ISSN 0081-5438, Proceedings of the Steklov Institute of Mathematics, 2015, Vol. 289, Suppl. 1, pp. S102–S110.

published in Trudy Instituta Matematiki i Mekhaniki UrO RAN, 2014, Vol. 20, No. 2.

Regularization and Normal Solutions of Systems of Linear Equations and Inequalities

A. I. Golikov[1] and Yu. G. Evtushenko[1]

Received January 14, 2014

Abstract—The paper provides some examples of mutually dual unconstrained optimization problems originating from regularization problems for systems of linear equations and/or inequalities. The solution of each of these mutually dual problems can be found from the solution of the other problem by means of simple formulas. Since mutually dual problems have different dimensions, it is natural to solve the unconstrained optimization problem of the smaller dimension.

Keywords: regularization, piecewise quadratic function, unconstrained optimization, mutually dual problems, generalized Newton method.

DOI: 10.1134/S0081543815050090

INTRODUCTION

I.I. Eremin is well known as the author of duality theory for improper linear optimization problems. He always paid much attention to the detection of duality in problems arising in different optimization methods. The authors of this paper were greatly impressed by the remark made by Eremin at one of the conferences "Mathematical Programming and Applications" that a dual unconstrained problem of quadratic programming can be considered as mutually dual for an original primal constrained problem of quadratic programming [1, 2].

Formally, unconstrained minimization problems have no Lagrange function, and, consequently, a dual problem cannot be directly constructed for them. Nevertheless, using additional variables, we can introduce artificial constraints and obtain an equivalent problem of nonlinear programming, for which the dual problem can be defined in the standard way. There exists a class of optimization problems for which mutually dual problems are problems of unconstrained optimization and the solution of each of these two problems can be expressed in terms of the solution of the other. These are problems of quadratic programming, which arise, for example, in the regularization of systems of linear equations and/or inequalities. Since mutually dual problems have different dimensions, it is natural to solve the unconstrained optimization problem of the smaller dimension. We present a typical result appearing under the regularization of a system of linear equations and inequalities [3] and in the SVM method of pattern recognition [4].

In Section 1, we consider a regularized problem of solving a system of linear equations. Throughout this paper, we use the Euclidean norm. For the regularized problem, which is a problem of unconstrained minimization of a strictly convex quadratic function, we give a mutually dual problem

[1]Dorodnitsyn Computing Center, Russian Academy of Sciences, ul. Vavilova 40, Moscow, 119333 Russia
email: evt@ccas.ru

of unconstrained maximization of a strictly concave quadratic function. We obtain simple formulas according to which the solution of each of these problems is found from the solution of the other problem. We also consider some approaches to finding a normal solution of a system of linear equations that are different from the regularization method.

In Sections 2 and 3, we consider similar mutually dual problems of finding normal solutions for systems of linear equations with nonnegative variables and systems of linear inequalities, respectively. Here, problems of unconstrained optimization of piecewise quadratic functions arise, for which the generalized Newton method, which is globally convergent in a finite number of steps, is especially efficient.

1. NORMAL SOLUTIONS OF SYSTEMS OF LINEAR EQUATIONS

Consider a consistent system of linear equations

$$Ax = b. \tag{1.1}$$

Here, A is a nonzero $m \times n$ matrix and the vector $b \in \mathbb{R}^m$ is such that $\|b\| \neq 0$. The regularization method involves a sequence of unconstrained minimization problems

$$\min_{x \in \mathbb{R}^n} F(x), \qquad F(x) = \frac{1}{2}(\|b - Ax\|^2 + \varepsilon\|x\|^2) \tag{1.2}$$

with a positive parameter ε tending to zero. The unique solution $x(\varepsilon)$ of problem (1.2) for fixed ε is expressed explicitly by the formula

$$x(\varepsilon) = (\varepsilon I_n + A^\top A)^{-1} A^\top b. \tag{1.3}$$

Here and below, I_k denotes the unit matrix of order k. The solution $x(\varepsilon)$ converges as $\varepsilon \to 0$ to a normal solution of system (1.1) [3]. In (1.3), the inverse matrix exists for any rank of the matrix A and for any $\varepsilon > 0$.

