equal and minimal travel cost. In this work we use implicit path enumeration via splitting rates at nodes. To solve the DUE problem we use heuristic gradient projection algorithms [7].

We applied our approach to the transport system of Nizhny Novgorod, the fifth largest city in Russia with a population of over 1.2 million. The road network graph consists of 1901 nodes, 5478 links and 264 zones. Our computational experiments show that satisfactory convergence level can be achieved in a reasonable number of iterations for moderate congestion. In contrast to static traffic assignment, dynamic traffic assignment models with spillback are extremely sensitive to the network parameters. The biggest problem is that gridlock situations may occur where the vehicles on the closed loop (e.g., roundabout) block each other, therefore convergence progress should be monitored during the calculation or some gridlock prevention algorithms should be used here. The computation time and the memory consumption for this medium-size network are in acceptable limits even on an ordinary modern computer. All computational procedures are designed to take full advantage of multiprocessor systems. In our future work, we aim at further increasing the overall computational performance of our modelling framework, which will make it possible to deal with more complex transport networks such as the Moscow agglomeration [8].

References

1. Ortuzar J. de D., Willumsen L.G. Modelling Transport. Oxford: Wiley-Blackwell, 2011.
2. Szeto W.Y., Wong S.C. Dynamic traffic assignment: model classifications and recent advances in travel choice principles // Cent. Eur. J. Eng. 2012. V. 2, N 1. P. 1–18.
3. Aliev A.S., Mazurin D.S., Maksimova D.A, Shvetsov V.I. Model of traffic flows based on the 4-step scheme with the chains of trips // Proc. of ISA RAS. 2016. V. 66, N 1. P. 3–9 (in russian).
4. Ben-Akiva M., Bierlaire M. Discrete choice models with applications to departure time and route choice. // Handbook of Transportation Science, 2nd edition. Kluwer, 2003. P. 7–38.
5. Yperman, I., Logghe, S., Immers, B. The Link Transmission Model: An Efficient Implementation of the Kinematic Wave Theory in Traffic Networks. // Proceedings of the 10th EWGT Meeting and 16th Mini-EURO Conference, Poznan, Poland. 2005. P. 122–127.
6. Tampére C.M., Corthout R., Cattrysse D., Immers L.H. A generic class of first order node models for dynamic macroscopic simulation of traffic flows // Transp. Res. B. 2011. V. 45, N 1. P. 289–309.
7. Gentile G. Solving a dynamic user equilibrium model based on splitting rates with gradient projection algorithms // Transp. Res. B. 2016. V. 92. P. 120–147.
8. Aliev A.S., Mazurin D.S., Maksimova D.A, Shvetsov V.I. The structure of the complex model of the transport system of Moscow // Proc. of ISA RAS. 2015. V. 65, N 1. P. 3–15 (in russian).

Some Adaptive Mirror Descent Algorithms for a Special Class Convex Constrained Optimization Problems*

Fedor S. Stonyakin, Alexander A. Titov and Mohammad S. Alkousa
V. I. Vernadsky Crimean Federal University, Moscow Institute of Physics and Technology, Simferopol, Moscow, Russian Federation

The report is devoted to a special Mirror Descent algorithm for problems of convex minimization with functional constraints (see [1], Section 3.3). The objective function may not satisfy the Lipschitz condition, but it must necessarily have the Lipshitz-continuous gradient. We assume, that the functional constraint can be non-smooth, but satisfying the Lipschitz condition. In particular, such functionals appear in the well-known Truss Topology Design problem [3, 5]. Also we have applied the technique of restarts in the mentioned version of Mirror Descent for strongly convex problems [2, 4]. Some estimations for a rate of convergence are investigated for considered Mirror Descent algorithms.

Let $(E, ||\cdot||)$ be a normed vector space and E^* be the conjugate space of E with the norm:

$$||y||_* = \max_x \{\langle y, x\rangle, ||x|| \leq 1\},$$

where $\langle y, x\rangle$ is the value of the continuous linear functional y at $x \in E$.

Let $X \subset E$ be a (simple) closed convex set. We consider two convex subdiffirentiable functions f and $g : X \to \mathbb{R}$. Also we assume that g is Lipschitz-continuous:

$$|g(x) - g(y)| \leq M_g ||x - y|| \; \forall x, y \in X. \tag{1}$$

We focus on the next type of convex optimization problems.

$$f(x) \to \min_{x \in \mathcal{X}}, \tag{2}$$

$$\text{s.t.} \quad g(x) \leq 0. \tag{3}$$

Let $d : X \to \mathbb{R}$ be a distance generating function (d.g.f) which is continuously differentiable and 1-strongly convex w.r.t. the norm $\|\cdot\|$, i.e.

$$\forall x, y \in X \;\; \langle \nabla d(x) - \nabla d(y), x - y\rangle \geq \|x - y\|^2,$$

*The reported study was supported by Russian Foundation for Basic Research according to the research project 18-31-00219.

and assume that $\min\limits_{x \in X} d(x) = d(0)$. Suppose, we have a constant Θ_0 such that $d(x_*) \leq \Theta_0^2$, where x_* is a solution of (2)–(3).

Note, that if there is a set of optimal points X_*, then we may assume, that

$$\min_{x_* \in X_*} d(x_*) \leq \Theta_0^2.$$

For all $x, y \in X$ consider the corresponding Bregman divergence

$$V(x, y) = d(y) - d(x) - \langle \nabla d(x), y - x \rangle.$$

Standard proximal setups, i.e. Euclidean, entropy, ℓ_1/ℓ_2, simplex, nuclear norm, spectahedron can be found, e.g. in [2]. Let us define the proximal mapping operator standardly:

$$\mathrm{Mirr}_x(p) = \arg\min_{u \in X} \big\{ \langle p, u \rangle + V(x, u) \big\} \text{ for each } x \in X \text{ and } p \in E^*.$$

We make the simplicity assumption, which means that $\mathrm{Mirr}_x(p)$ is easily computable.

Following [4], given a function f for each subgradient $\nabla f(x)$ at a point $y \in X$, we define

$$v_f(x, y) = \begin{cases} \left\langle \dfrac{\nabla f(x)}{\|\nabla f(x)\|_*}, x - y \right\rangle, & \nabla f(x) \neq 0 \\ 0 & \nabla f(x) = 0 \end{cases}, \quad x \in X. \tag{4}$$

The following adaptive Mirror Descent algorithm for Problem (2) – (3) was proposed by the first author.

Algorithm 1. Adaptive Mirror Descent, non-standard growth

```
Require: $\varepsilon, \Theta_0^2, X, d(\cdot)$
$x^0 = \underset{x \in X}{argmin}\, d(x)$
$I =: \emptyset$
$N \leftarrow 0$
repeat
    if$g(x^N) \leq \varepsilon \to$then
        $h_N \leftarrow \frac{\varepsilon}{||\nabla f(x^N)||_*}$
        $x^{N+1} \leftarrow Mirr_{x^N}(h_N \nabla f(x^N))$ (”productive steps”)
        $N \to I$
```

```
  else
      $(g(x^N) > \varepsilon) \to$
      $h_N \leftarrow \frac{\varepsilon}{||\nabla g(x^N)||_*^2}$
      $x^{N+1} \leftarrow Mirr_{x^N}(h_k \nabla g(x^N))$ ("non-productive steps")
  end if
  $N \leftarrow N + 1$
until $\Theta_0^2 \leqslant \frac{\varepsilon^2}{2}\left(|I| + \sum\limits_{k \notin I} \frac{1}{||\nabla g(x^k)||_*^2}\right)$
Ensure: $\bar{x}^N := \arg\min_{x^k, k \in I} f(x^k)$
```

Theorem 1. Let $\varepsilon > 0$ be a fixed positive number and Algorithm 1 work

$$N = \left\lceil \frac{2\max\{1, M_g^2\}\Theta_0^2}{\varepsilon^2} \right\rceil \tag{5}$$

steps. Then

$$\min_{k \in I} v_f(x^k, x_*) < \varepsilon. \tag{6}$$

We can apply Algorithm 1 to some class of problems with a special class of non-smooth objective functionals.

Corollary 1. Assume that $f(x) = \max\limits_{i=\overline{1,m}} f_i(x)$, where f_i is differentiable at each $x \in X$ and

$$||\nabla f_i(x) - \nabla f_i(y)||_* \leq L_i ||x - y|| \quad \forall x, y \in X.$$

Then after

$$N = \left\lceil \frac{2\max\{1, M_g^2\}\Theta_0^2}{\varepsilon^2} \right\rceil$$

steps of Algorithm 1 working the next estimate can be fulfilled:

$$\min_{0 \leq k \leq N} f(x^k) - f(x_*) \leq \varepsilon_f + \frac{L}{2} \cdot \varepsilon^2,$$

where

$$\varepsilon_f = \varepsilon \cdot \max_{i=\overline{1,m}} ||\nabla f_i(x_*)||_*, \quad L = \max_{i=\overline{1,m}} L_i.$$

Let us consider the following problem

$$f(x) \to \min, \;\; g(x) \leq 0, \;\; x \in X \tag{7}$$

with assumption (1) and additional assumption of strong convexity of f and g with the same parameter μ, i.e.,

$$f(y) \geq f(x) + \langle \nabla f(x), y - x \rangle + \frac{\mu}{2}\|y - x\|^2, \quad x, y \in X$$

and the same holds for g. We also slightly modify assumptions on prox-function $d(x)$. Namely, we assume, that $0 = \arg\min_{x \in X} d(x)$ and that d is bounded on the unit ball in the chosen norm $\|\cdot\|_E$, that is

$$d(x) \leq \frac{\Omega}{2}, \quad \forall x \in X : \|x\| \leq 1, \tag{8}$$

where Ω is some known number. Finally, we assume that we are given a starting point $x_0 \in X$ and a number $R_0 > 0$ such that $\|x_0 - x_*\|^2 \leq R_0^2$.

Algorithm 2. Algorithm for the strongly convex problem

```
Require: accuracy $\varepsilon > 0$;
strong convexity parameter $\mu$; $\Theta_0^2$ s.t. $d(x) \leq \Theta_0^2 \quad \forall x \in X : \|x\| \leq 1$;
starting point $x_0$ and number $R_0$ s.t. $\|x_0 - x_*\|^2 \leq R_0^2$
1: Set $d_0(x) = d\left(\frac{x - x_0}{R_0}\right)$
2: Set $p = 1$
3: repeat
4:     Set $R_p^2 = R_0^2 \cdot 2^{-p}$
5:     Set $\varepsilon_p = \frac{\mu R_p^2}{2}$
6:     Set $x_p$ as the output of partial adaptive version of Algorithm 1
       with accuracy $\varepsilon_p$, prox-function $d_{p-1}(\cdot)$ and $\Theta_0^2$
7:     $d_p(x) \leftarrow d\left(\frac{x - x_p}{R_p}\right)$
8:     Set $p = p + 1$
9: until $p > \log_2 \frac{\mu R_0^2}{2\varepsilon}$
Ensure: $x_p$
```

Consider the function $\tau : \mathbb{R}^+ \to \mathbb{R}^+$:

$$\tau(\delta) = \max\left\{\delta\|\nabla f(x_*)\|_* + \frac{\delta^2 L}{2};\ \delta\right\}.$$

It is clear that τ increases and therefore for each $\varepsilon > 0$ there exists

$$\hat{\varphi}(\varepsilon) > 0 : \ \tau(\hat{\varphi}(\varepsilon)) = \varepsilon.$$

Theorem 2. Assume that $f(x) = \max\limits_{i=\overline{1,m}} f_i(x)$, where f_i is differentiable at each $x \in X$ and

$$||\nabla f_i(x) - \nabla f_i(y)||_* \leq L_i||x - y|| \quad \forall x, y \in X. \tag{9}$$

Let f and g satisfy (9). If f, g are μ-strongly convex functionals on $X \subset \mathbb{R}^n$ and $d(x) \leq \theta_0^2 \quad \forall x \in X, \quad \|x\| \leq 1$. Let the starting point $x_0 \in X$ and the number $R_0 > 0$ be given and $\|x_0 - x_*\|^2 \leq R_0^2$. Then for $\widehat{p} = \left\lceil \log_2 \dfrac{\mu R_0^2}{2\varepsilon} \right\rceil$ $x_{\widehat{p}}$ is the ε-solution of Problem (2) – (3) (i.e. $f(x_{\widehat{p}}) - f(x_*) < \varepsilon$ and $g(x_{\widehat{p}}) < \varepsilon$), where

$$\|x_{\widehat{p}} - x_*\|^2 \leq \frac{2\varepsilon}{\mu}.$$

At the same time, the total number of iterations of partial adaptive version of Algorithm 1 does not exceed

$$\widehat{p} + \sum_{p=1}^{\widehat{p}} \frac{2\theta_0^2 \max\{1, M_g^2\}}{\varphi^2(\varepsilon_p)}, \quad \text{where } \varepsilon_p = \frac{\mu R_0^2}{2^{p+1}}.$$

Acknowledgments. The authors are very grateful to Yurii Nesterov, Alexander Gasnikov and Pavel Dvurechensky for fruitful discussions.

References

1. A. Bayandina, P. Dvurechensky, A. Gasnikov, F. Stonyakin, A. Titov (2017). *Mirror Descent and Convex Optimization Problems With Non-Smooth Inequality Constraints. In LCCC Focus Period on Large-Scale and Distributed Optimization, June 14-16, 2017. Lund, Sweden: Lund Center for Control of Complex Engineering Systems, Lund University* [Online]. Available: `https://arxiv.org/pdf/1710.06612.pdf`.
2. A. Ben-Tal and A. Nemirovski, *Lectures on Modern Convex Optimization.* Philadelphia: SIAM, 2001.
3. A. Ben-Tal and A. Nemirovski, "Robust Truss Topology Design via Semidefinite Programming", in *SIAM J. Optim.*, vol. 7, no. 4, pp. 991–1016, Nov., 1997.
4. Y. Nesterov. Introductory Lectures on Convex Optimization: a basic course. Kluwer Academic Publishers, Massachusetts, 2004.

5. Y. Nesterov. *Subgradient methods for convex functions with non-standard growth properties*, 2016. `http://www.mathnet.ru:8080/PresentFiles/16179/growthbmnesterov.pdf`.

Neural network approach to remodel stochastic freeway capacity[*]

A. Sysoev and J. Geistefeldt
Lipetsk State Technical University,
Ruhr University Bochum, Lipetsk, Russia, Bochum, Germany

Introduction. To study and then to predict the capacity of freeway section, it is necessary firstly to construct the mathematical model, that allows to estimate the dynamics of the system under consideration. There are many approaches to determine this parameter. But the main definition considers deterministic concept and says that capacity is the maximum number of vehicles, that a particular section of the transportation system is able to serve during specified time interval. This idea is a leading one in HCM 2010 (Highway Capacity Manual, USA), HBS 2015 (German Highway Capacity Manual, Germany) and ODM 218.2.020–2012 (Guidelines to estimate road capacity, Russia).

But it was proven, that the capacity rate of a freeway segment has a stochastic nature. The first paper [1] following this concept introduces an analogy of lifetime analysis, which is called the Product Limit method. The non-parametric estimation of a fit function in this method indicates the probability of exceeding the capacity of a freeway segment. Late it was found out [2], that Weibull distribution function fit the probability of breakdown the best. The ranges for shape and scale parameters of the distribution law were also investigated. These values correlate to the other research [3]. So, the stochastic capacity could be estimated using the formula

$$F(q) = 1 - \exp\left(-\left(\frac{q}{b}\right)^{a}\right), \tag{1}$$

here $F(q)$ is a capacity distribution function, q is a flow rate, a and b are shape and scale parameters of Weibull law respectively.

Capacity has a very special nature and could not be measured directly, except certain time intervals when traffic flow transits from fluid to congested regime. And, of course, the capacity during thaffic jam could

[*] The reported study is supported by the Russian Science Foundation within the project 18-71-10034.

be assumed as the number of vehicles which are within the freeway segment, i.e. traffic flow rate. The idea proposed in [1,2] supposes, that observations during both fluid and congested traffic are included in the capacity analysis. The capacity during congestion differs from the capacity before a breakdown. Thus, it is more appropriate to include only those values measured during fluid traffic conditions. So, time intervals are divided into "uncensored" (the observed volume causes a breakdown of traffic flow) and "censored" (traffic is fluent in the current interval and remains fluent in the following interval). The question was how to specify the transition point. The paper [2] proposes to use the speed as a criteria to determine time interval of that transition.

It should be noted, that the freeway capacity does not depend on the traffic flow rate, however, it depends on the number of other (including traffic flow) parameters. According to the HCM 2010, HBS 2015 and ODM 2012 they are geometric characteristics of the road, characteristics of the traffic flow (average speed, number of trucks, etc.), control conditions in case of traffic light control (e.g., cycle time, control parameters).

Remodeling Concept. The natural fact, that when analyzing the capacity in the current time interval, it is reasonable to include values of the selected parameters from the previous interval, to add the process dynamics. Regardless of which model is used to estimate the freeway capacity, this problem has to be fixed. The approach which could provide the new model to describe the capacity based on some existing models and "measured" capacity values in congested intervals is Mathematical Remodeling [4,6]. This is an approach to describe complex and/or composite systems based on the transition from mathematical or simulation models of one type to models of the other unified class. Depending on purposes and specific applied tasks, various interpretations of remodeling are possible. A theoretical model of some dependency built on the basis of its physical background, can have a structure which is quite complex and not appropriate for further analysis. In this case an array of dependency "input-output" data can be generated (which can be inaccessible under real conditions) and a simpler model of some unified structure with the required accuracy could be proposed. This is an approximation remodeling. To construct a new model a neural network and neuron-fuzzy models could be applied. In this case a remodeling has a neurostructural nature [5].

The following algorithm can be used to remodel freeway segment capacity.

Step 1. To divide the whole observed data into two subsets: (a) “congested” (uncensored) intervals, where the capacity rate could be estimated directly and (b) “fluid” (censored) intervals, where the capacity are obtained by using model (1). The criteria to separate intervals must be predefined at this step.

Step 2. Using the data sample obtained on Step 1, to train the neural network model with the predefined structure. On this step the analysis of the model accuracy must be done and corrections (in case of unsatisfied results) should be applied.

Step 3. Using the model obtained on Step 2, to estimate capacity rate within the new data set.

The proposed scheme is concerned to be a remodeling approach because it combines different ways to estimate capacity and presents a unified model to simulate this parameter.

Scope of Experiment. The initial for the modeling data obtained by the loop-detector system from German Autobahn A57. They are 1-minute intervals array from 01.01.2014 to 31.12.2014. Since only one segment of the freeway was involved in the study, some constant parameters (such as number of lanes, geometric characteristics, etc.) were neglected. The average speed in the current and previous time intervals, the percentage of trucks were considered as factors. Since the loop-detector system provides separate information on average speeds of personal vehicles and trucks, the average weighted speed indicator was used. Data were aggregated in 3, 5 and 10 min intervals and their different combinations to train and test neural network were applied.

To remodel the freeway capacity, various different structures of neural network were investigated and it was determined, that the best result demonstrated the network with one input layer, one hidden layer consisting of one neuron and the output layer. In its analytic form the model can be written as

$$y = \sigma\left(\sigma\left(w_{0,input} + \sum_{i=1}^{3} w_i x_i\right) + w_{0,hidden}\right). \tag{2}$$

Here y is freeway capacity (veh/h), $x_i = \{v_{cars}, tr, v_{cars}^{prev}\}$, v_{cars} is an average speed of vehicles in the current time interval (km/h), tr is a percentage of trucks and v_{cars}^{prev} is an average speed of vehicles in the previous time interval (km/h), w_i are weights, $w_{0,input}$ and $w_{0,hidden}$ are coefficients of the offset in the input and hidden layers respectively, $\sigma(\cdot)$ is the activation function. T

Numerical experiments shown that logistic function gives the best results

$$\sigma(net) = \frac{1}{1 + \exp(-net)}.$$

To train the neural network model (2), it was prepared a data sample including input parameters mentioned above and output capacity values obtained by the presented algorithm. The criteria to specify the time interval as congested was the average speed below 70 km/h (then the current flow rate was used as the actual capacity value). In other cases output values were simulated as a random Weibull-distributed value with defined parameters typical for this freeway segment. Different variants of neural network training and testing based on the duration of time intervals were studied. Presented numerical results demonstrate the duration of training intervals of 1 min and the duration of test intervals of 5 min, when the size of training sample was 393 120 intervals, the size of testing set was 26 496 intervals.

Figure 1 shows time series for flow rate and corresponding capacity. It is evident that the developed neural network well reacts to the increase of flow rate.

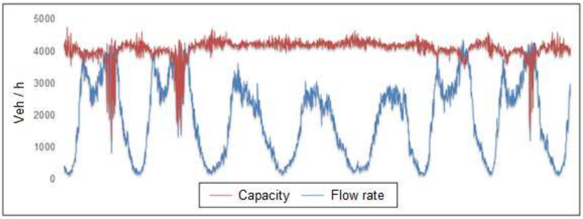


Fig. 1. Comparing flow rate and simulated capacity

Figure 2 shows the flow rate in comparison with the speed. Time intervals when traffic jam took place are specified. The accuracy of the neural network model (2) in prediction is 90.3% (of 81% of cases the traffic jam was identified).

