Proceedings include extended abstracts of reports presented at the III International Conference on Optimization Methods and Applications "Optimization and applications" (OPTIMA-2012) held in Costa da Caparica, Portugal, September 23–30, 2012. Edited by V.I. Zubov.


The Conference is organized by

University of Évora, CIMA-UE, Portugal

Dorodnicyn Computing Centre of RAS, Russia

Научное издание

© Федеральное государственное бюджетное учреждение науки
Вычислительный центр им. А. А. Дородницына
Российской академии наук, 2012
Organizing Committee

Vladimir Bushenkov, Chair, University of Evora, Portugal
Yury G. Evtushenko, Chair, Dorodnicyn Computing Centre of RAS, Russia

Alexander P. Afanasiev, Institute for System Analysis of RAS, Russia
Rasim Alguliev, Institute of Information Technology of ANAS, Azerbaijan
Anatoly S. Antipin, Dorodnicyn Computing Centre of RAS, Russia
Oleg Burdakov, Linkoping University, Sweden
Vladimir A. Garanzha, Dorodnicyn Computing Centre of RAS, Russia
Alexander I. Golikov, Dorodnicyn Computing Centre of RAS, Russia
Vladimir Goncharov, University of Evora, Portugal
Alexander Yu. Gornov, Institute System Dynamics and Control Theory, SB of RAS, Russia
Miloica Jacimovic, Montenegrin Academy of Sciences and Arts, Montenegro
Alexander V. Lotov, Dorodnicyn Computing Centre of RAS, Russia
Yuri Nesterov, CORE Universite Catholique de Louvain, Belgium
Vera Roshchina, University of Evora, Portugal
Yaroslav D. Sergeyev, University of Nizhni Novgorod, Russia, University of Calabria, Italy
Gueorgui Smirnov, University of Minho, Portugal
Tatiana Tchemisova, University of Aveiro, Portugal
Ivetta A. Zonn, Dorodnicyn Computing Centre of RAS, Russia
Vladimir I. Zubov, Dorodnicyn Computing Centre of RAS, Russia
The Conference is organized in cooperation with

The Continuous Optimization Working Group of the Association of European Operational Research Societies (EUROPT)

Associação Portuguesa de Investigação Operacional (APDIO)

University of Aveiro, CIDMA-UA, Portugal

Centro Internacional de Matemática, Portugal

Câmara Municipal de Almada, Portugal

Hotel Costa de Caparica, Portugal

Companhia Industrial de Resinas Sintéticas, S.A., Portugal

Fundação para a Ciência e a Tecnologia, Portugal
Contents

Vagif Abdullayev  Numerical solution to optimal control problems for loaded dynamic systems with integral conditions ........................................ 10

Alexander P. Abramov  On second eigenvalue in Leontiev’s model ........................................... 13

Oleg V. Abramov  Optimization Techniques for Parametric Synthesis of Engineering Systems ............................................................................. 16

Alexander Afanasyev, Elena Putilina  Maximizing the volume of three-dimensional bodies on the basis submetric transformation ......... 21

Kamil Aida-zade, Yegana Ashrafova  Optimal control problems without initial conditions ............................................................................. 21


Keyvan Amini, Masoud Ahookhosh, Somayeh Bahrami  A Conjugate Gradient Projection Algorithm for systems of Large-Scale Nonlinear Monotone Equations ............................................................................. 30

Anton Anikin  Software implementation of an algorithm for finding the optimal control using a graphics accelerator ........................................ 31

Maxim Anop, Yaroslava Katueva  The reduction of the optimal parametric synthesis to the linear programming problem ......................... 34

Anatoly Antipin  Boundary value games in optimal control ..................... 36

Dmitry Arkhipov, Alexander Lazarev, Elena Musatova  Minimization of maximum lateness for M stations with tree topology ..................... 42

Armen Beklaryan  Existence theorems for elliptic equations in unbounded domains ................................................................. 47

Vladimir Bushenkov, Bento Caldeira, Georgi Smirnov  On the determination of the earthquake slip distribution via linear programming techniques ............................................................................. 51

Miguel Constantino, Xenia Klimentova, Ana Viana  New Integer Programming formulations for the Kidney Exchange Problem .......... 54

V.V. Dikusar, E.S. Zasukhina  Identification of parameters in model of water transfer in soil ................................................................. 59
Anna Dorjieva  Improvement Technology for the Accuracy of Solution of Unconstrained Argument Problems ........................................ 63

Olga Druzhinina, Natalia Petrova On optimal stabilization with respect to a part of variables for multiply connected controlled systems ... 67

W. Dunin-Barkowski, L. Vyshinskiy Numerical experiments on computer model of a cerebellum .................................................. 72

Yu.G. Evtushenko, A.A. Tretyakov $P$-th order methods for solving nonlinear system ................................................................. 76

Evgeny R. Gafarov, Alexandre Dolgui, Alexander Lazarev Some Complexity Results for the Simple Assembly Line Balancing Problem ... 81

Shamil Galiev, Maria Lisafina, Vitalii Yudin Optimization of a multiple covering of a surface taking into account its relief ......................... 86

Alexander Gasnikov, Eugenia Gasnikova Stochastic subgradient barrier-multiplicative descent for entropy optimization .................... 91

Dinh Thanh Giang, Phan Thanh An, Le Hong Trang An Efficient Algorithm for Determining the Lower Convex Hull of a Finite Point Set in 3D ........................................ 93

Alexander I. Golikov LP projection algorithm and Newton method for solving dual LP problems ...................................................... 94

Evgenij Golshteijn Many-Person Games With Convex Structure ........... 99

Vladimir Goncharov, Fátima Pereira Proximal Analysis and Regularity of Viscosity Solution to some Hamilton-Jacobi Equation .......... 101

Aleksander Gornov Multimethod’s algorithm for parametric identification of nonlinear dynamic systems ........................................... 103

Anna D. Guerman Optimization Problems in Astrodynamics ............... 106

Nguyen Ngoc Hai, Phan Thanh An Blaschke Convergence Theorem for $G$-type Convex Sets in Metric Spaces ................................. 106

Niyaz Ismagilov, Farit Nasyrov Pathwise optimal control of diffusion type processes ............................................................. 111

M. Jaćimović, I. Krnić On accuracy of the regularization method of constrained ill-posed quadratic minimization problems .............. 115
Ruben V. Khachaturov Cubes Lattice’s properties investigation and possibilities of its application in Combinatorial Optimization .......... 118

Vladimir R. Khachaturov General Theory of Optimization on Finite Lattices .................................................. 122

Michael Khachay, Maria Poberii Hyperplane Covering Problems. Complexity and Approximation Issues ......................... 124

Elena Khoroshilova Leontief’s model as a boundary value problem in optimal control ........................................ 128

Konstantin Kobylkin Approximation to minimum committee problem for system of linear inequalities in \( \mathbb{R}^3 \) ......................... 133

Pavel Korenev, Alexander Lazarev Metric for the total tardiness minimization problem ................................. 134

Olga Kostyukova, Tatiana Tchemisova New CQ-free optimality criterion for convex SIP problems with polyhedral index sets .......... 139

Vladimir Krivonozhko, Finn Førsund, Andrey Lychev Optimization methods for measurement of returns to scale in the non-radial DEA models .................................................. 144

Ulyana Kulbida, Olga Kaneva Selection of the target audience by the leverage method in the expert system for advertising specialist .......... 147

Samir Kuliev Zonal control of lumped systems on different classes of feedback functions ........................................ 153

Maya Laskova, Alexander Lazarev, Elena Musatova The Heuristic Approach to movement optimization on single-track part of the railway net 156

Alexander V. Lotov, Georgij Kamenev Finite time-interval robustness study of dynamic systems with imprecisely identified parameters ... 159

A.A. Lukovenko, T.M. Tikhomirova Estimation of economic damage from human mortality by external causes on macro-, meso-, micro- levels 162

Eloísa Macedo, Adelaide Freitas Statistical Methods and Optimization in Data Mining ........................................ 164

Arsalan Mizhidon Analytic design of an optimal controller under permanent stochastic disturbances .......................... 175

Boris Mordukhovich Optimal control of the sweeping process ....... 178

Evgenii Murashkin Optimal deformation during the creep .......... 178

L.A. Muravey, V.M. Petrov, A.M. Romanenkov Modeling and optimization of ion-beam etching process ......................... 181

Elena Musatova, Alexander Lazarev, Nail Husnullin Special algorithm for Three-Stations Railway problem ........................ 187

Daniar Nurseitov, Maksim Shishlenin, Syrym Kasenov Numerical solution of two-dimensional inverse problem for the Helmgoltz equation . 192

Nataliya Obrosova, Alexander Shanin Production model in the conditions of unstable demand .................................. 194

Andrei Orlov, Sergei Pinigin Global search in bilinear separation problems .............................................................. 198

Valeriy Parkhomenko Ensemble calculations application for estimation and optimization of climate model parameters ................. 203

Sergey Perzhabinsky, Valery Zorkaltsev Interior point algorithms . 207

Lev F. Petrov Interactive optimization as a tool for finding the complex periodic solutions in nonlinear dynamics ......................... 213

M.R. Peyghami, H. Tavakoli, M. Ahmadian Attari On the Semidefinite Representation of the Maximum Optimal Rate Problem in LDPC Codes ................................................................. 218

Alexander Plakhov Problems of optimal resistance in Newtonian aerodynamics ......................................................... 219

Boris Polyak $L_1$ problems in control and numerical methods for their solution ................................................................. 220

Mikhail Posypkin, Izrael Sigal Object-oriented Framework for Dynamic Control of the Parallel Branch-and-Bound ......................... 221

Ekaterina Rassadnikova, Aida Valeeva, Nelli Magafurzyanova Graf of decision logistics making for problem of goods delivery ........ 223

Nataliya Sedova Discontinuous control and Lyapunov functions for nonlinear systems ................................................. 227
Simon Serovajsky Optimal control of nonlinear parabolic equations and the differentiability of the control-state mapping .................. 229

Iliyas Shakenev Inverse problems for parabolic equations with infinite horizon ................................................................. 234

Kanat Shakenev Solution of the parametric inverse problem of stochastic optimal control .................................................. 237

O. Shcherbina, A. Sviridenko Graph-based local elimination algorithms for sparse discrete optimization problems .................. 240

Yuri N. Sotskov, Omid Gholami, Frank Werner Heuristic Algorithms for a Job-Shop Problem with Minimizing Total Job Tardiness ... 245

A.A. Tretyakov P-regular nonlinear optimization. High order optimality conditions ............................................................... 250

Evgeniya A. Vorontsova Separating plane algorithm with additional clipping for convex optimization .............................. 254

V.P. Vrzheshch, N.P. Pilnik, I.G. Pospelov Equilibrium Model of the Russian Economy for the period of Global Financial Crisis ....... 256

R. Yousefpour A Modified Steepest Descent Method Based on BFGS Method for Locally Lipschitz Functions .......................... 257

Vitaly Zhadan The primal affine-scaling method for semidefinite programming with steepest descent ..................................... 262

Valery Zorkaltsev The point projections on linear manifold ............. 267

Vladimir Zubov, Alla Albu, Andrey Albu The effect of the setup parameters on the evolution of the substance crystallization process .... 270

Anna Zykina, Nikolay Melenchuk Convergence of the two-step extra-gradient method in a finite number of iterations .................. 274

Author index ........................................................................................................................................ 278
Numerical solution to optimal control problems for loaded dynamic systems with integral conditions

Vagif Abdullayev

1 Cybernetics Institute of ANAS, Baku, Azerbaijan; vaqif_ab@rambler.ru

Let us consider the following optimal control problem for the process described by a system, which is linear with respect to the phase variable, of loaded ordinary differential equations:

$$\dot{x}(t) = A(t,u)x(t) + \sum_{s=1}^{l_3} B^s(t)x(\tilde{t}_s) + C(t,u), \quad t \in (t_0, T],$$  \hspace{1cm} (1)

where $x(t) \in E^n$ is the phase variable; $u(t) \in U \subset E^r$ is the control vector-function from the class of piecewise continuous functions, admissible values of which belong to the given compact set $U$; the $(n \times n)$ matrix functions $A(t,u) \neq \text{const}$, $B^s(t), s = 1, \ldots, l_3$, and $n$-dimensional vector-function $C(t,u)$ are continuous with respect to $t$ and continuously differentiable with respect to $u$. The points of loading time $\tilde{t}_s \in [t_0, T], \quad \tilde{t}_{s+1} > \tilde{t}_s, \quad s = 1, 2, \ldots, l_3$ are given.

Nonseparated multipoint and integral conditions are given in the following form:

$$\sum_{i=1}^{l_1} \int_{\tilde{t}_{2i-1}}^{\tilde{t}_{2i}} \bar{D}_i(x(\tau))d\tau + \sum_{j=1}^{l_2} \bar{D}_j x(\tilde{t}_j) + \sum_{s=1}^{l_3} \bar{D}_s x(\tilde{t}_s) = L_0,$$  \hspace{1cm} (2)

where the continuous matrix function $\bar{D}_i(\tau)$ and scalar matrices $\bar{D}_j, \bar{D}_s$ have the dimension $(n \times n)$; $L_0$ is the $n$-dimensional vector; $\tilde{t}_i, \tilde{t}_j$ are the points of time belonging to $[t_0, T]; \quad \tilde{t}_{i+1} > \tilde{t}_i, \quad \tilde{t}_{j+1} > \tilde{t}_j, \quad i = 1, \ldots, 2l_1 - 1, \quad j = 1, \ldots, l_2 - 1, \quad l_1, l_2, l_3$ are given.

To be definite, without loss of generality, make an assumption that

$$\min (\tilde{t}_1, \tilde{t}_1) = t_0, \quad \max (\tilde{t}_{2l_1}, \tilde{t}_{l_2}) = T,$$  \hspace{1cm} (3)

and for all $i = 1, \ldots, 2l_1, \quad j = 1, \ldots, l_2, \quad s = 1, \ldots, l_3, \quad l_1, l_2, l_3$ the following condition holds
The target functional is as follows:

\[
J(u) = \Phi(x(\hat{t})) + \int_{t_0}^T f^0(x, u, t) dt \rightarrow \min_{u(t) \in U},
\]

where the function \(\Phi\) is continuous with respect to its arguments along with the private derivatives, and \(f^0(x, u, t)\) is continuously differentiable with respect to \((x, u)\), and continuous with respect to \(t\); \(\hat{t} = (\hat{t}_1, \hat{t}_2, ..., \hat{t}_{2l_1+l_2})\) is the ordered union of points of the sets \(\hat{t} = (\hat{t}_1, \hat{t}_2, ..., \hat{t}_{2l_1})\) and \(\bar{t} = (\bar{t}_1, \bar{t}_2, ..., \bar{t}_s)\), i.e. \(\hat{t}_j < \hat{t}_{j+1},\ j = 1, ..., 2l_1 + l_2 + l_3 - 1\).

Suppose that the problem (1) and (2) is solvable under any admissible control \(u(t) \in U \in E^r\).

**Theorem.** The gradient of the functional in the problem (1)-(5) is determined as follows:

\[
(\text{grad } J(u))^* = \frac{\partial f^0(x, u, t)}{\partial u(t)} - \psi^*(t) \left[ \frac{\partial A^*(t, u)}{\partial u(t)} x(t) + \frac{\partial C^*(t, u)}{\partial u(t)} \right].
\]

where the vector-function \(\psi(t) \in E^n\) and the vector \(\lambda \in E^n\) satisfy the following differential equation:

\[
\dot{\psi}(t) = -A^*(t, u)\psi(t) - \sum_{s=1}^{l_3} \delta \left( t - \bar{t}_s \right) \int_{t_0}^T B^{s*}(t)\psi(t) dt + \sum_{i=1}^{l_1} [\chi(\bar{t}_2i) - \chi(\bar{t}_{2i-1})] \tilde{D}^*(t)\lambda + \frac{\partial f^{0*}(x, u, t)}{\partial x(t)},
\]

the following boundary conditions

\[
\psi(t_0) = \begin{cases} 
(\frac{\partial \Phi(x(\bar{t}_1))}{\partial x(\bar{t}_1)})^* + \tilde{D}_1^* \lambda, & \text{for } t_0 = \bar{t}_1, \\
(\frac{\partial \Phi(x(\bar{t}_1))}{\partial x(\bar{t}_1)})^*, & \text{for } t_0 = \bar{t}_1,
\end{cases}
\]

\[
\psi(T) = \begin{cases} 
- (\frac{\partial \Phi(x(\bar{t}_2))}{\partial x(\bar{t}_2)})^* - \tilde{D}_{t_2}^* \lambda, & \text{for } \bar{t}_{t_2} = T, \\
- (\frac{\partial \Phi(x(\bar{t}_1))}{\partial x(\bar{t}_1)})^*, & \text{for } \bar{t}_{2l_1} = T,
\end{cases}
\]
the following jump conditions at the intermediate points $\tilde{t}_j$, for which $t_0 < \tilde{t}_j < T$,

$$
\psi^+(\tilde{t}_j) - \psi^-(\tilde{t}_j) = \left( \frac{\partial\Phi(x(\hat{t}))}{\partial x(\tilde{t}_j)} \right)^* + \tilde{D}^*_j \lambda, \quad j = 1, \ldots, l_2, \quad (10)
$$

the following jump conditions at the loading points $\tilde{t}_s$, for which $t_0 < \tilde{t}_s < T$,

$$
\psi^+(\tilde{t}_s) - \psi^-(\tilde{t}_s) = \left( \frac{\partial\Phi(x(\hat{t}))}{\partial x(\tilde{t}_s)} \right)^* + \tilde{D}^*_s \lambda, \quad s = 1, \ldots, l_3, \quad (11)
$$

and the following jump conditions at the points $\bar{t}_i, i = 1, \ldots, 2l_1$, for which $t_0 < \bar{t}_i < T$,

$$
\psi^+(\bar{t}_i) - \psi^-(\bar{t}_i) = \left( \frac{\partial\Phi(x(\hat{t}))}{\partial x(\bar{t}_i)} \right)^*, \quad i = 1, \ldots, 2l_1. \quad (12)
$$

Here "*" is the transposition sign; $\delta(\cdot)$ is the delta function; $\chi(t)$ is the Heaviside function.

For numerical solution to the problem, we propose to use standard procedures of first order optimization. To determine the value of the gradient by the formula (6), at each iteration, it is necessary to: 1) solve the problem (1) under the current control with multipoint and integral conditions (3) using the technique of convolving integral conditions into local conditions (here we mean to use the results of the work [1]); 2) solve the adjoint problem (7)-(11) using the generalized operation of iterated shifts, making a special emphasis on the participation of the parameters $\lambda$ in the conditions (10)-(11) (here we mean to use the results of the works [2, 3]). Following the shift of the conditions, we obtain an algebraic system of equations with the $n(l_3 + 2)$ unknowns $\lambda$, with the values of the phase trajectory at one of two ends of the interval, and with the loading points [4].

Results of numerical experiments obtained by solving the problems of
the form (1)-(5) are given in the presentation.

REFERENCES


On second eigenvalue in Leontiev’s model

Alexander P. Abramov

1 Computing Center RAS, Moscow, Russia; apabramov@list.ru

Consider the Leontiev simple dynamic model. In our case, the economy is closed and consists of $n$ sectors. The output vector $x(t) \in \mathbb{R}^n$ satisfies the inequality $Y x(t) \leq x(t - 1)$ at step $t$, $t = 1, 2, \ldots$, where $Y = \{y_{ij}\}$ is a technological matrix. Assume that the matrix $Y$ is primitive.

If we put in order the eigenvalues of the matrix $Y$, we get

$$|\lambda_1| \leq |\lambda_2| \leq \ldots \leq |\lambda_{n-1}| < |\lambda_n = \lambda_Y|,$$

where $\lambda_Y$ is the Frobenius eigenvalue of the matrix $Y$.

Denote by $p_Y$ and by $x_Y$ the left and the right Frobenius vectors of the matrix $Y$ such that $p_Y x_Y = 1$. Using these vectors, let us define the square matrix $L$ such that $L = x_Y p_Y$. Recall [1] that the sequence \((Y/\lambda_Y)^t\), $t = 1, 2, \ldots$ tends to $L$ as $t \to \infty$. In this case the degree of convergence is estimated by the formula

$$\left\|(Y/\lambda_Y)^t - L\right\|_{l_\infty} < C r^t, \quad (1)$$
where $r$ is any number such that $|\lambda_{n-1}|/\lambda_Y < r < 1$ and $C = C(r, Y)$ is some positive constant.

In [2] it was considered the model of decentralised economy with Leon-
tiev’s technologies such that the economic system may asymptotically reach the balanced growth. Denote by $x^p(1)$ the vector of the planned out-
put at step 1. According to the model, for the normed sequence of output, we get $\{ (Y/\lambda_Y)^t x^p(1) \}$. This sequence tends to $\vartheta x_Y$, where $\vartheta \equiv p_Y x^p(1)$.

Using (1), we get

$$
\left\| (Y/\lambda_Y)^t x^p(1) - \vartheta x_Y \right\|_\infty < \tilde{C} r^t,
$$

where $\tilde{C}$ is some positive constant.

In our case, the scalar $r$ in (1)-(2) is bounded below by $|\lambda_{n-1}|/\lambda_Y$. This bound depends on second eigenvalue $\lambda_{n-1}$. Let us consider the economic contents of this value. Further assume that the matrix $Y$ is positive.

Recall that the elements of a positive matrix comply with Hopf’s bound [1], so that

$$
\frac{|\lambda_{n-1}|}{\lambda_Y} \leq \frac{M - \mu}{M + \mu} < 1,
$$

where $M = \max_{i,j} y_{ij}$ and $\mu = \min_{i,j} y_{ij}$. This estimate shows that the decreasing of the range $(M - \mu)$ decreases the upper bound for the fraction $|\lambda_{n-1}|/\lambda_Y$.

Assume that $\lambda_{n-1} > 0$. Denote by $z$ some eigenvector corresponding to $\lambda_{n-1}$ such that

$$
Y z = \lambda_{n-1} z.
$$

Since the matrix $Y$ is irreducible it follows that the vector $z$ has the components of different signs. Let $x_k$ be a component such that $z_k < 0$.

We say that the technological process is the reverse one if this process extracts the resources from the final product. It is assumed that the volumes of these resources are exactly the same as the original input. Moreover, the consumer properties are identical for the recovered resources and for the normal products. We stress that the definition of the reverse process is speculative.

Using this definition, we see that $|z_k|$ on the right-hand side in (4) is equal to the input for the reverse process from some external source. At the same time the product $y_{ik} z_k$ is equal to the volume of the resource of type $i$ coming into the system as a result of the reverse process.
On the other hand, the negative value of the sum $\sum_i y_k i z_i$ means that the amount of resource of type $k$, which will be produced in all reverse processes, exceeds the consumption of this product in all "direct" processes. Thus, the sector $k$ such that $z_k < 0$, may be considered as a multi-product producer of resources. This producer consumes only the product of type $k$, which comes into the system from an external source. We stress that the consumption of these products should grow with the same rate as the rate of economic growth corresponding to the turnpike. In addition, a reverse process should occur immediately at the beginning of each step such that the released resources were available for using in all starting "direct" processes.

It is clear that the additional resources increase the rate of balanced growth in sectors with "direct" processes. This fact is expressed formally by the inequality $(1/\lambda_{n-1}) > (1/\lambda_Y)$. It should be stressed that the economic system can not be closed if it uses at least one reverse process.

It follows from bound (3) that the ratio $\lambda_{n-1}/\lambda_Y$ tends to zero as the parameter $\mu$ tends to $M$ from below. It means that the decreasing of the range $(M - \mu)$ increases the rate of balanced growth compared to the $1/\lambda_Y$, which can be achieved by the reversion of some technological processes.

This conclusion is obvious because the system may use some resources without limit from an external source, and the specific consumption of these resources is approximately the same in all sectors.

In addition, the high efficiency of the reverse processes in terms of increasing the balanced growth rate also positive effects on the rate of convergence of the normalized outputs (see (2)).

The author was supported by the Russian Foundation for Basic Research (project no. 11-07-00201).

REFERENCES

Optimization Techniques for Parametric Synthesis of Engineering Systems

Oleg V. Abramov¹

¹ Institute for Automation and Control Processes, Far Eastern Branch RAS, Vladivostok, Russia; abramov@iacp.dvo.ru

The synthesis of engineering systems consists of two basic parts: developing of structure (structural synthesis) and internal parameter values choosing (parametric synthesis). This paper proposes the approach and some algorithms for seeking a numerical solution of the parametric optimization problem (parametric synthesis) of analogous electronic circuit. The circuit design optimization process is confounded by three significant kinds of unfavorable complexity, namely
- the complexity of function and gradient evaluation, which can be extreme,
- the combinatorial complexity of approximation algorithms which are basically exponential in \( n \), the dimension of the design parameter space,
- the uncertainty of models used for analogous electronic circuits, and in terms of the statistical uncertainty of the values assumed by the parameters of these models. In parametric optimization, the topology of the circuit and component types are fixed.

In general the optimal parametric synthesis problem can be stated as follows [1].

Suppose that we have a circuit which depends on a set of \( n \) parameters \( \mathbf{x} = (x_1, \ldots, x_n) \). We will say that circuit is acceptable if \( \mathbf{Y}(\mathbf{x}) \) satisfy the conditions (1):

\[
\mathbf{a} \leq \mathbf{Y}(\mathbf{x}) \leq \mathbf{b},
\]

where \( \mathbf{Y}(\mathbf{x}) \), \( \mathbf{a} \) and \( \mathbf{b} \) are \( m \)-vectors of circuit responses (output parameters) and their specifications. The inequalities (1) define a region \( D_x \) in the space of design parameters

\[
D_x = \{ \mathbf{x} \mid \mathbf{a} \leq \mathbf{Y}(\mathbf{x}) \leq \mathbf{b} \}
\]

\( D_x \) is called the tolerance margin domain (region of acceptability) for the circuit. It is region in the input parameters space.

Let given the characteristics of random processes \( \mathbf{X}(t) \) of system parameters variations, a region of admissible deviation - \( D_x \) and a service
time $T$, find such a deterministic vector of parameter ratings (nominals) $x_r = (x_{1r}, \ldots, x_{nr})$ that the probability

$$P_r(x_r, T) = P_r \{ [X_1(t) - x_{1r}, \ldots, X_n(t) - x_{nr}] \in D_x, \forall t \in [0, T] \}$$

be maximized.

Any optimization technique requires, first, a method of objective function calculation and, second, an extremum searching method which allows to find a solution with a minimum cost.

The practical algorithm of the stochastic criterion calculation is based on the conventional Monte Carlo method. The Monte Carlo method approximates $P_r(x_r, T)$ by the ratio of number of acceptable realizations (falling in region $D_x$) – $N_a$ to the total number of trials – $N$. Unfortunately, often the region $D_x$ is unknown. It is given only implicitly through system’s equations and the systems response functions. If we do not know the region $D_x$, then a Monte Carlo evaluation of probability $P_r(x_r, T)$ at particular nominal value $x_r$ requires $N$ system analysis for each trial set of parameter $x_r$. Typically, hundreds of trials are required to obtain a reasonable estimate for $P_r(x_r, T)$. Optimization requires the evaluation of our probability $P_r(x_r, T)$ for many different values of the nominal values of the parameters $x_r$. Therefore to make practical the use of Monte Carlo techniques in statistical system design, it is necessary to reduce the number of system analysis required during optimization.

As a solution, the following two-steps technique and the corresponding algorithms can be used for practical reliability optimization.

The first step consists in replacing the original stochastic criterion with a certain deterministic one, allowing nearby optimum solutions to be obtained. The two such objective functions are possible. One of them is a so-called a "minimal serviceability reserve" that can be presented in the general form:

$$F(x) = \min_{i-1,m} [(a_i - Y_i(x))/w_i - 1],$$

where $Y_i(x)$ – the $i$-th output value, $a_i$ – the $i$-th constraint ($Y(x) \leq a$) and $w_i$ – the $i$-th weight coefficient. From this, we have a following optimization problem:

$$x_r = \arg \max_{x \in D} F(x).$$

It means that such a nominal point should be found that would have the largest distance from the acceptability region margins.
An other method, which can be used, for the reliability optimization is so-called ”equal densities method”. This method is of combined type, which uses statistical data and a deterministic optimization technique.

At the first step we should estimate distribution density function (DDF) for output parameter. As can be shown analytically, probability maximum will be achieved, if DDF will be shifted such, that both lower and upper constraints will cut equal densities on DDF [1].

Now the first design step is completed. At the same time the next design step must be made if the reliability index that was achieved by using deterministic methods is not high enough. This step is a direct probability optimization, i.e.; methods of stochastic optimization should be used here.

It should be pointed out that most of optimization methods have the highest convergence speed when they start at a ”good” initial point. Therefore, it would be most natural to get a previous solution as an initial point for the next design step.

Particularly effective way to decrease total design time on the phase of modelling and statistical optimization is to use modern supercomputing technologies and parallel processing techniques [2].

Note that evaluation of \( extrP_r(x_r, T) \) requires a global optimization. The simplest method of global optimization is scanning (full enumeration) method. However, such method is considered computationally inefficient. The effective way to decrease optimization time is to use modern supercomputing technologies and parallel algorithms.

The nominal values of the schematic components \( x_n \) commonly used for engineering systems should lie in the predefined set of values as it is required by various standards and technical recommendations, it is sometimes more preferable to search the optimal vector inside the discrete set of values that conforms to the standards and lies in the acceptable region \( D_x \).

Let us have the known internal parameters vector \( x_r \in D_x \). Therefore at the each point of discrete set \( D_r^{in} = \{x_r^{in}/x_r \in D_x \} \) we need to find the \( \hat{P}_r(x_r^{in}) \) estimate. The optimum nominal vector \( x_r \) we are looking for can be found as a solution of the following task

\[
x_r = \max_{x_r} \hat{P}_r(x_r^{in}) \tag{3}
\]

In the simplest case the solution can be found by complete check of each
element of the set $D_r^{in}$ with the probability estimation for each of them. The set $D_r^{in}$ building can be implemented as a preliminary procedure that puts the element values to the database. The optimum search process can be performed in parallel mode. This algorithm can be presented as a two-level distributed process that requires $RN$ processors for implementation (here $R$ means the number of elements in the set $D_r^{in}$).

Note an analogous method would apply to the general optimization problem by using the regionalization (discretization) approach. Regionalization consists of dividing the tolerance box into a finite number of non overlapping regions $D^J$, to form a grid. Then, the center or midpoint $c_j$ of each region $D^j$, is chosen to ”represent” entire region. The information on a variation of values of internal parameters can be presented as limits of their values, i.e.

$$x_{i\min} \leq x_i \leq x_{i\max}, \ i = 1, n$$

(4)

The area in space of internal parameters assigned by relations (4), represents $n$-dimensional the orthogonal parallelepiped, which we shall name as a beam of tolerances $B_\phi$:

$$B_\phi = \{x| x_{i\min} \leq x_i \leq x_{i\max}, \ i = 1, n\}$$

It is possible to define the area of acceptable values of parameters $D_x$ by methods based on multivariate exploration of tolerance region $B_\phi$. At multivariate exploration a beam (region) of tolerances $B_\phi$ can be represented by a finite number of sampling points. It is obvious, that in situations the discrete change of all parameters simultaneously is taken into account, and set of incompatible situations is sampled representation of a beam $B_\phi$. Each of situations is some sampling point representing appropriate subregion of a tolerance box (quantum-neighborhood).

For each of $R$ possible situations output parameters $Y(x)$ is computed, condition (1) is tested and discrete set of parameter nominals $D_r^{in} = \{x_r^{in}/x_r \in D_x\}$ is formed. The optimum nominal vector $x_r$ we are looking for can be found as a solution of the task (3). A second method of using parallel parallel processing techniques to maximize reliability is random search method.

On the basis of the proposed parallel methods and algorithms for region of acceptability location, modeling and discrete optimization a computer-
aided reliability-oriented distributed design (CARD) system has been developed [3]. The CARD system builds mathematical models and calculates ratings of component parameters so that achieve the highest precision, acceptability (manufacturing yield) or reliability of analog electronic circuits under design. The CARD system includes:

— the simulation module (it facilitates the use of a variety of simulation programs for electronic circuits design);
— the module for deterministic and statistical analysis;
— the module for objective function (reliability and/or manufacturing yield) calculation;
— the optimization module.

The system is organized from group of computers incorporated in a network. Such system allows using all advantages of client - server technology.

CARD system uses a widely distributed PSPICE 9.0 circuit simulation program that allows simulating a large class of analogous devices in direct current, frequency and time domains.

REFERENCES

Maximizing the volume of three-dimensional bodies on the basis submetric transformation

Alexander Afanasyev$^1$, Elena Putilina$^2$

$^1$ Institute for Systems Analysis RAS, Moscow, Russia; apa@isa.ru
$^2$ Institute for Systems Analysis RAS, Moscow, Russia; elena@isa.ru

The classical problem of maximizing the volume of three-dimensional bodies on the basis submetric transformation reduces to the of optimal control problem. Consider the following cases: bodies of revolution, cylinders, convex polyhedra.

Optimal control problems without initial conditions

Kamil Aida-zade$^1$, Yegana Ashrafova$^2$

$^1$ Azerbaijan State Oil Academy, Baku, Azerbaijan; kamil_aydazade@rambler.ru
$^2$ Cybernetics Institute of ANAS, Baku, Azerbaijan; y_aspirant@yahoo.com

One of the most important classes of problems of distribution of boundary regimes is the class of “problems without initial conditions”. If control of boundary regimes lasts long enough, then due to the friction inherent in any real physical system, the influence of initial data on the process’s behavior subsides with the course of time. Thus we naturally come to a problem without initial conditions.

Tikhonov A.N. was the first to study boundary-value problems without initial conditions for parabolic and hyperbolic equations in his work [1]. He gave the method of investigating problems without initial conditions, as well as their first rigorous solution [2]. In the well-known work [3], he investigated uniqueness of the solution to problems without initial conditions as applied to the heat conduction equation (Fourier problems).

In the present work, we investigate an optimal control problem without initial conditions, considering, as an example, the wave process arising in
hydrocarbon raw material pipeline transportation systems; we also investigate an optimal control problem for the heat conduction process without initial conditions.

Let the process be described by the following hyperbolic differential equation system [4]:

\[
\begin{align*}
\frac{\partial P(x,t)}{\partial x} &= \frac{\partial Q(x,t)}{\partial t} + aQ(x,t), \quad t \in [0,T], \; x \in [0,l], \\
\frac{\partial P(x,t)}{\partial t} &= c^2 \frac{\partial Q(x,t)}{\partial x},
\end{align*}
\]

(1)

where \(a\) is the friction coefficient; \(c\) is the velocity of sound in the environment; \((P(x,t), Q(x,t))\) is the process phase state, determined from the solution to the system (1)–(2) under the corresponding admissible value of the optimizable control vector-function \(u = (u_0(t), u_1(t))\).

Suppose that there are constraints, proceeding from technological conditions and technical requirements, on the control vector-functions, of the form

\[
u \leq u(t) \leq \bar{u}, \; t \in [0,T],
\]

(3)

Determining the process initial state belong to some admissible set of pairs of functions \(D = \{Q_0(x), P_0(x) : x \in [0,l]\}\), for each of which all the conditions of existence and uniqueness of the solution to the corresponding boundary-value problem are fulfilled.

The objective of the problem is to find such values of the boundary controls \(u_1(t), u_2(t), \; t \in (0,T]\), under which the following functional:

\[
J(u) = \frac{1}{\text{mes}D} \int_D \int_0^T \left[ \int [Q(x,T; u, Q_0, P_0) - q_T(x)]^2 + \\
[P(x,T; u, Q_0, P_0) - p_T(x)]^2 \right] \rho(Q_0) \rho(P_0) dx dQ_0 dP_0 \rightarrow \min
\]

(4)

takes its minimal value. Here \((Q(x,T; u, Q_0, P_0), P(x,T; u, Q_0, P_0))\) is the solution to the initial boundary-value problem (1), (2), and (4) for some chosen admissible initial conditions \(Q_0(x), P_0(x)\). The functional (4) determines the mean value of the deviation of the process state at \(t = T\)
from the given desired state \((q_T(x), p_T(x))\) for all possible initial conditions \((Q_0(x), P_0(x)) \in D; \rho(Q_0), \rho(P_0)\) are the density functions of the distribution of the initial values on the set \(D\). The time interval \([t_0, T]\), on which the process state does not depend on the values of the initial conditions given at \(t = 0\), plays one of the major roles in investigation of the optimal control and boundary-value problems.

We can use the method of variation of the optimizable parameters to obtain the formulas for the gradient of the functional [5].

Let \(\psi_i(x, t) = \psi_i(x, t; u, Q_0, P_0), i = 1, 2\) be the solution to the next adjoint initial boundary-value problem:

\[
\begin{aligned}
- \frac{\partial \psi_1(x,t)}{\partial x} &= \frac{\partial \psi_2(x,t)}{\partial t}, \quad x \in (0, l), \ t \in (0, T), \\
- \frac{\partial \psi_1(x,t)}{\partial t} &= c^2 \frac{\partial \psi_2(x,t)}{\partial x} - a \psi_1(x,t), \\
\psi_1(x,T) &= 2[Q(x,T) - q_T(x)], \\
\psi_2(x,T) &= 2[P(x,T) - p_T(x)], \quad x \in (0, l), \\
\psi_2(0, t) &= 0, t \in [0, T], \quad \psi_2(l, t) = 0, t \in [0, T].
\end{aligned}
\]  

(5)

(6)

(7)

Here \((P(x,T) = P(x, T; u, Q_0, P_0), \ Q(x,T) = Q(x, T; u, Q_0, P_0))\) is the solution to the initial boundary-value problem (1), (2), and (4) under any admissible \(u = u(t), Q_0 = Q_0(x), P_0 = P_0(x)\).

The formulas for the components of the gradient of the target functional with respect to the control functions \(u_0(t), u_1(t)\) are determined in the following form:

\[
\begin{aligned}
\text{grad}_{u_0(t)} J &= - \frac{1}{mes D} \int_{D} \psi_1(0, t) \rho(Q_0) \rho(P_0) dQ_0 dP_0, \ t \in [0, T], \\
\text{grad}_{u_1(t)} J &= \frac{1}{mes D} \int_{D} \psi_1(l, t) \rho(Q_0) \rho(P_0) dQ_0 dP_0, \ t \in [0, T].
\end{aligned}
\]

(8)

(9)

For numerical solution to the optimal control problem in distributed systems (1)-(5), we propose to use first order iterative optimization methods based on the application of the analytical formulas derived for the
gradient of the target functional with respect to the optimizable functions. For example, we can make use of gradient projection methods:

\[ u^{k+1} = \text{Pr}_{U}(u^k - \lambda_k \text{grad} J(u^k)), \quad k = 0, 1, \ldots, \]

or conjugate gradient projection methods [5]. Here \( u^0 = [u^0_0(t), u^0_1(t)] \) is some given initial value of the control; \( \text{grad} J(u) \) is the gradient of the target functional with respect to the optimizable vector-functions; \( \lambda_k \) is the step of one-dimensional search in the line of the anti-gradient of the target functional; \( \text{Pr}_{U}(\cdot) \) is the projection operator (this operator has a simple form for the positional constraints (3) [5]).

The formulas for the gradient of the target functional obtained above can also be used to formulate necessary optimality conditions (in the form of maximum principle in the variation form).

In the work, we will also consider the one-dimensional problem of optimal boundary control of the heating process without initial conditions. Formulas for the gradient of the target functional in this problem will be given. Results of numerical experiments of the solution to the optimal control problems will be given at the presentation.

**References**

New Nonsmooth Trust Region Method for Unconstraint Locally Lipschitz Optimization Problems

Z. Akbari\textsuperscript{1}, R. Yousefpour\textsuperscript{2}, M. R. Peyghami\textsuperscript{3}

\textsuperscript{1} Department of Mathematics, K.N. Toosi University of Technology, Tehran, Iran; z.akbari@dena.kntu.ac.ir
\textsuperscript{2} Department Mathematical Sciences, University of Mazandaran, Babolsar, Iran; yousefpour@umz.ac.ir
\textsuperscript{3} Department of Mathematics, K.N. Toosi University of Technology, Tehran, Iran; peyghami@kntu.ac.ir

Abstract

In this paper, a local model is presented for the locally Lipschitz functions. This local model is constructed by an approximation of the steepest descent direction. The steepest descent direction is an element of $\varepsilon$-subdifferential with minimal norm. In fact in the quadratic model, gradient is replaced by an approximation of the steepest descent direction. The classical trust region method is applied on this model. We prove that this algorithm is convergent by using the bounded positive definite matrices. The positive definite matrix is updated in each iterations by the BFGS method. Finally, the presented algorithm is implemented by MATLAB code.

Keywords: Trust region, Lipschitz functions, Local model, Steihaug method

Introduction

The nonsmooth unconstraint minimization problem is one of the important problems in the real world. For example in smooth case, the penalty and lagrangian functions are nonsmooth optimization problems. Also, these problems are used in control optimization. Therefore, solving these problems are attended.

The trust region (TR) method is an iterative method. In this method, the objective function is trusted by a local model. In each iteration, the model is reduced instead of objective function in the adequate region. If
$f : \mathbb{R}^n \to \mathbb{R}$ is continuously differentiable, then the local model is defined as follows

$$m(x_k, B_k)(p) = f(x_k) + \nabla f(x_k)^T p + 1/2 p^T B_k p,$$

where $B_k$ is adequately selected. If $f$ is twice continuously differentiable, then $B_k$ is the Hessian matrix. In some methods, $B_k$ is updated by the Quasi-Newton methods.

A local method, that can be practically implemented on the general local functions, is not presented. In this paper, we use the steepest descent direction to construct the local model. The steepest descent direction for the locally Lipschitz functions is an element of the Goldstein subgradient with minimal norm. Based on the method, that approximate this direction, several bundle algorithms were developed [1-6]. The efficiency of these algorithms depends on the approximation accuracy. To improve the efficiency of an algorithm, a larger number of subgradients must be computed to approximate the Goldstein subgradient efficiency and, this is time consuming. For example, in [6], the steepest descent direction is approximated by sampling gradients. This approximation is appropriate, but computing this approximation for large scale problems is very expensive. In [4], the steepest descent direction is iteratively approximated. This method computes a good approximation for the steepest descent direction by the less number of subgradients. The numerical results showed that this algorithm is more efficient than other bundle methods.

By an approximation of the steepest descent direction, we propose an quadratic model for the locally Lipschitz functions. We combine the Cauchy point and CG-Steihaug methods [7] to approximate the quadratic model solution. The numerical results show that the TR algorithm has better behavior by this combination. In this paper, we implement this algorithm by Matlab code and compare its efficiency by other methods.

**The nonsmooth trust region algorithm and its convergence**

In [8], the local model for locally Lipschitz functions is given as follow

$$m(x, p) = f(x) + \phi(x, p) + \frac{1}{2} p^T B p.$$
Based on some assumption on $\phi(.,.)$, the global convergent of TR was proved. The authors purposed the following function

$$\phi(x, p) = \max_{v \in \partial f(x)} <v, p>.$$ 

But by this definition, minimization of the local model is impractical. In this paper, we give another local model for the locally Lipschitz functions. To construct the local model for the locally Lipschitz functions, we try to substitute the gradient in (1) by a suitable element of $\partial \epsilon f(x)$.

Let $\epsilon > 0$, the steepest descent direction is computed by using $\partial \epsilon f(x)$. Consider the following function

$$v_0 = \arg \min_{v \in \partial \epsilon f(x)} \|v\|,$$ 

and let $d_0 = -\frac{v_0}{\|v_0\|}$. By Lebourg’s Mean Value Theorem, there exists $\xi \in \partial \epsilon f(x)$ such that

$$f(x + d_0) - f(x) = \epsilon \xi^T d_0 \leq -\epsilon v_0^T \frac{v_0}{\|v_0\|} = -\epsilon \|v_0\|.$$ 

In fact, $d_0$ is the steepest descent direction. But solving (3) is impractical, thus $\partial \epsilon f(x)$ is approximated by its finite subset, i.e., if $W \subset \partial \epsilon f(x)$ then $\text{conv} W$ is considered an approximation of $\partial \epsilon f(x)$. Consider the following problem

$$v_w = \arg \min_{v \in \text{conv} W} \|v\|,$$ 

let $d = -\frac{v_w}{\|v_w\|}$. If $f(x + \epsilon d) - f(x) \leq -c \epsilon \|v_w\|$ for some $c \in (0, 1)$, then $d$ can be an approximation of a steepest descent direction. Else by adding a new element of $\partial \epsilon f(x)$ in $W$, the approximation of $\partial \epsilon f(x)$ is improved. The method, how construct such a subset, is described in [4].

Suppose that $W_k \subseteq \partial \epsilon f(x_k)$ and $\text{conv} W_k$ is an approximation of $\partial \epsilon f(x_k)$. We consider the following problem

$$\|v_k\| = \arg \min_{v \in \text{conv} W_k} \|v\|,$$ 

and suppose that $f(x_k - \epsilon \frac{v_k}{\|v_k\|}) - f(x) \leq -c \epsilon \|v_k\|$ where $c \in (0, 1)$. In [4], an algorithm is presented for finding $W_k$ and $v_k$. Based on this subdifferential, $v_k \in \partial \epsilon f(x_k)$, we define the following quadratic model:

$$m(x_k, p) = f(x_k) + v_k^T p + \frac{1}{2} p^T B_k p,$$ 

27
where $B_k$ is a positive definite matrix. Based on this quadratic model, the trust region method is presented as follows.

**Algorithm 1.** (The nonsmooth trust region algorithm)

**Step 0:** Let $\Delta_0, \Delta_1 > 0$, $\theta_\Delta, \delta_1, \theta_\delta \in (0, 1)$, $x_1 \in \mathbb{R}^n$, $\xi_1 \in \partial f(x_1)$, $c_1, c_2, c_3 \in (0, 1)$, $c_4 > 1$, $B_1 = I$ and, $k = 1$.

**Step 1:** Apply Algorithm 2 in [4] at point $x_k$ with parameters $\epsilon = \Delta_k$, $\delta = \delta_k$ and $c = c_1$. Suppose Algorithm 2 in [4] finds a proper approximation of $\partial_\epsilon f(x_k)$, $\text{conv}W_k$, and a adequate subgradient, $v_k$, such that
$$v_k = \arg \min_{v \in \text{conv}W_k} \|v\|.$$

**Step 2:** If $\|v_k\| = 0$, then stop, else if $\|v_k\| \leq \delta_k$, then set $\Delta_{k+1} = \theta_\Delta \times \Delta_k$, $\delta_{k+1} = \delta_k \times \theta_\delta$, $x_{k+1} = x_k$, $k = k + 1$ and go to Step 1. Else set $\delta_{k+1} = \delta_k$ and go to Step 3.

**Step 3:** Solve the following quadratic subproblem:
$$\min_{p \in \mathbb{R}^n} m(x_k, p) = f(x_k) + v_k^T p + \frac{1}{2} p^T B_k p \quad \text{s.t.} \quad \|p\| \leq \Delta_k,$$
and set $p_k$ be its solution.

**Step 4:** If $f(x_k + p_k) - f(x_k) \leq c_1 v_k^T p_k$, then set $x_{k+1} = x_k + p_k$ and go to Step 5, else set $\Delta_{k+1} = \theta_\Delta \times \Delta_k$, $x_{k+1} = x_k$, $k = k + 1$ and go to Step 1.

**Step 5:** Define the following ratio
$$\rho_k = \frac{f(x_k + p_k) - f(x_k)}{Q(p_k) - Q(0)}.$$
If $\rho_k \geq c_3$ and $\|p_k\| = \Delta_k$ then, set $\Delta_{k+1} = \min\{\Delta_0, c_4 \times \Delta_k\}$ and, if $\rho \leq c_2$ then, set $\Delta_{k+1} = \Delta_k \times \theta_\Delta$. Else set $\Delta_{k+1} = \Delta_k$.

**Step 6:** Select a subgradient $\xi_{k+1} \in \partial f(x_{k+1})$, then update $B_k$ by the BFGS method. Set $k = k + 1$ and go to Step 1.

The following theorem proves the convergent of algorithm.
Theorem 1. Let $f : \mathbb{R}^n \rightarrow \mathbb{R}$ be a locally Lipschitz function. If the level set
\[
L := \{x : f(x) \leq f(x_1)\}
\]
is bounded, then either Algorithm 1 terminates finitely at some $k_0$ with $\|v_{k_0}\| = 0$, or the sequence $\{x_k\}$, generated by Algorithm 1, is convergent. If $x^* = \lim_{k \rightarrow \infty} x_k$, then $0 \in \partial f(x^*)$.

References

A Conjugate Gradient Projection Algorithm for systems of Large-Scale Nonlinear Monotone Equations

Keyvan Amini\textsuperscript{1}, Masoud Ahookhosh\textsuperscript{2}, Somayeh Bahrami\textsuperscript{3}

\textsuperscript{1} Department of Mathematics, Razi University, Kermanshah, Iran; kamini@razi.ac.ir
\textsuperscript{2} Department of Mathematics, Razi University, Kermanshah, Iran; aahoo.math@gmail.com
\textsuperscript{3} Department of Mathematics, Razi University, Kermanshah, Iran; Bahrami.somaye@gmail.com

Abstract

Systems of nonlinear equations generally are a family of problems that is so close to optimization problems and often arise in the applied sciences, technology and industry. In general, the system of nonlinear equations can be formulated mathematically by

\[ G(x) = 0, \quad \text{subject to} \quad x \in \mathbb{R}^n, \quad (1) \]

where \( G: \mathbb{R}^n \to \mathbb{R}^n \) is a continuous function. In particular, the nonlinear monotone equations are a class of nonlinear equations whenever \( G(x) \) satisfies the following monotonicity condition

\[ (G(x) - G(y))^T (x - y) \geq 0, \quad \text{for all } x, y \in \mathbb{R}^n, \]

guaranteeing that the solution set of (1) is a convex set.

We propose two derivative-free approaches for solving a large-scale nonlinear monotone system. The framework firstly generates a specific direction then employs a line search to construct a new point. If the new point doesn’t solve the problem, the projection technique constructs an appropriate hyperplane that separates the current iterate from the solutions of the problem. Then the projection of the new point onto the hyperplane will determine the next iterate. Thanks to the low memory requirement, we use two new conjugate gradient directions. The global convergence is established under appropriate conditions.
Software implementation of an algorithm for finding the optimal control using a graphics accelerator

Anton Anikin

1 ISDCT SB RAS, Irkutsk, Russia; anton.anikin@gmail.com

An optimal control problem with the system, linear in the control

\[ \dot{x} = f_1(x, t) + f_2(x, t) \cdot u \]  

and parallelepiped restrictions on the control action is considered. Need to find minimum of terminal functional \( \varphi(x(t_1)) \), where \([t_0, t_1]\) - the time interval of the process. In this formulation, with maintaining the regularity of the optimality conditions, controls which satisfies the Pontryagin maximum principle, have the relay character. The report proposes a method of finding the global extremum of the terminal objective function, based on the calculation of the switching points of optimal control.

Parameterization of the scalar control action that allows to construct a valid control with predetermined number of switching points is proposed. The first parametrization parameter includes not only the value of the first switching point, but also points to the border - the bottom or top, from which the constructed control begins. Because of this first parametrization parameter changes on then interval \([t_0, 2 \cdot t_1]\), the rest parameters - on the interval \([t_0, t_1]\). An example of such parametrization approach for case with 2 switching points presented on Fig. 1.

![Fig. 1](image)

Fig. 2 presents constructed relay and piecewise linear controls from some selected switching points. An algorithm for solving the optimal
control problem consists of a sequence of nonconvex unconstrained minimization problems with an increasing number of variables corresponding to the desired number of switching points. The optimal solutions of the auxiliary finite-dimensional problems make up a monotonic sequence of values converging to the minimum value of functional in problems with finite number of switching points in optimal control.

The search algorithm based on the sequential solution of nonconvex problems of one-dimensional search on a random direction is proposed to solve the unconstrained minimization problem on the hypercube. For each random direction is calculated interval of variation that would guarantee to find any solution within the allowable box and the problem is formulated as a one-dimensional search which is non-convex in the general case. Search for global minimum in the direction is performed using an algorithm based on a combination of spline-search proposed in [1], and reliable, but slow Strongin classical method [2]. An example of such global search for 2-point controls is presented on Fig. 3.

Parallelization of the algorithm performed with using CUDA technology [3] by accelerating multiple function calculation during on-dimensional searches. Such calculations require solutions of the Cauchy problem. Numerical experiments performed on Nvidia GPUs (Tesla and Fermi generation) with using a single (float) and double precision confirms the high potential of parallelism of the algorithm, results for some test problems presented in Table 1.

The authors were supported by the Russian Foundation for Basic Research (project no. 10-01-00595).
Table 1: Numerical experiment results; 1024 · 1024 integration

<table>
<thead>
<tr>
<th>Problem</th>
<th>GPU (s)</th>
<th>CPU (s)</th>
<th>CPU / GPU</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>8.6</td>
<td>149.0</td>
<td>17.3</td>
</tr>
<tr>
<td>2</td>
<td>21.0</td>
<td>383.1</td>
<td>18.2</td>
</tr>
<tr>
<td>3</td>
<td>14.6</td>
<td>241.5</td>
<td>16.5</td>
</tr>
</tbody>
</table>

REFERENCES


The reduction of the optimal parametric synthesis to the linear programming problem

Maxim Anop$^1$, Yaroslava Katueva$^2$

$^1$ Institute of Automation and Control Processes of FEB RAS, Vladivostok, Russia; manop@iacp.dvo.ru

$^2$ Institute of Automation and Control Processes of FEB RAS, Vladivostok, Russia; gloria@iacp.dvo.ru

The engineering system parameters are subject to random variations and the variations may be considered as non-stationary stochastic processes. The conventional methods for choosing parameters (parametric synthesis) generally do not take account of parameters field deviations from their design values. As a result the engineering systems designed in such a manner are not optimal in the sense of their gradual failure reliability.

Suppose that we have a system which depends on a set of $n$ input parameters $\mathbf{x} = (x_1, \ldots, x_n)^T$. The structure of the system determines the dependence of the output parameters of the internal parameters $\mathbf{y}(\mathbf{x})$. It is considered that the equations $\mathbf{y}(\mathbf{x})$ are described with a model that is given in any form such as analytical equations, algorithmic form or simulation model.

We will say that system is acceptable if $\mathbf{y}(\mathbf{x})$ satisfy the conditions (1):

$$a \leq \mathbf{y}(\mathbf{x}) \leq b,$$

where $\mathbf{y}$, $a$ and $b$ are $m$-vectors of system responses (output parameters) and their specifications, e.g. $y_1(\mathbf{x})$ – average power, $y_2(\mathbf{x})$ – delay, $y_3(\mathbf{x})$ – gain.

The inequalities (1) define a region $D_\mathbf{x}$ in the space of input (system) parameters

$$D_\mathbf{x} = \{ \mathbf{x} \in \mathbb{R}^n | a \leq \mathbf{y}(\mathbf{x}) \leq b \}$$

(2)

$D_\mathbf{x}$ is called the performance region for the system.

The engineering system parameters are subject to random variations (aging, wear, temperature) and the variations may be considered as stochastic processes:

$$X(t) = \{X_1(t), \ldots, X_n(t)\}.$$
In general parametric reliability optimization problem (optimal parametric synthesis) can be stated as follows [1].

The equations \( y(x) \), conditions of acceptability (1) and a service time \( T \) are given. The task is to find such a deterministic vector of parameter ratings (nominal values) \( x_{nom} = (x_{1nom}, x_{2nom}, \ldots, x_{nnom})^T \) that the reliability

\[
x_{nom} = \arg \max_P \{X(x_{nom}, t) \in D_x, \forall t \in [0, T]\}
\]

are maximal.

The practical algorithm of the stochastic criterion calculation is based on the conventional Monte-Carlo method [1]. In fact the distribution lows of system parameters variations and the characteristics of random parameters degradation processes \( X(t) \) are often unknown. The replacement of original stochastic criterion with a certain deterministic one is used in case of uncertainty conditions allows nearby optimum solutions to be obtained. It is a so-called a “minimal serviceability reserve”, the largest distance from the region margins and e.t.c. [1-2].

The region of acceptability, and its border are not analytically given generally. In this case, the problem (4) can be formulated as follows:

\[
x_{nom} = \arg \max (d(x_{nom}, \partial D_x), x_{nom} \in D_x)
\]

where \( d(x, \partial D_x) \) is a distance from \( x \) to the boundary \( \partial D_x \) measured with any way. The solution of the problem (5) is the center of the inscribed in the region \( D_x \) figure with maximum norm.

The modification of the simplicial approximation method offered by the S.W. Director and G.D. Hetchell [3] is discussing in the paper.

The first step in parametric synthesis problem is narrowing the search area in a space of internal parameters. A circumscribed parallelepiped is constructed for this purpose. Proposed in [4] the algorithm based on Monte-Carlo method enables to receive the points of contact with minimal \( K_i^− \) and maximal \( K_i^+ \) coordinates for each coordinate direction, belonging both to circumscribed box and region of acceptability.

The second step is the construction of the piecewise linear internal approximation \( \hat{D}_x \) for region of acceptability \( D_x \). Let’s assume that points \( p^1, \ldots, p^N \) belonging to the border of \( \partial D_x \) are received. Then the convex hull of this set \( \{p^j \in \partial D_x, j = 1, 2, \ldots, N\} \) will give required approximation \( \hat{D}_x \). It is possible to use points of a contact \( K_i^−, K_i^+, i = 1, 2, \ldots, 2^n \) as the set \( \{p^j \in \partial D_x, j = 1, \ldots, N\} \).
The convex polytope of contact’s point enables to reduce the problem (5) to a linear programming problem. The maximum figure (cube, ellipsoid) will be solution of this task. The center of it will be a required vector of nominal parameters.

The authors were supported by FEB RAS: Grants 12-I-OEMMPU -01 (Basic Research Program of PMMCP Branch RAS no.14).

REFERENCES


Boundary value games in optimal control

Anatoly Antipin

1 Computing Center of Russian Academy of Sciences, Moscow, Russia; 
asantip@yandex.ru

1. Statement of problem in finite dimensional space.

Consider a two-person game, which is the problem of computing a fixed point $x_0^* = (x_{10}^*, x_{20}^*)$ of extreme inclusions

$$x_{10}^* \in \text{Argmin}\{f_{10}(x_{10}, x_{20}^*) + \varphi_1(x_{10}) \mid C_{10}x_{10} \leq c_{10}, \ x_{10} \in X_{10}\},$$

$$x_{20}^* \in \text{Argmin}\{f_{20}(x_{10}^*, x_{20}) + \varphi_2(x_{20}) \mid C_{20}x_{20} \leq c_{20}, \ x_{20} \in X_{20}\}, \ (1)$$
where $X_{10} \subset R_1^{n_1}$ and $X_{20} \subset R_2^{n_2}$ are closed convex sets in finite dimensional Euclidean spaces. The objective functions $f_{10}(x_{10}, x_{20}) + \varphi_1(x_{10})$ and $f_{20}(x_{10}, x_{20}) + \varphi_2(x_{20})$ are defined on the Cartesian product of spaces $R_1^{n_1}$ and $R_2^{n_2}$. All functions are convex in own variables, i.e. first function is convex in the variable $x_{10}$, the second – in the variable $x_{20}$ for all $x_{10} \in X_{10}$ and $x_{20} \in X_{20}$. For the first player $x_{20} \in X_{20}$ is the parameter, for the second player $x_{10} \in X_{10}$, on the contrary, is the parameter. If both sets are compact, then there is always a solution $x^*_0 = (x^*_{10}, x^*_{20})$ of the game (1).

The meaning of the solution of this game lies in the fact that none of the players are not interested in breach of its state otherwise the value of its objective function can only increase. It seems convenient to scalarize the problem (1) and instead of the system of parametric optimization problems compute a fixed point of the extremal mapping.

For this purpose, we introduce a normalized function of the form

$$
\Phi_0(v_0, w_0) + \varphi_0(w_0) = f_{10}(z_{10}, x_{20}) + \varphi_1(z_{10}) + f_{20}(x_{10}, z_{20}) + \varphi_2(z_{20}),
$$

(2)

where $w_0 = (z_{10}, z_{20}), v_0 = (x_{10}, x_{20}), v_0, w_0 \in W_0 = X_{10} \times X_{20}$. In the new variables, two-person game with a Nash equilibrium is transformed into the problem of computing the fixed points of extremal mapping

$$
v^*_0 \in \text{Argmin}\{\Phi_0(v^*_0, w_0) + \varphi_0(w_0) \mid w_0 \in W_0\}. 
$$

(3)

By the separability of the function $\Phi(v, w)$ with respect to the variables $w_0 = (z_{10}, z_{20})$ solution of problem (3) is a solution of problem (1), but not vice versa.

If the neighborhood of the fixed point in problem (3) has a saddle structure, the saddle-point methods such as extraproximal or extragradient methods converge to the solution of this problem.

Let us consider the differential analogue of the problem (1) in functional space. This game is considered on a fixed time interval $[t_0, t_1]$ with free ends, and linear differential systems. On the sets of attainability generated by free right ends $(x_{11}, x_{21}) = (x_1(t_1), x_2(t_1))$ of the trajectories $x_1[u_1(t)], x_2[u_2(t)] = (x_1(t), x_2(t))$, the payoff functions are defined, $u_1(t), u_2(t) \in U_1 \times U_2 \subset PC[t_0, t_1]$ are set of piecewise continuous controls, $x_1(t), x_2(t) \in X_1 \times X_2 \subset PC[t_0, t_1]$ are set of piecewise continuously differentiable trajectories. A formal statement of the problem has the form:
the first player

\[ \frac{d}{dt} x_1(t) = D_1(t)x_1(t) + B_1(t)u_1(t), \quad x_{10}^* \in X_1(t_0), \quad (4) \]

\[ U_1 = \{ u_1(t) \in L^2_2[t_0, t_1] \mid u_1(t) \in [u_1^-, u_1^+] \}, \quad t_0 \leq t \leq t_1, \quad (5) \]

\[ x_{11}^* \in \text{Argmin}\{ f_1(x_{11}, x_{21}^*) + \varphi_1(x_{11}) \mid C_{11}x_{11} \leq c_{11}, \quad x_{11} \in X_1(t_1) \}, \quad (6) \]

the second player

\[ \frac{d}{dt} x_2(t) = D_2(t)x_2(t) + B_2(t)u_2(t), \quad x_{20}^* \in X_2(t_0), \quad (7) \]

\[ U_2 = \{ u_2(t) \in L^2_2[t_0, t_1] \mid u_2(t) \in [u_2^-, u_2^+] \}, \quad t_0 \leq t \leq t_1, \quad (8) \]

\[ x_{21}^* \in \text{Argmin}\{ f_2(x_{11}^*, x_{21}) + \varphi_2(x_{21}) \mid C_{21}x_{21} \leq c_{21}, \quad x_{21} \in X_2(t_1) \}, \quad (9) \]

where \( X_1(t_1) \subset R^{n_1}, X_2(t_1) \subset R^{n_2} \). Here \( x_{10}^*, x_{20}^* \) stand for the initial conditions, which form two-person game \((1)\) solution. The pair \( x_{11}^*, x_{21}^* \) is a solution of the terminal two-person game \((6),(9)\) with a Nash equilibrium.

The dynamics \((4),(5)\) and \((7),(8)\) takes the system \((4)–(9)\) from the initial state to the terminal one.

Both games are relatively independent: the interests of the players are connected only by payoff functions, but not by their dynamics. Note that the payoff functions describe the overall interest of each player: \( \varphi_1(x_{11}), \varphi_2(x_{21}) \) – interests, where the players are not going to make concessions, \( f_1(x_{11}, x_{21}), f_2(x_{11}, x_{21}) \) – interests, where the players are willing to make concessions to find a compromise.

2. The problem of calculating the fixed points of extremal mapping. Let us imagine a game of two persons \((4)–(9)\) in an aggregated form. For this purpose we introduce the new macro variables

\[ w(t) = (x_1(t), x_2(t))^T, \quad u(t) = (u_1(t), u_2(t))^T \]

and represent the controlled dynamics on the Cartesian product \( X_1(t) \times X_2(t) \) in the form

\[ \begin{pmatrix} \dot{x}_1(t) \\ \dot{x}_2(t) \end{pmatrix} = \begin{pmatrix} D_1(t) & 0 \\ 0 & D_2(t) \end{pmatrix} \begin{pmatrix} x_1(t) \\ x_2(t) \end{pmatrix} + \begin{pmatrix} B_1(t) & 0 \\ 0 & B_2(t) \end{pmatrix} \begin{pmatrix} u_1(t) \\ u_2(t) \end{pmatrix}, \]

\[ w(t) = \begin{pmatrix} x_1(t) \\ x_2(t) \end{pmatrix} \in \begin{pmatrix} X_1(t) \\ X_2(t) \end{pmatrix} = W(t), u(t) = \begin{pmatrix} u_1(t) \\ u_2(t) \end{pmatrix} \in \begin{pmatrix} U_1 \\ U_2 \end{pmatrix} = U. \]
Then we introduce the aggregate terminal variables \( w_1 = (z_{11}, z_{21})^T, v_1^* = (x_{11}^*, x_{21}^*)^T \) and the payoff function

\[
\Phi_1(v_1^*, w_1) + \varphi_1(w_1) = f_{11}(z_{11}, x_{21}^*) + f_{21}(x_{11}^*, z_{21}) + \varphi_1(z_{11}) + \varphi_2(z_{21}),
\]

and represent the aggregate terminal problem in the form

\[
\Phi_1(v_1^*, w_1) + \varphi_1(w_1) \leq \Phi_1(v_1^*, w_1) + \varphi_1(w_1),
\]

\[
C_1 = \begin{pmatrix} C_{11} & 0 \\ 0 & C_{21} \end{pmatrix} \begin{pmatrix} x_{11} \\ x_{21} \end{pmatrix} \leq \begin{pmatrix} c_{11} \\ c_{21} \end{pmatrix}; \quad \begin{pmatrix} x_{11} \\ x_{21} \end{pmatrix} \in \begin{pmatrix} X_1(t_1) \\ X_2(t_1) \end{pmatrix} = W_1.
\]

Now, the problem of calculating the fixed point \( v_1^* \in W_1 \) of extremal mappings has the form

\[
\frac{d}{dt} w(t) = D(t)w(t) + B(t)u(t), \quad v(t_0) = v_0^*, \quad u(t) \in U, \quad t_0 \leq t \leq t_1, \quad (10)
\]

\[
\Phi_1(v_1^*, v_1^*) + \varphi_1(v_1^*) \leq \Phi_1(v_1^*, w_1) + \varphi_1(w_1),
\]

\[
C_1 w_1 \leq c_1, \quad w_1 \in W_1 \subset R^{2n}. \quad (11)
\]

The extremal mapping can be written in explicit form, then

\[
\frac{d}{dt} w(t) = D(t)w(t) + B(t)u(t), \quad w(t_0) = v_0^*, \quad u(t) \in U,
\]

\[
v_1^* \in \text{Argmin}\{\Phi_1(v_1^*, w_1) + \varphi_1(w_1) \mid C_1 w_1 \leq c_1, \quad w_1 = w(t_1) \in W_1 \subset R^{2n}\}, \quad (12)
\]

Systems (10),(11) or (12) are aggregated form of the game (4)–(9). For a fixed parameter \( v_1 = v_1^* \) the resulting system is a convex programming problem, formulated in a functional space with respect to finite-dimensional \( w_1 = w(t_1) \in W_1 \) and functional \( w(t), u(t) \) variables.