Expression (1.3) for calculating $x(\varepsilon)$ can be represented in another form by means of the Sherman–Morrison–Woodbury formula [5]:

$$x(\varepsilon) = \frac{1}{\varepsilon}\big(I_n - A^\top(\varepsilon I_m + AA^\top)^{-1}A\big)A^\top b. \tag{1.4}$$

Note that, in this formula, a square matrix of order m is inverted, in contrast to formula (1.3), where a matrix of order n is inverted. Below, we derive one more formula (1.17) to calculate $x(\varepsilon)$, which also involves the inversion of a matrix of order m but with a fewer number of matrix multiplications than in formula (1.4).

It is possible to consider problem (1.2) from different points of view: as the quadratic penalty function method, the regularization of a linear programming problem with zero objective function, the least squares method, or multicriteria optimization. For example, (1.2) is an auxiliary problem of the penalty method with a penalty coefficient at the objective function for the following problem of quadratic programming:

$$\min_{x \in X} \frac{1}{2}\|x\|^2, \qquad X = \{x \in \mathbb{R}^n \colon Ax = b\}. \tag{1.5}$$

Problem (1.2) is equivalent to the regularized linear programming problem

$$\min_{x\in X}\{0_n^\top x+\frac{\varepsilon}{2}\|x\|^2\},\qquad X=\{x\in\mathbb{R}^n\colon Ax=b\}. \tag{1.6}$$

It is easy to show that the problem dual to (1.6) is the following unconstrained maximization of a quadratic function [1] (Eremin showed that these problems are mutually dual):

$$\max_{u\in\mathbb{R}^n}\{b^\top u-\frac{1}{2\varepsilon}\|A^\top u\|^2\}. \tag{1.7}$$

This problem, in turn, is a penalized version of the linear programming (LP) problem

$$\max_{u\in U} b^\top u,\qquad U=\{u\in\mathbb{R}^m\colon A^\top u=0_n\}.$$

Note that the solution $u(\varepsilon)$ of unconstrained maximization problem (1.7) can be used to calculate a solution of problem (1.6) by the formula

$$x(\varepsilon)=\frac{1}{\varepsilon}A^\top u(\varepsilon).$$

According to [6, 7], this formula gives a solution of problem (1.5) for any $\varepsilon>0$.

For a fixed parameter ε, we can consider problem (1.2) as the least squares method (the minimal residual method) applied to the inconsistent system

$$Ax=b,\qquad -\sqrt{\varepsilon}x=0_n. \tag{1.8}$$

The vector $x(\varepsilon)$ is a solution of unconstrained minimization problem (1.2) and a pseudosolution of system (1.8). By $z_1(\varepsilon)=b-Ax(\varepsilon)$ and $z_2(\varepsilon)=\sqrt{\varepsilon}x(\varepsilon)$, we denote the components of the minimal residual vector $z(\varepsilon)^\top=[z_1(\varepsilon)^\top,z_2(\varepsilon)^\top]$ for system (1.8).

According to the theorem of alternatives (see, for example, [8, 9]) for inconsistent system (1.8), we can construct a consistent alternative system of the form

$$A^\top u_1-\sqrt{\varepsilon}u_2=0_n,\qquad b^\top u_1=\rho>0. \tag{1.9}$$

Here, ρ is a fixed positive constant, and the vectors of unknowns are $u_1\in\mathbb{R}^m$ and $u_2\in\mathbb{R}^n$. According to [9], the normal vector $\tilde{u}(\varepsilon)^\top=[\tilde{u}_1(\varepsilon)^\top,\tilde{u}_2(\varepsilon)^\top]$ of alternative system (1.9) is expressed in terms of the minimal residual vector $z(\varepsilon)$ by the formulas

$$\tilde{u}_1(\varepsilon)=\frac{\rho z_1(\varepsilon)}{\|z(\varepsilon)\|^2},\qquad \tilde{u}_2(\varepsilon)=\frac{\rho z_2(\varepsilon)}{\|z(\varepsilon)\|^2}.$$