Conclusion and outlook. Many approaches allow estimating the value of freeway segment capacity. But it is naturally (and confirmed empirically) to treat this important parameter as a random value with a certain statistical distribution. Regardless of which approach to estimate capacity is used, there are time intervals where this parameter could be measured directly. It occurs when the traffic flow has a congested regime.

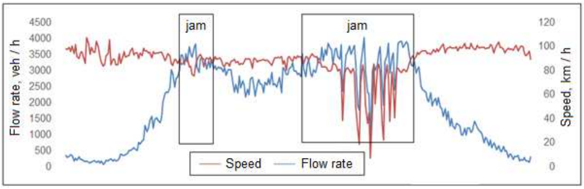


Fig. 2. Flow rate and speed in comparison

In the other time intervals capacity rate could be only estimated with some degree of probability. The proposed approach allows to remodel the estimation of freeway segment capacity, i.e. to build an analytical model representing the estimated parameter, based on data obtained by direct measurements in some cases and probabilistic estimates in others. The presented neural network model allows adding the dynamics to stochastic capacity in the current time interval by including flow parameters from the previous time intervals. It should also be mentioned, that the traffic flow parameters (primarily segment capacity) could be estimated at fixed in the meaning of time duration interval. The perspective problem is to find such a duration of time interval, using of which will give more adequate (in some sense) results.

References

1. van Toorenburg J. Praktijkwaarden voor de capaciteit. Rijkswaterstaat dienst Verkeerskunde, Rotterdam, the Netherlands, 1986.
2. Brilon W., Geistefeldt J. and Regler M. Reliability of Freeway Traffic Flow – A Stochastic Concept of Capacity // Proceedings of the 16th International Symposium on Transportation and Traffic Theory, College Park, Maryland, 2005, P. 125–144.
3. Bigazzi A. and Figliozzi M. An Analysis of the Relative Efficiency of Freeway Congestion Mitigation as an Emissions Reduction Strategy // Proceedings of the 90th Annual Meeting of the Transportation Research Board. Washington, D.C., 2011.
4. Sysoev A., Blyumin S. Saraev P. and Galkin A. Remodeling Approach as a Way to Automate Complicated Systems // Computer Science and Information Technologies, North America, Feb. 2018. Available at: `http://csit.ugatu.su/index.php/csit/csit2017/paper/view/5`.
5. Saraev P.V., Blyumin S.L., Galkin A.V. and Sysoev A.S. Neural Remodelling of Objects with Variable Structures. // Proceedings

of the Second International Scientific Conference "Intelligent Information Technologies for Industry" (IITI'17). Advances in Intelligent Systems and Computing, vol 679. Springer, Cham.

6. Galkin A., Sysoev A. and Saraev P. Variable Structure Objects Remodelling Based on Neural Networks // 2017 International Conference on Industrial Engineering, Applications and Manufacturing (ICIEAM), St. Petersburg, 2017, pp. 1-4.

Asymptotic analysis of complex stochastic systems

Stability conditions for constant retrial rate queueing system with a regenerative input flow[*]

L.G. Afanaseva

Department of Probability, Faculty of Mathematics and Mechanics, Lomonosov Moscow State University, Moscow, Russia

We consider a multiserver queueing system in which primary customers arrive according to a regenerative stream of rate $\lambda > 0$. The system has m stochastically identical servers. An arriving customer finding one or more servers idle obtains service immediately. Customers who find all servers busy go directly to the orbit. In the pioneering studies of retrial queues [1,2,3,4] it is assumed that each customer of the orbit generates a stream of repeated requests independently of the rest of the customers in the retrial group. Here we assume that the orbit works in the following way.

The requests for service appear through iid random intervals $\{\zeta_j\}_{n=1}^{\infty}$. If there are customers on the orbit and at least one server is idle then the service of one of these customers begins. Note, the orbit size (number of customers on the orbit) does not affect the orbit rate $\nu = (\mathsf{E}\zeta)^{-1}$.

[*] This research is partially supported by Russian Foundation for Basic Research grant 17-01-00468

This constant retrial policy was introduced by Fayolle [5] who modeled a telephone exchange system. Since this work, there was a rapid groth in the literature [6,7,8,9,10]. Note also that this retrial policy is a useful device for modeling the retrial phenomen in communication and computer networks. To the best of our knowledge, stability conditions for these systems have not yet been formally proved in general propositions with respect to input flow, service time distribution and retrial process. In the paper [11] $M|M|m$ retrial queue with a constant retrial rate and exponential distribution of interval ζ was studied by matrix-geometric methods. In particular the necessary and sufficient condition of stability of the system was obtained.

The main contribution of this paper is a proof of the stability criterion for the model with a regenerative input flow, general distributions of service times and intervals between requests from the orbit.

Now we discribe the model in more details. We consider an m-server queueing system S with a regenerative input flow $X(t)$ of rate λ. Let $\{\theta_j\}_{j=1}^{\infty}$ be a sequence of regeneration points for $X(t)$, $\tau_j = \theta_j - \theta_{j-1} (\theta_0 = 0)$ regeneration periods $(j = 1, 2, \ldots)$ and $\xi_j = X(\theta_j) - X(\theta_{j-1})$. Assume $\mathbf{E}\tau_1 < \infty$, $\mathbf{E}\xi_1 < \infty$ then w.p.1 $\lambda = \frac{X(t)}{t} = \frac{\mathsf{E}\xi_1}{\mathsf{E}\tau_1}$. For more details on these types of flows see, for instance, the works of Afanaseva and Bashtova [12], Thorisson [13]. We consider the retrial system with the constant retrial policy and the sequence $\{\zeta_j\}_{j=1}^{\infty}$ of iid random variables consists of time-intervals between requests from the orbit, $\nu^{-1} = \mathsf{E}\zeta$. Service times are defined by the sequence $\{\eta_n\}_{n=1}^{\infty}$ of iid random variables wth c.d.f. $B(x)$ and $\mathsf{E}\eta = b$.

Let $q(t)$ be the number of the customers and $n(t)$ the number of the occupied servers at the moment $t \geq 0$. We call $q(t)$ a stable process if there is

$$\lim_{t\to\infty} \mathsf{P}(q(t) \leq x) = \Phi(x) \tag{1}$$

and $\Phi(x)$ is a c.d.f. which does not depend on the initial state of the system.

Condition 1 $\mathsf{P}(\xi_1 = 0, \tau_1 > 0) + \mathsf{P}(\xi_1 = 1, \tau_1 - t_1 > \eta_1) > 0$*, where* $\theta_1 + t_1$ *is the arrival time of the customer on the first regeneration period and* η_1 *is its service time.*

Condition 2 *The random variable* ζ_n *has the second exponential phase. This means that* $\zeta_n = \zeta_n^{(1)} + \zeta_n^{(2)}$*, where* $\zeta_n^{(1)}$ *and* $\zeta_n^{(2)}$ *are independent random variables and* $\mathsf{P}(\zeta_n^{(2)} > x) = e^{-\delta x} (\delta > 0)$*.*

Under these conditions the process $q(t)$ is a regenerative one.

Also Conditions 1 and 2 provide the realization of the conditions of Theorem 1 from [14]. Therefore there are two possibilities for $q(t)$: $q(t)$ is a stable process or

$$q(t) \xrightarrow[t\to\infty]{P} \infty. \tag{2}$$

For the system S let $Y(t)$ be the number of served customers and $N(t)$ - the number of requests from the orbit up to time t.

Here we introduce an auxiliary retrial system $\tilde{S}$ with the same input flow $X(t)$, sequence $\{\eta_n\}_{n=1}^{\infty}$ of service times and the process $N(t)$ assuming that there are always customers in the orbit. Let $\tilde{Y}(t)$ and $\tilde{n}(t)$ be the processes $Y(t)$ and $n(t)$ respectively for $\tilde{S}$. Then the stochastic inequality

$$Y(t) \le \tilde{Y}(t) \tag{3}$$

takes place.

Although $X(t)$ and $\tilde{Y}(t)$ are dependent flows we may construct the common points of regeneration for them.

For instance, if Conditions 1 and 2 are fulfilled we define the common points of regeneration $\{\tilde{T}_n^{(0)}\}_{n=1}^{\infty}$ by the recursion

$$\begin{gathered}\tilde{T}_n^{(0)} = \min\{\theta_k \ge \tilde{T}_{n-1} : \bigcup_{k=1}^{\infty}\{\tilde{n}(\theta_k - 0) = 0\}\cap \\ \cap \bigcup_{s=0}^{\infty}\{\theta_k \in (Z_s + \zeta_{s+1}^{(1)}, Z_{s+1})\}\}, \\ \tilde{T}_0^{(0)} = 0.\end{gathered} \tag{4}$$

Here $Z_j = \zeta_1 + \ldots + \zeta_j, Z_0 = 0$. Since $\tilde{Y}(t)$ is a regenerative flow there is $\lim\limits_{t\to} \frac{\tilde{Y}(t)}{t} = \tilde{\lambda_Y}$.

Now let $\{\tilde{T}_n\}_{n=1}^{\infty}$ be a sequence of common regeneration points for $X(t)$ and $\tilde{Y}(t)$, $\tilde{\tau}_n = \tilde{T}_{n+1} - \tilde{T}_n$. Define the increments of $X(t)$, $Y(t)$ and $\tilde{Y}(t)$ on the regeneration period

$$\Delta_X(n) = X(\tilde{T}_{n+1}) - X(\tilde{T}_n), \Delta_Y(n) = Y(T_{n+1}) - Y(T_n), \tag{5}$$

$$\tilde{\Delta}_Y(n) = \tilde{Y}(T_{n+1}) - \tilde{Y}(T_n).$$

Then $\{\Delta_X(n), \tilde{\Delta}_Y(n)\}_{n=1}^{\infty}$ is a sequence of iid random vectors and

$$\lambda = \frac{\mathsf{E}\Delta_X(1)}{\mathsf{E}\tilde{\tau}_1}, \quad \tilde{\lambda}_Y = \frac{\mathsf{E}\tilde{\Delta}_Y(1)}{\mathsf{E}\tilde{\tau}_1}. \tag{6}$$

Consider the system S_0 with failures, m servers and a regenerative input flow $U(t) = X(t) + N(t)$, i.e. $Reg|G|m$. Service times are defined by the sequence $\{\eta_n\}_{n=1}^{\infty}$. As regeneration points for $U(t)$ we may take the sequence $\{\hat{T}_n\}_{n=1}^{\infty}$ defined by the recursion

$$\hat{T}_n = \min\{\theta_k > \hat{T}_{n-1} : \bigcup_{l=1}^{\infty}\{\theta_k \in (Z_l + \zeta_{l+1}^{(1)}, Z_{l+1})\}\}, \hat{T}_0 = 0. \qquad (7)$$

Then $\tilde{Y}(t)$ is the number of served customers up to time t and $\tilde{n}(t)$ is the number of busy servers at instant t in the system S_0.

Denote $\{t_k\}_{k=1}^{\infty}$ the sequential moments of jumps of the input flow $U(t)$. Since $\tilde{n}(t)$ is a regenerative process (under Conditions 1 and 2) then there exists $\lim\limits_{k\to\infty} \mathsf{P}(\tilde{n}(t_k) = j) = P_j(j = \overline{0, m})$.

Lemma 1 *Under Conditions 1 and 2*

$$\tilde{\lambda}_Y = \lim_{t\to\infty} \frac{\tilde{Y}(t)}{t} = (\lambda + \nu)(1 - P_m) \qquad (8)$$

The proof is based on the renewal theory.

We define the traffic rate for the system S as follows $\rho = \frac{\lambda}{\tilde{\lambda}_Y}$ where $\tilde{\lambda}_Y$ is given (6) or (8).

Condition 3 *The distribution of the service time has the second exponential phase.*

Theorem 1 *Let the Conditions 1 and 2 be fulfilled. If $\rho < 1$ then $q(t)$ is a stable process. If $\rho > 1$ or $\rho = 1$ and additionally Condition 3 is fulfilled then (2) takes place.*

For $\rho > 1$ the proof follows from the inequality (3) For the case $\rho = 1$ we construct the majorizing system and use results for a random walk with zero drift [13].

Consider the case $\rho < 1$. Without the loss of generality we assume that ζ_n has the first exponential phase. As common points of regeneration $\{\hat{T}_n^{(0)}\}_{n=1}^{\infty}$ for $X(t)$ and $\tilde{Y}(t)$ we take subsequence $\{\theta_{n_k}\}_{k=1}^{\infty}$ of the sequence $\{\theta_n\}_{n=1}^{\infty}$ such that θ_{n_k} gets into the first exponential phase of intervals $\{\zeta_n\}_{n=1}^{\infty}$ and $\tilde{n}(\theta_{n_k} - 0) = 0$ We define $\Delta_X^{(0)}(n)$, $\tilde{\Delta}_Y^{(0)}(n)$ and $\Delta_Y^{(0)}(n)$ by formulas (5) with $\tilde{T}_n^{(0)}$ instead of $\tilde{T}_n$.

If the convergence (2) takes place then for any $\epsilon > 0$ there is n_ϵ such that for $n > n_\epsilon$ we have $\mathbf{E}\Delta_Y^{(0)}(n) \geq \mathbf{E}\tilde{\Delta}_Y^{(0)}(1) - \epsilon$. Then for $\epsilon = (1-\rho)\mathsf{E}\tilde{\Delta}_Y^{(0)}(1)$ we obtain

$$\mathsf{E}q(\hat{T}_n^{(0)}) \leq \mathsf{E}q(\hat{T}_{n-1}^{(0)}) + \mathsf{E}\Delta_X^{(0)}(1) - \mathsf{E}\tilde{\Delta}_Y^{(0)}(1) + \epsilon = \mathsf{E}q(\hat{T}_{n-1}^{(0)})$$

that contradicts (2).

Theorem 2 *Let $X(t)$ and $N(t)$ be Poisson processes with rates λ and ν respectively. Then $q(t)$ is a stable process if and only if*

$$\frac{\lambda}{\lambda+\nu} < \sum_{j=0}^{m-1} \frac{\alpha^j}{j!} (\sum_{j=0}^{m} \frac{\alpha^j}{j!})^{-1} \tag{9}$$

where $\alpha = (\lambda+\nu)b$

Let us note that condition (9) is the same as obtained in [11] for a model with exponentially distributed service times.

References

1. Cohen J.W. Basic problems of telephone traffic theory and the influence of repeated calls// Philips Telecommunication Review. 1957. V. 18. P. 49–100.
2. Falin G.I. and Artalejo J.R. Approximations for multiserver queues with balking/retrial discipline// OR Spektrum. 1997. V. 17. P. 239–244.
3. Falin G.I. and Templeton J.G.C. Retrial queues// London: Chapman and Hall, 1997.
4. Wilkinson R.J. Theories for toll traffic engineering// USA Bell System Technical Journal. 1950. V. 35. P. 421–514.
5. Fayolle G. A simple telephone exchange with delayed feedbacks, in: Bpxma O.J.,Cohen J.W.,Tijms H.C.(Eds), in: Teletraffic Analysis and Computer Performance Evaluation, Elsevier, Amsterdam, 1986.
6. Artalejo J.R.: Stationary analysis of the characteristics of the $M|M|2$ queue with constant repeated attempts// Opsearch 33. 1996. P. 83–95.
7. Choi B.D., Shin Y.W., Ahn W.C. Retrial queues with collision arising from unslotted CSMA/CD protocol// Queueing Systems. 1992. V. 11. P. 335–356.

8. Choi B.D., Rhee K.H. and Pearce C.E.M. An $M|M|1$ retrial queue with control policy and general retrial times// Queueing Systems. 1993. V. 14. P. 275–292.
9. Choi B.D., Rhee K.H. and Park K.K. The $M|G|1$ retrial queue with retrial rate control policy// Probability in the Engineering and Informational Sciences. 1993. V. 7. P. 29–46.
10. Gomez-Corral A. and Ramalhoto M.F. On the stationary distribution of Markovian process arising in the theory of multiserver retrial queueing systems// Mathematical and Computer Modeling. 1999. V. 30. P. 141–158.
11. Artalejo J.R., Gomez-Corral A. and Neuts M.F. Analysis of multiserver queues with constant retrial rate.// European Journal of Operational Research. 2001. V. 135. P. 569–581.
12. Afanasyeva L.G. and Bashtova E.E.: Coupling method for asymptotic analysis of queues with regenerative input and unreliable server// Queueing Systems. 2014. V. 76, 2. P. 125–147.
13. Thorisson H. Coupling, Stationary and Regeneration// New-York: Springer, 2000.
14. Afanasyeva L. and Tkachenko A.: Multichannel queueing systems with regenerative input flow// Theory of Probability and Its Applications. 2014. V. 58, 2. P. 174–192.

Controlled reflected one-dimensional diffusions[*]

S.V. Anulova
V.A. Trapeznikov Institute of Control Sciences of Russian Academy of Sciences, Moscow, Russia

Given a standard probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t), P)$ and a one-dimensional $(\mathcal{F}_t)$ Wiener process $B = (B_t)_{t\geq 0}$ on it we considered controlled one-dimensional SDEs:

- a compact set $U \subset \mathbb{R}$ is a range of possible control values, and each control parameter is a measurable function $\alpha:\ \mathbb{R} \to U$;
- the coefficients for SDEs, the drift b and diffusion σ, are continuous bounded functions $U \times \mathbb{R} \to \mathbb{R}$, C^1 on $\mathbb{R}$, and σ is uniformly nondegenerate;

[*]This research is supported by RFBR grant No. 16-08-01285_a “Control over stochastic, deterministic and quantum systems in the fast motion phases”.

- for each control parameter α the corresponding SDE is

$$\begin{aligned} dX_t^\alpha &= b(\alpha(X_t^\alpha), X_t^\alpha)\, dt + \sigma(\alpha(X_t^\alpha), X_t^\alpha)\, dW_t, \quad t \geqslant 0, \\ X_0^\alpha &= x \in \mathbb{R}; \end{aligned} \tag{1}$$

  its (weak) solution exists [2] and under our conditions – one-dimensional, boundedness of all coefficients and uniform non-degeneracy of σ – is weakly unique;

- for ergodicity

$$\lim_{|x|\to\infty} \sup_{u\in U} x\, b(u,x) = -\infty. \tag{2}$$

  .

Given a running cost function $f : U \times \mathbb{R} \to \mathbb{R}$, continuous and bounded, for each $x \in \mathbb{R}$ and a control strategy $\alpha : \mathbb{R} \to U$ we define for the corresponding solution X^α the averaged cost function

$$\rho^\alpha(x) := \liminf_{T\to\infty} \frac{1}{T} \int_0^T \mathbb{E}_x f(\alpha(X_t^\alpha), X_t^\alpha)\, dt. \tag{3}$$

Maximizing this value for a fixed $x \in \mathbb{R}$, we discover a uniform α for all $x \in \mathbb{R}$, with all values equal: there exists $\rho := \rho^\alpha(x),\ x \in \mathbb{R}$ (see [3]).

The optimal control parameter α in [3] was an important generalization of results for this described problem with σ not depending on $u \in U$, see [4] (they have another object for ergodicity). The problem for SDEs with reflection was also solved in previous researches with this restriction. Now we generalize the result of [3] to reflection.

Now the controlled process is situated on $[0,\infty) \subset \mathbb{R}$ with reflection on 0.

Theorem. *The equation*

$$\begin{gathered} \max_{u\in U}\left[b(u,x)V^{'}(x) + \frac{1}{2}\sigma(u,x)V^{''}(x) + f(u,x) - \rho \right] = 0, \quad x \in (0,\infty), \\ V^{'}(0) = 0, \end{gathered} \tag{4}$$

holds true for ρ and some auxiliary function $V \in C^2$; solution of this equation is unique for ρ and for V up to a constant.

To prove this theorem we study again a full one-dimensional controlled process with coefficients: for all $x \in \mathbb{R}$

$$\bar{b}(u,x) = \begin{cases} b(u,x) & \text{for } x \geqslant 0, \\ -b(u,-x) & \text{for } x < 0 \end{cases} \quad \bar{\sigma}(u,x) = \begin{cases} \sigma(u,x) & \text{for } x \geqslant 0, \\ -\sigma(u,-x) & \text{for } x < 0 \end{cases}$$

$$\bar{f}(u,x) = \begin{cases} f(u,x) & \text{for } x \geqslant 0, \\ f(u,-x) & \text{for } x < 0. \end{cases}$$

For a given control α of the reflected process denote $\bar{\alpha}$ a control for the full process

$$\bar{\alpha}(x) = \begin{cases} \alpha(x) & \text{for } x \geqslant 0, \\ \alpha(-x) & \text{for } x < 0 \end{cases}$$

and $\bar{X}^{\bar{\alpha}}$ the corresponding full process. It is clear that for each control α the reflected process X^{α} equals the process $|\bar{X}^{\bar{\alpha}}|$ and hence the search of the optimal control comes to this for a full process. Although the optimal control is found in [3], we cannot use it for the reflected process: now the functions $\bar{b}, \bar{\sigma}$ are made not continuous by the point $x = 0$. Still we can generalize the result of [3] for this case, namely, we use again Theorem 1 of [1] because it is possible to prove the measurability in $U \times \mathbb{R}$ of (u,x) with u for x maximizing a consecutive function $v_n(x)$, $x \in \mathbb{R}, n = 0, 1, 2, ...$, described in the beginning of section "4 Main result" of [3].