In the regular case, the Lagrange function

\[
L_1(v_1^*, p_1, w_1, \psi(t), w(t), u(t)) = \Phi_1(v_1^*, w_1) + \varphi_1(w_1) + \langle p_1, C_1 w_1 - c_1 \rangle +
\]

\[
+ \int_{t_0}^{t_1} \langle \psi(t), D(t)w(t) + B(t)u(t) - \frac{d}{dt} w(t) \rangle dt
\]

defined for all \( p_1 \geq 0, w_1 \in W_1, \psi(t) \in PC[t_0, t_1], w(t) \in PC^1[t_0, t_1], u(t) \in U \subset PC[t_0, t_1], \) where \( p_1, \psi(t) \) are dual variables, and \((w(t), u(t))\) are primal variables, has a saddle point \((p_1^*, \psi^*(\cdot)), (v_1^*, v^*(\cdot)), u^*(\cdot))\). This point satisfies the system of inequalities

\[
\Phi_1(v_1^*, v_1^*) + \varphi_1(v_1^*) + \langle p_1, C_1 v_1^* - c_1 \rangle +
\]

39
\[
\int_{t_0}^{t_1} \langle \psi(\cdot), D(t)v^*(\cdot) + B(t)u^*(\cdot) - \frac{d}{dt}v^*(\cdot) \rangle dt \leq \Phi_1(v^*_1, v^*_1) + \langle p^*_1, C_1 v^*_1 - c_1 \rangle + \int_{t_0}^{t_1} \langle \psi^*(\cdot), D(t)v^*(\cdot) + B(t)u^*(\cdot) - \frac{d}{dt}v^*(\cdot) \rangle dt \leq \Phi_1(v^*_1, w_1) + \varphi_1(w_1) + \langle p^*_1, C_1 w_1 - c_1 \rangle + \int_{t_0}^{t_1} \langle \psi^*(\cdot), D(t)w(\cdot) + B(t)u(\cdot) - \frac{d}{dt}w(\cdot) \rangle dt
\]

for all \( p_1 \in \mathbb{R}^n_+ \), \( \psi(\cdot) \in PC^1[t_0, t_1] \), \( v_1, w(\cdot) \in PC^1[t_0, t_1] \), \( u(\cdot) \in U \subset PC[t_0, t_1] \).

From the resulting system of saddle-point inequalities by using relatively simple arguments, we can get the dual problem in form:

\[
\frac{d}{dt}\psi^*(t) + D^T(t)\psi^*(t) = 0, \quad \psi^*_1 = \nabla_2 \Phi(v^*_1, v^*_1) + \nabla \varphi_1(v^*_1) + C_1^T p^*_1, \\
\int_{t_0}^{t_1} \langle B^T(t)\psi^*(t), u(t) - u^*(t) \rangle dt \geq 0, \quad u(t) \in U.
\]

3. The boundary value problem and the method for its solution.

Combining the primal and dual problems, we obtain the boundary value problem

\[
\frac{d}{dt}v^*(t) = D(t)v^*(t) + B(t)u^*(t), \quad v(t_0) = v^*_0, \\
\langle p_1 - p^*_1, C_1 w^*_1 - c_1 \rangle \leq 0, \quad p_1 \geq 0, \\
\frac{d}{dt}\psi^*(t) + D^T(t)\psi^*(t) = 0, \quad \psi^*_1 = \nabla_1 \Phi(v^*_1, v^*_1) + \nabla \varphi_1(v^*_1) + C_1^T p^*_1, \\
\int_{t_0}^{t_1} \langle B^T(t)\psi^*(t), u(t) - u^*(t) \rangle dt \geq 0, \quad u(t) \in U.
\]

To solve this system, including differential equations and variational inequalities, we use the saddle-point method in the form of an extragradient process. This method can be treated as a controlled method of simple iteration. In this method, each iteration consists of two half-steps. The
first half-step can be interpreted as a control in the form of feedback. The
formulas of this iterative process have the form:

1) prediction half-step

\[
\frac{d}{dt} v^n(t) = D(t)v^n(t) + B(t)u^n(t), \quad v^n(t_0) = v_0^*,
\]

\[
\bar{p}_1^n = \pi_+ (p_1^n + \alpha(C_1 v_1^n - c_1)),
\]

\[
\frac{d}{dt} \psi^n(t) + D^T(t)\psi^n(t) = 0, \quad \psi^n_1 = \nabla_1 \Phi(v_1^n, v_1^n) + \nabla \varphi(v_1^n) + C^T_1 \bar{p}_1^n,
\]

\[
\bar{u}^n(t) = \pi_U(u^n(t) - \alpha B^T(t)\psi^n(t));
\]

(14)

2) basic half-step

\[
\frac{d}{dt} \bar{v}^n(t) = D(t)\bar{v}^n(t) + B(t)\bar{u}^n(t), \quad \bar{v}^n(t_0) = v_0^*,
\]

\[
p_{1}^{n+1} = \pi_+ (p_1^n + \alpha(C_1 \bar{v}_1^n - c_1)),
\]

\[
\frac{d}{dt} \bar{\psi}^n(t) + D^T(t)\bar{\psi}^n(t) = 0, \quad \bar{\psi}^n_1 = \nabla_1 \Phi(\bar{v}_1^n, \bar{v}_1^n) + \nabla \varphi(\bar{v}_1^n) + C^T_1 \bar{p}_1^n,
\]

\[
u^{n+1}(t) = \pi_U(u^n(t) - \alpha B^T(t)\bar{\psi}^n(t)).
\]

(15)

It follows from this process, that the differential equations are only used for the calculation of conjugate functions \(\psi^n(t), \bar{\psi}^n(t)\). Therefore, the process can be written in a more compact form

\[
\bar{p}_1^n = \pi_+ (p_1^n + \alpha(C_1 v_1^n - c_1)),
\]

\[
\bar{u}^n(t) = \pi_U(u^n(t) - \alpha B^T(t)\psi^n(t)),
\]

\[
p_{1}^{n+1} = \pi_+ (p_1^n + \alpha(C_1 \bar{v}_1^n - c_1)),
\]

\[
u^{n+1}(t) = \pi_U(u^n(t) - \alpha B^T(t)\bar{\psi}^n(t)),
\]

(16)

where \(\psi^n(t)\) and \(\bar{\psi}^n(t)\) are computed as solutions of the differential systems. It has been proven that the process converges monotonically in the norm of controls space to one of the solutions of original problem.

**Theorem.** If the set of solutions (13) is not empty and belongs to the subspace \(PC[t_0, t_1] \times PC_1[t_0, t_1]\), the functions \(\Phi_i(v_i^*, w_i) + \varphi_i(w_i), i = \)
1, 2, are positive semidefinite, and convex in the variables \( w_i \), differentiable with respect to these variables, whose gradients satisfy the Lipschitz conditions, then the sequence of approximations generated by the process (14), (15) with the choice of the parameter \( \alpha \) from the condition \( 0 < \alpha < \alpha_0 \), decreases monotonically in the norm on the Cartesian product of variables (controls, trajectories and variables of terminal problems). At the same time, any weakly converging subsequence of controls \( u^{n_i}(t) \) weakly converges to the optimal control \( u^*(t) \), and a corresponding subsequence of trajectories \( v^{n_i}(t) \) converges to optimal trajectory \( v^*(t) \) in the uniform norm \( C^n[t_0, t_1] \).

If the sequence of controls \( u^n(t) \) has a strong limit point in the norm of \( L^p_2 \), then the process \( (v^n(t), u^n(t)) \) converges to a solution \( (v^*(t), u^*(t)) \) monotonically in the norm spaces \( L^p_2 \times L^p_2 \).

In the method (14), (15) the vector \( v_0^* \) of initial conditions is used, and it must first be calculated by solving the equilibrium problem (1).

**Minimization of maximum lateness for M stations with tree topology**

Dmitry Arkhipov\(^1\), Alexander Lazarev\(^2\), Elena Musatova\(^3\)

\(^3\) Institute of Control Sciences, Moscow, Russia; miptrafter@gmail.com
\(^2\) Institute of Control Sciences, Moscow, Russia; lazarev@ipu.ru
\(^3\) Institute of Control Sciences, Moscow, Russia; nekolyap@mail.ru

**Minimization maximum weighted lateness for 2 stations.** The following problem of scheduling theory is considered. There is a set of orders (wagons) \( N \). Each order \( j \in N \) releases on the station \( A \) at the moment \( r_j \). Due date of order \( j \) equals \( d_j = r_j + \delta \). Each order has its own value \( w_j > 0 \). Wagons are delivered to the station \( B \) by train, which covers the distance between \( A \) and \( B \) in time \( p \). Each train can be departed after the time \( \alpha \) of the previous departure. Our goal is to transport all wagons on the station \( B \). The objective function is

\[
\min \max_{j \in N} (w_j L_j) \tag{1}
\]

We also formulate the ancillary problem with the objective function:

\[
\min C_{\max} |(wL)_{\max} | < y \tag{2}
\]
Schedule which holds (2) we would call $\Theta(N, y)$.

For each values $j$ and $y$ we can determine the moment $t'_j = r_j + \delta + \frac{w_j}{w_j} - p$. Order $j$ must depart from the station $A$ in the moment which belongs to the interval $[r_j, t'_j)$. We define a set of orders which must be transported to the station $B$ on the train which number is not exceeding $m$ as $S_m$. Note that on the first step of algorithm sets $S_1, \ldots, S_{m-1}$ are empty and set $S_m$ is full of orders $1, \ldots, n$.

**Property 1.** We consider the train $m$ which departs at the moment $t^m$ in the schedule $\Theta(N, y)$.

(i) If at the moment $t^m$ there are more than $k$ orders then we should transport on the train $m$ $k$ jobs with minimal $t'_j$.

(ii) Train $m$ can depart at the moment $t^m$ holds $t^m \geq \max(r_{km}, r(S_m), t^{m-1} + \alpha)$.

(iii) All orders $J_l$ which holds $t^m + \alpha \geq t'_l$ must be transported on the trains which numbers are not bigger than $m$, so $J_l \in S_m$.

(iv) When the train $m$ was departed, all orders from the set $S_m$ must had already depart.

**Algorithm 1.** On each step of algorithm we try to depart the train $m$ from the station $A$. Firstly, we choose the moment $t^m$ which holds (ii). Secondly, we choose $k$ orders which would be transported on train $m$ with help of (i). After that we check if (iii) holds. If there exists an order $J_l : r_l > t^m, t^m + \alpha \geq t'_l$, so according to (iii) we have to include this order $J_l$ into sets $S_m, S_{m+1}, \ldots, S_q$, then let us return to checking (ii). If there is no such order $J_l$, (iii) must hold, except the case when we have $x > k$ released orders from the set $S_m$ at the moment $t^m$ (this orders are .... $X^0$). We obtain $x - k$ orders from $X^0$, that have to be transported on the trains with number lower than $m$. After looking for orders $J_j, t'_j > t^m + \alpha$ in the sets $T_{m-1}, T_{m-2}, \ldots$ until we found $x - k$ orders hold this property (let the last one was founded in the train $s$). After we obtain three sets of orders: $T_s^{m-1} = T_s \cup \cdots \cup T_{m-1}, X'$ - set of jobs which holds $\{j \mid t'_j > t^m + \alpha, j \in T_s^{m-1}\}$ with minimal $x$ moments $t'$ from all such jobs $j$, and a set $X^0$. Let us consider the set $X = (T_s^{m-1} \setminus X') \cup X^0$. Orders from this set must be transported on trains $s, \ldots, m$. When we depart this orders we have to change $r_{ki}$ on $r(X_{(i-s+1)k})$ in the property (ii) because we can depart only orders from $X$. Property (i) holds because all orders which are not belongs to $X$ and released until this moment holds $t' > t^m + \alpha$. Properties (iii) and (iv) hold, except the cases when one of
sets $S_i$ changes, if we face it we should go to the next step - considering the train $i$. If we don’t face problems during the transporting set $X$ we should go to the next step - considering the train $m + 1$. This algorithm terminates if on any step we obtain the set $S_i$ with more than $k_i$ orders.

**Theorem 1.** Algorithm 1 constructs the schedule $\Theta(N,y)$ according to criterion $C_{\text{max}}|(wL)_{\text{max}} < y$. If algorithm 1 was terminated, there are no schedule $\pi$ holds $(wL)_{\text{max}} < y$.

**Lemma.** If there are exist two schedules $\Theta(N,y_1)$ and $\Theta(N,y_2)$ ($y_1 > y_2$) constructed with help of the algorithm 1, then for each $i = \{1,\ldots,q\}$ and a pair of sets $S_i(\Theta(N,y_1))$ and $S_i(\Theta(N,y_2))$ holds $S_i(\Theta(N,y_1)) \subseteq S_i(\Theta(N,y_2))$

**Algorithm 2.** Firstly we construct a schedule in which each train departs as soon as possible. Then we consider the order $j$ with maximal $w_jL_j$. Order $j$ transports on the train $m$. If we want to improve the objective function we must transport order $j$ on the train which number is lower than $m$, so $j \in S_{m-1}$. On the next step we construct the schedule $\Theta(N,w_jL_j)$. We should repeat this operation until we construct the schedule $\Theta(N,y')$ with the objective function $y_0$, when schedule $\Theta(N,y')$ doesn’t exist. On this step we note that $\Theta(N,y')$ is an optimal schedule with the objective function $y_0$.

**Theorem 2.** Algorithm 2 constructs the schedule $\pi$ which is optimal according to criterion $(wL)_{\text{max}}$ and has minimal $C_{\text{max}}$ among all schedules with the objective function $(wL)_{\text{max}}$.

**Minimization maximum lateness for 3 stations.** There are three stations $A, B, C$ and three sets of orders $N^{AB}, N^{AC}, N^{BC}$. The order $j \in N^{AC}$ releases at the moment $r^A_C$ on the station $A$ and must be transported to the station $C$. The due date of this order we define as $d^A_C = r^A_C + \delta^{AB} + \delta^{BC}$. Parameters of other orders defines similarly. Train covers the distance between $A$ and $B$ in time $p^{AB}$ and the distance between $B$ and $C$ in time $p^{BC}$. There are $k$ wagons in each train. Each train can be departed after time $\alpha$ of the previous departure. Let us suppose that $\delta^{AB} > p^{AB} + \alpha$ and $\delta^{BC} > p^{BC}$. The objective function is

$$\min\left(\max_{j \in N^{AB} \cup N^{BC} \cup N^{AC}} (L_j)\right)$$

We also formulate the ancillary problem with the objective function

$$\min(C_{\text{max}}) | L_{\text{max}} < y$$

44
The schedule which holds (4) we would call \( \Theta^3(N, y)\), \( N = N^{AB} \cup N^{AC} \cup N^{BC} \).

**Algorithm 3.** Firstly, we construct intervals for each order and each track \((AB\text{ and } BC)\). The order \( j \in N^{AB} \) must be transported on \( B \) before the moment \( r^{AB}_j + \delta^{AB} + y \), so we obtain that it’s interval on the track \( AB \) is \([r^{AB}_j, r^{AB}_j + \delta^{AB} + y - p^{AB}]\). Order \( j \in N^{BC} \), corresponds with the interval \([r^{BC}_j, r^{BC}_j + \delta^{BC} + y - p^{BC}]\) on the track \( BC \). The order \( j \in N^{AC} \) corresponds with the interval \([r^{AC}_j, r^{AC}_j + \delta^{AB} + \delta^{BC} + y - p^{AB} - p^{BC}]\) on the track \( AB \), because this order must be departed from \( B \) before the moment \( r^{AC}_j + p^{AB} \). We also obtain that order \( j \) corresponds with the interval \([r^{AC}_j + p^{AB}, r^{AC}_j + \delta^{AB} + \delta^{BC} + y - p^{BC}]\) on the track \( BC \). Each job must start it’s transportation on the track in time belongs to the interval which is corresponds with this track. When all interval are constructed we use the algorithm 1 to construct the schedule \( \Theta(N^{AB\cup AC}, y) \) corresponds with the track \( AB \) and the schedule \( \Theta(N^{BC\cup AC}, y) \) corresponds with the track \( BC \), subject to constructed intervals. If one of this schedules isn’t constructed successfully our algorithm terminates. If not we pay attention for ”bad” orders \( j \) which completion time \( C_j \) in the schedule \( \Theta(N^{AB\cup AC}, y) \) more than it’s time of departure in the schedule \( \Theta(N^{BC\cup AC}, y) \). If there are no ”bad” orders we should depart orders on each track according to it’s own schedule \( \Theta \). If there are exist some ”bad” orders we should find the first of them. We can use two methods to get out of the ”bad” order.

**Method 1.** We don’t change moments of departure of trains on the track \( AB \).
**Method 2.** We change the moments of departure of trains on the track \( AB \).

We use the following scheme to choose the right method.

If there are no "bad" orders in our pair of schedules \( \Theta(N_{AB \cup AC}, y) \) and \( \Theta(N_{BC \cup AC}, y) \), then we depart trains on the track \( AB \) according to \( \Theta(N_{AB \cup AC}, y) \) and on the track \( BC \) according to \( \Theta(N_{BC \cup AC}, y) \). As a result we obtain the schedule \( \Theta^3(N, y) \).

**Theorem 3.** Algorithm 3 constructs the schedule \( \Theta^3(N, y) \). If algorithm 3 terminates then there is no schedule which holds (4).

To construct the optimal schedule we use algorithm 2. The only difference is that we should use the schedule \( \Theta^3(N, y) \) instead of \( \Theta(N, y) \).

**M station with tree topology.** The formulation of this problem and the problem for 3 stations is the same. The only difference is that we deals with \( M \) stations with tree topology. Due to the tree topology there is only one way between each pair of stations. We also can enumerate stations from left to right (or from right to left).

**Algorithm 4.** We use algorithm 3 to get out of "bad" orders for each
station, from the left to the right follows the numeration (according to chosen direction of the train moving). When there are no “bad” orders on each station we obtain the schedule \( \Theta^M(N, y) \). After that we use algorithm 2 to construct the optimal schedule for \( M \) stations.

**Theorem 4.** Algorithm 4 constructs the optimal schedule according to criterion \( L_{\text{max}} \) in \( O(M^2n^4) \) operations.

The authors were supported by the Russian Foundation for Basic Research (project no. 11-08-13121).

---

**Existence theorems for elliptic equations in unbounded domains**

**Armen Beklaryan\(^1\)**

\(^1\) Lomonosov Moscow State University, Moscow, Russia; beklaryan@hotmail.com

We consider the first boundary value problem for elliptic systems defined in unbounded domains, which solutions satisfy the condition of finiteness of the Dirichlet integral also called the energy integral

\[
\int_\Omega |\nabla u|^2 dx < \infty.
\]

**Basic concepts**

Let \( \Omega \) is an arbitrary open set in \( \mathbb{R}^n \). As is usual, by \( W_{2, \text{loc}}^1(\Omega) \) we denote the space of functions which are locally Sobolev, i.e.

\[
W_{2, \text{loc}}^1(\Omega) = \{ f : f \in W_2^1(\Omega \cap B_\rho^x), \forall \rho > 0, \forall x \in \mathbb{R}^n \},
\]

where \( B_\rho^x \) – open ball with center at point \( x \) and with radius \( \rho \). If \( x = 0 \) then we will write \( B_\rho \). We will denote by \( \overline{W}_{2, \text{loc}}^1(\Omega) \) set of functions from \( W_{2, \text{loc}}^1(\mathbb{R}^n) \), which is the closure of \( C_0^\infty(\Omega) \) in the system of seminorms.
∥u∥_{W^1_2(Ω)}, where K ⊂ R^n are various compacts. Let denote by L^1_2(Ω) a space of generalized functions in Ω, which first derivatives belong to L^2_2(Ω) [4], in other words

\[ L^1_2(Ω) = \{ f \in D'(Ω) : \int_Ω |∇f|^2 dx < ∞ \}. \]

Let ω ⊂ R^n is an open set, K ⊂ ω is a compact. We will denote by \( Φ_ϕ(K,ω) \) the set of functions \( ψ \in C^\infty_0(ω) \) such that \( ψ = ϕ \) in the neighborhood of \( K \), or in other words \( ψ - ϕ \in W^1_2,loc(\mathbb{R}^n \setminus K) \).

Let’s define a capacitance of a compact \( K \) relative to the set \( ω \) [4]:

\[ \text{cap}_ϕ(K,ω) = \inf_{ψ \in Φ_ϕ(K,ω)} \int_ω |∇ψ|^2 dx. \]

The capacitance of arbitrary closed set \( E \subset ω \) in \( R^n \) is defined by the formula \( \text{cap}_ϕ(E,ω) = \sup_{K \subset E} \text{cap}_ϕ(K,ω) \). If \( ω = R^n \), then instead of \( \text{cap}_ϕ(E,R^n) \) we will write \( \text{cap}_ϕ(E) \).

**Problem statement**

Let \( L \) is a divergent operator

\[ L = \sum_{i,j=1}^n \frac{∂}{∂x_i} \left( a_{ij}(x) \frac{∂}{∂x_j} \right), \]

where \( a_{ij} \) are bounded measurable functions in \( R^n \) satisfying condition

\[ γ|ξ|^2 \leq \sum_{i,j=1}^n a_{ij}(x) ξ_i ξ_j, \quad ξ \in R^n, γ > 0. \]

The solution of the Dirichlet problem

\[
\begin{cases}
Lu = 0 \text{ in } Ω \\
u|_{∂Ω} = ϕ,
\end{cases}
\]

(1)

where \( ϕ \in W^1_{2,loc}(\mathbb{R}^n) \), is a function \( u \in W^1_{2,loc}(Ω) \) such that:
1) \( u - \varphi \in W^{1,\text{loc}}_2(\Omega) \), i.e. \((u - \varphi)\mu \in \dot{W}^1_2(\Omega)\) for any function \( \mu \in C_0^\infty(\mathbb{R}^n) \);

2) function \( u \) has bounded Dirichlet integral

\[
\int_{\Omega} |\nabla u|^2 dx < \infty;
\]

3) \[
\int_{\Omega} \sum_{i,j=1}^{n} a_{ij}(x) \frac{\partial u}{\partial x_j} \frac{\partial \psi}{\partial x_i} dx = 0
\]

for any function \( \psi \in C_0^\infty(\Omega) \).

**Basic results**

**Theorem 1.** Let’s \( \text{cap} \varphi - c(\mathbb{R}^n \setminus \Omega) < \infty \) for some constant \( c \in \mathbb{R}^n \). Then the problem (1) has a solution.

**Theorem 2.** Let the problem (1) has a solution and it is true that \( \int_{\mathbb{R}^n \setminus \Omega} |\nabla \varphi|^2 dx < \infty \).

Then there is such constant \( c \in \mathbb{R}^n \), that \( \text{cap} \varphi - c(\mathbb{R}^n \setminus \Omega) < \infty \).

**Theorem 3.** Let \( n \geq 3 \). Then \( \text{cap} \varphi - c(\mathbb{R}^n \setminus \Omega) < \infty \) if and only if

\[
\sum_{k=N}^{\infty} \text{cap} \varphi - c((B_{2k+1} \setminus B_{2k-1}) \cap (\mathbb{R}^n \setminus \Omega), B_{2k+2} \setminus \overline{B}_{2k-2}) < \infty
\]

for some \( N \in \mathbb{N} \).

**Particular cases**

Let consider the space \( \mathbb{R}^n \) with a set of coordinates \((x_1, x_2, \ldots, x_n)\) and let \( \varphi_\alpha = (1 + |x_1|)^\alpha \). Domain \( \Omega_{1,i} \) is upper half-plane relative to \( x_i \), where \( i \neq 1 \), in other words \( \Omega_{1,i} = \{(x_1, x_2, \ldots, x_n)|x_i \geq 0, i \neq 1\} \). Domain \( \Omega_2 \) is the outer part of the space formed by surface of revolution relative to \( x_1 \) of the curve from Fig.1.
\[ x_2 = |x_1|^\beta, \beta < 0 \]

**Fig. 1:** Domain \( \Omega_2 \)

**Corollary 1.** Let \( n \geq 2 \). Then for the domain \( \Omega_{1,i} \) and for bounded function \( \varphi_\alpha \) the existence of solutions of the problem (1) is equivalent to either an inequality \( \alpha < -\frac{1}{2} \) or \( \alpha = 0 \).

**Corollary 2.** Let \( n \geq 3 \). Then for the domain \( \Omega_2 \) and for bounded function \( \varphi_\alpha \) the existence of solutions of the problem (1) is equivalent to either an inequality \( \alpha < -\frac{1 + \beta(n - 3)}{2} \) or \( \alpha = 0 \).

**References**

On the determination of the earthquake slip distribution via linear programming techniques

Vladimir Bushenkov\textsuperscript{1}, Bento Caldeira\textsuperscript{2}, Georgi Smirnov\textsuperscript{3}

\textsuperscript{1} University of Évora, Évora, Portugal; bushen@uevora.pt
\textsuperscript{2} University of Évora, Évora, Portugal; bafcc@uevora.pt
\textsuperscript{3} University of Minho, Braga, Portugal; smirnov@math.uminho.pt

The description that one can have of the seismic source is the manifestation of an imagined model, obviously outlined from Physic Theories and supported by mathematical methods. In that context, the modelling of earthquake rupture consists in finding values of the parameters of the selected physics-mathematical model, through which it becomes possible to reproduce numerically the records of earthquake effects on the Earths surface. Actually, these effects are the elastic records at near field source and at far field source, and inelastic deformations recorded by geodetic techniques. The detail and accuracy level, with which the characteristic parameters for large earthquakes are computed, depends on the combination of two factors - the applied methods and the used data.

Under the hypothesis of constant slip direction and constant rise time of individual source time function, the problem of complete seismic slip time history and distribution reconstruction reduces to the solution of a system of linear equations. It is well-known that this inverse problem is ill-posed [6]. The usual regularization techniques [8] can hardly be applied in this case because of a very high dimension of this problem (see, e.g., [3]). The problem can be overcome by introducing some additional regularizing constraints. Some additional physical hypotheses, like no-backslip constraint, result in condition of non-negativeness of solutions to the system of linear equations.

The positivity that prohibits negative seismic moment values, is a constraint naturally assumed when used the Non Negative Least Squares algorithm (NNLS) [5] to inverts seismic waveforms to slip distribution (e.g., [7]).

We present and test a Linear Programming (LP) inversion in dual form, for reconstructing the kinematics of the rupture of large earthquakes through space-time seismic slip distribution on finite faults planes. The proposed method can be considered as a continuation of the work started
in [2]. The proposed algorithm uses strong ground motion waveforms, but it can also be used with other types of data as teleseismic waveforms as well as with geodesic data (static deformation). We test the method with data obtained by application to a synthetic model of rupture. To compare it, we rehearsed reconstructions with same data, but made by other strongly used algorithms. Green functions (see, e.g., [1]) were calculated by a finite differences method applied to a 3D structure model [4].

The hypothesis of constant slip direction in general is not verified and the "real" seismic slip time history and distribution reconstruction becomes an hard nonlinear problem. In this work we suggest an algorithm for seismic slip time history and distribution reconstruction allowing to solve the problem in its general setting. The solution of an auxiliary linear programming problem is an essential part of the developed method. To test the algorithm we use a synthetic displacement function for the fault model and perform the inversion.

The slip determination problem can be formalized in the frame of mathematical programming in the following way

\[ \langle c, x \rangle \rightarrow \min, \]
\[ A(\lambda)x = b, \]
\[ x \geq 0. \] (1)

Here \( x \) is the unknown vector of amplitudes and residuals (see [2]) and the vector \( \lambda \) represents the unknown rakes. Note that the displacement field models can be different but the mathematical formalization is always the same. If we fix the rake vector \( \lambda \), problem (1) becomes a linear programming problem. This observation is the key to an effective solution of problem (1). It turns out that the gradient of the minimized functional \( \langle c, x \rangle \) with respect to \( \lambda \) can be calculated in terms of the solution to the linear programming problem dual to (1).

The following algorithm describes the process.

**Algorithm:**

Given \( \lambda_0, \Delta > 0, \) and \( \epsilon > 0, \)

\[ \text{for } k = 0, 1, 2, \ldots \]

**Step 1.** Solve linear programming problem (1) with \( \lambda = \lambda_k \) and obtain \( x_k. \)
Step 2. Obtain search direction $\lambda_k$ and a step $\delta_k > 0$.

    if $\delta_k \| \lambda_k \| < \epsilon$ break

else

Step 3. Set $\lambda_{k+1} = \lambda_k + \delta_k \bar{\lambda}_k$.

end (for)

The second step of the algorithm is not trivial. The derivative is calculated using the dual linear programming problem. The latter has a very specific form:

$$
\langle c, x \rangle \rightarrow \min,
Bx \leq 0,
-1 \leq (x)_i \leq 1, \ i = 1, n,
$$

(2)

where $B$ is an $(m \times n)$-matrix with $m < n$. This special structure of the dual problem allows one to effectively find an admissible vertex.

This research is supported by the Portuguese Foundation for Science and Technologies (FCT), through the project PTDC/CTE-GIN/82704/2006.

References

New Integer Programming formulations for the Kidney Exchange Problem

Miguel Constantino\textsuperscript{1}, Xenia Klimentova\textsuperscript{2}, Ana Viana\textsuperscript{3}

\textsuperscript{1} Faculdade de Ci\'encias da Universidade de Lisboa, Lisboa, Portugal; mfconstantino@fc.ul.pt
\textsuperscript{2,3} Instituto de Engenharia de Sistemas e Computadores, Porto, Portugal; xenia.klimentova@inescporto.pt, aviana@inescporto.pt

Problem statement

Kidneys are the most demanded organs for transplantation, but finding a suitable kidney can be difficult because of their scarcity and of blood and/or tissue incompatibility between donor and patient. In recent years Kidney Exchange Programs [1,2,4,5,6,7] brought a new opportunity for patients who have someone willing to donate him/her a kidney but, because they are not physiologically compatible, the transplantation cannot be performed. The Kidney Exchange Problem (KEP) appears within the frame of these programs.

The problem can be represented as a directed graph $G(V,A)$, where the set of vertices $V$ is the set of incompatible donor-patient pairs. Two vertices $i$ and $j$ are connected by arc $(i,j) \in A$ if a donor from pair $i$ is compatible with the patient of pair $j$. A weight $w_{ij}$ can be associated to each arc; in the simplest case $w_{ij}$ equals to 0 or 1 $\forall (i,j) \in A$. A feasible “kidney exchange” in the whole graph is defined by a set of vertex-disjoint cycles and the objective associated to KEP is to determine the exchange that maximizes the overall number of transplants.

If there is no bound on the number of pairs that can be involved in a cycle, the problem turns into an assignment problem and can be solved in polynomial time. However, this problem is not of practical interest. First, because all operations in a cycle have to be performed simultaneously, the number of personnel and facilities needed for simultaneous operations bring several logistics problems. Second, because last-minute testing of donor and patient can elicit new incompatibilities that were not detected before causing the donation and all possible exchanges in that cycle to be canceled, it is preferred that cycles are of limited size. Summarizing the KEP can be formulated as follows:
Find a maximum weight packing of vertex-disjoint cycles with length at most \( k \).

In case \( k = 2 \) the problem is a maximum cardinality matching problem which can be solved in polynomial time [2,9], but when \( k \geq 3 \) and bounded the problem is NP-complete [1,7].

**Extended edge formulations – new compact formulations**

There are two known Integer Programming (IP) models which have been proposed independently in [1] and [10] — the so called edge and cycle formulations. Unfortunately, both of these formulations have exponential number constraints or variables, which can become a bottleneck for large scale problems defined on high density graphs. In [1] Column Generation with Branch-and-Bound is implemented for the cycle formulation. The authors report very good results and claim that they where able to solve problems with up to 10000 vertices in a graph. They also show that the cycle formulation dominates the edge formulation. This paper presents a new compact formulations for the problem, i.e. a formulations where both the number of variables and constraints grows polynomially with the number of pairs \(|V|\).

In the first formulation, we consider edge formulation variables \( x_{ij} \) for each arc \((i,j) \in A\) in the graph \( G = (V,A)\): \( x_{ij} = 1 \) if donor of pair \( i \) gives the kidney to a patient of pair \( j \). Additional assignment variables are used to prevent exponential number of constraints in the edge formulation. Let \( L \) be an upper bound on the number of cycles in any solution. Such an upper bound is e.g. \( L = |V| \), because each vertex can participate in at most one cycle, but more refined bounds could be considered. The cycles in the solution can be represented by an index \( l \) with \( 1 \leq l \leq L \). Consider the following assignment variables: \( y_{li}^{l} = 1 \) if and only if node \( i \) belongs to cycle \( l \).

To avoid multiplicity of solutions an order is imposed in cycles: a cycle \( l \) in the solution must have node \( l \), and any other nodes must have an index larger than \( l \). Moreover, variables \( y_{li}^{l} \) can be eliminated when \( l \) and \( i \) cannot be in the same cycle. Thus let \( d_{ij}^{l} \) denote the shortest path (in terms of number of arcs) between vertices \( i \in V \) and \( j \in V \) in graph \( G \), where \( i,j \geq l \). For each vertex \( l \in V \), build the set of vertices \( V^{l} = \{i \in V| i \geq l \land d_{li}^{l}+d_{il}^{l} \leq k\} \). Denote by \( L \) the set of indices \( l \) such that \( V^{l} \neq \emptyset \). One can write the following formulation with polynomial number of constraints.
\[
\max \sum_{(i,j) \in A} w_{ij} x_{ij} \tag{1}
\]
\[
\sum_{j: (j,i) \in A} x_{ji} = \sum_{j: (i,j) \in A} x_{ij} \quad \forall i \in V \tag{2}
\]
\[
\sum_{j: (i,j) \in A} x_{ij} \leq 1 \quad \forall i \in V \tag{3}
\]
\[
\sum_{i \in V^l} y_{i}^l \leq k \quad \forall l \in \mathcal{L} \tag{4}
\]
\[
\sum_{l \in \mathcal{L}} y_{i}^l = \sum_{j: (i,j) \in A} x_{ij} \quad \forall i \in V^l \tag{5}
\]
\[
y_{i}^l + x_{ij} \leq 1 + y_{j}^l \quad \forall (i,j) \in A, \quad \forall l \in \mathcal{L} \tag{6}
\]
\[
y_{i}^l \leq y_{i}^l \quad \forall i \in V^l, \quad l \in \mathcal{L} \tag{7}
\]
\[
x_{ij}, y_{i}^l \in \{0, 1\} \quad \forall i \in V^l, \quad \forall l \in \mathcal{L}, \quad \forall (i,j) \in A
\]

The objective function (1) maximizes the weighted sum of the exchange. Constraints (2) assure that the number of kidneys received by patient \( i \) is equal to the number of kidneys given by donor \( i \). Constraints (3) guarantee that a donor can only donate one kidney. Constraints (4) model cardinality of feasible cycles. Constraints (5) ensure that node \( i \) is in a cycle \( \left( \sum_{j: (i,j) \in A} x_{ij} = 1 \right) \) if and only if there is an assignment of \( i \) to some \( l \left( \sum_{l} y_{i}^l = 1 \right) \). Constraints (6) state that if node \( i \) is in cycle \( l \left( y_{i}^l = 1 \right) \) and donor \( i \) gives a kidney to recipient \( j \left( x_{ij} = 1 \right) \) then node \( j \) must also be in cycle \( l \left( y_{j}^l = 1 \right) \). Constraints (7) guarantee vertex \( l \) to appear in cycle \( l \) in the solution.

In the second formulation we consider the same idea of representing cycles by extra indices \( l \) but they are introduced in the decision variables \( x_{ij} \). The problem variables will be \( x_{ij}^l \) if arc \( (i,j) \) is selected to be in a cycle \( l \). In addition to the elimination procedures for variables \( y_{i}^l \) which were implemented above, one can also eliminate variables \( x_{ij}^l \). If there is no cycle of size at most \( k \) containing node \( l \) and an arc \( (i,j) \), \( \forall j \), with \( l < i, j \), then variable \( x_{ij}^l \) can be set to zero or eliminated from the model.

Summarizing, the application of the elimination procedures lead to the construction of a subgraph \( G^l = (V^l, A^l) \) for each index \( l \in \mathcal{L} \), where \( V^l \)
was defined above and $A_l = \{(i,j) \in A \mid i,j \in V^l$ and $d(l,i) + 1 + d(j,l) \leq k\}$. Then the whole reduced graph is $\bar{G} = (\bar{V}, \bar{A})$, where $\bar{V} = \bigcup_{l \in \mathcal{L}} V^l$, $\bar{A} = \bigcup_{l \in \mathcal{L}} A^l$. With this notation one can write the reduced extended edge formulation as follows ($x^l_{ij}$ will be built over $\bar{A}$).

$$\max \sum_{(i,j) \in \bar{A}} w_{ij} x^l_{ij},$$

(3)

$$\sum_{j:(j,i) \in A^l} x^l_{ji} = \sum_{j:(i,j) \in A^l} x^l_{ij} \quad \forall i \in V^l, \forall l \in \mathcal{L},$$

(4)

$$\sum_{l \in \mathcal{L}} \sum_{i:(i,j) \in A^l} x^l_{ij} \leq 1 \quad \forall j \in \bar{V},$$

(5)

$$\sum_{(i,j) \in A^l} x^l_{ij} \leq k \quad \forall l \in \mathcal{L},$$

(6)

$$\sum_{j:(i,j) \in A^l} x^l_{ij} \leq \sum_{j:(l,j) \in A^l} x^l_{ij} \quad \forall i \in V^l, \forall l \in \mathcal{L},$$

(7)

$$x^l_{ij} \in \{0, 1\} \quad \forall (i, j) \in A, \forall l \in \mathcal{L}$$

The objective (3) is to maximize the total weight of arcs in the set of all subgraphs of the graph. Constraints (4) guarantee that in each cycle $l$ the number of kidneys received by patient $i$ is equal to number of kidneys given by donor $i$. Constraints (5) make sure a donor donates only once. Constraints (6) state that in each cycle $l$ a maximum number of $k$ edges is allowed. Constraints (7) ensure that whenever an arc $(i, j)$ is in cycle $l$, there is a cycle containing a node with index $l$.

**Conclusions**

Computational experiments where carried out to compare the proposed and know formulations in terms of time needed to find an optimal solution and of the integrality gaps for upper bounds of linear relaxations of the models. CPU times and bounds were obtained with CPLEX 12.1. A wide variety of test instances was considered, with different graph densities. According to the results we can conclude that the non-compact cycle formulation can be very effective for low density graphs with small values
of $k$. However the importance of considering and using compact formulations becomes obvious for larger values of $k$, especially for problems with denser graphs.

This work is financed by the ERDF European Regional Development Fund through the COMPETE Programme (operational programme for competitiveness) and by National Funds through the FCT Fundação para a Ciência e a Tecnologia (Portuguese Foundation for Science and Technology) within project "KEP - New models for enhancing the kidney transplantation process /FCT ref: PTDC/EGE-GES/110940/2009."

REFERENCES

Identification of parameters in model of water transfer in soil

V.V. Dikusar¹, E.S. Zasukhina²

¹ Computing Center RAS, Moscow, Russia; dikussar@ccas.ru
² Computing Center RAS, Moscow, Russia; zasuhina@ccas.ru

A one-dimensional model of vertical water transfer in soil is considered. We assume that soil is an isothermal porous homogeneous medium. In that case a water transfer can be described by one-dimensional nonlinear parabolic equation.

Consider following initial-boundary value problem:

\[
\frac{\partial \theta}{\partial t} = \frac{\partial}{\partial z} \left( D(\theta) \frac{\partial \theta}{\partial z} \right) - \frac{\partial K}{\partial z}, \quad (z,t) \in Q,
\]

\[
\theta(z,0) = \varphi(z), \quad z \in (0,L),
\]

\[
\theta(L,t) = \psi(t), \quad t \in (0,T),
\]

\[
\left. \frac{\partial \theta}{\partial t} \right|_{z=0} = \left( D(\theta) \frac{\partial \theta}{\partial z} - K \right) \bigg|_{z=0} + R(t) - E(t), \quad t \in (0,T),
\]

\[
\theta_{\min} \leq \theta(0,t) \leq \theta_{\max}, \quad t \in (0,T).
\]

Here, \( z \) is the coordinate; \( t \) is time; \( \theta(z,t) \) is humidity at point \((z,t)\); \( Q = (0,L) \times (0,T) \); \( \varphi(z) \) and \( \psi(t) \) are given functions; \( D(\theta) \) is the coefficient of diffusion; \( K(\theta) \) is the hydraulic conductivity; \( R(t) \) is precipitation; \( E(t) \) is evaporation; \( \theta_{\min} \) and \( \theta_{\max} \) are minimal and maximal values of humidity respectively.

According to [1]-[2] \( D(\theta) \) and \( K(\theta) \) are given in the form:

\[
K(\theta) = K_0 S^{0.5} \left[ 1 - \left( 1 - S^{1/m} \right)^m \right]^2,
\]

\[
D(\theta) = K_0 \frac{1 - m}{\alpha m (\theta_{\max} - \theta_{\min})} S^{0.5 - 1/m} \times
\]

\[
\times \left[ (1 - S^{1/m})^{-m} + (1 - S^{1/m})^{-m} - 2 \right],
\]

where \( S = \frac{\theta - \theta_{\min}}{\theta_{\max} - \theta_{\min}} \), and \( K_0, \alpha, m \) are some parameters.

We call this problem a direct problem.

Sometimes it is difficult and expensive to determine the parameters \( \alpha \) and \( m \) experimentally. Our aim is to determine them as a result of solution of inverse problem.
Suppose that a function \( \hat{\theta}(z,t) \) is defined in \( Q_0 \subset Q \). This function can be interpreted as the experimental data. Suppose that the set \( Q_0 \) consists of finite number of elements \( q, q = (z,t) \). Consider following problem: determine \( \alpha \) and \( m \) so that the solution \( \theta(z,t) \) of the direct problem is close to given function \( \hat{\theta}(z,t) \) in \( Q_0 \). More precisely, the optimal control problem is to determine \( u_{opt} = (\alpha_{opt}, m_{opt}) \) and corresponding optimal solution \( \theta_{opt}(z,t) \) such that functional

\[
J(u) = \frac{1}{2} \sum_{q \in Q_0} \xi(q) \cdot [\theta(q) - \hat{\theta}(q)]^2
\]

is minimized. Here \( \xi(q), q \in Q_0 \), are weight multipliers such that

\[
\xi(q) \geq 0, \quad q \in Q_0, \quad \sum_{q \in Q_0} \xi(q) = 1.
\]

In order to solve optimal control problem numerically it is necessary to pass to its discrete analogue. Let us divide the time interval \( (0, T) \) and the space interval \( (0, L) \) into \( N \) and \( I \) subintervals respectively. The lengths of these subintervals are \( \tau = T/N \) and \( h = L/I \) respectively. Consider following finite-difference approximation of the direct problem:

\[
\frac{\theta_{i}^{n+1} - \theta_{i}^{n}}{\tau} = \frac{1}{h} \left( D_{i+1/2}^{n} \frac{\theta_{i+1}^{n+1} - \theta_{i}^{n+1}}{h} - K_{i+1/2}^{n} - D_{i-1/2}^{n} \frac{\theta_{i}^{n+1} - \theta_{i-1}^{n+1}}{h} + K_{i-1/2}^{n} \right),
\]

\[1 < i < I, \quad 0 \leq n < N,\]

\[\theta_{0}^{0} = \varphi_{i}, \quad 0 \leq i \leq I, \quad \theta_{I}^{n} = \psi^{n}, \quad 1 \leq n \leq N,\]

where \( \theta_{i}^{n}, D_{i+1/2}^{n}, K_{i-1/2}^{n} \) are values of the functions \( \theta(z,t), D(\theta(z,t)), K(\theta(z,t)) \) at the points \((ih, n\tau), ((i + 1/2)h, n\tau), ((i - 1/2)h, n\tau)\) respectively. On the left boundary

\[
\frac{\theta_{0}^{n+1} - \theta_{0}^{n}}{\tau} = \frac{2}{h} \left( D_{1/2}^{n} \frac{\theta_{1}^{n+1} - \theta_{0}^{n+1}}{h} - K_{1/2}^{n} + R^{n+1} - E^{n+1} \right),
\]

\[0 \leq n < N.\]

Here \( R^{n+1}, E^{n+1} \) are values of \( R(t) \) and \( E(t) \) at \( t = (n+1)\tau \).
Then the discrete analogue of the direct problem has the form:

\[
\Phi_n^0 = -\left(\frac{1}{\tau} + \frac{2}{h} D_{1/2}^{n-1}\right) \theta_0^n + \frac{2}{h} D_{1/2}^{n-1} \theta_n^1 + \frac{1}{\tau} \theta_{0}^{n-1} + \frac{2}{h} \left(-K_{1/2}^{n-1} + R^n - E^n\right) = 0,
\]

\[
\theta_{\min} \leq \theta_0^n \leq \theta_{\max}, \quad 1 \leq n \leq N,
\]

\[
\Phi_n^i = \frac{1}{h^2} D_{i-1/2}^{n-1} \theta_{i-1}^n + \frac{1}{h^2} D_{i+1/2}^{n-1} \theta_{i+1}^n - \left\{ \frac{1}{\tau} + \frac{1}{h^2} \left(D_i^{n-1} + D_{i-1/2}^{n-1}\right) \right\} \theta_i^n + \left\{ \frac{\theta_{i-1}^{n-1}}{\tau} + \frac{1}{h} \left(K_{i-1/2}^{n-1} - K_i^{n-1}\right) \right\} = 0,
\]

\[
1 \leq i \leq I - 1, \quad 1 \leq n \leq N,
\]

\[
\Phi_I^n = \theta_I^n - \psi^n = 0, \quad 1 \leq n \leq N,
\]

\[
\theta_i^0 = \varphi_i, \quad 0 \leq i \leq I.
\]

Suppose that \(Q_0\) is a set \(Q_0 = \{(z,t) : z = ih, t = kl\tau\}\), where \(0 \leq i < I, 1 \leq l \leq M, M = \lfloor N/k \rfloor (M \text{ is floor of } N/k), k \text{ is some natural number, } k \geq 1.\) Then the functional has a form:

\[
W(u) = \frac{1}{2} \sum_{i=0}^{I-1} \sum_{l=1}^{M} \xi_i^l \cdot (\theta_i^n - \hat{\theta}_i^n)^2,
\]

where \(n = kl.\) The discrete optimal control problem is to find optimal control \(u^{opt} = (\alpha^{opt}, m^{opt})\) and correspondent optimal solution of direct problem (1) such that the functional \(W(u)\) is minimized.

The system of equations (1) is solved from top to bottom, from first time layer to \(N\)-th one. On the each time layer the solution is determined by sweep method. After obtaining \(\theta_i^n, 0 \leq i \leq I,\) the value of \(\theta_i^n\) is compared with \(\theta_{\min}\) and \(\theta_{\max}.\) If \(\theta_i^n < \theta_{\min}\) or \(\theta_i^n > \theta_{\max}\) then \(\theta_i^n\) is set to \(\theta_{\min}\) or to \(\theta_{\max}\) respectively. Then calculations on the \(n\)-th layer are repeated.

To solve optimal control problem we determine gradient of \(W.\) Introduce the following notations: \(\Phi^T = [\Phi_0^1, ..., \Phi_I^1, ..., \Phi_0^N, ..., \Phi_I^N], \theta^T = [\theta_0^1, ..., \theta_I^1, ..., \theta_0^N, ..., \theta_I^N].\) According to methodology of fast automatic differentiation [3] the relations for determining the gradient of the functional
\( W \) are of the form:

\[
d\frac{W}{u} = W_u + \Phi_u^T p, \tag{2}
\]

\[
W_\theta + \Phi_\theta^T p = 0_n, \tag{3}
\]

The formula (3) of gradient \( d\frac{W}{u} \) contains Lagrange multiplier \( p \) determined by solving system (2). This system is linear with respect to \( p \). The basic matrix \( \Phi_\theta^T \) of the system (2) can be considered as a \( N \times N \) matrix of block elements. Each block has a size \( I \times I \). This block matrix is an upper bidiagonal matrix, i.e. all its nonzero blocks are concentrated on the main diagonal and on the closest diagonal above the main one. It is clear that system (2) can be split into \( N \) systems which can be solved sequentially from bottom to top. Each of these systems is solved by sweep method.

This approach was applied to numerical solution of parameter identification problem with following values of input parameters:

\[
K_0 = 100 \text{ cm/d}, \quad L = 100 \text{ cm}, \quad T = 30 \text{ d},
\]

\[
\varphi(z) = 0.3, \quad z \in (0, L),
\]

\[
\psi(t) = 0.3, \quad E(t) = 0, \quad t \in (0, T),
\]

\[
R(t) = \begin{cases} 
0.1, & t \in (0, T/2), \\
0, & t \in (T/2, T).
\end{cases}
\]

\[
\theta_{\text{min}} = 0.05, \quad \theta_{\text{max}} = 0.5,
\]

\[
\xi(q) = \frac{1}{I \cdot M}, \quad q \in Q_0.
\]

The initial-boundary value problem was considered in dimensionless form. The calculations were carried out on a grid with \( I = 10 \) and \( N = 30 \). Firstly the direct problem was solved with parameters \( \alpha = 0.01 \) and \( m = 0.2 \). Obtained function \( \theta(z,t), (z,t) \in Q_0 \), was considered as the prescribed function \( \hat{\theta}(z,t) \). Then solving optimal control problem we determine \( u^{\text{opt}} = (\alpha^{\text{opt}}, m^{\text{opt}}) \). We started with \( \alpha^{\text{init}} = 0.03 \) and \( m^{\text{init}} = 0.3 \). Numerical optimization was executed by gradient method. The gradient was determined by the described technique. Numerical experiments were performed with \( k = 1, 2, 3, 4 \). The results obtained are presented in the following table.
Obviously with increasing $k$ the optimization process becomes more difficult, and obtained parameters $\alpha^{opt}$ and $m^{opt}$ differ more from true values $\alpha = 0.01$ and $m = 0.2$.

The authors were supported by the Russian Foundation for Basic Research (project no. 10-08-00624 and project no. 11-07-00201).

REFERENCES


Improvement Technology for the Accuracy of Solution of Unconstrained Argument Problems

Anna Dorjieva

1 Institute for System Dynamics and Control Theory of SB RAS, Irkutsk, Russia; ikadark@list.ru

Argument problems form one of the most complicated classes of extremal problems. As is mentioned by some experts [1], today there are no reliable methods for numerical solution of general argument problems.

In this paper we consider two approaches for improvement of solutions of unconstrained argument problems. The first approach is based
on a specific transformation of the original cost function to an auxiliary one with the same points of minimum. The second approach refers to the so-called “Cauchy method”, which reduces our extremal problem to a search for a numeric solution of a system of ordinary differential equations (ODE). The mentioned approaches are specified by means of special computational techniques. Results of numerical experiments are given. Our experimental collection contains a family of Skokov-Nesterov functions, also called “generalized Rosenbrock functions” [1], along with continuously differentiable and non-differentiable functions proposed recently by Y.E. Nesterov [2]:

\[
f(x) = \sum_{i=1}^{n-1} (1 - x_i)^2 + c(x_{i+1} - x_i^2)^2, \quad c = 100, \quad (1)
\]

\[
f(x) = \sum_{i=1}^{n-1} |1 - x_i| + c|x_{i+1} - |x_i||, \quad c = 100, \quad (2)
\]

\[
f(x) = \frac{1}{4}(1 - x_1)^2 + \sum_{i=1}^{n-1} (1 + x_{i+1} - 2x_i^2)^2, \quad (3)
\]

\[
\frac{1}{4}|1 - x_1| + \sum_{i=1}^{n-1} |1 + x_{i+1} - 2|x_i||. \quad (4)
\]

As is well known, the most of local optimization algorithms do not prove themselves effective in minimization of functions of the form (3)-(4), even in the case of relatively small dimensions. By a “right solution” we mean a vector, whose deviation from the optimal value \((1, 1, \ldots, 1)\) is less than 1% in all variables. Given \((-1, 1, -1, 1, \ldots, 1)\) as an initial estimate, the optimal value of the cost function is equal to zero. In what concerns optimization of non-smooth Nesterov functions, we apply the well-known smoothing method \(|y| \approx \sqrt{y^2 + \varepsilon^2}\), and an approximation of the original problem by a sequence of smooth problems depending on a parameter \(\varepsilon\).

Consider the following optimization problem:

\[
f(x) \longrightarrow \min, \quad x \in E^n.
\]

Assume that \(f(x)\) is a unimodal function vanishing at points of its exact minimum. It is easy to see that the point of global minimum of the function \(f(x)\) coincides with one of the function \(g(x) = 1 - \frac{1}{1 + f(x)}\).
Introduce the function $F(x) = f(x) + \alpha g(x)$, where $\alpha$ is a scalar parameter, and consider optimization problem

$$F(x) \rightarrow \min, \quad x \in E^n. \quad (6)$$

Clearly, optimal solutions of problems (5) and (6) coincide.

To solve problem (6) numerically we apply method MSBH (see, e.g., [5]), which is implemented in OPTCON-A solver (see [4]).

The results of the numerical experiments are presented in Tables 1–3. Here, # denotes a number of a test function, $N$ is a dimension, $f^{rec}$ is an obtained estimation to the value of the objective function, $\Delta x = \max_{i=1,\ldots,N}(|x^{rec} - 1|)$ is the maximal deviation in arguments from the components of an optimal vector.

Table 2: MSBH, the original problem

<table>
<thead>
<tr>
<th>#</th>
<th>$N$</th>
<th>$f^{rec}$</th>
<th>$\Delta x$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>17</td>
<td>$5.66366E-14$</td>
<td>$1E-02$</td>
</tr>
<tr>
<td>2</td>
<td>16</td>
<td>$1.54296E-15$</td>
<td>$1E-02$</td>
</tr>
<tr>
<td>3</td>
<td>9</td>
<td>$4.04784E-16$</td>
<td>$1E-02$</td>
</tr>
<tr>
<td>4</td>
<td>10</td>
<td>$8.25361E-18$</td>
<td>$1E-03$</td>
</tr>
</tbody>
</table>

Table 3: MSBH, the reduced problem

<table>
<thead>
<tr>
<th>#</th>
<th>$N$</th>
<th>$f^{rec}$</th>
<th>$\Delta x$</th>
<th>$\alpha$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>26</td>
<td>$1.08420E-13$</td>
<td>$1E-02$</td>
<td>$1.00E+06$</td>
</tr>
<tr>
<td>2</td>
<td>30</td>
<td>$7.36295E-17$</td>
<td>$1E-03$</td>
<td>$1.00E+07$</td>
</tr>
<tr>
<td>3</td>
<td>16</td>
<td>$4.74014E-22$</td>
<td>$1E-02$</td>
<td>$1.00E+10$</td>
</tr>
<tr>
<td>4</td>
<td>16</td>
<td>$2.30513E-19$</td>
<td>$1E-02$</td>
<td>$1.00E+08$</td>
</tr>
</tbody>
</table>

Note that in our numerical experiments the values of the parameter $\alpha$ were given, though they were particular for different dimensions. In what concerns the setup of the parameter $\varepsilon$ (in the case of non-differentiable
functions), we simply put $\varepsilon = 1E - 12$. By varying $\alpha$ and $\varepsilon$ in a more sophisticated way, we succeeded in obtaining “right solutions” for problems of relatively high dimensions in a series of experiments.

Another approach for improving the accuracy of solutions of unconstrained argument problems consists in application of a computational technique with a qualifying local search algorithm. According to this technique, the original problem is reduced to the following Cauchy problem for the system of ODE:

\[
\begin{aligned}
\dot{y} &= -\nabla f(y), \quad t \in [t_0, t_1], \\
y_0 &= y(t_0).
\end{aligned}
\] (7)

Given $y^k$ we set $y^{k+1} = y^k + h_k \cdot \nabla f(y^k)$, where $h_k$ defines, as usual, the $k$th “step” of our numerical method, and $\nabla f(y^k)$ is calculated by means of a finite-difference scheme. The least of the values $f_k = f(y^k)$, $k = 1, K$, is said to be numerical solution of the original problem, while the corresponding values $y^k$ are called the points of minimum of the cost function.

<table>
<thead>
<tr>
<th>N</th>
<th>$f_{rec}$</th>
<th>$\Delta x$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>$2,00759E - 12$</td>
<td>$1E - 02$</td>
</tr>
<tr>
<td>2</td>
<td>$6,36721E - 15$</td>
<td>$1E - 02$</td>
</tr>
<tr>
<td>3</td>
<td>$4,58770E - 13$</td>
<td>$1E - 03$</td>
</tr>
<tr>
<td>4</td>
<td>$5,53190E - 15$</td>
<td>$1E - 02$</td>
</tr>
</tbody>
</table>

Our numerical experiments show that the proposed technique of the problem transformation to an “inverted” one allows us to increase considerably (plus 6 or more variables) the dimension of “rightly solved” problems. We also analyzed the dependence of the “quality” of the obtained solutions on values of the parameter $\alpha$. It was found out that by applying the qualifying local search, one can increase the dimension of rightly solved problems in 4 or more variables, depending on values of $h_k$. Note that in all the problems we solved numerically, the deviation in arguments of the objective function from their optimal values monotone increase as the number of a variable raises. Finally we conclude that the
proposed computational technologies allows one to increase appreciably the accuracy of solutions of some unconstrained argument problems.

The author was partially supported by the Russian Foundation for Basic Research (project no. 12-01-00193).

REFERENCES


On optimal stabilization with respect to a part of variables for multiply connected controlled systems

Olga Druzhinina¹, Natalia Petrova²

¹ Dorodnicyn Computing Center of RAS, Russia; ovdruzh@mail.ru
² Dorodnicyn Computing Center of RAS, Russia; lyza@ccas.ru

For nonlinear controlled systems methods of optimal stabilization are considered in [1]–[5] and other works. In this work we suggest a method for solving the problem of optimal stabilization to respect to a part of variables for multiply connected nonlinear controlled dynamic systems given in the form of nonlinear systems of the ordinary differential equations. The decision of the specified problem reduces to the solving of a problem of optimal stabilization in the sense of V.V. Rumyantsev [1], [2].
V.V. Rumyantsev in [1] proved a theorem about the optimal stabilization of stable (asymptotically stable) system of differential equations of perturbed motion with the additional forces under the condition of minimization of a functional characterizing the quality of control. The functional is given in the form of a definite integral with the upper infinite limit. Integrand function of the functional is defined in the proof of the theorem, in this case the known Lyapunov function for the system of differential equations of perturbed motion without the control becomes the optimal Lyapunov function for the specified system under the action of additional forces.

We consider multiply connected nonlinear controlled dynamic system

\[
\frac{dx_s}{dt} = f_s(t, x_s, u^\text{loc}_s, u^\text{glob}_s) + F_s(t, x_s, u^\text{loc}_s, u^\text{glob}_s) \equiv \Phi_s(t, x_s, u^\text{loc}_s, u^\text{glob}_s), \quad s = \overline{1,q},
\]

where \(x = (x^T_1, \ldots, x^T_q)^T\), \(x_s \in R^{n_s}\), \(R^{n_1} \oplus \ldots \oplus R^{n_q} = R^n\), \(u^\text{loc}_s(t,0) = 0\), \(u^\text{glob}_s(t,0) = 0\), \(\Phi_s(t,0,0,0) \equiv 0\).

It is accepted that right part of system (1) is defined in domain

\[
\Omega_1 = \{t, x, u^\text{loc}_s, u^\text{glob}_s : t \geq t_0 \geq 0, ||x|| < H, ||u^\text{loc}_s|| < \infty, ||u^\text{glob}_s|| < \infty, 0 < H = \text{const}, s = \overline{1,q}\},
\]

and conditions of existence and uniqueness of solution are satisfied. Let us assume than system (1) can be represented as

\[
\begin{align*}
\frac{dy_s}{dt} &= Y_s(t, y_s, z_s, u^\text{loc}_s) + \sum_{j=1}^{q} Y_{1sj}(t, y, z)u^\text{glob}_s, \\
\frac{dz_s}{dt} &= Z_s(t, y_s, z_s, u^\text{loc}_s) + \sum_{j=1}^{q} Z_{1sj}(t, y, z)u^\text{glob}_s,
\end{align*}
\]

where \(x_s = (y^T_s, z^T_s)^T\), \(x = (y^T, z^T)^T\), where \(y_s \in R^{k_s}\), \(z_s \in R^{n_s}\), \(k_s + m_s = n_s\), \(s = \overline{1,q}\). For system (3) domain (2) takes the form

\[
\Omega_2 = \{t, x, u^\text{loc}_s, u^\text{glob}_s : t \geq t_0 \geq 0, ||y_s|| < H_s, ||z_s|| \leq \infty, \\
||u^\text{loc}_s|| < \infty, ||u^\text{glob}_s|| < \infty, 0 < H = \text{const}, s = \overline{1,q}\},
\]

and each solution is z-extendible.
We consider the subsystems of the form
\[
\begin{align*}
\frac{dy_s}{dt} &= Y_s(t, y_s, z_s, u_{s}^{\text{loc}}), \\
\frac{dz_s}{dt} &= Z_s(t, y_s, z_s, u_{s}^{\text{loc}}), \quad s = 1, q. 
\end{align*}
\] (4)

Further, we will solve the problem of optimal $y_s$-stabilization of multiply connected dynamical systems of the form (3), $s = 1, q, y = (y_1^T, ..., y_q^T)^T$, using the method of Lyapunov vector-functions. In this case the strategy of solving the problem of stabilization is that each subsystem must be $y_s$-stabilized with the help of local controls $u_{s}^{\text{loc}}$, $s = 1, q$, i.e. it must be $y_s$-stabilized on the level of subsystems, and then the asymptotic $y_s$-stability of interconnected subsystems must be checked. The general scheme of a two-level stabilization scheme [6] is that the global control $u_{s}^{\text{glob}}$, $s = 1, q$, is added to the decentralized control in order to weaken the effect of interrelated subsystems. In this work the problem of optimal stabilization of multiply connected system is solved also using a two-level stabilization scheme with respect to a part of variables.

We consider the case when right parts of (4) can be written in the form
\[
\begin{align*}
Y_s(t, y_s, z_s, u_{s}^{\text{loc}}) &\equiv \overline{Y}_s(t, y_s, z_s) + b_1(t, y_s, z_s)u_{s}^{\text{loc}1}, \\
Z_s(t, y_s, z_s, u_{s}^{\text{loc}}) &\equiv \overline{Z}_s(t, y_s, z_s) + b_2(t, y_s, z_s)u_{s}^{\text{loc}2}, \quad s = 1, q, 
\end{align*}
\] (5)

where $b_1(t, y_s, z_s)$ and $b_2(t, y_s, z_s)$ are matrices of appropriate dimensions, and controls $u_{s}^{\text{loc}1}$ and $u_{s}^{\text{loc}2}$ are built considering the choice of Lyapunov vector-functions.

It was shown that equilibrium state of system (4) taking into account (5) is uniformly asymptotic $y_s$-stable. In this case system (3) can be represented in the form
\[
\begin{align*}
\frac{dy_s}{dt} &= \varphi_s(t, y_s, x_s) + Y_{1s}(t, \tilde{y}, \tilde{z})u_{s}^{\text{glob}}, \\
\frac{dz_s}{dt} &= \psi_s(t, y_s, z_s) + Z_{1s}(t, \tilde{y}, \tilde{z})u_{s}^{\text{glob}}, 
\end{align*}
\] (6)

where
\[
\begin{align*}
\varphi_s(t, y_s, z_s) &= \overline{Y}(t, y_s, z_s) + b_1(t, y_s, z_s)u_{s}^{\text{loc}}(t, y_s, z_s), \\
\psi_s(t, y_s, z_s) &= \overline{Z}_s(t, y_s, z_s) + b_2(t, y_s, z_s)u_{s}^{\text{loc}}(t, y_s, z_s).
\end{align*}
\]
For system (6) we consider the problem of optimal stabilization. Criterion of quality control we write in the integral form

$$J = \int_{0}^{\infty} w(t, y[t], z[t], u_{s}^{\text{glob}}[t]) dt,$$

in this case in the process of solution we define the function $w(t, x, u)$.

As optimal Lyapunov function for system (6) we choose the function

$$V(t, y, z) = \sum_{s=1}^{q} \alpha_{s} V_{s}(t, y_{s}, z_{s}),$$

where $\alpha_{s}$ are positive real constants, $V_{s}(t, x_{s})$ are Lyapunov functions which guarantee uniform asymptotic $y_{s}$-stability of systems (4), $s = 1, q$.

We introduce Krasovsky–Bellman function $B(t, x, u, v, u_{\text{glob}})$ with special component $\Psi(t, y, z)$ allowing to consider the function $w$ in integral (7) in the form

$$w(t, y, z, u_{\text{glob}}) = \Psi(t, y, z, u_{\text{glob}}) + \frac{1}{2} \sum_{s=1}^{q} (u_{s}^{\text{glob}})^{T} \theta_{s} u_{s}^{\text{glob}}.$$

According to Rumyantsev theorem function $B(t, x, u, v, u_{\text{glob}})$ is positive-definite with respect to $y$, and along optimal control $(u_{s}^{0})^{\text{glob}}$ we have that

$$\frac{\partial B}{\partial u_{s}^{\text{glob}}}(u_{s}^{0})^{\text{glob}} = 0, s = 1, q.$$

In result we obtain positive-definite function with respect to $y$-component of phase vector of system (3). In this case we can write the criterion of quality control in the form

$$J = \int_{t_{0}}^{\infty} \left( \sum_{s=1}^{q} \alpha_{s} W_{s}(t, y_{s}, z_{s}) + \sum_{s,j=1}^{q} \theta_{s,j}(u_{s}^{0})^{\text{glob}}(u_{j}^{0})^{\text{glob}} + \right.$$  

$$+ \sum_{s,j=1}^{q} \theta_{s,j} u_{s}^{\text{glob}} u_{j}^{\text{glob}} \right) dt.$$

Thus if (i) for systems (4) $y_{s}$-stabilizing (to uniform asymptotic stability) controls $u_{s}^{\text{loc}} = u_{s}^{\text{loc}}(t, x_{s})$ exist, $s = 1, q$, (ii) function $\Psi(t, y, z)$
is positive-definite with respect to vector \( y \) of system (3), then controls \((u_0^0)_{\text{glob}}\) defined from system (8) are the functions solving the problem of optimal \( y \)-stabilization of system (3) with respect to functional (9).

By the aid of the presented results with applications of results of works [7]–[11] the algorithms of optimal stabilization of multiply connected nonlinear controlled systems are constructed.

The authors were supported by the Russian Foundation for Basic Research (project no. 10-08-00826-a).

REFERENCES

Numerical experiments on computer model of a cerebellum

W. Dunin-Barkowski$^1$, L. Vyshinskiy$^2$

$^1$ Scientific Research Institute for System Analysis RAS, Moscow, Russia; wldbar@gmail.ru
$^2$ A.A. Dorodnitsyn Computing Center RAS, Moscow, Russia; wysh@ccas.ru

Computer modeling of neural networks of big dimension becomes the effective tool in studying of mechanisms of processing of information in a human brain. It is connected, on the one hand, with increase in speed of modern computers at the expense of application of parallel calculations, and on the other hand with achievements of neurophysiology which is close to understanding of many mechanisms of work of a brain. New possibilities allow to investigate various schemes and algorithms of processing of information which ”are invented” by the nature and, maybe, to find of them application for the solution of various applied tasks. It is very tempting ideas as real neural structures of live organisms, and first of all a human brain, successfully solve the most difficult problems of forecasting, management and control over various situations. In the report the researches which are carried out by authors in the field of mathematical modeling of neural structures of a cerebellum are presented ([1]).