Let the vector of variables $z\in\mathbb{R}^{m+n}$ be decomposed as $z^\top=[z_1^\top,z_2^\top]$, where $z_1\in\mathbb{R}^m$ and $z_2\in\mathbb{R}^n$. We write the problem of strictly concave quadratic programming

$$\max_{z\in Z}\Big\{b^\top z_1-\frac{1}{2}(\|z_1\|^2+\|z_1\|^2)\Big\},\qquad Z=\{z\in\mathbb{R}^{m+n}\colon A^\top z_1-\sqrt{\varepsilon}z_2=0_n\}. \tag{1.10}$$

This problem can be considered as a regularization of the LP problem

$$\max_{z\in Z}\{b^\top z_1+0_n^\top z_2\},\qquad Z=\{z\in\mathbb{R}^{m+n}\colon A^\top z_1-\sqrt{\varepsilon}z_2=0_n\},$$

which is mutually dual to the LP problem

$$\min_{x \in X} 0_n^\top x, \qquad X = \{x \in \mathbb{R}^n \colon Ax = b, \sqrt{\varepsilon} x = 0_n\}. \tag{1.11}$$

For $\varepsilon \neq 0$ and $\|b\| \neq 0$, the constraints in (1.11) are inconsistent, and the problem is improper of the first kind [10]. We can consider (1.2) as an auxiliary problem for the quadratic penalty function method applied to LP problem (1.11). It is known [1] that the problem of the quadratic penalty function method for the LP problem and the regularized LP problem are mutually dual; i.e., problems (1.2) and (1.10) are mutually dual.

In problem (1.10), we can eliminate the variables z_2 by expressing them in terms of z_1 and substitute $z_2 = (1/\sqrt{\varepsilon})A^\top z_1$ into the objective function of (1.10). Then, we come to the following equivalent problem of unconstrained maximization for a strictly concave quadratic function:

$$\max_{z_1 \in \mathbb{R}^m} H(z_1), \qquad H(z_1) = b^\top z_1 - \frac{1}{2\varepsilon}\|A^\top z_1\|^2 - \frac{1}{2}\|z_1\|^2. \tag{1.12}$$

Thus, this problem and problem (1.2) are mutually dual.

Theorem 1.1. *For any $\varepsilon > 0$, the unique solution $x(\varepsilon) = \operatorname{Arg} \min_{x \in \mathbb{R}^n} F(x)$ of problem* (1.2) *and the unique solution $z_1(\varepsilon) = \operatorname{Arg} \max_{z_1 \in \mathbb{R}^m} H(z_1)$ of problem* (1.12) *are related by the equations*

$$x(\varepsilon) = \frac{1}{\varepsilon}A^\top z_1(\varepsilon), \tag{1.13}$$

$$z_1(\varepsilon) = b - Ax(\varepsilon), \tag{1.14}$$

and the equality of the optimal values of the objective functions holds: $F(x(\varepsilon)) = H(z_1(\varepsilon))$.

Proof. For $\varepsilon > 0$, the strictly convex quadratic function F(x) on $\mathbb{R}^n$ is bounded from below by zero. Therefore, by the Frank–Wolfe theorem [11], problem (1.2) always has a unique solution.

The maximized quadratic function $H(z_1)$ for $\varepsilon > 0$ is strictly concave and is bounded from above on the whole space $\mathbb{R}^m$. Indeed, the following expressions are valid:

$$\begin{aligned} H(z_1) = b^\top z_1 - \frac{1}{2\varepsilon}\|A^\top z_1\|^2 - \frac{1}{2}\|z_1\|^2 &= \frac{1}{2}\|b\|^2 - \frac{1}{2}\|b\|^2 + b^\top z_1 - \frac{1}{2}\|z_1\|^2 - \frac{1}{2\varepsilon}\|A^\top z_1\|^2 \\ &= \frac{1}{2}\|b\|^2 - \frac{1}{2}\|b - z_1\|^2 - \frac{1}{2\varepsilon}\|A^\top z_1\|^2 \leq \frac{1}{2}\|b\|^2. \end{aligned}$$

Consequently, problem (1.10) for any $\varepsilon > 0$ always has a unique solution.