References

1. L.D. Brown, R. Purves. Measurable selections of extrema. *Ann. Stat.*, 1: 902–912, 1973.
2. N.V. Krylov. On the selection of a Markov process from a system of processes and the construction of quasi-diffusion processes. *Math. USSR Izv.* 7: 691–709, 1973.
3. S.V. Anulova, H. Mai, A.Yu. Veretennikov. On averaged expected cost control as reliability for 1D ergodic diffusions. *Reliability: Theory and Applications.* No. 4(47): Volume 12, December 2017.
4. A. Arapostathis, Vivek S. Borkar, Mrinal K. Ghosh. Ergodic control of diffusion processes. *Encyclopedia of Mathematics and its Applications. Cambridge: Cambridge University Press*, volume 143, 2011.

Evolutionary operator for supercritical branching random walk with different branching sources*

D.M. Balashova
Lomonosov Moscow State University, Moscow, Russia

We consider continuous-time branching random walk (BRW) with a finite number of branching sources situated at the points $x_1, x_2 ... x_N$ on multidimensional lattice $\mathbb{Z}^d$ $(d \geqslant 1)$. A random walk is defined by a matrix of transition intensities $A = (a(x,y))_{x,y \in \mathbb{Z}}$, that is symmetric and $a(x,y) = a(y,x) = a(0, y-x) = a(y-x)$ for all x and y. Thus, the transition intensities are spatially homogenous. The walk is described in terms of the function $a(z), z \in \mathbb{Z}^d$, where $a(0) < 0, a(z) \geqslant 0$ when $z \neq 0$ and $a(z) = a(-z)$. We assume that $\sum\limits_{z \in \mathbb{Z}^d} a(z) = 0$ and the matrix A is irreducible, i.e. for all $z \in \mathbb{Z}^d$ there exists a set of vectors $z_1, z_2, ..., z_k \in \mathbb{Z}^d$ such that $z = \sum\limits_{i=1}^{k} z_i$ and $a(z_i) \neq 0$ for $i = 1, 2, ..., k$.

The transition probability $p(t, \cdot, y)$ is treated as a function $p(t)$ in $l^2(\mathbb{Z}^d)$ depending on time t and the parameter y. For time $h \to 0$

$$p(h,x,y) = a(x,y)h + o(h) \text{ for } y \neq x,$$
$$p(h,x,x) = 1 + a(x,x)h + o(h).$$

Then according to [1] we can rewrite the evolution equation as the following differential equation in space $l^2(\mathbb{Z}^d)$:

$$\frac{dp}{dt} = \mathcal{A}t, \quad p(0) = \delta_y,$$

where δ is Kronecker delta and the operator $\mathcal{A}$ acts as

$$(\mathcal{A}u)(z) := \sum_{z \in \mathbb{Z}^d} z(z-z')u(z').$$

Branching occurs at some sources x_i and is defined by an infinitesimal generating functions $f_i(u) = \sum_{n=0}^{\infty} b_{i,n} u^n$ such that $\beta_{i,r} = f_i^{(r)}(1) < \infty$ for all $r \in \mathbb{N}$. The value $\beta_i \equiv \beta_{i,1}$ characterizing the intensity of x_i source.

The behaviour of the mean number of particles both at an arbitrary point and on the entire lattice can be described in terms of the evolutionary operator of a special type (e.g. [1]), which is a perturbation of

*This research is supported by RFBR Grant no. 17-01-00468.

the generator $\mathcal{A}$ of a symmetric random walk. This operator has form

$$\mathcal{H}_\beta = \mathcal{A} + \sum_{i=1}^{N} \beta_i \delta_{x_i} \delta_{x_i}^T, \quad x_i \in \mathbb{Z}^d,$$

where $\mathcal{A} : l^p(\mathbb{Z}^d) \to l^p(\mathbb{Z}^d)$, $p \in [1, \infty]$ is a symmetric operator and $\delta_x = \delta_x(\cdot)$ denotes a column vector on the lattice taking the value one at the point x and zero otherwise.

Green function of the operator $\mathcal{A}$ can be represented as the Laplace transform of the transition probability $p(t, x, y)$:

$$G_\lambda(x, y) := \int_0^\infty e^{-\lambda t} p(t, x, y) dt = \frac{1}{(2\pi)^d} \int_{[-\pi,\pi]^d} \frac{e^{i(\theta, y-x)}}{\lambda - \phi(\theta)} d\theta, \quad \lambda \geqslant 0,$$

where $\phi(\theta) = \sum_{z \in \mathbb{Z}^d} a(z) e^{i(\theta, z)}, \quad \theta \in [-\pi, \pi]$.

The average number of hits of the particle to the point y with the start of the process from the point x as time tends to infinity is $G_0(x, y)$. We denote $G_0 := G_0(0, 0)$.

If the condition of finite variance of jumps

$$\sum_{z \in \mathbb{Z}^d} |z|^2 a(z) < \infty, \tag{1}$$

where $|z|$ — the Euclidean norm of a vector z, $G_0 = \infty$ for $d = 1$ and $d = 2$ and $G_0 < \infty$ for $d \geqslant 3$.

If for all sufficiently large norm $z \in \mathbb{Z}^d$ it is satisfied the asymptotic property

$$a(z) \sim \frac{H\left(\frac{z}{|z|}\right)}{|z|^{d+\alpha}}, \quad \alpha \in (0, 2), \tag{2}$$

where $H(\cdot)$ — positive function, symmetric on the sphere $S^{d-1} = \{z \in R^d : |Z| = 1\}$, then $G_0 = \infty$ for $d = 1$, $\alpha \in [1, 2)$ and $G_0 < \infty$ for$d = 1$, $\alpha \in (0, 1)$ or for $d \geqslant 2$, $\alpha \in (0, 2)$. Condition (2) leads to a convergence series $\sum_{z \in \mathbb{Z}^d} |z|^2 a(z)$ and infinite variance of jumps.

Analysis of such operator in more general form was investigated in [1]. The perturbation of the form $\sum_{i=1}^{N} \beta_i \delta_{x_i} \delta_{x_i}^T$ of the operator $\mathcal{A}$ may result in the emergence of positive eigenvalues of the operator $\mathcal{H}_\beta$ and the multiplicity of each of them does not exceed N. In [2] it was proved that for the case of equal β_i and finite variance of jumps the total multiplicity of all eigenvalues (counting their multiplicity) does not exceed N and the multiplicity of each eigenvalue of the operator $\mathcal{H}_\beta$ does not exceed $N-1$. In [3] the case of infinite variance of jumps was analyzed and it was demonstrated that the appearance of multiple lower eigenvalues in the spectrum of the evolutionary operator can be caused by a simplex configuration of branching sources.

The main aim of the study is a numerical analysis of phase transitions in the supercritical case, where branching sources have negative and positive intensities and model with finite number of different sources in an arbitrary configuration.

Branching sources with positive and negative intensities

We consider branching random walk with $N = p + n$ sources, that are located in the vertices of a simplex on $\mathbb{Z}^d$. P sources $x_1 ... x_p$ have intensities $\beta > 0$ and n sources $x_{p+1} ... x_{p+n}$ have intensities $(-\beta)$. Source intensity is negative when degree of death prevails over the degree of birth. The evolutionary operator in this case has form

$$\mathcal{H}_\beta = \mathcal{A} + \beta\Delta_{x_1} + \beta\Delta_{x_2} + \cdots + \\ + \beta\Delta_{x_p} - \beta\Delta_{x_{p+1}} - \beta\Delta_{x_{p+2}} - \cdots - \beta\Delta_{x_{p+n}}.$$

For $p \geqslant 2$ denote $\beta_c := \beta_c(n,p)$ is such a minimal positive intensity that for $\beta > \beta_c$ operator $\mathcal{H}_\beta$ has positive eigenavalues and $\beta_{c_1} > \beta_c$ that for $\beta \in (\beta_c, \beta_{c_1})$ operator $\mathcal{H}_\beta$ has only one eigenavalue $\lambda_0(\beta)$.

Theorem 1 *The amount of eigenvalues $\lambda > 0$ of the evolutionary operator $\mathcal{H}_\beta$ (counting their multiplicity) does not exceed the amount of branching sources with positive intensities, the maximal of these eigenvalues is simple and $\beta_{c_1} > \beta_c$.*

Assume $|x_i - x_j| = s, i \neq j$. Then according to [1] $\lambda > 0$ is an eigenvalue of the operator $\mathcal{H}_\beta$ if and only if

$$(\beta G_\lambda - \beta G_\lambda(s) - 1)^{p-1}(\beta G_\lambda - \beta G_\lambda(s) + 1)^{n-1} \\ \times ((\beta G_\lambda)^2 + (p+n-2)\beta^2 G_\lambda G_\lambda(s) - (p+n-1) \\ \times (\beta G_\lambda(s))^2 + (p-n)\beta G_\lambda(s) - 1) = 0$$

and hence the values β_c and β_{c_1} can be found:

$$\beta_c = \frac{(n-p)\tilde{G}_0 + \sqrt{(n-p)^2(\tilde{G}_0)^2 + 4(G_0-\tilde{G}_0)(G_0+\tilde{G}_0(n+p-1))}}{2(G_0-\tilde{G}_0)(G_0+\tilde{G}_0(n+p-1))},$$

$$\beta_{c_1} = \frac{1}{G_0-\tilde{G}_0}.$$

Note that

$$\beta_{c_1} - \beta_c = \frac{8\tilde{G}_0 p(G_0+\tilde{G}_0(n+p-1))}{G_0-\tilde{G}_0} > 0.$$

Branching sources with different positive intensities

Let put N brancing sources with arbitrary positive intensities in an arbitrary configuration. We denote $\beta_{min} := \min_i\{\beta_i\}$ and $\beta_{max} := \max_i\{\beta_i\}$. Let $\beta_{a_{min}}$ is a minimal value such that for $\beta_{min} > \beta_{a_{min}}$ operator $\mathcal{H}_{\beta_1,\dots,\beta_N}$ contains a positive eigenvalue and $\beta_{a_{max}}$ is a maximal value such that for $\beta_{max} < \beta_{a_{max}}$ operator $\mathcal{H}_{\beta_1,\dots,\beta_N}$ is not contains positive eigenvalues.

Theorem 2 *Let BRW satisfy the condition of finite varience of jumps (1) or infinite varience of jumps (2). If* $G_0 = \infty$*, then BRW is supercritical and* $\beta_{a_{min}} = 0$ *for* $N \geqslant 1$*. If* $G_0 < \infty$*, then* $\beta_{a_{min}} = G_0^{-1}$ *for* $N = 1$ *and* $0 < \beta_{a_{max}} \leqslant \beta_{a_{min}} < G_0^{-1}$ *for* $N > 1$*.*

References

1. Yarovaya E.B. (2012) Spectral properties of evolutionary operators in branching random walk models. Mathematical Notes 92(1):115–131.
2. Antonenko E.A., Yarovaya E.B. (2016) On the Number of Positive Eigenvalues of the Evolutionary Operator of Branching Random Walk. Branching Processes and their Applications, Book. Lecture Notes in Statistics, Spriger, vol. 219, p. 41–55.
3. Yarovaya E.B. (2016) Positive discrete spectrum of the evolutionary operator of supercritical branching walks with heavy tails. Methodology and Computing in Applied Probability pp 1–17, DOI 10.1007/s11009-016-9492-9.

4. Horn R.A., Johnson C.R. (1989) Matrix analysis. Cambridge University Press, Cambridge. doi:10.1017/CBO9780511810817.
5. Gradshteyn I. S., Ryzhik I. M. (2000) Tables of Integrals, Series, and Products, 6th ed. San Diego, CA: Academic Press, p. 1101.

Stochastic counterparts for nonlinear parabolic systems*

Ya.I. Belopolskaya and A.O. Stepanova
PDMI RAS, SPbSUACE, St.Petersburg, Russia

We consider the Cauchy problem for systems of nonlinear second order parabolic equations describing conservation and balance laws in physics, chemistry, biology and other fields. We show that these systems allow a probabilistic interpretation as systems of forward Kolmogorov equations for corresponding Markov processes. Motivated by this interpretation we are interested in generalised solutions of the forward Cauchy problem for systems with fully nondiagonal second order terms (systems with cross-diffusion) as well as classical,. We construct stochastic representations for generalised solutions of the forward Cauchy problem solutions in terms of diffusion processes and their multiplicative functionals [1].

For a certain subclass of these systems which includes systems with diagonal principal part with different coefficients of the second order terms and nondiagonal first and/or zero order terms we reduce the construction of a generalised solution of the forward Cauchy problem for a PDE system to the correspondent stochastic problem. Namely, without appealing to an original nonlinear PDE solution we derive a closed stochastic problem and solve it under some suitable assumptions. Finally, we state conditions on the stochastic problem data that allow to verify the solution of the stochastic problem gives rise to the required generalised solution of the original PDE problem [2].

Recall that stochastic approach to investigation of a PDE or a system of PDEs includes three steps. The first is to find a stochastic representation of a solution to the problem under consideration. The second is to derive a closed stochastic system which we call a stochastic counterpart of the original PDE problem. The final step is to investigate the stochastic system derived at the second step and to prove that it yields a required solution of the PDE problem under consideration.

*This research is supported by RNF Grant 17-11-01136

To be more precise we consider two types of nonlinear parabolic systems

$$\frac{\partial u_m}{\partial t} = \frac{1}{2}\sum_{i,j=1}^{d} \frac{\partial}{\partial x_i \partial x_j}[F_{ij}^m(x)u_m] + \sum_{l=1}^{d_1} c_{ml}(u)u_l, \tag{1}$$

and

$$\frac{\partial u_m}{\partial t} + \sum_{i=1}^{d}\sum_{l=1}^{d_1} \frac{\partial}{\partial x_i}[B_i^{ml}(x,u)u^l] = \frac{1}{2}\sum_{i,j=1}^{d} \frac{\partial^2}{\partial x_i \partial x_j}[F_{ij}^m(x)u_m], \tag{2}$$

where $F_{ij}^m(x) = \sum_{k=1}^d A_{ik}^m(x)A_{kj}^m(x)$.

Our aim is to study the probabilistic interpretation of the Cauchy problem solution for them with initial data $u_m(0,x) = u_{0m}(x)$. We show that there exist stochastic processes which allow to construct stochastic representations of a solution to the required Cauchy problem. In other words we will treat these systems as systems of forward Kolmogorov equations for some Markov processes and thus we will be interested in generalised (weak) solutions of these Cauchy problems.

We start with reaction diffusion systems and construct its stochastic counterpart. To this end first we give two definitions of a weak solution to the system (1) which are equivalent in cases under consideration (see [4] lemma 1.1) but are useful for our purposes. Let $H = [L^2(R^d)]^{d_1}$, and $\mathcal{D} = [C_0^\infty(R^d)]^{d_1}$ be the space of infinite differentiable functions with compact supports, $W_{d_1}^k = [W^k(R^d)]^{d_1}$ be the space of k differentiable functions with square integrable derivatives up to k-th order. and $W_{d_1}^{-k}$ be its dual. We denote by $\langle u, h\rangle = \sum_m \int_{R^d} u_m(x)h_m(x)dx$ pairing between these spaces as well as the inner product in H.

We say that u is a weak solution of (1) if the integral identity

$$\frac{\partial}{\partial t}\langle u_l(t), h_l\rangle + \langle u_l(t), \frac{1}{2}\sum_{i,j=1}^{d} F_{ij}^l(x)\frac{\partial^2 h_l}{\partial x_i \partial x_j}\rangle + \tag{3}$$

$$+\langle u_l(t), \sum_{m=1}^{d_1} c_{lm}^u(x)h_m\rangle = 0$$

hold for $m = 1, 2$ and arbitrary $h_m \in \mathcal{D} \cap L^2(R)$ and we say that u is a weak solution of (2) if the integral identity

$$\frac{\partial}{\partial t}\langle u_l(t), h_l\rangle + \langle u_l(t), \frac{1}{2}\sum_{i,j=1}^{d} F_{ij}^l(x)\frac{\partial^2 h_l}{\partial x_i \partial x_j}\rangle + \tag{4}$$

$$+\langle u_l(t), \sum_{i=1}^{d}\sum_{m=1}^{d_1} B_i^{lm}(x,u)\frac{\partial}{\partial x_i}h_m\rangle = 0$$

hold for $l = 1,2$ and arbitrary $h_l \in \mathcal{D} \cap L^2(R)$.

In addition we need an alternative definitions, namely, we say that u is a weak solution of the system (1), when the integral identity

$$\langle u_{0l}, h_l(0)\rangle + \int_0^T \langle u_l(\theta), \left[\frac{\partial h_l(\theta)}{\partial\theta} + \frac{1}{2}\sum_{i,j=1}^{d} F_{ij}^m(x)\frac{\partial^2 h_l(\theta)}{\partial x_i \partial x_j}\right]\rangle d\theta + \tag{5}$$

$$+\int_0^T \langle u_l(\theta), \sum_{m=1}^{d_1} c_{lm}^u h_m(\theta)\rangle d\theta = 0$$

hold for $l = 1,2$ and $\forall h_m \in C_0^{1,k}([0,T]\times R^d) \cap L^2([0,T]\times R^d)$.

In the same manner we say that u is a weak solution of the system (2), when the integral identity

$$\langle u_{0l}, h_l(0)\rangle + \int_0^T \langle u_l(\theta), \left[\frac{\partial h_l(\theta)}{\partial\theta} + \frac{1}{2}\sum_{i,j=1}^{d} F_{ij}^l(x)\frac{\partial^2 h_l(\theta)}{\partial x_i \partial x_j}\right]\rangle d\theta + \tag{6}$$

$$+\int_0^T \langle u_l(\theta), \sum_{i=1}^{d}\sum_{m=1}^{d_1} B_i^{lm}(x,u)\frac{\partial h_m(\theta)}{\partial x_i}\rangle d\theta = 0$$

hold for $l = 1,2$ and $\forall h_l \in C_0^{1,k}([0,T]\times R^d) \cap L^2([0,T]\times R^d)$.

Note that the second couple of definitions allows us to obtain the form of a generator of a Markov process we are looking for. It appears to be of crucial importance when we deal with fully nondiagonal systems of parabolic equations [3],[4].

Below we consider some particular cases of systems of the form (1) and (2) and restrict ourself to the cases $d = 1, d_1 = 2$.

The systems which we analyse in this paper are a reaction-diffusion system of the form

$$\frac{\partial u_1}{\partial t} = \sigma_1^2\frac{\partial^2 u_1}{\partial x^2} + \left[-(b+1) + u_1 u_2 + \frac{a}{u_1}\right]u_1, \quad u_1(0,x) = u_{01}(x), \tag{7}$$

$$\frac{\partial u_2}{\partial t} = \sigma_2^2\frac{\partial^2 u_2}{\partial x^2} + [bu_1 - u_1^2]u_2, \quad u_2(0,x) = u_{02}(x) \tag{8}$$

called the Brusselator system and the MHD-Burgers system

$$\frac{\partial u_1}{\partial t} + \frac{\partial (u_1 u_2)}{\partial x} = \frac{\sigma^2}{2} \frac{\partial^2 u_1}{\partial x^2}, \quad u_1(0,x) = u_{10}(x) \tag{9}$$

$$\frac{\partial u_2}{\partial t} + \frac{1}{2} \frac{\partial (u_1^2 + u_2^2)}{\partial x} = \frac{\mu^2}{2} \frac{\partial^2 u_2}{\partial x^2}, \quad u_2(0,x) = u_{20}(x) \tag{10}$$

where σ_1, σ_2, a, b are given constants.

The above systems can be treated as systems of forward Kolmogorov equations for specific Markov processes which will be described below. Due to this interpretation it is natural to look for generalised solutions of the Cauchy problem for them.

Let $(\Omega, \mathcal{F}, P)$ be a given probability space and $w(t) \in R$ be a standard Wiener process. Consider a process $\hat{\xi}(t) = x - \sigma w(t)$ and denote by $\psi_{0,t} : x \to \hat{\xi}(t)$ a random map acting in R, called a stochastic flow generated by $\hat{\xi}(t)$.

A stochastic counterparts of the Cauchy problem (7), (8) and (9), (10) are presented in the following theorems.

Theorem 1.*Assume that there exists a unique generalized solution $u \in \mathcal{W}_T^2 \times \mathcal{W}_T^2$ of the Cauchy problem (7), (8). Then it admits a probabilistic representation of the form*

$$u_m(t,x) = E[\tilde{\eta}(t) u_{0m}(\hat{\xi}_m(t))]. \tag{11}$$

where

$$\hat{\xi}_m(t) = x - \sqrt{2}\sigma_m w(t), \quad \tilde{\eta}(t) = \exp\{\int_0^t c_m^u(\psi_{\theta,t}(x)) d\theta\}, \tag{12}$$

$$c_1^u(x) = -(b+1) + u_1 u_2 + \frac{a}{u_1}, \quad c_2^u(x) = \frac{b u_1}{u_2} - u_1^2,$$

and $\psi_{0,t}^m$ are stochastic flows generated by $\hat{\xi}^m(t)$.

Theorem 2.*Assume that there exists a unique generalized solution $u \in \mathcal{W}_T^2 \times \mathcal{W}_T^2$ of the Cauchy problem (9), (10). Then it admits a probabilistic representation of the form*

$$u^m(t) = E\left[\tilde{\eta}_m(t) u_{0m} \circ \psi_{0,t}^m\right], \quad m = 1,2, \tag{13}$$

where

$$d\hat{\xi}^m(\theta) = -\sigma_m dw(\theta), \tag{14}$$

$$\tilde{\eta}^m(t) = exp\left\{\int_0^t C^u_m(\psi_{\theta,t}(x))dw(\theta) - \frac{1}{2}\int_0^t [C^u_m]^2(\psi_{\theta,t}(x))d\theta\right\}, \quad (15)$$

$$C^u_1(x) = \frac{1}{\sigma}u_2(\theta,x), \quad C^u_2(x) = \frac{1}{2\mu}\left[u_2(\theta,x) + \frac{u_1^2(\theta,x)}{u_2(\theta,x)}\right],$$

and $\psi^m_{0,t}$ are stochastic flows generated by $\hat{\xi}^m(t)$.