Cerebellum - one of the most studied independent structures of a brain. The cerebellum contains more than all other structures of a brain of neurons. He receives a large volume of information from the outside. About 40 million nervous fibers are connected by a cerebellum with a bark of big hemispheres and approximately as with peripheral nervous system. At a cerebellum rather homogeneous neural structure which speaks about existence of some general mechanism of calculation of millions the parameters necessary in the solution of various problems of movements control and not only movement. This mechanism as a first approximation includes some stages: processing of arriving entrance information from a bark of big hemispheres and peripheral nervous system, control (training) of neural structures of a cerebellum according to teams arriving from the outside, calculation of values of the parameters necessary for solution of problems of control of an organism and their transfer to a bark of big hemispheres or to peripheral nervous system. Problems of calculation of
a large number of multidimensional functions in a cerebellum are solved by means of homogeneous structure with feedback which includes various layers of nervous cells. On the drawing (Fig. 1) the block diagram of a cerebellum model which is studied in the real work is shown.

![Diagram of a cerebellum model](image)

To each layer of cells presented on the chart there corresponds the set of the equations describing behavior of this layer.

\[ M(t) = K(X,t), \]  
\[ T_G dG_i/dt = -G_i(t) + \Sigma j g_{ij} M_j(t), \quad i \in 1,..,ng, j \in 1,..,nm, \]  
\[ T_P dP_i/dt = -P_i(t) + \Sigma j p_{ij} w_{ij}(t) G_j(t), \quad i \in 1,..,np, j \in 1,..,ng, \]  
\[ T_N dN_i/dt = -N_i(t) + \Sigma j n_{ij} P_j(t), \quad i \in 1,..,nn, j \in 1,..,np, \]  
\[ T_{CF} dC_i/dt = \Psi(C_i(t), \Sigma j c_{ij} N_j(t) + Y_i(t)), \quad i \in 1,..,nc, j \in 1,..,nn, \]  
\[ \Psi(C(t), F(t)) = \text{if} (C(t)<H_0) \text{then} (F(t) - C(t)) \text{else} (-H_0 \cdot \delta(t)), \]  
\[ dw_{ij}/dt = \varepsilon_{CF} \Sigma_k e_{ki} \chi(t - \tau_k(t)) G_j(t), \]  
\[ i \in 1,..,np, j \in 1,..,ng, k \in 1,..,nc, \]  
\[ Z(t) = H(N(t)). \]
Multidimensional function $K(X,t)$ models transformation of information coming of external structures in signals on mossy fibers of a cerebellum. By means of this function it is possible to set various ways of coding of entrance information. In the report some options of coding are considered and compared.

The group of the equations (2) describes work of a layer of granular cells which perceive and will transform signals from mossy fibers. The layer of granular cells is the most numerous in cerebellum structures. The quantity of cells in this layer can reach hundred thousands. Entrances of granular cells are connected to mossy fibers and perceive signals coming from the outside through these fibers. At granular cages small number of entrances (5-10). Connections with mossy fibers (a matrix $|g_{ij}|$) have casual character, however various hypotheses here can be used.

The equations (3) are the equations of Purkinye’s cells, the main calculator in a cerebellum. Purkinye’s cells are multidimensional arithmetic adders of the weighed signals from granular cells. The quantity of entrances of Purkinye’s cells can make tens of thousands. Conductivity of communications from granular cages to synapse Purkinye’s cells (synaptic weight) cope fiber cells. Synaptic weights are the instrument of control of feedback in a cerebellum.

The layer of small nuclear cells (the equations 4) provides feedback in structure from Purkinye’s cells to climbing fiber cells.

The equations (5) describe behavior of climbing fiber cells. Signals come to climbing fiber cells from Purkinye’s cells through small nuclear cells and external signals which play a role of operating (training) signals models. The climbing fiber cell can be described as pulse system which periodically gives out the impulses correcting synaptic weights on entrances of cells of Purkinye. This process closes feedback in a cerebellum and allows regulating and adjusting its work on performance of various functions.

Processes of change of values of synaptic weights are described by the equations (6). The fiber cell submits an impulse on synoptic knots to the moments of a relaxation of function $\Psi$. It changes weight for all active communications. In the equation (6) $\tau$ – the last moment of a relaxation a fiber cell of the connected with concrete synaptic knot. The function $\chi(t)$ sets an impulse form.
Multidimensional function \( H(N) \) (7) forms the target information which is transferring from a cerebellum in a bark of big hemispheres and in peripheral nervous system.

For studying of this model the special research stand was developed. This stand is interactive program for modeling the equations (1) - (7) for various ways of formation of matrixes of communication \( ||g_{ij}||, ||p_{ij}||, \) \( ||n_{ij}||, ||c_{ij}|| \), and various values of parameters of the equations.

In the report results of the computing experiments which have been carried out at this stand will be presented. Entrance variables which hand over external information in cerebellum model, are divided into two canals – the X-th and Y. On the channel X-th information gets through mossy fibers and granular cells on Purkinje’s cells. On the channel Y of information arrives on climbing fiber cells and influences control of synaptic weights. Therefore the channel Y is considered as suited “management” or ”training”. In examples of figure 2 some tasks which decided in experiments with cerebellum model are shown.

\[
\begin{align*}
\text{a)} \quad &X(t) \\
&Y(t) = F(X(t)) \\
&\text{CB} \\
&\rightarrow Z(t) = G(X(t), u, t) \sim Y(t)
\end{align*}
\]

\[
\begin{align*}
\text{b)} \quad &X(t) = F(X(t-\Delta)) \\
&Z(t) = G(X(t), u, t) \sim Y(t+\Delta)
\end{align*}
\]

\[
\begin{align*}
\text{c)} \quad &\frac{dX}{dt} = F(X, Z, u, t) \\
&X(t) \\
&\text{CB} \\
&\rightarrow Z(t) = G(X(t), u, t) \sim Y(t)
\end{align*}
\]

Fig. 2
a) approximation of functions; b) forecasting of numerical sequences; c) modeling of dynamic systems and training of control of dynamic systems. (variable “u” on schemes is the switch of model’s operating modes – training or check)

Program realization of the interactive stand is executed with the help of “The generator of projects” ([2]).

The authors were supported by the Russian Foundation for Basic Research (project no. 10-07-00206).

REFERENCES


P-th order methods for solving nonlinear system

Yu.G. Evtushenko¹, A.A. Tretyakov¹,²,³

¹ Dorodnicyn Computing Center RAS, Moscow, Russia;
² System Res. Inst., Polish Acad. Sie, Warsaw, Poland;
³ University of Podlasie, Siedlce, Poland;

evt@ccas.ru, tret@ap.siedlce.pl

For a system of nonlinear equations, we derive a formula for a family of its approximate solutions, which provides an elementary proof of the tangent direction theorem and the implicit function theorem. Iterative methods with a local $p$-order convergence rate are constructed for solving systems of equations. An elementary proof of the Lagrange theorem on necessary extremum conditions in an equality constraint nonlinear programming problem is given as an application.
Given a vector function $F(x)$:

$$F(x) : \mathbb{R}^n \to \mathbb{R}^m, \quad F^T(x) = [f^1(x), f^2(x), \ldots, f^m(x)],$$

where $f^i(x) : \mathbb{R}^n \to \mathbb{R}^1, i = 1, 2, \ldots, m, m \leq n$, $F(x) \in C^3(\mathbb{R}^n)$, we define the feasible set

$$X = \{x \in \mathbb{R}^n \mid F(x) = 0_m\}. \quad (1)$$

The rectangular $m \times n$ Jacobi matrix $F_x(x)$ and its kernel and image at a point $x \in \mathbb{R}^n$ are defined as

$$\text{Ker} F_x(x) = \{h \in \mathbb{R}^n \mid F_x(x)h = 0_m\},$$

$$\text{Im} F_x(x) = \{\eta \in \mathbb{R}^m \mid \eta = F_x(x)y, \ y \in \mathbb{R}^n\} = (\text{Ker}^T F_x(x))^\perp. \quad (2)$$

Let the rank of $F_x(x)$ be $m$. Then there exists a nonsingular $m \times m$ Gram matrix $G(x) = F_x(x)F_x^T(x)$ and a pseudoinverse (right inverse) $F_x^+(x) = F_x^T(x)G^{-1}(x)$ of $F_x(x)$; i.e., $F_x(x)F_x^+(x) = I_m$. Define two $n$-dimensional vectors

$$N(x) = F_x^+(x)F_x(x), \quad p(x) = -\frac{1}{2} F_x^+(x)hF_{xx}(x*)h^T. \quad (3)$$

Here, $h \in \mathbb{R}^n$ and $h^T F_{xx}(x*)h$ is an $m$-dimensional vector whose $i$th component is $h^T f_{xx}^i(x*)h$.

**Theorem 1.** Let a point $x* \in X$ be such that the columns of the matrix $F_x^T(x*)$ are linearly independent and $h \in \text{Ker} F_x(x*)$.

Then in $\mathbb{R}^n$ there is a family of arcs

$$x(\alpha) = x_0 + \alpha h + \alpha^2 p(x_0), \quad (4)$$

that issue from $x_0$ and belong to the set $X$ up to second-order quantities.

Based on formulas (3) and (4), we can construct an iterative method for solving the system $F(x) = 0_m$ with $m \leq n$. For the mapping $F(x)$ to be approximated up to an $O(\|x - x_*\|^3)$ error in (4), we use, as $x(\alpha), x_*, \alpha h$, and $\alpha^2 p(x_0)$, the vectors

$$x_{i+1}, \quad x_i, \quad -N(x_i), \quad -(1/2)F_x^+(x_i)N^T(x_i)F_{xx}(x_i)N(x_i),$$

respectively. As a result, we obtain the iterative method

$$x_{i+1} = x_i - N(x_i) - \frac{1}{2} F_x^+(x_i)N^T(x_i)F_{xx}(x_i)N(x_i). \quad (5)$$
Theorem 2. Suppose that $F \in C^2(\mathbb{R}^n)$ and there exists a point $x_* \in X$ at which the rank of the matrix $F_x(x_*)$ is $m$.

Then there exists a neighborhood $U(x_*)$ of $x_*$ such iterative process (5) converges at a cubic rate for any point $x_i \in U(x_*)$; moreover,

$$\text{dist}(x_{i+1}, X) \leq c \text{dist}(x_i, X)^3,$$

where $c > 0$ is an independent constant.

Note that $\{x_k\}$ converges to a point $\bar{x} \in X$ so as to satisfy the convergence rate estimate

$$\|x_{i+1} - \bar{x}\| \leq c \|x_i - \bar{x}\|^3.$$

Consider the special case of system (1) where $F : \mathbb{R}^n \to \mathbb{R}^n$. Assume that the square matrix $F_x(x)$ is nonsingular. Then the computational formulas simplify, specifically, $F_x^+(x) = F_x^{-1}(x)$, $N(x) = F_x^{-1}(x)F(x)$, and iterative method (5) becomes

$$x_{i+1} = x_i - F_x^{-1}(x_i) \left( F(x_i) + \frac{1}{2} N^T(x_i)F_{xx}(x_i)N(x_i) \right).$$

(6)

Theorem 3. Suppose that $F \in C^2(\mathbb{R}^n)$ and there exists a point $x_* \in X$, where the matrix $F_x(x_*)$ is nonsingular.

Then iterative method (6) locally converges to $x_*$ at a cubic rate:

$$\|x_{i+1} - x_*\| \leq c \|x_i - x_*\|^3,$$

where $c$ is a constant.

In the special case of $n = m = 1$, it follows from (6) that

$$x_{i+1} = x_i - \frac{f(x_i)}{f_x(x_i)} - \frac{f^2(x_i) f_{xx}(x_i)}{2 f_x^3(x_i)}.$$

This method for finding the roots of the equation $f(x) = 0$ was proposed in 1838 by Chebyshev [1]. Iterative methods converging faster than quadratically were studied in [2, 3].

Consider the case $p \geq 3$ for solving following system:

$$F(x) = 0_n, \quad x \in \mathbb{R}^n.$$  

(10)
Assume that the point \(x\) belongs to a sufficiently small neighborhood of \(x_*\). Then use the Taylor formula we obtain

\[
F(x_*) = 0_n = F(x) - F'(x)h + \frac{1}{2} F^{(2)}(x)[h]^2 + \cdots + \frac{(-1)^p}{p!} F^{(p)}(x)[h]^p + O_n(\|h\|^{p+1}).
\]

(11)

Here \(h(x) = x - x_*\), \([h(x)]^p\) — multivector, \(O_n(\|\cdot\|^{p+1})\) — \(n\)-dimensional vector with \((p+1)\)-order approximation by \(h\) of each component, \(F'(x) = F_x(x)\). Suppose that the matrix \(F'(x_*)\) nondegenerate. Then there exists neighborhood \(U_\varepsilon\) of the point \(x_*\) such that \(F'(x)\) is nondegenerate and we define vector \(N(x) = [F'(x)]^{-1}F(x)\).

Let us multiple (11) by \([F'(x)]^{-1}\) we obtain

\[
x - x_* = h_1(x) + [F'(x)]^{-1}\left[\frac{1}{2} F^{(2)}(x)[h]^2 + \cdots + \frac{(-1)^p}{p!} F^{(p)}(x)[h]^p\right] + O_n(\|h\|^{p+1}),
\]

(12)

where \(h_1(x) = N(x)\). Denote

\[x_1(x) = x - h_1(x).\]

Omitting the second order terms we obtain from (12)

\[\|x_1(x) - x_*\| = O_n(\|x - x_*\|^2).\]

If we take \(x_1(x)\) as approximate solution to the system (10), then we obtain

\[
F(x_1(x)) = f(x) - F'(x)h_1(x) + O_n(\|x - x_*\|^2) = O_n(\|x - x_*\|^2).
\]

(13)

Thus vector \(x_1(x)\) define approximation solution of system (10) with second order explicitly of \(\|x - x_*\|\).

In the right side of (12) omit the terms of third order we obtain

\[
x - x_* - h_2(x) + O_n(\|x - x_*\|^3), \quad x_2 = x - h_2(x),
\]

(14)

where

\[
h_2(x) = h_1(x) + \frac{1}{2} [F'(x)]^{-1} F^{(2)}(x)[h_1(x)]^2.
\]

(15)
Denote by $x_2(x) = x - h_2(x)$. Take place the estimations analogous to (14), (15)

$$\|F(x_2)\| \approx O_n(\|x - x_*\|^3), \quad \|x_2(x) - x_*\| \approx O_n(\|x - x_*\|^3).$$

Analogously, omit in (3) the terms of fourth order we find

$$x - x_* = h_3(x) - O_n(\|x - x_*\|^4), \quad x_3(x) = x - h_3(x), \quad (16)$$

where

$$h_3 \triangleq h_1 + [f'(x)]^{-1} \left\{ \frac{1}{2} F^{(2)}(x)[h_2]^2 - \frac{1}{3!} F^{(3)}(x)[h_2]^3 \right\}$$

and

$$\|F(x_3)\| \approx O(\|x - x_*\|^4), \quad \|x_3(x) - x_*\| \approx O(\|x - x_*\|^4).$$

Continue this process we are coming to following formulas:

$$x - x_* = h_p(x) - O_n(\|x - x_*\|^{p+1}), \quad x_p(x) = x - h_p(x),$$

where

$$h_p(x) = h_1(x) + [F'(x)]^{-1} \left\{ \frac{1}{2} F^{(2)}(x)[h_{p-1}]^2 + \cdots + \right.$$ 

$$+ \frac{(-1)^p}{p!} F^{(p)}(x)[h_{p-1}]^p \right\}.$$  \hspace{1cm} (17)

Generalize all what has been said above we can assert that will be true following result.

**Theorem 4.** Suppose that $F \in C^{p+1}(\mathbb{R}^n, \mathbb{R}^n)$, there exists a point $x_* \in X$ where the matrix $F'(x_*)$ is nonsingular, $x_i \in U_\varepsilon(x_*)$ and $\varepsilon > 0$ sufficiently small.

Then for the point $x_{i+1}$

$$x_{i+1} = x_i - [F'(x_i)]^{-1} \left\{ F(x_i) + \frac{1}{2} F''(x_i)[h_{p-1}]^2 + \cdots + \right.$$ 

$$+ \frac{(-1)^p}{p!} F^{(p)}(x_i)[h_{p-1}]^p \right\}.$$  \hspace{1cm} (18)
will be hold the following estimations:

\[ \|F(x_{i+1})\| \leq c \|x_i - x_*\|^{p+1}, \quad \|x_{i+1} - x_*\| \leq c \|x_i - x_*\|^{p+1}, \]  

(19)

where \( c > 0 \) independent constant.

The authors were supported by the Russian Foundation for Basic Research (grant No 11-01-00786a) and the Council for the State Support of Leading Scientific Schools (grant 5264.2012.1).

REFERENCES


Some Complexity Results for the Simple Assembly Line Balancing Problem

Evgeny R. Gafarov\(^1\), Alexandre Dolgui\(^2\), Alexander Lazarev\(^3\)

\(^1\) Ecole Nationale Superieure des Mines, FAYOL-EMSE, CNRS:UMR6158, LIMOS, F-42023 Saint-Etienne, France; axel73@mail.ru

\(^2\) Ecole Nationale Superieure des Mines, FAYOL-EMSE, CNRS:UMR6158, LIMOS, F-42023 Saint-Etienne, France; dolgui@emse.fr

\(^3\) Institute of Control Sciences of the Russian Academy of Sciences, Profsoyuznaya st. 65, 117997 Moscow, Russia; jobmath@mail.ru

Introduction

We consider the simple assembly line balancing problem (SALBP-1) which is formulated as follows.
Given a set \( N = \{1, 2, \ldots, n\} \) of operations and \( K \) stations (machines) \( 1, 2, \ldots, K \). For each operation \( j \in N \) a processing time \( t_j \geq 0 \) is defined. The cycle time \( c \geq \max\{t_j, j \in N\} \) is given. Furthermore, finish-start precedence relations \( i \to j \) are defined between the operations according to an acyclic directed graph \( G \). The objective is to assign each operation \( j, j = 1, 2, \ldots, n \), to a station in such a way that:

- number \( m \leq M \) of stations used is minimized;
- for each station \( k = 1, 2, \ldots, m \) a total load time \( \sum_{j \in N_k} t_j \) does not exceed \( c \), where \( N_k \) – a set of operations assigned to a station \( k \);
- given precedence relations are fulfilled, i.e. if \( i \to j, i \in N_{k_1} \) and \( j \in N_{k_2} \) then \( k_1 \leq k_2 \).

A survey on results for NP-hard in the strong sense SALBP-1 is presented, e.g., in [1,2].

**Partition problem.** Given is a set \( N = \{b_1, b_2, \ldots, b_n\} \) of numbers \( b_1 \geq b_2 \geq \cdots \geq b_n > 0 \) with \( b_i \in \mathbb{Z}_+ \), \( i = 1, 2, \ldots, n \), and a number \( A \in \mathbb{Z}_+ \) with \( A < \sum_{j \in N} b_j \). Is there a subset \( N' \subset N \) such that \( \sum_{j \in N'} b_j = A \)?

The worst case running time of \( B&B \) algorithms for the well-known Knapsack Problem is analyzed, e.g., in [4]. In these papers authors choose to use only special cases, for which it is fairly easy to find an optimal solution with a \( B&B \) algorithm. However to prove its optimality, almost all feasible solutions should be considered. We use a similar idea to construct a special case of SALBP-1 for which each \( B&B \) algorithm with no matter what polynomial time computed Lower Bound has an exponential run time. This makes algorithms ineffective for instances with \( n \geq 60 \) operations.

**Modified instance of the Partition problem.** Given is a set \( \overline{N} = \{\overline{b}_1, \overline{b}_2, \ldots, \overline{b}_{2n}\} \) of numbers \( \overline{b}_1 \geq \overline{b}_2 \geq \cdots \geq \overline{b}_{2n} > 0 \) with \( \overline{b}_i \in \mathbb{Z}_+ \), \( i = 1, 2, \ldots, 2n \), and a number \( \overline{A} \in \mathbb{Z}_+ \) with \( \overline{A} < \sum_{j \in \overline{N}} \overline{b}_j \). The numbers \( \overline{b}_i, \ i = 1, 2, \ldots, 2n \) are denoted as follows:

\[
\overline{b}_{2n} = 1, \quad \overline{b}_{2i} = 2 \cdot \sum_{j=i+1}^{n} \overline{b}_{2j-1}, \quad i = n-1 \ldots, 1, \quad \overline{b}_{2i-1} = \overline{b}_{2i} + b_i, \quad i = n \ldots, 1,
\]

where \( b_1, b_2, \ldots, b_n \) – numbers from the initial instance. Let \( \overline{A} = \sum_{i=1}^{n} \overline{b}_{2i} + A \). Without lost of generality let us assume \( A = \frac{1}{2} \sum_{i=1}^{n} b_i \) and as consequence \( \overline{A} = \frac{1}{2} \sum_{i=1}^{2n} b_i \). The question is: "Is there a subset \( \overline{N}' \subset \overline{N} \) such that \( \sum_{j \in \overline{N}'} \overline{b}_j = \overline{A}' \)? If for the initial instance of the Partition
Problem the answer is "YES" (and the same answer has the modified instance) then \( N' \) contains one and only one number \( \bar{b}_i \) from each pair \( \{\bar{b}_{2i-1}, \bar{b}_{2i}\} \), \( i = 1, 2, \ldots, n \). If the number \( b_i \) is included in the set \( N' \) then \( \bar{b}_{2i-1} \) is included in \( N' \), otherwise the number \( \bar{b}_{2i} \in N' \).

In the special case of SALBP-1 there are \( 2n \) operations. Let \( w' = \min\{w|10^w \geq 2A\} \). Let us

\[ t_i = 10^{w'} + \bar{b}_i, \quad i = 1, 2, \ldots, 2n \]

and \( c = \frac{1}{2} \sum_{i=1}^{2n} t_i \). There are no precedence relations between operations.

It is obvious that if and only if for the modified instance of the Partition Problem the answer is "YES" then the minimal number of stations \( m^* = 2 \), otherwise \( m^* = 3 \). As a consequence, if \( NP \neq P \), there is no polynomial time computed Lower Bound with a relative error equal or less than \( \frac{3}{2} \). That means, for any set of polynomial time computed Lower Bounds \( \{LB_1, LB_2, \ldots, LB_X\} \), there is a modified instance of the Partition Problem with an answer "NO", for which \( LB_x = 2 \), \( i = 1, 2, \ldots, X \), although \( m^* = 3 \). For the special case of SALBP-1, any feasible solution is optimal. However, to prove its optimality almost all feasible solutions must be considered.

Let us estimate the possible number of feasible solutions. On the first station there could be processed at least \( n - 1 \) operations. Thus, there are at least \( \binom{2n}{n-1} \) possible loads of the first station, i.e. the number of feasible solutions which have to be considered is greater than \( \binom{2n}{n-1} = \frac{n+1}{n} \binom{2n}{n} \approx \frac{n+1}{n} \frac{2^{2n}}{\sqrt{n\pi}} \). To solve such the instance of SALBP-1 with \( 2n = 60 \) a computer must perform more than \( 2^{60} \) operations. Let us assume that the fastest known computer performs \( 2^{30} \) operations per second, or less than \( 2^{47} \) operations per day. Then a run time of an algorithm will be more than \( \frac{2^{13}}{10} > 800 \) days! That means there are instances of SALBP-1 for which any \( B&B \) algorithm with polynomial time computed Lower Bounds has an inappropriate running time.

We can conclude the following. Despite the best known algorithm \( B&B \) [3] solves all benchmark instances in less than 1 second per instance, known \( B&B \) algorithms for SALBP-1 remain exponential and can not solve some instances with the size \( n > 60 \) in an appropriate time. That is why we consider exact algorithms for the general case of the problem unpromising. Researchers can concentrate on special cases or on
essentially new solution schemes.

Maximization of Number of Stations

To propose an essentially new solution scheme for SALBP-1, it is necessary to investigate properties of optimal solutions. We can investigate not only properties of good solutions to try imitate their character but properties of poor solutions as well to avoid solutions with their aspects. Here, in contrast to standard SALBP-1, where the number of stations used should be minimized we consider an optimization problem with the opposite objective criteria, in other words the maximization of the number of stations. The investigation of a particular problem with the maximum criterion is an important theoretical task [5]. To make the maximization problem not trivial we assume that all stations (instead the last one) should be maximal loaded, i.e. for two stations \( m_1, m_2, m_1 < m_2 \) there is no operation \( j \) assigned on the station \( m_2 \) which can be assigned on station \( m_1 \) without violation of precedence constraints or the feasibility’s condition ”total load time of the station does not exceed the cycle time”.

Denote the maximization problem by \( \text{max-SALBP-1} \).

Theorem 1. \( \text{max-SALBP-1} \) is NP-hard in the strong sense (by reduction from the 3-Partition Problem).

Theorem 2. \( \text{max-SALBP-1} \) is not approximated with an approximation ratio \( \leq \frac{3}{2} \) unlike \( P = NP \).

An experimental study of maximal number of stations for benchmark instances published on http://www.assembly-line-balancing.de was done. The results show that the maximal founded deviation \( m^{\text{max}} - m^{\text{min}} \) does not exceed 20%.

Flat Graph of Precedence Relations

In [6] authors propose a transformation of graph \( G \) of precedence relations to planar one for the well-known Resource-Constrained Project Scheduling Problem. The same idea can be used for SALBP-1.

Theorem 3. For any instance of SALBP-1 with \( n \) operations and \( v \) precedence relations, there exists an analogous instance with a flat graph \( G' \) with \( n' \) operations and \( v' \) relations, where \( n + v \geq n' + v' \).

We obtain an analogous instance from the original one by adding ”dummy” operations (with \( t_j = 0 \)) and deleting all the unnecessary relations. According to the well-known Euler’s Theorem, \( v' \leq 3n' - 6 \) in such the planar graph.
The number of precedence relations influences running time and the theoretical complexity of solution algorithms. The number of precedence relations is estimated by different authors as $O(n^2)$ (i.e., the "Order strength" on http://www.assembly-line-balancing.de is estimated according to the number $n \cdot (n - 1)$ of precedence relations). If we consider only instances with planar graphs then the number of relations is $\leq 3n - 6$, i.e. $O(n)$. So, the fact mentioned in Theorem 3 allows us to reduce the run time of algorithms (by reduction of unnecessary relations) and estimate the complexity exacter.

The authors were supported by the Russian Foundation for Basic Research (projects no. 11-08-01321, 11-08-13121).

REFERENCES

Optimization of a multiple covering of a surface taking into account its relief

Shamil Galiev$^1$, Maria Lisafina$^2$, Vitalii Yudin$^3$

$^1$ National Research Technical University; Kazan, Russia; sh.galiev@mail.ru
$^2$ National Research Technical University; Kazan, Russia; e.marie@mail.ru
$^3$ National Research Technical University; Kazan, Russia; zeratustro.ru@gmail.ru

We propose a numerical technique for optimization of a $k$-fold ($k \geq 1$) covering of a bounded part of a given surface $G$ by some figures taking into account the relief of the surface $G$. The $k$-fold covering implies that each point of the surface $G$ is observable from at least $k$ centres of covering figures, while the segment connecting the observer with the observable object contains no points that lie below the surface $G$. For $k > 1$ we additionally require that the distance between centres of figures is not less than some given value. We construct models and propose algorithms for determining the minimal possible number of figures and the location of centres under the indicated conditions and the additional requirement. The obtained numerical results demonstrate the efficiency of the proposed technique.

The problem of covering a bounded plane parts by circles of minimal radii or that of covering a bounded plane parts by the least number of circles of a given radius widely studied in the literature. This is connected both with the mathematical interest to the mentioned combinatorial optimization problem and with various applications such as the problem of the choice of the quantity and the location of various stations (for example, those of cellular communication, medical ambulance, technical services, etc). One can find possible applications of covering problems and their solution methods, for example, in [1].

In problems of determining the quantity and location of components of forest fire control systems such as Fire Watch [5], Lesnoy Dozor (Forest Patrol) [6], or elements of lidar (LIght Detection and Ranging) systems [2] it is necessary to take into account the lay of land. The problem of determining the quantity and location of video cameras in security systems is also close (in a sense) to the mentioned problem. Thus, in many cases one encounters covering problems that take into account the observability
of the controllable object. One can treat such problems as problems of covering the surface taking into account its relief.

In navigation systems or image recognition ones it is necessary to provide a multiple covering by observation zones. Moreover, there is no need in placing more than one station or camera at one and the same point, because this does not help to solve the stated problems. Therefore, one has to impose an additional constraint on the minimal distance between centres of figures.

Let $E^3$ be a three-dimensional Euclidean space with a Cartesian coordinate system with axes $0x$, $0y$, and $0z$; let $P_1$ be a bounded domain in the plane $x0y$. Assume that a function $g(x, y), (x, y) \in P_1$, is Lipschitz on $P_1$ and its Lipschitz constant is known. We denote the surface defined by the function $g(x, y)$ by $G$. We treat a point $p \in G$ as observable from a point (a centre) $s \in G$, if the segment $[p, s]$ contains no points that lie below the surface $G$ and the length of the segment does not exceed a given value $r_1$. Let us state the following problem.

**Problem Z.** Determine the least number of centres $s_{q_1}, s_{q_2}, \ldots, s_{qm}$, $m \geq 1$, and their location on $G$ so as to make any point $p \in G$ observable from at least $k$ ($1 \leq k \leq m$) centres. Moreover, the distance between the observable point $p$ and the observing centre should not exceed $r_1$, and with $k > 1$ the minimal distances between these centres should be not less than the given value $r_2$.

The set of points $p \in G$ observable from $s_j$ form some figure $K_j$ in $G$; we treat $s_j$ as the centre of $K_j$, $1 \leq j \leq m$. If $G$ is a plane, this figure represents a circle of radius $r_1$.

We introduce a discrete analogue of the covering problem for $G$ and treat its solution as an approximate solution of the initial problem. On $P_1$ we construct a rectangular grid with some step $\Delta x$ along axes $0x$ and $0y$. Let $pp_1, \ldots, pp_n$ be the set of nodes of the constructed grid on $P_1$ ($pp_j \in P_1$, $1 \leq j \leq n$) and let points $pp_j$ have coordinates $x_j, y_j, z_j$ ($z_j = 0$), $1 \leq j \leq n$. Let points $p_j$ have the same coordinates $x_j, y_j$ as points $pp_j$, and let the coordinate $z_j$ equal $z_j = g(x_j, y_j) + \alpha_j$, $\alpha_j \geq 0$, $1 \leq j \leq n$. Now instead of the covering of the whole surface $G$ we consider the covering of the set of points $G_n = \{p_1, p_2, \ldots, p_n\}$. We assume that centres are located at some points of the set $S = \{s_1, s_2, \ldots, s_n\}$, where points $s_j$ have the same coordinates $x_j, y_j$ as points $pp_j$, and the coordinate $z$ equals $zs_j = g(x_j, y_j) + \beta_j$, $\beta_j \geq 0$, $1 \leq j \leq n$. We introduce values $\alpha_j$.
and $\beta_j$, $1 \leq j \leq n$, in order to take into account the fact that an observer (for example, an antenna or a video camera) can be located above the surface $G$, as well as the fact that the observable object can be located above $G$.

Let $d(s, t)$ be the Euclidean distance between points $s, t \in E^3$, and let $r_1$ and $r_2$ be some given positive values. For the set of points $p_i$ and $s_j$ we calculate distances $d_{ij} = d(p_i, s_j)$, $1 \leq i, j \leq n$. We find values $a_{ij}$, $1 \leq i, j \leq n$, by the following procedure.

Procedure W1:
1. Choose points $p_i$ and $s_j$ in sets $G_n$ and $S$, respectively, $1 \leq i, j \leq n$.
2. If $d(p_i, s_j) > r_1$, then we set $a_{ij} = 0$ and choose the next pair of points $p_i$ and $s_j$ in sets $G_n$ and $S$, respectively, $1 \leq i, j \leq n$.
3. If $d(p_i, s_j) \leq r_1$, then we find out whether at least one point of the segment $[p_i, s_j]$ is located below the surface $G$. If there are no such points, then we set $a_{ij} = 1$, otherwise we do $a_{ij} = 0$ and choose the next pair of points $p_i$ and $s_j$ in sets $G_n$ and $S$, respectively, $1 \leq i, j \leq n$.
4. If all possible points $p_i$ and $s_j$ of sets $G_n$ and $S$, respectively, $1 \leq i, j \leq n$, are considered already, then stop.

As a result of the implementation of Procedure W1, we find all values of $a_{ij}$, $1 \leq i, j \leq n$, which form an $n \times n$ matrix $A = (a_{ij})$ consisting of zeros and units.

Procedure W2.
1. Choose points $s_i$ and $s_j$ in the set $S$, $1 \leq i, j \leq n$.
2. If $d(s_i, s_j) \leq r_2$, then we set: a) $b_{ii} = m_i + 1$, where $m_i$ is a number of point $s_j$, $1 \leq j \leq n$, for which $d(s_i, s_j) \leq r_2$; b) $b_{ij} = 1$ if $i \neq j$; otherwise (for $d(s_i, s_j) > r_2$) we do $b_{ij} = 0$ choose the next pair of points $s_i$ and $s_j$ in the set $S$, $1 \leq i, j \leq n$.
3. If all possible points $s_i$ and $s_j$ of the set $S$, $1 \leq i, j \leq n$, are considered already, then stop.

As a result of the implementation of Procedure W2, we find all values of $b_{ij}$, $1 \leq i, j \leq n$, which form an $n \times n$ matrix $B = (b_{ij})$ consisting of zeros, units and $m_i + 1$, $1 \leq i \leq n$.

Introduce the following denotations:

$$t_j = \begin{cases} 
1, & \text{if the center of the } j\text{th figure is located at the point } s_j, \\
0 & \text{otherwise}.
\end{cases}$$
\[ t = (t_1, t_2, \ldots, t_n)^T, \text{ while } I \text{ and } V \text{ are } n\text{-dimensional vectors, namely, } I = (1, 1, \ldots, 1)^T \text{ and } V = (k, k, \ldots, k)^T. \]

Let us state the problem

\[ \min \left\{ \sum_{i=1}^{n} t_i : At \geq V, \ t \in \{0, 1\}^n \right\}. \] (1)

Problem (1) is the well-known covering problem, and in this case it means the determination of the least quantity of figures that form a covering of the set \( G_n \), taking into account the relief of the surface \( G \), and the determination of centres of covering figures on the set \( S \). In our talk we describe the solvability conditions for problem (1), using the results obtained in [4] and [3]. We also consider other variants of the problem, including the weighted one.

The problem that takes into account the additional requirement

\[ \text{minimal distances between centres is not less than } r_2 \] (2)

is stated analogously. In order to take into account the conditions (2), one has to impose the constraint \( Bt \leq M \), where \( M = (m_1 + 1, m_2 + 1, \ldots, m_n + 1)^T \). Note that we have not succeeded in finding such a condition for the covering problem in the related literature. We introduce this constraint for a \( k \)-fold \((k > 1)\) covering, but, if necessary, one can introduce it even for \( k = 1 \).

For the realization of the step 3 of Procedure W1 we used the Lipschitz property of the function \( g(x, y) \). Then for solving problem (1) we applied a well-known method based on the relaxation, the construction of the core problem, and its solution by one of known methods.

For solving Problem Z we developed a computer program (using CPLEX-11.2) and obtained numerical results. Consider the case when the domain \( P_1 \) is the unit square \( S \). Assume that the function \( g(x, y) = 0 \), i.e., the surface \( G \) coincides with \( S \). Note that in the \( k \)-fold covering problem, condition (2) can be fulfilled with certain values of \( r_2 \). For example, in the 2-fold covering of \( S \) by five circles, it is possible that the minimal radius of circles \( r_1 \approx 0.527 \), while condition (2) is fulfilled with \( r_2 \leq 0.336 \); the 3-fold covering of \( S \) by eight circles takes place with \( r_1 \approx 0.518 \), while condition (2) is fulfilled with \( r_2 \leq 0.264 \). But in many cases one should introduce condition (2) by using the matrix \( B \). The location of four circles...
of the radius $\sqrt{2}/4$ that form an 1-fold covering of the square $S$ is known. One can make sure that a 2-fold covering of the square $S$ by eight circles of the same radius is realized, when each of four circles is replaced with two ones. The centres of introduced circles coincide with those of the initial ones. Introduce constraint (2). With the same radius of circles as before, the 2-fold covering with $r_2 = 0.2$ requires 10 circles. If $r_2 = 0.28$, then we obtain a 2-fold covering of $S$ by eight circles of the radius $r_1 = 0.380$; with $r_2 = 0.2$ we obtain that radii of eight circles that ensure a 2-fold covering of $S$ should equal $r_1 = 0.367$. Therefore, the constraint imposed on the distance between centres leads to the increase of either the quantity of circles, or their radii. The relief of the surface $G$ is also essential; it strongly affects the number of circles even with the increase of their radii.

According to results of numerical experiments, the proposed models and their solution methods are rather efficient.

References

Stochastic subgradient barrier-multiplicative descent for entropy optimization

Alexander Gasnikov\textsuperscript{1,2}, Eugenia Gasnikova\textsuperscript{2}

\textsuperscript{1} Premolab MIPT, Moscow, Russia; \textsuperscript{2} FAMC MIPT, Moscow, Russia;

Consider an entropy-linear programming problem in the following general form:

\[
\begin{aligned}
- \sum_{k=1}^{m} x_k \ln(x_k/e) &\to \max_{\vec{x} \in A \cap \mathbb{R}^m_+} \\
A: \Lambda(\vec{x}) &= \vec{q} - T\vec{x} = \vec{0}_l; \\F(\vec{x}) &= \vec{d} - G\vec{x} \geq \vec{0}_w, l + w \leq m
\end{aligned}
\]

A family of dual barrier-multiplicative algorithms for finding a unique solution of (1) was suggested in [1]. This family contains Shelekhovskij, MART, GISM, Popkov’s algorithms and has the following form:

\[
\begin{aligned}
x^n_k &= \exp\left\{- \sum_{p=1}^{l} t_{pk}^n \lambda^n_p - \sum_{q=1}^{w} g_{qk}^n \mu^n_q \right\}, \quad k = 1, \ldots, m \\
\lambda^{n+1}_p &= \lambda^n_p - \gamma g_p^\lambda q_p - \sum_{k=1}^{m} t_{pk} x^n_k, \quad p = 1, \ldots, l \\
\mu^{n+1}_q &= \mu^n_q - \alpha \mu^n_q g^\mu_q d_q - \sum_{k=1}^{m} g_{qk} x^n_k, \quad q = 1, \ldots, w
\end{aligned}
\]

where $\gamma, \lambda > 0$ are sufficiently small constants,

\[
\{g_p^\lambda(\cdot)\}_{p=1}^{l}, \{g_q^\mu(\cdot)\}_{q=1}^{w} \in C^2
\]

are monotone increasing functions, equal to zero on a null vector.

A typical choice for those functions is

\[
g_p^\lambda(y) = - \ln(1 - y/q_p), \quad p = 1, \ldots, l; \quad g_q^\mu(y) = y/d_q, \quad q = 1, \ldots, w.
\]

\textbf{Theorem 1.} Assume that

1. $\exists \vec{z} \geq \vec{0}_m; \vec{q} - T\vec{z} = \vec{0}_l, \vec{d} - G\vec{z} \geq \vec{0}_w,$

2. Rows of matrices $T = \|t_{pk}\|_{p,k=1}^{l,m}$ and $G = \|g_{qk}\|_{q,k=1}^{w,m}$, if taken together, are independent.
Then the iterative process (2) is globally convergent.

The second assumption is usually not fulfilled in practice. So we will present a new proposition.

**Theorem 2.** Assume that there $\exists \vec{z} \geq \vec{0}_m : \vec{q} - T\vec{z} = \vec{0}_l, \vec{d} - G\vec{z} \geq \vec{0}_w$. Then there exist sufficiently small $\gamma, \lambda > 0$, such that the iterative process (2) is globally convergent.

Iterative algorithm (2) is particularly effective, if $T$ and $G$ are sparse matrices or if $l + w << m$. Then the algorithm (2) will have complexity $O(m + l + w)$ or $O(m \cdot (l + w))$ at the each step.

General formulation of (2) allows us to formulate also the stochastic versions of it. In practice, sometimes it is more computationally effective to zero out via some probabilistic algorithm most of the $\{g^\lambda_p(\cdot)\}^l_{p=1}, \{g^\mu_q(\cdot)\}^w_{q=1}$ components. This probabilistic algorithm will base on a current state of (2) and have a logarithmic complexity (see [4]). One should remember to scale the resulting vector in order to keep its norm close to the gradient norm (see [5]). So we end up with sort of a stochastic subgradient descent in a dual space (see [6], [7]).

The work was supported by RFBR 11-01-00494-a. The first author is partially supported by Laboratory for Structural Methods of Data Analysis in Predictive Modeling, MIPT, RF government grant, ag. 11.G34.31.0073.

**References**

6. http://www2.isye.gatech.edu/~nemirovs/
An Efficient Algorithm for Determining the Lower Convex Hull of a Finite Point Set in 3D

Dinh Thanh Giang\textsuperscript{1}, Phan Thanh An\textsuperscript{2}, Le Hong Trang\textsuperscript{3}

\textsuperscript{1} CEMAT, Instituto Superior Técnico, Av. Rovisco Pais 1049-001 Lisboa, Portugal; dtgiang@math.ist.utl.pt

\textsuperscript{2} Institute of Mathematics, 18 Hoang Quoc Viet Road, 10307 Hanoi, Vietnam; thanhan@math.ac.vn

\textsuperscript{3} Department of Electrical Engineering, K.U.Leuven, Kasteelpark Arenberg 10, 3001 Leuven, Belgium; le.hongtrang@esat.kuleuven.be

To date, the lower convex hull of a finite point set is determined from the entire convex hull of the set. There arises a question “Can we determine the lower convex hull of a finite point set without rely on the entire convex hull?”

In this paper, some properties of lower facets and lower convex hulls are given. Among the either, we show that the lower convex hull is wrapped by lower facets starting from an extreme edge of the lower convex hull. Then an efficient algorithm for determining the lower convex hull of a finite point set in 3D without the entire convex hull is presented. The actual run times on the set of random points (in uniform distribution) on a paraboloid show that our algorithm runs significantly faster than the incremental convex algorithm implemented in \cite{O'Rourke} and some versions of the gift-wrapping algorithm (see \cite{An} and \cite{Preparata}).

References


LP projection algorithm and Newton method for solving dual LP problems

Alexander I. Golikov

1 Computing Center RAS, Moscow, Russia; gol@ccas.ru

Large-scale LP problems usually have more than one solution. Such techniques as the simplex methods, interior point methods make it possible to obtain different solutions in the case of non-uniqueness. For example, the simplex method yields a solution belonging to a vertex of polyhedron. Some variants of the interior point method converge to a solution satisfying the strict complementary slackness condition.

LP projection method is close to the quadratic penalty function method and to the modified Lagrangian function method. This method yields the exact projection of a given point on the solution set of primal LP problem as a result of the single unconstrained maximization of an auxiliary piecewise quadratic concave function for any sufficiently large values of the penalty parameter. A generalized Newton method with a stepsize chosen using Armijo’s rule was used for unconstrained maximization. The proof of globally convergent finitely terminating generalized Newton method for piecewise quadratic function was giving in [1], [2]. LP projection method solves LP problems with a very large ($\approx 10^7$) number of variables and moderate ($\approx 10^5$) number of constraints [3], [4].

In a similar way, the exact projection of a given point on the solution set of the dual LP problem can be obtained by nonnegative constrained maximization of auxiliary quadratic function for sufficiently large but finite values of the penalty parameter [5]. Unfortunately the Newton method can not be applied to nonnegativity constrained maximization problem directly but there exists a very simple way to overcome this shortcoming.

Consider the primal linear program in the standard form

$$f_* = \min_{x \in X} c^\top x, \quad X = \{x \in \mathbb{R}^n : Ax = b, \ x \geq 0_n\}$$  \hspace{1cm} (P)$$

together with its dual

$$f_* = \max_{u \in U} b^\top u, \quad U = \{u \in \mathbb{R}^m : A^\top u \leq c\},$$  \hspace{1cm} (D)$$
where $A \in \mathbb{R}^{m \times n}$, $c \in \mathbb{R}^n$ and $b \in \mathbb{R}^m$ are given, $x$ is a primal variable and $u$ is a dual variable, $0_i$ denotes the $i$-dimensional zero vector.

Consider the problems of finding the least 2-norm projection $\hat{u}\_*$ of the point $\hat{u}$ on the solution set $U\_*$: \[
\frac{1}{2}\|\hat{u}_* - \hat{u}\|^2 = \min_{u \in U_*} \frac{1}{2}\|u - \hat{u}\|^2, \tag{1}
\]

\[
U_* = \{u \in \mathbb{R}^m : A^\top u \leq c, b^\top u = f_*\}.
\]

Here, the Euclidian norm of vectors is used, and $f_*$ is an a priori unknown optimal value of the objective function of the original LP problems $(P)$ and $(D)$.

Consider the following maximization problem on the positive orthant $\mathbb{R}_+^n$: \[
\max_{y \in \mathbb{R}_+^n} S_1(y, \alpha, \hat{u}), \tag{2}
\]

\[
S_1(y, \alpha, \hat{u}) = -c^\top y + \hat{u}^\top Ay - \frac{1}{2}\|\alpha b - Ay\|^2
\]

Here the scalar $\alpha$ is fixed. A solution $y(\alpha)$ to problem (2) is closely connected with the solution $\hat{u}_*$ to problem (1).

**Theorem 1.** [5] There exists $\alpha_*$ such that for all $\alpha \geq \alpha_*$ the unique least 2-norm projection $\hat{u}_*$ of a point $\hat{u}$ on $U_*$ is given by

\[
\hat{u}_* = \hat{u} + \alpha b - Ay(\alpha),
\]

where $y(\alpha)$ is a point maximizing $S_1(y, \alpha, \hat{u})$ on $\mathbb{R}_+^n$.

Theorem 1 makes it possible to replace problem (1), which contains an a priori unknown value $f_*$, by problem (2), which involves the half-interval $[\alpha_*, +\infty]$ instead of this value. This essentially simplifies the calculations. Note that the value $\alpha_*$ may be negative. This occur when the projection of the point $\hat{u}$ on the solution set $U_*$ coincides with the projection of this point on the feasible set $U$. The estimation of the threshold value $\alpha_*$ is given in [5].

**Theorem 2.** [5] For all $\alpha > 0$ and all $\hat{u} = u_* \in U_*$ an exact solution to primal problem $(P)$ is given by $u_* = y(\alpha)/\alpha$, where $y(\alpha)$ is a point maximizing $S_1(y, \alpha, u_*)$ on $\mathbb{R}_+^n$.

Unfortunately the Newton method can not be applied to nonnegativity constrained maximization problem (2) directly. By incorporating the
nonnegativity constraint $y \geq 0_n$ into the objective function of (2) as a penalty term, we have unconstrained maximization problem

$$\max_{y \in \mathbb{R}^n} \{-c^\top y + \hat{u}^\top Ay - \frac{1}{2}\|\gamma b - Ay\|^2 - \frac{\tau}{2}\|(-y)_+\|^2\},$$

where $\tau > 0$ is a penalty parameter. In this case we obtained the optimal solution $y$ only in limit as $\tau \to +\infty$. This property complicates the computation. There exists a very simple way to overcome this shortcoming.

Let a vector $w \in \mathbb{R}^{m+n}$ consists of two vectors $w^\top = [u^\top, v^\top]$, where $u \in \mathbb{R}^m$, $v \in \mathbb{R}^n$. Consider the following LP problem

$$f_* = \max_{w \in W} b^\top u,$$

where $W = \{u \in \mathbb{R}^m$, $v \in \mathbb{R}^n : A^\top u + v = c$, $v \geq 0_n\}$, which is equivalent to dual problem ($D'$). The solution set of this problem is denoted by $W_* = [U_* \times V_*]$. For a given point $\hat{w}$ we find the least 2-norm projection $\hat{w}_*$ on $W_*$ as a solution to following minimization problem

$$\frac{1}{2}\|\hat{u}_* - \hat{u}\|^2 + \frac{1}{2}\|\hat{v}_* - \hat{v}\|^2 = \min_{w \in W_*} \left\{ \frac{1}{2}\|u - \hat{u}\|^2 + \frac{1}{2}\|v - \hat{v}\|^2 \right\},$$

where $\hat{W}_* = \{u \in \mathbb{R}^m$, $v \in \mathbb{R}^n : A^\top u + v = c$, $v \geq 0_n$, $b^\top u = f_*\}$. Similarly to the approach which was used above we come to following unconstrained maximization problem

$$\max_{y \in \mathbb{R}^n} S^2(y, \gamma, \hat{w}),$$

where $S^2(y, \gamma, \hat{w}) = -c^\top y + \hat{u}^\top Ay - \frac{1}{2}\|\gamma b - Ay\|^2 - \frac{1}{2}\|(-y)_+\|^2$.

The following Theorems hold [7].

**Theorem 3.** There exists $\gamma_*$ such that for all $\gamma \geq \gamma_*$ the unique least 2-norm projection $\hat{w}_*^\top = [\hat{u}_*^\top, \hat{v}_*^\top]$ of a point $\hat{w}^\top = [\hat{u}^\top, \hat{v}^\top]$ on $W_*$ is given by

$$\hat{u}_* = \hat{u} + \gamma b - Ay(\gamma),$$

$$\hat{v}_* = (\hat{v} - y(\gamma))_+,$$

where $y(\gamma)$ is a solution to unconstrained maximization problem (3).
Theorem 4. For all $\gamma > 0$ and $\hat{w} = w_0 \in W^*$ an exact solution to primal problem $(P)$ is given by $x^*_\gamma = y(\gamma)/\gamma$, where $y(\gamma)$ is a solution to unconstrained problem $(3)$.

To solve the primal and dual LP problems simultaneously, one can use the following iterative process:

$$y_{s+1} \in \arg \max_{y \in \mathbb{R}^n} \{-c^T y + u_s^T A y - \frac{1}{2} \| \gamma b - A y \|^2 - \frac{1}{2} \| (v_s - y) + \|^2 \},$$

$$u_{s+1} = u_s + \gamma b - A y_{s+1},$$

$$v_{s+1} = (v_s - y_{s+1})_+.$$  
(4) 
(5) 
(6)

Theorem 5. For all $\gamma > 0$ and for arbitrary starting point $w_0$ the iterative process (4)–(6) converges to $w^*_\gamma \in W^*$ in a finite number of iterations $\sigma$. The formula $x^*_\gamma = y_{\sigma+1}/\gamma$ gives an exact solution to the primal problem $(P)$.

The goal function $S^2(y, \gamma, \hat{w})$ of unconstrained maximization problem $(3)$ is piecewise quadratic concave function. Therefore one can define its generalized Hessian which is $m \times m$ symmetric negative semi-definite matrix:

$$\frac{\partial^2}{\partial y^2} S^2(y, \gamma, \hat{w}) = -A^T A - D(z).$$

Here, $D(z)$ denotes the $n \times n$ diagonal matrix with diagonal elements $z^i, i = 1, 2, \ldots, n$. If $(\hat{v} - y)^i > 0$ then $z^i = 1$, if $(\hat{v} - y)^i \leq 0$ then $z^i = 0$. Now for solving problem $(3)$ one can use generalized Newton method.

There is an important difference between problem $(2)$ and problem $(3)$. In the first case we look for the projection of a given point $\hat{u}$ on $U^*$, in the second case we project a point $\hat{W} = [\hat{u}, \hat{v}]$ on the solution set $W^*$. Let $\hat{u}^1_*$ and $\hat{u}^2_*$ denote the projections of point $\hat{u}$ and $[\hat{u}, \hat{v}]$ in the first and second cases, respectively. Then following inequality holds:

$$\| \hat{u}^1_* - \hat{u} \| \leq \| \hat{u}^2_* - \hat{u} \|.

The comparison of LP projection methods (which were implemented in MATLAB) with some well-known commercial (CPLEX) and research software packages showed that they are competitive with the simplex and the interior point methods [4].

Several parallel versions of the generalized Newton method for solving linear programs based on various data decomposition schemes of matrix

97
A (column, row, and cellular schemes) were implemented (see [6]). The resulting parallel algorithms were successfully used to solve large-scale LP problems (up to several dozens of millions of variables and several hundreds of thousands of constraints) for a relatively dense matrix $A$. The computational experiments were performed on the cluster consisting of two-processor nodes based on 1.6 GHz Intel Itanium 2 processors connected by Myrinet 2000. For example, for LP problem with one million variables and 10000 constraints, the cellular scheme for 144 processors of the cluster accelerated the computations approximately by a factor of 50, and the computation time was 28 s. LP problem with two million variables and 200000 constraints was solved in about 40 min. on 80 processors. Another LP problem with 60 million variables and 4 thousand constraints was solved by column scheme in 140 s. on 128 processors.

This research was supported by the Programs of Fundamental Research of Russian Academy of Sciences P-15, P-18, by the Russian Foundation for Basic Research (grant no.11-01-00786), and by the Leading Scientific Schools Grant no.5264.2012.1

REFERENCES

Many-Person Games With Convex Structure

Evgenij Golshtejn

1 CEMI RAS, Moscow, Russia; golshtn@emi.rssi.ru

Let $X_i$ be a nonempty subset of a Euclidean space $E_i$, for $1 \leq i \leq k$, $k \geq 2$, and $\hat{X}_i$ the direct product of the subsets $X_j$ for $1 \leq j \leq k$, $j \neq i$, $X = X_i \times \hat{X}_i$, $E = E_1 \times \cdots \times E_k$; and, consider real-valued functions $f_i$ is defined on $X$, for $1 \leq i \leq k$.

Introduce a $k$-person non-cooperative game $\Gamma$ defined by the $i$-th player’s strategy set $X_i$ and payoff function $f_i$, for each $1 \leq i \leq k$. We say that the game $\Gamma$ satisfies condition $C_1$ if $X$ is a convex compact, and the payoff function $f_i$ is continuous over $X$, concave and Lipschitz with respect to $x_i \in X_i$ for any (fixed) $\hat{x}_i \in \hat{X}_i$ with the Lipschitz constant independent on $\hat{x}_i \in \hat{X}_i$, for any $1 \leq i \leq k$.

If condition $C_1$ holds, the game $\Gamma$ has a nonempty set $X^*$ of Nash equilibrium points, and each payoff function $f_i$ is super-differentiable with respect to $x_i \in X_i$. Further, define a point-to-set mapping $T_\Gamma: X \to 2^E$ as

$$T_\Gamma(x) = \{t = (t_1, \ldots, t_k): -t_i \in \partial_{x_i}f_i(x), \ 1 \leq i \leq k\}, \ x \in X,$$

where $\partial_{x_i}f_i(x)$ is the super-differential of $f_i$ with respect to $x_i$ calculated at the point $x$.

Since the set $X^*$ of Nash equilibrium points of the game $\Gamma$ coincides with the set of solutions to the variational inequality

$$t \in T(x), \ \langle t, x' - x \rangle \geq 0, \ \forall x' \in X$$

if $T = T_\Gamma$, then the methods solving variational inequalities can be used to find the elements of the set $X^*$. In our previous paper [1], we described an efficient numerical algorithm to solve variational inequalities. Its convergence for $T = T_\Gamma$ is guaranteed if $\Gamma$ satisfies condition $C_1$ and the mapping $T_\Gamma$ is monotone.

We say that a non-cooperative game $\Gamma$ satisfying condition $C_1$ has a convex structure if the mapping $T_\Gamma$ is monotone. Any antagonistic game with condition $C_1$ valid has a convex structure; however, for non-antagonistic games it is not true in general.
A game $\Gamma$ is said to satisfy condition $C_2$ whenever condition $C_1$ holds and, apart from that, the $i$-th player’s payoff function $f_i$ is convex with respect to $\hat{x}_i \in \hat{X}_i$ for any $x_i \in X_i$, $1 \leq i \leq k$, whereas the total payoff $f(x) = \sum_{i=1}^{k} f_i(x)$ of the game $\Gamma$ is a concave function over $x \in X$.

**Theorem 1.** If the payoff functions $f_i$ of the game $\Gamma$ can be represented as $f_i(x) = f_0^i(x) + \varphi(\hat{x}_i)$, $x \in X$, $1 \leq i \leq k$, where the functions $f_0^i(x)$, $x \in X$, $1 \leq i \leq k$, define a game satisfying condition $C_2$, then the game $\Gamma$ has a convex structure.

Consider a finite non-cooperative game with $k$ players, where player $i$ has $n_i$ strategies, and his payoff being determined by a $k$-dimensional table $A_i = (a_i^{(s_1\ldots s_k)})$; here $a_i^{(s_1\ldots s_k)}$ is the $i$-th player’s payoff when player $\alpha$ chooses strategy $s_\alpha$, $1 \leq \alpha \leq k$. Having extended the players’ strategy sets by allowing mixed strategies, we come to the game $\Gamma$, in which

$$X_i = \{x_i = (x_{i1}, \ldots, x_{in_i}): \sum_{j=1}^{n_i} x_{ij} = 1, \ x_{ij} \geq 0\},$$

$$f_i(x) = \sum_{s_1\ldots s_k} A_i^{(s_1\ldots s_k)} x_{i1}s_1 \times \cdots \times x_{ik}s_k,$$

$$1 \leq i \leq k, \ x = (x_1, \ldots, x_k).$$

**Theorem 2.** Let $\Gamma$ be a finite non-cooperative game with mixed strategies defined by tables $A_i$, $1 \leq i \leq k$. The game $\Gamma$ has a convex structure if and only if the tables $A_i$ can be represented in the following form $A_i = \sum_{j=1}^{k} A_{ij}$, $1 \leq i \leq k$, where the entries of the $k$-dimensional table $A_{ij}$ depend only upon the indices $s_i$ and $s_j$ when $i \neq j$, but do not depend on the index $s_j$ for $i = j$. Moreover, for any $i \neq j$, all entries of the table $A_{ij} + A_{ji}$ are zero.

Theorem 2 implies that the requirements to a game $\Gamma$ listed in Theorem 1 guaranteeing that $\Gamma$ has a convex structure, are not only sufficient but are also necessary if the game $\Gamma$ belongs to the class of finite non-cooperative many-person games with mixed strategies. Furthermore, for this class of games, the assumptions that the payoff functions $f_0^i$ are concave with respect to $x_i$ and convex by $\hat{x}_i$, whereas the function $f^0 = \sum_{i=1}^{k} f_0^i$ is concave with respect to $x$, turn out to be equivalent to the following properties: the functions $f_0^i$ are linear by $x_i$ on $E_i$ and affine with respect to $\hat{x}_i$ over $\hat{X}_i$, while $f^0$ is affine on $X$. 

100
Proximal Analysis and Regularity of Viscosity Solution to some Hamilton-Jacobi Equation

Vladimir Goncharov\textsuperscript{1}, Fátima Pereira\textsuperscript{2}

\textsuperscript{1} Universidade de Évora, CIMA-UE. Évora, Portugal; goncha@uevora.pt
\textsuperscript{2} Universidade de Évora, CIMA-UE. Évora, Portugal; fmfp@uevora.pt

Our talk is devoted to the Hamilton-Jacobi equation of a special form:

\begin{equation}
\rho_F (\nabla u(x)) - 1 = 0,
\end{equation}

where $F$ is a closed convex bounded subset of a Hilbert space $(H, \| \cdot \|)$, containing the origin in its interior, and $\rho_F(\cdot)$ is the Minkowski functional (gauge function) associated to $F$,

\[ \rho_F (\xi) := \inf \{ \lambda > 0 : \xi \in \lambda F \}. \]

This type equations arise, e.g., in geometric optics such as the famous eikonal equation

\[ \| \nabla u(x) \| - a = 0, \]

$a > 0$, in $H = \mathbb{R}^3$, or, in general, the elliptic equation

\[ \sum_{i=1}^{3} c_i^2 u_{x_i}^2 + \frac{2}{c} \langle \overrightarrow{v}, \nabla u \rangle - 1 = 0, \]

describing the propagation of a light wave from a point source placed at the origin in anisotropic medium moving with a constant velocity $\overrightarrow{v}$. Here
are the (constant) coefficients of refraction of light rays parallel to the coordinate axes, and \( c \) means the speed of the light in a vacuum.

We are interested in regularity properties of the viscosity solution \( \hat{u}(\cdot) \) to the equation (1) on an open region \( \Omega \subset H \) with the boundary data \( \hat{u}|_{\partial \Omega} = \theta \) where the function \( \theta(\cdot) \) satisfies a kind of slope condition with respect to \( F \). It is well known that in this case the viscosity solution takes the form

\[
\hat{u}(x) = \inf_{y \in C} \{ \rho_{F^0}(x-y) + \theta(y) \},
\]

where \( C := H \setminus \Omega \), and it can be interpreted also through the value function in a minimum-time problem with a dynamics associated to the polar set \( F^0 \).

Since, in general, viscosity solution is not smooth everywhere out of \( C \), our goal is to study the (Fréchet) differentiability of \( \hat{u}(\cdot) \) just near \( C \) (or near the boundary \( \partial \Omega \)).

Notice that the differentiability of the function (2) strongly relates with the existence, uniqueness and regularity of minimizers in the respective (mathematical programming) problem. In the case of compact \( C \) it follows, for instance, from the representation of the Clarke’s subdifferential of a marginal function through Radon measures supported on the set of minimizers (see [1, Section 2.8]). So, denoting by \( \pi_{F,\theta}^{C}(x), x \notin C \), the set of minimizers in (2) we find first rather general conditions guaranteeing that \( \pi_{F,\theta}^{C}(x) \) is a singleton (Lipschitz) continuous near a given point \( x_0 \in \partial C \). These conditions involve a balance between the (proximal) subgradients of the restriction \( \theta|_{C} \) and the normals to \( F \).

Then, under the same (local) hypotheses assuming, in addition, the smoothness either of \( \theta|_{C} \) or of \( F \) we prove that the function \( \hat{u}(\cdot) \) is (Fréchet) differentiable in \( (x_0 + \delta B) \setminus C \) for some \( \delta > 0 \) (with eventually Hölder continuous gradient \( \nabla \hat{u}(\cdot) \)).

To achieve our goals we apply the mixed technique of Convex Analysis with some tools of Proximal Calculus (see [2, Chapter 1]) such as the fuzzy sum rule and the Ekeland’s variational principle. Observe that the basic local hypothesis naturally generalizes one of the geometric conditions obtained in [3, 4] for the well-posedness of the minimum-time problem
with a constant convex dynamics (the case $\theta \equiv 0$). Some illustrative examples in finite dimensional spaces are given.

The authors were supported by Fundação para Ciência e Tecnologia (FCT), the Portuguese institutions COMPETE, QREN and the European Regional Development Fund (FEDER) (project PTDC/MAT/111809/2009 “Variational Analysis: Theory and Applications”).

**References**


**Multimethod’s algorithm for parametric identification of nonlinear dynamic systems**

Aleksander Gornov

\[ ISDCT SB RAS, Irkutsk, Russia; gornov@icc.ru \]

The problem of parametrical identification of dynamic systems continues to remain an actual problem of mathematical modeling. The main missions arising at the solution of task of parametrical identification of dynamic systems, problems of nonconvexity of optimized functionals and possible stiffness of studied system of the differential equations continue to remain.

The report examines the multimethod’s algorithms are oriented to the solution of applied problems arising in the numerical study of dynamical models. As basis of work the previous researches of author published in [1-6].
For numerical integration of dynamic systems the set of algorithms including both options of the elementary methods of type Runge-Kutta, and the adaptive program technologies developed by experts DOPRI5 and DOPRI8 for nonstiff systems, RADAU and RADAU5 for stiff is realized. At digitization of system will be applied a grid with the non-uniform step, generated by applied algorithms individually to each of the Cauchy problem. Accuracy of integration is supervised by the specialized algorithm, capable to estimate both an error of digitization of continuous system and an error of algorithms of integration.

One of the main installations applied by development of new algorithms, refusal from traditional in global optimization, but low-constructive for this class of problems of hypotheses was considered. In particular, due to lack of the possibility of obtaining analytical formulas for the derivatives of the algorithms required for adaptive technology assessment of the difference gradients. On the other hand, calculation of the functionals defined on solutions of the Cauchy problem, does almost impossible use of reliable interval methods and doesn’t allow to receive aprioristic estimates of the Lipschitz constant.

To solve the nonconvex optimization problems formulated in the identification system, developed a standardized set of optimization algorithms, including both the modification of known methods - multistart, simulated annealing, genetic search, differential evolution, MSBH (Monotonic Sequence Basin-Hopping), and new algorithms by taking into account the specific characteristics of this class of problems.

Among the new algorithms may be mentioned the ”algorithm of parabolas,” based on the nonlocal one-dimensional search algorithm with parabolic convex triples and refinement of descent, ”the tunneling algorithm”, scanning the search space by means of spline approximation to a random direction, ”pass-algorithm” that implements the search in the neighborhood extreme points of the new shutter with a concave profile, and the modification of the Powell algorithm, which allows jump out from local extrema.

For specification of the found local extrema some local algorithms, allowing to specify the decision without use of the gradients which are inevitably containing integration noise are realized. The set of the described base algorithms formed a basis for creation of multimethod computing schemes.

The realized interface system allows to make trial calculations in a
dialogue mode and to design for each task the unique algorithm, capable to make active the best lines of each base algorithm. In particular, in many tasks the combination of algorithms of differential evolution, local descent on the Cauchy method and algorithm of MSBH perfectly proved.

The proposed algorithms are standardized, equipped with algorithmic parameters and realized in language C in the form of a scientific prototype of the program complex functioning under control of OS Windows. The efficiency of the algorithms studied on a number of test problems. With application of the offered algorithms a number of applied tasks from ecology and nanophysics area is solved. The maximum size of system of the differential equations for which problems of parametrical identification are solved, makes 550 phase variables.

The authors were supported by the Russian Foundation for Basic Research (project no. 12-01-00193, no. 10-01-00595).

References

Optimization Problems in Astrodynamics

Anna D. Guerman

1CAST - Centre for Aerospace Science and Technologies, University of Beira Interior, Covilha, PORTUGAL; anna@ubi.pt

Planning of a space mission has to take into account several limitations on the available resources and restrictions on possible solutions, so large spectrum of optimization problems naturally arise in the context. Since the beginning of the Space Era such problems as optimal reorientation maneuver or minimum time trajectory transfer have been in the focus of attention of researchers; their successful solution made possible a number of breathtaking missions in early 60th and still are the source of inspiration for many space scientists and engineers. Nowadays, the equipment and energy available on-board offer one much more options for in-orbit maneuvers, and many new optimization methods have been developed to make the best of new technology.

We present a review of a number of problems related to analysis of modern space missions, from attitude maneuvers to orbital control.