For mutually dual problems (1.2) and (1.12), by the weak duality theorem, the following inequality holds for any $z_1 \in \mathbb{R}^m$, $x \in \mathbb{R}^n$, and $\varepsilon > 0$:

$$b^\top z_1 - \frac{1}{2\varepsilon}\|A^\top z_1\|^2 - \frac{1}{2}\|z_1\|^2 \leq \frac{1}{2}(\|b - Ax\|^2 + \varepsilon\|x\|^2).$$

By the duality theorem, the optimal values of the objective functions of these problems coincide:

$$b^\top z_1(\varepsilon) - \frac{1}{2\varepsilon}\|A^\top z_1(\varepsilon)\|^2 - \frac{1}{2}\|z_1(\varepsilon)\|^2 = \frac{1}{2}(\|b - Ax(\varepsilon)\|^2 + \varepsilon\|x(\varepsilon)\|^2). \tag{1.15}$$

It is easily seen that expressions (1.13) and (1.14) satisfy the necessary and sufficient optimality conditions for problem (1.2)

$$-A^\top(b - Ax(\varepsilon)) + \varepsilon x(\varepsilon) = 0_n$$

and the necessary and sufficient optimality conditions for problem (1.12)

$$b - z_1(\varepsilon) - \frac{1}{\varepsilon}AA^\top z_1(\varepsilon) = 0_m.$$

The theorem is proved.

Note that, in view of relations (1.13) and (1.14) between the solutions $x(\varepsilon)$ and $z(\varepsilon)$, equality (1.15) of the optimal values of the objective functions for mutually dual problems (1.2) and (1.12) is representable in the following two forms, each containing the solution of only one of the problems:

$$b^\top z_1(\varepsilon) = \|z(\varepsilon)\|^2,$$
$$b^\top(b - Ax(\varepsilon)) = \|b - Ax(\varepsilon)\|^2 + \varepsilon\|x(\varepsilon)\|^2. \tag{1.16}$$

As follows from (1.16), $\varepsilon\|x(\varepsilon)\|^2 = z_1(\varepsilon)^\top Ax(\varepsilon)$. Hence, we obtain a known result in the least squares method: $z_1(0) \perp Ax(0)$ for $\varepsilon = 0$. Note that, if system (1.1) is inconsistent, then $z_1(0) \neq 0$.

Problem (1.12) for any $\varepsilon > 0$ and for any rank of the matrix A can be solved explicitly:

$$z_1(\varepsilon) = \varepsilon(\varepsilon I_m + AA^\top)^{-1}b.$$

Substituting this expression into (1.13), we obtain one more formula to calculate the solution $x(\varepsilon)$ of problem (1.2):

$$x(\varepsilon) = A^\top(\varepsilon I_m + AA^\top)^{-1}b. \tag{1.17}$$

This formula involves the inversion of a matrix of order m, in contrast to formula (1.3), where it is required to invert a matrix of order n. Therefore, if $m < n$ in problem (1.1), then it is reasonable to use formula (1.17) or formula (1.13) for calculating $x(\varepsilon)$. In the case when formula (1.13) is applied, one must solve unconstrained optimization problem (1.12) with m unknowns.

Expression (1.17) for calculating $x(\varepsilon)$ can be represented in another form by means of the Sherman–Morrison–Woodbury formula

$$x(\varepsilon) = \frac{1}{\varepsilon}A^\top(I_m - A(\varepsilon I_n + A^\top A)^{-1}A^\top)b.$$

In this formula, it is required to invert a matrix of order n, in contrast to formula (1.17), where a matrix of order m is inverted.