Theorem 3. Given bounded strictly positive square integrable initial functions u_{0m} there exists an interval $[0,T]$ such that for all $t \in [0,T]$ there exist a solution to the system (11), (12) and the function $u = (u_1,u_2)$ of the form (11) is a weak solution of (7), (8).

Theorem 4. Given bounded strictly positive square integrable initial functions u_{0m} there exists an interval $[0,T]$ such that for all $t \in [0,T]$ there exist a solution to the system (13)– (15) and the function $u = (u_1,u_2)$ of the form (13) is a weak solution of (9), (10).

See detailed proof of the theorems 2 and 4 in [6]. The proof of theorems 1 and 3 will be given in a forthcoming paper.

References

1. Belopolskaya Ya. I. Dalecky Yu.L. Stochastic equations and differential geometry. Kluwer Academic Publishers. 1990. 260.
2. Belopolskaya Ya., Woyczynski W. Generalized solution of the Cauchy problem for systems of nonlinear parabolic equations and diffusion processes.// Stochastics and dynamics. 2012. V. 11, 1 P. 1–31.
3. Belopolskaya Ya. I. Probabilistic models of the dynamics of the growth of cells under contact inhibition // Mathematical Notes. 2017. V. 101, 3, P. 346–358.
4. Belopolskaya Ya. I. Probabilistic representation of the Cauchy problem solutions for systems of nonlinear parabolic equations // Global and Stochastic Analysis. 2016. V. 3, 1, P. 25–32.
5. Bogachev V., Röckner M., Shaposhnikov S. On uniqueness problems related to the Fokker-Planck-Kolmogorov equations for measures. J. Math. Sci. 2011. Vol. 179, No. 1, pp. 7–47.
6. Belopolskaya Ya., Stepanova A. Stochastic interpretation of MHD-Burgers system. Zapiski nauchn.sem PDMI. 2017. Vol. 446, pp. 7–29.

Topological, metric (Harris) and Poincare recurrences for general Markov chains

M.L. Blank

Russian Academy of Sci. Inst. for Information Transmission Problems,and National Research University Higher School of Economics, Moscow, Russia

In this work we discuss a collection of questions related to the idea of recurrence in general general Markov chains. The recurrence property is well known in theory describing two very different classes of random systems: lattice random walks and ergodic theory of continuous selfmaps. On the other hand, the recurrence property is next to being neglected in general theory of Markov chains (perhaps except for a few notable exceptions which we will discuss in detail).

In literature one finds two approaches to the definition of a recurrent point: topological and metric (defined in very different contexts). Since our aim is to introduce their analogies for general Markov chains let us recall basic definitions.

By an inhomogeneous Markov chain one means a random process $\xi_t : (\Omega, \mathcal{F}, P) \to (X, \mathcal{B}, m)$ acting on a Borel $(X, \mathcal{B})$ space with a finite reference measure m (not connected to ξ_t). This process is completely defined by a family of *transition probabilities*

$$Q_s^t(x, A) := P(\xi_{s+t} \in A | \xi_s = x), \ A \in \mathcal{B},$$

having standard properties:

- For fixed s, t, x the function $Q_s^t(x, \cdot)$ is a probability measure on the σ-algebra $\mathcal{B}$.
- For fixed s, t, A the function $Q_s^t(\cdot, A)$ is $\mathcal{B}$-measurable.
- For $t = 0 \quad Q_s^t(x, A) = \delta_x(A)$.
- For each $s, 0 \leqslant t \leqslant t'$ and $A \in \mathcal{B}$ we have

$$Q_s^{t'}(x, A) = \int_X Q_s^t(x, dy) Q_t^{t'-t}(y, A).$$

The process ξ_t induces the action on measures:

$$Q_s^t \mu(A) := \int Q_s^t(x, A) \ d\mu(x)$$

and the action on functions:

$$Q^t_s\phi(x) := \int \phi(y)Q^t_s(x,dy).$$

In particular, the well known Feller property in terms of the action on functions means that $Q^t_s : C^0 \to C^0 \;\; \forall s,t \geqslant 0$.

A Borel measure μ is said to be *invariant* or *stationary* for the Markov chain ξ_t if it is a solution to the equation

$$Q^t_s\mu = \mu \;\; \forall s,t.$$

By the *t-preimage* with $t \geqslant 0$ of a set $B \in \mathcal{B}$ under the action of the homogeneous Markov chain ξ_t we call the set of points

$$Q^{-t}(B) := \{x \in X : \quad Q^t(x,B) > 0\}.$$

Now we are ready to return to the notion of recurrence. Observe that since the phase space is equipped if the Borel σ-algebra, it is equipped with the corresponding topology as well. We start from the well known in the field of topological dynamics (see e.g. [1]) notion of the topological recurrence.

A point $x \in X$ is called *topologically recurrent* if for any open neighborhood $U \ni x$ for each s there exists an (arbitrary large) $t = t(x,U,s)$ such that $Q^t_s(x,U) > 0$ (i.e. a trajectory eventually returns to U with positive probability).

A point $x \in X$ is called *metrically recurrent* if for any set $V \ni x$ of positive m-measure and any s there exists an (arbitrary large) $t = t(x,V,s)$ such that $Q^t_s(x,V) > 0$ (i.e. a trajectory eventually returns to V with positive probability).

This is our modification of the metric version of the recurrence notion proposed by T.E. Harris in [2] in order to get reasonably general assumptions guaranteeing the existence of an invariant measure. In fact, T.E. Harris used this property only in the case when the reference measure m is invariant with respect to the process. Another weak point of this approach is that whence a point x is metrically recurrent, the corresponding trajectory (realization of the process) needs to visit any set of positive measure with positive probability, which looks away too excessive.

The 3d approach to this notion is well known and studied in ergodic theory of deterministic dynamical systems, but again only in the case when the measure m is dynamically invariant.

A point $x \in X$ is called *Poincare recurrent* with respect to a $\mathcal{B}$-measurable set $A \ni x$ if for each s there exists an (arbitrary large) $t = t(x, A, s)$ such that $Q_s^t(x, A) > 0$ (i.e. a trajectory eventually returns to A with positive probability).

The main question of interest for us is to say how large is the set of recurrent points?

The celebrated Poincare recurrence theorem (see e.g. [3]) states that for a measurable discrete time dynamical system (T, X) and any measurable set A the set of non Poincare recurrent points is of zero measure with respect to any T-invariant measure. There is a number of problems for the generalization of this result for a general Markov chain: continuous time, inhomogeneity, and more important a principal absence of invariant measures (stationary distributions) for inhomogeneous Markov chains. Nevertheless our main result in this direction is as follows.

Theorem 1. *Let m be a finite measure on $(X, \mathcal{B})$ and let Q_s^t does not depend on s. Then the property that for each set $A \in \mathcal{B}$ the set of Poincare recurrent points in A is of full m-measure (i.e. its complement in A is of zero m-measure) is equivalent to the existence of a constant $\gamma > 0$ such that*

$$\sum_{n \geqslant 1} m(Q^{-n\gamma}(A) \cap A) = \infty \quad \forall A \in \mathcal{B}: \; m(A) > 0.$$

An important observation here is that if Q_s^t does depend on s, then (under a slight modification of the preimage definition) the direct statement in theorem 1 is still correct, but the inverse one fails. Namely, it is possible that the above sum in is infinite, but for some s

$$\limsup_{t \to \infty} P(\xi_{s+t} \in A | \xi_s \in A) = 0.$$

The reason for this is that despite the chain of "(pre)images" of the set A inevitably intersect itself an infinite number of times, but the original set needs not to be included to the intersections. A sketch of the counterexample gives a finite state discrete time Markov with 4 states, whose graph of transitions is given by the following diagram:

$$x_3 \leftarrow x_3 \leftarrow x_0 \leftarrow x_1 \longleftrightarrow x_2.$$

To be more precise, we consider a deterministic version of the process and denoting the only image of a state x at time s by $T_s^1 x$ and its unique preimage by $T_s^{-1}x$ we have

$$A := x_0 = T_1^{-1}x_1,\ T_2^{-1}x_1 = x_2,\ T_3^{-1}x_2 = x_1,\ T_1^1 x_0 = x_3,\ T_i^1 x_3 = x_3\ \forall i.$$

Observe that for the state x_0 the relation (*) holds true for the uniform measure, but this set is non-recurrent.

The situation with other types of recurrence is much more subtle, in particular we demonstrate that a very "good" Markov chain for which all points are topologically recurrent with respect to the "standard" topology may change drastically when one chooses a different topology instead. Under this new topology all points might be non-recurrent.

The following result gives sufficient conditions for the topological recurrence of "typical" points.

Theorem 2. *Let the space $(X, \mathcal{T})$ be compact with respect to the topology $\mathcal{T}$ compatible with the σ-algebra $\mathcal{B}$. Then under the assumptions of Theorem 1 m-almost every point $x \in X$ is topologically and metrically recurrent.*

References

1. Nitecki Z. Differentiable dynamics. An introduction to the orbit structure of diffeomorphisms. Cambridge: The MIT Press. 1974
2. Harris T.E. The existence of stationary measures for certain Markov processes // Matematika, 1960, Volume 4, Issue 1, 131–143.
3. Cornfeld I.P., Fomin S.V., Sinai Y.G. Ergodic Theory. New York: Springer-Verlag. 1982.

Asymptotic analysis of some applied probability systems*

E.V. Bulinskaya

Lomonosov Moscow State University, Moscow, Russia

Introduction

The aim of presentation is asymptotic analysis of the mathematical models describing the real-life systems pertaining to insurance, inventories or queueing. It is necessary for establishing the systems stability with

*This research is supported by Russian Foundation for Basic Research grant 17-01-00468

respect to small parameters fluctuations and perturbations of underlying distributions and providing the optimal control.

We begin by treating a discrete-time insurance system (or other organization) which is interested in short-term credits (or bank loans). It is supposed that at the beginning of each period (year, month or week) it is possible to apply to a bank in order to obtain a credit card valid for a fixed number of periods. The card is provided immediately. The upper limit z of the credit is chosen by the applicant who pays bank at once the amount cz where c is the interest rate. The loan is used to satisfy the claims flow described by a sequence of nonnegative i.i.d. random variables $\{\xi_n\}_{n\geqslant 1}$. We assume that each claim ξ has a known distribution function $F(t)$ possessing a density $\varphi(t) > 0$ for $t > 0$ and a finite expectation. If a claim amount ξ is larger than the cash amount u available for payment then another loan is obtained at the interest rate p, $p > c$, its size is $\xi - u$. The amount $u - \xi$, not used for payment before the card expiration term, is lost. Moreover, if the credit is taken in a currency differing from that of the claims, the financial loss of applicant is equal to $k(u - \xi)$. Here the constant k depends on the exchange rate. The fixed transaction cost K may be taken into account as well. Our goal is to determine the optimal n-period strategy of applicant. Optimality means the minimization of expected discounted costs entailed by the n-step credit strategy. Below we formulate the results only for the simplest case.

One-period credit

Assume that the credit is valid for one period only. That means, the money not used for payment during the period cannot be used later. Denote by $f_n(x)$ the minimal expected discounted costs incurred by the implementation of n-period credit strategy. Here x is the cash amount available initially for claims payment if $x > 0$ and $|x|$ is the debt amount if $x < 0$. Put $H_1(y) = cy + L(y)$ with $L(y) = p\int_y^\infty (s-y)\varphi(s)\,ds + k\int_0^y (y-s)\varphi(s)\,ds$ and $y = x + z$ where z is the credit limit. If we do not take into account the transaction costs (in other words, put $K = 0$) then the following statements are valid.

Lemma 1. *For any x,*

$$f_1(x) = -cx + \min_{y\geqslant x} H_1(y).$$

If $p > c$ then there exists the critical level $\bar{x}_1$ defined by the relation

$$F(\bar{x}_1) = (p-c)(p+k)^{-1}$$

such that the optimal credit limit is given by $z_1(x) = \max(0, \bar{x}_1 - x)$. *The function* $f_1(x)$ *is twice differentiable and convex, whereas*

$$f_1'(x) = \begin{cases} -c, & x \leqslant \bar{x}_1, \\ L'(x), & x \geqslant \bar{x}_1. \end{cases}$$

Turning to multi-period case we introduce a discount factor $\alpha \in (0, 1)$ and establish the following results.

Theorem 1. *The function* $f_n(x)$ *specified by*

$$f_n(x) = -cx + \min_{y \geqslant x} H_n(y), \tag{1}$$

where $H_n(y)$ *has the form*

$$H_n(y) = H_1(y) + \alpha f_{n-1}(0)F(y) + \alpha \int_y^\infty f_{n-1}(y - s)\varphi(s)\, ds, \tag{2}$$

is twice differentiable and convex for all $n > 1$. *There exists* $\bar{x} > \bar{x}_1$ *such that the optimal credit limit* $z_n(x) = \max(0, \bar{x} - x)$ *for any* x *and* $n > 1$. *The critical level* $\bar{x}$ *is defined by the relation*

$$F(\bar{x}) = (p - c(1 - \alpha))(p + k + \alpha c)^{-1}. \tag{3}$$

According to Bellman optimality principle (see, e.g., Bellman [1]) it is possible to conclude that equations (1) and (2) are valid. Further proof is carried out by induction.

Thus, for any $n > 1$, the optimal credit strategy is determined by a single critical number $\bar{x}$, whereas for one-step case it is necessary to use $\bar{x}_1$ instead of $\bar{x}$.

Corollary. *The critical level* $\bar{x}_1$ *is an increasing function of* p *and a decreasing function of* c *and* k, *whereas* $\bar{x}$ *increases in* p *and* α *and decreases in* c *and* k.

Asymptotic analysis

Now we use the introduced short-term credit model to show how one performs the asymptotic analysis of multi-step processes. First of all we establish the limit behavior of the minimal costs as the planning horizon tends to infinity.

Theorem 2. *If* $\alpha < 1$ *the functions* $f_n(x)$ *defined in Theorem 1 converge uniformly, as* $n \to \infty$, *to a function* $f(x)$ *satisfying the following functional equation*

$$f(x) = -cx + \min_{y \geqslant x}[H_1(y) + \alpha f(0)F(y) + \alpha \int_y^\infty f(y - s)\varphi(s)\, ds].$$

In order to establish existence of $f(x) = \lim_{n\to\infty} f_n(x)$ we consider $u_n(x) = f_{n+1}(x) - f_n(x)$. According to (1) and (2) one easily gets the inequality $\delta_n = \sup_x |u_n(x)| \leqslant d\alpha^{n-1}$ where $d = 3H_1(\bar{x}) + c(\bar{x} + \mathrm{E}\xi)$. This enables us to prove that $f_n(x)$ converge uniformly on the whole real line to the limit $f(x)$ and estimate the convergence rate. Hence, $f(x)$ satisfies the functional equation.

Next, suppose that there exist two different claim distributions F and G, and we would like to estimate the difference between the corresponding optimal cost functions. So, the index F (or G) will be added to all the functions arising in our study under assumption that claim distribution is F (resp. G). It is clear that $L^F(y) := L(y)$, $H_k^F(y) := H_k(y)$ and $f_k^F(x) := f_k(x)$, $k \geqslant 1$, since previously we had only one distribution F. Definition of analogous functions with index G is obvious, namely, in all the formulas we write G instead of F.

We need also to introduce some probability metrics. Thus, (see, e.g. [2]) Kantorovich (Wasserstein) metric is defined as

$$\kappa(F, G) = \int_{-\infty}^{\infty} |F(x) - G(x)|\, dx.$$

The main result concerning the system stability is as follows.

Theorem 3. *Let distributions F and G be such that $\kappa(F, G) < \varepsilon$ then, for any $n \geqslant 1$ and $\alpha \in (0, 1)$,*

$$\gamma_n = \sup_{-\infty < x < \infty} |f_n^F(x) - f_n^G(x)| < D\varepsilon$$

where $D = (\max(k, p) + \alpha c)(1 - \alpha)^{-1}$.

Turning to the case $\alpha = 1$ we note that minimal n-step costs tend to ∞, as $n \to \infty$, for any initial capital x and it is impossible to establish the estimate (not depending on n) for the difference between costs corresponding to claim distributions F and G. However we can employ another objective function, namely, long-run average costs per period.

Furthermore, we introduce the following

Definition. A policy $\hat{y}(x) = \{\hat{y}_n(x), n \geqslant 1\}$ is asymptotically optimal if

$$\lim_{n\to\infty} \frac{1}{n}\hat{f}_n(x) = \lim_{n\to\infty} \frac{1}{n} f_n(x)$$

where $\hat{f}_n(x)$ are the costs obtained by applying the policy $\hat{y}(x)$ and $f_n(x)$ are the minimal n-step costs.

Theorem 4. *The policy $\hat{y}(x)$ with $\hat{y}_n(x) = \max(\bar{x}, x)$ for all $n \geqslant 1$ is asymptotically optimal. Moreover,*

$$\lim_{n\to\infty} \frac{1}{n}\hat{f}_n(x) = D(\bar{x})$$

where $D(\bar{x}) = c\bar{x}F(\bar{x}) + c\int_{\bar{x}}^{\infty} s\,dF(s) + L(\bar{x})$ and $\bar{x}$ is given by (3).

Proof is carried out in two steps. *Step* 1. We establish that

$$\frac{1}{n}(\hat{f}_n(x) - f_n(x)) \to 0, \qquad \text{as } n \to \infty.$$

Step 2. We obtain the explicit form of $D(\bar{x})$.

Remark. It is possible to strengthen the result of Theorem 4 establishing that under the asymptotically optimal policy not only expected average costs tend to limit $D(\bar{x})$ but (random) average costs converge with probability one to the same limit. For this purpose one has to use Wald's identity (see, e.g. Wald [3]), the strong law of large numbers and other properties of renewal processes.

Other research directions

Now we briefly mention the other topics which will be treated in the presentation. First, we investigate the credit models with k-period validity ($k \geqslant 2$) and discrete-time insurance models introduces in Bulinskaya [4,5]. For these models we establish the stability with respect to small perturbations of underlying distributions and carry out the sensitivity analysis with respect to small fluctuations of parameters using the methods of Saltelli et al. [6], see also Bulinskaya [7].

Second, we deal with continuous-time insurance models involving reinsurance, dividends, investment and taxes. New optimization criterions such as Gerber-Shiu function and Parisian ruin, permitting bankruptcy implementation delays are used. Necessary definitions can be found, e.g., in Bulinskaya [8].

References

1. Bellman R. Dynamic programming. Princeton University Press, 1957.
2. Rachev S.T., Klebanov L.B., Stoyanov S.V., Fabozzi F.J. The methods of distances in the theory of probability and statistics. Springer, 2013.
3. Wald A. Some generalizations of the theory of cumulative sums of random variables // Annals of Mathematical Statistics. 1945. V. 16. P. 287–293.

4. Bulinskaya E. Stochastic Insurance Models: Their Optimality and Stability // Ch.H. Skiadas (ed.) Advances in Data Analysis. Boston, Basel, Berlin: Birkhäuser, 2010. P. 129–140.
5. Bulinskaya E., Gusak J. Optimal Control and Sensitivity Analysis for Two Risk Models // Communications in Statistics – Simulation and Computation. 2016, V. 45, N 5. P. 1451–1466.
6. Saltelli A., Ratto M., Campolongo T., Cariboni J., Gatelli D., Saisana M. and Tarantola S. Global Sensitivity Analysis. The Primer. Wiley. 2008.
7. Bulinskaya E.V. Sensitivity analysis of some applied models // Pliska Studia Mathematica Bulgarica. 2007, V. 18. P. 57–90.
8. Bulinskaya E. New research directions in modern actuarial sciences // Panov V. (ed.) Modern problems of stochastic analysis and statistics – - selected contributions in honor of Valentin Konakov. Springer Series in Mathematics and Statistics 208, 2017. P. 349–408.

Queueing systems in which customer requires a random number of servers*

S. Grishunina
Department of Probability, Faculty of Mathematics and Mechanics, Lomonosov Moscow State University; Moscow Institute of Electronics and Mathematics, National Research University Higher School of Economics, Moscow, Russia

In this paper we study the stability conditions of the system with m identical servers in which customers arrive according to a regenerative input flow $X(t)$ [1]. An arrived customer requires service from $i \leq m$ servers simultaneously with probability α_i $(1 \leq i \leq m)$. A customer who arrives when the queue is empty begins service immediately if the number of servers he requires is available. Otherwise, when a customer becomes first in a queue, servers are held as they free up and service begins immediately when the number he requires is available. If a customer arrives to the system when the queue is not empty he goes to the end of the queue. When service begins, each server's completion time is independent of all other servers. Thus servers assigned to the same customer may free up individually. We consider two models according to the service time of the customer on each server: in the first model S_1 service

*This research is supported by Russian Foundation for Basic Research grant 17-01-00468

time has an exponential distribution $B_1(x) = 1 - e^{-\mu x}$ with the mean $\frac{1}{\mu}$. In the second model service time is a constant τ. Customer j occupies ζ_j servers simultaneously for independent times $\overline{\eta}_j = (\eta_{j1}, \dots, \eta_{j\zeta_j})$ with distribution function $B_1(x)$ in the system S_1 and for the time τ in the system S_2. The sequence $\{\zeta_j\}_{j=1}^{\infty}$ consists of independent identically distributed (iid) random variables with given distribution

$$\alpha_j = \mathsf{P}(\zeta = j),\ j = 1, \dots, m,\ \sum_{j=1}^{m} \alpha_j = 1.$$

We introduce an auxiliary process $Z_i(t)(i = 1, 2)$ for the systems S_1 and S_2 that is the number of service completions by all m servers up to time t under the assumption that there are always customers for service. This means that $Z_i(t)$ is defined by the sequence of service times $\{\overrightarrow{\eta}_n\}_{n=1}^{\infty}$ and $\{\zeta_n\}_{n=1}^{\infty}$ and does not depend on the input flow $X(t)$.