Blaschke Convergence Theorem for G-type Convex Sets in Metric Spaces

Nguyen Ngoc Hai, Phan Thanh An

1Department of Mathematics, International University, Vietnam National University, Ho Chi Minh City, Vietnam; nnhai@hcmiu.edu.vn
2Institute of Mathematics, Vietnam Academy of Science and Technology, 18 Hoang Quoc Viet Road, 10307 Hanoi, Vietnam thanhan@math.ac.vn

The Blaschke convergence theorem is a classical theorem in convex analysis. It says that a uniformly bounded infinite collection of closed
convex sets in a finite dimensional space contains a sequence which converges to a nonempty compact convex set. There are generalizations of this theorem for infinite dimensional normed linear spaces (see [7]) or for sets that are not convex in the usual sense such as star-shaped sets (see [5]). In [4], we dealt with analytic properties of geodesic convex sets in a simple polygon. Among the others, Blaschke’s convergence theorem for geodesic convex sets was presented. The aim of this paper is to develop some basic ideas in [4] to get a generalization of Blaschke’s convergence theorem in metric spaces. First, we will introduce the notion of $G$-type convexity. The usual convexity in linear spaces and geodesic convexity in uniquely geodesic spaces are special cases of $G$-type convexity. Some properties of convex sets are still true for $G$-type convex sets (Propositions 1 and 3 (ii)). We then get a generalization of Blaschke’s convergence theorem for metric spaces (Theorem 7). In particular, in a proper uniquely geodesic space, every uniformly bounded sequence of nonempty geodesically convex sets contains a subsequence which converges to a nonempty compact geodesically convex set (Corollary 8). Theorem 7 also reduces to the classical Blaschke convergence theorem when $X$ is a finite dimensional normed linear space and $G(x, y)$ is the line segment joining $x$ and $y$.

**Definition 1.** A generalized geodesic space is a metric space $(X, d)$ together with a set-valued mapping $G : X \times X \to 2^X$ that satisfy the following conditions:

1. $G(x, y) \neq \emptyset$ for all $x, y \in X$;
2. If $(x_n)$ and $(y_n)$ are sequences in $X$, $x_n \to x$ and $y_n \to y$, then

$$d_H(G(x_n, y_n), G(x, y)) \to 0 \quad \text{as} \quad n \to \infty$$

where the metric on a normed linear space $(X, \| \cdot \|)$ is the usual metric $d(x, y) = \|x - y\|$ and $d_H$ is the Hausdorff metric. We denote this generalized geodesic space by $(X, d, G)$. Then, any normed linear space is a generalized geodesic space with $G(x, y) = [x, y]$, any nonempty convex set in a normed linear space is a generalized geodesic space.

**Definition 2.** Let $X$ be a set and $G : X \times X \to 2^X$ a set-valued mapping that takes on nonempty values. A set $A \subset X$ is called $G$-type convex if $G(x, y) \subset A$ whenever $x, y \in A$.

If $X$ is a linear space and $G(x, y) = [x, y]$, then $G$-type convexity coincides
with the usual convexity.

**Proposition 1.** The intersection of an arbitrary family of $G$-type convex sets is $G$-type convex.

**Proposition 2.** If $A$ is a nonempty $G$-type convex set in a generalized geodesic space $(X, d, G)$ then $(A, d, G)$ is a generalized geodesic space, too.

**Proposition 3.** Let $(X, d, G)$ be a generalized geodesic space.

(i) If $x_n, y_n \in X$, $x_n \to x$, $y_n \to y$ and $z \in G(x, y)$ then there is $z_n \in G(x_n, y_n)$, $n = 1, 2, \ldots$ such that $z_n \to z$ as $n \to \infty$.

(ii) If $A$ is a $G$-type convex subset of $X$, then so is its closure $\bar{A}$.

**Corollary 4.** (See [6], p. 68) Let $X$ be a proper uniquely geodesic space and let $A$ be a geodesically convex subset of $X$. Then its closure $\bar{A}$ is geodesically convex.

Consider now the collection $G(X) = \{G(x, y) : x, y \in X\}$.

**Theorem 5.** Let $(X, d, G)$ be a compact generalized geodesic space.

(i) For any sequence $(G(x_n, y_n))$, there exists a subsequence $(G(x_{n_j}, y_{n_j}))$ such that $d_H(G(x_{n_j}, y_{n_j}), G(x, y)) \to 0$ for some $x, y \in X$. So if $G(u, v)$ is closed for all $u, v \in X$ then $(G(X), d_H)$ is a compact metric space.

(ii) If $A$ is a nonempty compact subset of $X$ and $d_H(G(x_n, y_n), A) \to 0$, then $A = \overline{G(x, y)}$ for some $x, y \in X$.

**Proposition 6.** Let $(A_n)$ be a sequence of nonempty $G$-type convex subsets of a generalized geodesic space $(X, d, G)$ and let $A$ be a nonempty closed subset of $X$. If $d_H(A_n, A) \to 0$, then $A$ is also $G$-type convex.

Let us denote by $GC(X)$ the collection of all nonempty $G$-type convex closed subsets of a generalized geodesic space $(X, d, G)$. A sequence of sets in $X$ is said to be uniformly bounded if there exists some ball in $X$ that contains every member of the sequence (see [10], p. 96). We can now formulate our main result of this paper.

**Theorem 7.** Let $(X, d, G)$ be a generalized geodesic space.

(i) If $X$ is compact, then so is $(GC(X), d_H)$.

(ii) If $X$ is proper, then every uniformly bounded sequence of nonempty $G$-type convex sets in $X$ contains a subsequence which converges to some
nonempty compact $G$-type convex set in $X$.

A particular case of Theorem 7 (i) when $X$ is a simple polygon $P$ on the plane and $G(x, y)$ is the shortest segment connecting $x$ and $y$ was proved in [4]. We now have a direct and important consequence of Theorem 7.

**Corollary 8.** If $X$ is a proper uniquely geodesic space then every uniformly bounded sequence of nonempty geodesically convex subsets of $X$ contains a subsequence which converges to some nonempty compact geodesically convex subset in $X$.

In fact, in Theorem 7 and Corollary 8 $X$ may be any metric space provided that there is a compact set $Y$ containing all members of the sequence $(A_n)$. The proof of Theorem 7 (ii) remains valid for this case.

**Proposition 9.** Let $(A_n)$ be a sequence of nonempty $G$-type convex subsets of a generalized geodesic space $(X, d, G)$. If there exists a compact set that contains all members of $(A_n)$, then there exists a subsequence of $(A_n)$ which converges to some nonempty compact $G$-type convex set in $X$.

**Corollary 10.** (See [7]) If $C$ is a nonempty compact convex set in a normed linear space, then every sequence of nonempty compact convex subsets of $C$ contains a subsequence which converges to some nonempty compact convex set in $C$.

Finally we can obtain the classical Blaschke convergence theorem for convex sets in $\mathbb{R}^n$, which is a special case of Theorem 7 (ii).

**Corollary 11.** (See [10], p. 97) Every uniformly bounded sequence of nonempty compact convex sets in $\mathbb{R}^n$ contains a subsequence which converges to some nonempty compact convex set in $\mathbb{R}^n$.

**References**

Pathwise optimal control of diffusion type processes

Niyaz Ismagilov\(^1\), Farit Nasyrov\(^2\)

\(^1\) Ufa State Aviation Technical University, Ufa, Russia; niyaz.ismagilov@gmail.com
\(^2\) Ufa State Aviation Technical University, Ufa, Russia; farsagit@yandex.ru

Introduction

In the work, we consider a stochastic optimal control problem of diffusion type processes with pathwise cost functional, that is, the problem of finding a control function such that it minimizes cost for every single trajectory of state variable.

We introduce new method of solving problems of pathwise cost minimization. The main idea of the method is that original stochastic control problem can be reduced to deterministic control problem. Solution to the latter gives pathwise optimal solution to the original problem.

Problem Statement

Let \((\Omega, F, P)\) be a probability space. Consider a stochastic differential equation defined on this space describing dynamics of some system

\[
\frac{dx_t}{dt} = b(t, x_t, u_t) dt + \sigma(t, x_t, u_t) \ast dW_t, \quad x(0) = x_0.
\]  

In the above equation \(t \in [0, T] \subset \mathbb{R}\); state variable function \(x_t\) is a real-valued stochastic process; control function \(u_t\) is a stochastic process taking values in \(V \subset \mathbb{R}\) with paths in class of piecewise continuous functions; \(b(t, x, u)\) and \(\sigma(t, x, u) : [0, T] \times \mathbb{R} \times V \to \mathbb{R}\) are known determined functions, which we will call drift coefficient and diffusion coefficient respectively; \(W_t(\omega)\) is a standard Wiener process and its differential \(\ast dW_t\) in right-hand side of (1) is interpreted in the sense of Stratonovich integral.

We introduce a cost functional to measure performance of control

\[
J = \int_0^T f^0(t, x_t, u_t) dt.
\]

The problem is to find control that minimizes the functional (2) subject to dynamics equation (1).
Previous work on optimal control of diffusion processes has mainly been concerned with problem of finding control that minimizes “mean” value of cost

\[ \bar{J} = E \int_0^T f^0(t, x_t, u_t) dt \to \min, \]

where \( E \) denotes expectation \([1],[2]\). In contrast, stated above problem is concerned with minimization of functional (2), which represents cost for every single path of state variable. Hence pathwise optimality is the distinguishing feature of this work.

In the work we consider two particular cases of problem (1), (2):

A  Control function only affects on drift coefficient, i.e. instead of (1) dynamics of system is

\[ dx_t = b(t, x_t, u_t) \ dt + \sigma(t, x_t) * dW_t, \quad x(0) = x_0. \]  \( (3) \)

The problem is to minimize functional (2) subject to (3).

B  Control function consist of two components \( u^1_t \) and \( u^2_t \) which affect on drift coefficient and diffusion coefficient respectively, and the diffusion coefficient is linear w.r.t. \( u^2_t \), i.e. instead of (1) dynamics of system is

\[ dx_t = b(t, x_t, u^1_t) \ dt + \sigma(t, x_t)u^2_t * dW_t, \quad x(0) = x_0. \]  \( (4) \)

The problem is to minimize functional

\[ J = \int_0^T f^0(t, x_t, u^1_t, u^2_t) dt. \]  \( (5) \)

**Solution to the problems**

**Solution to problem A.** It’s known ([3]) that \( x_t \) is a solution to the equation (3) if and only if it can be represented in the form

\[ x_t = \varphi^*(t, W_t + y_t), \]  \( (6) \)

where \( \varphi^*(t, v) \) is known function determined by diffusion coefficient and not depending on control or state variable, and \( y_t \) is a solution to the following Cauchy problem

112
\[
\frac{dy_t}{dt} = \frac{b(t, \varphi^*(t, W_t + y_t), u_t) - (\varphi^*)'(t, W_t + y_t)}{\sigma(t, \varphi^*(t, W_t + y_t))}, \quad (7)
\]
\[
y(0) = y_0 : \quad \varphi^*(0, W_0 + y_0) = x_0, \quad (8)
\]

Here \((\varphi^*)'(t, W_t + y_t)\) is partial derivative of \(\varphi^*(t, v)\) w.r.t. \(t\) evaluated at point \((t, W_t + y_t)\). Hence (3) is equivalent to (6)–(8).

As function \(\varphi^*(t, v)\) doesn’t depend on control we conclude that all the affect of control on state variable \(x_t = \varphi^*(t, W_t + y_t)\) is focused on \(y_t\). The latter allows to consider constrain (7) instead of (3) leading us to the new problem with state variable \(y\), same control function, dynamics equation (7) and cost functional

\[
J = \int_0^T f^0(t, \varphi^*(t, W_t + y_t), u_t) dt, \quad (9)
\]

that is (6) substituted in (2).

Thus, problem (3), (2) is equivalent to problem (7)–(9) which has an advantage of having non-stochastic differential equation as dynamics of system. That is, despite having random functions (these are \(W_t(\omega), y_t(\omega)\) and \(u_t(\omega)\)) in the right-hand side of equation (7), for almost all \(\omega \in \Omega\) the equation is an ordinary differential equation of the form \(dy_t/dt = f(t, y_t, u_t)\). Hence the problem (7)–(9) is deterministic optimal control problem and can be solved using classical deterministic methods as e.g. maximum principle.

**Solution to problem B.** The idea of solution to problem B is similar to the one presented for problem A, but with a slightly different result. Solution to equation (4) is presented as \(x_t = \varphi^*(t, W_t u_t^2 + y_t)\) (compare to (6)), where \(\varphi^*(t, v)\) is again known function determined by diffusion coefficient \(\sigma\) and not depending of control or state variable, and \(y_t\) is a solution to the following Cauchy problem

\[
\begin{cases}
\frac{dy_t}{dt} = \frac{b(t, \varphi^*(t, W_t u_t^2 + y_t), u_t^2) - (\varphi^*)'(t, W_t u_t^2 + y_t)}{\sigma(t, \varphi^*(t, W_t u_t^2 + y_t)) u_t^2} - W_t \nu_t, \\
\frac{du_t^2}{dt} = \nu_t.
\end{cases} \quad (10)
\]

\[
y(0) = y_0, \quad u^2(0) = u_0^2. \quad (11)
\]
In the same way as in solution to problem A, a new problem with dynamics given by (10), (11) and cost functional

$$J = \int_0^T f^0(t, \varphi^*(t, W_t u^2_t + y_t), u^1_t, u^2_t) dt,$$

is equivalent to problem (4), (5).

The difference to the case of problem A is that in reduced problem (10)–(12) instead of single equation of dynamics there is a system of two differential equations (10) with generalized derivatives, and $u^2_t$ is treated as a state variable among with $y_t$. Control functions are $u^1_t$ and $\nu_t$, the latter is an impulse control function, i.e. represented as sum of absolutely continuous function and linear combination of Dirac delta functions

$$\nu_t = \nu^c_t + \sum_i h_i \delta(t - \tau_i).$$

Thus problem B reduces to deterministic optimal impulse control problem (10)–(12) which can be solved using corresponding methods.

**Conclusion**

The method introduced in work provides capability of solving pathwise stochastic optimal control problems. Reduction of initial stochastic problem to the new deterministic problem lies at the heart of the method. It is shown that problems with control affecting only drift coefficient reduce to classical deterministic optimal control problems and problems with control affecting both drift and diffusion coefficients reduce to deterministic optimal impulse control problems.

Reduced problems can be solved using deterministic methods as e.g. maximum principle. The main difficulty in applying deterministic methods to reduced problems is that coefficients of the equations (7), (10) may have unbounded variation in $t$, as a result be not differentiable. Furthermore, optimal control found by this method not always turns out to be non-anticipating.

**References**

On accuracy of the regularization method of constrained ill-posed quadratic minimization problems

M. Jačimović¹, I. Krnić²

¹ Montenegrin Academy of Sciences and Arts, Podgorica, Montenegro; milojica@jacimovic.me
² Department of Mathematics, University of Montenegro, Podgorica, Montenegro

Let $H$ and $F$ be real Hilbert spaces, $A : H → F$ - bounded linear operator from $H$ to $F$, $f ∈ F$ - a fixed element, and $U ⊆ H$ - a closed convex set. We will deal with the minimization problem

$$J(u) = \frac{1}{2} \|Au - f\|^2 \rightarrow \inf, \ u ∈ U.$$  (1)

which is is equivalent to the variational inequality

$$\text{find } u ∈ U : \langle A^*Au - A^*f, v - u \rangle ≥ 0, \ ∀v ∈ U.$$  (2)

In case of $U = H$, it is equivalent to the operator equation

$$A^*Au = A^*f.$$  (3)

Problems of the type (1) with infinite-dimensional spaces $H$ and $F$ are usually present in linear optimal control problems. We will shortly present one example of such problems (see [5]).

Given $T > 0$, find $u = u(·) ∈ U = \{u ∈ L_2[0, T] : \int_0^T u^2(t)dt ≤ r^2\}$, such that

$$J(u) = \int_0^T \|x(T, s; u) - f_0(s)\|^2 ds +$$

115
\[
\int_0^l |x'_t(T; s; u) - f_1(s)|^2 ds,
\]
be minimal. Here \(f_0(\cdot) \in H_0^1[0, 1]\), \(f_1(\cdot) \in L_2[0, l]\) are given functions and \(x = x(t, s; u)\) is the solution of the equation

\[
\frac{\partial^2 x(t, s)}{\partial t^2} = \frac{\partial^2 x(t, s)}{\partial s^2} + u(t),
\]

\[
x(t, 0) = x(t, l) = 0, 0 < t < T;
\]

\[
x(0, s) = 0, \frac{\partial x(t, 0)}{\partial t} = 0, 0 < s < l.
\]

Problem (1) in literature (see [4]) is regularly studied under assumption that instead of the exact operator \(A\) and instead of the element \(f\) one actually deals with their approximations \(A_\eta \in \mathcal{L}(H, F), f_\delta \in F\), such that

\[
\|A - A_\eta\| \leq \eta, \|f - f_\delta\| \leq \delta, \|A_\eta\|^2 \leq a,
\]

where \(\eta > 0, \delta > 0\) are small positive real numbers and \(a > 0\).

In general case, problems (1) and (3) are ill-posed. This fact generates the necessity to apply some methods of regularization [1], [3], [5], [4], that will produce good approximate solutions of the problems.

Usually the bounds of the accuracy of the regularization methods for solving ill-posed problems (1) and (3) were obtained for the problems that satisfy so-called source conditions or sourcewise representable conditions. In this paper, we will consider the power source condition (see [1],[2], [3] and [4])

\[
u_\infty = |A|^p h_*, \text{ where } h_* \in H, |A|^p = (A^*A)^{\frac{p}{2}}, p > 0,
\]

and projected power source conditions (see [3]),

\[
u_* = \pi_U(|A|^p h_*), h_* \in H, p > 0.
\]

that were used widely for obtaining the estimates of the convergence rate of regularization methods for solving linear operator equations. Here, \(u_\infty\) and \(u_*\) are the solutions of (3) and (1) with minimal norms (so-called normal solutions). The source condition seems quite natural for problem (3) if we have in mind that \(u_\infty \in \overline{R(A^*)}\), where \(R(A^*)\) is the range of the
operator $A$ and $\overline{R(A^*)}$ its closure in norm of the space $H$. It means that the solution $u_\infty$ is densely surrounded by the elements from $R(A^*)$. In a general case, the normal solution $u_*$ of (1) does not belong to $R(A^*)$, but $u_* \in \pi_U(R(A^*))$.

As a regularized approximate solution of problem (1), we will consider the solution $u = u_\alpha$ of the variational inequality

$$\langle g_\alpha(A_\eta^*A_\eta)u - A_\eta f_\delta, v - u \rangle \geq 0, \forall v \in U,$$

(7)

where $g_\alpha : [0, a] \mapsto R$, $(\alpha > 0)$, are Borel measurable functions satisfying the following conditions:

$$1 - tg_\alpha(t) \geq 0, t \in [0, a],$$

(8)

$$\frac{1}{t + \beta_\alpha} \leq g_\alpha(t) \leq \frac{1}{\beta_\alpha}, t \in [0, a], \beta > 0,$$

(9)

$$(\exists p_0 > 0) \forall p \in [0, p_0] \sup_{0 \leq t \leq a} t^p(1 - tg_\alpha(t)) \leq \gamma_\alpha p,$$

(10)

where $\gamma = \gamma(p_0)$. Here, $a$ is the constant from (4). Number $p_0$ is called qualification of the family $\{g_\alpha : \alpha > 0\}$.

Note that the functions $g_\alpha(t) = (t + \alpha)^{-1}$ and $g_\alpha(t) = \sum_{j=0}^{m-1} \frac{\alpha^j}{(t + \alpha)^p} = t^{-1}(1 - (1 + t)^{-m})$ (that defines Tikhonov methods of regularization and its iterated variant) satisfy these conditions. Similar methods of regularization of operator equations were used widely in literature (see, for example [1], [4]).

The next theorem contains the conditions of the convergence of the regularized solution $u_\alpha$ to normal solution $u_*$ of (1).

**Theorem.** Suppose conditions (4) and (8)-(10) are satisfied.

(a) If the parameter $\alpha$ in (7) is chosen such that $\alpha = \alpha(\eta, \delta) \to 0$ and $\eta^2 + \delta^2 \to 0$ as $\eta, \delta \to 0$, then $u_\alpha \to u_*$ as $\eta, \delta \to 0$.

(b) If one of the conditions (5) or (6), and $g_\alpha(t) = 1/(t + \alpha)$ is satisfied, and

$$\alpha = \alpha(\eta, \delta) = d(\eta + \delta)^2/(p + 2), d = \text{const},$$

then

$$\|u_\alpha - u_*\| = O(\eta + \delta)^p/(p + 2), 0 \leq p \leq 2p_0 - 1,$$

$$\|A_\eta(u_\alpha - u_*)\| = O(\eta + \delta).$$
Cubes Lattice’s properties investigation and possibilities of its application in Combinatorial Optimization

Ruben V. Khachaturov¹

¹ Dorodnicyn Computing Centre of RAS, Moscow, Russia; rv_khach@yahoo.ie

A new type of lattices — Lattice of Cubes (Cubes Lattice) [1, 2] is under the investigation. It is shown, that the number of all sub-cubes of the $M$-dimensional cube is equal to $3^M$, and the set of all such sub-cubes (with a corresponding choice of union and intersection operations) forms a lattice named a Lattice of Cubes [1, 2]. Algorithms of construction of such lattices are described, results of work of these algorithms for various dimensions of lattices are illustrated (fig. 1, 2). The total number of elements of such lattice (including the empty set) is equal to $3^M + 1$, where $M$ is dimension of the lattice. It is proved, that the Lattice of Cubes is a lattice with relative supplement, that allows to use effective algorithms [3, 4, 5] for solving on it the problems of minimization and maximization of supermodular functions. Concrete examples of such functions are given.
Optimization algorithms and possibilities of setting and solving a new class of problems on the Cubes Lattices are discussed.

**Lemma.** The number $|K_r|$ of all $r$-dimensional subcubes ($0 \leq r \leq m$) in the cube $C^m$ is equal to

$$|K_r| = 2^{m-r}C_r^m = 2^{m-r}\frac{m!}{r!(m-r)!}.$$

**Theorem 1.** The total number of subcubes of all dimensions in an $m$-dimensional cube $C^m$ is $3^m$.

The proofs of these Lemma and Theorem can be found in [1, 2, 5].

Let’s define the union and intersection operations for the Cubes Lattices.

**Definition 1.** If $C_1, C_2 \in K$ — elements of Cubes Lattice $K$ and $C_1 = (\omega_1^1; \omega_1^2), C_2 = (\omega_2^1; \omega_2^2), \omega_1^1 \subset \omega_1^2, \omega_2^1 \subset \omega_2^2$, then

$$C_1 \cup C_2 = C = (\omega_1^1 \cap \omega_1^2; \omega_1^1 \cup \omega_1^2),$$

$$C_1 \cap C_2 = \begin{cases} C(\omega_1^1 \cup \omega_1^2; \omega_2^1 \cap \omega_2^2), & \text{if } (\omega_1^1 \cup \omega_1^2) \subset (\omega_2^1 \cap \omega_2^2), \\ \emptyset, & \text{if } (\omega_1^1 \cup \omega_1^2) \not\subset (\omega_2^1 \cap \omega_2^2). \end{cases}$$

**Theorem 2.** The Cubes Lattice $K$ is a lattice with relative supplement.

**Proof.** To prove this theorem it is enough to show, that for any element of the lattice $C_1 \in K$ there will be at least one element $C_2 \in K$ such that $C_1 \cup C_2 = (0; I)$, $C_1 \cap C_2 = \emptyset$.

Indeed, let $C_1 = (\omega_1^1; \omega_1^2)$, where $\omega_1^1 \subset \omega_1^2$, and $C_2 = (\omega_2^1; \omega_2^2)$, where $\omega_2^1 \subset \omega_2^2$. Let’s show, that the element of the lattice $C_2$, at which $\omega_1^2 = 0$ and $\omega_2^2 = I \setminus \omega_1^2$, is relative supplement to the element $C_1$.

It can be proved by direct check, using the above-stated definitions of union and intersection operations for elements of the Cubes Lattice $K$, as follows:

$$C_1 \cup C_2 = (\omega_1^1; \omega_2^2) \cup (0; I \setminus \omega_2^2) = (0; I).$$

$$C_1 \cap C_2 = (\omega_1^1; \omega_1^2) \cap (0; I \setminus \omega_2^2) = (\omega_1^1; 0) = \emptyset.$$ 

The theorem is proved.

Note: the relative supplement to each element may not be unique.

Union operation allows to introduce a partial order between the elements $C \in K$. The following important definition is based on this operation:
Definition 2. \( C_1 = (\omega_1^1; \omega_1^2) \subset C_2 = (\omega_2^1; \omega_2^2) \) if \( C_1 \cup C_2 = C_2 \), i.e. when \( \omega_1^1 \supset \omega_2^1 \) and \( \omega_1^2 \subset \omega_2^2 \).

Two algorithms for constructing diagrams of a set of all \( C \in K \) (i.e. — Lattice of Cubes), based on this definition, have been worked out.

First algorithm. At a zero level of the Lattice of Cubes diagram there is always only one element — an empty subset (\( \emptyset \)). At the first level we shall arrange all \( 2^m \) sub-cubes containing only one element, i.e. cubes \( C^0 = (\omega_1; \omega_2) \), where \( \omega_1 = \omega_2 \). Then, using the union operation, we shall construct elements of the second level. At this level we shall arrange only cubes \( C^1 \), and so on. This process of constructing the Cubes Lattice is similar to the modified algorithm of successive calculations \([3, 4]\) with modifications corresponding to the union operation. In the process of constructing the next levels of the cubes diagram, we construct only a part of elements, as it is done in the algorithm of successive calculations, rather than construct simultaneously all elements of each level. This allows to effectively apply the corresponding rejection rules while solving the optimization problems \([3, 4, 5]\).

Note: The total number of elements of \( m \)-dimensional Cubes Lattice, including an empty set \( \emptyset \), is \( 3^m + 1 \).

![Diagram of the Lattice of Cubes for \( m = 3 \)](image-url)

Fig. 1. The Lattice of Cubes diagram for \( m = 3 \)
Figure 1 shows the Lattice of Cubes (Cubes Lattice) constructed using the first algorithm for $m = 3$.

**Second algorithm.** In the beginning the minimal set $\omega_1 = (0, \ldots, 0)$ is fixed and all sets $\omega_2 \supset \omega_1$ are looking through under the order of increase in binary system of calculation from $\omega_2 = \omega_1$ up to $\omega_2 = I$. During this we construct the corresponding lattice elements and the ribs connecting each constructed element with its neighborhood. Then we add 1 to $\omega_1$ in binary system, and look through all $\omega_2 \supset \omega_1$ again from $\omega_2 = \omega_1$ up to $\omega_2 = I$. This is repeated until $\omega_1$ becomes equal to $I$. Thus, we obtain all elements of the Cubes Lattice, passed through all $2^m$ of independent ways, that allows to solve quickly some kinds of problems of discrete optimization. In addition, this algorithm allows to arrange elements of the lattice on a plane in accordance with $2^m$ basis vectors.

Figure 2 shows the Lattice of Cubes (Cubes Lattice) constructed using the second algorithm for $m = 4$.

![Fig. 2. The Lattice of Cubes diagram for $m = 4$](image)

**REFERENCES**

General Theory of Optimization on Finite Lattices

Vladimir R. Khachaturov

1 Dorodnicyn Computing Centre of RAS, Moscow, Russia; rv_khach@yahoo.ie

The report consists of the Introduction and five main sections. In Section 2 the basic concepts, definitions and optimization problems setting are given. In Section 3 we describe the problems of minimization of supermodular functions on the different types of lattices: Boolean lattices, lattices with relative supplements (division lattices, lattices of vector subspaces of finite-dimensional vector space, geometrical lattices), lattices equal to Cartesian product of chains. The previously obtained theoretical results, on the basis of which the problems of minimization of supermodular functions on these lattices have been solved, are described. It’s noted, that these results have been extended to the distributive lattices.

Section 4 is devoted to the elaboration of the new basic propositions of the theory of maximization of supermodular functions on Boolean lattices (they were worked out only for the problems of minimization before) and establishing of the relation between the global minimum and maximum of supermodular functions for the main types of lattices. In Section 5 a new type of lattices — Lattice of Cubes, is defined and described [1, 2]. The
problems of minimization and maximization of supermodular functions are also considered on it.

In Section 6 we describe a general approach to the optimization on lattices with use of atomic lattices. We propose to map the atomic lattice into the corresponding Boolean lattice and then perform the optimization on this more ample Boolean lattice. If the properties of supermodularity of the function, defined on the atomic lattice, are obeyed on the Boolean lattice, then for the optimization it is possible to use all the theoretical results from Sections 3–5. Therefore, the elaboration of different effective algorithms of representation of high-dimensional Hyper-Cubes is of big significance. We describe the original combinatorial algorithms of automated representation of high-dimensional Hyper-Cubes (Booleans) on a plane in the form of different projections and diagrams, keeping the properties of Boolean as a partially ordered set of its vertexes. This gives us the ample opportunities for construction of various schemes of looking through the elements of atomic lattices and for visualization of the optimization process. The examples of various types of applied problems that have been solved using the elaborated optimization methods [3] are described. Although the majority of these problems are NP-hard, solution of a great amount of applied tasks demonstrated the high practical effectiveness of the elaborated methods and algorithms [3, 4]. (Analogy with the simplex-method can be seen.)

Thus, the obtained theoretical results and a great amount of optimization problems for lattices with concrete types of supermodular functions allow to consider the methods and algorithms for solving the problems of optimization of supermodular functions on lattices as a new field of mathematical programming — supermodular programming.

REFERENCES


---

**Hyperplane Covering Problems. Complexity and Approximation Issues**

Michael Khachay$^1$, Maria Poberii$^2$

$^1$ *Institute of Mathematics and Mechanics, UB of RAS, Ekaterinburg, Russia; mkhachay@imm.uran.ru*

$^2$ *Institute of Mathematics and Mechanics, UB of RAS, Ekaterinburg, Russia; maschas_briefen@mail.ru*

Settings of geometric covering problem and related problems are usual in various operations research domains [1-3]: optimal facility location theory, cluster analysis, pattern recognition, etc. Mathematically, family of such problems can be participated into two classes.

The first one contains special cases of well-known abstract Set Cover problem. The main general feature of all these problems is the finiteness of the initial family of subsets, for which it is required to find a subfamily (or just prove its existence) covering some target set and satisfying given optimality conditions. There are many papers studying problems from this class (see survey at [4]). The classical papers [5-7] seem to be the most important among them. First two papers contain intractability proof of Set Cover problem and two main design patterns for constructing approximation algorithms for this problem. The last paper proves the optimality of these patterns, unless $P = NP$.

The second class consists of problems without the mentioned above finiteness constraint. Usually, the initial family of subsets is given here implicitly in terms of some geometric property characterizing its elements. For instance, for a given set it is required to find a minimal cardinality cover by straight lines, circles of a given radii, etc.
In the paper, a series of hyperplane covering problems for given finite sets in finite-dimensional vector spaces of fixed dimension \( d > 1 \) is considered. The first element of this family (for \( d = 2 \)), also known as Point Covering on the plane (2PC) problem was studied by N.Megiddo and A.Tamir [8] who proved its intractability in the strong sense.

We extend this result on to the case of appropriate fixed dimensionality \( d > 1 \) and prove that all these problems are Max-SNP-hard and consequently have no PTAS, unless \( P = NP \).

**Problem 1.** ‘Point covering by lines on the plane’ (2PC). A finite subset \( P = \{p_1, \ldots, p_n\} \subset \mathbb{Z}^2 \) and natural number \( B \) are given. Is there exists a cover \( C \) of \( P \) by straight lines such that \(|C| \leq B|\)?

Obviously, in the particular case when the set \( P \) is in the general position, i.e. each triple of its points does not belong to the same line, the 2PC problem has a trivial solution (‘Yes’ whether \( B \geq \lceil |P|/2 \rceil \) and ‘No’ otherwise), which can be found in a polynomial time. But in the general case this problem is intractable.

**Theorem 1 [8].** The 2PC problem is NP-complete in the strong sense.

Let us consider the more general problem settings.

**Problem 2.** ‘Hyperplane covering in \( d \)-dimensional space’ (dPC). For a fixed \( d > 1 \), a finite subset \( P = \{p_1, \ldots, p_n\} \subset \mathbb{Z}^d \) and natural number \( B \) are given. Is there exists a cover \( C \) of \( P \) by hyperplanes such that \(|C| \leq B|\)?

**Problem 3.** ‘Minimal hyperplane covering in \( d \)-dimensional space’ (Min-dPC). Let a finite subset \( P = \{p_1, \ldots, p_n\} \subset \mathbb{Z}^d \) be given. It is required to find a minimum cardinality partition \( J_1, \ldots, J_L \) of a set \( \mathbb{N}_n = \{1, \ldots, n\} \) such that for each \( i \in \mathbb{N}_L \) there is a hyperplane \( H_i \) and

\[
\{p_j \in P : j \in J_i\} \subset H_i.
\]

We start with construction of polynomial-time reduction of \((d - 1)\text{PC}\) to dPC problem. Let an instance of \((d - 1)\text{PC}\) be given by subset \( P = \{p_1, \ldots, p_n\} \subset \mathbb{N}_M^{d-1} \) and \( B \in \mathbb{N} \). We use a natural isomorphic embedding of \((d - 1)\)-dimensional into \( d \)-dimensional vector space:

\[
x \in \mathbb{R}^{d-1} \mapsto [x, 0] \in \mathbb{R}^d.
\]
Map any point $p_i \in P$ into couple of points in $\mathbb{Z}^d$ by the formula

$$\bar{p}_{2i-1} = [p_i, -w_i], \bar{p}_{2i} = [p_i, w_i],$$

where

$$w_i = (K + 2)^{i-1} \text{ and } K = \left\lceil (d-1)\frac{d-1}{2} (M-1)^{d-1} \right\rceil.$$ 

Such a way, we construct the subset $\bar{P} \subset \mathbb{Z}^d$ and the setting $(\bar{P}, B)$ of the $d$PC problem.

It is evident, that any hyperplane cover of $P$ induces the equivalent cover (with the same number of hyperplanes) of $\bar{P}$ in $\mathbb{R}^d$. The converse statement should be proved.

Denote by $\pi_0$ the hyperplane $\{[x, 0] : x \in \mathbb{R}^{d-1}\}$. Let $\text{Pr}_{\pi_0} Q$ be an orthogonal projection of the subset $Q \subset \mathbb{R}^d$ onto $\pi_0$.

**Lemma 1.** Let subsets $Q \subset P$ and $\bar{Q} \subset \bar{P}$ be related by $Q = \text{Pr}_{\pi_0} \bar{Q}$ and the following inequalities be valid

$$|\bar{Q}| \geq d + 1,$$

$$\dim \text{aff} \bar{Q} \leq d - 1.$$

Then $\dim \text{aff} Q \leq d - 2$.

**Lemma 2.** Let $\bar{\Pi} = \{\bar{\pi}_1, \ldots, \bar{\pi}_t\}$ be a hyperplane cover of subset $\bar{P}$. The subset $P$ also has a hyperplane cover $\Pi$ such that $|\Pi| \leq t$.

**Lemma 3.** The described above reduction $(d-1)\text{PC}$ to $d\text{PC}$ can be done in polynomial time of $\text{Length}((d-1)\text{PC})$.

On the basis of these lemmas we can prove the following

**Theorem 2.** For an arbitrary fixed $d > 1$, the $d\text{PC}$ problem is NP-complete (and the Min-$d\text{PC}$ problem is NP-hard) in the strong sense.

Now we show that the supposed above $(d-1)\text{PC}$ to $d\text{PC}$ reduction can be reformulated as $L$-reduction [9] from Min-$(d-1)\text{PC}$ to Min-$d\text{PC}$ problem.

**Definition 1.** Let sets $\mathcal{I}$ and $S$, set-valued map $F : \mathcal{I} \to 2^S$ and some target function $c : \bigcup_{I \in \mathcal{I}} F(I) \to \mathbb{R}^+$ be given. The quadruple $A = (\mathcal{I}, S, F, c)$, where each $I \in \mathcal{I}$ is mapped to optimization problem

$$\min \{c(s) : s \in F(I)\},$$

126
is called a combinatorial minimization problem.

W.o.l.g., any \( I \in \mathcal{I} \) is called an instance of the problem \( A \) and its optimum value is denoted by \( OPT(I) \).

**Definition 2.** Consider problems \( A \) and \( B \) of combinatorial minimization. It is called, that there is an \( L \)-reduction from \( A \) into \( B \), if there are two \( \text{LSPACE} \)-computable functions \( R \) and \( S \) and positive constants \( \alpha \) and \( \beta \) such that the following conditions are valid:

1. for each instance \( I \) of the problem \( A \), \( R(I) \) is an instance of \( B \) and
   \[ OPT(R(I)) \leq \alpha OPT(I); \]
2. for each feasible solution \( z \) of \( R(I) \), \( S(z) \) is a feasible solution of \( I \) such that
   \[ c_A(S(z)) - OPT(I) \leq \beta(c_B(z) - OPT(R(I))), \]
   where \( c_A, c_B \) are target functions of \( A \) and \( B \) correspondingly.

**Theorem 3.** For each fixed \( d > 2 \), there is an \( L \)-reduction of Min-\((d - 1)PC\) to Min-\(dPC\) problem.

Taking into account the following known result

**Theorem 4 [11].** Min-2PC problem is Max-SNP-hard.

one can formulate the last

**Theorem 5.** For each fixed \( d > 1 \), the Min-\(dPC\) problem is Max-SNP-hard.

Consequently, Min-\(dPC\) problem has no polynomial-time approximation schema (PTAS) for each fixed \( d > 1 \), unless \( P = NP \).

The authors were supported by the Russian Foundation for Basic Research (projects no. 10-01-00273 and 10-07-00134) and Ural Branch of RAS (projects no. 12-P-1-1016 and 12-C-1-1017/1).

**References**

Leontief’s model as a boundary value problem in optimal control

Elena Khoroshilova

1 Lomonosov Moscow State University, Moscow, Russia; khorelena@gmail.com

1. Statement of problem. The optimal control problem with free right end on a fixed interval is considered in this paper. The dynamics of the process is described by a system of ordinary differential equations

\[ \frac{d}{dt} x(t) = D(t)x(t) + B(t)y(t), \quad t_0 \leq t \leq t_1, \quad x(t_0) = x_0 = 0, \quad (1) \]
with the trajectories
\[ x(\cdot) \in PC^1([t_0, t_1], \mathbb{R}^n), \]
and the controls \( y(\cdot) \in Y. \)

\[ Y = \{ y(\cdot) \in PC([t_0, t_1], \mathbb{R}^n) \mid y_i(t) \in [y_i^-, y_i^+], i = 1, \ldots, n \}, \tag{2} \]

\[ D(\cdot), B(\cdot) \in C([t_0, t_1], \mathbb{R}^{n \times n}). \]

Let the right ends \( x_1 = x(t_1), y_1 = y(t_1) \) of the trajectories and controls satisfy the constraints

\[ x_1 = A_1 x_1 + y_1, \quad y_1 \geq 0, \tag{3} \]

\[ A_1 \in \mathbb{R}^{n \times n}. \]

Complementing the system (1) with a control \( y(\cdot) \in Y \) and solving it, we find the trajectory \( x(\cdot) \). When control changes within the set \( Y \) the right-hand ends \( x_1 \) of trajectories describe the attainable set, on which the following objective function is defined

\[ \varphi(x_1, y_1) = \varphi_1(x_1) + \varphi_2(y_1), \tag{4} \]

where \( \varphi_1(x_1) \) and \( \varphi_2(y_1) \) are convex and differentiable in the variables \( x_1 \) and \( y_1 \), respectively.

We need to determine the optimal control \( y^*(\cdot) \in Y \) and the corresponding trajectory \( x^*(\cdot) \in PC^1([t_0, t_1], \mathbb{R}^n) \), subject to the system (1).

At the same time their right ends \( y_1^* \) and \( x_1^* \) have to minimize the objective function (4) under constraints (2)–(3).

It is assumed that a solution exists, but not unique. Similar formulations of the problem, but without the terminal control, were examined in [1],[2]. In case \( \varphi(x_1, y_1) \equiv \varphi(y_1) \), the optimization problem is similar to the input-output economic model [3].

2. Lagrangian, dual problem and boundary-value problem.

Consider the Lagrangian for the optimization problem

\[ \mathcal{L}(p_1, \psi(\cdot); x(\cdot), y(\cdot)) = \varphi_1(x_1) + \varphi_2(y_1) + \]

\[ + \langle p_1, (I - A_1)x_1 - y_1 \rangle + \int_{t_0}^{t_1} \langle \psi(t), D(t)x(t) + B(t)y(t) - \frac{d}{dt} x(t) \rangle dt, \]

defined for all \( x(\cdot) \in PC^1([t_0, t_1], \mathbb{R}^n), x_1 \in \mathbb{R}^n, y(\cdot) \in Y, y_1 \geq 0, p_1 \in \mathbb{R}^n, \psi(\cdot) \in PC^1([t_0, t_1], \mathbb{R}^n)' \) — the dual space.

\[ ^1 \text{Here } PC^1([t_0, t_1], \mathbb{R}^n) \text{ is the class of continuous vector-valued functions: } [t_0, t_1] \rightarrow \mathbb{R}^n \text{ with piecewise continuous derivatives; } PC([t_0, t_1], \mathbb{R}^n) \text{ is the class of piecewise continuous vector-valued functions: } [t_0, t_1] \rightarrow \mathbb{R}^n. \]
The point \((p^*_1, \psi^*(\cdot); x^*(\cdot), y^*(\cdot))\) is called a saddle point of Lagrange function if for all \(x(\cdot) \in PC^1([t_0, t_1], \mathbb{R}^n), x_1 \in \mathbb{R}^n, y(\cdot) \in Y, y_1 \geq 0, p_1 \in \mathbb{R}^n, \psi(\cdot) \in PC^1([t_0, t_1], \mathbb{R}^n)'\) the following system holds

\[
\mathcal{L}(p_1, \psi(\cdot); x^*(\cdot), y^*(\cdot)) \leq \mathcal{L}(p^*_1, \psi^*(\cdot); x^*(\cdot), y^*(\cdot)) \leq \mathcal{L}(p^*_1, \psi^*(\cdot); x(\cdot), y(\cdot)).
\]

(5)

Here \((x^*(\cdot), y^*(\cdot))\) and \((p^*_1, \psi^*(\cdot))\) are called direct and dual variables.

According to the Kuhn-Tucker theorem, there exist \(p^*_1\) and \(\psi^*(\cdot)\), such that if the pair \((x^*(\cdot), y^*(\cdot))\) is a solution, then \((p^*_1, \psi^*(\cdot); x^*(\cdot), y^*(\cdot))\) is a saddle point of the Lagrangian. Converse is also true: if the pair \((x^*(\cdot), y^*(\cdot))\) satisfies the saddle point system (5), then it is a solution to the original problem.

Transforming the Lagrangian, as well as the saddle point system to conjugate form, and using the right-hand inequality of the system, we obtain the dual problem. Combining the direct and dual problems, we come to a boundary value problem

\[
\frac{d}{dt} x^*(t) = D(t)x^*(t) + B(t)y^*(t), \quad x^*(t_0) = x_0, \quad y^*(\cdot) \in Y, \quad t_0 \leq t \leq t_1,
\]

\[
x^*_1 = A_1 x^*_1 + y^*_1, \quad y^*_1 \geq 0,
\]

\[
\frac{d}{dt} \psi^*(t) + D^T(t)\psi^*(t) = 0, \quad \psi^*_1 = \nabla \varphi_1(x^*_1) + (I - A_1^T)p^*_1,
\]

\[
\varphi_2(y_1) - \varphi_2(y^*_1) - \langle y_1 - y^*_1, p^*_1 \rangle + \int_{t_0}^{t_1} \langle B^T(t)\psi^*(t), y(t) - y^*(t) \rangle dt \geq 0.
\]

3. Method of solution. The method of simple iteration is the simplest of the known numerical methods. However, in this case we are dealing with a saddle problem, for which the simple iteration method, generally speaking, not converge. Therefore, to solve the problem, we use an extra-proximal approach [4]–[5]:

1) prediction half-step

\[
\frac{d}{dt} x^k(t) = D(t)x^k(t) + B(t)y^k(t), \quad x^k_0 = x_0,
\]

\[
p^*_1^k = p^*_1^k + \alpha((I - A_1)x^k_1 - y^k_1),
\]

\[
\frac{d}{dt} \psi^k(t) + D^T(t)\psi^k(t) = 0, \quad \psi^k_1 = \nabla \varphi_1(x^k_1) + (I - A_1^T)p^*_1^k,
\]

130
\[
(\bar{y}_1^k, \bar{y}^k(\cdot)) = \arg\min \left\{ \frac{1}{2} |y_1^k - y_1^{k-1}|^2 + \alpha \langle \nabla \varphi_2(y_1), p_1^k, y_1 - y_1^k \rangle + \frac{1}{2} \int_{t_0}^{t_1} |y(t) - y^k(t)|^2 dt + \alpha \int_{t_0}^{t_1} \langle B^T(t) \psi^k(t), y(t) - y^k(t) \rangle dt \mid y_1 \geq 0, y(\cdot) \in Y \right\};
\]

2) basic half-step

\[
\frac{d}{dt} \bar{x}^k(t) = D(t) \bar{x}^k(t) + B(t) \bar{y}^k(t), \quad \bar{x}^k(t_0) = x_0,
\]

\[
p_1^{k+1} = p_1^k + \alpha((I - A_1) \bar{x}_1^k - \bar{y}_1^k),
\]

\[
\frac{d}{dt} \bar{\psi}^k(t) + D^T(t) \bar{\psi}^k(t) = 0, \quad \bar{\psi}^k_1 = \nabla \varphi_1(\bar{x}_1^k) + (I - A_1^T) \bar{p}_1^k,
\]

\[
(y_1^{k+1}, y_1^{k+1}(\cdot)) = \arg\min \left\{ \frac{1}{2} |y_1 - y_1^{k-1}|^2 + \alpha \langle \nabla \varphi_2(y_1), \bar{p}_1^k, y_1 - y_1^k \rangle + \frac{1}{2} \int_{t_0}^{t_1} |y(t) - y^k(t)|^2 dt + \alpha \int_{t_0}^{t_1} \langle B^T(t) \bar{\psi}^k(t), y(t) - y^k(t) \rangle dt \mid y_1 \geq 0, y(\cdot) \in Y \right\}.
\]

In the first half-step of each iteration we find the "forward" vectors \( \bar{p}_1^k, \bar{y}_1^k, \bar{y}^k(\cdot) \) (those that should be at the next step in the simple iteration method). Then in the second half-step we use them to find the direction of future movement on the \((k + 1)\)-th iteration. From the current position in the \(k\)-th iteration, we take a step in that direction, and calculate \( y_1^{k+1}, y_1^{k+1}(\cdot) \). Thus, when we choose the direction of movement, then do not take into account the current gradient, but the gradient of the next step.

The following theorem on the convergence of the method was proved.

**Theorem 1.** Let the set of optimal trajectories \( x^*(\cdot) \) of the problem (1)–(4) be not empty and belong to the subspace \( PC^1([t_0, t_1], \mathbb{R}^n) \). If \( \varphi_1(x_1) \) and \( \varphi_2(y_1) \) are convex and differentiable functions with gradients satisfying a Lipschitz condition with constants \( L_1, L_2, D(t), B(t) \) are continuous matrices, \( Y \) is a set of the form (2). Then the sequence

\[
\{ |y_1^k - y_1^*|_{\mathbb{R}^n}^2 + |p_1^k - p_1^*|_{\mathbb{R}^n}^2 + \|y^k(\cdot) - y^*(\cdot)\|_{L_2}^2 \},
\]

generated by method (6) with the choice of parameter \( \alpha \) from the condition \( 0 < \alpha < \alpha_0 \), where \( \alpha_0 = \min \left( \frac{1}{\sqrt{2}L_2}, \frac{1}{\sqrt{2}L_2^2 + B_{max}^2 C_1}, \frac{1}{B_{max} \sqrt{C_1}} \right) \),

131
$B_{\text{max}}, C_1, C_2$ – some calculated constants, decreases monotonically. Moreover, every weakly converging subsequence of controls $\{y^{k_i}(\cdot)\}$ converges weakly to the optimal control $y^*(\cdot)$, and the corresponding subsequence of trajectories $\{x^{k_i}(\cdot)\}$ converges to the optimal trajectory $x^*(\cdot)$ in the uniform norm $C^n[t_0, t_1]$.

If the sequence of controls $\{y^k(\cdot)\}$ has at least one strong limit as $k \to +\infty$ then the process $\{x^k_1, y^k_1, p^k_1; x^{k}(\cdot), y^{k}(\cdot), \psi^{k}(\cdot)\}$ converges to a solution of the problem $x^*_1, y^*_1, p^*_1; x^*(\cdot), y^*(\cdot), \psi^*(\cdot)$, with respect to the variables $y(\cdot), y^1, p^1$ – monotonically.

4. Conclusions. The optimal control problem with free right end and linear differential equations constraints is considered. The right-hand ends of controls and trajectories generate a finite-dimensional Cartesian product, on which a minimum of the objective function is defined under constraints such as input-output model of Leontief.

To solve this problem we suggest an iterative method of extra-proximal type, consisting in the construction of the functional sequences of trajectories and controls. We have proved that the sequences of controls, trajectories, conjugate trajectories, as well as the sequences in finite-dimensional spaces of primal and dual variables converge monotonically in norm to solution of the problem.

The author was supported by the Russian Foundation for Basic Research (project no. 12-01-00783) and by the Program "Leading Scientific Schools" (NSh-5264.2012.1).

References

Approximation to minimum committee problem for system of linear inequalities in $\mathbb{R}^3$

Konstantin Kobylkin

1 Institute of Mathematics and Mechanics, Ural Branch of RAS, Ekaterinburg, Russia; kobylkinks@gmail.com

An NP-hard minimum committee problem for infeasible three-dimensional system of homogeneous linear inequalities (MC(3)) is considered. Well known minimum covering problem for finite set on the plane using straight lines is reducible to MC(3) [1]. A useful approximation to MC(3) is studied in the form of minimum committee problem for infeasible system

$$(c_j, h) > b_j, j = 1, \ldots, m, c_j, h \in \mathbb{R}^2, b_j \in \mathbb{R}, \text{rank}\{c_i, c_j\} = 2, i \neq j.$$ (1)

Definition 1. [2] A finite subset $K$ of $\mathbb{R}^2$ is called a committee of a system (1) iff for each $j \in N_m = \{1, \ldots, m\}$ there exists a subset $K(j) \subset K$ with $|K(j)| > |K|/2$ whose elements satisfy $j$th inequality of (1). A committee of (1) having the least cardinality is called minimum committee.

Definition 2. Let $i, j \in N_m, i \neq j$. We call $i$th inequality of (1) inessential with respect to its $j$th inequality (relative to system (1)) iff $i$th inequality belongs to each maximum feasible subsystem of (1) containing $j$th inequality.

Denote by $\mathcal{J}(j, \Delta b_j)$ the system obtained from (1) by excluding inequality $(c_j, h) > b_j$ and adding inequality $(c_j, h) > b_j + \Delta b_j$ instead (with the same number $j$), where $\Delta b_j \in \mathbb{R}$. Let $\rho_1(i, j)$ be exact lower bound of $\Delta b_j$ such that $i$th inequality is inessential with respect to $j$th inequality relative to the system $\mathcal{J}(j, \Delta b_j)$. Also let $\rho_2(i, j)$ be exact lower bound of $\Delta b_j$ such that there exists infeasible subsystem of three inequalities of $\mathcal{J}(j, \Delta b_j)$ which contains inequalities $(c_i, h) > b_i$ and $(c_j, h) > b_j + \Delta b_j$. 

133
**Theorem.** If \( \min\{\rho_1(i,j), \rho_2(i,j)\} \leq 0 \) for every distinct \( i, j \in N_m \), then minimum committee cardinality for system (1) is the same as that for homogeneous system

\[
(c_j, h) > 0, \ j \in N_m.
\]  

MC problem for system (2) is known to be \( O(m) \) hard [2]. As a consequence MC problem for (1) is polynomially solvable under theorem conditions which can also be checked in polynomial time.

The authors were supported by Presidium of Ural Branch of RAS (projects no. 12-P-1-1016 and 12-S-1-1017/1), and by Russian Foundation for Basic Research (projects no. 10-01-00273 and 10-07-00134).

**REFERENCES**


**Metric for the total tardiness minimization problem**

Pavel Korenev\(^1,2\), Alexander Lazarev\(^1,2,3\)

1 Institute of Control Sciences of the Russian Academy of Sciences, Russia;

2 Moscow State University, Moscow, Russia;

3 National Research University Higher School of Economics, Moscow, Russia;

jobmath@mail.ru, pkorenev@rambler.ru

**Introduction**

Suppose that we have a set \( N = \{1, 2, \ldots, n\} \) of \( n \) jobs to be processed on a single machine. Preemptions are not allowed. The machine is available
since time $t_0 = 0$ and can handle only one job at a time. Job $j \in N$ is available for processing since its release date $r_j \geq 0$, its processing requires processing time $p_j \geq 0$ time units and should ideally be completed before its due date $d_j$. We will call an instance the set of given parameters: release dates, processing times, and due dates. We will use superscripts to distinguish parameters belonging to different instances. Note that an instance $A = \{r^A_1, \ldots, r^A_n, p^A_1, \ldots, p^A_n, d^A_1, \ldots, d^A_n\}$ can be considered as a vector in $3n$-dimensional space.

Let $S_j(\pi)$ and $C_j(\pi)$ be the starting and the completion time of job $j \in N$ in schedule $\pi$, respectively. We will consider only early schedules, i.e., if $\pi = \{j_1, \ldots, j_n\}$, then $S_{j_1} = \max\{0, r_{j_1}\}$, $S_{j_k} = \max\{r_{j_k}, C_{j_{k-1}}\}$, $k = 2, 3, \ldots, n$, and $C_j(\pi) = S_j(\pi) + p_j, j \in N$. Thus an early schedule is uniquely determined by a permutation of the jobs of set $N$. Then let $T_j(\pi) = \max\{0, C_j(\pi) - d_j\}$ be a tardiness of job $j$ in schedule $\pi$.

The objective is to find an optimal schedule $\bar{\pi}$ which minimizes the total tardiness, i.e., objective function is $F(\pi) = \sum_{j \in N} T_j(\pi)$. The problem is denoted by $1|r_j|\sum T_j.$

In the paper we propose a new approach for the total tardiness minimization problem. The approach is to construct a polynomially solvable instance $B$ and apply its solution to the given instance $A$. To evaluate the error of the solution we construct a metric for the considered problem. For the problem $1|r_j|\sum T_j$ we propose a metric $\rho(A, B)$.

$$\rho(A, B) = n \cdot \max_{j \in N} |r^A_j - r^B_j| + n \cdot \sum_{j \in N} |p^A_j - p^B_j| + \sum_{j \in N} |d^A_j - d^B_j|.$$  

This function can be considered as a metric for the problem and bounds difference between optimal values of objective functions of instances $A$ and $B$.

**Metrical approach**

**Theorem 1.** The function

$$\rho(A, B) = n \cdot \max_{j \in N} |r^A_j - r^B_j| + n \cdot \sum_{j \in N} |p^A_j - p^B_j| + \sum_{j \in N} |d^A_j - d^B_j|.$$  

satisfies the metric axioms.
Theorem 2. Let $\bar{\pi}^A$ and $\bar{\pi}^B$ be an optimal schedules for instances $A$ and $B$, respectively. Moreover, let $\tilde{\pi}^B$ be an approximate schedule, subject to

$$\sum_{j \in N} T_j^B(\tilde{\pi}^B) - \sum_{j \in N} T_j^B(\bar{\pi}^B) \leq \delta.$$  

Then

$$\sum_{j \in N} T_j^A(\tilde{\pi}^B) - \sum_{j \in N} T_j^A(\bar{\pi}^A) \leq 2\rho(A, B) + \delta.$$  

The idea of the metrical approach is to find the least distanced in the metric from the given instance $A$ polynomially solvable instance $B$. Then, by applying known polynomial algorithm to the instance $B$, one obtains a schedule $\pi^B$ which can be used as an approximate solution for instance $A$ with error no greater than $2\rho(A, B)$. One can also use approximate solution for the instance $B$ with an absolute error $\delta$ as an approximate solution for instance $A$, in this case the error is not greater that $2\rho(A, B) + \delta$.

Thereby, the problem 1$|r_j|\sum T_j$ is reduced to the function $\rho(A, B)$ minimization problem.

Let us search for the instance $B$ in the polynomially solvable class defined by the system of linear inequalities

$$\mathbf{A} \cdot \mathbf{R}^B + \mathbf{B} \cdot \mathbf{P}^B + \mathbf{C} \cdot \mathbf{D}^B \leq \mathbf{H},$$

where $\mathbf{R}^B = (r_1^B, \ldots, r_n^B)^T$, $\mathbf{P}^B = (p_1^B, \ldots, p_n^B)^T$, $\mathbf{D}^B = (d_1^B, \ldots, d_n^B)^T$, $p_j^B \geq 0, r_j^B \geq 0$, $j \in N$, $T$ is transposition symbol, $\mathbf{A}, \mathbf{B}, \mathbf{C} - m \times n$ matrices, and $\mathbf{H}$ – a column of $m$ elements.

Then the problem of finding the least distanced from $A$ instance of the given polynomially solvable class can be formulated as follows

$$\text{minimize } f = n \cdot (y^r - x^r) + n \cdot \sum_{j \in N} (y_j^p - x_j^p) + \sum_{j \in N} (y_j^d - x_j^d),$$
subject to

\[ x^r \leq r^A_j - r^B_j \leq y^r, \]
\[ x^p_j \leq p^A_j - p^B_j \leq y^p, \]
\[ x^d_j \leq d^A_j - d^B_j \leq y^d, \]
\[ r^B_j \geq 0, p^B_j \geq 0, j \in N, \]
\[ A \cdot R^B + B \cdot P^B + C \cdot D^B \leq H. \]

It is the problem of the linear programming, with \(7n + 2\) variables: \(r^B_j, p^B_j, d^B_j, x^p_j, y^p, x^d_j, y^d, x^r, y^r, j = 1, \ldots, n.\)

However, it is not necessary to use algorithms of the linear programming, if there are less complicated ways.

The metrical approach can be applied to other scheduling problems, if a metric function with required properties is constructed.

**Lemma 1.** Consider the scheduling problem with following objective function

\[ F(\pi) = \sum_{j \in N} \phi_j(\pi, r_1, \ldots, r_n, p_1, \ldots, p_n, d_j). \]

Then the function

\[ \rho(A, B) = \sum_{j \in N} \sum_{i \in N} (R_{ji}|r^A_j - r^B_j| + P_{ji}|p^A_j - p^B_j|) + \sum_{j \in N} D_j|d^A_j - d^B_j|, \]

where \(R_{ji} \geq |\frac{\partial \phi_i}{\partial r_j}|, P_{ji} \geq |\frac{\partial \phi_i}{\partial p_j}|, D_j \geq |\frac{\partial \phi_i}{\partial d_j}|\), can be used as a metric for the problem, and the metrical approach can be applied to find an approximate solution of the problem.

**Computational experiments**

We used three polynomially solvable classes in computational experiments. These classes are \(\{PR : p_j = p, r_j = r, j \in N\}, \{PD : p_j = p, d_j = d, j \in N\}, \{RD : r_j = r, d_j = d, j \in N\}\). In the optimal schedules for these classes jobs are processed in the increasing order the free parameter.

**Lemma 2.** Minimum of the metric function \(\rho(A, B)\), where \(B \in \{PR, PD, RD\}\) can be found in \(O(n)\) operations.
Table 5: Average experimental error in percentage of the theoretical error

<table>
<thead>
<tr>
<th>n</th>
<th>PR</th>
<th>PD</th>
<th>RD</th>
</tr>
</thead>
<tbody>
<tr>
<td>4</td>
<td>2,5%</td>
<td>4,6%</td>
<td>20,8%</td>
</tr>
<tr>
<td>5</td>
<td>2,6%</td>
<td>4,8%</td>
<td>23,1%</td>
</tr>
<tr>
<td>6</td>
<td>2,6%</td>
<td>4,6%</td>
<td>24,6%</td>
</tr>
<tr>
<td>7</td>
<td>2,6%</td>
<td>4,7%</td>
<td>26%</td>
</tr>
<tr>
<td>8</td>
<td>2,5%</td>
<td>4,6%</td>
<td>27%</td>
</tr>
<tr>
<td>9</td>
<td>2,4%</td>
<td>4,7%</td>
<td>27,9%</td>
</tr>
<tr>
<td>10</td>
<td>2,4%</td>
<td>4,6%</td>
<td>28,6%</td>
</tr>
</tbody>
</table>

To evaluate approximate solutions for both cases we have run computational experiments. For each value of \( n \) and each of used polynomially solvable classes 10000 instances were generated. Experiments were performed for \( n = 4, 5, \ldots, 10 \). For each instance, processing times \( p_j \) were generated randomly in the interval \([1, 100]\), due dates \( d_j \) were generated in the interval \([p_j, \sum_{j \in N} p_j]\), and release dates \( r_j \) were generated in the interval \([0, d_j - p_j]\). We used proposed approach to find an approximate solution with value of objective function \( F_a \) for each instance, and branch & bound algorithm to find an optimal solution with value of objective function \( F_o \). After we estimated experimental error \( \delta = F_a - F_o \) in percentage of the theoretical error, which is doubled value of function \( \rho(A, B) \).

All obtained distributions are bell-shaped. Obtained average errors are shown in Table 1. In the PR-case experimental errors averages near 2,5% of the theoretical, in PD-case average error is near 4,5% and in RD-case error grows from 20% to 30% with increasing of \( n \).

**Conclusion**

In the paper we have proposed the new approach to the total tardiness minimization problem. The approach is based on search for the polynomially solvable instance which has a minimal distance in the metric from the original instance. In further research we are going to improve the approach by constructing new metrics and finding new polynomially solvable cases of scheduling problems.
The authors were supported by the Russian Foundation for Basic Research (projects no. 11-08-01321 and 11-08-13121).

References


New CQ-free optimality criterion for convex SIP problems with polyhedral index sets

Olga Kostyukova\textsuperscript{1}, Tatiana Tchemisova\textsuperscript{2}

\textsuperscript{1} Institute of Mathematics, National Academy of Sciences of Belarus, Minsk, Belarus; kostyukova@im.bas-net.by

\textsuperscript{2} University of Aveiro, Aveiro, Portugal; tatiana@ua.pt

Consider a convex Semi-Infinite Programming (SIP) problem in the form

\begin{align*}
(P) & : \quad \min_{x \in \mathbb{R}^n} c(x), \\
\text{s.t.} & \quad f(x, t) \leq 0 \quad \forall t \in T = \{t \in \mathbb{R}^s : h_k^T t \leq \Delta h_k, \ k \in K\},
\end{align*}

\[139\]
where the objective function $c(x), x \in \mathbb{R}^n$, is convex, the constraint function $f(x, t), x \in \mathbb{R}^n, t \in T$, is linear w.r.t. $x$; $h_k \in \mathbb{R}^s, \Delta h_k \in \mathbb{T}, k \in K, |K| < \infty$. Notice that the index set $T$ in $(P)$ is a convex polyhedron.

Let $X$ be the feasible set of problem $(P)$: $X = \{x \in \mathbb{R}^n : f(x, t) \leq 0, \forall t \in T\}$. Suppose that $f(x, t)$ is sufficiently smooth w.r.t. $t$ for all $x \in X$ and $t \in T$.

Given $t \in T$, denote by $K_a(t) \subset K$ the set of active indices in $t$, $K_a(t) := \{k \in K : h_k^T t = \Delta h_k\}$, and by $L(t)$ the set of feasible directions in $T$ starting from $t$,

$$L(t) := \{l \in \mathbb{R}^s : h_k^T l \leq 0, k \in K_a(t)\}. \tag{3}$$

Given $x \in X$, the set of active indices in $x$ is $T_a(x) := \{t \in T : f(x, t) = 0\}$.

**Definition 1.** Let us say that an index $\bar{t} \in T$ is immobile in problem $(P)$, if $f(x, \bar{t}) = 0$ for all $x \in X$.

Denote by $T^*$ the set of all immobile indices in problem $(P)$. It is evident that $T^* \subset T_a(x)$ for all $x \in X$.

**Definition 2.** The constraints of problem $(P)$ satisfy the Slater condition if there exists $\bar{x} \in X$ such that $f(\bar{x}, t) < 0, \forall t \in T$.

In [4], it is proved that a convex SIP problem with $X \neq \emptyset$ satisfies the Slater condition if and only if the set of immobile indices is empty. Thus the emptiness of the set $T^*$ can be considered as a constraint qualification (CQ) equivalent to the Slater-type condition for SIP.

**Definition 3.** An immobile index $\bar{t} \in T^*$ has the order of immobility $q(\bar{t}, \bar{l})$ along a nontrivial feasible direction $\bar{l} \in L(\bar{t})$ if

1. $\frac{d^i f(x, \bar{t} + \alpha \bar{l})}{d\alpha^i} \bigg|_{\alpha=+0} = 0, \forall x \in X, i = 0, \ldots, q(\bar{t}, \bar{l})$,

2. there exists a feasible $\bar{x} \in X$ such that $\frac{d^{q(\bar{t}, \bar{l})+1} f(\bar{x}, \bar{t} + \alpha \bar{l})}{d\alpha^{q(\bar{t}, \bar{l})+1}} \bigg|_{\alpha=+0} \neq 0$.

Given $\bar{t} \in T$, it is easy to see that the set $L(\bar{t})$ defined in (3) is a convex polyhedral cone in $\mathbb{R}^s$. Then, according to the known results on the convex polyhedral cone’s decomposition (see [3]), there exist a finite set of vectors $b_i, i \in \{1, \ldots, p\}$, $a_i, i \in I$, such that $L(\bar{t})$ admits a finite representation in the parametric form:

$$L(\bar{t}) = \{l \in \mathbb{R}^s : l = \sum_{i=1}^{p} \beta_i b_i + \sum_{i \in I} \alpha_i a_i, \alpha_i \geq 0, i \in I\}. \tag{4}$$
with \( p = s - \text{rank}(h_k, k \in \mathcal{K}), |I| < \infty, \beta_i \in \mathbb{R}, i \in \{1, \ldots, p\}, \alpha_i \in \mathbb{R}, i \in I. \)

Vectors \( b_i, i \in \{1, \ldots, p\} \) satisfy the conditions \( h_k^T b_i = 0, \ i = 1, \ldots, p, \ k \in K(\bar{t}), \) and are usually referred to as bidirectional extremal rays. Vectors \( a_i, i \in I, \) in turn, satisfy the inequalities \( h_k^T a_i \leq 0, \ i \in I, \ k \in K(\bar{t}), \) and are called unidirectional extremal rays. The extremal rays can be found explicitly (see [2]).

\textbf{Remark 1.} In the case of a pointed cone \( L(\bar{t}) \), the set of vectors \( b_i, i = 1, \ldots, p, \) is empty. If \( \bar{t} \in \text{int } T, \) then the set of vectors \( a_j, j \in I, \) is empty and \( b_i = e_i, i = 1, \ldots, p = s. \)

Suppose now that \( \bar{t} \in T^* \subset T \) is an immobile index in \( (P). \) Consider the corresponding sets \( \bar{L} = L(\bar{t}), \bar{K} = K_a(\bar{t}), \) and suppose that the extremal rays in \( \bar{L} \) are defined explicitly. Given a sufficiently small \( \varepsilon > 0, \) denote by \( T_\varepsilon(\bar{t}) \) an \( \varepsilon \)-neighborhood of \( \bar{t} \) in \( T: T_\varepsilon(\bar{t}) = \{ t \in T: ||\bar{t} - t|| \leq \varepsilon \}. \) From the parametric representation \( (4) \) of the cone of feasible directions \( \bar{L}, \) it follows that the local constraints \( f(x, t) \leq 0, \forall t \in T_\varepsilon(\bar{t}), \) can be presented in the form of the following modified constraints:

\[
\bar{f}(x, (\beta, \alpha)) \leq 0, \forall (\beta, \alpha), \ \alpha \geq 0, \ ||(\beta, \alpha)|| \leq \varepsilon, \quad (6)
\]

where \((\beta, \alpha)^T \in \mathbb{R}^{p+|I|}, \ \bar{f}(x, (\beta, \alpha)) := f(x, \bar{t} + B\beta + A\alpha), \) and the columns of matrices \( B \in \mathbb{R}^s \times p \) and \( A \in \mathbb{R}^s \times |I| \) are presented by the bidirectional and unidirectional rays respectively.