It is shown in [12] that, for any matrix A from system (1.1), its pseudoinverse matrix can be defined as follows:

$$A^+ = \lim_{\varepsilon\to 0}(\varepsilon I_n + A^\top A)^{-1}A^\top = \lim_{\varepsilon\to 0}A^\top(\varepsilon I_m + AA^\top)^{-1},$$

and, for any vector b (in particular, for b that makes system (1.1) inconsistent), $\tilde{x}_* = A^+b$ is a vector with minimum norm among all vectors minimizing $\|b - Ax\|^2$.

In [9], another method for finding a normal solution of system (1.1) is proposed. This method is based on the application of theorems on alternatives. In this case, system (1.1) has the inconsistent alternative system

$$A^\top u = 0_n, \qquad b^\top u = \rho \neq 0,$$

for which the following problem of minimization of its residuals is solved:

$$\min_{u\in\mathbb{R}^m} \frac{1}{2}\{\|A^\top u\|^2 + (\rho - b^\top u)^2\}. \tag{1.18}$$

Here, ρ is an arbitrary nonzero constant.

Let u_* be a solution of unconstrained minimization problem (1.18). Then, the normal solution of original system (1.1) is expressed by the formula

$$\tilde{x}_* = \frac{A^\top u_*}{\rho - b^\top u_*}. \tag{1.19}$$

If the rank of the matrix A in the original system is m, then the solution of problem (1.18) is found analytically as

$$u_* = \rho(AA^\top + bb^\top)^{-1}b.$$

Substituting this formula into (1.19), we obtain one more expression for the normal solution of (1.1):

$$\tilde{x}_* = \frac{A^\top(AA^\top + bb^\top)^{-1}b}{1 - b^\top(AA^\top + bb^\top)^{-1}b}.$$

2. REGULARIZATION OF SYSTEMS OF LINEAR EQUATIONS WITH NONNEGATIVE VARIABLES

Consider now the consistent system of linear equations with nonnegative variables

$$Ax = b, \qquad x \geq 0_n,$$

and its regularized problem

$$\min_{x\in\mathbb{R}^n_+} F(x), \qquad F(x) = \frac{1}{2}(\|b - Ax\|^2 + \varepsilon\|x\|^2) \tag{2.1}$$

with a positive parameter ε tending to zero.

The unconstrained maximization problem

$$\max_{u\in\mathbb{R}^m} H(u), \qquad H(u) = b^\top u - \frac{1}{2\varepsilon}\|(A^\top u)_+\|^2 - \frac{1}{2}\|u\|^2 \tag{2.2}$$

is dual to regularized problem (2.1). Here and below, a_+ denotes a vector a in which all negative components are replaced by zeros.

Theorem 2.1. *For any $\varepsilon > 0$, the unique solution $x(\varepsilon) = \operatorname{Arg}\min_{x\in\mathbb{R}^n_+} F(x)$ of problem* (2.1) *and the unique solution $u(\varepsilon) = \operatorname{Arg}\max_{u\in\mathbb{R}^m} H(u)$ of problem* (2.2) *are related by the equations*

$$x(\varepsilon) = \frac{1}{\varepsilon}(A^\top u(\varepsilon))_+,$$
$$u(\varepsilon) = b - Ax(\varepsilon),$$

and the equality of the optimal values of the objective functions holds: $F(x(\varepsilon)) = H(u(\varepsilon))$.

Proof. The minimand in problem (2.1) for $\varepsilon > 0$ is a strictly convex quadratic function bounded from below by zero on $\mathbb{R}^n_+$. Therefore, by the Frank–Wolfe theorem [11], problem (2.1) always has a solution, which is unique.

In problem (2.2), the maximized piecewise quadratic function for $\varepsilon > 0$ is strictly concave and bounded from above on the whole space $\mathbb{R}^m$. By the Frank–Wolfe theorem, problem (2.2) always has a unique solution.