Then we construct the control Markov Chains for the processes $Z_i(i = 1, 2)$. For the model S_1 we define the stochastic process $U_1(t)$ which is the number of occupied servers at time t. For the model S_2 the stochasic process $U_2(t)$ is the number of empty servers at time t. In the both cases we assume that there are always customers for service. In the system S_2 process $U_2(t)$ may change the state only in the moments $n\tau$ where $n \in \mathbb{N}$. So we introduce the process $\tilde{U}_2(n) = U_2(t), t \in [n\tau, (n+1)\tau)$.

For the both processes there is a non-periodic class $\mathsf{K}_i(i = 1, 2)$ of communicating states. Therefore there are limits

$$\lim_{t \to \infty} \mathsf{P}(U_i(t) = \overrightarrow{k}) = \mathsf{P}_i(\overrightarrow{k}) > 0 \text{ for } \overrightarrow{k} \in \mathsf{K}_i(i = 1, 2). \qquad (1)$$

We note that $Z_i(t)(i = 1, 2)$ is a regenerative flow and we may take the moments of hits $U_1(t)$ and $\tilde{U}_2(n)$ into some fixed state $\overrightarrow{k} \in \mathsf{K}_i(i = 1, 2)$ as points of regeneration. Since a regeneration period of $Z_i(t)$ has a finite mean there exist the limits

$$\lim_{t \to \infty} \frac{Z_i(t)}{t} = \lambda_{Z_i} \text{ w.p.1 } (i = 1, 2)$$

The rates $\lambda_{Z_i}(i = 1, 2)$ may be defined as follows

$$\begin{aligned} \lambda_{Z_1} &= \mu \sum_{j=1}^{m} j\mathsf{P}_1(j), \\ \lambda_{Z_2} &= \frac{1}{\tau} \sum_{j=0}^{m-1} (m-j)\mathsf{P}_2(j) \end{aligned} \tag{2}$$

Now we introduce the process

$$V(t) = \sum_{j=1}^{X(t)} \zeta_j$$

that is the necessary number of servers for customers arrived to the system up to time t. It is evident that the rate

$$\lambda_V = \lim_{t \to \infty} \frac{V(t)}{t} = \mathsf{E}\zeta\lambda_X \quad w.p.1.$$

We define the traffic rate ρ_i for the system $S_i(i = 1, 2)$. as follows

$$\rho_i = \frac{\lambda_V}{\lambda_{Z_i}} = \frac{\mathsf{E}\zeta\lambda_X}{\lambda_{Z_i}}(i = 1, 2). \tag{3}$$

Then we construct the common points of regeneration for the both processes $X(t)$ and $Z_i(t)$. For the first system S_1 $Z_1(t)$ is a strongly regenerative flow. In the second system S_2 we suppose that the input flow $X(t)$ is strongly regenerative. Strong regeneration means that the regeneration period of the process may be considered as the sum of two independent random phases where the first phase has an exponential distribution.

For the system S_1 common points or regeneration are those points of regeneration of the input flow which hit the exponential phase of regeneration period of $Z_1(t)$. For the system S_2 these points are those points of regeneration of $Z_2(t)$ which hit the exponential phase of regeneration period of $X(t)$. The periods of regeneration for these sequence of points of regeneration have finite means for both systems $S_i(i = 1, 2)$.

Now we may formulate the stability criterion:

Theorem 1 *Let $q_i(t)$ be the number of customers at the system $S_i(i = 1, 2)$ at instant t. Then $q_i(t)$ is a stable process if and only if*

$$\rho_i < 1.$$

The proof of the Theorem is based on the relations between the auxiliary process $Z(t)$ and the real process of the number of service completions up to the time t and results in [2,3].

For the system S_1 stability criterion is

$$\lambda_X < \mu(\sum_{j=1}^{m} \frac{1}{m-j+1} \mathsf{P}(\zeta_1 \geq j))^{-1}.$$

This criterion is the same as obtained by Gillent and Latouche [4] for a queueing system with a Poisson input flow and exponential distribution of the service time.

For the system S_2 we introduce the function $h(n) = \sum\limits_{k=1}^{n} \mathsf{P}(\zeta_1 + \ldots + \zeta_k = n), n \leq m$ which is the renewal probability for the renewal process for $\{\zeta_k\}_{k=1}^{\infty}$.

Then the transition probabilities for $\tilde{U}_2(n)$ are defined as follows

$$\mathsf{P}_{ji} = \sum_{s=j+1}^{m-i} \tilde{\alpha}_{js} h(m-s-i)\beta_i,$$

$$\mathsf{P}_{00} = \sum_{s=1}^{m} \alpha_s h(m-s) = h(m),$$

$$\mathsf{P}_{j0} = \sum_{s=j+1}^{m} \tilde{\alpha}_{js} h(m-s),$$

where $\beta_i = \sum\limits_{j=i+1}^{m} \alpha_j$, $\tilde{\alpha}_{js} = \frac{\alpha_s}{\sum\limits_{i=j+1}^{m} \alpha_i}$. The class of states K_2 for the process $\tilde{U}_2(n)$ is finite so it is always possible to find the limit distribution for the process $\tilde{U}_2(n)$ and, hence, λ_{Z_2} and stability criterion.

But in the case $m > 2$ servers it is quite difficult to estimate the limit distribution so here we consider only the case when the number of servers $m = 2$.

The stability criterion for the system S_2 for the case $m = 2$ is

$$\lambda_X = \frac{\lambda_{Z_2}}{\mathsf{E}\zeta} = \frac{1}{\tau} \frac{2 + \alpha - \alpha^2}{(2-\alpha)(1 + \alpha(1-\alpha))} < 1.$$

The stability criterion for the system S_1 for the case $m = 2$ is

$$\lambda_X = \frac{\lambda_{Z_1}}{\mathsf{E}\zeta} = \mu \frac{2}{3 - 2\alpha} < 1.$$

We consider that the means of the service times for the both systems are the same that is $\mu = \frac{1}{\tau}$.

Hence $\frac{\lambda_{Z_2}}{\lambda_{Z_1}} \geq 1$. Therefore, the constant service time is better than the exponential service time.

References

1. Afanasyeva L.G., Bashtova E.E. Coupling method for asymptotic analysis of queues with regenerative input and unreliable server// Queueing Systems. 2014. V. 76, 2. P. 125–147.
2. Afanasyeva L., Tkachenko A. Multichannel queueing systems with regenerative input flow// Theory of Probability and Its Applications. 2014. V. 58, 2. P. 174–192.
3. Thorisson H.: Coupling, Stationary and Regeneration. New York: Springer, 2000.
4. Gillent L., Latouche G. Semi-explicit solutions for $M|PH|1$-like queueing systems// European Journal of Operations Research. 1983. V. 13, 2. P. 151–160.

Asymptotic properties of service and control operations in tandem systems with cyclic algorithms with prolongation

V.M. Kocheganov and A.V. Zorine

N.I. Lobachevsky National Research State University of Nizhny Novgorod, Nizhny Novgorod, Russia

Operations research is a field of mathematics which deals with finding optimal way of *operating* on different objects. The nature of these operations might contain very complex stochastic grain. Once the objects are customers or demands and the operation is service and control, queuing theory methods can be used. Queuing theory approaches can be divided into two groups. The first group of methods is classical and assumes homogeneity of all customers as well as that of all operations on the customers. For the first time such assumptions were considered in the early XX century by F. Johanssen, A.K. Erlang, A.Ya. Khinchine, F. Pollaczek, C. Palm, D. Kendall. During the second half of the XX century this sort of customers and operations on them have been investigated by A.N. Kolmogorov, B.V. Gnedenko, T.D. Saati, Yu.V. Prokhorov, E.S. Ventsel et al.

However, there are a lot of cases where homogeneity assumptions do not hold: information flows processing in local-area computer networks and telecommunication networks, control of conflicting flows of aircrafts at passage of intersecting air lanes, conflicting transport flows control at intersections with complicated intersection geometry etc. The second group of queuing theory methods deals with nonhomogeneous customers and nonhomogeneous operations on these customers. This is done via solving the following fundamental problems: 1) classification of nonhomogeneous customers and description of flows of nonhomogeneous customers; 2) formation and development of control algorithms of conflicting flows of nonhomogeneous customers. In this paper we describe one of the problems that can be solved with second group of queuing theory methods.

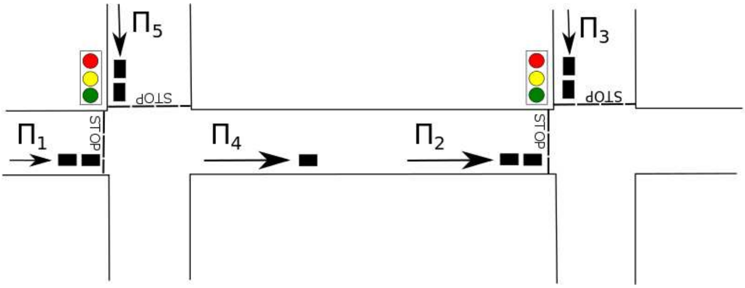


Fig. 1. A tandem of crossroads

Consider a real-life system of tandem of two consecutive crossroads (Fig. 1). The input flows are flows of vehicles. The flows Π_1 and Π_5 at the first crossroad are conflicting; Π_2 and Π_3 at the second crossroad are also conflicting. Every vehicle from the flow Π_1 after passing the first road intersection joints the flow Π_4 and enters the queue O_4. After some random time interval the vehicle arrives to the next road intersection. Such a pair of crossroads is an instance of a more general queuing model described below.

Consider a queuing system with four input flows of customers Π_1, Π_2, Π_3, and Π_4 entering the single server queueing system (Π_5 flow has no effect on system behaviour and is omitted in remaining discussion). Customers in the input flow Π_j, $j \in \{1, 2, 3, 4\}$ join a queue O_j with an unlimited capacity. For $j \in \{1, 2, 3\}$ the discipline of the queue O_j is FIFO

(First In First Out). Discipline of the queue O_4 will be described later. The input flows Π_1 and Π_3 are generated by an external environment, which has only one state. Each of these flows is a nonordinary Poisson flow. Denote by λ_1 and λ_3 the intensities of bulk arrivals for the flows Π_1 and Π_3 respectively. The probability generating function of number of customers in a bulk in the flow Π_j is $f_j(z) = \sum_{\nu=1}^{\infty} p_\nu^{(j)} z^\nu, \quad j \in \{1, 3\}$. We assume that $f_j(z)$ converges for any $z \in \mathbb{C}$ such that $|z| < (1 + \varepsilon)$, $\varepsilon > 0$. Here $p_\nu^{(j)}$ is the probability of a bulk size in flow Π_j being exactly $\nu = 1, 2, \ldots$. Having been serviced the customers from O_1 come back to the system as the Π_4 customers. The Π_4 customers in turn after service enter the system as the Π_2 ones. The flows Π_2 and Π_3 are conflicting in the sense that their customers can't be serviced simultaneously. This implies that the problem can't be reduced to a problem with fewer input flows by merging the flows together.

In order to describe the server behavior positive integers d, n_0, $n_1, \ldots$, n_d are fixed and a finite set $\Gamma = \{\Gamma^{(k,r)} \colon k = 0, 1, \ldots, d; r = 1, 2, \ldots n_k\}$ of states server can reside in is introduced. At the state $\Gamma^{(k,r)}$ the server stays during a constant time $T^{(k,r)}$. We will assume, that for each fixed k^* cycle subset $\{\Gamma^{(k^*,r)} \colon r = 1, 2, \ldots n_k^*\} = C_{k^*}^{\mathrm{N}} \cup C_{k^*}^{\mathrm{O}} \cup C_{k^*}^{\mathrm{I}}$, that is consists of three disjoint sets called neutral, output and input sets of states. In more details server is described in [1].

In general, service durations of different customers can be dependent and may have different laws of probability distributions. So, saturation flows will be used to define the service process. The saturation flow Π_j^{sat}, $j \in \{1, 2, 3, 4\}$, is defined as a virtual output flow under the maximum usage of the server and unlimited number of customer in the queue O_j. The saturation flow Π_j^{sat}, $j \in \{1, 2, 3\}$ contains a non-random number $\ell(k, r, j) \geqslant 0$ of customers in the server state $\Gamma^{(k,r)}$.

The queuing system under investigation can be regarded as a cybernetic control system, it helps to rigorously construct a formal stochastic model [2]. There are following blocks present in the system: 1) the external environment with one state; 2) input poles of the first type — the input flows Π_1, Π_2, Π_3, and Π_4; 3) input poles of the second type — the saturation flows Π_1^{sat}, Π_2^{sat}, Π_3^{sat}, and Π_4^{sat}; 4) an external memory — the queues O_1, O_2, O_3, and O_4; 5) an information processing device for the external memory — the queue discipline units δ_1, δ_2, δ_3, and δ_4; 6) an internal memory — the server (OY); 7) an information processing device for internal memory — the graph of server state transitions; 8) output poles — the output flows Π_1^{out}, Π_2^{out}, Π_3^{out}, and Π_4^{out}.

Let us introduce the following variables and elements along with their value ranges. To fix a discrete time scale consider the epochs $\tau_0 = 0$, τ_1, τ_2, $\ldots$ when the server changes its state. Let $\Gamma_i \in \Gamma$ be the server state during the interval $(\tau_{i-1}; \tau_i]$, $\varkappa_{j,i} \in \mathbb{Z}_+$ be the number of customers in the queue O_j at the instant τ_i, $\eta_{j,i} \in \mathbb{Z}_+$ be the number of customers arrived into the queue O_j from the flow Π_j during the interval $(\tau_i; \tau_{i+1}]$, $\xi_{j,i} \in \mathbb{Z}_+$ be the number of customers in the saturation flow Π_j^{sat} during the interval $(\tau_i; \tau_{i+1}]$, $\overline{\xi}_{j,i} \in \mathbb{Z}_+$ be the actual number of serviced customers from the queue O_j during the interval $(\tau_i; \tau_{i+1}]$, $j \in \{1, 2, 3, 4\}$. The server changes its state according to the following rule: $\Gamma_{i+1} = h(\Gamma_i, \varkappa_{3,i})$ where the mapping $h(\cdot, \cdot)$ is defined in paper [1]. Lets define function $\psi(\cdot, \cdot, \cdot)$: $\psi(k; y, u) = C_y^k u^k (1-u)^{y-k}$. $\psi(k; y, u)$ is probability of arrival of k Π_2-customers given O_4 has y customers and server is in state $\Gamma^{(k,r)}$, that is $u = p_{k,r}$. If $0 \leqslant k \leqslant y$ does no hold we put $\psi(k; y, u)$ equal 0. Mathematical model in more details can be found in work [3].

We now present several results regarding asymptotic behaviour of described system. Consider stochastic sequences:

$$\{(\Gamma_i(\omega), \varkappa_{3,i}(\omega)); i = 0, 1, \ldots\}, \tag{1}$$

$$\{(\Gamma_i(\omega), \varkappa_{1,i}(\omega), \varkappa_{3,i}(\omega)); i = 0, 1, \ldots\}, \tag{2}$$

$$\{(\Gamma_i(\omega), \varkappa_{1,i}(\omega), \varkappa_{3,i}(\omega), \varkappa_{4,i}(\omega)); i = 0, 1, \ldots\}, \tag{3}$$

which include number of customers $\varkappa_{1,i}(\omega)$, $\varkappa_{3,i}(\omega)$ and $\varkappa_{4,i}(\omega)$ in the queues O_1, O_3 and O_4 respectfully.

Theorem 1. Let $\Gamma_0 = \Gamma^{(k,r)} \in \Gamma$ and $\varkappa_{3,0} = x_{3,0} \in \mathbb{Z}_+$ be fixed. Then the sequence (1) is Markov chain.

Theorem 2. Let $\Gamma_0 = \Gamma^{(k,r)} \in \Gamma$ and $(\varkappa_{1,0}, \varkappa_{3,0}) = (x_{1,0}, x_{3,0}) \in \mathbb{Z}_+^2$ be fixed. Then the sequence (2) is Markov chain.

Theorem 3. For Markov chain (1) to have a stationary distribution it is sufficient to satisfy the following inequalitiy

$$\min_{k=\overline{1,d}} \frac{\sum_{r=1}^{n_k} \ell(k,r,3)}{\lambda_3 f_3'(1) \sum_{r=1}^{n_k} T^{(k,r)}} > 1.$$

Theorem 4. For Markov chain (2) to have a stationary distribution it is sufficient to satisfy the following inequalities

$$\min_{k=\overline{0,d}} \frac{\sum_{r=1}^{n_k} \ell(k,r,1)}{\lambda_1 f_1'(1) \sum_{r=1}^{n_k} T^{(k,r)}} > 1, \quad \min_{k=\overline{1,d}} \frac{\sum_{r=1}^{n_k} \ell(k,r,3)}{\lambda_3 f_3'(1) \sum_{r=1}^{n_k} T^{(k,r)}} > 1.$$

Theorem 1 and Theorem 3 concern the low-priority queue which is also described in [1,3,4]. Theorem 2 and Theorem 4 concern primary input flow queues which are referenced in [5].

Theorem 5. Assume assumptions of Theorem 4 and the inequality

$$\min_{k=\overline{0,d},r=\overline{1,n_k}} \{p_{k,r}\} > 0.$$

Then the stochastic sequence (3) is bounded.

Proof Put $A_i(w_1, w_3, w_4, \gamma) = \{\omega\colon \varkappa_{1,i} = w_1, \varkappa_{3,i} = w_3, \varkappa_{4,i} = w_4, \Gamma_i = \gamma\}$. Let $(\gamma, x_3) \in \Gamma \times \mathbb{Z}_+$ and $\Gamma^{(\tilde{k},\tilde{r})} = h(\gamma, x_3)$. Since $\varkappa_{4,i+1} = \varkappa_{4,i} + \min\{\xi_{1,i}, \varkappa_{1,i} + \eta_{1,i}\} - \eta_{2,i}$, one has:

$$\begin{aligned} E[\varkappa_{4,i+1}|A_i(w_1, w_3, w_4, \gamma)] &= \\ = E[w_4 - \eta_{2,i} + \min\{\xi_{1,i}, w_1 + \eta_{1,i}\}|A_i(w_1, w_3, w_4, \gamma)] &\leqslant \\ \leqslant E[w_4 - \eta_{2,i} + \xi_{1,i}|A_i(w_1, w_3, w_4, \gamma)] &= \\ = w_4 + \ell(\tilde{k}, \tilde{r}, 1) - E[\eta_{2,i}|A_i(w_1, w_3, w_4, \gamma)]&. \end{aligned}$$

From the definition of $\psi(\cdot;\cdot,\cdot)$ we get $E[\eta_{2,i}|A_i(w_1, w_3, w_4, \gamma)] = w_4 p_{\tilde{k},\tilde{r}}$. Hence it is true that $E[\varkappa_{4,i+1}|A_i(w_1, w_3, w_4, \gamma)] \leqslant w_4(1-p_{\tilde{k},\tilde{r}})+\ell(\tilde{k}, \tilde{r}, 1)$. Using the law of total expectation one gets:

$$\begin{aligned} E[\varkappa_{4,i+1}] &= \sum_{w_1=0}^{\infty} \sum_{w_3=0}^{\infty} \sum_{w_4=0}^{\infty} \sum_{\gamma\in\Gamma} E[\varkappa_{4,i+1}|A_i(w_1, w_3, w_4, \gamma)]\times \\ &\times \mathbf{P}(A_i(w_1, w_3, w_4, \gamma)) \leqslant \sum_{w_3=0}^{\infty} \sum_{w_4=0}^{\infty} \sum_{\gamma\in\Gamma} (w_4(1 - p_{\tilde{k},\tilde{r}}) + \ell(\tilde{k}, \tilde{r}, 1))\times \\ &\times \mathbf{P}(A_i(w_1, w_3, w_4, \gamma)) \leqslant (1 - \min\{p_{\tilde{k},\tilde{r}}\})\times \\ &\times \sum_{w_4=0}^{\infty} w_4 \mathbf{P}(\varkappa_{4,i} = w_4) + \max\{\ell(\tilde{k}, \tilde{r}, 1)\}) \sum_{w_4=0}^{\infty} \mathbf{P}(\varkappa_{4,i} = w_4) = \\ &= (1 - \min\{p_{\tilde{k},\tilde{r}}\}) E[\varkappa_{4,i}] + \max\{\ell(\tilde{k}, \tilde{r}, 1)\}). \end{aligned}$$

The sequence

$$M_0 = E[\varkappa_{4,0}], \qquad M_{i+1} = (1 - \min\{p_{\tilde{k},\tilde{r}}\}) M_i + \max\{\ell(\tilde{k}, \tilde{r}, 1)\})$$

dominates the sequence $E[\varkappa_{4,i+1}]$ and under the theorem assumptions is limited. It implies O_4 queue size $\varkappa_{4,i}$ to be limited. Since conditions of the Theorem 4 are satisfied then Markov chain (2) has stationary distribution and queue sizes $\varkappa_{1,i}$ and $\varkappa_{3,i}$ are limited as well. QED.