Without loss of generality we can use here the maximum norm given as \( ||y|| = \max_{i=1,\ldots,n} |y_i| \) for \( y \in \mathbb{R}^n. \) Then the modified constraints \( (6) \) can be considered as the box constraints w.r.t. variables \((\alpha, \beta), \ \alpha \in \mathbb{R}^{|I|}, \ \beta \in \mathbb{R}^p. \)

From the definition of the immobility index \( \bar{t}, \) it follows that for any \( x \in X, \) the vector \( \bar{t} \) maximizes the function \( f(x, t), \) or equivalently, vector \((\beta = 0, \bar{\alpha} = 0) \) is a solution of a so called lower level problem:

\[
\max_{(\beta, \alpha)} \bar{f}(x, (\beta, \alpha)), \ \text{s.t. } \alpha \geq 0. \quad (8)
\]

The first and the second order optimality conditions for the vector \((\beta = 0, \bar{\alpha} = 0) \) in problem \( (8) \) can be formulated as follows:

\[
\frac{\partial^T f(x, \bar{t})}{\partial t} b_i = 0, \ i = 1, \ldots, p; \quad \frac{\partial^T f(x, \bar{t})}{\partial t} a_i \leq 0, \ i \in I, \ \forall x \in X, \quad (9)
\]

\[
(\beta^T, \alpha^T)(B, A)^T \frac{\partial^2 f(x, \bar{t})}{\partial t^2} (B, A) \begin{pmatrix} \beta \\ \alpha \end{pmatrix} \leq 0, \quad (10)
\]

141
for all \((\beta^T, \alpha^T) \in \mathbb{R}^{p+|I|}\) such that \(\alpha_i = 0\) if \(\frac{\partial f_T(x, \bar{t})}{\partial t} a_i < 0\), and \(\alpha_i \geq 0\) if \(\frac{\partial f_T(x, \bar{t})}{\partial t} a_i = 0, i \in I\).

Assumption 1. Suppose that \(X \neq \emptyset\), the set \(T\) is bounded and \(q(l, t) \leq 1, \forall l \in L(t) \setminus \{0\}, \forall t \in T^*\).

It can be showed that Assumption 1 implies the finiteness of the set of immobile indices: \(T^* = \{t_j^*, j \in J_*\}\) with \(|J_*| < \infty\), and the existence of \(\bar{x} \in X\) such that \(|T_a(\bar{x})| < \infty\).

Suppose that the set of immobile indices and their immobility orders along the corresponding extremal rays are known ([5]). Denote:

\[
I_* := \{i \in I : q(\bar{t}, a_i) = 0\} = \{i \in I : \exists x^{(i)} \in X : \frac{\partial T f(x^{(i)}, \bar{t})}{\partial t} a_i < 0\},
\]

\[
I_0 := I \setminus I_* = \{i \in I : \frac{\partial T f(x, \bar{t})}{\partial t} a_i = 0, \forall x \in X\}.
\]

Taking into account Assumption 1, we get \(q(\bar{t}, b_i) = 1, i = 1, \ldots, p;\) \(q(\bar{t}, a_i) \geq 1, i \in I_0;\) \(q(\bar{t}, a_i) = 0, i \in I_*\). Then from conditions (9), (10), we conclude that for all \(x \in X\) and \((\beta, \alpha_0)^T \in \mathbb{R}^p \times \mathbb{R}^{|I_0|}\) it holds

\[
\frac{\partial T f(x, \bar{t})}{\partial t} b_i = 0, i = 1, \ldots, p; \quad \frac{\partial T f(x, \bar{t})}{\partial t} a_i = 0, i \in I_0, \quad (11)
\]

\[
\frac{\partial T f(x, \bar{t})}{\partial t} a_i \leq 0, i \in I_*; (\beta^T, \alpha_0^T) (B, A_0) \frac{\partial^2 f(x, \bar{t})}{\partial t^2} (B, A_0) \begin{pmatrix} \beta \\ \alpha_0 \end{pmatrix} \leq 0, \quad (12)
\]

where \(A_0 = (a_i, i \in I_0), \alpha_0 = (\alpha_i, i \in I_0)\).

Taking into account that for all \(\bar{t} \in T^*\) and any \(x \in X\) the relations (11), (12) are satisfied, and repeating the considerations made in [4] for the case of the box constrained index set \(T\), we prove the following implicit optimality criterion.

Theorem 1. Under Assumption 1, a vector \(x^0 \in X\) is optimal in the convex SIP problem (P) with polyhedral index set \(T\), if and only if there exists a finite set of indices \(\{t_j, j \in J_a(x^0)\} \subset T_a(x^0) \setminus T^*\) with \(|J_a(x^0)| \leq n\), such that \(x^0\) is optimal in the following auxiliary problem:

\[
(P_{aux}) : \min_{x \in \mathbb{R}^n} c(x)
\]

\[\text{s.t.} \quad f(x, t_j) \leq 0, \ j \in J_a(x^0),\]

142
\[ f(x, t_j^*) = 0, \frac{\partial T f(x, t_j^*)}{\partial t} B(j) = 0, \frac{\partial T f(x, t_j^*)}{\partial t} A_0(j) = 0, \frac{\partial T f(x, t_j^*)}{\partial t} A_*(j) \leq 0, \]

\[ (\beta^T(j), \alpha_{0}^T(j))(B(j), A_0(j)) \frac{\partial^2 f(x, t_j^*)}{\partial t^2}(B(j), A_0(j)) \begin{pmatrix} \beta(j) \\ \alpha_0(j) \end{pmatrix} \leq 0, \]

where \((\beta(j), \alpha_0(j))^T \in \mathbb{R}^{p(j)} \times \mathbb{R}^{I_0(j)}\), \(j \in J_*, B(j) = (b_i(j), i = 1, \ldots, p(t_j^*))\), and \(A_0(j) = (a_i(j), i \in I_0(t_j^*)), A_*(j) = (a_i(j), i \in I_*(t_j^*))\).

Notice that the auxiliary problem \((P_{aux})\) is also a SIP problem but it is more easy to study and solve than the original problem \((P)\) since

1. the infinite constraints in \((P_{aux})\) are quadratic w.r.t. multi-dimensional indices \((\beta(j), \alpha_0(j))^T \in \mathbb{R}^{p(j)} \times \mathbb{R}^{I_0(j)}\), hence this problem can be considered as a light generalization of the common semidefinite (SDP) problem (see [1]);
2. due to Assumption 1, the constraints of \((P_{aux})\) satisfy the Slater type condition, i.e. there exists a vector \(x \in X\) such that for all \(t_j^* \in T^*, j \in J_*\), it is satisfied:

\[ (\beta^T(j), \alpha_{0}^T(j))(B(j), A_0(j)) \frac{\partial^2 f(x, t_j^*)}{\partial t^2}(B(j), A_0(j)) \begin{pmatrix} \beta(j) \\ \alpha_0(j) \end{pmatrix} < 0, (1) \]

\[ \forall (\beta(j), \alpha_0(j))^T \in \mathbb{R}^{p(j)} \times \mathbb{R}^{I_0(j)}\), \( (\beta(j), \alpha_0(j))^T \neq 0; \]

3. explicit optimality conditions for SDP-type problems satisfying the Slater condition can be easy formulated and can be efficiently applied to theory and practice of SIP.

The novelty of the approach presented here consists in use of the fact that the immobile indices solve the lower level problem for all feasible \(x\). The analysis of the optimality conditions for the lower level problem allows one to form a new set of constraints that should be satisfied by all \(x \in X\), and to formulate new optimality conditions (implicit or explicit) for the original SIP problem \((P)\) in the form of CQ-free optimality criterion for a special auxiliary problem \((P_{aux})\) that has a more simple structure. Notice that in the convex case, such new optimality conditions are more strong than the known ones for SIP (see for example, [1,6]).

The authors were partially supported by the state program "Mathematical Models 13" of fundamental research in Republic of Belarus; and by the
Center for Research and Development in Mathematics and Applications (University of Aveiro) and the Portuguese Foundation for Science and Technology (“FCT”) (project no. PEst-C/MAT/UI4106/2011).

REFERENCES

2. Chernikova N.V. “Algorithm for discovering the set of all the solutions of a linear programming problem,” U.S.S.R. Computational Mathematics and Mathematical Physics, 8, No.6, 282-293 (1968).

Optimization methods for measurement of returns to scale in the non-radial DEA models

Vladimir Krivonozhko¹, Finn Førsund², Andrey Lychev³

¹ National University of Science and Technology “MISiS”, Moscow, Russia; KrivonozhkoVE@mail.ru
² University of Oslo, Oslo, Norway; finn.forsund@econ.uio.no
³ National University of Science and Technology “MISiS”, Moscow, Russia; lychev@misis.ru

The non-radial DEA models [1] possess some specific features. First, multiple reference sets may exist for a production unit. Second, multiple supporting hyperplanes may occur on optimal units of the frontier. Third,
multiple projections (a projection set) may occur in the space of input and output variables. All these features cause certain difficulties under measurement of returns to scale of production units.

The non-radial DEA model can be written in the following form [1, 2]

$$\max \ h = (C^+ T S^+ + C^- T S^-)$$

subject to

$$\sum_{j=1}^{n} X_j \lambda_j + S^- = X_o, \quad \sum_{j=1}^{n} Y_j \lambda_j - S^+ = Y_o,$$

$$\sum_{j=1}^{n} \lambda_j = 1, \quad S^+ \geq 0, \quad S^- \geq 0, \quad \lambda_j \geq 0, \quad j = 1, \ldots, n,$$

here $X_j = (x_{1j}, \ldots, x_{mj})$ and $Y_j = (y_{1j}, \ldots, y_{rj})$ represent the observed inputs and outputs of production units $(X_j, Y_j), \ j = 1, \ldots, n$. $S^- = (s_1^-, \ldots, s_m^-)$ and $S^+ = (s_1^+, \ldots, s_r^+)$ are vectors of slack variables. The superscript “T” indicates a vector transpose. The components of the objective-function vectors $C^+$ and $C^-$ are specified as follows:

$$c^-_k = (m + r)^{-1} (\max \{x_{kj} | j = 1, \ldots, n\} - \min \{x_{kj} | j = 1, \ldots, n\})^{-1},$$

$$c^+_i = (m + r)^{-1} (\max \{y_{ij} | j = 1, \ldots, n\} - \min \{y_{ij} | j = 1, \ldots, n\})^{-1},$$

$k = 1, \ldots, m, \quad i = 1, \ldots, r.$

In the model (1), an efficiency score for unit $(X_o, Y_o)$ is evaluated, where $(X_o, Y_o)$ is any production unit from the set $(X_j, Y_j), \ j = 1, \ldots, n$. If the optimal value $h^*$ of the model is equal to zero, then unit $(X_o, Y_o)$ is considered efficient, if $h^* > 0$, then the unit is inefficient [1].

Banker et al. [1] proposed a two-stage approach to determine returns to scale in these models. Sueyoshi and Sekitani [2] showed that this approach may generate incorrect results in some cases. An interesting approach was proposed for measurement of returns to scale based on using strong complementary slackness conditions (SCSC) in the non-radial DEA models.
The SCSC/NM model is written in the following form

\[
\max \eta \quad \left\{ \begin{array}{l}
\theta X_0 - \sum_{j=1}^n \lambda_j X_j \geq 0, \quad \sum_{j=1}^n \lambda_j Y_j \geq Y_o, \quad \sum_{j=1}^n \lambda_j = 1, \\
\lambda_j \geq 0, \quad j = 1, \ldots, n, \\
v^T X_0 = 1, \quad -v^T X_j + u^T Y_j + u_0 \leq 0, \quad j = 1, \ldots, n, \\
v \geq 0, \quad u \geq 0, \\
\theta = u^T Y_o + u_0, \\
\lambda_j + v^T X_j - u^T Y_j - u_0 \geq \eta, \quad j = 1, \ldots, n, \\
v - \sum_{j=1}^n \lambda_j X_j + \theta X_0 \geq \eta, \\
u + \sum_{j=1}^n \lambda_j Y_j - Y_o \geq \eta, \quad \eta \geq 0
\end{array} \right. 
\] (2)

The first six conditions are from the primal model (1), the next three conditions are from the dual problem, the tenth condition provides the equality of the objective functions of the primal and dual problems. The last three conditions express the SCSC constraints. In order to secure that strong complementarity is obtained the variable \( \eta \) is entered as the objective function in (2).

Our theoretical consideration and computational experiments show that the SCSC/NM method may not be efficient from the computational point of view. Model SCSC/NM generates ill-conditioned basic matrices during the solution process, which results in “strange results” that do not coincide with the optimal solution of the corresponding non-radial DEA model. This naturally contradicts the optimization theory.

In our work we propose a two-stage approach to measure returns to scale in the non-radial DEA models. At first stage, an interior point, belonging to the optimal face, is found using a special elaborated method. In our previous work [3] we proved that any interior point of a face has the same returns to scale as any other interior point of this face. At the second stage, we propose to determine the returns to scale at the interior point found in the first stage with the help of Banker and Thrall’s method [4] or using the direct method of Førsund et al. [5].
Our computational experiments documented that the proposed approach is reliable and efficient for solving real-life DEA problems.

The authors were supported by the Russian Foundation for Basic Research (project no. 11-07-00698-a).

References


Selection of the target audience by the leverage method in the expert system for advertising specialist

Ulyana Kulbida¹, Olga Kaneva ²

¹ Omsk State Technical University, Omsk, Russian; uni_form@mail.ru
² Omsk State Technical University, Omsk, Russian okaneva@yandex.ru

Advertising positioning is an independent branch within the general theory of positioning, and it determines the optimal way of presenting information about a trademark to a particular target audience.

The market position of a trademark defines the location of goods on the market, while the advertising position defines the individuality of the
trademark in the communications area. Well-known researchers of mar-
keting and advertising Rossiter and Percy proposed to conduct advertising
positioning of the trademark in levels: macro-level \((X - YZ)\), meso-level
\((I - D - U)\), micro-level \((a - b - e)\) \([1]\). Further we are going to examine
the above mentioned levels more detailed.

**Macro-level.** This level determines the place of the trademark within
appropriate product category, selects a target audience and the way of
the trademarks positioning, taking into account particular qualities of the
consumer or the product. The scheme of macro level can be expressed by
the following formula: “Product \(X\) provides to people \(Y\) assistance \(Z\)”.
In this formula \(X\) equals the trademark, \(Y\) equals the target audience,
while \(Z\) equals the benefits of the trademark.

**Meso-level.** On this level, it is to be decided what benefits should be
determined while positioning (the model \(I - D - U\)). Benefits emphasized
in advertising, should correspond with the following three main conditions:

1) Importance: correspondence between the benefits and the motive
that drives the one who buyes the trademark.

2) Delivery: the consumer’s subjective opinion about the trademarks
ability to provide benefits.

3) Uniqueness: the perceived ability of the trademark to provide ben-
efits better than other trademarks do.

**Micro-level.** This level defines the way of focusing on the main bene-
fits. The technique of focusing on the benefits of the trademark is based
on distinguishing characteristics (physical properties) of the product, its
benefits (what the buyer wants) and emotions (feelings of the buyer caused
by buying or by using the product). The model was named after the first
letters of these terms \(a - b - e\), in which \(a\) stands for attribute, \(b\) stands
for benefit, and \(e\) stands for emotion \((e^+\) positive, \(e^-\) negative).

Making decision on the macro-, meso- and micro-levels of positioning
allows us to make a conclusion about the position of the trademark. This
conclusion presents the trademarks creative strategy. Obtaining a detailed
scheme of such a conclusion is the decision of the unformalized problem
in the field of advertising.

Work “Expert system for positioning and brand advertising strategies”
\([2]\) proposes the algorithm of the solution of the first problem of the macro-
level (determination of the trademarks place within appropriate product
category) for the developed expert system. This algorithm is based on
the use of attractiveness and competitiveness indicators. The results of
the numerical experiments are given.

The next problem of the macro-level is the selection of the target audience. Target audience is a group of people to which the specific campaign advertising or sales promotion is aimed at [1]. According to the Rossiter and Percy there are five groups of customers:

1) New consumers of the product category. This group of people buying our product, gets acquainted with the category.
2) Loyal consumers, who buy our product on a regular basis.
3) Fickle consumers, who buy our products as well as other brands.
4) Fickle consumers of other brands, who buy products from other brands, but not ours.
5) Loyal to another brands consumers, who regularly buy someone else’s product brand.

Generally the target audience it is only one group of consumers. However, sometimes it is useful to define the primary and secondary target audiences.

Frequently, the purpose of the advertising campaign is not only to attract new customers from other groups, but to keep people who are already loyal to the brand. The group of loyal customers of the own brand it is usually the secondary target audience of new campaigns which are aimed at other groups of buyers.

It is necessary to estimate distributing capacity (the quantity of items or services that the company can sell) of each company using the following correlation: potential growth in sales (in monetary terms) to the value of the event of the advertising campaign, which can provide this growth in sales. This correlation is called “leverage”. If we know the number of the customers in the certain group and the cost value for realization of the advertising campaign in this group we can count the common value of the leverage for the group:

\[ Profit \ leverage = \frac{number \ of \ the \ customers \times \ growth \ of \ profit}{cost \ value \ of \ the \ advertising \ campaign}. \] (1)

Obviously, in the usual case a group of customers can be considered as the target audience only if the profit leverage of this group exceeds one unit. This means that the effect of the advertising campaign will surpass the cost of its implementation. The higher profit leverage, the larger group of buyers corresponds to the role of the target audience. The character of the changing of the leverage with time is almost equal for the
third and fourth groups of customers, consequently leverage is calculated not for five, but four groups of buyers. On the assumption of the values of the leverage we can determine the primary and sometimes secondary target audiences.

Cost calculation of the advertising campaign (the values in the denominator of the formula (1)) is a nontrivial problem. Lots of professionals are based on personal experience, common sense and simple interdependences in the calculation of the advertising budget. Recently there were more sophisticated methods of calculating the advertising budget, but also their evaluation and application should not be divorced from practice.

In the developed expert system, there is an approach that uses the index of $MEF$ - the minimum effective frequency (the model of the optimization of the minimum effective frequency of Rossiter-Danaher). The index $MEF$ shows the number of the consumers contacts with the advertisement, which is optimal for the influence on the target audience. Measured in absolute units [1]:

$$MEF = 1 + AMA \times (TA + BA + CA + PI),$$  \hspace{1cm} (2)

where 1 – initial level $MEF$ in one advertising contact;

$AMA$ – corrective factor “attention to the means of advertising”;

$TA$ – corrective factor “target audience”;

$BA$ – corrective factor “brand awareness”;

$CA$ – corrective factor “character of the advertisement”;

$PI$ – corrective factor “personal influence”.

Table 1 gives the approximate values for the quantities involved in the calculating of $MEF$.

The algorithm for calculating of the cost of the campaign

1) Let $i$ – number of the customers group $(i = 1, 4)$, $j$ – number of advertising media vehicle which is used $(j = 1, m)$.

2) Define the vectors: $\alpha = (\alpha_1, \ldots, \alpha_m)^T$ and $\beta = (\beta_1, \ldots, \beta_m)^T$, where $\alpha_j$ – cost of the developing promotional material for the media vehicle $j$, $\beta_j$ – the cost of placement of the one advertising exposure at the media vehicle $j$.

3) For each group of customers are counting the value of $\nu_i = MEF_i$, $i = 1, 4$, where $MEF_i$ – the minimum effective frequency for $i$-th group of customers (2). In the result we get vector quantity $\nu = (\nu_1, \nu_2, \nu_3, \nu_4)^T$. 150
4) Let $x_{ij}$ – number of the advertising exposure at $j$-th media vehicle for $i$-th target audience. We obtain a matrix of variables:

$$X = \begin{pmatrix}
x_{11} & x_{12} & \ldots & x_{1m} \\
x_{21} & x_{22} & \ldots & x_{2m} \\
x_{31} & x_{32} & \ldots & x_{3m} \\
x_{41} & x_{42} & \ldots & x_{4m}
\end{pmatrix}.$$

5) To determine the advertising budget of each customer group it is required to solve the following problem of integer programming:

$$\begin{align*}
\sum_{j=1}^{m} \beta_j x_{ij} & \to \min, \\
\sum_{j=1}^{m} x_{ij} & \geq \nu_j, & x_{ij} & \geq \gamma_j, & x_{ij} & \in \mathbb{Z}, & j = 1, m,
\end{align*}$$

where $\gamma_j \geq 0$ – the prescribed values, indicating the required number of advertising exposures at the $j$-th media vehicle for the $i$-th target audience. The quantity on the objective function is equal to the cost of placement of the advertising at the media vehicles for the $i$-th group of customers.

6) After solving integer programming problems for each group of customers we will get the values of $C_1, C_2, C_3$ and $C_4$ campaign by the formula:

$$C_i = \sum_{j=1}^{m} \beta_j x_{ij} + \sum_{j=1}^{m} \alpha_j.$$

Then we use the formula (1) for the counting the value of leverage for each customer group and determination which of the groups will be the target audience.

In the future we plan refinement algorithm by entering into the model (3) indicators of the effectiveness of advertising on each media vehicle.
### Table 1: The values of correction factors

<table>
<thead>
<tr>
<th>Factor</th>
<th>Value of the factor</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0</td>
</tr>
<tr>
<td>Attention to the means of advertising</td>
<td>Strong attention</td>
</tr>
<tr>
<td>Target audience</td>
<td>Loyal consumers of our brand</td>
</tr>
<tr>
<td>Brand awareness</td>
<td>Advertisement for brand recognition</td>
</tr>
<tr>
<td>Character of the advertisement</td>
<td>Information advertisement</td>
</tr>
<tr>
<td>Personal influence</td>
<td>Strong personal influence</td>
</tr>
</tbody>
</table>
Zonal control of lumped systems on different classes of feedback functions

Samir Kuliev\textsuperscript{1}

\textsuperscript{1} Cybernetics Institute of ANAS, Baku, Azerbaijan; copal@box.az

In this work we investigate a class of feedback control problems for dynamic, in the general case, nonlinear objects involving lumped parameters. For synthesized control actions we introduce the notion of zonality that means constancy of the synthesized control parameters’ values in each of the subsets (zones). These subsets are obtained by partitioning the set of all possible states of the object investigated. Formulas for the gradient of the target functional with respect to the optimizable parameters of the synthesized controls are obtained. These formulas can be used to build numerical solution schemes on basis of first order iterative optimization methods. Results of numerical experiments obtained by solving some test problems are given.

Let the controlled process be described by the following nonlinear differential equations system:

\begin{align*}
\dot{x}(t) &= f(x(t), u(t), p), \quad t > 0, \\
x(0) &= x^0 \in X^0 \subset \mathbb{R}^n, \quad p \in P \subset \mathbb{R}^m,
\end{align*}

where $x(t)$ is the $n$-dimensional vector function of the process state; $u(t) \in U$ is the $r$-dimensional control vector function; $U \subset \mathbb{R}^r$ is the closed set of the control actions’ admissible values; $p$ is the $m$-dimensional vector of the process’s constant parameters, the values of which are uncertain, but there is a set of their possible values $P$ and the density (weighting) function $\rho_P(p) \geq 0$ defined on $P$; $X^0$ is the set of possible values of the process’s initial states with the density (weighting) function $\rho_{X^0}(x^0) \geq 0$ given.

Control of the process (1) is realized with the use of feedback; the state vector $x(t)$ may be measured fully or partially. Observations of the process state may be carried out at discrete points of time or continuously. To control the process, we propose to choose the values of the synthesized control actions according to a subset (zone) the measured current process state belongs to. The subsets are obtained by partitioning the set of all possible phase states of the object.
The objective of the feedback control problem considered is to determine the values of the parameters of the zonal control actions \( u(t) \) minimizing the following functional

\[
J(u) = \int_{X^0} \int_{P} I(u, T; x^0, p) \rho_{X^0}(x^0) \rho_{P}(p) dP dX^0 / (mesX^0 \cdot mesP),
\]

(3)

\[
I(u, T; x^0, p) = \int_0^T g(x(t), u(t)) dt + \Phi(x(T), T).
\]

(4)

Here \( x(t) = x(t; x^0, p, u) \) is the solution to the system (1) under the admissible control \( u(t) \), initial state \( x^0 \), and the values of the parameters \( p; T = T(x^0, P) \) is the corresponding completion time of the process, which can be either a fixed quantity \( T = T(x^0, P) = const = \bar{T} \), or an optimizable function of the values of the initial state and of the object’s parameters \( T = \{T(x^0, P): T(x^0, P) \leq \bar{T}, x^0 \in X^0, p \in P\} \), where \( \bar{T} \) is given. The latter case arises in, as a rule, speed-in-action problems for control systems. We considered both cases.

The functional (3) and (4) defines the quality of control which is optimal on the average with respect to the admissible values \( x^0 \in X^0 \) and \( p \in P \). Denote by \( X \subset R^n \) the set of all possible states of the object under different admissible initial states \( x^0 \in X^0 \), the values of the parameters \( p \in P \), and the controls \( u(t) \in U \) for \( t \in [0, \bar{T}] \).

Let the set \( X \) be partitioned into given number \( L \) of open subsets \( X^i \) such that

\[
\bigcup_{i=1}^{L} X^i = X, \ X^i \cap X^j = \emptyset, \ i \neq j, \ i, j = 1, 2, ..., L,
\]

where \( X^i \) is the closure of the set \( X^i \). In the work we consider the following four types of feedback control problems, which differ in organization of feedback with the object and, therefore, in formation of the control actions’ values.

**Problem 1.** There are points of time \( \tau_j \in [0, \bar{T}], \ j = 0, 1, ..., N, \ \tau_0 = 0 \) given, at which it is possible to observe the current state of the process \( x(\tau_j) \in X \). The frequency of these observations is such that when the process state belongs to some subset, it is observed at least once. The values of the control \( u(t) \), which are constant for \( t \in [\tau_j, \tau_{j+1}) \), are
assigned depending on the value of the last observed current process state, namely, depending on the subset \( X^i, i = 1, 2, ..., L \) of the phase space \( X \) which the measured (observed) current state belongs to. Therefore

\[
u(t) = v^i = \text{const, } x(\tau_j) \in X^i, \ t \in [\tau_j, \tau_{j+1}),
\]

\( v^i \in U \subset R^r, i = 1, 2, ...L, \ j = 0, 1, ...N - 1, \ \tau_N = \bar{T}. \)

It is required to determine zonal values of the control \( v^i, \ i = 1, 2, ...L, \) optimizing the functional (3).

Problem 2. The control actions are defined in the form of a linear function of the results of observations of the state variables at given discrete points of time \( \tau_i \in [0, \bar{T}], \ i = 0, 1, ..., N: \)

\[
u(t) = K_1^i \cdot x(\tau_j) + K_2^i, \ t \in [\tau_j, \tau_{j+1}), \ x(\tau_j) \in X^i, \ t \in [0, \bar{T}],
\]

\( i = 1, 2, ..., L, \ j = 0, 1, ..., N - 1. \) (6)

Here \( K_1^i \) is the \((r \times n)\) matrix and \( K_2^i \) is the \(r\)-dimensional vector which are constant for \( t \in [\tau_{j-1}, \tau_j) \). The problem is to determine the values \( K_1^i, \ K_2^i, \ i = 1, 2, ..., L, \) optimizing the functional (3).

Problem 3. Continuous observation of the process state is carried out; the control actions take zonal values of the control:

\[
u(t) = w^i = \text{const}, \ x(t) \in X^i, \ t \in [0, \bar{T}],
\]

\( w^i \in U \subset R^r, i = 1, 2, ..., L. \) (7)

It is required to determine the zonal values of the control \( w^i, \ i = 1, 2, ...L, \) optimizing the functional (3).

Problem 4. Continuous observation of the process state is carried out; the control actions are defined by a linear function of the measured current values of the process variables:

\[
u(t) = L_1^i \cdot x(t) + L_2^i,
\]

\( x(t) \in X^i, \ t \in [0, \bar{T}], i = 1, 2, ..., L, \ j = 0, 1, ..., N - 1. \) (8)

Here \( L_1^i \) is the \((r \times n)\) matrix and \( L_2^i \) is the \(r\)-dimensional vector which are constant for each subset \( X^i \), i.e. while \( x(t) \in X^i \). It is required to determine the values \( L_1^i, \ L_2^i, \ i = 1, 2, ..., L, \) optimizing the functional (3).
Note that in all the four problems, the synthesized controls are defined by the finite-dimensional constant vectors and matrices.

To solve the optimization problems stated above numerically and to determine the control actions from the classes (5)-(8), we propose to use first order optimization methods and the corresponding standard software [1]. For this purpose, we obtained formulas for the gradient of the target functional using the technique of the target functional increment obtained at the expense of the optimizable arguments increment [2]. Results of numerical experiments carried out by the example of the solution to several model problems are given.

REFERENCES


The Heuristic Approach to movement optimization on single-track part of the railway net

Maya Laskova\textsuperscript{1}, Alexander Lazarev\textsuperscript{2}, Elena Musatova\textsuperscript{3}

\textsuperscript{1} The Institute of Control Sciences V. A. Trapeznikov Academy of Sciences, Moscow, Russia; laskovayamaya@moscow-index.ru
\textsuperscript{2} The Institute of Control Sciences V. A. Trapeznikov Academy of Sciences, Moscow, Russia; jobmath@mail.ru
\textsuperscript{3} The Institute of Control Sciences V. A. Trapeznikov Academy of Sciences, Moscow, Russia; nekolyap@mail.ru

Abstract: This paper represents our solution for the problem of movement organization based on timetable optimization on the problematic part of railway system, i.e. single-track line. The approximate solution of this problem was founded on the heuristic method. The method gives the exact results in the case of limited amount of parameters and also can
be used in the case with huge number of parameters due to reasonable computational time.

Key words: scheduling theory, algorithm, single-track railway problem.

Introduction

Single track railways are of great interest to scheduling theory because nowadays they are the most weak chain in the railroad transportation all over the world and especially in Russia. The article tend to solving bottleneck problems in transport network. Bottleneck is a part of the way with low bandwidth compare with other parts of the same road. Often it is the railway line with a limited number of tracks. It can also be a narrow bridge, tunnel or narrow causeway. Presence of such railway parts often cause delays and Timetable failure, and it is necessary to obtain good solution by optimization of schedule (Timetable). This task with single line track is known to be NP-hard; we made an attempt to create an Heuristic algorithm which will help to minimize reasonable computational time for solving this problem.

Problem formulation

We consider that arrival numerouse applications to stations 1 and 2 are known in advance. Considering the trains with following parameters: $N = N_1 \cup N_2$ — set of trains;

$N_1 = \{1, 2, ..., n\}$ — set of trains arrived at the station 1;

$N_1 = \{1, 2, ..., m\}$ — set of trains arrived at the station 2;

$r_1^i$ — planning time of $i$ train to station 1;

$r_2^j$ — planning time of $j$ train to station 2;

$d_1^i$ — the due date of arrival $i$ train $i \in (1, n)$ at the station 1 to the station 2;

$d_2^j$ — the due date of arrival $j$ train $j \in (1, m)$ at the station 2 to the station 1;

$p$ — the average time of movement the train ($p = \text{const}$);

$\delta$ — headway between trains.

Fact data

$C_1^1$ — real arrival time of $i$ train to station 2;

$C_2^1$ — real arrival time of $j$ train to station 1;
$S_i^1$ — real departure time of the $i$ train;
$S_j^2$ — real departure time of the $j$ train.

The objective function is following: $T\sum(\pi) = \sum_{i=1}^{N} \max\{0, C_i - d_i\}$

Our task was to create the train schedule optimizing the total delay of movement trains on the single-track part of the railway net. (Fig.1)

\[\text{Fig. 1. Bottleneck problem}\]

**Note 1.** If $\delta = p$, i.e. on the railway at the time can be no more than one train. The problem reduces to $(m + n)$ service requirements on a single device to the agreed arrival time and the due date. There are number of polynomial algorithms for solving such problems [1-3].

**Note 2.** Further, we assume that $\delta = 0$, i.e. the delay time of each train is negligible compared to the distance between stations.

The heuristic algorithm

The algorithm consists of two parts: direct flow and indirect flow.

**Definition 1.** $Batch(n, i)$ — a set, containing $n$ trains, departed from the station 2 at the same moment $C_n^i = \max\{r_{2n}^i, r_{1n}^i + p\}$.

Under the direct flow we create $Batch(n, i)$ (see def.1). For $Batch(n, i)$ we determine all possible departure moments $S_n^i$. In this way the trains from the station 1 with $r_{1i}^i \in (S_n^i - p; S_n^i + p)$ will depart at the moment $t = S_n^i + p$. At this stage we analyze the getting results of objective function $F(S_n^k)$; and afterwards we choose the smallest meaning. Thus
we got the optimal departure time $S_n$ for the least train. Similarly we find the optimal departure time $S_{n-1}, S_{n-2}, \ldots S_1$.

**Definition 2.** $Batch(m, j)$ — a set, containing $m$ trains, departed from the station 1 at the same moment $S_j^m = \max\{r^1_m, r^2_j + p\}$, $j \in (1, n)$.

In the calculation of the indirect flow we create $Batch(m, j)$.

The development of this algorithm allows to simplify the task of schedule optimization and it is especially useful for failure movement.

**References**


Finite time-interval robustness study of dynamic systems with imprecisely identified parameters

**Alexander V. Lotov**$^1$, **Georgij Kamenev**$^2$

$^1$ Dorodnicyn Computing Center of RAS, Moscow, Russia; avlotov@gmail.com
$^2$ Dorodnicyn Computing Center of RAS, Moscow, Russia; GKK@mail.ru

New technique for analysis of dynamic models of complex systems (economic, biological, etc.) is proposed. The models are characterized by ambiguity of behavior because of inaccuracy of identification of their parameters. In this case, quantitative analysis of the robustness of trajectories of the model at finite time intervals with respect to imprecision of parameters is needed. It is based herein on approximating the trajectory tubes. The robustness study can be used, for example, in the framework of multi-criteria decision support. The technique is based on methods for approximating the reachable sets of dynamical models and methods.
for approximating the feasible criterion set in multi-criteria problems (see [1]) as well as on application of these methods in visual identification of parameters [2].

Let the dynamics of the system under study to be described by the system of difference equations

\[
\frac{x^{k+1} - x^k}{\tau} = f(x^k, u^k, c^k), \quad k = 0, \ldots, N - 1,
\]

where \( x^k \in \mathbb{R}^n \) is the state vector, \( u^k \in \mathbb{R}^r \) is the control vector, \( c^k \in \mathbb{R}^s \) is the vector of unknown parameters, at the time-moment \( k \). The initial state \( x^0 = x_0 \) is assumed to be given. The technique consists of two steps.

Step 1. Parameter identification. It is assumed that there is information about the behavior of the dynamical system under study in the past that allows to draw conclusions about possible values of the parameter vector \( c \). The parameter identification method is based on visualization of the multi-dimensional graph of error function [2], which helps to analyze the stability of the solution of the error function minimizing problem. If the solution is not stable, the expert points out such a region \( C \) in the parameter space \( \mathbb{R}^s \), that the solution of the parameter identification has the form \( c \in C \). In such an approach, the model parameters are identified by using a synthesis of observations and non-formal experience of the expert. Further, the study examines the case when the region \( C \) contains more than one point.

Step 2. Stability analysis

2a. Let us consider the case when the vector \( c \in C \) is constant in time, i.e. \( c^k = c, k = 0, \ldots, N - 1 \). In this case, for a given control \( \hat{u}^k, k = 0, \ldots, N - 1 \), the system (1) allows constructing the trajectory of the system for each vector \( c \in C \). The output for the entire set \( C \), i.e. the set of uncertainty, can be approximated for the time moment \( k = N \) as well as for any intermediate time moment. It can be done using the method for approximating the set of feasible objective values for nonlinear systems described in [1]. The method is based on the constructing the trajectories for a large number of random vectors from \( C \) and further covering the set of the ends of the trajectories, i.e. the points \( x^N \), by a relatively small number of parallelotops, i.e. polytopes which hyperplanes are parallel to the coordinate planes. The method is provided with an estimate of the quality of such an approximation.

The constructed sets of uncertainty are approximations of cross-sections of the trajectory tubes of the system (1) while \( c \in C \) with the given control.
\( \hat{u}^k, k = 0, \ldots, N - 1 \). In total, they approximate the trajectory tube of the system (1), which we denote by \( X \). Visualization of the cross-sections allows quantitatively analyzing the stability of the trajectories of the system (1) with respect to the uncertainty in the parameter vector \( c \in C \). It supports evaluating the possibility of using the tube of trajectories in the decision problem of the selecting the control.

2b. Let the vector \( c \) change in time. Let us consider a possible deviation of a trajectory from the "unperturbed" vector-valued function \( \hat{c}^k, k = 0, \ldots, N - 1 \). Under the condition \( c \in C \). If it is not too large, it is possible to use the linearization of the equation (1). Suppose, moreover, that the set \( C \) is convex. For given function \( \hat{u}^k, k = 0, \ldots, N - 1 \) and \( \hat{c}^k, k = 0, \ldots, N - 1 \), the trajectory of (1) is constructed, which we denote by \( \hat{x}^k, k = 0, \ldots, N \). Then, the equation (1) is linearized in the neighborhood of \( \hat{x}^k, k = 0, \ldots, N, \hat{u}^k, k = 0, \ldots, N - 1 \) and \( \hat{c}^k, k = 0, \ldots, N - 1 \). Next, by using the methods of polyhedral approximation of the reachable sets in the convex case (see [1]), a collection of sets of uncertainty for the linearized system for \( N \) steps, as well as for intermediate points in time, is constructed. In the same way as in the case 2a, this set of approximations allows us to study the tube of trajectories of the linearized system, which is also denoted by \( X \). Just as in the previous case, the visualization of the cross-sections allows to quantitatively analyze the stability of the trajectories of the system with respect to the uncertainty in the parameter vector \( c \) and to evaluate the possibility of using the tube of trajectory tubes in the decision problems.

Assume that the difference system turned out to be sufficiently robust to the imprecision of parameters. Let us consider a decision problem characterized by \( m \) criteria, denoted by \( y \) and a finite number control alternatives \( u_1, \ldots, u_M \) where \( u_j = (u_j^k, k = 0, \ldots, N - 1) \). For the simplicity, let us assume that \( y = f(x^N) \) and \( c \in C \) does not change in time. By using the method for approximating the feasible set of criterion vectors for nonlinear systems mentioned in section 2a, one can approximate the feasible criterion set \( Y^{(j)} \) related to any given control \( u_j \) and all possible vectors \( c \in C \). By his the problem is reduced to selecting a control on the basis of comparing the sets \( Y^{(j)}, j = 1, \ldots, M \). Selecting a control can be based on the reasonable goals method [1]. For the case of the imprecise outputs of alternatives, this method was proposed in [3] and studied in [4]. In the framework of the method, the convex hull of the best points of the
polyhedra that approximate the sets $Y^{(j)}, j = 1, \ldots, M$, is constructed. The interactive visualization of the convex hull is used for studying its multi-dimensional Pareto frontier. By specifying the most preferred point of the Pareto frontier (reasonable goal), the user expresses his/her preferences. Then, a small number of such alternatives is found that are in line with the specified goals.

The authors were supported by the Russian Foundation for Basic Research (projects no.10-01-00199 and 11-01-12136-ophi-m) and Programs of fundamental research of Russian Academy of Sciences Pi-15 and Pi-18.

REFERENCES


Estimation of economic damage from human mortality by external causes on macro-, meso-, micro-levels

A.A. Lukovenko1, T.M. Tikhomirova2

1 Plekhanov Russian University of Economics, Moscow, Russia; gelya-14@yandex.ru

2 Plekhanov Russian University of Economics, Moscow, Russia; t_tikhomirova@mail.ru

In modern Russia there is no universal accepted methods for estimation of economic damage from human untimely mortality. Damage from
mortality by external causes may be evaluated in particular as: GDP per employed person, annual salary, consumption, etc. But in practice using any methods leads to a number of difficulties and criticisms, which essentially is that every economic level has its own goals and objectives that affect the incoming flow of socioeconomic factors.

In order to estimate economic damage from untimely mortality it is necessary to analyze contributing factors and develop particular methods inherent in every level, which may work as in the system as separately for individual appraisals.

It is worth to mentioning that on macrolevel the estimation of damage from mortality by external causes should be maximum aggregated and correlated with social prosperity of the country or group of countries. The estimation will allow to determine the amount of short-received GDP due to disposal of the economically active population and to reflect the amount of social benefits and payments in the event of unexpected death.

On meso-level definition of economic damage estimation affords to reflect the system of measures in order to level differentiation between regions on the socioeconomic development, also improve the targeting to allocate investments, required to improve the living standards of regions.

On microlevel the damage from unexpected human death should be detailed to the level of an objective appraisal of individual life insurance.

Depending on the level the estimation of economic damage may be characterized by the influence of a various set of factors, such as: human life cost, death rate by external causes, inflation, crime, activities risk-bearing for life - on macrolevel; gender and age population structure, marriage and divorce rates, migration, educational level, Human Development Index - on meso-level; gender, age, marital condition, area of employment, chronic diseases, income level, stress level in region of residence, educational level - on microlevel.

The application of economic and mathematical approaches and models plays an important role in development of methods for definition of economic damage from external causes for every level. For macrolevel application of factor models is most characteristic, where the estimation
of damage is influenced by macroeconomic and socio-demographic factors.

On meso-level object of observation is a city, region or group of regions. The estimation of damage strongly depends on both the microeconomic factors that reflect the geographical features and the specifics of the region and the macroeconomic factors that regulate the economy in the region. To develop methods for economic damage definition there will be required a probabilistic model, where the independent variables will be socioeconomic factors.

On microlevel object of observation is a person or group of people, so that all the influencing factors are taken into account with a certain probability, which allows the application of probabilistic models.

REFERENCES

2. V.I. Orlov “Differentiated Estimation of health and economic losses due to untimely mortality”, Diss...C.m.s., FRI HOI, Moscow, (2009).

Statistical Methods and Optimization in Data Mining

Eloísa Macedo¹, Adelaide Freitas²

¹ University of Aveiro, Aveiro, Portugal; macedo@ua.pt
² University of Aveiro, Aveiro, Portugal; adelaide@ua.pt

The increasing number of the sequenced genomes has created new challenges in several scientific domains, namely statistics, optimization and computer sciences. Various numerical transformations related to the sequenced genomes (e.g., frequency of each nucleotide, association between consecutive genomic symbols) have been proposed in order to take advantage of statistical methodologies available for quantitative data. It is
expected that such numerical data sets contain useful information about mathematical properties of DNA sequences. An important issue associated with data sets where each individual is characterized by a high-dimensional vector of variables consists in the identification of patterns or homogeneous groups. Since high dimensionality turns the visualization and analysis of data into a complex problem, the space reduction and the features subset selection techniques are aimed to facilitate the visualization and capture the important and relevant relationships existing in data.

To detect the existence of patterns in a data matrix \( (n \text{ objects} \times p \text{ variables}) \), it is often desirable to partition the data sets according to some similarity criteria. This task is related to the data mining technique of partitioning data sets into groups of objects with some similar properties (clusters) called clustering. There exists a variety of clustering techniques designed for several data types, applied in many areas such as pattern recognition, image segmentation and bioinformatics [1,4,5]. Clustering problems are usually formulated as mixed-integer problems, or \((0,1)\)-semidefinite and semi-infinite programming problems that in turn can be reduced to nonsmooth and nonconvex nonlinear problems [2,4].

While dimensionality reduction of objects is usually achieved by clustering techniques, the dimensionality reduction of the variable space can be provided applying statistical techniques such as Principal Component Analysis (PCA), to detection of a lower number of uncorrelated variables (components) able to explain the maximum variability of the data. The reduction of objects and variables can be obtained applying the two techniques sequentially. Recently, a new technique called Clustering and Disjoint Principal Component Analysis (CDPCA) was suggested in [3] to solve the clustering of objects and the partition of variables using PCA simultaneously. This technique permits to cluster objects along a set of centroids and to partition of variables along a reduced set of components, in order to maximize the between cluster deviance of the components in the reduced space. The model obtained is a quadratic mixed continuous and integer optimization problem. In [3], this model is solved by an alternating least-squares (ALS) algorithm that can be considered as an heuristic that divides the model solving iteratively in four steps, modifying in each step certain parameters of the data. The methods of Mixed-Integer Programming are used on the basic steps of the algorithm. In [3], the CDPCA algorithm was tested for two data sets, one with 20 objects and
The main objective of this work is to test the ability of this new technique on biological data sets to make possible visual representation of relevant characteristics for data interpretation. For this purpose, we implemented CDPCA in R language, which is an open source software widely used in statistics, with a lot of specific packages for efficient data treatment [6].

Let us introduce the notations that will be used.

— $X = [x_{ij}]$ is the data matrix with $I$ objects and $J$ variables (variables are supposed to be normalized);
— $P$, $Q$ are the desirable numbers of clusters of objects and variables, respectively;
— $E$ is the $I \times J$ error matrix;
— $U = [u_{ip}]$ is a $I \times P$ binary matrix and row stochastic defining a partition of objects into $P$ clusters where $u_{ip} = 1$ if the $i$-th object belongs to cluster $p$, otherwise, $u_{ip} = 0$;
— $V = [v_{jq}]$ is a $J \times Q$ binary matrix and row stochastic defining a partition of variables into $Q$ clusters where $v_{jq} = 1$ if the $j$-th variable belongs to cluster $q$, otherwise, $v_{jq} = 0$;
— $A$ is the $J \times Q$ matrix of the coefficients of the linear combination, such that $\text{rank}(A) = Q$ and each row (variable) contributes to a single column (component);
— $Y = [y_{iq} = \sum_{j=1}^{J} a_{jq} x_{ij}]$ is the $I \times Q$ component score matrix where $y_{iq}$ is the value of the $i$-th object for the $q$-th component $y_q$ (common information of a subset of variables);
— $\bar{X}$ is the $P \times J$ matrix of individual centroids in the space of the observed variables;
— $\bar{Y}$ is the $P \times Q$ matrix of individual centroids in the reduced space.

The model associated to CDPCA minimizes the norm of the error matrix

$$E = X - U\bar{Y}A^T$$

w.r.t. parameters representing $U$, $\bar{Y}$ and $A$ subject to certain constraints. According to [3], $A$ can be decomposed in the form $A = BV$, where $B$ is a $J \times J$ diagonal matrix of the form

$$B = \sum_{q=1}^{Q} \text{diag}(v_q)\text{diag}(c_q),$$
where $v_q$ is the vector corresponding to column $q$ in matrix $V$ and $c_q$ is the eigenvector associated to the largest eigenvalue of the matrix

$$\text{diag}(v_q) \tilde{X}^T U^T U \tilde{X} \text{diag}(v_q).$$

We can formulate the problem as follows.

$$\max_{U, \tilde{X}, B, V} \|U \tilde{X} BV\|^2 = \max_{v, c, \tilde{x}, u} \sum_{p=1}^P \sum_{q=1}^Q \left( \sum_{j=1}^J v_{jq} c_q \tilde{x}_{pj} \right)^2 \sum_{i=1}^I u_{ip}$$

s. t. $\sum_{p=1}^P u_{ip} = 1, \quad u_{ip} \in \{0, 1\}, \quad i = 1, \ldots, I; p = 1, \ldots, P,$

$$\sum_{q=1}^Q v_{jq} = 1, \quad v_{jq} \in \{0, 1\}, \quad j = 1, \ldots, J; q = 1, \ldots, Q, \quad (P)$$

$$\sum_{j=1}^J c_{jq}^2 = 1, \quad q = 1, \ldots, Q,$$

$$\sum_{j=1}^J c_{jq} c_{jr} = 0, \quad q = 1, \ldots, Q - 1; r = q + 1, \ldots, Q.$$

The alternating least-squares algorithm suggested in [3] alternates four basic steps: update $V$ (allocation of variables), update $B$ (the PCA step), update $U$ (allocation of objects) and update $\tilde{X}$ (centroid matrix), and it is summarized in the following box. Here, the estimates of the matrices are denoted by $\hat{\cdot}$.
ALS Algorithm for CDPCA

input: numeric data matrix $X$ and tolerance $\varepsilon$

Generate (e.g. randomly) $\hat{U}$ and $\hat{V}$, considering the constraints of problem (P);

Compute $\hat{X} = \left(\hat{U}^T \hat{U}\right)^{-1} \hat{U}^T X$. Set $k=1$;

while $F_{k+1}(\hat{B},\hat{U},\hat{X},\hat{V}) - F_k(\hat{B},\hat{U},\hat{X},\hat{V}) < \varepsilon$:

Update $B$: Given $\hat{X},\hat{U},\hat{V}$, calculate $\hat{B} = \sum_{q=1}^{Q} \text{diag}(v_q) \text{diag}(c_q)$.

Update $V$: Given $\hat{B},\hat{X},\hat{U}$, for $j = 1, \ldots, J$, set:

$$\hat{v}_{jq} = \begin{cases} 1, & \text{if } F(\hat{c}_q,\hat{U},\hat{X},[v_{jq}]) = \max_{r=1,\ldots,Q} \{F(\hat{c}_r,\hat{U},\hat{X},[v_{jr} = 1])\} \\ 0, & \text{otherwise.} \end{cases}$$

where $F(\hat{c}_q,\hat{U},\hat{X},\hat{V}) = \|\hat{U} \hat{X} \hat{B} \hat{V}\|^2$.

Update $U$: Given $\hat{B},\hat{X},\hat{V}$, for $i = 1, \ldots, I$, set:

$$\hat{u}_{ip} = \begin{cases} 1, & \text{if } \|\hat{V}^T \hat{B} x_i - \hat{V}^T \hat{B} \hat{x}_p\|^2 = \min_{s=1,\ldots,P} \{\|\hat{V}^T \hat{B} x_i - \hat{V}^T \hat{B} \hat{x}_s\|^2\} \\ 0, & \text{otherwise.} \end{cases}$$

Update $\hat{X}$: Given $\hat{B},\hat{U},\hat{V}$, calculate $\hat{X} = \left(\hat{U}^T \hat{U}\right)^{-1} \hat{U}^T X$.

Compute $F_k(\hat{B},\hat{U},\hat{X},\hat{V}) = \|\hat{U} \hat{X} \hat{B} \hat{V}\|^2$.

do $k = k + 1$;

The algorithm stops when the difference between consecutive computations of the values of the objective function of problem (P) is smaller than a specified threshold $\varepsilon > 0$. According to [3], since $F(B,U,X,V)$ is bounded above, the algorithm converges to a stationary point, which is a local maximum of problem (P). To guarantee that the algorithm finds the global minimum, the authors of the heuristic in [3] suggest to apply the algorithm repetitively for different initial values of matrices $U$ and $V$, that are randomly chosen.

In order to test the ability of the CDPCA to reveal and visualize biologically meaningful patterns in a 2-dimensional reduced space, we have implemented the algorithm using R and carried out an experimental study involving several real data sets extracted from molecular biology domain. Besides the matrices $U, V, A$, the implementation of CDPCA suggested in this work returns a pseudo-confusion matrix and draws two scatterplots where the data are displayed in the 2-dimensional reduced space, one where the objects are labelled according to the real classification and other
with the classification found by CDPCA. The pseudo-confusion matrix indicates the number of objects introduced in each cluster (the real and that found by CDPCA).

On the basis of the realized numerical tests we conclude that the implementation of the CDPCA algorithm in R is efficient for the tested data sets. The main advantage of this technique is that each component is characterized by a disjoint set of variables. This offers a promising approach for the clustered visual representation of data. On the other hand, it permits to overcome the difficulties on the interpretability of the data in the reduced space. The proposed heuristic can be improved, since we can update the parameters of problem (P) simultaneously using optimization methods that efficiently use the structure and properties of this problem. This is a subject of further research.

This work was supported by FEDER founds through COMPETE-Operational Programme Factors of Competitiveness and by Portuguese founds through the Center for Research and Development in Mathematics and Applications (CIDMA, University of Aveiro) and the Portuguese Foundation for Science and Technology (FCT).

REFERENCES

On control with coefficients for high order partial differential equations

Igor E. Mikhailov\textsuperscript{1}, L.A. Muravey\textsuperscript{2}

\textsuperscript{1} Dorodnicyn Computing Centre of RAS, Moscow, Russia; mikh_igor@mail.ru
\textsuperscript{2} MATI Russian State Technological University, Moscow, Russia; pm@mati.ru

1. On classes of existence of optimization problems of form

Let’s consider problem formulation on an example of form optimization of a thin plate under the influence of the distributed loading described by the model equation

\[
\Delta(D(x, y)\Delta u) = q(x, y), \quad (x, y) \in \Omega \subset \mathbb{R}^2,
\]

(1.1)

where $\Delta$ is the two-dimensional Laplace operator, $\sqrt[3]{D(x, y)} = h(x, y)$ is the distribution of a thickness of the plate, $u(x, y)$ is the form of the plate deflection, $q(x, y)$ is the density of the distributed loading, the area $\Omega$ is the plate base. For definiteness we will consider the case of hinged fastening plate, with the following conditions on the boundary $\Gamma = \partial \Omega$

\[
u|\Gamma = \Delta u|\Gamma = 0.
\]

(1.2)

It would appear reasonable that the equation (1.1) with boundary conditions (1.2) is named as the state equation of plate L (it contains the elliptic operator of the fourth order). The optimization problem is to minimize the integral functional

\[
F(h) = \int\int_{\Omega} f(x, y, u, \Delta u) \, dx \, dy,
\]

(1.3)

where the control parameter $h$ (a thickness of the plate) belongs to some class of functions $K$. An examples are the functional

\[
F_0(h) = \int\int_{\Omega} [u(x, y) - z(x, y)]^2 \, dx \, dy = \|u - z\|_{L_2(\Omega)}^2,
\]

(1.4)

expressing a deviation in norm $L_2(\Omega)$ of the plate form $u(x, y)$ with thickness $h(x, y)$ from the set form $z(x, y)$, or the functional

\[
F_1(h) = \int\int_{\Omega} q(x, y)u(x, y) \, dx \, dy = \int\int_{\Omega} D(x, y)(\Delta u)^2 \, dx \, dy,
\]

(1.5)
expressing the work required for a deflection the plate of the thickness 
$h(x, y)$ under the influence of force with density $q(x, y)$. Let’s notice that the state equation $L$ can be reduced to a system of the elliptic equations of the second order

\[
\Delta u = \frac{v}{D}, \quad (x, y) \in \Omega, \tag{1.6}
\]

\[
\Delta v = q, \quad (x, y) \in \Omega, \tag{1.7}
\]

\[
u|_{\Gamma} = v|_{\Gamma} = 0. \tag{1.8}
\]

The equation (1.7) is a special case of the equilibrium equation of the fixed elastic membrane

\[
\text{div}(k(x, y)\nabla v) = q, \quad (x, y) \in \Omega, \tag{1.9}
\]

\[
v|_{\Gamma} = 0, \tag{1.10}
\]

where function $k(x, y)$, named by a tension factor, belongs to some class $K$ which consists from measured in $\Omega$ functions $k(x, y)$, satisfying almost everywhere in $\Omega$ inequalities $0 < k_0 \leq k(x, y) \leq k_1 < \infty$. Then at $q \in L_2(\Omega)$ $v(x, y)$ is the generalized solution of the problem (1.7), (1.8) which belongs to the Sobolev space $W^{1,2}_0(\Omega)$.

Let’s notice that for the analogue of the functional (1.5) $F_1(k) = \int_{\Omega} k(x, y)|\nabla v|^2 \, dx \, dy$ the minimization problem in the class $K$ has the solution, that was shown by M. Goebel [1] (the Goebel’s method does not depend on space dimension).

At the same time for functional $F_0(k) = \|v - z\|_{L_2(\Omega)}$ the minimization problem in the class $K$, named the Lions problem, has no solution (see counter-examples of M. Murat [2] and D. Korsakova [3]). From these examples it follows that for resolvability of an optimization problem it is required to establish certain connection between the state equation $L$, minimized functional $F$ and a class $K$ of controlling functions, i.e. of coefficients in the equation (1.9). As was shown by E. De Giorgi, S. Spagnolo and T. Tartar [4], [5] for the solution of the Lions problem it is necessary to replace the equation (1.9) with the equation

\[
\text{div}(A(x, y)\nabla v) = q(x, y), \quad (x, y) \in \Omega,
\]
where the symmetric positively defined matrix $A(x, y)$, such that
\[ 0 < \alpha|\xi|^2 \leq (A\xi, \xi) \leq \beta|\xi|^2, \quad \alpha < \beta, \quad \xi \in \mathbb{R}^2, \]
belongs to the class $K$ of controlling matrixes $A$ if its own numbers $\alpha \leq \lambda_1(x, y) \leq \lambda_2(x, y) \leq \beta$ satisfy to relation $\lambda_1 = \alpha \beta / (\alpha + \beta - \lambda_2)$.

Establishment of conditions on functional $F(h)$ from (1.3), allowing to solve considered minimization problems in the case of the state equations of the fourth order in a class of scalar functions
\[ K = \{ k(x, y): 0 \leq h_0 \leq h(x, y) \leq h_1 < \infty, \ (x, y) \in \Omega \}, \quad (1.11) \]
that is dictated by physical sense of the equation ($h(x, y)$ is a plate thickness), has been made by L. Muravey, I. Ismailov, E. Eyniev, I. Mikhailov in [6], [7]. These conditions for the functional (1.3) are the continuity the function $f(x, y, t, \tau)$ with respect to all of the variables, growth with respect to variable $t$ and convexity on variable $\tau$.

These conditions, in particular, allowed to solve the minimization problem both for the functional $F_0(h)$, and for the functional $F_1(h)$, as well as to develop effective numerical methods of construction of the approximate solution of the optimization problem. As an example we will fulfill the calculation of a rectangular plate with the sides $a = 1, b = 2$. Restrictions in (1.12) were $h_0 = 0; h_1 = 3$. As $z(x, y)$ in the functional $F_0(h)$ was the function $z(x, y) = \frac{27}{128}x(a - x)y^2(b - y)$. For this function $z_y(x, 0) = 0$.

Fig. 1

Fig. 1 shows the received distribution $h(x, y)$. It is seen from Fig. 1 that the thickness of a plate sharply increases near the line $y = 0$ and reaches the top restriction $h_1$.

It confirms that the considered model adequately reflects the phenomenon, well-known from the elasticity theory, that at hinged fastening of the loaded plate at $y = 0$ the maintenance of a zero deflection of a plate $u_y(x, 0) = 0$ is possible only if its thickness is infinite in the vicinity $y = 0$.

Fig. 2 shows the plate deflection $u(x, y)$. 172
Thus the offered mathematical model and its numerical realization qualitatively correctly describe a plate deflection.

2. Solution of the optimization problem in case of the general equilibrium equation of a thin plate of a variable thickness

In this point as the state equation we will consider the equation \([8]\)

\[
\Delta(D(x,y)\Delta u) - (1 - \nu)(D_{yy}u_{xx} - 2D_{xy}u_{xy} + D_{xx}u_{yy}) = q(x,y),
\]

\((x,y) \in \Omega \subset \mathbb{R}^2,\)

(2.1)

with boundary conditions

\[
u|_\Gamma = \Delta u|_\Gamma = 0,
\]

(2.2)

where the Poisson number \(\nu\) can vary in the limit from 0 to 1/2 (in practice \(\nu = 1/3\), or \(\nu = 1/2\) are taken).

The object of investigation is minimization of the functionals

\[
F_0(h) = \|u - z\|^2_{L_2(\Omega)}
\]

(2.3)

\[
F_1(h) = \iint_{\Omega} q(x,y)u(x,y) \, dx \, dy =
\]

\[
\iint_{\Omega} [D(\Delta u)^2 - (1 - \nu)(D_{yy}u_x^2 + 2D_{xy}u_xu_y + D_{xx}u_y^2)] \, dx \, dy.
\]

(2.4)

Let’s notice that, as a plate is thin then the class \(K_1\) of controlling functions \(D(x,y)\) has the form \(\{D(x,y) \in C^2(\Omega), \ 0 < \alpha \leq D(x,y) \leq \beta < \infty, \ |D_{xx}| + |D_{xy}| + |D_{yy}| \leq \delta\}\), where \(\delta > 0\) is sufficiently small. It is not difficult to see that

\[
(1 - \nu)\iint_{\Omega} (D_{yy}u_x^2 + 2D_{xy}u_xu_y + D_{xx}u_y^2) \, dx \, dy \leq (1 - \nu)\delta\|\nabla u\|_{L_2(\Omega)} \leq
\]

173
\[ \leq \varepsilon \| \nabla u \|_{L^2(\Omega)} + C(\varepsilon, \delta) \| u \|_{L^2(\Omega)}, \]

where \( \varepsilon \) is an arbitrary positive number, \( C(\varepsilon, \delta) = \text{const.} \). It follows that the problem of the minimization of the functionals (2.3) and (2.4) can be solved in the class \( K_1 \), a similar class \( K \) in item 1.

At the same time the aim of work was to obtain numerical estimates of errors in the computation of the deformation and the thickness of the plates with the use of real equilibrium equation (2.1) in comparison with the model equation (1.1). The calculations have shown, that upon minimizing the functional the real deformation of the plate described by the equation (2.1) with \( \nu = 1/2 \), differed by more than 20% from the deformation of a plate described by the equation (1.1). The difference of thickness of plates was about 6%.

A separate work will be devoted to more detail study of the problem of minimization the functional \( F_1(h) \).

References

Analytic design of an optimal controller under permanent stochastic disturbances

Arsalan Mizhidon

East Siberia State university of Technology and Management, Ulan-ude, Russia; miarsdu@mail.ru

In [1] there is a problem of analytic design of an optimal controller under permanent stochastic disturbances for the case of functionals with an integrand of the form $x'Qx + u'Ru$. In this paper the problem of analytic design of an optimal controller under permanent stochastic disturbances was generalized to the case of functionals with an integrand of the form $x'Qx + x'Pu + u'Ru$. The problem of analytic design of an optimal controller under permanent determinate disturbances for such case of functionals was considered in [2].

Consider a linear system with constant coefficients

$$
\dot{x} = Ax + Bu + Gz,
$$

where $x$ — $n$-dimensional vector of state variables; $u$ — $r$-dimensional control vector; $A$ — $(n \times n)$ - matrix; $B$ — $(n \times r)$ -matrix; $z$ — $p$-dimensional vector of stochastic disturbances; $G$ — $(n \times p)$ -matrix.

Suppose control constraint is not imposed. In addition, suppose that $z(t)$ is a persistent bounded stochastic function of time, which can be represented as the solution of a linear system of stochastic equations

$$
\dot{z} = Dz + \xi(t).
$$

Here $D$ — $p \times p$ - matrix, $\xi(t)$ — vector stochastic process as white noise.

Stochastic process modeling, as white noise transmission through linear system (2), matches to stochastic stationary process modeling with a known spectral energy density. Obviously matrix $D$ is stable, otherwise the solution of the equation (2) would be an unlimited stochastic process.

It is required to define control $u(x, z)$ minimizing following functional

$$
J(u) = \lim_{T \to \infty} \frac{1}{2T} \int_{0}^{T} (x'Qx + x'Pu + u'Ru)dt,
$$

where $x$ — $n$-dimensional vector of state variables; $u$ — $r$-dimensional control vector; $A$ — $(n \times n)$ - matrix; $B$ — $(n \times r)$ -matrix; $z$ — $p$-dimensional vector of stochastic disturbances; $G$ — $(n \times p)$ -matrix.

Suppose control constraint is not imposed. In addition, suppose that $z(t)$ is a persistent bounded stochastic function of time, which can be represented as the solution of a linear system of stochastic equations

$$
\dot{z} = Dz + \xi(t).
$$

Here $D$ — $p \times p$ - matrix, $\xi(t)$ — vector stochastic process as white noise.

Stochastic process modeling, as white noise transmission through linear system (2), matches to stochastic stationary process modeling with a known spectral energy density. Obviously matrix $D$ is stable, otherwise the solution of the equation (2) would be an unlimited stochastic process.

It is required to define control $u(x, z)$ minimizing following functional

$$
J(u) = \lim_{T \to \infty} \frac{1}{2T} \int_{0}^{T} (x'Qx + x'Pu + u'Ru)dt,
$$

where $x$ — $n$-dimensional vector of state variables; $u$ — $r$-dimensional control vector; $A$ — $(n \times n)$ - matrix; $B$ — $(n \times r)$ -matrix; $z$ — $p$-dimensional vector of stochastic disturbances; $G$ — $(n \times p)$ -matrix.

Suppose control constraint is not imposed. In addition, suppose that $z(t)$ is a persistent bounded stochastic function of time, which can be represented as the solution of a linear system of stochastic equations

$$
\dot{z} = Dz + \xi(t).
$$

Here $D$ — $p \times p$ - matrix, $\xi(t)$ — vector stochastic process as white noise.

Stochastic process modeling, as white noise transmission through linear system (2), matches to stochastic stationary process modeling with a known spectral energy density. Obviously matrix $D$ is stable, otherwise the solution of the equation (2) would be an unlimited stochastic process.

It is required to define control $u(x, z)$ minimizing following functional

$$
J(u) = \lim_{T \to \infty} \frac{1}{2T} \int_{0}^{T} (x'Qx + x'Pu + u'Ru)dt,
$$

where $x$ — $n$-dimensional vector of state variables; $u$ — $r$-dimensional control vector; $A$ — $(n \times n)$ - matrix; $B$ — $(n \times r)$ -matrix; $z$ — $p$-dimensional vector of stochastic disturbances; $G$ — $(n \times p)$ -matrix.

Suppose control constraint is not imposed. In addition, suppose that $z(t)$ is a persistent bounded stochastic function of time, which can be represented as the solution of a linear system of stochastic equations
where \( Q \) — nonnegative-defined constant \( n \times n \) - matrix; \( P \) — constant \( n \times r \) - matrix; \( R \) — positive-defined constant \( r \times r \) - matrix; \( M(.) \) — mathematical expectation. In addition, suppose matrices \( Q, P, R \) satisfy the requirement: matrix \( W \), consisting of these matrices,

\[
W = \begin{pmatrix}
Q & \frac{1}{2}P \\
\frac{1}{2}P' & R
\end{pmatrix},
\]

is positive-defined matrix.

Let us introduce expanded state vector \( y = (x, z)' \). Because of (1), (2) this vector satisfies the system of differential equations

\[
\dot{y} = \tilde{A}y + \tilde{B}u + f(t).
\]

Here

\[
\tilde{A} = \begin{pmatrix} A & G \\ 0 & D \end{pmatrix}, \quad \tilde{B} = \begin{pmatrix} B \\ 0 \end{pmatrix}, \quad f(t) = \begin{pmatrix} 0 \\ \xi(t) \end{pmatrix}.
\]

Functional (3) will become

\[
J(u) = \lim_{T \to \infty} \frac{1}{2T} M \int_0^T (y'\tilde{Q}y + y'\tilde{P}u + u'Ru)dt,
\]

where

\[
\tilde{Q} = \begin{pmatrix} Q & 0 \\ 0 & 0 \end{pmatrix}, \quad \tilde{P} = \begin{pmatrix} P & 0 \\ 0 & 0 \end{pmatrix}.
\]

It was shown that the solution of stochastic optimal control problem of system (4) on a finite interval of observation with the performance measure

\[
J(u) = \frac{1}{2} M \int_0^T (y'\tilde{Q}y + y'\tilde{P}u + u'Ru)dt,
\]

results in the following.