Introducing additional variables $y \in \mathbb{R}^m$ and constraints $Ax + y = b$, we rewrite problem (2.1) in the equivalent form

$$\min_{y\in\mathbb{R}^m} \min_{x\in\mathbb{R}^n_+} \frac{1}{2}\{\|y\|^2 + \varepsilon\|x\|^2\}, \tag{2.3}$$
$$Ax + y = b.$$

For this problem, we introduce the Lagrange function

$$L(y, x, u) = \frac{1}{2}\|y\|^2 + \frac{\varepsilon}{2}\|x\|^2 + u^T(b - Ax - y).$$

Here, $u \in \mathbb{R}^m$ are Lagrange multipliers for problem (2.3). The problem dual to (2.3) has the form

$$\max_{u\in\mathbb{R}^m} \min_{y\in\mathbb{R}^m} \min_{x\in\mathbb{R}^n_+} L(y, x, u). \tag{2.4}$$

Let us write the minimum condition in y and x for the inner problem in (2.4):

$$L_y(y(\varepsilon), x(\varepsilon), u) = y(\varepsilon) - u = 0_m,$$
$$L_x(y(\varepsilon), x(\varepsilon), u) = \varepsilon x(\varepsilon) - A^\top u \ge 0_n, \qquad x^\top(\varepsilon)(\varepsilon x(\varepsilon) - A^\top u) = 0, \qquad x(\varepsilon) \ge 0_n.$$

From these conditions, we easily find solutions of the inner minimization problem in (2.4):

$$y(\varepsilon) = u, \tag{2.5}$$
$$x(\varepsilon) = \frac{1}{\varepsilon}(A^\top u)_+. \tag{2.6}$$

Substituting solutions (2.5) and (2.6) into the Lagrange function $L(y, x, u)$, after simple transformations, we obtain the dual function for problem (2.4)

$$H(u) = b^\top u - \frac{1}{2\varepsilon}\|(A^\top u)_+\|^2 - \frac{1}{2}\|u\|^2;$$

i.e., we come to dual problem (2.2) for problem (2.3) and, hence, for problem (2.1). Duality theory implies the equality of the optimal values of the objective functions in problems (2.1) and (2.2).

The outer problem in (2.4) consists in the unconstrained maximization of $H(u)$ with respect to u. The necessary and sufficient maximum condition for this problem is $H_u(u(\varepsilon)) = b - \frac{1}{\varepsilon}A(A^\top u(\varepsilon))_+ - u(\varepsilon) = b - Ax(\varepsilon) - u(\varepsilon) = 0_m$. Hence, by (2.6), we have

$$u(\varepsilon) = b - \frac{1}{\varepsilon}A(A^\top u(\varepsilon))_+ = b - Ax(\varepsilon).$$

The theorem is proved.

It follows from the theorem that, if the number of rows in the $m \times n$ matrix A is $m < n$, then, instead of minimization problem (2.1), it is reasonable to solve dual problem (2.2), which is a concave piecewise quadratic problem of unconstrained maximization. Problem (2.2) can be solved very efficiently by the generalized Newton method. The function $H(u)$, which is maximized in problem (2.2), is concave, piecewise quadratic, and differentiable. The usual Hessian matrix does not exist for this function. Indeed, the gradient of the function $H(u)$

$$H_u(u) = b - \frac{1}{\varepsilon} A(A^\top u)_+ - u$$

is not differentiable. However, for this function, we can define the generalized Hessian matrix, which is a nondegenerate $(m \times m)$-matrix of the form

$$H_{uu} = -\Big(\frac{1}{\varepsilon} A D(z) A^\top + I_m\Big),$$

where $D(z)$ denotes the diagonal $(n \times n)$-matrix with its ith diagonal element z^i equal to 1 if $(A^\top u)^i > 0$ and equal to 0 if $(A^\top u)^i \le 0$, $i = 1, \dots, n$. The proof of the finite global convergence of the generalized Newton method for the unconstrained optimization of a piecewise quadratic function with the step size chosen by the Armijo rule can be found in [4, 13, 14]. The generalized Newton method makes it possible to effectively solve problems on uniprocessor computers for $n \approx 10^6$ and $m \approx 10^4$ and on multiprocessor computer systems for n of the order of tens of millions and m of the order of hundreds of thousands [15].