References

1. Kocheganov V.M., Zorine A.V. Low-Priority Queue and Server's Steady-State Existence in a Tandem Under Prolongable Cyclic Service // DCCN. Ser. Communications in Computer and Information Science. Springer, Cham. 2016. V. 678. P. 210–221.
2. Zorin A.V. Stability of a tandem of queueing systems with Bernoulli noninstantaneous transfer of customers // Theory of Probability and Mathematical Statistics. 2012. V. 84. P. 173–188.
3. Kocheganov V.M., Zorine A.V. Low-Priority Queue Fluctuations in Tandem of Queuing Systems Under Cyclic Control with Prolongations // DCCN. Ser. Communications in Computer and Information Science. 2016. V. 601. P. 268–279.
4. Kocheganov V.M., Zorine A.V. Sufficient condition of low-priority queue stationary distribution existence in a tandem of queuing systems // Bulletin of the Volga State Academy of Water Transport, 2017, vol. 50, pp. 47–55. (In Russian)
5. Kocheganov V., Zorine A. Primary input flows in a tandem under prolongable cyclic service // DCCN. 2017. Pp. 517–525.

On the ruin probability of a joint-stock insurance company in the Sparre Andersen model

A.A. Muromskaya

Lomonosov Moscow State University, Moscow, Russia

Let us consider the Sparre Andersen risk model. According to this model, the surplus of an insurance company at time t is as follows:

$$X_t = x + ct - \sum_{i=1}^{N_t} X_i,\ t \geqslant 0, \qquad (1)$$

where x is the initial capital of the company, premiums are acquired continuously at the rate c, N_t is a renewal process. The claim amounts $\{X_i\}$ are non-degenerate i.i.d. random variables with distribution function $F(y)$ such that $F(0) = 0$. In its turn, the function $G(y)$ is a distribution function of the intervals between the claim times $\{T_j\}$. In addition, the variables $\{X_i\}$ and the process N_t are supposed to be independent too.

One of the important results related to the Sparre Andersen risk model is the determination of an upper bound for the ruin probability of the company. Let $\tau = \inf[t \geqslant 0 : \ X_t < 0]$ be the time of ruin of the insurance company whose surplus at time t has the form (1). Then $\psi(x) = P(\tau < \infty | X_0 = x)$ is the probability of ruin. We also assume that there exists a unique positive root R of the equation

$$\int_0^\infty e^{ry} dF(y) \int_0^\infty e^{-crt} dG(t) = 1 \tag{2}$$

called the adjustment coefficient or Lundberg's exponent. The following estimate for the ruin probability is valid in this case:

$$\psi(x) \leqslant e^{-Rx}. \tag{3}$$

This inequality is an analogue of the famous Lundberg's inequality proved for the particular case of the Sparre Andersen risk model, namely, the Cramer-Lundberg model, according to which N_t is a Poisson process [1].

However, it is worth noting that the classical Sparre Andersen risk model does not take into account the fact that the insurance company can be a joint-stock company, while dividend payment is of great importance in calculating the probability of ruin. The surplus of the insurance company paying dividends is defined as $U_t = X_t - D_t$, where D_t denotes the amount paid to shareholders as dividends by the time t. Dividends are paid by the company in accordance with a barrier strategy with a barrier level b_t, namely, no dividends are paid whenever $U_t < b_t$, and dividends are paid out to the shareholders with intensity $c - db_t$ whenever $U_t = b_t$. If $U_t > b_t$, then the total amount $U_t - b_t$ is immediately paid as dividends (note that the inequality $U_t > b_t$ can hold only for $t = 0$). The random variable $T = \inf[t \geqslant 0 : \ U_t < 0]$ is then the ruin time of the joint-stock insurance company, and $\psi^{div}(x) = P(T < \infty | U_0 = x)$ is its probability of ruin. Thus, of special interest are studies that focus on the research of the function $\psi^{div}(x)$. A key role in the research is played by the choice of the barrier function b_t and the choice of the distribution of the interclaim times. In the case when this distribution is exponential, a number of interesting results have been obtained for various functions b_t (many of them are mentioned in the review article [2]). It is known, for example, that if the barrier b_t does not change over time, then the probability of the company's ruin will be equal to 1. However, very few papers ([3] and [4]) are devoted to the problems of searching and estimating $\psi^{div}(x)$ in the case when the distribution of the intervals between the claim times $\{T_j\}$ is more general than the exponential one, and the barrier b_t is not constant. In the article [3] we can find estimates of the ruin probability $\psi^{div}(x)$ for all distribution functions $F(y)$ and $G(y)$ for which the adjustment coefficient R exists, and for a barrier strategy with a step barrier function, according to which the barrier level b_t is equal to b_i on half-intervals of the form $[T_{i-1}, T_i)$, $i \geqslant 1$ ($T_0 = 0$). In [4] it is assumed that the interclaim times follow a generalized Erlang(n) distribution i.e. these random variables can be represented as a convolution of n independent exponentially distributed random variables with parameters $\lambda_1, \dots, \lambda_n$. A linear barrier strategy with barrier level $b_t = b + at$, where $b \geqslant 0$, $0 < a < c$, is chosen as a dividend strategy. Under the given conditions of the model, integro-differential equations for the survival probability $\delta^{div}(x, b) = 1 - \psi^{div}(x, b)$ are derived in [4], however, it turns out to be very difficult to solve the equations explicitly and the authors of [4] give only a detailed solution algorithm for the case $n = 2$ and the exponential distribution of claims $\{X_i\}$.

Thus, the ruin probability in models with barrier dividend strategies with non-constant (and, in particular, linear) barrier functions b_t has not been fully explored to date. In this paper, we obtain an upper bound for $\psi^{div}(x, b)$ provided that the insurance company uses a linear barrier strategy and the intervals between the claim times $\{T_j\}$ have a gamma distribution with the density function $g(y) = \frac{\lambda^\alpha}{\Gamma(\alpha)} y^{\alpha-1} e^{-\lambda y}$, $y \geqslant 0$, $\alpha > 1$, $\lambda > 0$. Since $\psi^{div}(x, b) = \psi^{div}(b, b)$ for $b < x$, we assume without loss of generality that the initial capital x does not exceed b. We also continue to suppose that there exists an adjustment coefficient R, which is a positive root of the equation (2). In addition to the coefficient R, we also need the coefficient Q, which is defined in the following lemma.

Lemma 1. *There exists a unique root $Q > 0$ of the equation*

$$\int_0^\infty e^{-qy} dF(y) \int_0^\infty e^{-t(Ra+qa-qc)} dG(t) = 1,$$

moreover $\frac{Ra}{c-a} < Q < \frac{Ra+\lambda}{c-a}$.

Let us now turn to the main theorem of the paper.

Theorem 1. *The following inequality holds for the ruin probability $\psi^{div}(x, b)$ of a joint-stock insurance company using a linear barrier divi-*

dend strategy in the case of a gamma distribution of the interclaim times:

$$\psi^{div}(x,b) \leqslant e^{-Rx} + Ke^{-(R+Q)b}e^{Qx},$$

where

$$K = \frac{2^{\alpha-1}(\lambda+Rc)^{\alpha-1}R}{(\lambda+Ra+Qa-Qc)^{\alpha-1}Q} + \frac{2^{\alpha-1}\left((\lambda+Rc)^{\alpha}-(\lambda+Ra)^{\alpha}\right)}{(\lambda+Ra)^{\alpha}-(\lambda+Ra+Qa-Qc)^{\alpha}}.$$

Theorem 1 implies inequality (3) in the case of absence of dividend payments.

References

1. Kalashnikov V.V., Konstantinidis D.G. Ruin probability // Fundamentalnaya i prikladnaya matematika. 1996. V. 2, N 4. P. 1055–1100.
2. Avanzi B. Strategies for dividend distribution: a review // North American Actuarial Journal. 2009. V. 13, N 2. P. 217–251.
3. Muromskaya A.A. Generalization of Lundberg's inequality for the case of stock insurance company // Moscow University Mathematics Bulletin. 2017. V. 72, N 1. P. 31–34.
4. Albrecher H., Hartinger J., Thonhauser S. On exact solutions for dividend strategies of threshold and linear barrier type in a Sparre Andersen model // ASTIN Bulletin. 2007. V. 37, N 2. P. 203–233.

On asymptotic insensitivity of reliability systems*

V.V. Rykov[1,2] and D.V. Kozyrev[2,3]

[1] *Gubkin Russian State University of Oil and Gas,*
[2] *Peoples' Friendship University of Russia (RUDN University),*
[3] *V.A.Trapeznikov Institute of Control Sciences of Russian Academy of Sciences, Moscow, Russia*

In this talk we give a short review of classic results about strong and asymptotic insensitivity of stochastic systems to input distributions of their elements as well as some latest investigations on the topic.

Introduction and Motivation

*This investigation has been supported with the RFBR grants No.17-07-00142 and No. 17-01-00633.

Investigation of stability issues of output characteristics of different systems to the changes in initial data, parameters or outside factors is a key problem of all natural sciences. For stochastic systems stability often means insensitivity or weak sensitivity of their output characteristics to the shapes of their elements' input distributions.

Some investigations by B.A. Sevast'anov (1957), I.N. Kovalenko (1975), B.V. Gnedenko and A.D. Solov'ev (1964—1970) have been devoted to the strong and asymptotic insensitivity of stochastic models. Recently some new results about asymptotic insensitivity of systems' reliability characteristics have been found by authors with colleagues. They have been represented at several conferences, partially published without proofs in their proceedings (for the review see for example [1]). In current talk a review of these investigations and some their generalisation is proposed.

Problem setting and the notations

Consider a heterogeneous hot double redundant repairable system. Life times of its components are supposed to be independent identically distributed (i.i.d) random variables (r.v.) A_i $(i = 1, 2)$ with exponential distribution with parameters α_i $(i = 1, 2)$. The repair times of components B_i $(i = 1, 2)$ are also supposed to be i.i.d. r.v.'s with absolute continuous cumulative distribution functions (c.d.f.) $B_i(x)$ $(i = 1, 2)$ and probability density functions (p.d.f.) $b_i(x)$ $(i = 1, 2)$ correspondingly. Denote by $\tilde{b}(s) = \int_0^\infty e^{-sx} b(x)dx$ the moment generation function (m.g.f.) of r.v.'s B_i (its p.d.f Laplace Transform — LT).

Denote by $E = \{0,\ 1,\ 2,\ 3\}$ the system set of states, which means: 0—both components are working, j $(j = i, 2)$—the j-th component is under repair, and the other one is working, and 3 —both components are in down states, system has failed. Denote also by $J = \{J(t\ t \geqslant 0)\}$ the process that takes value $J(t) = j$, if at time t the system is in the state $j \in E$.

The main reliability characteristics of the system are its *reliability function*:

$$R(t) = \mathbf{P}\{T > t\}, \tag{1}$$

where $T = \inf\{t :\ J(t) = 3\}$ means the system lifetime, and the *steady state probabilities* (s.s.p)

$$\pi_j = \lim_{t\to\infty} \mathbf{P}\{J(t) = j\}. \tag{2}$$

However there are no systems that exist for infinitely long time, so the most interesting characteristics of the system reliability are so called *quasi-stationary probabilities* (q.s.p.), which are the probabilities of a system to be in any state before its failure,

$$\hat{\pi}_i = \lim_{t\to\infty} \mathbf{P}\{J(t) = i|t \leqslant T\} = \lim_{t\to\infty} \frac{\pi_i(t)}{R(t)}. \tag{3}$$

To investigate the system's reliability we introduce two-dimensional Markov process $Z = \{Z(t) = (J(t), X(t)),\ t \geqslant 0\}$, where $J(t)$ is the system state, and an additional variable $X(t)$ means the time, spent by the process in the state $J(t)$ after its last occurrence in it before time t. At that the process state space will have the following form: $\mathcal{E} = \{0, (1, x), (2, x), (3, x)\}$, which meaning is evident. The process state probabilities are denoted by $\pi_0(t)$, $\pi_1(t; x)$, $\pi_2(t; x)$, $\pi_3(t; x)$, and corresponding limiting s.s.p., which existence is provided by the fact that the process Z has a positive atom, by $\pi_0 = \lim_{t\to\infty} \pi_0(t)$, $\pi_i(x) = \lim_{t\to\infty} \pi_i(t; x),\ (i = 1, 2, 3)$.

Reliability function

To calculate the system-level reliability the state 3 should be considered as an absorbing one. By using Kolmogorov's forward system of partial differential equations for the time dependent process probabilities the following theorem can be proved.

Theorem 1. LT $\tilde{\pi}_(s)$ $(i \in \{0, 1, 2, 3\})$ and $\tilde{R}(s)$ of the state probabilities $\pi_i(t)$ $(i \in \{0, 1, 2, 3\}$ and the reliability function $R(t)$ of the system are:

$$\begin{aligned}
\tilde{\pi}_0(s) &= \frac{1}{s + \psi(s)}, \\
\tilde{\pi}_i(s) &= \frac{\phi_i(s)}{(s + \alpha_i*)(s + \psi(s))},\ \ (i = 1, 2) \\
\tilde{\pi}_3(s) &= \left[\frac{\alpha_1}{s + \alpha_1}\phi_2(s) + \frac{\alpha_2(s)}{s + \alpha_1}\phi_1(s)\right] (s(s + \psi(s)))^{-1}, \\
\tilde{R}(s) &= \frac{(s + \alpha_1)(s + \alpha_2) + (s + \alpha_1)\phi_1(s) + (s + \alpha_2)\phi_2(s)}{(s + \alpha_1)(s + \alpha_2)(s + \psi(s))},
\end{aligned} \tag{4}$$

where $i^* = 2$ for $i = 1$, $i^* = 1$ for $i = 2$ and the following notations are used

$$\phi_i(s) = \alpha_i(1 - \tilde{b}_i(s + \alpha_{i^*})),\ (i = 1, 2), \tag{5}$$

$$\psi(s) = \phi_1(s) + \phi_2(s). \tag{6}$$

As a consequence to the theorem the mean system lifetime can be found as

$$m = \mathbf{E}[T] = \tilde{R}(0) = \frac{\alpha_1\alpha_2 + \alpha_1(1-\tilde{b}_1(\alpha_2))\alpha_2(1-\tilde{b}_2(\alpha_1))}{\alpha_1\alpha_2[\alpha_1(1-\tilde{b}_1(\alpha_2)) + \alpha_2(1-\tilde{b}_2(\alpha_1))]}. \qquad (7)$$

The above result demonstrates the evident dependence of the reliability function on the components' repair time distributions. However under rare components' failures this dependence becomes negligible.

Theorem 2. The reliability function of the system scaled to the mean system lifetime $m = \mathbf{E}[T]$ for $q = \max\{\alpha_1, \alpha_2\} \to 0$ converges to the exponent

$$\lim_{q\to 0} \mathbf{P}\left\{\frac{T}{m} > t\right\} = e^{-t}.$$

Steady state probabilities

For the renewable system different rules for the system repair are possible. Consider one of them — the full system repair after its failure during some random time, say B_3 with c.d.f. $B_3(t)$, as a result of which the system becomes as a new one and moves to state 0.

Theorem 3. The s.s.p. of the system with full repair are:

$$\begin{aligned} \pi_1 &= \frac{\alpha_1}{\alpha_2}(1-\tilde{b}_1(\alpha_2))\pi_0, && (8)\\ \pi_2 &= \frac{\alpha_2}{\alpha_1}(1-\tilde{b}_2(\alpha_1))\pi_0, && (9)\\ \pi_3 &= [\alpha_1(1-\tilde{b}_1(\alpha_2)) + \alpha_2(1-\tilde{b}_1(\alpha_2))]b_3\pi_0, && (10) \end{aligned}$$

where $b_3 = \mathbf{E}[B_3] = \int_0^\infty (1 - B_3(x))dx$ and π_0 is:

$$\pi_0 = \left[1 + (1-\tilde{b}_1(\alpha_2))\left(\frac{\alpha_1}{\alpha_2} + \alpha_1 b_3\right) + (1-\tilde{b}_2(\alpha_1))\left(\frac{\alpha_2}{\alpha_1} + \alpha_2 b_3\right)\right]^{-1}, \qquad (11)$$

Quasi-stationary probabilities

When studying the system's behavior during its life cycle the most interesting characteristics are their q.s.p., which are defined by the formulas (3). Using results of the theorem 1 one can prove the following theorem.

Theorem 4. The q.s.p. of the model under consideration has the following form:

$$\begin{aligned}
\hat{\pi}_0 &= \left[1+\frac{\alpha_1}{\alpha_2-\gamma}(1-\tilde{b}_1(\alpha_2-\gamma))+\frac{\alpha_2}{\alpha_1-\gamma}(1-\tilde{b}_2(\alpha_1-\gamma))\right]^{-1}, \\
\hat{\pi}_1 &= \alpha_1\frac{(1-\tilde{b}_1(\alpha_2-\gamma))(\alpha_1-\gamma)}{(\alpha_1-\gamma)(\alpha_2-\gamma)+\alpha_1\phi_1(-\gamma)+\alpha_2\phi_2(-\gamma)}, \\
\hat{\pi}_2 &= \alpha_2\frac{(1-\tilde{b}_2(\alpha_1-\gamma))(\alpha_2-\gamma)}{(\alpha_1-\gamma)(\alpha_2-\gamma)+\alpha_1\phi_1(-\gamma)+\alpha_2\phi_2(-\gamma)}, \qquad (12)
\end{aligned}$$

where $-\gamma$ is the root of the equation $\psi(s)=-s$.

All the above results demonstrate an evident sensitivity of the system characteristics to the shape of its components' repair time distributions. However under rare failures of the components this sensitivity becomes negligible.

Sensitivity analysis

The s.s.p. of the system under the full repair scenario are given by the given above formulas (8). By using their Tailor expansion with $\max\{\alpha_1, \alpha_2\}\to 0$ the following theorem can be proved.

Theorem 5. The asymptotic behavior of the s.s.p. for the system with full repair under rare failures of its components are:

$$\begin{aligned}
\pi_0 &\approx [1+\rho_1(1+\alpha_2 b_3)+\rho_2(1+\alpha_1 b_3)]^{-1}, \\
\pi_1 &\approx \frac{\rho_1}{1+\rho_1(1+\alpha_2 b_3)+\rho_2(1+\alpha_1 b_3)}, \\
\pi_2 &\approx \frac{\rho_2}{1+\rho_1(1+\alpha_2 b_3)+\rho_2(1+\alpha_1 b_3)}, \\
\pi_3 &\approx \frac{(\rho_1\alpha_2+\rho_2\alpha_1)b_3}{1+\rho_1(1+\alpha_2 b_3)+\rho_2(1+\alpha_1 b_3)}. \qquad (13)
\end{aligned}$$

For the q.s.p.'s analogous result also can be obtained by Tailor expansion of formulas (12) in the neighbourhood of points $\alpha_i-\gamma$ when $q=\max\{\alpha_i,\ i=1,2\}\to 0$ taking into account that $\gamma<\min\{\alpha_1, \alpha_2\}$.

Theorem 6. Under rare failures of the considered system components its q.s.p.'s have the form:

$$\pi_0\approx(1+\rho_1+\rho_2)^{-1},\ \pi_1\approx\frac{\rho_1}{1+\rho_1+\rho_2},\ \pi_2\approx\frac{\rho_2}{1+\rho_1+\rho_2}. \qquad (14)$$

The results of these two last theorems show an asymptotic insensitivity of s.s.p.'s and q.s.p.'s of the system states to the shapes of the system components' repair times.

Conclusions and further investigations

The problem of sensitivity analysis of redundant systems' output reliability characteristics with exponentially distributed lifetimes of their components to the shape of their repair time distributions is considered. The generalization of the obtained results to the more complicated models with general distributions of the components lifetimes as well as to the systems with dependent components' failures are among the topics of our further investigations.

References

1. Rykov V.V., Kozyrev D.V. Analysis of renewable reliability systems by Markovization method // Analytical and Computational Methods in Probability Theory. ACMPT 2017. Lecture Notes in Computer Science, Volume 10684. Springer, Cham, Pp. 210 – 220, 2017. DOI: 10.1007/978-3-319-71504-9_19

Synergetic effects in multiserver loss systems*

G.Sh. Tsitsiashvili[1], M.A. Osipova[1],
K.E. Samouylov[2], and Yu.V. Gaidamaka[2]
[1]*Institute of Applied Mathematics,*
Far Eastern Branch of the Russian Academy of Sciences,
[1]*Far Eastern Federal University,*
[2]*Peoples' Friendship University of Russia (RUDN University),*
[2]*Federal Research Center "Computer Science and Control"*
of the Russian Academy of Sciences,
[1]*Vladivostok,* [2]*Moscow, Russia*

Introduction

The queuing theory as a modern area of applied probability theory was developed in the framework of operations research, being one of the main

*The publication has been prepared with the support of the "RUDN University Program 5-100" and funded by RFBR according to the research projects 17-07-00177, 18-07-00576.

tools for analysing the performance of telecommunication systems. The queuing theory plays the special role in the performance analysis of the future generations of telecommunication networks. For example, in 5th generation network, despite its high bandwidth it is necessary to share the finite amount of resources between different applications and users.