The presence of white noise \( \xi(t) \) in the system (4) does not change the solution of determinate problem of designing an optimal controller [2], except for increase in minimum value of performance measure. Thus, optimal control in problem with performance measure (6) is given by

\[
u^*(y, t, T) = -\left( R^{-1} \tilde{B}'K(t, T) + \frac{1}{2}R^{-1}\tilde{P}' \right)y,
\]
where $K(t, T)$ — solution of matrix differential equation

$$\dot{K} + K\dot{A} + \dot{A}'K - K\tilde{B}R^{-1}\tilde{B}'K - \frac{1}{2}K\tilde{B}R^{-1}\tilde{P}' - \frac{1}{2}\tilde{P}R^{-1}\tilde{B}'K - \frac{1}{4}\tilde{P}R^{-1}\tilde{P}' + \tilde{Q} = 0,$$

(7)

with a condition on the right end

$$K(T) = 0.$$

Based on research of solutions $K(t, T)$ of matrix differential equation (7) with a zero condition on the right end as $T \to \infty$ there were obtained conditions of existence of following limit

$$\lim_{T \to \infty} K(t, T) = \tilde{K},$$

where $\tilde{K}$ — constant symmetric positive-defined matrix and $\tilde{K}$ is a solution of matrix algebraic equation

$$\tilde{K}\dot{A} + \dot{A}'\tilde{K} - \tilde{K}\tilde{B}R^{-1}\tilde{B}'\tilde{K} - \frac{1}{2}\tilde{K}\tilde{B}R^{-1}\tilde{P}' - \frac{1}{2}\tilde{P}R^{-1}\tilde{B}'\tilde{K} - \frac{1}{4}\tilde{P}R^{-1}\tilde{P}' + \tilde{Q} = 0.$$

It was proved that optimal control in problem (4)-(5) is given by

$$u^*(y, t) = -\left( R^{-1}\tilde{B}'K + \frac{1}{2}R^{-1}\tilde{P}' \right)y.$$ 

The author was supported by the Russian Foundation for Basic Research (project no. 12-08-00309).

**References**


Optimal control of the sweeping process

Boris Mordukhovich

Wayne State University, Detroit, USA; aa1086@wayne.edu

We formulate and study an optimal control problem for the sweeping (Moreau) process, where control functions enter the moving sweeping set. To the best of our knowledge, this is the first study in the literature devoted to optimal control of the sweeping process. We first establish an existence theorem of optimal solutions and then derive necessary optimality conditions for this optimal control problem of a new type, where the dynamics is governed by discontinuous differential inclusions with variable right-hand sides. Our approach to necessary optimality conditions is based on the method of discrete approximations and advanced tools of variational analysis and generalized differentiation. The final results obtained are given in terms of the initial data of the controlled sweeping process and are illustrated by nontrivial examples.

This talk is based on the joint work with G. Colombo, R. Henrion, and N. Hoang

Optimal deformation during the creep

Evgenii Murashkin

Institute for Automation and Control Processes Far Eastern Branch Russian Academy of Sciences, Vladivostok, Russia; murashkin@dvo.ru

In the design of ship hulls and aircraft are widely used panels and profiles of hardly-deformed aluminum alloys at normal temperatures. Traditional methods of formation of structural elements of such alloys often leads to the appearance of the plastic breaks, cracks and other damage. In this regard, effective way is formation under high temperature and low deformation under creep conditions. The use of such process ensures the production of parts with high accuracy, which reduces the complexity of
assembly and welding, and also reduces the formation of effort to improve
the residual service life and quality of construction. At the same time it
is necessary to solve problems of modeling the process of formation, the
timing and strength of deformation regimes, the determination of the ge-
ometry of the light snap of an elastic response after removal of the loading
effort. In this case there is a necessity of calculation of such processes in
the model with finite irreversible deformations and complicated rheo-
logical properties of materials. Consideration be carried out in the model of
finite elastoplastic deformations [1, 3], the basic kinematic relations one
in the Cartesian system (Eulerian coordinates) can be written in the form
of
\[
\frac{De_{ij}}{Dt} = \varepsilon_{ij} - \gamma_{ij} - \frac{1}{2}\left((\varepsilon_{ik} - \gamma_{ik} + z_{ik})e_{kj} + e_{ik}(\gamma_{kj} - \varepsilon_{kj} - z_{kj})\right),
\]
\[
\frac{Dp_{ij}}{Dt} = \gamma_{ij} - p_{ik}\gamma_{kj} - \gamma_{ik}p_{kj},
\]
\[
\frac{Dn_{ij}}{Dt} = \frac{dn_{ij}}{dt} - r_{ik}n_{kj} + n_{ik}r_{kj}, \quad r_{ij} = w_{ij} + z_{ij}(\varepsilon_{ij}, \varepsilon_{ij}),
\]
\[
\varepsilon_{ij} = \frac{1}{2}(v_{i,j} + v_{j,i}), \quad w_{ij} = \frac{1}{2}(v_{i,j} - v_{j,i}),
\]
\[
d_{ij} = \frac{1}{2}(u_{i,j} + u_{j,i} - u_{i,k}u_{k,j}) =
\]
\[
e_{ij} + p_{ij} - \frac{1}{2}e_{ik}e_{kj} - e_{ik}p_{kj} - p_{ik}e_{kj} + e_{ik}p_{km}e_{mj}.
\]
Here \(u_i, v_i\) – components of displacement and velocity vectors; \(e_{ij}^c = e_{ij} - 0.5e_{ik}e_{kj}\) and \(p_{ij}\) – reversible and nonreversible parts of total Almansi
strains; \(D/Dt\) – an objective derivative; source \(\gamma_{ij}\) in the equation of
tensor transfer, is a tensor of velocities of nonreversible deformations, \(z_{ij}\)
– non-linear part of rotation tensor \(r_{ij}\), completely written out in [2],
defining its difference from a rigid rotation tensor \(w_{ij}\). This assumption
allows us to write the stress-strain relations like Murnagan’s relations in
theory of non-linear elasticity.
\[
\sigma_{ij} = -p\delta_{ij} + \frac{\partial W}{\partial e_{ik}}(\delta_{kj} - e_{kj}),
\]
\[
W = (\alpha - \mu)J^1 + \alpha J^2 + \beta J^2_1 - \kappa J^1_1 J^2_2 - \zeta J^2_1,
\]
where $W(J_1, J_2)$ is an elastic potential and $\alpha, \mu, \beta, \kappa, \zeta$ are elastic constants.

In the course of deforming, anticipating plastic flow, and in the course of unloading, source $\gamma_{ij}$ in the equation of tensor transfer (1) for tensor of nonreversible strains $p_{ij}$ is identified with a velocity of creep strains $\gamma_{ij} = \varepsilon_{ij}^v$ for which we accept the power law (Norton-Bailey law)

$$\varepsilon_{ij}^v = \frac{\partial V(\Sigma)}{\partial \sigma_{ij}}, \quad V(\Sigma) = B \Sigma^n (\sigma_{ij}),$$

$$\Sigma = \sqrt{\frac{3}{2}} \left( ((\sigma_1 - \sigma)^2 + (\sigma_2 - \sigma)^2 + (\sigma_3 - \sigma)^2) \right).$$

Here $\sigma_1, \sigma_2, \sigma_3$ are principal values of Cauchy stress and $B, n$ are creep parameters.

We formulate the optimal deformation problem in the creep: we are necessary to define a way of deformation of elasto-creeping material within a specified period of time to the final point of unloading to obtain the specified values of residual strains with minimal damage parameter. Thus the problem reduces to a multi-criteria optimization problem [4,5]

$$J_0 = w_1 \left( \int_V (A(\varepsilon, \sigma, \gamma) + A(\varepsilon^*, \sigma^*, \gamma))dV - \int_S p_i u_i dS \right) + w_2 \left( \int_V \sigma_{ij} \gamma_{ij} dV \right),$$

$$A(\varepsilon, \sigma, \gamma) = \frac{1}{2} (\sigma_{ij} \varepsilon_{ij} - \sigma_{ij} \gamma_{ij}).$$

Where $w_1, w_2$ are weights, $p_i$ are boundary loads, $\varepsilon^*, \sigma^*$ are residual strains and stresses. The proposed model can be applied to process optimization in terms of reduction of residual stress and to identify the mechanism of "healing" of microdefects at improving performance properties of the finished product.

The authors were supported by the Grant of the President of the Russian Federation (project no. MK-776.2012.1).

REFERENCES

Modeling and optimization of ion-beam etching process

L.A. Muravey\textsuperscript{1}, V.M. Petrov\textsuperscript{1}, A.M. Romanenkov\textsuperscript{1}

\textsuperscript{1} MATI–RSTU, Moscow, Russia; pm@mati.ru

**Introduction.** The problem of optimal controlling the ion-beam etching (IBE) process with the purpose of minimizing the geometrical sizes of the elements being etched is solved by changing the angle of incidence of the ion beam with respect to the target. The evolution of the surface having an arbitrary form in the ion bombardment process is described by the equation:

\[ z_t(t,x) + \nu(\theta)\sqrt{1 + z_x^2(t,x)} = 0, \quad (1) \]

where \( z(t,x) \) is the height of the etched form at the time moment \( t \) in \( x \) position, \( \nu(\theta) \) is the material etching velocity, depending on the angle

![Fig. 1](image1.png)  
![Fig. 2](image2.png)
θ, formed by the angle of incidence and the normal to the surface being etched. In particular, if the ion beam is perpendicular to axis \( x \), then 
\[
\theta = \arctg \frac{\partial z}{\partial x}(t, x).
\]

The specific feature of all existing resists is the non-monotone nature of function \( \nu(\theta) \), i.e. the existence of a certain \( \theta = \theta^* \), at which \( \nu(\theta) \) is maximum. As a rule, the function \( \nu(\theta) \) is determined experimentally and normally has the form illustrated in Fig. 1. The direction of the ion beam can vary with time, then the angle \( \theta \) can be presented in the following form:
\[
\theta = \arctg \frac{\partial z}{\partial x} + \alpha(t),
\]
where \( \alpha(t) \) is the angle between the direction of the ion beam incidence and the axis \( y \). \( \alpha(t) \) will be considered as the control, upon which the natural restrictions have been imposed (see Fig. 2).
\[
0 \leq \alpha(t) \leq \alpha_{\text{max}}.
\]

**IBE equation derivation.** The ion-beam etching equation will be obtained in the assumption that during the etching process the sample is rotating with the angle velocity \( \omega \) with respect to the vertical straight line, passing through the semicircle centre. It should be noted that the two-dimensional and three-dimensional cases differ only by form for \( \theta \) angle, and the evolution process of the being etched profile is described by the same equation. Also, it is possible to show, that when \( \omega = 0 \), then we again obtain the two-dimensional case.

Let \( z = z(x, t) \) to be a smooth function, describing the etched surface evolution, \( P_t = \{x, z(x, t)\} \) is the etching profile. Let’s consider the point \( Q \in P_t \), we will draw the \( \bar{n}_Q \) normal vector from the given point:
\[
\bar{n}_Q = \frac{1}{\sqrt{1 + z_x^2(x, t)}} \{-z_x(x, t), 0, 1\}. \quad (4)
\]

As the mask rotates with the angular velocity \( \omega \), then (see Fig. 3) we have
\[
\gamma = \{\sin \alpha \sin \omega t, \sin \alpha \cos \omega t, \cos \alpha\}
\]

Note, that \( |\bar{n}_Q| = |\gamma| = 1 \), therefore (see Fig. 2) \( (\bar{n}_Q, \gamma) = \cos \theta \). That is
\[
\cos \theta = \frac{\cos \alpha - z_x(x, t) \sin \alpha \sin \omega t}{\sqrt{1 + z_x^2(x, t)}}. \quad (5)
\]
Through the time period $\Delta t$ the point $Q$ will move to the point $M(x_1, z_1)$ on the profile $P_{t+\Delta t}$. This movement will be executed along the vector $-\overrightarrow{\mathbf{n}}_Q$ with $\nu_Q$ velocity. Let’s consider the vector $\overrightarrow{QM} = \{x_1 - x, z_1 - z\}$. Based on the etching physical law it can be stated that the vector $\overrightarrow{QM}$ can be determined by the following formula:

$$\overrightarrow{QM} = -\nu(\theta) \Delta t \overrightarrow{\mathbf{n}}_Q = -\Delta t \frac{\nu(\theta)}{\sqrt{1 + z_x^2(x, t)}} \{-z_x(x, t), 0, 1\}.$$ 

Using the theorem of the finite increments we will obtain:

$$-\Delta t \frac{\nu(\theta)}{\sqrt{1 + z_x^2(x, t)}} = \Delta t \frac{\nu(\theta)}{\sqrt{1 + z_x^2(x, t)}} z_x^2(x, t) + z_t(x, t) \Delta t + o(\Delta t).$$

After division by $\Delta t$, and directing it to 0 and after the elementary transformations we have equation (1).

**On the evolution of the smooth convex profile. Effect of fracture appearance.** Note that if the mask initial form is smooth and convex, then in its evolution process the appearance of the angular points is possible. Let’s consider this effect in more detail.

Let us consider a simple example $\alpha = 0$, $\overrightarrow{\mathbf{\gamma}} \parallel Oz$ and the initial profile in the form of a semicircle

$$z \big|_{t=0} = \sqrt{1 - x^2}, |x| \leq 1,$$  \hspace{1cm} (6)

where $\overrightarrow{\mathbf{\gamma}}$-ion beam direction. In this case the solution of problem (7) can be represented in the analytical form. Let us give the unit normal vector profile $S_t$ in the form $\psi = \{\sin \theta, \cos \theta, |\theta| \leq \pi/2\}$ and the initial profile $S_0$ in the form $S_0 = \{(x, z) : x = \sin \theta, z = \cos \theta, |\theta| \leq \pi/2\}$ and introduce the functions

$$F'(\theta) = v(\theta) \sin \theta + v'(\theta) \cos \theta, \quad G'(\theta) = v(\theta) \cos \theta - v'(\theta) \sin \theta.$$  \hspace{1cm} (7)

**Theorem 1.** Then there exists the time $T > 0$, for which on the interval $0 < t < T$ the profile $S_t$ will remain smooth convex and will be represented as

$$S_t = \{(x, z) : x(\theta, t) = \sin \theta - tF(\theta), z(\theta, t) = \cos \theta - tG(\theta), |\theta| \leq \pi/2\}.$$  \hspace{1cm} (8)
So, the time $T$ may be found from relations

$$\begin{align*}
\frac{\partial x}{\partial \theta} &= \cos \theta - t \cos \theta (v(\theta) + v''(\theta)) = 0, \\
\frac{\partial z}{\partial \theta} &= -\sin \theta + t \sin \theta (v(\theta) + v''(\theta)) = 0,
\end{align*}$$

i.e. $1/T = \max R(\theta), \ |\theta| \leq \frac{\Pi}{2}$, where $R(\theta) = \nu(\theta) + \nu''(\theta)$. In our test example $\nu(\theta) = (1 + 2 \sin^2 \theta) \cos \theta$ we have $R(\theta) = 4 \cos \theta (4 \cos^2 \theta - 3)$ and $\max R(\theta) = R(\Pi/3) = 4$, hence $T = 1/4$ (see Fig. 4).

Consider the case of non-smooth convex profile.

$$C(S_t, \psi) = \max(x \psi_1 + z \psi_2), \ (x, z) \in S_t. \quad (9)$$

Let us consider the gradient

$$C'(S_0, \psi) = \left[ \frac{\partial C}{\partial \psi_1}(S_0, \psi), \frac{\partial C}{\partial \psi_2}(S_0, \psi) \right] \quad (10)$$

and give $[u(\theta), v(\theta)] = C'(S_0, \psi)|_{\psi=(\sin \theta, \cos \theta)}$.

Every non-smooth initial profile can be approximated be means of a smooth one according to the following procedure. Let us consider the rectangular profile $\Pi = \{ |x| \leq a, |y| \leq b \}$ with the support function $C(\Pi, \psi) = a|\psi_1| + b|\psi_2|$. The smoothing profile $\Pi_\mu$ is determined by means of the support function

$$C(\Pi_\mu, \psi) = \sqrt{a^2 \psi_1^2 + \mu^2 \psi_2^2} + \sqrt{\mu^2 \psi_1^2 + b^2 \psi_2^2}.$$
where $\mu > 0$ is a small parameter. Then due to (9), (10) we have

$$S_{\mu,t} = \{(x,z): x = u_{\mu}(\theta) - tF(\theta), z = v_{\mu}(\theta) - tG(\theta), |\theta| \leq \pi/2\}$$

(see Fig. 5).

One of the methods to avoid distortion of the geometrical sizes (see Fig. 6) is the possibility to change the ion beam direction in the etching process, i.e. of the angle change in time. The equation for this velocity of $x_0(t)$ (see Fig. 6) can be easily derived from the geometrical considerations. Namely

$$\frac{dx_0}{dt} = \nu (\alpha + \arctg \frac{\partial z}{\partial x} \bigg|_{z=0}) \sqrt{1 + \left(\frac{\partial z}{\partial x}\right)^2 \bigg|_{z=0}}$$

(11)

So, the optimal control problem consists in the following. To determine the functions $z(t,x)$ and $a(t)$, satisfying the equations (1), (11) the restrictions (3), the initial conditions

$$z(0,x) = g(x),$$

(12)

where $g(x)$ is the convex function, and the condition

$$z(T,1) = -H.$$  

(13)

In this connection it is necessary that at the finite moment of time $t = T$ the function $x_0(t)$ should take the maximum value, i.e.

$$J = x_0(T) \rightarrow \text{max}.$$  

(14)

Note that the formulated problem is the problem with the non-fixed time $T$. This creates some additional difficulties, because in addition to
the functions $\alpha(t)$ and $z(t, x)$ also, it is necessary to look for the process termination moment $T$. In our case this difficulty can be easily avoided. For this the function $h(t)$ – the depth of the working layer current etching – will be introduced. This function is monotone and varies within $0 \leq \tau = h(t) \leq H$ limits. Then the equations (1) and (11) will have the following form

$$\frac{\partial z(\tau, x)}{\partial \tau} + \frac{v(\theta + \alpha)}{v(\alpha)} \sqrt{1 + \left(\frac{\partial z}{\partial x}(\tau, x)\right)^2} = 0, \quad (15)$$

$$\frac{dx_0}{d\tau} = \frac{v(\theta + \alpha)}{v(\alpha)} \frac{\sqrt{1 + \left(\frac{\partial z}{\partial x}\right)^2|_{z=0}}}{(\frac{\partial z}{\partial x})|_{z=0}}, \quad (16)$$

with the conditions (12) and

$$x_0(0) = 1, \quad 0 \leq x \leq 1, \quad 0 \leq \tau \leq H. \quad (17)$$

So, our problem of the optimal control has been reduced to the search for the functions $\alpha(\tau)$ and $z(\tau, x)$ satisfying the equations (15), (16), conditions (17) and providing the maximum (14) with restriction (3).

**Solution method.** If we consider numerical solutions $z(\tau_j, x_i)$ and $x_0(\tau_j), x_i = i/m, i = 0, m, \tau_j = Hj/n, j = 0, n$ of problem (15)–(17), then due to Pontryagin maximum principle for every $\tau_j$ we can see the optimal value $\hat{\alpha}_j$ of control function $\alpha_j$ as $\min_k v(\theta_j + \alpha_k)/v(\alpha_k), 1 \leq k \leq n$, where

$$\theta_j = \arctg \frac{\partial z}{\partial x_i}|_{(z=0, x_0(\tau_j)=i/m)}.$$
Special algorithm for Three-Stations Railway problem

Elena Musatova¹, Alexander Lazarev², Nail Husnullin³

¹ Institute of Control Sciences RAS, Moscow, Russia; nekolyap@mail.ru
² Institute of Control Sciences RAS, Moscow State University, National Research University – Higher School of Economics, Moscow, Russia; jobmath@mail.ru
³ Institute of Control Sciences RAS, Moscow, Russia; nhusnullin@gmail.com

Problem statement

The paper is devoted to the problem of railway transportation. The railway network consists of stations between which freight cars are transported. Let us $S$ be a set of stations. Every station $s$ has a set $N^s = \{J_1, \ldots, J_{n^s}\}$ of orders to deliver. $N = \cup_{s \in S} N^s$ is a common set of orders. Each order represents one freight car. If an order consists of $k$ cars, we will consider it as $k$ different orders. We know realise time $r^s_j$ and due date $d^s_j$ of delivery for each car $J_j \in N^s$. Let us $p^s_j$ be traversing time for the car $J_j \in N^s$ and $w^s_j$ be its weight (importance). Our goal is to design freight trains and work out their schedule. Objective functions can be the following:

- minimizing the weight total tardiness
  \[ \min \sum_{s \in S} \sum_{j \in N^s} w^s_j \max\{0, C^s_j - d^s_j\}; \]

- minimizing the total completion time
  \[ \min \sum_{s \in S} \sum_{j \in N^s} C^s_j; \]

- minimizing the maximal lateness
  \[ \min \max_{j \in N^s, s \in S} \{C^s_j - d^s_j\}; \]

- in the set $N$ find subset $\overline{N} \subseteq N$ of cars that can be delivered on time:
  \[ \max \sum_{j \in \overline{N}} w^s_j - \sum_{j \in N \setminus \overline{N}} z^s_j. \]
These problems are large-scale and difficult to solve. Therefore, we proposed to divide them into subproblems which are easier to solve and consider special cases which would help to find important structural properties which are hard to recognize in the general case. We have suggested a number of railway basic models (with two stations, with chain of stations and so on) that gives us an opportunity to develop special exact algorithms which can be used in general railway problems. Some algorithms for two-stations railway problems can be found in [1]. In this paper we propose an algorithm for the special case of three stations.

Three-Stations Railway problem

Consider the problem with 3 stations that are connected by a railroad and one locomotive. A valid arrangement is shown on fig. 1. Arrows indicate a possible route of the locomotive from one station to another.

![Fig. 1 Possible arrangement of stations](image)

We have to implement a set of orders. $N^{ij}$ is a set of orders that should be delivered from the station $i$ to the station $j$. So $N^1 = N^{12} \cup N^{13}$ etc. Let us assume that $p$ is a traversing time from one station to another, $q$ is a capacity of a train, $r^s_i$ is a release time of $J^s_i \in N^s$, $s \in \{1, 2, 3\}$, $n_s$ is a common number of orders at the station $s$, $n_{si}$ is a number of orders at station $S$ that are should be delivered to the station $i$.

Objective function of the problem is following:

$$
\min \sum_{J^1_i \in N^1} C^1_i + \sum_{J^2_i \in N^2} C^2_i + \sum_{J^3_i \in N^3} C^3_i. \tag{1}
$$

It is easy to see the locomotive has the following strategies when he arrives to a station:

1. staying at the station and waiting a new order;
2. idling to the next station;
3. idling to the previous station;
4. moving to the next station with the largest possible number of cars;
5. moving to the previous station with the largest possible number of cars.

Only for this objective function (1) we can reduce the number of possible solutions if we consider the following:

- idling to the next or previous station is the same as moving without the wagons;
- if we can move at the second station with \( q \) wagons it is more preferable for us than stay at the station;
- idling to another station is preferable, if all orders have been delivered from station \( s \).

It is obviously that in an optimal schedule the train begins his movement from a station \( s \) only at the moments of its arrival to this station or at the moment of appearance of a new order, i.e. at the moment \( r^s_j \). So times points at which the train begins and ends movement between stations belong to \( T = \{ t : \exists r^s_j, \exists l \in \{1, \ldots, (n_1 + n_2 + n_3)\}, t = r^s_j + lp \} \).

Let us denote by

\[
S(s, t, k_{12}, k_{13}, k_{21}, k_{23}, k_{31}, k_{32})
\]

the state at the moment \( t \in T \), where \( s \) is the number of the station where the locomotive is, \( k_{12} \) is the number of delivered orders from the first to the second station, \( k_{23} \) the number of delivered orders from the second station to the third one, etc. Let us assume that \( P(s, t, k_{12}, k_{13}, k_{21}, k_{23}, k_{31}, k_{32}) \) is the smallest total delivery time in the scheduling which leads to state \( S(s, t, k_{12}, k_{13}, k_{21}, k_{23}, k_{31}, k_{32}) \). For the objective function (1) the optimal solution of the problem is

\[
\min_{s, t} P(s, t, n_{12}, n_{13}, n_{21}, n_{23}, n_{31}, n_{32}).
\]

Example 1. Consider the following problem. Let us assume that \( r^{12} = (1, 2, 3), r^{23} = (2, 3, 4), r^{13} = (3, 4, 5), r^{21} = (2, 3), r^{32} = (1, 2), \)}
\( r^{31} = (2, 3), \ p = 2, \ q = 2 \). One of the possible scheduling solutions is shown on fig. 2.

![Fig. 2 One of the possible scheduling solutions](image)

Algorithm 1 describes the calculation of the number of cars that the train can take to the next station in (2).

```
Algorithm 1
1: function GetPossibleOrders(s, n, futureS)
2:     futureQ ← 0
3:     while \( r_s^i < t \) do
4:         if \( i > n \) \&\& \( i - n \leq q \) then futureQ ← futureQ + 1
5:     end if
6:     end while
7:     return futureQ
8: end function
```

Let us introduce the following denotations:

- \( t \) — current time;
- \( s \) — station number, where the locomotive is located at the time \( t \);
- \( j \) — number of wagons, which the locomotive can take at the current time;
- \( n \) — number of delivered orders from station \( s \);
- \( q \) — max number of wagons that the locomotive can carry at a time;
- \( N[s][futureS] \) — array, which contain number of delivered orders from station \( s \) to \( futureS \);
- \( Runner \) — entry point, which execution of the program begins with;
- \( existsCarsOnStay \) — function, which returns false, if all orders have been delivered from station \( s \).
Algorithm 2 creates nodes of the tree and allows to move from one station to another.

```
Algorithm 2
1: function Runner
2:   int[,] N ← new int[s,s]
3:   N ← 0
4:   newS ← S(1,0,N)
5:   BuildTree(newS)
6: end function
7:
8: function BuildTree(prevS)
9:   j ← 1
10:  q ← 0
11:  s ← prevS.s
12:  t ← prevS.t
13:  N ← prevS.N
14:  while j..3 do
15:     if j = 1 then futureS ← s ⊕ 1
16:     end if
17:     if j = 2 then futureS ← s ⊖ 1
18:     end if
19:     if j = 3 then futureS ← s
20:     end if
21:     q ← GetPossibleOrders(s, N[s, futureS])
22:     N[s, futureS] ← N[s, futureS] + q
23:     if j <> 3 then
24:        t ← t + p
25:     else
26:        t ← t + 1
27:     end if
28:     newS ← S(futureS, t, N)
29:     if existsCarsOnStay(newS) then
30:        BuildTree(newS)
31:     end if
32:  end while
33: end function
```

As we have $O((n_1 + n_2 + n_3)^2)$ possible time moments running time of the proposed algorithm is $O((n_1 + n_2 + n_3)^2n_{12}n_{13}n_{21}n_{23}n_{31}n_{32})$.

The authors were supported by the Russian Foundation for Basic Research (project no. 11-08-13121).

REFERENCES


191
Numerical solution of two-dimensional inverse problem for the Helmgoltz equation

Daniar Nurseitov\textsuperscript{1}, Maksim Shishlenin\textsuperscript{2}, Syrym Kasenov\textsuperscript{3}

\textsuperscript{1} K.Satpaev KazNTU Information and Space Technologies National Scientific Laboratory, Almaty, Kazakhstan; ndb80@mail.ru
\textsuperscript{2} Sobolev Institute of Mathematics of the Siberian Branch of the Russian Academy of Sciences, Novosibirsk, Russia; mshishlenin@ngs.ru
\textsuperscript{3} Institute for Master and PhD programs Kazakh National Pedagogical University named after Abai, Almaty, Kazakhstan; syrym.kasenov@mail.ru

We consider the Helmholtz equation in $\Omega = (-b, b) \times (0, L)$ initial-bondary problem:

$$u_{zz} + u_{yy} - r^2(y, z)u = 0 \quad (y, z) \in \Omega \quad (1)$$

$$u_y(-b, z) = 0, \quad u_y(b, z) = 0, \quad z \in [0, L] \quad (2)$$

$$u_z(y, 0) = g(\omega, y), \quad u_z(y, L) = 0, \quad y \in [-b, b] \quad (3)$$

here $r^2(y, z) = \frac{\omega^2}{v^2(y, z)}$ $v$—the velocity of wave propagation.

Inverse Problem: find the coefficients $v(y, z)$ of equation (1) using additional information about the solution to the initial-boundary problem (1)-(3):

$$u(y, 0, \omega) = f(y, \omega), \quad y \in [-b, b] \quad (4)$$

This problem (1)-(4) is ill-posed. We write the problem (1)-(4) in the operator form $Au = f$. We solve numerically the equation $Au = f$, we minimize the objective functional

$$J(v) = \langle Au - f, Au - f \rangle = \sum_{\omega} \int_{-b}^{b} [u(y, 0, \omega; v) - f(y, \omega)]^2 dy \quad (5)$$

We minimize the quadratic functional (5) using the steepest descent method [1]:

$$v^{n+1} = v^n - \alpha_n J'(v^n) \quad n = 0, 1, \ldots, \quad (6)$$

where $v^0$ is an initial approximation, $J'(v^n)$ is the gradient of the objective functional, and the descent parameter $\alpha_n$ is determined from the condition
\[ \alpha_n = \arg \min_{\alpha} J(\nu^n - \alpha J'(\nu^n)). \]

Let us consider adjoint problem:

\[ \psi_{zz} + \psi_{yy} - r^2(y, z)\psi = 0 \quad (y, z) \in \Omega \]
\[ \psi_y(-b, z) = 0, \quad \psi_y(b, z) = 0, \quad z \in [0, L] \]
\[ \psi_z(y, 0) = 2(u(y, 0, \omega) - f(y, \omega)), \quad \psi_z(y, L) = 0, \quad y \in [-b, b] \]

Gradient of functional:

\[ J(\nu) = \sum_{\omega} \frac{\omega^2}{v^4(y, z)} u(y, z, \omega) \psi(y, z, \omega) \]

**Algorithm for solving the inverse problem** is

1. Choose an initial approximation \( \nu^0 \);
2. Solve the direct problem (1) (3) with we get \( u(y, 0, \omega) \);
3. We calculate the boundary condition of adjoint problem (9);
4. Solve the adjoint problem (11)-(13) and we get \( \psi(y, z, \omega) \);
5. Find the gradient by the formula (10);
6. We calculate the \( \nu^1 \) by the formula (6);
7. We verify value of the functional(5), when he reached the minimum, then the problem is solved, otherwise the return to step 1. using the \( \nu^0 = \nu^1 \).

**References**

2. *S.I. Kabanikhin, M.A. Bektemesov, A.T. Nurseitova* Iterative methods for solving Inverse and Ill-Posed problems with the data on the part of the boundary, PF International Fund for inverse problems, Almaty, Novosibirsk, 2006, pp.315. (in Russian)
Production model in the conditions of unstable demand

Nataliya Obrosova¹, Alexander Shananin²

¹ CC RAS, Moscow, Russia; nobrosova@ya.ru
² MIPT, Moscow, Russia, CC RAS, Moscow, Russia; alexshan@ya.ru

Russia’s accession to the World Trade Organization aggravated problems of product competitiveness of the Russian production as within the country, and in the world market. Thus industries of the Russian economy appear in various conditions. Domestic raw materials corporations were included long ago into the world energy market. Products of an oil and gas complex since the Soviet period are competitive. More difficult is a situation in processing industries. Historically production capacities of this sector were created in the conditions of the closed economy and absence of the competition to the import goods. Up to now there is a technological backwardness of processing sector remained. Products of sector lose in the competition to better import analogs. As a result the producer has delays with sales of products and current assets deficit which becomes covered or at the expense of bank loans, or at the expense of the state grants. The inefficiency of processing sector essentially influences economic indicators of production, and also economy indicators as a whole. The program implementation of upgrade of processing sector shall be carried out after the detailed analysis of their economic consequences. Such analysis shall be carried out taking into account feedback and potential influence of change of indicators of an industry on state of the economy as a whole. The adequate tool for carrying out such researches are the mathematical models of economy constructed on the basis of a system approach [1] to the analysis of economic events and allowing analyzing a consequence of large economic decisions taking into account their indirect consequences.

In 2005-2007 based on a system approach the model of economy of Russia intended for the purposes of short-term and mid-term forecasting [2] was developed. In the model the description of activities of processing sector taking into account the current assets deficit developed in [3] for the first time was used. The calculations for model showed [2] that without an inefficiency of processing sector growth rate of economy appears over-
estimated by 2-3 %. Thus features of processing sector need additional
detailed research on the basis of adequate mathematical models. Results
of this research are provided in the report.
Let’s assume that demand for a made product is unstable. Sale of
a product occurs during the random moments of time forming a Poisson
process with parameter \( \lambda \). As a result the producer is forced to accumulate
some amount of products in a warehouse in hope of sale. If sale doesn’t
come the producer has a current assets deficit which it is possible to cover
at the expense of the short-term credit line \( K(t) \) under percent \( r \).
Let’s designate \( \tau \) the maximum time which is profitable to the pro-
ducer to use the credit line \( K(t) \) in the conditions of absence of sale. Let
\( Y^* \)- restriction of trade infrastructure which is understood as the greatest
possible consignments. Let’s designate: \( Y_0 \) - the current product stock in
a warehouse of the producer, \( \eta \) - production capacity, \( y \) - cost value of a
product, \( p \) - the product price. Then
\[
K(t) = \begin{cases} y\eta, & 0 \leq t \leq \tau \\ 0, & t > \tau \end{cases},
\]
the loan debt \( L(t) \) changes owing to the equation
\[
\left\{ \begin{array}{l}
\frac{dL(t)}{dt} = K(t) + rL(t) \\
L(0) = 0
\end{array} \right., \text{i.e. } L(t) = \frac{y\eta}{r}(e^{rt} - e^{r(t-\tau)+}).
\]
The product output by the time \( t \) is
\[
Y(t) = \begin{cases} Y_0 + \eta t, & 0 \leq t \leq \tau \\
Y_0 + \eta \tau, & t > \tau \end{cases}.
\]
The purpose of the owner of manufacture is maximization of the in-
come \( W(Y_0) \) discounted with coefficient \( \Delta \geq r \) on the unrestricted horizon
choosing time \( \tau_0 \) of credit using. The discounted income of the owner of
manufacture \( W(Y_0) \) is the solution of the following Bellman equation
\[
W(Y_0) = \max_{\tau \geq 0} \int_0^\infty \lambda e^{-(\lambda+\Delta)t} [p \min(Y(t),Y^*) - L(t) + W((Y(t) - Y^*)_+)] dt.
\]
From the economic point of view \( W(Y_0) \) characterizes firm cost in case of
\( Y_0 \) a product stock.
Theorem 1. The solution $W(Y_0)$ of the equation (4) exists and is unique in a class $G[0, +\infty)$ of the continuous, non-negative, not decreasing, concave functions limited together with the derivative on a semi-interval $[0, +\infty)$, i.e.

$$G[0, +\infty) = \begin{cases} w(x) | x \in [0, +\infty), w \in C[0, +\infty), 0 \leq w(x) \leq \frac{\lambda}{\Delta} pY^*, \\ 0 \leq \frac{dw}{dx} \leq p, \ w(\alpha x + (1-\alpha)y) \geq \alpha w(x) + (1-\alpha)w(y) \end{cases} \bigg\| \forall x, y \in [0, +\infty), \alpha \in [0, 1].$$

If the first derivative in some point for function $w$ isn’t determined (owing to monotony the first derivative can have only discontinuity of the first kind and a set of discontinuity points is not more than countable), as a derivative we will understand a function derivative on the left in all points, except 0, and in 0 - a function derivative on the right.

It is proved that the solution of the equation (4) satisfies the condition

$$W(0) \leq \frac{\eta}{\Delta} \frac{\lambda}{\lambda + \Delta} \left(p - \frac{y(\lambda + \Delta)}{\lambda + \Delta - r}\right)_+.$$

We suppose everywhere further that the condition of profitability of production $p - \frac{y(\lambda + \Delta)}{\lambda + \Delta - r} > 0$ is fulfilled.

The equation (4) was reduced to the integral equation that allowed to prove the following theorem.

Theorem 2. The optimal period for using credit $\tau_0 = \left(\tau_2^0 + \frac{Y^*-Y_0}{\eta}\right)_+$, where $\tau_2^0 = \arg \max_{\tau \geq 0} \int_0^\tau \left[W'(\eta t) - \frac{y(\lambda + \Delta)}{\lambda + \Delta - r}\right]_+ e^{-(\lambda + \Delta)t} dt$.

Denote

$$\varsigma_0 = \frac{\tau_2^0 \eta}{Y^*}, \ Y_0 = \varsigma Y^*, \ \chi = (\lambda + \Delta) \frac{Y^*}{\eta}, \ \beta = \frac{\lambda + \Delta}{\lambda + \Delta} \frac{p - \frac{y(\lambda + \Delta)}{\lambda + \Delta - r}}{\lambda + \Delta (p - \frac{y(\lambda + \Delta)}{\lambda + \Delta - r})},$$

$$\Psi_0 = \left(\frac{\Delta}{\eta} W(0) - \frac{\lambda}{\lambda + \Delta} \left(p - \frac{y(\lambda + \Delta)}{\lambda + \Delta - r}\right)\right) \left(\frac{\lambda}{\lambda + \Delta} \left(p - \frac{y(\lambda + \Delta)}{\lambda + \Delta - r}\right)\right)^{-1}.$$

Parameter $0 < \beta \leq 1$ is connected to the profitability of production. In force of restriction on $W(0)$ the value $\Psi_0 \in [-1, 0]$.
Theorem 3. The function $W(Y_0) \in G[0, +\infty)$ is the solution of the equation (4) if and only if the function

$$
\hat{\Phi}(\varsigma) = \left( \frac{\lambda}{\lambda + \Delta} \left( p - \frac{y(\lambda + \Delta)}{\lambda + \Delta - r} \right) \right)^{-1} e^{-(\lambda + \Delta)\varsigma} \left[ W'(\varsigma Y^*) - \frac{y(\lambda + \Delta)}{(\lambda + \Delta - r)} \right]
$$

(5)

is the solution of the system

$$
\hat{\Phi}(\varsigma) = \begin{cases} 
\beta e^{-\varsigma} + \Psi_0, & \text{if } 0 \leq \varsigma \leq 1, \\
\Psi_0 + e^{-\varsigma} - (1 - \beta) e^{-\varsigma} - \frac{\lambda}{\lambda + \Delta} \chi e^{-\varsigma} \int_0^{\varsigma - 1} \Phi_+(\xi) d\xi + \frac{\lambda}{\lambda + \Delta} \Phi_-(\varsigma - 1), & \text{if } \varsigma > 1,
\end{cases}
$$

(6)

where $\varsigma_0, \Psi_0$ is the solution of the system

$$
\begin{cases} 
\hat{\Phi}(\varsigma_0) = 0, \\
\Psi_0 + e^{-\varsigma_0} = \frac{\lambda}{\lambda + \Delta} \chi e^{-\varsigma_0} \int_0^{\varsigma_0} \Phi(\xi) d\xi.
\end{cases}
$$

(7)

If the system (7) doesn’t have solution, then $\varsigma_0 = 0$, $\Psi_0 = -e^{-\xi}$.

Corollary 1. Let $W(Y_0) \in G[0, +\infty)$ is the solution of (4). Then the following statements are true.

1. If profitability condition taking into account sales volume restriction $\beta > e^{-\chi}$ is true, then

$$
\tau_2^0 = \arg \max_{\tau \geq 0} \int_0^\tau \left[ W'(\eta t) - \frac{y(\lambda + \Delta)}{\lambda + \Delta - r} \right] e^{-(\lambda + \Delta)t} dt > 0.
$$

2. If $\beta \leq e^{-\chi}$, then $\tau_2^0 = 0$, i.e. manufacture works only before achievement of a stock $Y^*$.

By the help of steps method the solution of the system (6) is obtained in an explicit form [4].

Corollary 2. For $0 < \varsigma_0 \leq 1$ it is necessary and sufficient that $e^{-\chi} < \beta \leq \beta_1$, where $\beta_1 = \left( 1 + \frac{\lambda}{\lambda + \Delta} - \frac{\lambda}{\lambda + \Delta} (1 + \chi) e^{-\chi} \right)^{-1}$.

The stock of product in the warehouse at the moment of time $t$ is $\varsigma(t)Y^*$. Change of value $\varsigma(t)$ is a random process. The analysis of the process allowed finding average loading of manufacture.
Theorem 4. In the case of $\varsigma_0 \in [0,1)$ average loading of manufacture is $u_0 = 1 - \exp(-\lambda_0 Y^*/\eta^*)$.

The developed model of production can be used for the description of features of functioning of processing sector in model of Russia economy.

The authors were supported by the Russian Foundation for Basic Research (projects N11-07-00162, 11-01-12084-ofi-m, PFI OMN RAN N3 (project N3.14), PFI of Presidium RAN N15 (project N106).

REFERENCES


Global search in bilinear separation problems

Andrei Orlov\textsuperscript{1}, Sergei Pinigin\textsuperscript{2}

\textsuperscript{1} Institute for System Dynamics and Control Theory, Siberian Branch of the Russian Academy of Sciences, Irkutsk, Russia; anor@icc.ru

\textsuperscript{2} Institute for System Dynamics and Control Theory, Siberian Branch of the Russian Academy of Sciences, Irkutsk, Russia; grey@irkutsk.ru

1. Introduction. The work is devoted to one of the problems of the so-called generalized separability [1, 2]. This problem can be arisen in many practical areas, where we need procedures for classifying objects.
The practical applications concern, for example, technical problems (diagnostic materials and nanomaterials), economic problems (banks and credits), medical problems (diagnosis and prognosis), and so on.

The purpose of this procedure is the classification of objects by investigation of available statistic data to certain classes by constructing a special rule (discriminant function). The simplest discriminant function is linear function [1]. But, as known, sets are linearly separable if and only if the intersection of their convex envelopes is empty. So, the practical demands are not limited by the linear separability only, and one often needs a more general concept of separability, e.g. bilinear separability [2].

2. Problem formulation and its reduction. Consider two non-empty finite sets \( \mathcal{A} \) and \( \mathcal{B} \) from the space \( \mathbb{R}^n \). These sets include \( m \) and \( k \) points, respectively, so that \( \mathcal{A} \) and \( \mathcal{B} \) can be represented by the matrices \( \mathbf{A} \in \mathbb{R}^{m \times n} \) and \( \mathbf{B} \in \mathbb{R}^{k \times n} \). The cells of these matrices are the coordinates of points from \( \mathcal{A} \) and \( \mathcal{B} \), respectively.

The problem of bilinear separation is to find two hyperplanes
\[
H_1(\omega^1, \gamma^1) = \{ x \in \mathbb{R}^n \mid \langle \omega^1, x \rangle = \gamma^1, \omega^1 \in \mathbb{R}^n, \gamma^1 \in \mathbb{R} \},
\]
and
\[
H_2(\omega^2, \gamma^2) = \{ x \in \mathbb{R}^n \mid \langle \omega^2, x \rangle = \gamma^2, \omega^2 \in \mathbb{R}^n, \gamma^2 \in \mathbb{R} \},
\]
separating given sets \( \mathcal{A} \) and \( \mathcal{B} \) in the following sense.

Definition 1. The sets \( \mathcal{A} \) and \( \mathcal{B} \) are said to be bilinearly separable by two hyperplanes \( H_1(\omega^1, \gamma^1) \) and \( H_2(\omega^2, \gamma^2) \) if for each point \( A^i \) \( (i = 1, m) \) of the set \( \mathcal{A} \), and for each point \( B^j \) \( (j = 1, k) \) of the set \( \mathcal{B} \) one of the following systems of inequalities holds:

\[
\begin{align*}
\langle A^i, \omega^1 \rangle > \gamma^1, \\
\langle A^i, \omega^2 \rangle > \gamma^2, \\
\langle B^j, \omega^1 \rangle < \gamma^1 \text{ or } \langle B^j, \omega^2 \rangle < \gamma^2; \tag{1}
\end{align*}
\]

\[
\begin{align*}
\langle B^j, \omega^1 \rangle > \gamma^1, \\
\langle B^j, \omega^2 \rangle > \gamma^2, \\
\langle A^i, \omega^1 \rangle < \gamma^1 \text{ or } \langle A^i, \omega^2 \rangle < \gamma^2; \tag{2}
\end{align*}
\]

\[
\begin{align*}
\langle A^i, \omega^1 \rangle > \gamma^1, \\
\langle A^i, \omega^2 \rangle > \gamma^2, \\
\langle B^j, \omega^1 \rangle < \gamma^1 \text{ or } \langle B^j, \omega^2 \rangle < \gamma^2; \tag{3}
\end{align*}
\]

It can be readily seen that the cases (1) and (2) are equivalent if the roles of \( \mathcal{A} \) and \( \mathcal{B} \) are interchanged.

199
The problem of bilinear separation in the sense of the system (1) can be reduced to the following equivalent nonconvex bilinear optimization problem (non-symmetrical case) [2]:

\[
F(z^1, z^2) = \langle z^1, z^2 \rangle \downarrow \min_{z^1, \omega^1, \gamma^1, z^2, \omega^2, \gamma^2} \begin{cases}
- A\omega^1 + \gamma^1 e_m + e_m \leq 0, \\
B\omega^1 - \gamma^1 e_k + e_k \leq z^1, \\
z^1 \geq 0,
\end{cases}
\]

\[
(z^1, \omega^1, \gamma^1) \in Z_1 \triangleq \begin{cases}
(z^1, \omega^1, \gamma^1) \in \mathbb{R}^K \;& \begin{cases}
- A\omega^1 + \gamma^1 e_m + e_m \leq 0, \\
B\omega^1 - \gamma^1 e_k + e_k \leq z^1, \\
z^1 \geq 0,
\end{cases}
\end{cases}
\]

\[
(z^2, \omega^2, \gamma^2) \in Z_2 \triangleq \begin{cases}
(z^2, \omega^2, \gamma^2) \in \mathbb{R}^K \;& \begin{cases}
- A\omega^2 + \gamma^2 e_m + e_m \leq 0, \\
B\omega^2 - \gamma^2 e_k + e_k \leq z^2, \\
z^2 \geq 0,
\end{cases}
\end{cases}
\]

where \( K = k + n + 1 \), \( e_m \triangleq (1, 1, \ldots, 1) \in \mathbb{R}^m \), \( e_k \triangleq (1, 1, \ldots, 1) \in \mathbb{R}^k \), \( z^1 \) and \( z^2 \) are auxiliary variables \( (z^1_j = \max \{0, \langle \omega^1, B^j \rangle - \gamma^1 \} \), \( z^2_j = \max \{0, \langle \omega^2, B^j \rangle - \gamma^2 \} \), \( j = 1, k \) \). These variables determine the errors of separation. If \( z^1 = 0 \) and \( z^2 = 0 \) then sets \( \mathcal{A} \) and \( \mathcal{B} \) are non-symmetrically bilinearly separable. So, we have the following theorem.

**Theorem 1.** [2] The sets \( \mathcal{A} \) and \( \mathcal{B} \) are bilinearly separable in the space \( \mathbb{R}^m \) in the sense of the system (1) if and only if the value of the problem \( (\mathcal{B}L\mathcal{P}_1) \) is zero. In that case, the components \( (\omega^1, \omega^2, \gamma^1, \gamma^2) \) of global minimum point determine separating hyperplanes \( \langle \omega^1, x \rangle = \gamma^1 \) and \( \langle \omega^2, x \rangle = \gamma^2 \).

The problem \( (\mathcal{B}L\mathcal{P}_1) \) is the bilinear optimization problem with disjoint constraints [3, 4].

In turn, the problem of bilinear separation in the sense of the system (3) can be reduced to the following optimization problem (symmetrical case) [1]:

\[
\langle (y^1 + y^2), (y^3 + y^4) \rangle + \langle (z^1 + z^2), (z^3 + z^4) \rangle \downarrow \min_{y^l, z^l, \omega^1, \omega^2, \gamma^1, \gamma^2} \begin{cases}
- A\omega^1 + \gamma^1 e_m + e_m \leq y^1, \\
- A\omega^2 + \gamma^2 e_m + e_m \leq y^2, \\
A\omega^1 - \gamma^1 e_m + e_m \leq y^3, \\
A\omega^2 - \gamma^2 e_m + e_m \leq y^4, \\
B\omega^1 - \gamma^1 e_k + e_k \leq z^1, \\
B\omega^2 + \gamma^2 e_k + e_k \leq z^2, \\
- B\omega^1 + \gamma^1 e_k + e_k \leq z^3, \\
B\omega^2 - \gamma^2 e_k + e_k \leq z^4, \\
y^l \geq 0, \quad z^l \geq 0, \quad l = \overline{1, 4},
\end{cases}
\]

\[(\mathcal{B}L\mathcal{P}_3)\]

where \( y^l \) and \( z^l \) \( (l = \overline{1, 4}) \) are auxiliary variables.

Similarly to Theorem 1, the sets \( \mathcal{A} \) and \( \mathcal{B} \) are symmetrically bilinearly separable if and only if the value of the problem \( (\mathcal{B}L\mathcal{P}_3) \) is zero, and the
components \((\omega_1^*, \omega_2^*, \gamma_1^*, \gamma_2^*)\) of global minimum point determine separating hyperplanes.

The problem \((BLP_3)\) is more difficult bilinear optimization problem with joint constraints \([4]\).

3. **Local and global search.** It is clear that the nonconvexity of the problems \((BLP_1)\) and \((BLP_3)\) occurs due to the scalar products of independent variables in the goal functions. In addition, it can be shown that these functions can be represented as a difference of two convex functions, i.e. the goal functions are d.c. functions \([5]\). Therefore, for the development of methods for solving these problems one can applies the Global Search Theory in d.c. optimization \([5]\).

Global Search Theory consist of two principal stages \([5]\):

1) a special local search methods, which takes into account the structure of the problem under scrutiny;

2) the procedures, based on Global Optimality Conditions \([5]\), which allow to improve the point provided by the Local Search Method.

For example, consider the problem \((BLP_1)\). Since the objective function \(F\) has a bilinear structure, so it is affine w.r.t. each of its variables when the other variable is fixed, for the implementation of local search we use well known idea of consecutive solving of the auxiliary linear programming problems, starting with fixed point \((z_1^s, z_2^s)\) \([3, 4]\):

\[
\langle z^1, z_2^s \rangle \downarrow \min_{(z^1, \omega^1, \gamma^1)} \left( z^1, \omega^1, \gamma^1 \right) \in Z_1,
\]

\[
\langle z_1^s, z^2 \rangle \downarrow \min_{(z^2, \omega^2, \gamma^2)} \left( z^2, \omega^2, \gamma^2 \right) \in Z_2.
\]

These auxiliary problems can be efficiently solved using known software packages. It is proved that Local Search procedure converges to the critical point, which is partially global solution to the problem \((BLP_1)\) w.r.t. each group of variables \((z^1, \omega^1, \gamma^1)\) and \((z^2, \omega^2, \gamma^2)\). In practical implementation, after a finite number of iterations when some stopping criterion is satisfied can be obtained a critical point with the required accuracy.

Next, using the following d.c. decomposition:

\[
F(z^1, z^2) = g(z^1, z^2) - f(z^1, z^2),
\]

where \(f(z^1, z^2) = \frac{1}{4} \| z^1 - z^2 \|^2\), \(g(z^1, z^2) = \frac{1}{4} \| z^1 + z^2 \|^2\), and special features of problem \((BLP_1)\) (see Theorem 1), an algorithm for global

---

201
search separating hyperplanes based on the method of solving bilinear programming of the general form [3, 4] was elaborated. It is necessary to emphasize that for the construction of level surface approximation of \( f(z^1, z^2) \), determining the basic nonconvexity in the problem \( \mathcal{BLP}_1 \) (this stage of the global search is one of the crucial stages), a new set of directions, which are more effective than the standard sets [3, 4] is proposed.

As far as solving of the problem \( \mathcal{BLP}_3 \) is concerned, we develop the same ideas of local and global search for the more difficult bilinear problems with joint constraints.

4. Generation of test problems and computational simulation. To analyze the efficiency of new methods for solving optimization problems it is often required to be able to find or generate test problems with known solutions and properties. Sometimes, libraries of test problems are available, in some other cases, the problem under consideration has a solution for any initial data (for example, this is the case of bimatrix games [4]). In bilinear separation problems, one cannot guarantee that a solution exists for arbitrary initial data. Moreover, no representative test libraries are available for these kind of problems. For that reason we propose a new method for generating bilinear separation test problems with known global solutions [6].

For example, to test the algorithms for local and global search in the problem of \( \mathcal{BLP}_1 \) more than 20,000 of various dimension (from 2 up to 200) test problems were generated by proposed method. The programs that implement the proposed algorithms of local and global search were elaborated, and computational simulation was carried out. In total more than 99% of the generated problems were solved with required accuracy.

In general, we can conclude that computational testing of the elaborated methods has shown the efficiency of the proposed approach to the problems of bilinear separability.

References

Ensemble calculations application for estimation and optimization of climate model parameters

Valeriy Parkhomenko

Institution of Russian Academy of Sciences
Dorodnicyn Computing Centre of RAS,
Moscow, Russia; parhom@ccas.ru

A simplified climate model is presented which includes a fully three-dimensional, frictional geostrophic ocean component [1, 2] but retains an integration efficiency considerably greater than extent climate models with three-dimensional, primitive-equation ocean representations (3000 years of integration can be completed in about 5 hours on a PC). The model also includes energy and moisture balance atmosphere and a dynamic and thermodynamic sea-ice model.

Climate models incorporate a number of adjustable parameters whose values are not always well constrained by theoretical or observational studies of the relevant processes. Even the nature of the processes may be unclear and dependent upon resolution, as sub-grid scale mixing parameterizations, particularly for coarse resolution models, may represent a wide variety of different physical processes (eddies and unresolved motions, inertia-gravity waves, tides etc.). In such cases parameter values would ideally be chosen by optimizing the fit of model predictions to observational data.
Efficient models have the potential to perform large numbers of integrations and hence explore larger regions of their parameter space. Where the parameters have clear physical interpretations or close equivalents in higher resolution models, the results may be of more general relevance. Efficient models are also essential for the understanding of very long-term natural climate variability, in which case the optimal inter-component balance of model complexity may depend on the timescale range of interest.

Frictional geostrophic ocean models [1, 2] are applicable with arbitrary bottom topography in a global setting, but significantly simpler than the primitive equation dynamics. We describe here the combination of the latter model, an energy and moisture balance atmosphere and a dynamic and thermodynamic sea-ice model. At a resolution of 72 by 72 cells in the horizontal [3], and given the extremely simple representation of the atmosphere, the resulting coupled model is highly efficient. We perform an initial investigation of the space generated by the simultaneous variation of model parameters by analyzing a set of model runs. We do not attempt to produce well-converged statistical analyses, our aim is to investigate the extent to which both model parameters and model predictions of global change are constrained by quantitative comparison with data [4]. Thus we commence our analysis by defining and applying an objective measure of model error and discuss the modeled climate in the low-error runs.

The frictional geostrophic ocean model principal governing equations [1, 2] are similar to classical general circulation models, with the neglect of momentum advection and acceleration. In the vertical there are normally 8 depth levels on a uniformly logarithmically stretched grid with vertical spacing increasing with depth from 175 m to 1420 m. The maximum depth is set to 5 km. The horizontal grid is uniform in the longitude and sin of latitude coordinates giving boxes of equal area in physical space. The horizontal resolution is normally 72 by 72 cells [3].

We use an energy and moisture balance model of the atmosphere, similar to that described in [5]. The prognostic variables are surface air temperature and surface specific humidity for which the governing equations can be written. The short-wave solar radiative forcing represents seasonally changed conditions. In a further departure from that model, the relevant planetary albedo is given by a surface type properties. Over sea ice the albedo is temperature-dependent. Heat absorption by water vapor, dust, ozone, clouds, etc. is parameterized by the constant value. The remaining heat sources and sinks describe the long-wave radiation
and latent heat processes. For anthropogenically forced experiments a greenhouse warming term is added which is proportional to the log of the relative increase in carbon dioxide concentration as compared to an arbitrary reference value. The sensible heat flux depends on the air-surface temperature difference and the surface wind speed (derived from the ocean wind-stress data), and the latent heat release is proportional to the precipitation rate [5].

It is used an implicit scheme to integrate the atmospheric dynamical equations. The scheme comprises an iterative, semi-implicit predictor step followed by a corrector step which renders the scheme exactly conservative. Sea ice dynamical equations are solved for the fraction of the ocean surface covered by sea ice in any given region and for the average height of sea ice. In addition a diagnostic equation is solved for the surface temperature of the ice. Following [6, 7] thermodynamic growth or decay of sea ice in the model depends on the net heat flux into the ice from the ocean and atmosphere. Sea-ice dynamics simply consist of advection by surface currents and Laplacian diffusion with constant coefficient.

Fixing the distribution of drag, we have a set of 10 model parameters related to mixing and transport, augmented to 12 if we allow for variation of the width and slope and of the atmospheric diffusivity. If we vary these parameters individually, as in conventional, single parameter studies, we visit only a very restricted region of parameter space. We therefore allow all 12 parameters to vary at once within specified ranges [4].

Note that it may be appropriate to use larger values of frictional and diffusive parameters than in higher resolution models. In our semi-random approach, we generate an ensemble by uniformly spanning the range of each individual parameter, but choose combinations of parameters at random. This is equivalent to an equal subdivision of probability space if the probability distributions for the parameters are uniform. Thus with M runs and N parameters, each parameter takes M, uniformly (or logarithmically) spaced values between its two extrema, but the order in which these values are taken is defined by a random permutation. Each run is a separate, 2000-year integration from a uniform state of rest under standard forcing. We are primarily concerned with the effects of variations of model parameters, thus we fix the initial ocean temperature at 20 C unless otherwise stated. This results in a rapid, convectively driven start to the oceanic adjustment process. To process the results of a large number of runs we have to define an objective measure of model
error. To do so it is used a weighted root mean square error over the set of all dynamical variables in the ocean and atmosphere, as compared to interpolated observational datasets, namely NCEP surface (1000 mb) atmospheric temperature and specific humidity, averaged over the period 1948 to 2002, and ocean temperature and salinity [8].

By analyzing a randomly generated set of runs, each 2000 years in length, we have considered the uncertainty in 12 mixing and transport parameters. Constructing a quantitative measure for the model error allowed us to address both the inverse problem of estimation of model parameters, and the direct problem of model predictions. Results represent an attempt at tuning a 3-D climate model by a strictly defined procedure which nevertheless considers the whole of the appropriate parameter space. Modelling approach is thus to match model outputs to observations while model inputs (parameters) are initially only weakly constrained.

The author was supported by the Russian Foundation for Basic Research (projects no. 11-01-93003, 11-01-00575, 11-07-00161) and Presidium RAS Basic Research Program (no. 14).

REFERENCES

Interior point algorithms

Sergey Perzhabinsky\textsuperscript{1}, Valery Zorkaltsev\textsuperscript{2}

\textsuperscript{1} Melentiev Energy Systems Institute SB RAS, Irkutsk, Russia; smper@isem.sei.irk.ru
\textsuperscript{2} Melentiev Energy Systems Institute SB RAS, Irkutsk, Russia; zork@isem.sei.irk.ru

Introduction. The algorithms for mathematical programming problems are considered in the report. Solution improvement in these algorithms is carried out inside a set of strongly feasible points. The strongly feasible points are points which satisfy inequality constraints in strict form. The pioneer results of the algorithm studies were obtained by S. Antsyz, I. Dikin \cite{1}, Yu. Evtushenko \cite{2}, V. Zhadan, V. Zorkaltsev. The given interior point algorithms are effectively used for realization of an applied models since 70th years of the last century. The interesting properties of algorithms were discovered on the basis of the theoretical and experimental studies of the algorithms. In particular, it was shown that the given algorithms are produced relative interior point of the optimal solution set in case of non-uniqueness of the optimal solution. The relative interior point has minimal set of the active constraints. The interior point algorithms attract the heightened interest since the middle of 80th years of the last century. It was demonstrated in the independent studies that the interior point methods are more efficient for linear programming problems comparing with the simplex-method. In addition the algorithms for solving nonlinear programming problems can be easy realized on the basis of the interior point method.

The results of theoretical justification and experiments of the interior point algorithms for solving linear are presented in the report. The ways
of improving of computational efficiency of the interior point algorithms are considered. The results of some studies of the interior point algorithms for nonlinear programming are discussed.

Let’s consider the primal and dual linear programming problems

$$c^T x \rightarrow min, \quad X = \{ x \in R^n : Ax = b, \ x \geq 0 \};$$

$$b^T u \rightarrow max, \quad U = \{ u \in R^m : g(u) \geq 0 \}. \quad (1)$$

Here $g(u) = c - A^T u$ is linear vector-function with components $g_j(u),$ $j = 1, \ldots, n$. The vectors $x \in R^n, u \in R^m$ are variables of problems (1), (2). The vectors $c \in R^n, b \in R^m$ and the matrix $A$ of order $n \times m$ are given.

The vectors $x \in X, u \in U$ are feasible solutions of problems (1), (2). The sets of optimal solutions of these problems are denoted

$$X = \text{Argmin} \{ c^T x : x \in X \}, U = \text{Argmax} \{ b^T u : u \in U \}.$$ 

The problem (1) is the primal linear programming problem. The problem (2) is the dual linear programming problem.

The problem (1) is called non-degenerate, if there is no more one vector $u \in R^m$ for which complementary conditions hold for any $x \in X$:

$$x_j g_j(u) = 0, \ j = 1, \ldots, n. \quad (3)$$

The problem (2) is called non-degenerate, if there is the unique $x \in R^n$ such that $Ax = b$ and condition (3) holds for any $u \in U$.

The interior region of the convex set $Q \subset R^n$ relative to minimal linear manifold which contains this region, is called relative interior $riQ$ [3].

The class of interior point algorithms. The initial approximation $x^0 > 0$ is given. The iterative process is considered

$$x^{k+1} = x^k + \lambda_k \Delta x^k, \ k = 0, 1, 2, \ldots. \quad (4)$$

Here the vector $\Delta x^k$ is the direction of improving the solution on iteration $k$, the value $\lambda_k$ is the positive step in this direction. The vector $\Delta x^k$ is the solution of the auxiliary problem

$$c^T \Delta x + \frac{1}{2} \sum_{j=1}^{n} \frac{(\Delta x_j)^2}{d_j^k} \rightarrow min, \quad \Delta x \in R^n, \quad (5)$$
\[ A \Delta x = r^k, \]  

(6)

where

\[ r^k = b - A x^k. \]  

(7)

In addition the values \(1/d_j^k\) should be interpreted as penalty coefficients which prevent breaking the borders of the feasible solution set. The inverse value of penalty coefficient will be called weighting coefficient.

The interval and axiomatic approaches to defining of coefficients \(d_j^k\) were introduced in [4]. The inequalities must be hold for the weighting coefficient

\[ \sigma_j(x_j^k) \leq d_j^k \leq \sigma_j(x_j^k), \ j = 1, \ldots, n \]  

(8)

where \(\sigma, \sigma\) are some continuous non-decreasing functions of positive argument which satisfies the two conditions

\[ 0 < \sigma(t) \leq \sigma(t), \ \forall t > 0, \]  

(9)

\[ \sigma(t) \leq Mt \]  

(10)

for some \(\varepsilon > 0, M > 0\) and all \(t \in (0, \varepsilon]\). The rules of definition of values \(d_j^k\) which satisfy (8) - (10) forms the class of algorithms. In particular, the weighting coefficients can be given in the form of functions of the components of vector of variables

\[ d_j^k = (x_j^k)^p, \ j = 1, \ldots, n \]  

(11)

where \(p \geq 1\). Let’s denote \(D_k = diag\ (d^k)\) where \(d^k\) is the vector of \(R^n\) with components \(d_j^k, j = 1, \ldots, n\).

The step size is computed by the formula

\[ \lambda_k = \gamma_k \min_{j: \Delta x_j^k < 0} \frac{-x_j^k}{\Delta x_j^k} \]  

(12)

as \(\gamma_k \in (0, 1)\). The meanings of parameter \(\gamma_k\) doesn’t decrease on iterations \(\gamma_k \leq \gamma_{k+1}\) for all \(k\). The moving from point \(x^k\) by the direction \(\Delta x^k\) is realized on the part of way which is equal \(\gamma_k\) to the bound of the region \(x \geq 0\). It is possible to change the parameter \(\gamma_k\) at each iteration. In particular, the parameter \(\gamma_k\) can converge to 1. This rule can be used for the acceleration of the computational process.

209
Using the Lagrange method of multipliers for solving the problem (5), (6), we obtain
\[ \Delta x^k = -D_k g(u^k) \]
where the vector \( u^k \) is the solution of the system of linear equations
\[ AD_k A^T u = r^k + AD_k c. \]  
(13)
Thus the vector of the Lagrange multipliers of the auxiliary problem is found at each iteration. This vector is an approximation of the solution of the dual linear programming problem.

The iterative process of the improving of the solution of problem (1) consists of the two stages. The first stage is the input to the feasible region. The absolute values of the residuals of the equality constraints monotonically decrease on the iterations at this stage. It follows from (4), (6), (7) that \( r^{k+1} = (1 - \lambda_k) r^k \). Therefore if \( r^k \neq 0 \) then
\[ \lambda_k = \min \{1, \tilde{\lambda}_k\} \]  
(14)
where \( \tilde{\lambda}_k \) is computed by (12).

The optimization in the feasible region is the second stage of the iterative process. All constraints of the problem (1) are satisfied and the objective function value monotonically decreases in this case. In addition \( r^k \) is equal 0 in (13).

The theoretical justification of the given computing process in the feasible region is presented in [5]. This theorem is proved under the non-degeneracy condition and for the different weighting coefficients.

**Theorem 1.** Let the problem (1) is non-degenerate, \( \overline{X} \neq \emptyset, \overline{X} \neq X \) and \( x^0 \in X \) then the following statements are true for the given algorithm.
1. There are \( \overline{x} \in \overline{X}, \overline{u} \in ri\overline{U} \) such that
   \[ x^k \to \overline{x}, \ u^k \to \overline{u} \quad \text{as} \quad k \to \infty. \]
2. If the condition holds for the weighting coefficients \( d_j^k \)
   \[ \overline{\sigma}(\alpha)/\overline{\sigma}(\beta) = O(\alpha/\beta) \]  
(15)
then \( \overline{x} \in ri\overline{X} \), the values \( \|u^k - \overline{u}\|, \|x^k - \overline{x}\| \) converge to 0 no slower than arcwise, and the following relation takes place for some positive values \( P_1, P_2, P_3, P_4, \rho \in (0, 1) \) for all \( k \)
\[ \|u^k - \overline{u}\| \leq P_1 L_k, \ L_k \leq P_2 T_k, \ T_k \leq \|x^k - \overline{x}\| \leq P_3 T_k, \ T_k \leq P_4 (\rho)^k \]
where

\[
L_k = \max_{j: g_j(\overline{u}) \neq 0} d_j^k, \quad T_k = \max_{j: g_j(\overline{u}) \neq 0} x_j^k.
\]

3. If the following condition holds for the weighting coefficients \(d_j^k\)

\[
\sigma(\alpha) / \sigma(\beta) = o(\alpha/\beta) \tag{16}
\]

then

\[
\left\| u_k - \overline{u} \right\| / \left\| x_k - \overline{x} \right\| \to 0 \text{ as } k \to \infty \tag{17}
\]

for some \(P_5 > 0\) and all \(k\)

\[
T_{k+q+1} \leq P_5 T_k \max\{ (1 - \gamma) : \tau \in \{k, \ldots, k + q\}\}.
\]

Here \(q\) is amount of numbers \(j\) such that \(g_j(\overline{u}) \neq 0\).