Queuing theory is currently developing precisely in this direction, having as its basis [1-6], while the essential role is played by the results obtained in the work on the mathematical tele traffic theory [7-10]. A feature of this work is an attempt to apply the asymptotic methods to the analysis of telecommunication networks [12-15]. Asymptotic behaviours of the blocking probability and parameters of the Equivalent Random Theory method was analysed in [13] for the case when both the number of servers and the load tend to infinity. It is also interesting to establish convergence for different types of input flow [15].

In this paper, we consider n - server loss system under the assumption that the intensity of the input flow is proportional to n. We investigate the convergence of the blocking probability in this system to zero at $n \to \infty$. A similar problem arises in the design of modern data transmission systems [7, Chapter 2]. A specific of suggested asymptotic results is that we did not obtain accuracy formulas or solutions of optimization problems for the transmission systems. Considered asymptotic formulas allow to establish convergence.

Asymptotic relations

Consider queuing system $A_n = M|M|n|0$ with intensity of input flow $n\lambda$ and intensities of service at all n servers μ, $\rho = \lambda/\mu$. System A_n is an aggregation of n systems $A_1 = M|M|1|0$. The number of customers in the system A_n is described by the process of death and birth $x_n(t)$ with birth and death rates $\lambda_n(k) = n\lambda,\ 0 \le k < n,\ \mu_n(k) = k\mu,\ 0 < k \le n$.

Denote $P_n(\rho)$ the stationary blocking probability in the system A_n at a given ρ. Let $a_n,\ b_n,\ n \ge 1$, be two real sequences. For $n \to \infty$ we assume that $a_n \succeq b_n$, if $\limsup\limits_{n\to\infty} \dfrac{a_n}{b_n} \ge 1$, let's say $a_n \sim b_n$, if $b_n \succeq a_n \succeq b_n$.

Theorem 1. *The following limit ratio is true:* $P_n(1) \sim \sqrt{\dfrac{2}{\pi n}},\ n \to \infty.$

Proof. Let $\varepsilon > 0$, consider the function $f(x) = 1 - x - \exp(-(1+\varepsilon)x)$. The $f(x)$ function satisfies the following relations: $f(0) = 0,\ f(1) < 0,$

$$f'(x) > 0,\ 0 < x < \frac{\ln(1+\varepsilon)}{1+\varepsilon},\ f'(x) < 0,\ \frac{\ln(1+\varepsilon)}{1+\varepsilon} < x \le 1.$$

Therefore, on the segment $[0,1]$ there exists a single $x(\varepsilon)$, satisfying the $f(x(\varepsilon))=0$ and means the inequalities $1-x\geq \exp(-(1+\varepsilon)x)$, $0\leq x\leq x(\varepsilon)<1$. Let $p_n(k)=\lim\limits_{t\to\infty}P(x_n(t)=k)$, $0\leqslant k\leqslant n$, then in force [16, Chapter 2, $\mathbb{S}$ 1]

$$p_n(n-1)=p_n(n)\frac{\mu}{\lambda}\,\frac{n}{n}\,,\;p_n(n-2)=p_n(n)\left(\frac{\mu}{\lambda}\right)^2\;\frac{n(n-1)}{n^2}\,,\ldots$$

Hence, the stationary blocking probability in virtue of the integral theorems of recovery and the law of large numbers for the recovery process [1, Chapter 9, $\mathbb{S}$ 4, 5] satisfies the equality

$$P_n(\rho)=p_n(n)=\left(\sum_{k=0}^{n}\rho^{-k}\prod_{j=0}^{k-1}\left(1-\frac{j}{n}\right)\right)^{-1}, \qquad (1)$$

where $\prod_{j=0}^{-1}$ equals 1. From Formula (1) we obtain the inequality

$$P_n^{-1}(1)\geq\sum_{0\leq k\leq nx(\varepsilon)}\prod_{j=0}^{k-1}\left(1-\frac{j}{n}\right)\geq\sum_{0\leq k\leq nx(\varepsilon)}\prod_{j=0}^{k-1}\exp(-(1+\varepsilon)j/n)\geq$$

$$\geq\sum_{1\leq k\leq nx(\varepsilon)}\exp(-(1+\varepsilon)k^2/2n).$$

This implies that

$$P_n^{-1}(1)\geq\int_1^{nx(\varepsilon)}e^{-(1+\varepsilon)x^2/2n}dx=\sqrt{\frac{n}{1+\varepsilon}}\int_{\sqrt{\frac{1+\varepsilon}{n}}}^{x(\varepsilon)\sqrt{n(1+\varepsilon)}}e^{-y^2/2}dy,$$

consequently

$$P_n(1)\sqrt{n}\leq(1+\varepsilon)\left(\int_{\sqrt{\frac{1+\varepsilon}{n}}}^{x(\varepsilon)\sqrt{n(1+\varepsilon)}}e^{-y^2/2}dy\right)^{-1}\to(1+\varepsilon)\sqrt{\frac{2}{\pi}},\;n\to\infty.$$

and so $\limsup\limits_{n\to\infty}P_n(1)\sqrt{\dfrac{\pi n}{2}}\leq 1+\varepsilon$.

Using Formula (1) and the inequality $1-x\log\exp(-x)$, $0\leq x\leq 1$, we obtain:

$$P_n^{-1}(1)\leq\sum_{1\leq k\leq n}e^{-k(k-1)/2n}\leq\sum_{1\leq k\leq n}e^{-(k-1)^2/2n}\leq\int_0^\infty e^{-x^2/2n}dx,$$

whence it follows that $1 \leq \liminf\limits_{n\to\infty} P_n(1)\sqrt{\frac{\pi n}{2}}$. Obtained above inequalities for upper and lower limits of leads to the statement of Theorem 1.

Analogously it is possible to obtain the following statements.

Theorem 2. *At $\rho < 1$ following relations are valid*

$$e^{-n\ln^2\rho/2}\sqrt{\frac{2}{\pi n}}\sqrt{\frac{\rho}{8}} \preceq P_n(\rho) \preceq (e^{-n\ln^2\rho/2})^{(\rho-1)/\ln\rho}\sqrt{\frac{2}{\pi n}}\sqrt{\frac{\ln\rho}{\rho-1}}. \quad (2)$$

Remark. For $\rho = 1-\gamma$, $\gamma \to 0$, the upper and lower bounds for the blocking probability $P_n(\rho)$ are close because the multiplier $\frac{\rho-1}{\ln\rho} \to 1$. Moreover, the multiplier $n\ln^2\rho$ most strongly affects the blocking probability $P_n(\rho)$.

Corollary. *Let $\rho = 1-n^{-\gamma}$, $\gamma > 0$, then Formulas (2) will be rewritten:*

$$\frac{1}{2}\sqrt{\frac{1}{\pi n}} \preceq P_n(\rho) \preceq \sqrt{\frac{2}{\pi n}},\ \gamma \geq \frac{1}{2},$$

$$\frac{1}{2}\sqrt{\frac{1}{\pi n}} \preceq P_n(\rho)\exp\left(\frac{n^{1-2\gamma}}{2}\right) \preceq \sqrt{\frac{2}{\pi n}},\ \gamma < \frac{1}{2}.$$

Theorem 3. *At $\rho > 1$, the following limit ratio is true*

$$P_n(\rho) \to 1-\mu/\lambda,\ n\to\infty.$$

Unification of multiserver loss systems

Suppose that we have m independent Poisson flows of customers with intensities $\lambda = \lambda_1 = \ldots = \lambda_m$ and parallel servers with the intensity of service at each of them equal to μ. We assume that the service of the k-th flow customer is realized on c_k servers, $1 \leq k \leq m$, and want to distribute the servers between the flows so that the blocking probabilities $P_n^{(k)}(1)$ for each of the flow $k = 1,\ldots,m$ are about the same. This problem arises in the design of modern communication systems.

Let the number of servers in the k-th subsystem be nn_k, based on Theorem 1, require that the basic equations $n_1/c_1 = \ldots = n_m/c_m$ are fulfilled. We rewrite these equations in the form

$$n_2 = n_1\cdot c_2/c_1,\ldots,n_m = n_1\cdot c_m/c_1.$$

Assume that the numbers $c_2/c_1, \ldots, c_m/c_1$ are rational and rewrite them as $c_2/c_1 = p_2/q_2, \ldots, c_m/c_1 = p_m/q_m$, where pairs of positive integers $p_2, ; q_2; \ldots ; p_m, q_m$ consist of mutually prime numbers. Then for the numbers $n_2, \ldots, n_m$ to be integers, it requires that number n_1 is a multiple of $q_2, \ldots, q_m$. Therefore the number n_1 should be divided by the smallest common multiple of l of the numbers $q_2, \ldots, q_m$. Thus, all possible values of the numbers $n_1, \ldots, n_m$, satisfying the basic equality, look like this:

$$n_1 = nL,\ n_2 = np_2L/q_2, \ldots, n_2 = np_mL/q_m,\ n = 1, \ldots$$

Let us consider now the case when the intensity of the input flows $\lambda_1, \ldots, \lambda_m$ differ and denote $\rho_k = \lambda_k/\mu,\ k = 1, \ldots, m$. In this case, it is natural to replace the base equality by the equality

$$(\ln^2 \rho_1) \cdot n_1/c_1 = \ldots = (\ln^2 \rho_m) \cdot n_m/c_m$$

and hold similar reviews.

References

1. Borovkov A.A. Probability theory. Springer. Universitext. 2013.
2. Borovkov A.A. Asymptotic Methods in Queuing Theory. Chichester: Wiley, 1984.
3. Gnedenko B.V., Korolev V.Yu. Random Summation: Limit Theorems and Applications. Boca Raton: CRC Press, 1996.
4. Afanasyeva L.G., Bulinskaya E.V. Certain Asymptotic Results for Random Walks in a Strip. Society for Industrial and Applied Mathematics (United States): Theory of Probability and its Applications. 1985. Volume 29, issue 4. P. 677–693.
5. Afanasyeva L.G., Bashtova E.E., Bulinskaya E.V. Limit Theorems for Semi Markov Queues and Their Applications. Communications in Statistics: Simulation and Computation. 2012. Volume 41, issue 6. P. 688–709.
6. Yarovaya E.B. Branching Random Walks with Several Sources. Taylor Francis (United Kingdom): Mathematical Population Studies. 2012. Volume 20. P. 14–26.
7. Basharin G.P. Lectures on mathematical theory of tele traffic. Moscow: Russian Peoples Friendship University, 2009. (In Russian).
8. Kelly F. Blocking Probabilities in Large Circuit-Switched Networks. Advances in Applied Probability. 1986. Volume 18. P. 473–505.

9. Ross K. Multiservice Loss Models for Broadband Telecommunication Networks. London: Springer. 1995.
10. Vishnevsky V.M., Semenova O.V. Polling Systems: Theory and Applications for Broadband Wireless Networks. London: Academic Publishing. 2012.
11. Whitt W. Stochastic-Process Limits: An Introduction to Stochastic-Process Limits and Their Application to Queues. New York: Springer Science and Business Media. 2002.
12. Naumov V.A. On the behavior of the parameters of the Equivalent Random Theory method at low load // Numerical methods and informatics. Moscow: UDN Publisher. 1988. (In Russian).
13. Tsitsiashvili G.Sh. Quantitative evaluation of decomposition effects of complex systems // Advances in Modelling and Analysis. 1995. Volume 47, issue 1. P. 27–30.
14. Tsitsiashvili G.Sh. Cooperative Effects in Multi-Server Queueing Systems // Mathematical Scientist. 2005. Volume 30, issue 1. P. 17–24.
15. Tsitsiashvili G.Sh. Osipova M.A. Synergetic effects for number of busy servers in multiserver queuing systems // Communications in Computer and Information Science. 2015. Volume 564. P. 404–414.
16. Ivchenko G.I., Kashtanov V.A., Kovalenko I.N. Queuing theory. Moscow: Vishaya Shkola, 1982. (In Russian).

Branching walks with a finite set of branching sources and pseudo-sources[*]

E. Yarovaya
Lomonosov Moscow State University, Moscow, Russia

We present results for continuous-time *branching random walks* on the lattice $\mathbf{Z}^d$, $d \geq 1$, with a finite number of particle generation centers called *branching sources*. The goal of the study is to analyze phase transitions for a branching random walk with different-type branching sources without any assumptions on a variance of jumps for the underlying random walk.

Consider particles living on $\mathbf{Z}^d$ independently of each other and of their history. Each particle walks on the lattice $\mathbf{Z}^d$ until it reaches a source where its behavior changed. Branching sources are of three types,

[*] Supported by the Russian Foundation for Basic Research (RFBR), project No. 17-01-00468.

depending on whether branching or violation of symmetry of the walk takes place or not. In sources of the first type, particles die or are born with keeping the random walk symmetry, see, e.g., [1-3]. In sources of the second type, walk symmetry is violated by increasing the degree of dominance of branching or walk, see, e.g., [4]. Sources of the third type should be called "pseudo-sources," because in them only the walk symmetry, without birth or death of particles, is violated. BRWs with r sources of the first type, k of the second type, and m of the third type are denoted BRW/r/k/m and introduced in [5].

In BRW/r/k/m more general multi-point perturbations of the self-adjoint operator $\mathcal{A}$ generated of the symmetric random walk are used than in BRW/r/0/0 or in BRW/0/k/0, see, e.g., [6]. This follows from the statement, see [5], that the mean numbers of particles $m_1(t) = m_1(t, \cdot, y)$ at a point $y \in \mathbf{Z}^d$ in BRW/r/k/m are governed by:

$$\frac{dm(t)}{dt} = \mathcal{Y} m_1(t), \quad m_1(0) = \delta_y, \tag{1}$$

where

$$\mathcal{Y} = \mathcal{A} + \left(\sum_{s=1}^{r} \beta_s \varDelta_{z_s} \right) + \left(\sum_{i=1}^{k} \zeta_i \varDelta_{x_i} \mathcal{A} + \sum_{i=1}^{k} \eta_i \varDelta_{x_i} \right) + \left(\sum_{j=1}^{m} \chi_j \varDelta_{y_j} \mathcal{A} \right). \tag{2}$$

Here, $\mathcal{A} : l^p(\mathbb{Z}^d) \to l^p(\mathbb{Z}^d)$, $p \in [1, \infty]$, is a symmetric operator, $\varDelta_x = \delta_x \delta_x^T$, and $\delta_x = \delta_x(\cdot)$ denotes a column-vector on the lattice taking the unit value at the point x and vanishing at other points, β_s, ζ_i, η_i, and χ_j are some constants. The same equation is also valid for a mean number of particles (a mean for a particle population size) over the lattice $m_1(t) = m_1(t, \cdot)$ with the initial condition $m_1(0) = 1$ in $l^\infty(\mathbb{Z}^d)$. The operator (2) can be written as

$$\mathcal{Y} = \mathcal{A} + \sum_{i=1}^{k+m} \zeta_i \varDelta_{u_i} \mathcal{A} + \sum_{j=1}^{k+r} \beta_j \varDelta_{v_j}. \tag{3}$$

In each of the sets $U = \{u_i\}_{i=1}^{k+m}$, and $V = \{v_j\}_{j=1}^{k+r}$, the points are pairwise distinct, but U and V may have a nonempty intersection. The points from $V \setminus U$ correspond to r sources of the first type; those from $U \cap V$ to k sources of the second type; and those from $U \setminus V$ to m sources of the third type.

Denote the highest positive eigenvalue of the operator Y by λ_0. We assume that $\zeta_i \geq 0$ and $\beta_j \geq 0$ in (3). Under this assumption we obtain that if λ_0 exits then λ_0 simple strictly positive and guarantees the exponential growth of the first moment m_1 of the numbers of particles both at an arbitrary point $\mu_t(y)$ and on the entire lattice $\mu_t = \sum_{y\in\mathbf{Z}^d}\mu_t(y)$. As in BRW/$r$/0/0, the same condition $\lambda_0 > 0$ implies the exponential growth of the higher-order moments. For all $n \in \mathbb{N}$ and $x, y \in \mathbb{Z}^d$, if:

$$m(n,x,y) = \lim_{t\to\infty}\frac{E_x\mu_t^n(y)}{m_1^n(t,x,y)} = \lim_{t\to\infty}\frac{m_n(t,x,y)}{m_1^n(t,x,y)},$$
$$m(n,x) = \lim_{t\to\infty}\frac{E_x\mu_t^n}{m_1^n(t,x)} = \lim_{t\to\infty}\frac{m_n(t,x)}{m_1^n(t,x)},$$

then, the limits

$$\lim_{t\to\infty}\mu_t(y)\,e^{-\lambda_0 t} = \xi\psi(y), \qquad \lim_{t\to\infty}\mu_t\,e^{-\lambda_0 t} = \xi, \tag{4}$$

where $\psi(y)$ is the eigenfunction corresponding to the eigenvalue λ_0 and ξ is a nondegenerate random variable, are valid for multiple sources in the sense of moment convergence. Eq. (4) reflects the exponential growth of the total number of particles both at an arbitrary point and on the entire lattice with the parameter λ_0. Additionally, by the Carleman criterion, if the growth rate $m(n,x)$ is limited by the condition $\sum_{n=1}^{\infty} m(n,x)^{-1/(2n)} = \infty$, then the moments define the distribution ξ uniquely. In this case, relations Eq. (4) are valid in the sense of convergence in distribution, too. The proof of 4 based on joint work with I. Christolubov supported by RFBR, project No. 17-01-00468.

References

1. Albeverio, S., Bogachev, L., and Yarovaya, E. Asymptotics of branching symmetric random walk on the lattice with a single source. // Comptes-rendus de l'Académie des Sciences, Paris, séries I. 1998. Vol. 326, P. 975–980.
2. Bogachev, L.V. and Yarovaya, E.B. A limit theorem for a supercritical branching random walk on $\mathbf{Z}^d$ with a single source.// Uspehi Matematicheskih Nauk. 1998. Vol. 53, P. 229–230.
3. Yarovaya, E.B. Branching random walks in a heterogeneous environment. //Center of Applied Investigations of the Faculty of Mechanics and Mathematics of the Moscow State University, Moscow, 2007.
4. Vatutin, V. A., Topchiĭ, V. A., and Yarovaya, E. B. Catalytic branching random walks and queueing systems with a random

number of independent servers. // Teor. Ĭmovīr. Mat. Stat. 2003. Vol. 69, P. 1–15.

5. Yarovaya E.B. Spectral properties of evolutionary operators in branching random walk models. // Mathematical Notes. 2012. Vol. 92:1, P. 115–131.
6. Yarovaya E.B. Positive Discrete Spectrum of the Evolutionary Operator of Supercritical Branching Walks with Heavy Tails // Methodology and Computing in Applied Probability. 2017. Vol. 19:4, P. 1151-1167.

Convergence rate for some extended Erlang–Sevastyanov queueing system[*]

G.A. Zverkina

Moscow State University of Railway Engineering (MIIT), V. A. Trapeznikov Institute of Control Sciences of Russian Academy of Sciences, Moscow, Russia

Consider the queueing system with infinitely many servers. Let t_1, t_2,…, t_n,…$(t_i > t_{i-1})$ be the time moments of incoming of 1-th, 2-th,…n-th customer correspondingly; $\tau_i \stackrel{\text{def}}{=} t_i - t_{i-1}$, $t_0 \stackrel{\text{def}}{=} 0$. Let ξ_1, ξ_1,…ξ_n be the service length of 1-th, 2-th,…n-th customer correspondingly. The distribution of the random variables $\{\tau_i\}_{i\in\mathbb{N}}$ and $\{\xi_i\}_{i\in\mathbb{N}}$ will be described by the *intensities*, and these intensities will be dependent on the *full queueing system state.*

Namely, the full queueing system state consists of changeable number of variables: $x_t^{(0)} \stackrel{\text{def}}{=} t - \max\{t_i:\ t_i \leqslant t\}$ (the time from the last arrival of the customer), and $x_t^{(i)}$ – the the elapsed time of service of i-th customer (in order of input) from n_t customers which are in the service at the time t. So, the behaviour of the queueing system is described by the stochastic process $X_t = \left(n_t, x_t^{(0)}; x_t^{(1)}, x_t^{(2)}, \dots, x_t^{(n_t)}\right)$ – for convenience, the variable n_t added here. The state space of the process X_t is

$$\mathcal{X} \stackrel{\text{def}}{=} \bigcup_{i=0}^{\infty} \mathcal{S}_i,$$

where $\mathcal{S}_i \stackrel{\text{def}}{=} \{i\} \times \prod\limits_{j=0}^{i} \mathbb{R}_+$, $i \in \mathbb{Z}_+$; the set $\mathcal{S}_0$ is a set of idle states of the system.

[*] This research is supported by RFBR grant No 17-01-00633 A

The intensity of input flow is $\lambda(X_t)$, and the intensity of the service of i-th customer from n_t customers current in the system.