Let’s note that the condition (15) holds for the weighting coefficients (11) as \(p = 1\). If \(p > 1\) then the stronger condition (16) holds. The algorithm with the weighting coefficients (11) as \(p = 2\) is also known in the western publications as affine scaling method. The weighting coefficients \(d_j^k = (x_j^k)^2\) and the step \(\lambda_k = \left( (g(u^k))^T D_k^{-1} g(u^k) \right)^{1/2}\) are used in Dikin interior point algorithm. If \(\overline{X} \neq \emptyset\) then vectors \(x^k\) converge arcwise to a point of \(ri\overline{X}\) in this algorithm [6].

The theoretical justification of the algorithms without the non-degenerate assumption is very important for the theory and practice. The given algorithms with the weighting coefficients (11) for all \(p \in (1, 3]\) were theoretically justified without the non-degenerate assumption [7]. The additional condition on \(\gamma_k\) was used for this justification

\[
\gamma_k \in (0, 2/(p + 1)), \quad k = 0, 1, 2, \ldots. \tag{18}
\]

**Theorem 2.** Let \(X \neq \emptyset, \overline{X} \neq X\) and \(x^0 \in X\) then the following statements are true for the algorithm (4) - (12) as \(p \in (1, 3]\) and the condition (18).

1. If \(\overline{X} \neq \emptyset\) then there are \(\overline{x} \in ri\overline{X}, \overline{u} \in ri\overline{U}\) such that

\[
x^k \to \overline{x}, \quad u^k \to \overline{u},
\]

\[
\left\| x^{k+1} - \overline{x} \right\| / \left\| x^k - \overline{x} \right\| \to (1 - \gamma),
\]

\[
\left\| u^k - \overline{u} \right\| / \left\| x^k - \overline{x} \right\| \to 0 \text{ as } k \to \infty.
\]
2. If \( \overline{X} = \emptyset \) then \( \lim_{k \to \infty} c^T x^k = -\infty \) or the all components of the vector \( \Delta x^k \) are positive of some iteration \( k \).

**Conclusion.** The condition (8) allows us to use different techniques of computing of the weighting coefficients. It is not necessary to have specific expressions for the functions \( \sigma, \sigma' \). It is enough to have the proof of the existence of such functions and execution of their properties. In particular, the following weighting coefficients are effective under \( k \geq 1 \)

\[
d^k_j = \frac{x^k_j}{\max\{\epsilon, g^k_j\}}, \quad j = 1, \ldots, n
\]

where \( \epsilon \) is small positive value. The interior point algorithms with the weighting coefficients (19) have superlinear rate of convergence [5].

The execution of the condition (10) for the weighting coefficients is enough for convergence of computational process accordingly to Theorem 1. The execution of the condition (15) is necessary for superlinear convergence of computational process and the receiving of the relative interior point of the optimal solution set. The execution of the condition (16) and convergence \( \gamma_k \) to one as \( k \to \infty \) provide superlinear rate of convergence of the given algorithm.

If the relation (17) holds then vectors of dual variables converge to the optimal solution of the dual problem rather than the primal variables converge to their optimal meanings.

**References**

Interactive optimization as a tool for finding the complex periodic solutions in nonlinear dynamics

Lev F. Petrov

1 Plekhanov Russian University of Economics, Department of Economics and Mathematics, Moscow, Russia; lfp@mail.ru, lefp@rambler.ru

Introduction. We consider essentially nonlinear dynamical systems with the ability to implement a chaotic behavior and deterministic solutions of various kinds. Among the deterministic solutions, we will highlight a variety of periodic solutions of different periods. Problems of control of dynamic regimes in such systems discussed in [1]. This work is devoted to numerical algorithms for constructing and analyzing the stability of periodic solutions of strongly nonlinear dynamical systems.

Problem Statement. We will consider the strongly nonlinear system, which places no restrictions on the value of the individual components. In the framework of this approach we can analyze the linear and quasi-linear system, but the focus will be given to an essentially nonlinear systems of general form. The only requirement that we make to a dynamic system is the ability to construct numerical solutions of the Cauchy problem with the required precision.

We will use a general approach to the problem of constructing periodic solutions of nonlinear systems of ordinary differential equations. This approach H. Poincare [2] formulated as follows: Suppose that

$$\frac{dx_i(t)}{dt} = X_i \quad (i = 1, 2, ..., n)$$

(1)

is the system of differential equations, where $X_i$ - data, unambiguous function of the variables $x_1, x_2, ..., x_n$, and maybe, time $t$.

Suppose now that

$$x_1 = \varphi_1(t), \quad x_2 = \varphi_2(t), ..., x_n = \varphi_n(t)$$

(2)
is particular solution of this system.

Imagine that at time $T$ $n$ variables $x_i$ take their initial values, so that $\varphi_i(0) = \varphi_i(T)$. It is clear that at time $T$, we will be in the same conditions as at time 0, and hence for any $t$ $\varphi_i(t) = \varphi_i(t + T)$. In other words, the functions $\varphi_i(t)$ are periodic functions of $t$.

Variant 1. The system (1) is autonomous, that is, the right parts $X_i$ are not depend on time $t$. In this case, the period $T$ of the solution is unknown.

Variant 2. The system (1) is non-autonomous, that is, the right parts $X_i$ are depend on time $t$:

$$X_i = X_i(t, x_1, x_2, ..., x_n) \quad (3)$$

In this case, the period $T$ of the system (1) is known:

$$X_i(t, x_1, x_2, ..., x_n) = X_i(t + T, x_1, x_2, ..., x_n) \quad (4)$$

A periodic solution can have a multiple of the period $kT, k = 1, 2, ...$

$$\varphi_i(t) = \varphi_i(t + kT) \quad (5)$$

H. Poincare for finding periodic solutions of the system implies the existence of a small parameter. We use his approach on the initial conditions of the periodic solution, but let’s not assume the existence of a small parameter in the system (1).

Problem Statement of constructing periodic solutions of strongly non-linear autonomous dynamical system (1) (variant 1):

Find the initial conditions $\varphi_i(0), (i = 1, ..., n)$, corresponding to the periodic solution and the period $T$ of this solution: $\varphi_i(0) = \varphi_i(T), (i = 1, ..., n)$, and therefore $\varphi_i(t) = \varphi_i(t + T), (i = 1, ..., n)$.

Problem Statement of constructing periodic solutions of strongly non-linear autonomous dynamical system (1) (variant 2):

Find the initial conditions $\varphi_i(0), (i = 1, ..., n)$, corresponding to the $kT$-periodic solution: $\varphi_i(0) = \varphi_i(kT), (i = 1, ..., n, k = 1, 2, ...)$, and therefore $\varphi_i(t) = \varphi_i(t + kT), (i = 1, ..., n)$.

Note that the dimension of variant 1 is $n + 1$, the dimension of variant 2 is $n$. After finding the initial conditions of the periodic solution it is built using the numerical integration in one period.

Consider the algorithm for determining the initial conditions of the periodic solution of nonautonomous nonlinear dynamics problems (variant
2). The period $T$ of the system is known. The period of solutions given by integer parameter $k$ at the start. Note that the increase in the integer parameter $k$ does not significantly restrict the form of the solution. For example, for $k = 12$ in the search box, the solution with $k = 1, 2, 3, 4, 6, 12$ are included. (Fig.1).

![Fig. 1](image1)

Denote the unknown initial conditions, corresponding to $kT$-periodic solution $Y_i = x_i(0)$. Obviously, for periodic solutions to satisfy the equality

$$x_i(kT) = Y_i, i = 1, 2, ..., n$$  \hfill (6)

This is done only when the initial conditions corresponding to a periodic solution.

Consider the function (Fig.2)

$$F(Y_1, Y_2, ..., Y_n) = \sqrt{\sum_{i=1}^{n} [Y_i - x_i(kT)]^2}$$  \hfill (7)

![Fig. 2](image2)

215
This function defines the discrepancy in fulfilling the conditions of periodicity. Obviously, for a periodic solution

\[ F(Y_1, Y_2, \ldots, Y_n) = 0 \]  

(8)

Therefore, to determine the initial conditions corresponding to a periodic solution, it is possible to use optimization algorithms with the objective function

\[ F(Y_1, Y_2, \ldots, Y_n) - \rightarrow \text{min} \]  

(9)

Comparative numerical experiments have shown that the more effective is another way to determine the initial conditions \(Y_i, i = 1, 2, \ldots, n\) corresponding to a periodic solution.

To find the initial conditions \(Y_1, Y_2, \ldots, Y_n\) corresponding to a periodic solution we use a system of nonlinear algebraic equations

\[ Y_i - x_i(kT) = 0, \quad i = 1, 2, \ldots, n \]  

(10)

This system is not divided into separate equations, since the quantities \(x_i(kT)\) are determined from the original nonlinear system of ordinary differential equations (1) by numerical calculation of the Cauchy problem on the interval \([0, kT]\). To solve this system we have used Newton’s method. In the computer implementation of this algorithm includes the possibility of interactive control of calculations. This allows us to specifically control the calculations to find the most interesting of periodic solutions.

For an autonomous system of ordinary nonlinear differential equations (1) (variant 1) the period \(T\) of the solution is also unknown. In this case, since the initial time is arbitrary, we assume that \(Y_n = 0\). Then to find the initial conditions of \(Y_1, Y_2, \ldots, Y_{n-1}\) and period \(T\) of solutions we have the system of nonlinear algebraic equations

\[ Y_i - x_i(T) = 0, \quad i = 1, 2, \ldots, n - 1 \]  

\[ x_n(T) = 0 \]  

(11)

This system is solved using Newton’s method. After finding the initial conditions \(Y_1, Y_2, \ldots, Y_{n-1}\) and period \(T\) of solution we numerically calculate periodic solution itself.

Note that the dimensions of systems of nonlinear algebraic equations for variant 1 (11) and variant 2 (10) are the same and equal to \(n\).
To determine the stability of periodic solutions found, we construct the variational system and calculate the multipliers.

Note that the presented algorithm for finding periodic solutions of strongly nonlinear dynamical systems is iterative, and stability analysis algorithm is finite.

**Interactive algorithm.** The complexity of the dynamic system behavior is described by periodic solutions of strongly nonlinear systems of ordinary differential equations determines the branching structure of the algorithm for constructing periodic solutions. Here we investigate the evolution of these solutions when changing parameters of the dynamical system.

Note that the convergence of Newton’s method depend on starting initial conditions. These conditions can be defined either randomly or on the results of the previous step on the parameter of the dynamical system. The dynamical system may not have periodic solutions of a certain period. To overcome these and other computational problems, our algorithm for finding periodic solutions of nonlinear systems of ordinary differential equations is realized in an interactive mode (Fig. 3). Interim results of finding periodic solutions are shown to the user in real time. The user can intervene in the computation and change the parameters of dynamic systems, finding solution numerical methods, tactics of searching.

This approach allows us to find most complex periodic solutions of nonlinear dynamical systems with several degrees of freedom even in the field of dynamical chaos, where all the periodic solutions are unstable. Also, this approach allows us to investigate bifurcation of periodic solutions.

![Fig. 3](image-url)
On the Semidefinite Representation of the Maximum Optimal Rate Problem in LDPC Codes

M.R. Peyghami\textsuperscript{1}, H. Tavakoli\textsuperscript{2}, M. Ahmadian Attari\textsuperscript{3}

\textsuperscript{1} Department of Mathematics, K.N. Toosi University of Technology, P.O. Box 16315-1618, Tehran, Iran; peyghami@kntu.ac.ir
\textsuperscript{2} Department of Electrical Engineering, K.N. Toosi University of Technology, Tehran, Iran; m.ahmadian@kntu.ac.ir
\textsuperscript{3} Department of Electrical Engineering, K.N. Toosi University of Technology, Tehran, Iran; tavakoli@ee.kntu.ac.ir

Abstract

Achieving and approaching to the channel capacity are two major important properties of designing good LDPC codes. For designing good LDPC codes in infinite mode, it is important to solve the maximum optimal rate problem. The Binary Erasure Channel (BEC) consists of the simplest form of the optimal rate problem, and achieving and approaching to the capacity for BEC have been studied in literature by many researchers. Although, the structure of the optimal rate problem for BEC is the simplest form, but finding an answer to satisfy the non-linear constraints of the problem is a hard part of the problem. In this talk, we will discuss about the semidefinite reformulation of the optimal rate code design in BEC.

Keywords: Semidefinite Optimization, Binary Erasure Channel, Maximum Optimal Rate Problem.
Problems of optimal resistance in Newtonian aerodynamics

Alexander Plakhov

1 University of Aveiro, Aveiro, Portugal and Institute for Information Transmission Problems, Moscow, Russia; plakhov@ua.pt

A body moves in a rarefied medium of resting particles and at the same time very slowly rotates (somersaults). Each particle of the medium is reflected elastically when hitting the body boundary (multiple reflections are possible). The resulting resistance force acting on the body depends on the time; we are interested in minimizing the time-averaged value of resistance (which is called $R$). The value $R(B)$ is well defined in terms of billiard in the complement of $B$, for any bounded body $B \subset \mathbb{R}^d$, $d \geq 2$ with piecewise smooth boundary.

Let $C \subset \mathbb{R}^d$ be a bounded convex body and $C_1 \subset C$ be another convex body with $\partial C_1 \cap \partial C = \emptyset$. It would be interesting to get an estimate for

$$R(C_1,C) = \inf_{C_1 \subset B \subset C} R(B). \quad (1)$$

If $\partial C_1$ is close to $\partial C$, problem (1) can be referred to as minimizing the resistance of the convex body $C$ by "roughening" its surface. We cannot solve problem (1); however we can find the limit

$$\lim_{\text{dist}(\partial C_1,\partial C) \rightarrow 0} \frac{R(C_1,C)}{R(C)}. \quad (2)$$

It will be explained that problem (2) can be solved by reduction to a special problem of optimal mass transportation, where the initial and final measurable spaces are complementary hemispheres, $X = \{x = (x_1,\ldots,x_d) \in S^{d-1} : x_1 \geq 0\}$ and $Y = \{x \in S^{d-1} : x_1 \leq 0\}$. The transportation cost is the squared distance, $c(x,y) = \frac{1}{2} |x - y|^2$, and the measures in $X$ and $Y$ are obtained from the $(d-1)$-dimensional Lebesgue measure on the equatorial circle $\{x = (x_1,\ldots,x_d) : |x| \leq 1, x_1 = 0\}$ by parallel translation along the vector $e_1 = (1,0,\ldots,0)$. Let $\mathcal{C}(\nu)$ be the total cost corresponding to the transport plan $\nu$ and let $\nu_0$ be the transport plan generated by parallel translation along $e_1$; then the value $\frac{\inf \mathcal{C}(\nu)}{\mathcal{C}(\nu_0)}$ coincides with the limit in (2).
Surprisingly, this limit does not depend on the body $C$ and depends only on the dimension $d$. In particular, if $d = 3$ ($d = 2$), it equals (approximately) $0.96945$ ($0.98782$). In other words, the resistance of a 3-dimensional (2-dimensional) convex body can be decreased by 3.05\% (correspondingly, 1.22\%) at most by roughening its surface.

The author was supported by FEDER funds through COMPETE–Operational Programme Factors of Competitiveness and by Portuguese funds through the Center for Research and Development in Mathematics and Applications (CIDMA) and the Portuguese Foundation for Science and Technology (FCT), within project PEst-C/MAT/UI4106/2011 with COMPETE number FCOMP-01-0124-FEDER-022690; by the FCT research project PTDC/MAT/113470/2009; and by the Grant of President of Russia for Leading Scientific Schools NSh-5998.2012.1.

$L_1$ problems in control and numerical methods for their solution

Boris Polyak$^1$

$^1$ V.A. Trapeznikov Institute of Control Sciences RAS, Moscow, Russia; boris@ipu.ru

$L_1$ techniques is highly popular in many fields, it suffices to mention Lasso regression, exact penalties and basis pursuit in optimization, Least absolute values methods in estimation, Compressed Sensing in signal and image processing, SVM methods in classification and recognition. However this approach was not exploited in control theory.

We consider various problems in linear control theory which can be treated in the framework of $L_1$-optimization. They include optimal control with $L_1$-performance index, $L_1$-filtering, feedback stabilization with reduced number of outputs or states available. For them we study the properties of the solutions and provide appropriate numerical methods.
The Branch-and-Bound (B&B) is a general name for a family of methods to split an initial problem into subproblems which are sooner or later eliminated by bounding rules. Bounding rules determine whether a subproblem can yield a solution better than the best solution found so far. Today there is a big variety of Branch-and-Bound methods applicable to solving discrete, mixed-integer and continuous global optimization problems and even to multiobjective problems[1,2].

B&B methods for different problems share the common structure. Thus it is possible to select the set of parameters driving the resolution process common for different problems. Below we outline the most important parameters:

- the subproblem selection strategy (width-first, best-first, depth-first or a combined search strategy);
- heuristic to improve the incumbent solution (0 for none, or the positive number for selecting the respective heuristic);
- the bounding strategy (0 for none and positive number for selecting the respective bounding strategy).

In [1] it was shown that the proper selection of the mentioned parameters can speed up the resolution process or improve the obtained solution. It is possible to select these parameters statically or adjust them dynamically during computations. In the latter case we can express the control of computations as a extended finite state machine (EFSM) that accepts the characteristics of the computation process as an input and issues commands driving the computation process, e.g. ”set depth-first strategy”.

In the parallel implementation parallel processes exchange the subproblems and incumbents with each other. So besides the commands
applicable to the sequential case we have some more commands specific to the parallel implementation:

- send N subproblems to the process P;
- send incumbent to the process P;
- send control command to the process P;
- receive information (subproblems, incumbent or control command) from the process P.

The parallel B&B method’s logic can be defined as a EFSM that issues one the commands described above as a reaction on certain events. The examples of events are:

- the requested number of B&B steps performed;
- the incumbent was updated;
- data arrived from some process.

It is worth noting that this set of commands is problem-independent. And thus it is possible to separate the logic of the B&B method defined by the EFSM and the problem specific implementation of those commands. Such separation is important for several reasons. First, it saves efforts when implementing new problem because only the problem-specific part has to be implemented and the B&B logic is reused. Second, general part strictly defined using EFSM language can be a subject for a separate study. For instance it is possible to compare different load balancing strategies on a simulator or check the correctness of the parallel algorithm, e.g. identify possible deadlocks.

The proposed approach was implemented in BNB-Solver library[3]. BNB-Solver is written in C++. The class hierarchy can be roughly split into several different groups:

- basic numerical subroutines;
- communication subroutines supporting data exchange over MPI;
- problem specific modules;
- computations control modules (schedulers);
• solver classes.

The concrete application for solving certain optimization problem is obtained by combining three classes: the scheduler class, the problem-specific resolver class and the communicator class. The scheduler class manages the computations by issuing commands to either resolver or communicator. The resolver performs the requested number the computation steps. The communicator class carries out data exchange between different processes. Currently this approach was tested on the knapsack problem, mixed-integer programming problems and multiobjective optimization.

This work was supported by Russian Foundation for Basic Research (project no. 11-07-00360).

REFERENCES


Graf of decision logistics making for problem of goods delivery

Ekaterina Rassadnikova¹, Aida Valeeva², Nelli Magafurzyanova³

¹ Ufa State Aviation Technical University, Ufa, Russia; rassadnikova.ekaterina@gmail.com
² Ufa State Aviation Technical University, Ufa, Russia; Aidavaleyeva@yandex.ru
³ Ufa State Aviation Technical University, Ufa, Russia; nellimag.1990@gmail.com

The largest industry of our country is oil refining industry. Currently in Russia there are 30 large oil refineries with a total production capacity
of 261.6 million tons, and 80 mini-refineries with total refining capacity of 11.3 million tons. What are costs faced by the oil refinery? The cost of preparation of oil products connected with the specific conditions of production, training, transportation and refining of oil in the country. That is to say costs of oil products preparation was minimal, it is necessary not only to optimize the processes associated with the manufacture of products in a particular enterprise, but also the processes occurring in the accompanying enterprises. Namely optimization of an accompanying enterprise is considered in this article.

A company produces two types of a reagent used to neutralization of hydrogen sulphide and mercaptan in the gaseous and liquid mediums. Two types of raw materials are required for the production of the reagent, which the company purchases in several Russian cities. The reagent is produced on the installation a maximum production power of 25 tons per day. Consumers of the reagent are refineries in Russia. Warehouse is used for the storage of raw materials and finished products, this warehouse located on the territory of the enterprise (storage capacity is 200 tons).

The company is interested in the organization of an effective manage the entire supply chain of the reagent, which includes categories: procurement and supply management, inventory management, production, warehousing, product packing, material handling, order processing.

All categories presented above are related. In order to obtain an effective supply chain management these categories represented in the form of a graph (Figure 1). This graph was compiled on the basis of the graph logistics solutions presented in [1].

Currently our group has considered following categories of decision-making logistics Demand Forecasting (vertex 8), Product packing (vertex 32) and Transportation (vertices 23-30).

Category Transportation consists of the following tasks:
23. Transportation modes
24. Types of carriers
25. Carriers
26. Degree of consolidation
27. Transportation fleet mix
28. Assignment of customers to vehicles
29. Vehicle routing and scheduling
Fig. 1 - Graf logistics solutions for the enterprise that produces and supplies the reagent. Vertices - logistical problems and directed edges - connections between them.

Let us consider how these categories are related. We can make further decisions about the desired level of reserves (12) and a transportation fleet mix (27) based on the forecasting demand. Adopted decisions in its turn influence on the subsequent decisions, directly or indirectly. Thus, the desired inventory level directly affects the reserve stock (13). And decision about inventory control strategy (9) affects the relative importance of stocks (10) indirectly, through a decision about the supplier (20). Relative importance of stocks (10) impacts on management methods (11). Thus in addition to decisions within the logistics system, we consider some factors that ”come” from outside.

Influence of tasks on the transportation problems category (vertices 23-30) can be seen on Figure 2.
Fig. 2 - Influence of tasks on the transportation problems category (vertices 23-30).

The selected package (vertex 32) influences warehouse layout (vertex 41), order entry procedures (vertex 45) and material handling (vertex 39).

Consider an example delivery of petrochemical product (reagent). It is necessary to deliver the reagent produced in Ufa to Astrakhan. Thus, the problem of delivery to consumers has the following optimality criteria:

1) Demand forecasting
2) Choice of way and means of transportation
3) Optimal route
4) Choice of container type
5) Exploitation of the loading space.

You can see solution of this problem on Figure 3.

This article was examine the problems associated with the delivery of the goods in the containers by road routes in different cities of Russia. We have presented a problem in the form of a graph making logistical decisions.

The authors were supported by the Russian Foundation for Basic Research (project no. 11-07-00579, 12-07-00631).
Results:

Criterion 1: Demand forecasting of order is 17 tons.

Criterion 2: Methods of delivery by automobiles

\[ DV = \text{tonnage} \times 17 \text{ tons.} \]

Criterion 3: Ufa-Orenburg-Samara-Saratov - Volgograd - Astrakhan = 1754 km.

Criterion 4: Ufa-Samara-Saratov-Volgograd-Astrakhan = 1486 km.

Criterion 5: Solution of real problem

Fig. 3 - Solution of real problem

References


Discontinuous control and Lyapunov functions for nonlinear systems

Nataliya Sedova

1 Ulyanovsk State University, Ulyanovsk, Russia; sedovano@ulsu.ru

Some controlled systems (in particular, non-holonomic systems) cannot be stabilized by continuous static feedback at the origin. Stabilizing
feedback laws have been found for several such systems using either time-varying feedback or dynamic feedback. Another approach is to construct a piecewise continuous feedback laws, which are widely used in the optimal control theory also.

In this paper piecewise continuous control laws for nonlinear time-invariant systems are considered. Piecewise smooth Lyapunov functions or control Lyapunov functions can be used to construct the control law and to study behavior of trajectories of the controlled system.

Let $\mathbb{R}^n$ be the real $n$-vector space with a norm $|\cdot|$, $\mathbb{R}^+ = [0, +\infty)$, $\dot{x} = \dot{x}(t)$ be the right derivative at $t$ of a function $x : \mathbb{R}^+ \to \mathbb{R}^n$, and $x(t, x_0)$ denote the solution of a differential equation with $x(0) = x_0$.

Let $G \subset \mathbb{R}^n$ be a connected open set, $0 \in \bar{G}$, $f : (\{0\} \cup G) \to \mathbb{R}^n$, $f(0) = 0$. Also, let $M$ be an open subset of $\mathbb{R}^n$ such that $0 \in \bar{M} \subset (\{0\} \cup G)$. We say that origin $0 \in \mathbb{R}^n$ is asymptotically stable for the system $\dot{x} = f(x)$ on the set $M$, if for any $\delta > 0$ there exists $\varepsilon > 0$ such that for $x_0 \in M \cap \{|x| < \delta\}$ we have $x(t, x_0) \in M \cap \{|x| < \varepsilon\}$ for all $t \geq 0$ and $|x(t, x_0)| \to 0$ as $t \to +\infty$ (in particular, $\bar{M}$ is $f$-invariant, i.e. every trajectory starting at $x_0 \in \bar{M}$ remains in $\bar{M}$ for all $t > 0$).

Let $G_1, G_2$ be open connected subsets of $\mathbb{R}^n$ such that $G_1 \cup G_2 = \mathbb{R}^n \setminus \{0\}$, $f^i \in C(G_i, \mathbb{R}^n)$, $i = 1, 2$. Also assume that there exists a separating hypersurface $S$ with $0 \in S$ and $S \setminus \{0\} \subset G_1 \cap G_2$.

Let $M_1, M_2$ be connected components of $\mathbb{R}^n \setminus S$, $M_i \subset G_i$, and $f_i$ points towards $M_i$ on $S$ for $i = 1, 2$. If the the origin is asymptotically stable for $\dot{x} = f^1(x)$ on $G_1$ and for $\dot{x} = f^2(x)$ on $G_2$, then the zero solution of the system

$$
\dot{x} = \begin{cases} 
  f^1(x) & \text{if } x \in (S \setminus \{0\}) \cap M_1 \\
  f^2(x) & \text{if } x \in M_2 \\
  0 & \text{if } x = 0 
\end{cases}
$$

(1)

is asymptotically stable [1].

Sufficient conditions of asymptotic stability for system (1) can be stated in terms of two Lyapunov functions $V_1$ and $V_2$ defined on $G_1$ and $G_2$ respectively. This idea can be applied to controlled systems of the form $\dot{x} = f(x) + g(x)u$ with an input $u \in \mathbb{R}^m$: two control Lyapunov functions $V_i : \bar{G}_i \to \mathbb{R}^+$, $V_i \in C^1(G_i, \mathbb{R}^+)$, $V_i(0) = 0$, and the stabilizing feedback control of the form

$$
u = \begin{cases} 
u^1(x) & \text{if } x \in S \cap M_1 \\
  \nu^2(x) & \text{if } x \in M_2 
\end{cases}
$$

228
can be constructed.

We also consider systems of the form

\[
\dot{x} = \begin{cases} 
    f_1(x) & \text{if } x \in M_1 \\
    f_2(x) & \text{if } x \in M_2 \\
    \alpha f_1(x) + (1 - \alpha)f_2(x) & \text{if } x \in S
\end{cases}
\]  

(2)

without invariance assumption. There are three basic types of solutions behavior along \( S \): a transversal motion with solution passing from one region \( M_i \) to another, the sliding mode, and the case of non-uniqueness. Lyapunov functions for such systems are defined and some stability and attraction results for the system (2) are stated. Some applications of the results to the optimal control problem are given.

The author was supported by the Russian Foundation for Basic Research (project no. 11-01-00541).

REFERENCES


Optimal control of nonlinear parabolic equations and the differentiability of the control-state mapping

Simon Serovajsky\(^1\)

\(^1\) al-Farabi Kazakh National University, Almaty, Kazakhstan

serovajskys@mail.ru

Practical methods for solving optimization problems used, as a rule, necessary conditions of optimality or gradient methods. These methods require finding the derivative of the minimizing functional. Therefore we have the necessity to prove the differentiability of the control-state mapping. This property can be substantiated with using Inverse Function
Theorem or Implicit Function Theorem. These results are true whenever the derivative of the state operator is invertible. This assumption and analogues of Lustenik’s condition use in known methods of the general extremum theory. Unfortunately these suppositions are broken for the large class of nonlinear infinite dimensional systems. Moreover the control-state mapping can be non-differentiable in this case. It is true without any nonsmooth terms in the cost functional and the state equation. So the well-known methods of nonsmooth optimization are inapplicable for these problems. We will overcome these difficulties by means of the extended differentiation theory.

Let $\Omega$ be an open bound set of the space $\mathbb{R}^n$ with boundary $\Gamma$, $T > 0$, $Q = \Omega \times (0, T)$, $\Sigma = \Gamma \times (0, T)$. Consider nonlinear parabolic equation

$$y' - \Delta y + |y|^\rho y = v_Q + f_Q, \ (x, t) \in Q$$

with boundary conditions

$$y = 0, \ (x, t) \in \Sigma,$$

$$y(x, 0) = v_\Omega + f_\Omega, \ x \in \Omega,$$

where $y' = \partial y/\partial t$, $\rho > 0$. Determine $W = L_2(0, T; H^1_0(\Omega))$ and its adjoint space $W' = L_2(0, T; H^{-1}(\Omega))$. Consider the spaces $X = W \cap L_q(Q)$ and $X' = W' + L_{q'}(Q)$, where $q = \rho + 2$, $1/q + 1/q' = 1$. The functions $f_Q \in X'$ and $f_\Omega \in L_2(\Omega)$ are known. The pair $v = (v_Q, v_\Omega)$ from the space $V = L_2(Q) \times L_2(\Omega)$ is chosen as a control. By monotone operators theory (see [1], Chapter VI, Theorem 1.1), for any $v \in V$ the problem (1) – (3) has a unique solution $y = y[v]$ from the space $Y = \{y | y \in W, y' \in W'\}$, besides the map $y[.] : V \rightarrow Y$ is weakly continuous.

Consider the functional

$$I(v, y) = \int_Q F_Q(\xi, y(\xi), \nabla y(\xi), v_Q(\xi))dQ + \int_\Omega F_\Omega(x, y(x, T), v_\Omega(x))d\Omega,$$

where $F_Q$ and $F_\Omega$ are known functions. It is given the set of admissible controls $U = U_Q \times U_\Omega$, where $U_Q$ and $U_\Omega$ are nonempty convex closed bound subsets of the spaces $L_2(Q)$ and $L_2(\Omega)$. Determine the functional $J(v) = I(v, y[v])$.

Problem P. Minimize the functional $J$ on the set $U$. 

230
The solvability of this problem is guaranteed by the following result.

**Theorem 1.** Suppose $F_Q : Q \times \mathbb{R}^{n+2} \to \mathbb{R}$ and $F_\Omega : \Omega \times \mathbb{R}^2 \to \mathbb{R}$ are Carathéodory functions, which satisfy inequalities

$$F_Q(\xi, \varphi, \psi) \geq \Phi(|\psi|) \quad \forall \xi \in Q, \varphi \in \mathbb{R}, \psi \in \mathbb{R}^{n+1},$$

$$F_\Omega(x, \eta) \geq \Psi(|\eta|) \quad \forall x \in \Omega, \eta \in \mathbb{R}^2,$$

where $\Phi : \mathbb{R}_+ \to \mathbb{R}_+$ and $\Psi : \mathbb{R}_+ \to \mathbb{R}_+$ are increasing convex coercive semicontinuous functions; besides $F_Q(\xi, \varphi, .)$ and $F_\Omega(x, .)$ are convex on the sets $\mathbb{R}^{n+2}$ and $\mathbb{R}^2$ for all $\xi \in Q, \varphi \in \mathbb{R}, x \in \Omega$. Then Problem $P$ is solvable.

The necessary condition of the minimum of the functional $J$ on the set $U$ at the point $u$ is the variational inequality

$$J'(u)(v - u) \geq 0 \quad \forall v \in U,$$

(4)

where $J'(u)$ is the derivative of $J$ at the point $u$. The application of this result for our problem requires the justification of the differentiability of the control-state mapping. However we have the following assertion.

**Lemma 1.** The map $y[.] : \left( X' \times L_2(\Omega) \right) \to \mathbb{Y}$ is not Gateaux differentiable for large enough $\rho$ and $n$.

Thus necessary conditions of the optimality for Problem $P$ can be obtained only for small enough values of the nonlinearity parameter $\rho$ and the dimension $n$ of the set. The differentiability of the control-state mapping can be proved actually with using of Inverse Function Theorem for this case. Note that the known results for optimization control systems described by nonlinear parabolic equations use small enough values of the nonlinearity parameter and the dimension of the set only (see for example, [2], Chapter IV, Theorem 2.6; [3], Chapter 1, Theorem 3.2; [4], Chapter 2, Theorem 8.1). However by Theorem 1, our problem is solvable for all values of these parameters. So it will be interesting to analyze it without any parameters constraints. The desirable result can be obtained by means of an extension of operator derivative [5].

**Definition.** An operator $L : V \to Y$, where $V$ and $Y$ are linear topological spaces, is called $(V_0,Y_0;V_*,Y_*)$-extended differentiable
at a point $u \in V$, if there exist linear topological spaces $V_0, Y_0, V_*, Y_*$
with continuous embeddings $V_* \subset V_0 \subset V, Y \subset Y_0 \subset Y_*$ and a linear
continuous operator $D : V_0 \to Y_0$ such that $[L(u + \sigma h) - Lu]/\sigma \to Dh$
in $Y_*$ for all $h \in V_*$ as $\sigma \to 0$.

For any point $u \in V$ determine the spaces

$$X_0(u) = \{y | y \in W, |y[u]|^\rho y \in L_2(Q)\},$$

$$X'_0(u) = \{\mu + |y[u]|^{\rho/2} \eta | \mu \in W', \eta \in L_2(Q)\},$$

$$Y_0(u) = \{y | y \in X_0(u) y' \in X'_0(u)\}.$$

**Lemma 2.** The map $y[\cdot] : V \to Y$ is $(V, Y_0(u); V, W)$-extended differ-
entiable at the arbitrary point $u \in V$. Its extended derivative $y'[u]$ satisfies
the equality

$$\int_Q \mu Q y'[u]hdQ + \int_\Omega \mu \Omega (y'[u]h)|_{t=T}d\Omega$$

$$= \int_Q p[\mu]hdQ + \int_\Omega p[\mu]|_{t=0}h_\Omega d\Omega \forall \mu \in (X'_0(u) \times L_2(Q)),$$

where $\mu = (\mu_Q, \mu_\Omega)$, $p[\mu]$ is the solution of the boundary problem

$$-p[\mu'] - \Delta p[\mu] + (\rho + 1)|y[u]|^\rho p[\mu] = \mu_Q, (x, t) \in Q,$$

$$p[\mu] = 0, (x, t) \in \Sigma,$$

$$p[\mu](x, T) = \mu_\Omega, x \in \Omega.$$

Besides $\{(|y[u + \sigma h]|_{t=T} - y[u]|_{t=T})/\sigma \to (y'[u]h)|_{t=T} \in L_2(\Omega) \text{ as } \sigma \to 0$.

Then we can prove the differentiability of the minimizing functional.

**Lemma 3.** Suppose the assumptions of Theorem 1 are true, and
the functions $F_Q(\xi, .)$ and $F_\Omega(x, .)$ have continuous derivatives $F_{Q0}, ..., F_{Q(n+1)}$ and $F_{\Omega 1}, F_{\Omega 2}$; besides

$$|F_Q(\xi, \eta)| \leq a_Q(\xi) + b_Q \sum_{i=0}^{n+1} |\eta_i|^2, |F_{Qj}(\xi, \eta)| \leq a_{Qj}(\xi) + b_{Qj} \sum_{i=0}^{n+1} |\eta_i|,$$
\[ |F_{\Omega}(x, \zeta)| \leq a_{\Omega}(x) + b_{\Omega}(|\zeta_1|^2 + |\zeta_2|^2), \quad |F_{\Omega l}(x, \zeta)| \leq a_{\Omega l}(x) + b_{\Omega l}(|\zeta_1| + |\zeta_2|) \]

for all \( \eta \in \mathbb{R}^{n+2}, \zeta \in \mathbb{R}^2, \xi \in Q, x \in \Omega, \) where \( a_Q \in L_1(Q), a_{Q_j} \in L_2(Q), \)
\( a_{\Omega} \in L_1(\Omega), a_{\Omega l} \in L_2(\Omega), b_Q > 0, b_{Q_j} > 0, b_{\Omega l} > 0, j = 0, \ldots, n + 1, \)
\( l = 1, 2. \) Then the functional \( J \) has Gateaux derivative at the arbitrary point \( u \in V \) such that

\[
J'(u)h = \int_{Q} (F_{Q_l}u + p)h_Q dQ + \int_{\Omega} (F_{\Omega l}u + p|_{t=T})h_{\Omega} d\Omega \quad \forall h \in V,
\]

where \( p \) is the solution of the adjoint system

\[
-p' - \Delta p + (\rho + 1)|y[u]|^\rho p = F_{Qy} - \text{div} F_{Q\nabla y}, (x, t) \in Q,
\]

\[
p = 0, (x, t) \in \Sigma,
\]

\[
p(x, T) = F_{\Omega y}, x \in \Omega.
\]

Besides

\[
F_{Qy}(\xi) = F_{Q0}(\xi, y[u](\xi), y[u](\xi), u_Q(\xi)),
\]

\[
F_{Q\nabla y}(\xi) = \left( F_{Qi}(\xi, y[u](\xi), y[u](\xi), u_Q(\xi)) \right), \quad i = 1, \ldots, n,
\]

\[
F_{Q_l}u(\xi) = F_{Q(n+1)}(\xi, y[u](\xi), y[u](\xi), u_Q(\xi)),
\]

\[
F_{\Omega y}(x) = F_{\Omega_1}(x, y[u](x, T), u_{\Omega}(x)),
\]

\[
F_{\Omega l}u(x) = F_{\Omega_2}(x, y[u](x, T), u_{\Omega}(x)).
\]

Using Lemma 3 and the formula (4), we get necessary conditions of optimality.

**Theorem 2.** Under the conditions of Lemma 3, the optimal control \( u = (u_Q, u_{\Omega}) \) satisfies the variational inequalities

\[
\int_{Q} (F_{Q} + p)(v_Q - u_Q) dQ \geq 0 \quad \forall v_Q \in U_Q,
\]

\[
\int_{\Omega} (F_{\Omega} + p|_{t=T})(v_{\Omega} - u_{\Omega}) d\Omega \geq 0 \quad \forall v_{\Omega} \in U_{\Omega}.
\]
Thus necessary conditions of optimality for Problem $P$ involve the state system (1) – (3) for $v = u$, adjoint system and variational inequalities.

REFERENCES


Inverse problems for parabolic equations with infinite horizon

Iliyas Shakenov$^1$

$^1$ al-Farabi Kazakh National University, Almaty, Kazakhstan;
ilias.shakenov@gmail.com

Practical problems often lead to difficulty that we call inverse problems. For example, if you need to know the temperature of soil at the depth of several meters while it is possible to measure the temperature only on the surface. In this type of problems there is a lot of information (over-determination) on one side of the boundary but no any data at the other side.

We consider one of such problems with the following mathematical problem definition.

Let we are given a one-dimensional heat conduction equation and
initial-boundary problem

\[ u_t = u_{xx} + f(x, t), \quad 0 < x < L, \quad 0 < t < \infty, \quad (1) \]

\[ u|_{t=0} = \varphi(x), \quad (2) \]

\[ u_x|_{x=0} = b(t), \quad (3) \]

\[ u_x|_{x=L} = y(t). \quad (4) \]

Right boundary function \( y(t) \) is unknown and has a meaning of heat flow. To determine this value we can use an additional information \( u(0, t) = a(t) \). To solve the problem numerically we cannot use infinite time interval, so we replace the original problem with its finite approximation. That is, \( 0 < t < T \), where \( T \) becomes larger and larger. We convert this problem to optimization, in which we are required to minimize the functional on (1) – (4):

\[
I(y) = \int_0^T \left( u(0, t; y) - a(t) \right)^2 dt \to \min
\]

If we find control \( y(t) \) which gives a zero value to the functional then the additional information is fulfilled.

We use one familiar method to solve the problem by constructing an iterative process.

\[ y_{n+1}(t) = y_n(t) - \alpha_n I'(y_n(t)) , \]

where \( \alpha_n > 0 \).

The first problem is to determine what is \( I' \) and how to find it? Using an apparatus of Gato derivative we derive the theorem, [1]:

**Theorem.** The Gato derivative of functional \( I \) in point \( y(t) \) is equal to \( \psi(L, t) \), where \( \psi(x, t) \) is a solution of adjoint problem

\[
\psi_t + \psi_{xx} = 0, \quad (5)
\]

\[ \psi(x, T) = 0, \quad (6) \]

\[ \psi_x(L, t) = 0, \quad (7) \]
\[ \psi_x(0, t) = -2(u(0, t; y) - a(t)). \] (8)

**Solution algorithm.**

1. **Parameters setting.** Such as deviation of functional from zero, step by time and step by \(x\)-axis, initialization of known functions. Choose initial approximation \(y_0(t)\).

2. **Direct problem solution.** We solve (1) – (4) and get \(u(0, t; y)\).

3. **Calculation of the functional.** If \(I(y) < \varepsilon\), then the algorithm terminates and results are shown. If \(I(y) > \varepsilon\), then go to item 4.

4. **Adjoint problem solution.** Here we get \(I'(y) = \psi(L, t)\).

5. **We choose \(\alpha\)-value.** It might be a constant or it can be calculated by specific formulas, [2].

6. **Next approximation construction.** We use the mentioned formula:

\[ y_{n+1}(t) = y_n(t) - \alpha_n I'(y_n(t)). \]

Go to item 2.

Detailed description of all algorithm steps are given in the work, [3].

Once we have constructed a computational algorithm, we can do a lot of experiments (accompanied with graphs and tables) and draw conclusions about the following issues:

1. First of all, having in hand exact solution, we can compare it with the derived by computational algorithm.

2. We analyze an influence of steps by time and \(x\)-axis, closeness of functional to zero on the solution of the problem.

3. The inverse problems like this has an effect of insolvability on the right-hand side of time-interval. We consider how it might be resolved and of course we touch the problem about in what fraction of entire time-interval we can get sufficient accuracy.

4. Different ways to chose parameter \(\alpha\) has been applied and compared. Some of them – traditional methods – turned out very inefficient.

5. We elongate the time-interval and observe solutions behavior. Also we examine some more methods of constructing successive approximations \(y_{n+1}(t)\) when time increases and compare them with each other. The main idea is in calculating the solution on small interval and filling the rest values with zero. This solution is used as initial approximation when time interval becomes larger.
6. Attempts to transform an equation using small parameter are done and the results compared with original method.

7. Theoretical researches about convexity of the functional, uniqueness of the solution, applicability of the maximum principle are also submitted.

Additionally we give an outline of the future works to improve and develop this research, which we believe has a great applicability and concerns about new unknown effects in inverse problem theory.

REFERENCES


Solution of the parametric inverse problem of stochastic optimal control

Kanat Shakenov

1 al-Farabi Kazakh National University, Almaty, Kazakhstan;
shakenov2000@mail.ru

Problem definition.

According to [1] we consider a problem of estimation of unknown density parameter when additional control parameter presents and can be chosen randomly. Let \( f(y|x, z) \) is a distribution density with respect to measure \( \mu(dy) \) on a line with dependence on parameters \( x \) and \( z \). Let assume that \( x, y, z \) are one–dimensional: \( x \in X \subset E_1, y \in E_1, z \in Z \subset E_1 \), where \( E_1 \) is Euclid space with dimension of 1. Parameter \( x = x_0 \) is unknown and must be estimated by means of samples with density \( f \). Parameter \( z \) has a meaning of control and it can be chosen in two ways:
first, it can be constant; second, \(i\)-th observation implies the following value:

\[
Z_i = Z_i(Y_1, \ldots, Y_{i-1}) = Z_i(Y^{(i-1)}).
\]

(1)

More precisely about (1): simultaneous distribution of \(Y_1, \ldots, Y_n\) can be determined by formula

\[
f_n(Y^{(n)}, x_0) = \prod_{i=1}^{n} f(y_i|x_0, z_i(Y^{(i-1)})),
\]

where \(Y^{(n)} = (y_1, \ldots, y_n)\).

We shall denote \(Z_i = Z_i(Y_1, \ldots, Y_{i-1}) = Z_i(Y^{(i-1)})\) and call it as acceptable control or plan of experiment or just plan if a sequence \(Z_1, Z_2, \ldots\) satisfies (1) and such that \(Z_i \in Z, \ i = 1, 2, \ldots\) The problem is to chose such a control (plan) which provides the best quality of estimation in sense of a given quality criteria. We take this criteria as \(E(x_n - x_0)^2\) for a while.

**Solution existence. Problems.**

Let for all \(x \in X, z \in Z\) information amount \(I(x, z) = \int \left( \frac{\partial f(y|x, z)}{f(y|x, z)} \right)^2 \nu(dy)\) of density \(f\) exists, where \(f(y|x, z)\) is a PDF with respect to measure \(\nu(dy)\) on line (\(x\) and \(z\) are parameters). Assume that \(E_x(x_n - x)^2 \geq \left( E_x \sum_{i=1}^{n} I(x, Z_i(Y^{(i-1)})) \right)^{-1}\) and \(I_{\sup}(x) = \sup_{z \in Z} I(x, z) < \infty, \ x \in X,\)

where \(Z_i = Z_i(Y_1, \ldots, Y_{i-1}) = Z_i(Y^{(i-1)})\) is a sample. It is proven that for any experiment plan the following inequality is true \(E_x(x_n - x)^2 \geq \frac{1}{n I_{\sup}(x)}\) and asymptotically optimal plan \(Z_1, \ldots, Z_n, \ldots\) exists. For this sequence of controls there is a sequence of estimations \(x_n = x_n(Y^{(n)})\) such that \(\sqrt{n}(x_n - x) \sim \mathcal{N}(0, \frac{1}{I_{\sup}(x_0)}), \ n \to \infty\). Indicated problem can have an exact solution if estimating parameter has probability distributions. Solution of such a problem lies in the theory of statistical solutions. (See for [2] or [3] and others). But this solution not always good for practice. We need in a priori distribution. It’s a difficult problem if a sample size \(n\) is small. But for our case we can neglect it as we construct estimations for \(n \to \infty\). For quite common assumptions Bayesian estimation with any prior density \(\pi(x), \ (\pi(x) \neq 0, \ x \in X)\), is asymptotically optimal if \(x_0\) is not random value. You can see [4] for this purposes. There is a much more difficult problem. If you want to obtain more precise solution, it will
take a lot of numerical computations. This problem is considered in [3], chapter IV.

**Problem of estimating with independent samples.**

Let’s consider a problem of estimation without control. Let \( Y_1, Y_2, \ldots, Y_n \) are independent random vectors (sample) from \( E_i \), each of which has a distribution density of \( f(y, x) \) with respect to \( \sigma \)-finite measure \( \nu(\cdot) \). Parameter \( x \) takes values from open set \( X \in E_k \).

**Problems.**

By values of \( Y_1, Y_2, \ldots, Y_n \) and \( f(y, x) \) find the value of \( x \). We have no any prior information about \( x \), except of \( x \in X \). We try to construct an estimation \( x_n(Y_1, Y_2, \ldots, Y_n) \) of \( x \). \( N_n \) is any measurable random value, where \( N_n \) is minimal \( \sigma \)-algebra of events, and all the values of \( Y_1, Y_2, \ldots, Y_n \) are measurable with respect to it.

Let for all \( x \in X, z \in Z \) information amount \( I(x, z) = \int \left( \frac{\partial f(y|x, z)}{\partial x} \right)^2 f(y|x, z) \nu(dy) \) of density \( f \) exists, where \( f(y|x, z) \) is a PDF with respect to measure \( \nu(dy) \) on line ( \( x \) and \( z \) are parameters), [5]. Assume that \( E_x(x_n - x)^2 \geq \left( E_x \sum_{i=1}^{n} I(x, Z_i(Y^{(i-1)})) \right)^{-1} \) and \( I_{sup}(x) = \sup_{z \in Z} I(x, z) < \infty, x \in X \), where \( Z_i = Z_i(Y_1, \ldots, Y_{i-1}) = Z_i(Y^{(i-1)}) \) is a sample. It is proven that for any experiment plan the following inequality is true \( E_x(x_n - x)^2 \geq \frac{1}{n I_{sup}(x)} \) and asymptotically optimal plan \( Z_1, \ldots, Z_n, \ldots \) exists. For this sequence of controls there is a sequence of estimations \( x_n = x_n(Y^{(n)}) \) such that \( \sqrt{n}(x_n - x) \sim N(0, \frac{1}{I_{sup}(x_0)}) \), \( n \to \infty \).

**References**


239
Graph-based local elimination algorithms for sparse discrete optimization problems

O. Shcherbina¹, A. Sviridenko²

¹ University of Vienna, Vienna, Austria; oleg.shcherbina@univie.ac.at
² Tavrian National University, Simferopol, Ukraine; d2rk@firejet.org

1. Introduction
The use of discrete optimization (DO) models and algorithms makes it possible to solve many practical problems, since the discrete optimization models correctly represent the nonlinear dependence, indivisibility of objects, consider the limitations of logical type and all sorts of technology requirements, including those that have qualitative character. To meet the challenge of solving large scale DO problems (DOPs) in reasonable time, there is an urgent need to develop and study new decomposition approaches. Among decomposition approaches appropriate for solving sparse DO problems we mention local elimination algorithms (LEA) [1], [2], which can exploit sparsity in the interaction graph of a DOP and allow to compute a solution in stages such that each of them uses results from previous stages. LEAs compute global information using local computations (i.e., computations of information about elements of neighborhoods of variables or constraints - usually, solving subproblems).

The algorithmic scheme of LEA is defined by an elimination tree [3] whose vertices are associated with subproblems and whose edges express information on interdependence between subproblems. The structure of a DO problem can be defined either by an interaction graph of initial elements (variables and constraints), or by various derived structures, e.g., block structures, block-tree structures defined by a so called condensed graph.

There are various computational schemes for realizing LEA, including the LEA elimination of variables, block-elimination algorithm, LEA based on tree decomposition.

2. Block local elimination scheme
2.1. Partitions, clustering, and quotient graphs
Consider the integer linear programming (ILP) problem with binary variables
\[ f(X) = CX = \sum_{j=1}^{n} c_j x_j \rightarrow \max \]

subject to constraints

\[ \sum_{j=1}^{n} a_{ij} x_j \leq b_i, i = 1, 2, \ldots, m, \]

\[ x_j = 0, 1, j = 1, 2, \ldots, n. \]

The local elimination procedure can be applied to elimination of not only separate variables but also to sets of variables and can use the so called elimination of variables in blocks, which allows to eliminate several variables in block.

Applying the method of merging variables into meta-variables allows to obtain condensed or meta-DOPs which have a simpler structure. If the resulting meta-DOP has a nice structure (e.g., a tree structure) then it can be solved efficiently.

An ordered partition of a set \( X \) is a decomposition of \( X \) into ordered sequence of pairwise disjoint nonempty subsets whose union is all of \( X \).

In general, graph partitioning is NP-hard. Since graph partitioning is difficult in general, there is a need for approximation algorithms. A popular algorithm in this respect is \textbf{MeTiS}\textsuperscript{1}, which has a good implementation available in the public domain.

An important special case of partitions are so-called blocks. Two variables are indistinguishable if they have the same closed neighborhood. A block is a maximal set of indistinguishable vertices. The blocks of \( G \) partition \( X \) since indistinguishability is an equivalence relation defined on the original vertices. The corresponding graph is called condensed graph, which is a merged form of original graph.

An equivalence relation on a set induces a partition on it, and also any partition induces an equivalence relation. Given a graph \( G = (X, E) \), let \( X \) be a partition on the vertex set \( X : X = x_1, x_2, \ldots, x_p, \ p \leq n \), where \( x_l = X_{K_l} \) (\( K_l \) is a set of indices corresponding to \( x_l, \ l = 1, \ldots, p \)). For this ordered partition \( X \), the DOP (1) – (3) can be solved by the LEA using quotient interaction graph \( G \).

\textsuperscript{1}http://www-users.cs.umn.edu/~karypis/metis

241
That is, $\bigcup_{i=1}^{p} x_i = X$ and $x_i \cap x_k = \emptyset$ for $i \neq k$. We define the quotient graph of $G$ with respect to the partition $X$ to be the graph

$$G = G/X = (X, E),$$

where $(x_i, x_k) \in E$ if and only if $Nb_G(x_i) \cap x_k \neq \emptyset$.

The quotient graph $G(X, E)$ is an equivalent representation of the interaction graph $G(X, E)$, where $X$ is a set of blocks (or indistinguishable sets of vertices), and $E \subseteq X \times X$ be the edges defined on $X$. A local block elimination scheme is one in which the vertices of each block are eliminated contiguously. As an application of a clustering technique we consider below a block local elimination procedure where the elimination of the block (i.e., a subset of variables) can be seen as the merging of its variables into a meta-variable.

2.2. Block local elimination algorithm

A. Forward part

Consider first the block $x_1$. Then

$$\max_X \{C_NX_N|A_{iS_i}X_{S_i} \leq b_i, \ i \in M, \ x_j = 0, 1, \ j \in N\} =$$

$$\max_{X_{K_2}, \ldots, X_{K_p}} \{C_{N-K_1}X_{N-K_1} + h_1(Nb(X_{K_1}))|A_{iS_i}X_{S_i} \leq b_i, \ i \in M - U_1, \ x_j = 0, 1, \ j \in N - K_1\}$$

where $U_1 = \{i : S_i \cap K_1 \neq \emptyset\}$ and

$$h_1(Nb(X_{K_1})) = \max_{X_{K_1}} \{C_{K_1}X_{K_1}|A_{iS_i}X_{S_i} \leq b_i, \ i \in U_1, \ x_j = 0, 1, \ x_j \in Nb[x_1]\}.$$  

The first step of the local block elimination procedure consists of solving, using complete enumeration of $X_{K_1}$, the following optimization problem

$$h_1(Nb(X_{K_1})) = \max_{X_{K_1}} \{C_{K_1}X_{K_1}|A_{iS_i}X_{S_i} \leq b_i, \ i \in U_1, \ x_j = 0, 1, \ x_j \in Nb[x_1]\},$$

and storing the optimal local solutions $X_{K_1}$ as a function of the neighborhood of $X_{K_1}$, i.e., $X_{K_1}^*(Nb(X_{K_1}))$.

The maximization of $f(X)$ over all feasible assignments $Nb(X_{K_1})$, is called the elimination of the block (or meta-variable) $X_{K_1}$. The optimization problem left after the elimination of $X_{K_1}$ is:

$$\max_{X - X_{K_1}} \{C_{N-K_1}X_{N-K_1} + h_1(Nb(X_{K_1}))|A_{iS_i}X_{S_i} \leq b_i, \ i \in M - U_1, \ x_j \in Nb[x_1]\}.$$  

242
Note that it has the same form as the original problem, and the tabular function \( h_1(Nb(X_{K_1})) \) may be considered as a new component of the modified objective function. Subsequently, the same procedure may be applied to the elimination of the blocks – meta-variables \( x_2 = X_{K_2}, \ldots, x_p = X_{K_p} \), in turn. At each step \( j \) the new component \( h_{x_j} \) and optimal local solutions \( X^*_j \) are stored as functions of \( Nb(X_{K_j} \mid X_{K_1}, \ldots, X_{K_{j-1}}) \), i.e., the set of variables interacting with at least one variable of \( X_{K_j} \) in the current problem, obtained from the original problem by the elimination of \( X_{K_1}, \ldots, X_{K_{j-1}} \). Since the set \( Nb(X_{K_p} \mid X_{K_1}, \ldots, X_{K_{p-1}}) \) is empty, the elimination of \( X_{K_p} \) yields the optimal value of objective \( f(X) \).

B. Backward part.

This part of the procedure consists of the consecutive choice of \( X^*_{K_p}, X^*_{K_p-1}, \ldots, X^*_{K_1} \), i.e., the optimal local solutions from the stored tables \( X^*_{K_1}(Nb(X_{K_1})), X^*_{K_2}(Nb(X_{K_2} \mid X_{K_1})), \ldots, X^*_{K_p} \mid X_{K_p}, \ldots, X_{K_1} \).

Underlying DAG of the local block elimination procedure contains nodes corresponding to computing of functions \( h_{x_i}(Nb_{G^{(i-1)}}(x_i)) \) and is a generalized elimination tree.

3. Comparative computational experiment

Among extremely important research questions about the effectiveness of local elimination algorithms (LEA), the next one causes special interest: “Is the use of LEA in combination with a discrete optimization (DO) algorithm (for solving problems in the blocks) consistently more efficient than the standalone use of the DO algorithm?” [4].

The computational capabilities of the LEA in combination with a modern solver were tested by using SYMPHONY\(^2\) as the implementation framework. SYMPHONY is part of the COIN-OR\(^3\) project and it can solve mixed-integer linear programs (MILP) sequentially or in parallel. We chose this framework since it is open-source and supports warm restarts, which implement postoptimal analysis (PA) of ILP problems.

All experimental results were obtained on an Intel Core 2 Duo at 2.66 GHz machine with 2 GB main memory, and running Linux, version 2.6.35-24-generic. SYMPHONY 5.4.1\(^4\) was used for the LEA implementation. The maximum solving time is denoted by \( TIMEOUT \), and is equal to 2

\(^2\)https://projects.coin-or.org/SYMPHONY  
\(^3\)http://www.coin-or.org  
\(^4\)http://www.coin-or.org/download/source/SYMPHONY/
hours. All the ILP problems with binary variables from a given experiment have artificially generated quasi-block structures. All the blocks from a single problem have the same number of variables, and also the same number of variables in separators between them. This is required in order to evaluate the impact of the PA on the time to solve the problem by increasing the number of variables.

The test problems were generated by specifying the number of variables, the number of constraints and the size of the separators between blocks. The objective function and constraint matrix coefficients, and the right-hand sides for each of the block were generated by using a pseudorandom-number generator.

Each test problem was solved by using three algorithms, a) the basic MILP SYMPHONY solver with the OsiSym interface, b) the LEA in combination with SYMPHONY, c) the LEA in combination with SYMPHONY and with PA (warm restarts). In all the cases SYMPHONY used preprocessing.

The computational experiments are described in details in [5] and show that LEA combined with SYMPHONY for solving quasi-block problems with small separators outperforms the stand alone SYMPHONY solver. Additionally, by increasing the size of the separators in the problems for the same number of variables and block sizes LEA becomes less efficient due to the increased number of iteration for solving the block subproblems.

References

Heuristic Algorithms for a Job-Shop Problem with Minimizing Total Job Tardiness

Yuri N. Sotskov\textsuperscript{1}, Omid Gholami\textsuperscript{1}, Frank Werner\textsuperscript{2}

\textsuperscript{1} United Institute of Informatics Problems, 220012 Minsk, Belarus; sotskov@newman.bas-net.by, gholamiiran@yahoo.com
\textsuperscript{2} Fakultät für Mathematik, Otto-von-Guericke-Universität Magdeburg, PSF 4120, 39016 Magdeburg, Germany; frank.werner@ovgu.de

In practice, it is often required to process a set of jobs without operation preemptions satisfying temporal and resource constraints. Temporal constraints say that some jobs have to be finished before some others can be started. Resource constraints say that operations processed on the same machine cannot be processed simultaneously. The objective is to construct a schedule specifying when each operation starts such that both temporal and resource constraints are satisfied and the given objective function has a minimum value. One can model such a scheduling process via the following job-shop problem.

There are \( n \) jobs \( \mathcal{J} = \{J_1, J_2, \ldots, J_n\} \) to be processed on \( m \) machines \( \mathcal{M} = \{M_1, M_2, \ldots, M_m\} \). Operation preemptions are not allowed, and the machine routes \( O^i = (O_{i1}, O_{i2}, \ldots, O_{in_i}) \) for processing the jobs \( J_i \in \mathcal{J} \) may be given differently for different jobs. The time \( p_{ij} > 0 \) needed for processing an operation \( O_{ij} \) of a job \( J_i \) on the corresponding machine \( M_v \in \mathcal{M} \) is known before scheduling. A job \( J_i \in \mathcal{J} \) is available for processing from time-point \( r_i \geq 0 \). The time-point \( d_i > r_i \) defines a due date for completing job \( J_i \). It is assumed that machine \( M_k \in \mathcal{M} \) can process a job \( J_i \in \mathcal{J} \) at most once. Consequently, any two operations \( O_{ij} \) and \( O_{ik}, j \neq k, \) of the same job \( J_i \in \mathcal{J} \) have to be processed by different machines and \( n_i \leq m \) (such a scheduling problem is called a classical job-shop). We consider the objective of finding a schedule minimizing the sum \( \sum_{i=1}^{n} T_i \) of the tardiness times \( T_i = \max\{0, C_i - d_i\} \) for the jobs \( J_i \in \mathcal{J} \). Hereafter, \( C_i \) denotes the completion time of a job \( J_i \in \mathcal{J} \). According to the three-field notation \( \alpha|\beta|\gamma \) used for machine scheduling problems, the above problem is denoted as \( J|r_i|\sum T_i \).

Problem \( J|r_i|\sum T_i \) arises, e.g., in train scheduling for a single-track railway network: To determine a schedule for a set of trains that does not violate the single-track capacities and the train timetable. In a single-
track railway, a pair of sequential stations can be connected by a single-track (railroad section) only. Specifically, this is the case for most railway networks in countries of the Middle East.

In a job-shop approach to train scheduling, trains and railroad sections are synonymous with the jobs \( J \) and the machines \( M \), respectively. An operation \( O_{ij} \) is regarded as a movement of the train \( J_i \in J \) across the railroad section \( M_v \in M \), where machine \( M_v \) has to process operation \( O_{ij} \). The positive number \( p_{ij} \) denotes the time required for train \( J_i \in J \) to travel through section \( M_v \in M \). The non-negative number \( r_i \) denotes the earliest possible departure time for the train (release time of the job) \( J_i \in J \) from the original station in the route \( O^i \). The positive number \( d_i \) denotes the official arrival time of the train (due date for completing the job) \( J_i \in J \) at the terminal station in the route \( O^i \). It should be noted that for train scheduling, the inequality \( m > n \) holds and each machine \( M_k \in M \) can process a job \( J_i \in J \) at most once.

Our aim was to find an algorithm for the problem \( J|r_i|\sum T_i \) to be fast even for a large size of the input data (this is the case for a real-world railway scheduling problem). It is clear that an exact branch and bound method creates a lot of branches in the solution tree for a large input data, and so it is not possible to use a branch and bound method for most real-world job-shop problems with large sizes. Heuristic methods like a genetic algorithm are basically rather slow. Furthermore, using algorithms like Lagrangian relaxation or simulated annealing can reduce the computational time only a bit. A lot of methods like tabu search need many calculations and cannot satisfy our aim as well.

Problem \( J|r_i|\sum T_i \) is very complicated in the computational sense since even its special cases belong to the class of binary (or unary) NP-hard problems [1]. In order to achieve a practical size of a classical job-shop problem, which can be solved heuristically within a reasonable time, we first coded a shifting bottleneck algorithm, which was originated in [2] for a job-shop problem \( J||C_{max} \) with the makespan criterion \( C_{max} = \max\{C_i : J_i \in J\} \). We tested the program realizing the shifting bottleneck algorithm for the problem \( J|r_i|\sum T_i \) as one of the most famous heuristic algorithms for job-shop problems [2] (this algorithm was improved in [3]). However, we obtained unsatisfactory large CPU-times for randomly generated instances \( J|r_i|\sum T_i \) with large and even moderate numbers \( m \) of machines provided that \( m > n \). Therefore, we were forced to look for other heuristic algorithms for the problem \( J|r_i|\sum T_i \), which
will run essentially faster than the shifting bottleneck algorithm and will provide sufficiently close objective function values when \( m > n \) and each machine \( M_v \in \mathcal{M} \) may process at most one operation from the route \( O^i \) of a job \( J_i \in \mathcal{J} \).

We observed that many recursive functions are needed to calculate data like a local due date, a critical path and a local release time for each vertex (i.e., operation) in the mixed graph \( G = (Q, A, E) \) representing a job-shop [4]. In the mixed graph \( G = (Q, A, E) \), the vertex set \( Q \) is the set of all operations including a source operation \( O \) and a sink operation \( O_i \) for each job \( J_i \in \mathcal{J} \). The arc set \( A \) defines temporal constraints given by the routes \( O^i \) of the jobs \( J_i \in \mathcal{J} \). The edge set \( E \) defines resource constraints given by the machine set \( M = \{M_1, M_2, \ldots, M_m\} \) and by the routes \( O^i \) for the jobs \( J_i \in \mathcal{J} \) (see [4] for details).

The earliest start time \( r_{ij} \) of an operation \( O_{ij} \in Q \) may be defined as the length of the longest path from a start vertex \( O \in Q \) to vertex \( O_{ij} \) in the digraph \( (Q, A, \emptyset) \) obtained from the mixed graph \( G \) via deleting all the edges \( E \). We call \( r_{i1} \) the release time of a job \( J_i \) provided that \( O_{i1} \in Q \) is the first operation of a job \( J_i \) in the route \( O^i \). Since the feasible digraph \( (Q, A, \emptyset) \) has no circuits, all the earliest start times \( r_{i1}, i \in \{1, 2, \ldots, n\} \), are finite and can be calculated in linear time of the sum \( |Q| + |A| \).

We focus on the shortest release times of the operations \( Q \) in the following algorithm which is called SRT-algorithm. The shortest release time of an operation is used as a priority (SRT-priority) in the SRT-algorithm. In contrast to the shifting bottleneck algorithm, which examines a bottleneck machine at each iteration [2,3], a critical job is examined at each iteration of the SRT-algorithm.

In the initial step of the SRT-algorithm, the earliest start times \( r_{ij} \) of all operations \( O_{ij} \in Q \) have to be computed due to the following recursion:

\[
    r_{ij} = r_{i,j-1} + p_{i,j-1}.
\]

The release time of the source operation \( O \) is equal to \( \min\{r_i : J_i \in \mathcal{J}\} \).

The first job to be examined is a job \( J_i \in \mathcal{J} \), whose last operation \( O_{in_i} \) in the route \( O^i \) is the next to the last one in the critical path of the digraph \( (Q, A, \emptyset) \). At the first iteration of the SRT-algorithm, the following two steps are realized.

**Step 1.** The SRT-algorithm finds the first request (i.e., operation \( O_{i1} \in Q \)) of job \( J_i \) for the machine \( M_v \in \mathcal{M} \) processing operation \( O_{i1} \). Then the algorithm compares the release time \( r_{i1} \) with the release times of all operations \( O_{ig} \in Q \) of the other jobs on the same machine \( M_v \in \mathcal{M} \).
processing operation $O_{i1}$. To resolve conflicts of jobs for the same machine, the SRT-priority is used as follows. If the release time $r_{i1}$ is not greater than that of operation $O_{jg}$ for the same machine $M_v \in \mathcal{M}$, then the arc $(O_{i1}, O_{jg})$ starting from operation $O_{i1}$ and ending in operation $O_{jg}$ has to be added to the digraph $(Q, A, \emptyset)$. Otherwise, the symmetric arc $(O_{jg}, O_{i1})$ has to be added to the digraph $(Q, A, \emptyset)$. If an arc is added to the digraph $(Q, A, \emptyset)$, some local release times may be changed after the second step of the SRT-algorithm.

**Step 2.** The release time of the destination vertex of an arc and the other vertices related to this vertex must be checked and the corresponding release times must be modified if it is necessary. A job priority will be defined by an arc between two operations requesting the same machine, one is the starting vertex of the arc (let this vertex be $O_{km}$) and the other one is an end vertex of the arc (let it be $O_{ij}$). A new release time of the operation $O_{ij}$ has to be calculated as follows: $r_{ij} := \max\{r_{ij}, r_{km} + p_{km}\}$, where the maximum has to be taken over all arcs $(O_{km}, O_{ij})$ belonging to the digraph already constructed. The above equation must be recursively applied to each vertex of the digraph that has an incoming arc from the vertex $O_{ij}$ until the sink vertex $O_i$. If the release time was not changed, the calculation of the recursive function is stopped.

Steps 1 and 2 are repeated for operation $O_{i2}$, then for operation $O_{i3}$, and so on until operation $O_{in_i}$ of the route $\mathcal{O}^i$. As a result, the mixed graph $G$ is transformed into a mixed graph denoted as $G_i = (Q, A \cup A_i, E \setminus E_i)$.

The second job to be examined is the “second critical” job, i.e., a job $J_u \in \mathcal{J} \setminus \{J_i\}$, whose last operation in the route $\mathcal{O}^u$ has the largest completion time in the digraph $(Q, A \cup A_i, \emptyset)$ among all jobs from the set $\mathcal{J} \setminus \{J_i\}$. So, at the second iteration, steps 1 and 2 are executed for the job $J_u$, the digraph $(Q, A \cup A_i, \emptyset)$, and the mixed graph $G_i$.

The process is continued for the “third critical” job, then for the “fourth critical” job and so on, until all jobs $\mathcal{J}$ have been accounted. As a result, the mixed graph $G$ is transformed into the digraph $G_h = (Q, A \cup A_h, E \setminus E_h)$, where $E \setminus E_h = \emptyset$ and job $J_h$ was examined at the $h$-th iteration. It is easy to convince that using the SRT-priority cannot generate a circuit in the digraphs $(Q, A \cup A_i, \emptyset)$, ..., $(Q, A \cup A_h, \emptyset)$ constructed from the first to the last iterations. A circuit-free digraph $G_h$ uniquely determines a semiactive schedule [4], which may be built via the critical path method in $O(n + |A_h|)$ time.

248
Both the shifting bottleneck algorithm and the SRT-algorithm have been coded in Delphi and tested on a laptop computer. We compared the CPU-time taken by the SRT-algorithm and by the shifting bottleneck algorithm to solve the same randomly generated problems $J|r_i|\sum T_i$ heuristically for different values $n \leq 20$ and $m \leq 20$. The SRT-algorithm runs considerably faster than the shifting bottleneck algorithm if $m > n$. In the experiments, we compared the objective function values for randomly generated problems $J|r_i|\sum T_i$ with different products $n \times m$.

The computational results showed that there is no meaningful difference between the quality of the schedules obtained by the two algorithms tested for randomly generated problems $J|r_i|\sum T_i$. The computational experiments also evaluated the effect of adding either new jobs or new machines to the CPU-time required to solve randomly generated problems $J|r_i|\sum T_i$. When we increased the number of jobs (machines, respectively) of the randomly generated instances $J|r_i|\sum T_i$, the CPU-time needed for the SRT-algorithm increased considerably (very slowly).

The SRT-algorithm used a smaller CPU-time than the shifting bottleneck algorithm especially when the number of machines $m$ was essentially larger than the number of jobs $n$. Since the total tardiness values $\sum T_i$ of the schedules constructed by the shifting bottleneck algorithm and the SRT-algorithm are rather close, we can claim that the SRT-algorithm is a good heuristic for a job-shop problem $J|r_i|\sum T_i$ with large $m > n$.

As observed from computational experiments, the SRT-algorithm is a good choice for the classical job-shop problem $J|r_i|\sum T_i$ when the number of machines is much larger than the number of jobs.

References

Consider the following singular optimization problem

\[ \min \phi(x), \quad (1) \]

subject to

\[ F(x) = 0, \quad (2) \]

where \( F : X \to Y, X, Y \) — Banach spaces, and \( \phi : X \to R, F \in C^{p+1}(X), \phi \in C^2(X) \) and at the solution point \( x^* \) we have

\[ \text{Im}F'(x^*) \neq Y. \quad (3) \]

Elements of \( p \)-regularity theory

Let us recall the basic constructions of \( p \)-regularity theory which is used in solving of singular problems. The construction of the \( p \)-factor-operator. Suppose that the space \( Y \) is decomposed into a direct sum

\[ Y = Y_1 \oplus \cdots \oplus Y_p, \quad (4) \]

where \( Y_1 = \text{Im}F'(x^*), Z_1 = Y \). Let \( Z_2 \) be closed complementary subspace to \( Y_1 \) (we assume that such closed complement exists), and let \( P_{Z_2} : Y \to Z_2 \) be the projection operator onto \( Z_2 \) along \( Y_1 \). By \( Y_2 \) we mean the closed linear span of the image of the quadratic map \( P_{Z_2}F^{(2)}(x^*)[\cdot]^2 \). More generally, define inductively,

\[ Y_i = \text{span} \text{Im}P_{Z_i}F^{(i)}(x^*)[\cdot]^i \subseteq Z_i, \quad i = 2, \ldots, p - 1, \]

where \( Z_i \) is a chosen closed complementary subspace for \( (Y_1 \oplus \cdots \oplus Y_{i-1}) \) with respect to \( Y, i = 2, \ldots, p \), and \( P_{Z_i} : Y \to Z_i \) is the projection
operator onto $Z_i$ along $(Y_1 + \cdots + Y_{i-1})$ with respect to $Y$, $i = 2, \ldots, p$. Finally, $Y_p = Z_p$.

The order $p$ is chosen as the minimum number for which (4) holds. Let us define the following mappings

$$F_i(x) = P_{Y_i}F(x), \quad F_i : X \to Y_i, \quad i = 1, \ldots, p,$$

where $P_{Y_i} : Y \to Y_i$ is the projection operator onto $Y_i$ along $(Y_1 + \cdots + Y_{i-1} + Y_{i+1} + \cdots + Y_p)$ with respect to $Y$, $i = 1, \ldots, p$.

\textbf{Definition 1.} The linear operator $\Psi_p(h) \in \mathcal{L}(X, Y_1 + \cdots + Y_p)$, $h \in X$, $h \neq 0$,

$$\Psi_p(h) = F'_1(x^*) + F''_2(x^*)h + \cdots + F^{(p)}_p(x^*)[h]^{p-1},$$

is called the $p$-factor operator.

\textbf{Definition 2.} We say that the mapping $F$ is $p$-regular at $x^*$ along an element $h$, if

$$\text{Im}\Psi_p(h) = Y.$$

\textbf{Remark.} The condition of $p$-regularity of the mapping $F(x)$ at the point $x^*$ along $h$ is equivalent to

$$F^{(p)}_p(x^*)[h]^{p-1} \circ \text{Ker}\Psi_{p-1}(h) = Y_p.$$

\textbf{Definition 3.} We say that the mapping $F$ is $p$-regular at $x^*$ if it is $p$-regular along any $h$ from the set

$$H_p(x^*) = \left\{ \bigcap_{k=1}^{p} \text{Ker}^k F^{(k)}_k(x^*) \right\} \setminus \{0\},$$

where $k$-kernel of the $k$-order mapping $F^{(k)}_k(x^*)$ is as follows

$$\text{Ker}^k F^{(k)}_k(x^*) = \{ \xi \in X : F^{(k)}_k(x^*)[\xi]^k = 0 \}.$$

For a linear surjective operator $\Psi_p(h) : X \mapsto Y$ between Banach spaces we denote by $\{\Psi_p(h)\}^{-1}$ its \textit{right inverse}. Therefore $\{\Psi_p(h)\}^{-1} : Y \mapsto 2^X$ and we have

$$\{\Psi_p(h)\}^{-1}(y) = \{ x \in X : \Psi_p(h)x = y \}.$$
We define the norm of \( \{ \Psi_p(h) \}^{-1} \) via the formula

\[
\| \{ \Psi_p(h) \}^{-1} \| = \sup_{y} \inf \{ \| x \| : x \in [\Psi_p(h)]^{-1} (y) \}.
\]

We say that \( \{ \Psi_p(h) \}^{-1} \) is bounded if \( \| \{ \Psi_p(h) \}^{-1} \| < \infty \).

The following theorem gives a description of a solution set in degenerate case.

**Theorem 1 (Generalized Lyusternik Theorem).** Let \( X \) and \( Y \) be Banach spaces and \( U \) be a neighborhood of \( x_0 \in X \). Assume that \( F : X \to Y, \quad F \in C^p(U) \) is \( p \)-regular at \( x_0 \). Then

\[
T_1 M(x^*) = H_p(x^*).
\]

**Optimality conditions for \( p \)-regular optimization problems**

We define \( p \)-factor Lagrange function

\[
\mathcal{L}_p(x, \lambda, h) = \varphi(x) + \left( \sum_{k=1}^{p} F_k^{(k-1)}(x)[h]^{k-1}, \lambda \right),
\]

where \( \lambda \in Y^* \) and

\[
\bar{\mathcal{L}}_p(x, \lambda, h) = \varphi(x) + \left( \sum_{k=1}^{p} \frac{2}{k(k+1)} F_k^{(k-1)}(x)[h]^{k-1}, \lambda \right).
\]

**Definition 4.** The mapping \( F \) is called strongly \( p \)-regular at the point \( x^* \) if there exists \( \gamma > 0 \) such that

\[
\sup_{h \in H_\gamma} \left\| \{ \Psi_p(h) \}^{-1} \right\| < \infty,
\]

where

\[
H_\gamma = \left\{ h \in X : \left\| F_k^{(k)}(x^*)[h]^k \right\|_{Y_k} \leq \gamma, \quad i = 1, p, \quad \| h \| = 1 \right\}.
\]

Let us recall the following basic theorems of the \( p \)-regularity theory.

**Theorem 2 (Necessary and sufficient conditions for optimality).** Let \( X \) and \( Y \) be Banach spaces, \( \varphi \in C^2(X), \quad F \in C^{p+1}(X), \quad F : \)
Suppose that $h \in H_p(x^*)$ and $F$ is $p$-regular along $h$ at the point $x^*$. If $x^*$ is a local solution to the problem (1)–(2), then there exist multipliers, $\lambda^*(h) \in Y^*$ such that
\[
L'_p(x^*, \lambda^*(h), h) = 0. \tag{5}
\]
Moreover, if $F$ is strongly $p$-regular at $x^*$, there exist $\alpha > 0$ and a multiplier $\lambda^*(h)$ such that (5) is fulfilled and
\[
\bar{L}_{pxx}(x^*, \lambda^*(h), h)[h]^2 \geq \alpha\|h\|^2 \tag{6}
\]
for every $h \in H_p(x^*)$, then $x^*$ is a strict local minimizer to the problem (1)–(2).

For our purposes will be useful the following modification of Theorem 1.

**Theorem 3.** Let $X$ and $Y$ be Banach spaces, $\varphi \in C^2(X)$, $F \in C^{p+1}(X)$, $F : X \to Y$, $\varphi : X \to \mathbb{R}$. Suppose that $h \in H_p(x^*)$ and $F$ is $p$-regular along $h$ at the point $x^*$. If $x^*$ is a local solution to the problem (1)–(2), then there exist multipliers $\bar{\lambda}_i(h) \in Y_i$, $i = 1, \ldots, p$, such that
\[
\varphi'(x^*) + F'(x^*)^* \bar{\lambda}_1(h) + \cdots + \left(F^{(p)}(x^*)[h]^{(p-1)}\right)^* \bar{\lambda}_p(h) = 0 \tag{7}
\]
and
\[
\left(F^{(k)}(x^*)[h]^{(k-1)}\right)^* \bar{\lambda}_i(h) = 0, \quad i = \{1, \ldots, p\} \setminus k, \tag{8}
\]
$k = 1, \ldots, p$.

Moreover if $f$ is strongly $p$-regular at $x^*$, there exist $\alpha > 0$ and multipliers $\bar{\lambda}_i(h)$, $i = 1, \ldots, p$, such that (7)–(8) hold and
\[
\varphi''(x^*) + \frac{1}{3} F''(x^*) \bar{\lambda}_1(h) + \cdots + \frac{2}{p(p+1)} F^{(p+1)}(x^*)[h]^{p-1} \bar{\lambda}_p(h) [h]^2 \geq \alpha\|h\|^2 \tag{9}
\]
for every $h \in H_p(x^*)$, then $x^*$ is a strict local minimizer to the problem (1)–(2).

The author were supported by the Russian Foundation for Basic Research (grant No 11-01-00786a) and the Council for the State Support of Leading Scientific Schools (grant 5264.2012.1).
Separating plane algorithm with additional clipping for convex optimization

Evgeniya A. Vorontsova

1 Institute of Automation and Control Processes, Far Eastern Branch of the Russian Academy of Sciences; Far Eastern Federal University, Vladivostok, Russia; vorontsovaea@gmail.com

We consider the solvable unconditional optimization problem

$$\min_{x \in E_n} f(x) = f(x^*)$$

where $E_n$ - $n$-dimensional Euclidean space and $f$ is a nonsmooth convex objective function from $E_n$ into $E_1$.

There are many areas of applications for such algorithms because this kind of problems occurs frequently in the industry. Moreover, the area of large-scale optimization will have more advantages of any improvements in the methods for minimization of nondifferentiable functions.

Research in this direction has resulted in several efficient methods [1-5]. This paper presents an algorithm for solving the problem (1). The algorithm belongs to a class of separating plane algorithms [4-5].

The basic idea of separating plane algorithms (SPA) is to use the next to trivial identity of convex analysis

$$\min f(x) = f(x^*) = -f^*(0)$$

where $f^*(g) = \sup_x \{xg - f(x)\}$ - the Fenchel-Moreau conjugate of the function $f$.

In this way the problem (1) can be reformulate as a problem of computing $f^*(0)$. The optimal point $x^*$ can be obtained as a subgradient of $f^*$: $x^* \in \partial f^*(0)$.

The SPA algorithms construct sequences of outer and inner approximations of the epigraph of $f^*$ (epi $f^* = \{(\nu, g) : \nu \geq f^*(g)\}$). At each iteration of the algorithm the approximations are gradually refined. Eventually we obtain converging lower and upper bounds for $f^*(0)$.

As the SPA algorithm [5] has no guarantee of monotony, the following improvement is suggested. At each iteration we execute an additional step
which removes the upper part of epi $f^*$:

$$\sup_{(g, \epsilon) \in \text{epi } f^*; \epsilon \leq \bar{v}} \{gx - \epsilon\}$$  \hspace{1cm} (2)

where estimate $\bar{v}$ is a solution to a problem $\bar{v} = \min \epsilon$ under the condition:

$$(0, \epsilon) \in \text{co} ((g^k, f^*(g^k)), k = 1, 2, ...) + 0 \times \mathbb{R}_+.$$

Using standard duality for (2), we reduce it to a line-search problem:

$$\min_{0 < \theta \leq 1} \frac{1}{\theta} (f(\theta x) + \bar{v}) = \min_{0 < \theta \leq 1} \phi(\theta).$$  \hspace{1cm} (3)

As $f$ is a convex function, the function of a single variable $\phi(\theta)$ is convex.

It is suggested to solve the one-dimensional nonsmooth optimization problem (3) by means of a new modification of line-search algorithm for nonsmooth convex optimization (see [6] for the original scheme).

Numerical experiments demonstrated quite satisfactory computational performance of the modified separating plane algorithm with additional clipping.

The author was supported by the Far Eastern Branch of the Russian Academy of Sciences (project no. 12-III-A-01N-014).

References

Equilibrium Model of the Russian Economy for the period of Global Financial Crisis

V.P. Vrzheshch¹, N.P. Pilnik², I.G. Pospelov³

¹ Faculty of Computational Mathematics and Cybernetics MSU, Moscow, Russia; valentin.vrzheshch@gmail.com
² High School of Economy, Moscow, Russia; u4d@yandex.ru
³ Dorodnicyn Computer Centre of RAS, Moscow, Russia; pospeli@yandex.ru


The model successfully reproduces dynamics of the main macroeconomic indicators (GDP and its components, inflation, loans, deposits, etc.) for ten years, including the period of the global crisis. The identified model demonstrates the strong turnpike property: even the agents are allowed to know the future, this information has little effect on their optimal behavior.

REFERENCES

A Modified Steepest Descent Method Based on BFGS Method for Locally Lipschitz Functions

R. Yousefpour

1 Department Mathematical Sciences, University of Mazandaran, Babolsar, Iran; yousefpour@umz.ac.ir

Abstract

In this paper, the steepest descent method is modified based on the BFGS method. At first, the descent direction is approximated based on the Goldstein subgradient and a positive definite matrix. Then the Wolfe condition is generalized for the locally Lipschitz functions. Next, an algorithm is developed to compute a step length satisfying the generalized Wolfe condition. At each iteration, the positive definite matrix is updated by the BFGS method. By the presented line search algorithm, a pair of subgradients are computed for updating the positive definite matrix such that the secant equation is satisfied. Finally, an algorithm is devolved based on the BFGS method. This algorithm is implemented with MATLAB codes.

keywords: Lipschitz functions, Armijo and Wolfe conditions, Line search, BFGS method.

Introduction

The steepest descent direction for the locally Lipschitz functions is computed based on an element of the Goldstein subgradient with minimal norm. By an approximation of this direction, several bundle methods were developed [1-6]. The efficiency of an algorithm, that developed based on an approximation of steepest descent direction, depends on the approximation accuracy. To improve the accuracy of an approximation, we need to compute a larger number of subgradient and, this is time consuming. For example, in [6], the steepest descent direction is approximated by sampling gradients. This approximation is appropriate, but computing this approximation for large scale problems is very expensive. In [4], the
steepest descent direction is iteratively approximated. This method computes a good approximation for the steepest descent direction by the less number of subgradients. The numerical results showed that this algorithm is more efficient than other methods.

If the steepest descent method is suitably approximated, then the constructed algorithm based on this direction will be efficient. On the other hand, the BFGS method has good behavior for some nonsmooth functions. It can be expected that the combination of the steepest descent me and BFGS methods will be efficient. The first step of combining the steepest descent and BFGS methods is that compute an approximation of the steepest descent direction by the Goldstein subgradient and, a positive definite matrix. In each iteration of the main algorithm, the positive definite matrix must be updated by the BFGS method thus, the Wolfe step length condition is generalized based on the Goldstein subgradient. Based on this generalization, we prove that there exist step lengths satisfying this generalization for a descent direction. We generalized the line search algorithms [7] and, prove that these generalize algorithms compute a step length satisfying the generalized Wolfe condition. These algorithms also compute two subdifferentials for updating the approximation of Hessian matrix.

By using the similar ideas to [4], the Goldstein subgradient is iteratively approximated until a descent direction is found based on a positive definite matrix. By these ideas, the Goldstein subgradient is efficiently approximated. We prove that this procedure finds a descent direction after finitely many iterations. This descent direction is an approximation of the steepest descent direction based on positive definite matrix. The function is reduced along this direction. In each iteration of a minimization algorithm, the positive matrix is updated by the BFGS method. The presented line search algorithm finds two subgradients such that the positive definite matrix can be updated and, the secant equation is satisfied. Finally, the algorithm is implemented with Matlab and, the results are compared with other methods.

**Generalized Armijo and Wolfe conditions**

In this paper, we suppose that $H$ is a positive definite matrix. Suppose that $f : \mathbb{R}^n \to \mathbb{R}$ is a locally Lipschitz function, $v \in \mathbb{R}^n$ and, $g = -Hv$. If
there exists $\alpha > 0$ such that
\[ f(x + t g) - f(x) \leq -t \| v \|_H, \ \forall t \in (0, \alpha), \]
then $g$ is a descent direction. Now, the Armijo condition is given as follows.

**Definition 1.** Let $f : \mathbb{R}^n \rightarrow \mathbb{R}$ be a locally Lipschitz function and, $v \in \mathbb{R}^n$. Suppose that $g = -Hv$ is a descent direction at $x$. A step length $\alpha > 0$ satisfies in the Generalized Armijo Condition (GAC), if
\[ f(x + \alpha g) - f(x) \leq -c_1 \alpha \| v \|_H, \]
where $c_1 \in (0, 1)$. Equivalently $A(\alpha) \leq 0$, where
\[ A(\alpha) := f(x + \alpha g) - f(x) + c_1 \alpha \| v \|_H. \]

**Definition 2.** Let $f : \mathbb{R}^n \rightarrow \mathbb{R}$ be a locally Lipschitz function, $v \in \mathbb{R}^n$ and, $g = -Hv$ be a descent direction at $x$. If $W(\cdot)$ is not decreasing on a neighborhood of $\alpha$ and, the GAC satisfies in $\alpha$ along direction $g$ at $x$, then we say that the GWC satisfies at $\alpha$ along direction $g$ at $x$.

**Computing descent direction**

Here, we discuss on finding a descent direction for a locally Lipschitz function $f$ based on $\partial_\varepsilon f(x)$ and, a given positive definite matrix $H$. Consider the following problem and, suppose that $0 \notin \partial_\varepsilon f(x)$,
\[ \xi_0 = \arg \min_{\xi \in \partial_\varepsilon f(x)} \| \xi \|_H. \tag{1} \]
Let $g = -H\xi_0$. We have $f^\circ(x, g) \leq -\| \xi_0 \|_H < 0$ and, by the Leburg Mean Value Theorem,
\[ f(x + t g) - f(x) \leq -t \| \xi_0 \|_H, \ \forall t \in (0, \epsilon). \tag{2} \]
Thus, $g$ is a descent direction. But, solving this problem often is impractical. So, $\partial_\varepsilon f(x)$ is approximated by the convex hull of its some elements. More exactly, if
\[ W_k = \{ v_1, v_2, \ldots, v_k \} \subset \partial_\varepsilon f(x), \]
then we consider $\text{conv} W_k$ as an approximation of $\partial_\varepsilon f(x)$. Therefore, we solve the following problem, which is an approximation of (1),
\[ v_0 = \arg \min_{v \in \text{conv} W_k} \| v \|_H. \]
Minimization Algorithm

In the previous Subsection, we found a descent direction based on a positive definite matrix and, an approximation of $\partial \epsilon f(x)$. To update the inverse of Hessian matrix approximation, we need a pair of subgradients such that the secant equation is satisfied. Thus, by applying the line search algorithm, a step length satisfying in the GWC and, a pair of subgradients for updating the inverse of Hessian matrix approximation are computed. Now, we present the nonsmooth version of BFGS algorithm.

Algorithm 1.

Step 0 (Initialization)
Let $x_1 \in \mathbb{R}^n$, $v_1 \in \partial f(x_1)$, $\theta_\epsilon$, $c_1$, $\theta_\delta$, $\epsilon_1$, $\delta_1 \in (0, 1)$, $H_1 = I_{n \times n}$, $c_2 \in (c_1, 1)$ and, set $k = 1$, where $I_{n \times n}$ is the identical matrix.

Step 1 (Set new parameters)
Set $m = 1$, $H_k^m = H_k$ and, $x_k^m = x_k$.

Step 2 (Compute descent direction)
Apply Algorithm 2 in [4] at point $x_k^m$, with $H = H_k^m$, $v = v_k^m$, $\delta = \delta_k$ and, $\epsilon = \epsilon_k$. Let $n_k^m$ be the number of iterations needed for termination of the algorithm and, let $\|w_k^m\|_{H_k^m} = \min\{\|w\|_{H_k} : w \in \text{conv } W_k^m\}$. If $\|w_k^m\|_{H_k} = 0$ then Stop else let $g_k^m = -H_k^m w_k^m$ be the descent direction.

Step 3 (Line search)
If the stopping condition (2) is not satisfied then go to Step 5, else apply the line search algorithm. If the algorithm terminates successfully, then $\alpha$ is the line search parameter satisfying the GWC and, $v_{k+1}^m \in \partial \epsilon f(x_k^m + \alpha g_k^m)$ is a vector such that $<(v_{k+1}^m), g_k^m > + c_2\|w_k^m\|_{H_k^m} > 0$, else $\alpha$ is the line search parameter satisfying the GAC. Construct the next iterate $x_{k+1}^m = x_k^m + \alpha g_k^m$ and, go to Step 4.

Step 4 (BFGS update)
If the line search algorithm terminates successfully, then set $s = \alpha g_k^m$ and, $y = v_{k+1}^m - v_k^m$ and

$$H_k^{m+1} = H_k^m - \frac{H_k^m y y^T H_k^m}{<y, H_k^m y>} + \frac{ss^T}{<y, y>},$$
Theorem 1. Let \( f : \mathbb{R}^n \rightarrow \mathbb{R} \) be a locally Lipschitz function. If the level set \( M = \{ x : f(x) \leq f(x_1) \} \), is bounded, then either Algorithm 1 terminates finitely at some \( k_0 \) and, \( m_0 \) with \( \{ w_{m_0} \} = 0 \) or every cluster point of the sequence \( \{ x_k \} \), generated by Algorithm 1, belongs to the set \( X = \{ x \in \mathbb{R}^n : 0 \in \partial f(x) \} \).

References

The primal affine-scaling method for semidefinite programming with steepest descent

Vitaly Zhadan

1 Dorodnicyn Computing Centre of RAS, Moscow, Russia; zhadan@ccas.ru

In this paper we consider the variant of a feasible affine scaling primal method for solving linear semidefinite programming problem. This variant has a local convergence at a linear rate, when all points in the iterative process belong to the relative interior of the feasible set. But our aim is to show how it is possible to modify the method in order to have the opportunity to apply the steepest descent approach for choosing the step size.

Let \( S^n \) denote the space of symmetric matrices of order \( n \), and let \( S^n_+ \) be the cone in \( S^n \), consisting of positive semidefinite matrices. We use also the inequality \( M \succeq 0 \) to indicate that a matrix \( M \) belongs to \( S^n_+ \). The space \( S^n \) is finite-dimensional, and its dimension is equal to so-called triangular number \( k_\Delta(n) = n(n + 1)/2 \). The inner product of matrices \( M_1 \) and \( M_2 \) of the same size is defined as the trace of the matrix \( M_1^T M_2 \) and is denote by \( M_1 \cdot M_2 \).

Consider the linear semidefinite programming problem

\[
\min \ C \cdot X, \quad A_i \cdot X = b^i, \quad 1 \leq i \leq m, \quad X \succeq 0, \quad (1)
\]

where the matrices \( C, X \) and \( A_i \ (1 \leq i \leq m) \) belong to \( S^n \). The dual problem of (1) has the form

\[
\max \ b^T u, \quad \sum_{i=1}^{m} u^i A_i + V = C, \quad V \succeq 0, \quad (2)
\]

where \( b = (b^1, \ldots, b^m)^T, \ V \in S^n \). It is assumed that both problems (1) and (2) are solvable and the matrices \( A_i \ (1 \leq i \leq m) \) are linearly independent. We denote the feasible set of primal problem (1) by \( \mathcal{F}_P \) and its relative interior by \( \mathcal{F}^0_P \).

If \( X_* \) and \( V_* \) are optimal solutions of problems (1) and (2), respectively, then \( X_* \cdot V_* = 0 \) and the matrices \( X_* \) and \( V_* \) must commute. Hence, there exists an orthogonal matrix \( Q \) such that

\[
X_* = Q \text{Diag}(\lambda_*) Q^T, \quad V_* = Q \text{Diag}(\mu_*) Q^T.
\]

262
where $\lambda_* = [\lambda_*^1, \ldots, \lambda_*^n]$ and $\mu_* = [\mu_*^1, \ldots, \mu_*^n]$ are the eigenvalues of $X_*$ and $V_*$ respectively. The eigenvalues $\lambda_*^i$ and $\mu_*^i$ satisfy the complementarity conditions $\lambda_*^i \mu_*^i = 0$, $1 \leq i \leq n$. The strict complementarity condition means that one of the values $\lambda_*^i$ or $\mu_*^i$ is strictly positive.

Since we assumed that both problems (1) and (2) have solutions, the following system of equalities and inequalities

\[
\begin{align*}
X \cdot V &= 0, \\
A_i \cdot X &= b_i, \quad 1 \leq i \leq m, \\
V &= C - \sum_{i=1}^{m} u^i A_i, \\
X &\succeq 0, \quad V \succeq 0,
\end{align*}
\]

is necessarily solvable.

Denote by $X \ast V$ the symmetrized product of square matrices $X$ and $V$ defined by the formula $X \ast V = (XV + V^T X^T)/2$.

**Statement.** For symmetric matrices $X \succeq 0$ and $V \succeq 0$, the equality $X \ast V = 0_{nn}$ is possible iff $XV = VX = 0_{nn}$.

Using the statement, we can replace the first equality in (3) by the following one: $X \ast V = 0_{nn}$.

Let the symbol vec$X$ denote the direct sum of the columns of $X \in S^n$, that is, the column vector of dimension $n^2$ that consists of the columns of $X$ written one after another from top to bottom. For symmetric matrices, it is more convenient to deal with the column vector vech$X$ of dimension $k_{\Delta}(n)$. It also consists of the columns of $X$ written one after another; however, these are not the entire columns but their parts beginning with the diagonal entry. The operation vec$X$ is defined similarly. It differs from the preceding operation vech$X$ only in that the off-diagonal entries of $X$ are multiplied by two before placing into vec$X$.

Let also $L_n$ and $D_n$ are the elimination and duplicated matrices respectively [1]. The matrix $L_n$ for an arbitrary square matrix $X$ effects the transformation: $L_n vecX = vechX$. By contrast, the matrix $D_n$ acts on an arbitrary symmetric matrix $X$ so that $D_n vechX = vecX$.

With the help of vectorization operators vec$X$ and vecs$X$, the optimality conditions (3) can be rewritten as

\[
\begin{align*}
\tilde{X} \otimes \text{vech}V &= 0_{k_{\Delta}(n)}, \\
A_{\text{vecs}} \text{vech}X &= b, \\
\text{vech}V &= \text{vech}C - A_{\text{vech}}^T u, \\
X &\succeq 0, \quad V \succeq 0,
\end{align*}
\]
where \( \tilde{X} \otimes = L_n X \otimes D_n \), and \( X \otimes = [X \otimes I_n + I_n \otimes X]/2 \) is the Kronecker sum of a matrix \( X \), \( I_n \) is the identity matrix of order \( n \). By \( A_{\text{vech}} \) and \( A_{\text{vecs}} \) we denote the \( m \times n^2 \) matrices with \( \text{vech} A_i \) and \( \text{vecs} A_i \) respectively as their rows, \( 1 \leq i \leq m \).

Let us consider the primal affine scaling method for solving problem (1), based on some approach for solving the system (4). For its deriving from (4) we substitute the expression for \( \text{vech} V \) from the third equality to the first one and multiply both sides of that equality by the matrix \( A_{\text{vecs}} \) from the left. As a result, we obtain the following system of linear algebraic equations with respect to the vector \( u \):

\[
\Gamma(X) u = A_{\text{vecs}} \tilde{X} \otimes \text{vech} C, \tag{5}
\]

where \( \Gamma(X) = A_{\text{vecs}} \tilde{X} \otimes A_{\text{vech}}^T \).

If the matrix \( \Gamma(X) \) is nonsingular, then solving system (5), we obtain

\[
u(X) = \Gamma^{-1}(X) A_{\text{vecs}} \tilde{X} \otimes \text{vech} C.
\]

Denote \( V(u) = C - \sum_{i=1}^{m} u^i A_i \), \( V(X) = V(u(X)) \). After substitution \( V(X) \) into the first equality from (4) we obtain the nonlinear system of equations with respect to \( X \):

\[X \ast V(X) = 0_{nn}.
\]

Applying the fixed point method for solving this system, we have

\[
X_{k+1} = X_k - \alpha_k X_k \ast V_k, \quad V_k = V(X_k), \tag{6}
\]

where \( X_0 \in \mathcal{F}_P^0 \), \( \alpha_k > 0 \).

Let \( \mathcal{R}_A \) be the subspace of \( \mathcal{S}_n \) spanned by the matrices \( A_i \), \( 1 \leq i \leq m \), and let \( \mathcal{R}_A^\perp \) be its orthogonal complement. Let also \( \mathcal{T}_X \) be the tangent space of \( \mathcal{S}_n^+ \) at the point \( X \). Following [2], we give the definition of a nondegenerate points \( X \) and \( V \).

**Definition.** A feasible point \( X \) of the primal problem (1) is nondegenerate, if \( \mathcal{R}_A^\perp + \mathcal{T}_X = \mathcal{S}_n^+ \). Similarly, a feasible point \( V \) of dual problem (2) is nondegenerate, if \( \mathcal{R}_A + \mathcal{T}_V = \mathcal{S}_n^+ \).

**Lemma.** Let \( X \in \mathcal{F}_P \) be a nondegenerate point. Then the matrix \( \Gamma(X) \) is nonsingular.

In what follows, we assume that the problem (1) is nondegenerate, i.e. all points from \( \mathcal{F}_P \) are nondegenerate.
Theorem [3]. Let the solutions $X_*$ and $V_*$ of primal and dual problems (1), (2) are strictly complementary. Let also points $X_*$ and $V_*$ are nondegenerate. Then for $\alpha_k$ sufficiently small the iterative process (6) locally converges to $X_*$ at a linear rate.

Denote $\Delta X_k = X_k * V(X_k)$. The iterative process (6) possesses the following properties:

$$A_i \cdot \Delta X_k = 0, \quad 1 \leq i \leq m; \quad C \cdot X_{k+1} \leq C \cdot X_k.$$ 

Thus, in order to obtain the maximal decreasing of the value of the objective function $C \cdot X$, the step size $\alpha_k$ should be taken as large as possible provided that $X_{k+1} \succeq 0$. It leads to application of steepest descent approach for choosing the step size. But at boundary points $X_k \in \mathcal{F}_P$ we have $\Delta X_k \in \mathcal{T}_{X_k}$. Hence, we must modify the right hand side in (8), when $X_k$ is a boundary point of $\mathcal{F}_P$.

Assume that $X$ is a boundary point of the feasible set $\mathcal{F}_P$, and assume also that the rank of $X_k$ is $r < n$. Then the matrix $X$ can be represented in the form:

$$X = Q \text{Diag}(\lambda^1, \ldots, \lambda^r, 0, \ldots, 0) Q^T,$$

where $Q$ is an orthogonal matrix and $\lambda^i > 0$, $1 \leq i \leq r$. Let $Q_B$ and $Q_N$ be the submatrices formed of the first $r$ and the last $n-r$ columns of the matrix $Q$. Let also $\lambda_B = [\lambda^1, \ldots, \lambda^r]$.

At the non-optimal point $X \in \mathcal{F}_P$ the matrix $V(X)$ is not positively semidefinite. Therefore the matrix $V^Q(X) = Q^T V(X) Q$ is sign indefinite. Suppose that its principal submatrix $V^Q_{NN} = Q^T_N V(X) Q_N$ is also sign indefinite. Without loss of generality we may regard $V^Q_{NN}$ as a diagonal matrix. Then some of its diagonal elements are negative.

Let $y$ be a $(n-r)$-dimensional vector consisting from nonnegative elements with $\sum_{i=1}^{n-r} y^i = 1$. Moreover let positive elements of $y$ dispose at places corresponding to negative diagonal elements of $V^Q_{NN}$. Denote be $Y$ the diagonal matrix such that $Y^2 = \text{Diag}(y)$.

We take some $\varepsilon > 0$ and instead of the direction $\Delta X = X * V(X)$ consider the following one:

$$\Delta X = \Delta X_B + \varepsilon \Delta X_N, \quad (7)$$

where

$$\Delta X_B = Q_B \Lambda_B Q_B^T, \quad \Lambda_B = \text{Diag}(\lambda_B) \ast (Q_B^T Q_B (\tilde{u}) Q_B),$$

and

$$\Delta Y = Q_N \Lambda_N Q_N^T, \quad \Lambda_N = \text{Diag}(\lambda_N) \ast (Q_N^T Q_N (\tilde{u}) Q_N).$$
\[ \Delta X_N = Q_N \Lambda_N Q_N^T, \quad \Lambda_N = Y^2 \ast (E_{N-} Q_N^T V(\tilde{u}) Q_N E_{N-}). \]  \tag{8}

In (7) the first direction \( \Delta X_B \) belongs to the minimal face of the set \( \mathcal{F}_P \) containing the point \( X \). The second direction \( \Delta X_N \) belongs to the conjugate face of the minimal face. The matrix \( E_{N-} \) in (8) is a diagonal matrix with the indicator vector of \( y \) on its diagonal.

The dual vector \( \tilde{u} \) is taken in such a way that the following system of linear equalities

\[ A_i \cdot (\Delta X_B + \varepsilon \Delta X_N) = 0, \quad i = 1, \ldots, m. \]

is fulfilled.

It can be verified that the direction (7) leads to decreasing of the value of the objective function \( C \cdot X \). Using the steepest descent for choosing \( \alpha_k \), we obtain the updated point \( X_{k+1} \). This point again belongs to the boundary of the feasible set \( \mathcal{F}_P \). The behavior of the method is similar to behavior of the simplex method. Certainly, because of the fact that the cone \( S^n_+ \) is non-polyhedral, infinitely many iterations are required for solving (1).

The author was supported by the Programs of Fundamental Research of Russian Academy of Sciences P-15, P-18, by the Russian Foundation for Basic Research (project no.11-01-00786), and by the Leading Scientific Schools Grant no.5264.2012.1

REFERENCES

The point projections on linear manifold

Valery Zorkaltsev

1 Melentiev Energy Systems Institute SB RAS, Irkutsk, Russia; 
zork@isem.sei.irk.ru

Many mathematical models can be reduced to the geometric problem of finding the least distance points of the linear manifold from the origin. The problems of estimating of the parameters of the linear regression, the problems of the regularization of solutions in the balance models, the problems of the search of pseudo-solutions in models with conflicting conditions are considered in the report as examples.

The given geometric problem appears as a component part of the some computational algorithms. It is discussed in the report as applied to solving of the systems of nonlinear equations by the iterative linearization method and conformably to the interior point algorithms for solving of the linear and nonlinear mathematical programming problems.

The main goal of this report is to present the recent results and results published in [1], in the study of properties and correlations of the least distant from the origin points of the linear manifold under the different definitions of closeness.

Let’s consider some concretizations of the given geometric problem. Let $L$ be the linear manifold in $R^n$, set $J(x)$ be set of numbers of nonzero components of the vector $x$ from $R^n$ (carrier of the vector). The symbol $\subset$ denotes strict inclusion. The set of vectors $R^n$ with all positive components we denote $R^n_{++}$. The convex hull of set and closure of set $X$ from $R^n$ are denoted $coX$ and $clX$.

1. Let’s introduce the set of vectors of the linear manifold with the minimal carriers

$$B = \{x \in L : \neg \exists y \in L, J(y) \subset J(n)\}.$$

It is known [1] that the amount of vectors in $B$ is finite and no more than $C^n_m$, where $m$ is dimension of manifold $L$.

2. Lets denote the set of vectors of the linear manifold with Pareto-minimal absolute meanings of the components

$$Q = \{x \in L : \neg \exists y \in L, \sum |y_i| < \sum |x_i|, |y_i| \leq |x_i|, \ i = 1, \ldots, n\}.$$
The convex hull $B$ is minimal convex set, which contains $Q$ (see [1]),

$$coQ = coB.$$ 

The set $Q$ is bounded according to this relation. In addition to this we note that the set $Q$ can be nonconvex, but it is tie set and it is union of the finite amount of the polytopes.

3. Let’s introduce the Holder, octahedral and Euclidean projection of a point on the linear manifold. We obtain for the Holder norm the following expressions when the vector of a weighted coefficients $h \in R^+_n$ and the degree coefficient $p \in (1, \infty)$ are given

$$y(p, h) = \arg \min \left\{ \sum h_i |y_i|^p : y \in L \right\},$$

$$Y(1, h) = \text{Arg min} \left\{ \sum h_i |y_i| : y \in L \right\},$$

$$y(\infty, h) = \arg \text{lex min} \left\{ \max_i h_i |y_i| : y \in L \right\}.$$

The Holder projection is Euclidean projection under $p = 2$.

Let us imply that the Chebyshev projection $y(\infty, h)$ is determined as a result of solving of the following multicriteria problem of lexicographic optimization (the symbol lex denotes this). We find the vector from $L$ with minimal meanings of the maximal values $h_i |y_i|$. Moreover, we find the relative interior point of solutions of this problem. We fix the meanings of a components of the vector $y$ under which the maximal meaning of the given values is obtained (it is true for all solutions of the given problem). On the next stage we solve the minimization problem of maximal meanings of the values $h_i |y_i|$ for the remaining variables. When we found a relative interior point of the optimal solutions of this problem, we fix variables $y_i$ on their optimal level under which the optimal value of objective function is achieved on the second stage problem. If the solution is not unique then we’ll repeat the process. We obtain the unique solution after the finite stages of computation. This solution can be used as the Chebyshev projection of the origin on the linear manifold $L$.

The sets of all octahedral, Chebyshev, Euclidean projections with the different weighted coefficients and the set of the Holder projections are denoted

$$P_1 = \bigcup_{h \in R^+_n} Y(1, h),$$
\[ P_{\infty} = \{ y(\infty, h) : h \in \mathbb{R}^{n}_{++} \}, \]
\[ P_2 = \{ y(2, h) : h \in \mathbb{R}^{n}_{++} \}, \]
\[ P = \{ y(p, h) : h \in \mathbb{R}^{n}_{++}, p \in (1, \infty) \}. \]

The following relations was earlier proved
\[ P_2 = P, \; clP_2 = Q. \]

Let’s announce the statements
\[ Y(1, h) \cap B \neq \emptyset, \forall h \in \mathbb{R}^{n}_{++}, \]
\[ P_1 = Q, \; P_{\infty} = P_2. \]

The octahedral projections can have more than one meaning for the given vector of weighted coefficients \( h \). Moreover, the vectors of the linear manifold with the minimal carrier always exist among the octahedral projections with the given vector of weighted coefficients according to the first adduced relations.

In particular, any Holder and Chebyshev projections of point on the linear manifold can be received by the least-squares method thanks to choice of the vector of weighted coefficients. Any octahedral, Euclidean projections and any vector of the linear manifold with Pareto-minimal absolute meanings of components (including any vector from \( B \)) can be obtained with any given accuracy according to the technique too. These statements follow from the above relations.

REFERENCES

The effect of the setup parameters on the evolution of the substance crystallization process

Vladimir Zubov¹, Alla Albu², Andrey Albu³

¹ Dorodnicyn Computing Centre of RAS, Moscow, Russia; zubov@ccas.ru
² Dorodnicyn Computing Centre of RAS, Moscow, Russia; alla.albu@mail.ru
³ Dorodnicyn Computing Centre of RAS, Moscow, Russia; 7488148@mail.ru

The process of solidification in metal casting is considered. The crystallization of molten metal is one of the important stages of foundry practice. On that how the crystallization process proceeded, depends the quality of the obtained model.

Liquid metal is placed into the work cavity of a mould with a prescribed configuration (see Fig. 1). A special setup is used to crystallize the metal (see Fig. 2). It consists of upper and lower parts. The upper part consists of a furnace with a mold moving inside. The lower part is a cooling bath consisting of a large tank filled with liquid aluminum whose temperature is somewhat higher than the aluminum melting point. The mould is slowly moving among the furnace and begins to dip into liquid aluminum that has a relatively low temperature and thus is cooling the metal. At the same time the mould receives heat from the furnace and this heat doesn’t allow the crystallization process to run too fast. The crystallization process is affected by different phenomena such as heat loss due to heat radiation, gain of energy by the mould due to expanse of
heat radiation from aluminum and furnace, heat exchange between liquid aluminum and the mould.

Different technological requirements can be imposed on the evolution of the phase boundary to obtain a detail of the desired quality. In this work certain ones are considered: the shape of the phase boundary should be close to a plane and its law of motion should be close to a preset one. One of the consequences of satisfying these requirements would be absence of "bubbles" of liquid metal. These "bubbles" are essentially areas of liquid metal that are surrounded by already solidified metal. If such areas are present during the solidification process, the quality of the detail would be unacceptable for sure.

The process is described by a three-dimensional unsteady heat equation (see [1]). The complication of the problem is that the metal can be present simultaneously in two phases: solid and liquid. Density of materials, their heat capacity and thermal conductivity all depend on the temperature and are discontinuous in the border between the metal and the mould.

The evolution of the solidification front is affected by numerous parameters (for example, by the furnace temperature, the liquid aluminum temperature, the depth to which the object is immersed in the liquid aluminum, the velocity of the mold relative to the furnace, etc.) that could be controlled, but in practice most often the speed $u(t)$ at which the mold moves relative to the furnace is used as the control. To find a control $u(t)$ that will satisfy the mentioned technological requirements, we state the optimal control problem (see [2]) and write down the following cost functional:

$$I(u) = \frac{1}{t_2(u) - t_1(u)} \int_{t_1(u)}^{t_2(u)} \int \int_S [Z_{pl}(x, y, t, u) - z_*(t)]^2 dx dy dt.$$  

Here $t_1$ is the time, when the crystallization front arises; $t_2$ is the time, when the crystallization of metal completes; $(x, y, Z_{pl}(x, y, t, u))$ are the real coordinates of the interface at the time $t$; $(x, y, z_*(t))$ are the desired coordinates of the interface at the time $t$; $S$ is the largest cross section of the mold that is filled with metal.

In this paper we investigate the effect of the furnace temperature and the maximum depth of immersion of the mold into the cooler on the evolution of the substance crystallization process. A large series of calculations
of the substance crystallization problem was carried out in which these parameters were varied within wide limits. As a starting point in this series was selected a point with the following parameters: maximum allowable depth of immersion into the cooler $DISPL = 360 \text{ mm}$ (immersion up to the fourth stage), and the furnace temperature $T_{sou} = 1840 \degree K$ (here and below the immersion is measured in millimetres, the furnace temperature is measured in Kelvin degrees and the speed of moving object is measured in mm/min). These are the actual values of the parameters of the substance crystallization process that are used in the real setup shown in Fig. 2. The aim of the research was to answer the following question: is it possible by changing the oven temperature $T_{sou}$ and the maximum depth of immersion $DISPL$ to achieve a more preferable crystallization process?

The study was conducted for the case when the velocity of the object relative to the furnace was constant. The value of this velocity is varied in the range from 2 to 50 (mm/min). The results of the numerical studies presented below are divided into four groups. Each group includes the results of calculations which were obtained at the same value of the maximum depth of immersion $DISPL$. Various versions in each group differ by temperature $T_{sou}$ and velocity $u(t)$ of the object relative to the furnace.

$DISPL = 360 \text{ (immersion up to the 4th stage)}$

The calculation results show that when using this immersion depth the metal solidification process is admissible (technological requirements are violated within acceptable limits), if the temperature of the furnace is in the range $1830 < T_{sou} < 1850$ and the constant speed of movement is in the range $18 < u(t) < 25$. If speed $u(t)$ is out of this range then it leads to formation of liquid metal ”bubbles” during the crystallization process for all considered values of the furnace temperature $T_{sou}$. A similar situation arises if the furnace temperature $T_{sou} < 1830$ for all considered values of velocity. If the temperature $T_{sou} > 1850$, the ”bubbles” of liquid metal do not form, but at the same time either crystallization process doesn’t come to an end or it takes an unacceptably long time. Note also that a decrease in speed and increase in the oven temperature (in the specified range) leads to a crystallization process that is closer to the desired one.
\( \text{DISPL} = 280 \) (immersion up to the 3th stage)

Decrease in the maximum depth of immersion of the object into the cooler, as shown by calculations, leads to a deterioration of the characteristics of the crystallization process. Neither change in the velocity of the mold \( u(t) \), nor change in the oven temperature \( T_{sou} \) allowed to get rid of liquid ”bubbles” in the metal. It should also be noted that in this case the crystallization of the metal runs unacceptably long (or isn’t happening at all) for even lower temperatures starting from \( T_{sou} > 1810 \).

\( \text{DISPL} = 420 \) (immersion up to the 5th stage)

The increase in the maximum depth of immersion has a positive effect on the characteristics of the crystallization process. The ranges of variable parameters \((u(t), T_{sou})\), for which the crystallization process is admissible, are as follows: \((10 < u(t) < 25, 1900 < T_{sou} < 1950)\). In this case even smaller speeds \( u(t) \) and higher temperatures \( T_{sou} \) are acceptable than when \( \text{DISPL} = 360 \). For temperatures \( T_{sou} < 1900 \), as in the case of \( \text{DISPL} = 360 \), there appear liquid metal ”bubbles”, and for \( T_{sou} > 1950 \) the process of solidification is unacceptably long. It should be noted that if the object is immersed up to the fifth stage \( \text{DISPL} = 420 \), then the best results (in terms of the objective function) occur at the lowest allowable speed \( u(t) = 10 \). In this case, as before, the value of the cost functional decreases when the furnace temperature \( T_{sou} \) increases.

\( \text{DISPL} = 450 \) (almost complete immersion)

With this depth of immersion the following ranges of parameters were acceptable: \((5 < u(t) < 25, 1970 < T_{sou} < 2015)\). Qualitatively, the characteristics of the crystallization process behave similarly to the case when \( \text{DISPL} = 420 \). The objective functional \( I(u) \) decreases with the increase of the furnace temperature and its minimum value is achieved when the object’s velocity \( u(t) \approx 10 \). We’d like to note that this minimum is the smallest among all obtained in the series of calculations with varying depth \( \text{DISPL} \).

The results of calculations performed for different depths of immersion \( \text{DISPL} \) allow us to conclude that in each case there is a range of acceptable furnace temperatures \( T_{sou} \), the limits of which increase with depth \( \text{DISPL} \). The minimum is delivered to cost functional \( I(u) \) at the upper boundary of this range. There is also a range for the speed \( u(t) \)
of the object in the oven. This range gets wider with an increase in the maximum depth of immersion. The value of the functional $I(u)$ reaches its minimum value at the lower boundary of this range for small values of $DISPL$ and gets inside this range for higher ones. The value of $u(t)$ at which $I(u)$ is minimal decreases when $DISPL$ increases.

This work was supported by the Russian Foundation for Basic Research (project No. 12-01-00572- and No. 11-01-12136-ofi-m-2011), by the Program for Fundamental Research of Presidium of RAS P18, and by the Program ”Leading Scientific Schools” (NSh-5264.2012.1).

REFERENCES


Convergence of the two-step extragradient method in a finite number of iterations

Anna Zykina¹, Nikolay Melenchuk²

¹ Omsk State Technical University, Omsk, Russian; avzykina@mail.ru
² Omsk State Technical University, Omsk, Russian; melenchukanv@gmail.com

The extragradient methods [1] are an effective tool of solving of the variational inequalities. For convergence they require fewer conditions than the gradient methods need. The monotone convergence in the norm to one of the solutions of the variational inequalities is known for the extragradient methods. In the present paper the two-step method of solving of the variational inequalities is considered. Convergence of this method
in a finite number of the iterations is shown if the severity condition is fulfilled.

To solve the variational inequality means to find a vector \( z^* \in \Omega \) that satisfies the following conditions:

\[
\langle H(z^*), z - z^* \rangle \geq 0, \quad \forall z \in \Omega.
\]  

(1)

Here \( H : \mathbb{R}^n \to \mathbb{R}^n \), \( \Omega \) – a convex closed set \( \Omega \subset \mathbb{R}^n \).

The expressions

\[
\begin{align*}
\tilde{z}^k &= P_\Omega(z^k - \alpha H(z^k)), \\
\bar{z}^k &= P_\Omega(\tilde{z}^k - \alpha H(\tilde{z}^k)), \\
z^{k+1} &= P_\Omega(z^k - \alpha H(\bar{z}^k)).
\end{align*}
\]

(2)

are the recurrence relations for the two-step extragradient method of solving the variational inequalities \([2]\).

A condition providing convergence of a computational scheme (2) in a finite number of iterations is the severity condition, which holds (is fulfilled) for the solutions \( z^* \) of the variational inequality with some \( \gamma > 0 \):

\[
\langle H(z), z - z^* \rangle \geq \gamma \|z - z^*\|, \quad \forall z \in \Omega.
\]  

(3)

We demonstrate that the condition (3) means an uniqueness of the solution of the variational inequality. We consider two solutions of the variational inequality \( z^*_1 \) and \( z^*_2 \). Let they are different, and the severity condition (3) is fulfilled for each solution:

\[
\begin{align*}
\langle H(z), z - z^*_1 \rangle &\geq \gamma \|z - z^*_1\|, \quad \forall z \in \Omega, \\
\langle H(z), z - z^*_2 \rangle &\geq \gamma \|z - z^*_2\|, \quad \forall z \in \Omega.
\end{align*}
\]

(4)

The point \( z^*_2 \) belongs to the domain \( \Omega \) of finding of a solution, then it can be substituted in the inequality (3) instead of \( z \). We obtain

\[
\langle H(z_2^*), z_2^* - z_1^* \rangle \geq \gamma \|z_2^* - z_1^*\|,
\]

(5)

From the inequality (1) at \( z^* = z_2^* \) and \( z = z_1^* \) we have

\[
\langle H(z_2^*), z_1^* - z_2^* \rangle \geq 0.
\]

(6)
The inequalities (5) and (6) are filled simultaneously at $z_1^* = z_2^*$. Consequently, the uniqueness of solution of the variational inequality (1) holds if the condition of severity (3) is fulfilled. This fact narrows significantly a class of problems satisfying these conditions.

Under the condition of severity (3) for the variational inequality (1) the theorem is proved.

**Theorem.** Let:
\begin{enumerate} 
    \item $\Omega$ is a closed convex set;
    \item $H(z)$ is a monotone operator: $\langle H(z) - H(v), z - v \rangle \geq 0$, $\forall z, v \in \Omega$, satisfying the Lipshitz condition with a constant $L > 0$:
    \[ \|H(z) - H(v)\| \leq L\|z - v\|, \forall z, v \in \Omega; \]
    \item there exists a solution $z^*$ of the variational inequality (1);
    \item the severity condition (3) with a constant $\gamma > 0$ is fulfilled;
    \item $0 < \alpha < \frac{1}{\sqrt{3L}}$.
\end{enumerate}
Then the sequence $\{z^k\}$, defined by the recurrence relations (2), converges to the solution $z^* \in \Omega$ of the variational inequality (1) in a finite number of iterations.

As an example a question of convergence in a finite step number of the two-step extragradient method for a linear programming problem is considered [3]:
\begin{align*}
    \bar{x}^k &= [x^k + \alpha(Ay^k - c)]^+, \\
    \bar{y}^k &= [y^k - \alpha(ATx^k - b)]^+, \\
    \tilde{x}^k &= [\tilde{x}^k + \alpha(A\tilde{y}^k - c)]^+, \\
    \tilde{y}^k &= [\tilde{y}^k - \alpha(AT\tilde{x}^k - b)]^+, \\
    x^{k+1} &= [x^k + \alpha(A\tilde{y}^k - c)]^+, \\
    y^{k+1} &= [y^k - \alpha(AT\tilde{x}^k - b)]^+.
\end{align*}

Here a pair of the dual linear programming problems is given as follows:
\begin{align*}
    \langle c, x \rangle &\rightarrow \text{min}, & \langle b, y \rangle &\rightarrow \text{max}, \\
    ATx &\geq b, & Ay &\leq c,
\end{align*}
\[ x \geq 0, \quad y \geq 0, \]

\[ [p]^+ = \max\{0, p\} \] is the projection on the positive semiaxis for the scalar quantity \( p \) and \([p]^+ = ([p_1]^+, [p_2]^+, \ldots, [p_l]^+)\) is the projection on the positive orthant for the vector \( p = (p_1, p_2, \ldots, p_l) \).

The linear programming problem (8) is a special case of a variational inequality. Interest to this problem is explained by a possibility to demonstrate the results in the practice. In the case of an one-step extragradient method the question of convergence in a finite step number for a linear programming problem (8) has been studied by A.S. Antipin in [4].

The authors were supported by the Russian Foundation for Basic Research (project no. 12-07-00326).

REFERENCES


### Author index

<table>
<thead>
<tr>
<th>Author</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Abdullayev Vagif</td>
<td>10</td>
</tr>
<tr>
<td>Abramov Alexander P.</td>
<td>13</td>
</tr>
<tr>
<td>Abramov Oleg V.</td>
<td>16</td>
</tr>
<tr>
<td>Afanasyev Alexander</td>
<td>21</td>
</tr>
<tr>
<td>Ahookhosh Masoud</td>
<td>30</td>
</tr>
<tr>
<td>Aida-zade Kamil</td>
<td>21</td>
</tr>
<tr>
<td>Akbari Z.</td>
<td>25</td>
</tr>
<tr>
<td>Albu Alla</td>
<td>270</td>
</tr>
<tr>
<td>Albu Andrey</td>
<td>270</td>
</tr>
<tr>
<td>Amini Keyvan</td>
<td>30</td>
</tr>
<tr>
<td>An Phan Thanh</td>
<td>93, 106</td>
</tr>
<tr>
<td>Anikin Anton</td>
<td>31</td>
</tr>
<tr>
<td>Anop Maxim</td>
<td>34</td>
</tr>
<tr>
<td>Antipin Anatoly</td>
<td>36</td>
</tr>
<tr>
<td>Arkhipov Dmitry</td>
<td>42</td>
</tr>
<tr>
<td>Ashafova Yegana</td>
<td>21</td>
</tr>
<tr>
<td>Attari M. Ahmadian</td>
<td>218</td>
</tr>
<tr>
<td>Bahrami Somayeh</td>
<td>30</td>
</tr>
<tr>
<td>Beklaryan Armen</td>
<td>47</td>
</tr>
<tr>
<td>Bushenkov Vladimir</td>
<td>51</td>
</tr>
<tr>
<td>Caldeira Bento</td>
<td>51</td>
</tr>
<tr>
<td>Constantino Miguel</td>
<td>54</td>
</tr>
<tr>
<td>Dikusar V.V.</td>
<td>59</td>
</tr>
<tr>
<td>Dolgui Alexandre</td>
<td>81</td>
</tr>
<tr>
<td>Dorjievu Anna</td>
<td>63</td>
</tr>
<tr>
<td>Druzhinina Olga</td>
<td>67</td>
</tr>
<tr>
<td>Dunin-Barkowski W.</td>
<td>72</td>
</tr>
<tr>
<td>Evtushenko Yu.G.</td>
<td>76</td>
</tr>
<tr>
<td>Freitas Adelaide</td>
<td>164</td>
</tr>
<tr>
<td>Førsund Finn</td>
<td>144</td>
</tr>
<tr>
<td>Gafarov Evgeny R.</td>
<td>81</td>
</tr>
<tr>
<td>Galiev Shamil</td>
<td>86</td>
</tr>
<tr>
<td>Gasnikov Alexander</td>
<td>91</td>
</tr>
<tr>
<td>Gasnikova Eugenia</td>
<td>91</td>
</tr>
<tr>
<td>Gholami Omid</td>
<td>245</td>
</tr>
<tr>
<td>Giang Dinh Thanh</td>
<td>93</td>
</tr>
<tr>
<td>Golikov Alexander I.</td>
<td>94</td>
</tr>
<tr>
<td>Golshtejn Evgenij</td>
<td>99</td>
</tr>
<tr>
<td>Goncharov Vladimir</td>
<td>101</td>
</tr>
<tr>
<td>Gornov Aleksander</td>
<td>103</td>
</tr>
<tr>
<td>Guerman Anna D.</td>
<td>106</td>
</tr>
<tr>
<td>Hai Nguyen Ngoc</td>
<td>106</td>
</tr>
<tr>
<td>Husnullin Nail</td>
<td>187</td>
</tr>
<tr>
<td>Ismagilov Niyaz</td>
<td>111</td>
</tr>
<tr>
<td>Jaćimović M.</td>
<td>115</td>
</tr>
<tr>
<td>Kamenev Georgij</td>
<td>159</td>
</tr>
<tr>
<td>Kaneva Olga</td>
<td>147</td>
</tr>
<tr>
<td>Kasenov Syrym</td>
<td>192</td>
</tr>
<tr>
<td>Katueva Yaroslava</td>
<td>34</td>
</tr>
<tr>
<td>Khachaturov Ruben V.</td>
<td>118</td>
</tr>
<tr>
<td>Khachaturov Vladimir R.</td>
<td>122</td>
</tr>
<tr>
<td>Khachay Michael</td>
<td>124</td>
</tr>
<tr>
<td>Khoroshilova Elena</td>
<td>128</td>
</tr>
<tr>
<td>Klimentova Xenia</td>
<td>54</td>
</tr>
<tr>
<td>Kobylkin Konstantin</td>
<td>133</td>
</tr>
<tr>
<td>Korenev Pavel</td>
<td>134</td>
</tr>
<tr>
<td>Kostyukova Olga</td>
<td>139</td>
</tr>
<tr>
<td>Krivonozhko Vladimir</td>
<td>144</td>
</tr>
<tr>
<td>Krnić I.</td>
<td>115</td>
</tr>
<tr>
<td>Kulbida Ulyana</td>
<td>147</td>
</tr>
<tr>
<td>Kuliev Samir</td>
<td>153</td>
</tr>
<tr>
<td>Laskova Maya</td>
<td>156</td>
</tr>
<tr>
<td>Name</td>
<td>Page Numbers</td>
</tr>
<tr>
<td>-----------------------------</td>
<td>--------------------</td>
</tr>
<tr>
<td>Lazarev Alexander</td>
<td>42, 81, 134, 156, 187</td>
</tr>
<tr>
<td>Lisafina Maria</td>
<td>86</td>
</tr>
<tr>
<td>Lotov Alexander V.</td>
<td>159</td>
</tr>
<tr>
<td>Lukovenko A.A.</td>
<td>162</td>
</tr>
<tr>
<td>Lychev Andrey</td>
<td>144</td>
</tr>
<tr>
<td>Macedo Eloísa</td>
<td>164</td>
</tr>
<tr>
<td>Magafurzyanova Nelli</td>
<td>223</td>
</tr>
<tr>
<td>Melenchuk Nikolay</td>
<td>274</td>
</tr>
<tr>
<td>Mikhailov Igor E.</td>
<td>170</td>
</tr>
<tr>
<td>Mizhidon Arsalan</td>
<td>175</td>
</tr>
<tr>
<td>Mordukhovich Boris</td>
<td>178</td>
</tr>
<tr>
<td>Murashkin Evgenii</td>
<td>178</td>
</tr>
<tr>
<td>Muravey L.A.</td>
<td>170, 181</td>
</tr>
<tr>
<td>Musatova Elena</td>
<td>42, 156, 187</td>
</tr>
<tr>
<td>Nasyrov Farit</td>
<td>111</td>
</tr>
<tr>
<td>Nurseitov Daniar</td>
<td>192</td>
</tr>
<tr>
<td>Obrosova Natalia</td>
<td>194</td>
</tr>
<tr>
<td>Orlov Andrei</td>
<td>198</td>
</tr>
<tr>
<td>Parkhomenko Valeriy</td>
<td>203</td>
</tr>
<tr>
<td>Pereira Fátima</td>
<td>101</td>
</tr>
<tr>
<td>Perzhabinsky Sergey</td>
<td>207</td>
</tr>
<tr>
<td>Petrov Lev F.</td>
<td>213</td>
</tr>
<tr>
<td>Petrov V.M.</td>
<td>181</td>
</tr>
<tr>
<td>Petrova Natalia</td>
<td>67</td>
</tr>
<tr>
<td>Peyghami M. R.</td>
<td>25, 218</td>
</tr>
<tr>
<td>Pilnik N.P.</td>
<td>256</td>
</tr>
<tr>
<td>Pinigin Sergei</td>
<td>198</td>
</tr>
<tr>
<td>Plakhov Alexander</td>
<td>219</td>
</tr>
<tr>
<td>Pobertii Maria</td>
<td>124</td>
</tr>
<tr>
<td>Polyak Boris</td>
<td>220</td>
</tr>
<tr>
<td>Pospelov I.G.</td>
<td>256</td>
</tr>
<tr>
<td>Posypkin Mikhail</td>
<td>221</td>
</tr>
<tr>
<td>Putilina Elena</td>
<td>21</td>
</tr>
<tr>
<td>Rassadnikova Ekaterina</td>
<td>223</td>
</tr>
<tr>
<td>Romanenko A.M.</td>
<td>181</td>
</tr>
<tr>
<td>Sedova Nataliya</td>
<td>227</td>
</tr>
<tr>
<td>Serovajsky Simon</td>
<td>229</td>
</tr>
<tr>
<td>Shakenov Iliyas</td>
<td>234</td>
</tr>
<tr>
<td>Shakenov Kanat</td>
<td>237</td>
</tr>
<tr>
<td>Shaninin Alexander</td>
<td>194</td>
</tr>
<tr>
<td>Shcherbina O.</td>
<td>240</td>
</tr>
<tr>
<td>Shishlenin Maksim</td>
<td>192</td>
</tr>
<tr>
<td>Sigal Izrael</td>
<td>221</td>
</tr>
<tr>
<td>Smirnov Georgi</td>
<td>51</td>
</tr>
<tr>
<td>Sotskov Yuri N.</td>
<td>245</td>
</tr>
<tr>
<td>Sviridenko A.</td>
<td>240</td>
</tr>
<tr>
<td>Tavakoli H.</td>
<td>218</td>
</tr>
<tr>
<td>Tchemisova Tatiana</td>
<td>139</td>
</tr>
<tr>
<td>Tikhomirova T.M.</td>
<td>162</td>
</tr>
<tr>
<td>Trang Le Hong</td>
<td>93</td>
</tr>
<tr>
<td>Tretyakov A.A.</td>
<td>76, 250</td>
</tr>
<tr>
<td>Valeeva Aida</td>
<td>223</td>
</tr>
<tr>
<td>Viana Ana</td>
<td>54</td>
</tr>
<tr>
<td>Vorontsova Evgeniya A.</td>
<td>254</td>
</tr>
<tr>
<td>Vrzheshch V.P.</td>
<td>256</td>
</tr>
<tr>
<td>Vyshinskii L.</td>
<td>72</td>
</tr>
<tr>
<td>Werner Frank</td>
<td>245</td>
</tr>
<tr>
<td>Yousefpour R.</td>
<td>25, 257</td>
</tr>
<tr>
<td>Yudin Vitalii</td>
<td>86</td>
</tr>
<tr>
<td>Zasukhina E.S.</td>
<td>59</td>
</tr>
<tr>
<td>Zhadan Vitaly</td>
<td>262</td>
</tr>
<tr>
<td>Zorkaltsev Valery</td>
<td>207, 267</td>
</tr>
<tr>
<td>Zubov Vladimir</td>
<td>270</td>
</tr>
<tr>
<td>Zykin Anna</td>
<td>274</td>
</tr>
</tbody>
</table>

279