This means, that:

$$\mathbf{P}\left\{\begin{array}{l} n_{t+\Delta} = n_t + 1,\ x_{t+\Delta}^{(0)} = x_{t+\Delta}^{(n_{t+\Delta})} \in (0;\Delta), \\ \\ \text{and } x_{t+\Delta}^{(i)} = x_t^{(i)} + \Delta \text{ for all } i = 1, \dots, n_t \end{array}\right\} = \lambda(X_t)\Delta + o(\Delta);$$

$$\mathbf{P}\{n_{t+\Delta} = n_t - 1\} = \sum_{i=1}^{n_t} h_i(X_t)\Delta + o(\Delta);$$

$$\mathbf{P}\left\{\begin{array}{l} n_{t+\Delta} = n_t - 1,\ \text{and :} \\ \\ 1. \text{ for all } j < i,\ x_{t+\Delta}^{(j)} = x_t^{(j)} + \Delta, \\ \\ 2. \text{ for all } j > i\ x_{t+\Delta}^{(j)} = x_t^{(j+1)} + \Delta \end{array}\right\} = h_i(X_t)\Delta + o(\Delta)$$

$$\text{– for all } i = 1, 2, \dots, n_t;$$

$$\mathbf{P}\left\{n_{t+\Delta} = n_t; x_{t+\Delta}^{(i)} = x_t^{(i)} + \Delta \text{ for all } i = 0, \dots, n_t\right\} = \\ = 1 - \left(\lambda(X_t) + \sum_{i=1}^{n_t} h_i(X_t)\right)\Delta + o(\Delta).$$

In these conditions, the process X_t is regenerative, and its regenerative points are the times when $X_t = (1, 0; 0)$, i.e. the jumps from the idle state to the busy state.

In [1], this extended Erlang-Sevastyanov queueing system was studied in the conditions:

$$0 < \lambda_0 \leqslant \lambda(X_t) \leqslant \Lambda < \infty, \qquad h_i(X_t) \geqslant \frac{K}{1 + x_t^{(i)}}, \qquad K > 2. \tag{1}$$

In [1], the convergence rate of the distribution $\mathcal{P}_t$ of X_t to the stationary one $\mathcal{P}$ was estimated, namely:

If the conditions (1) are true, then there exists the calculated constant $C(\Lambda, \lambda_0, K, k)$ such that for all $k \in [0, K-1)$ the inequality

$$\|\mathcal{P}_t - \mathcal{P}\|_{TV} \leqslant \frac{C(\Lambda, \lambda_0, K, k)}{(1+t)^k}$$

is true. The algorithm of the calculation of $C(\Lambda, \lambda_0, K, k)$ was given in [1].

Now, the generalization of this fact is proved.

Theorem. *If*

$$0 < \frac{\lambda_0}{1+x_t^{(0)}} \leqslant \lambda(X_t) \leqslant \Lambda < \infty, \qquad h_i(X_t) \geqslant \frac{K}{1+x_t^{(i)}}, \qquad K > 2,$$

then the distribution $\mathcal{P}_t$ of the process X_t weakly converges to the stationary distribution $\mathcal{P}$, and for all $k \in [0, K-1)$ there exists computable constant $\hat{C}(\Lambda, \lambda_0, K, k)$ such that

$$\|\mathcal{P}_t - \mathcal{P}\|_{TV} \leqslant \frac{\hat{C}(\Lambda, \lambda_0, K, k)}{(1+t)^k}.$$

In the algorithm of the calculation of the constatnt $\hat{C}(\Lambda, \lambda_0, K, k)$ the coupling constant $\varkappa$ (see [1]) is less than one in [1].

References

1. Zverkina G.A. Simple bounds for the convergence rate of $M|G|\infty$ queueing system // Analytical and computational methods in probability theory and its applications (ACMPT-2017) Proceedings of the International Scientific Conference 23–27 October 2017, Moscow, Russia. Moscow: RUDN, 2017. P. 613–618 (in rusian).

Short abstracts

Continuous optimization problems

On existence of minima of lower semicontinuous functions and solvability of nonlinear equations*

A.V. Arutyunov and S.E. Zhukovskiy
Peoples' Friendship University of Russia, Moscow, Russia

In [1], the following sufficient condition for existence of minima of functions defined on a metric space was proved.

Let (X, ρ) be a complete metric space, $U : X \to \mathbb{R}$ be a lower semicontinuous function bounded from below by some $\gamma \in \mathbb{R}$, i.e. $U(x) \geq \gamma$ $\forall\, x \in X$. *Given* $k > 0$, *assume that*

$$\forall\, x \in X : \;\; U(x) > \gamma \;\; \exists\, x' \in X \setminus \{x\} : \;\; U(x') + k\rho(x, x') \leq U(x). \qquad (1)$$

Then for every $x_0 \in X$ *there exists a point of minimum* $\bar{x} \in X$ *of the function* U *such that* $U(\bar{x}) = \gamma$ *and* $\rho(x_0, \bar{x}) \leq k^{-1}(U(x_0) - \gamma)$.

In the talk we discuss application of this result to equations in metric spaces. We present sufficient conditions for a differentiable mapping $F : \mathbb{R}^n \to \mathbb{R}^k$ to be surjective. In particular, this result is similar to the Hadamard theorem on homeomorphism (see, for example, Theorem 5.3.10 in [2]).

References

1. Arutyunov A.V. Caristi's condition and existence of a minimum of a lower bounded function in a metric space. Applications to the

*The first author is supported by a grant from the Russian Science Foundation (project no. 17-11-01168). The second author is supported by a grant of the President of the Russian Federation (project no. MK-2085.2017.1).

theory of coincidence points // Proc. Steklov Inst. Math. 2015. V. 291. Iss. 1. P. 24–37.

2. Ortega J.M., Rheinboldt W.C. Iterative solution of nonlinear equations in several variables. New York: Academic Press, 1970.

High-order methods for variational inequities with coupled constraints

M. Jaćimović and N. Mijajlović
University of Montenegro, Podgorica, Montenegro

We will study high-order methods for solving a class of the variational inequalities with coupled constraints when the changeable set is described by translation of a fixed, closed and convex set. We will present continuous and iterative variants of the second-order gradient-type projection method, establish sufficient conditions for the convergence of the proposed methods and derive a estimate of the rates of the convergence.

References

1. Antipin A.S., Jacimovic M., Mijajlovic N. A second-order continuous method for solving quasi-variational inequalities. Comp. Math. Math. Phys. 51(11):1856-1863, (2011)
2. Antipin A.S., Jacimovic M., Mijajlovic N. A second-order iterative method for solving quasi-variational inequalities. Comp. Math. Math. Phys. 53(3): 258-264, (2013)
3. Ryazantseva I. P., Second-order methods for some quasivariational inequalities, Differ. Equations 44, 1006-1017 (2008)

OR in biology, medicine, physics and ecology

Control of large dynamic systems using a hierarchical distributed MPC approach

O.Y. Maryasin
Yaroslavl State Technical University, Yaroslavl, Russia

An approach, based on using predictive models for managing various dynamic systems, is gaining much attention across the Western scientific community. This approach has acquired the name of Model Predictive Control (MPC) and has long proven itself in applications within such areas as Chemistry and Petrochemistry. Lately, it has extended its use to such areas as Energy Engineering, Transport as well as some other spheres.

To control large dynamic system consisting of multiple interconnected subsystems, distributed versions of MPC are used, such as distributed MPC and hierarchical distributed MPC [1]. Of these, only hierarchical, distributed MPC allows to achieve a minimum of the global quality criterion for the whole system, with due regard to the interrelations between the subsystems. When implementing a hierarchical distributed MPC-algorithm, within each step of its implementation, there arises the problem of coordinated solution for the mathematical programming problems for each of the subsystems. For the criterion of optimality in the form of the norm l_2, these are going to be the problems of quadratic programming.

The authors propose a method for solving the global mathematical programming problem, performed at each step of the MPC algorithm, based on the decomposition method via resource sharing [2]. The authors prove that under certain assumptions on sets of admissible solutions for local problems, if the local optimization problems have a solution, then the coordination problem will have an admissible optimal solution. Since the target function of the coordinating task may turn out to be non-differentiable at individual points, the Bundle method is used to solve it. If random factors influence the fulfillment of constraints in local problems, then a one-step stochastic programming problem with rigid and / or soft constraints can be put in place [3].

The authors used the proposed method for solving the problem of managing energy consumption and the microclimate of large multi-zone buildings. The mathematical model of the microclimate of a large multi-zone building is based on the equations of thermal and material balance, and is described by a system of ordinary differential equations. To ensure the required climate in the building, various types of energy resources, including renewable ones, can be used. For each type of energy resources, there are restrictions for both individual zones and the entire building. Therefore, the energy input for managing the microclimate within a building will depend on the consumption of energy resources for other, including domestic, needs. Since the consumption of energy resources for domestic needs is accidental, the restrictions on the consumption of energy resources for managing the climate will be accidental as well.

The results of numerical experiments showed the advantages of using the method for controlling the microclimate of large multi-zone buildings proposed by the authors. The availability of various types of energy resources allows, for example, in case of a sharp increase in the household needs consumption of thermal energy during peak hours, to increase the input of electricity or gas to maintain the required microclimate. At the same time, on the whole, a minimum level of energy input is secured, with regard to the implementation of global constraints and the cost of energy resources at current tariffs.

References

1. Scattolini R. Architectures for distributed and hierarchical Model Predictive Control. - A review. Journal of Process Control, 19, 2009, P. 723–731.
2. Lasdon L. Optimization of large systems. Moscow: Nauka, 1975.–432 p.
3. Ostrovsky G.M., Ziyatdinov N.N., Lapteva T.V. Optimization of technical systems. Moscow: KNORUS, 2012.–432 p.

Game-theoretic models

Collective action and the evolution of social norm internalization

S. Gavrilets
National Institute for Mathematical and Biological Synthesis,
Center for the Dynamics of Social Complexity, Knoxville, TN, USA

Human behavior is strongly affected by culturally transmitted norms and values. Certain norms are internalized (i.e., acting according to a norm becomes an end in itself rather than merely a tool in achieving certain goals or avoiding social sanctions). Humans' capacity to internalize norms likely evolved in our ancestors to simplify solving certain challenges – including social ones. Here we study theoretically the evolutionary origins of the capacity to internalize norms. In our models, individuals can choose to participate in collective actions as well as punish free riders. In making their decisions, individuals attempt to maximize a utility function in which normative values are initially irrelevant but play an increasingly important role if the ability to internalize norms emerges. Using agent-based simulations, we show that norm internalization evolves under a wide range of conditions so that cooperation becomes "instinctive." Norm internalization evolves much more easily and has much larger effects on behavior if groups promote peer punishment of free riders. Promoting only participation in collective actions is not effective.

Typically, intermediate levels of norm internalization are most frequent but there are also cases with relatively small frequencies of “oversocialized” individuals willing to make extreme sacrifices for their groups no matter material costs, as well as “undersocialized” individuals completely immune to social norms. Evolving the ability to internalize norms was likely a crucial step on the path to large-scale human cooperation.

Analysis of political processes and corruption

Optimal majority thresholds for different distributions in voting in a stochastic environment

V.A. Malyshev

V. A. Trapeznikov Institute of Control Sciences of Russian Academy of Sciences, Moscow, Russia

There is voting paradox consisting in the fact that democratic decisions taken by the majority systematically reduce social wealth if proposals are generated by the stochastic environment. It has been described in [1] (the ”pit of losses” paradox). The simplest approach to ”neutralize” this nuisance is increasing the majority threshold (the percentage of society that should support a proposal to accept it). However, society can miss many profitable proposals using an excessively high majority threshold in a favorable environment. In this paper, we show how to choose the optimal majority threshold to maximize average capital increment of a member of the society.

We use the ViSE (Voting in Stochastic Environment [2]) model. The society consists of n participants (egoists). Every participant supports the proposals that increase his/her own capital. The behavior of the voters corresponds to the Downsian concept [3]. A proposal of the environment is a vector of proposed capital increases of the participants. The proposals are successively put to a general vote. If the proposal is supported by a proportion of the society exceeding the majority threshold, then it is accepted (the voting procedure is ”α-majority” [4, 5, 6]) and the participants’ capitals receive their increments. Otherwise, the capital values remain the same. In accordance with the ViSE model the capital increments (that form the proposal of the environment) are the realizations of independent identically distributed random variables. A very similar model with randomly generated proposals is presented in [7].

We obtain the expressions for the optimal majority threshold for different distributions (a general expression and its specializations for the normal, continuous uniform, and symmetrized Pareto distributions) in such a kind of society. Moreover, we show that there is no ”pit of losses” (a negative average capital increment) in the case of the optimal majority threshold. In an unfavorable environment (with a negative expected value), the threshold is usually greater than 0.5 and less than 0.5 in a favorable one. The curves of the optimal threshold are different for different distributions. This work presents some generalization of [1].

References

1. Chebotarev P.Yu., Malyshev V.A., Tsodikova Ya.Yu., Loginov A.K., Lezina Z.M., Afonkin V.A. The Optimal Majority Threshold as a Function of the Variation Coefficient of the Environment // Automation Remote Control. 2018. To appear.
2. Borzenko V.I., Lezina Z.M., Loginov A.K., Tsodikova Ya.Yu., Chebotarev P.Yu. Strategies of Voting in Stochastic Environment: Egoism and Collectivism // Automation Remote Control. 2006. V. 67. N 2. P. 311–328.
3. Downs A. An Economic Theory of Democracy. – New York: Harper and Brothers, 1957.
4. Nitzan S., Paroush J. Optimal Decision Rules in Uncertain Dichotomous Choice Situations // International Economic Review. 1982. V. 23. N 2. P. 289–297.
5. Nitzan S., Paroush J. Are Qualified Majority Rules Special? // Public Choice. 1984. V. 42. N 3. P. 257–272.
6. Rae D.W. Decision-Rules and Individual Values in Constitutional Choice // American Political Science Review. 1969. V. 63. N 1. P. 40–56.
7. Compte O. and Jehiel Ph. On the Optimal Majority Rule. // CEPR Discussion Paper. 2017. N DP12492.

The typical models for congested traffic

The physics of empirical nuclei for spontaneous traffic breakdown in free flow at highway bottlenecks

B.S. Kerner[1], M. Koller[2], S.L. Klenov[3], H. Rehborn[2], and M. Leibel[4]

[1] *Physik von Transport und Verkehr, Universität Duisburg-Essen, 47048 Duisburg, Germany*

[2] *Daimler AG, 71063 Sindelfingen, Germany*

[3] *Moscow Institute of Physics and Technology, Department of Physics, 141700 Dolgoprudny, Moscow Region, Russia*

[4] *Karlsruhe University of Applied Sciences, 76133 Karlsruhe, Germany*

Based on an empirical study of real field traffic data measured in 1996–2014 through road detectors installed on German freeways [1], in this presentation we reveal physical features of empirical nuclei for spontaneous traffic breakdown in free flow at highway bottlenecks. A microscopic stochastic three-phase traffic model of the nucleation of spontaneous traffic breakdown presented in the talk explains the empirical findings. It turns out that in the most cases a nucleus for the breakdown occurs through an interaction of one of waves in free flow with an empirical permanent speed disturbance localized at a highway bottleneck (Fig. 1). The wave is a localized structure in free flow, in which the total flow rate is larger and the speed averaged across the highway is smaller than outside the wave. The waves in free flow appear due to oscillations in the percentage of slow vehicles; these waves propagate with the average speed of slow vehicles in free flow. Any of the empirical waves exhibits a two-dimensional asymmetric spatiotemporal structure: Wave's characteristics are different in different highway lanes.

References

1. Kerner B.S. The Physics of Traffic. Berlin, New York: Springer, 2004.

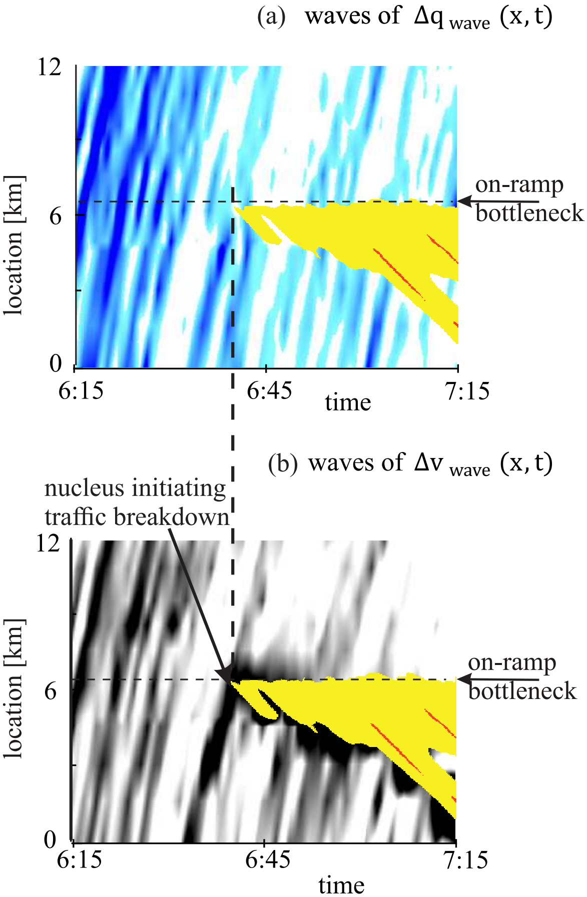


Fig. 1. Empirical nucleus in free flow. Empirical flow rate waves Δq_{wave} (a) and the speed waves Δv_{wave} (b). Real field traffic data measured by road detectors on three-lane freeway A5-South in Germany on April 15, 1996 (Monday).

Asymptotic analysis of complex stochastic systems

Moment asymptotics for branching random walks with immigration*

D. Han[1], Yu. Makarova[2], S. Molchanov[1,3], and E. Yarovaya[2]
[1] *University of North Carolina at Charlotte, Charlotte, NC, USA,*
[2] *Lomonosov Moscow State University,* [3] *Higher School of Economics ,*
Moscow, Russia

The evolution of populations with birth, death and migration can be described in terms of branching random walks, see details, e.g., in [2].

We consider a symmetric continuous-time branching random walk with generation of particles at every point of a multidimensional lattice and infinite number of initial particles. We allow immigration in our model. It can help to stabilize the population when the birth rate is less than the mortality rate. Such a model may describe the demographic situations associated with immigration in different Europian countries. This approach was suggested by Han, Molchanov and Whitmeyer in [1], but only for the case of the binary splitting, i.e. when one particle can produce one offspring at the moment of birth.

In [3] we considered the case when particles can produce an arbitrary number of offsprings and investigated the asymptotic behaviour of the first two moments of particle numbers as $t \to \infty$. In this paper, we present analysis of high-order moments. These results make it possible to obtain the limit theorem on behavior of the numbers of particles in the branching random walk with immigration.

The subject of our study is the particle field $n(t,x)$, $t \geqslant 0$, $x \in \mathbb{Z}^d$. In the initial moment $t = 0$ we assume that $n(0,x)$ are independent identically distributed random variables with finite exponential moments. The evolution of the field includes four opportunities.

Firstly, each particle can jump from the point x to the point $x + z$ with probability $a(z)$. We assume that $a(z) = a(-z)$, $\sum_{z \neq 0} a(z) = 1$ and $a(0) = -1$. The intensity of jumps is denoted by κ. Then the probability to jump from the point x to the point $x + z$ during the small time dt is $\kappa a(z)dt$. Moreover, we assume that the random walk is irreducible with finite variance of jumps. Secondly, each particle can die with the

*Yu. Makarova and E. Yarovaya were supported by the Russian Foundation for Basic Research (RFBR), project No. 17-01-00468. S. Molchanov was supported by the Russian Science Foundation (RSF), project No. 17-11-01098.

mortality rate μ. Thirdly, each particle, independently on others, can produce n new offsprings (or we can say that it produces $n-1$ new particles and still stays at the same point on the lattice). Let b_n, $n \geqslant 2$, be the intensity to produce n offsprings and $\beta := \sum_{n=2}^{\infty}(n-1)b_n$. Finally, we allow immigration. It means that at any point on the lattice, during a small time interval $(t, t+dt)$, a new particle can come from outside with the probability kdt, where k is the rate of immigration.

We consider the moments $m_n(t, x_1, ..., x_n) = E\,(n(t, x_1) \cdot ... \cdot n(t, x_n))$. In [7] the asymptotics of the first two moments $m_1(t, x_1)$ and $m_2(t, x_1, x_2)$ were obtained as $t \to \infty$. Our main result is the asymptotic behavior of $m_n(t, x_1, ..., x_n)$, as $t \to \infty$, for every $n \geq 3$, $k > 0$ and $\mu > \beta$.

Based on the results for the moments it is shown that for $t \to \infty$ the limit sequence of moments $\{m_n(t, x_1, ..., x_n)\}_{n=1}^{\infty}$ uniquely determines the limit distribution of some random variable.

References

1. Han D., Molchanov S., and Whitmeyer J. Population processes with immigration //In book: Modern problems of stochastic analysis and statistics — selected contributions in honor of Valentin Konakov, V. Panov (ed). Springer, 2017. p. 411–434.
2. Yarovaya E.B. Branching random walks in a heterogeneous environment. Moscow: Center of applied investigations of the Faculty of Mechanics and Mathematics of the Mocsow State University, 2007.
3. Han D., Makarova Y., Molchanov S., Yarovaya E. Branching Random Walks with Immigration. In: Rykov V., Singpurwalla N., Zubkov A. (eds) Analytical and Computational Methods in Probability Theory. ACMPT 2017. Lecture Notes in Computer Science, vol 10684. Springer, Cham, 2017.

Author index