3. REGULARIZATION OF SYSTEMS OF LINEAR INEQUALITIES

The regularization of systems of linear inequalities is similar to the procedure considered in the preceding section. Let a consistent system of linear inequalities

$$Ax \ge b$$

be given. The regularized problem has the form

$$\min_{x \in \mathbb{R}^n} F(x), \qquad F(x) = \frac{1}{2}(\|(b - Ax)_+\|^2 + \varepsilon \|x\|^2) \tag{3.1}$$

with a positive parameter ε tending to zero. Then, the following maximization problem on the positive orthant

$$\max_{u \in \mathbb{R}^m_+} H(u), \qquad H(u) = b^\top u - \frac{1}{2\varepsilon}\|A^\top u\|^2 - \frac{1}{2}\|u\|^2 \tag{3.2}$$

is dual to regularized problem (3.1).

In this case, we have a theorem similar to Theorem 2.1.

Theorem 3.1. *For any $\varepsilon > 0$, the unique solution $x(\varepsilon) = \operatorname{Arg}\min_{x \in \mathbb{R}^n} F(x)$ of problem* (3.1) *and the unique solution $u(\varepsilon) = \operatorname{Arg}\max_{u \in \mathbb{R}^m_+} H(u)$ of problem* (3.2) *are related by the equations*

$$x(\varepsilon) = \frac{1}{\varepsilon} A^\top u(\varepsilon),$$
$$u(\varepsilon) = (b - Ax(\varepsilon))_+\,,$$

and the equality of the optimal values of the objective functions holds: $F(x(\varepsilon)) = H(u(\varepsilon))$.

Unfortunately, it is difficult to apply the Newton method directly to problem (3.2), in contrast to problem (2.2) [7, 16].

ACKNOWLEDGMENTS

This work was supported by the Russian Foundation for Basic Research (project no. 14-07-00805), by the Program for State Support of Leading Scientific Schools of the Russian Federation (project no. NSh-4640.2014.1), and by Program P-18 of the Presidium of the Russian Academy of Sciences.

REFERENCES

1. I. I. Eremin, "On quadratic and complete quadratic problems of convex programming," Russ. Math. (Izv. VUZ) **42** (12), 20–26 (1998).
2. I. I. Eremin, *Duality Theory in Linear Optimization* (Izd. Yuzhno-Ural'sk. Gos. Univ., Chelyabinsk, 2005) [in Russian].
3. A. N. Tikhonov, "On ill-posed problems in linear algebra and a stable method for their solution," Dokl. Akad. Nauk SSSR **163** (4), 591–594 (1965).
4. O. L. Mangasarian, "A finite Newton method for classification," Optim. Methods Softw. **17** (5), 913–929 (2002).
5. J. M. Ortega and W. C. Reinbolt, *Iterative Solution of Nonlinear Equations in Several Variables* (Academic, San Diego, 1970; Mir, Moscow, 1975).
6. A. I. Golikov and Yu. G. Evtushenko, "Finding the projection of a given point on the set of solutions of a linear programming problem," Proc. Steklov Inst. Math., Suppl. 2, S68–S83 (2008).
7. A. I. Golikov and Yu. G. Evtushenko, "Generalized Newton method for linear optimization problems with inequality constraints," Proc. Steklov Inst. Math. **284** (Suppl. 1), S96–S107 (2014).
8. O. L. Mangasarian, *Nonlinear Programming* (McCrow-Hill, New York, 1969).
9. A. I. Golikov and Yu. G. Evtushenko, "Theorems on alternatives and their application to numerical methods," Comp. Math. Math. Phys. **43** (3), 338–358 (2003).
10. I. I. Eremin, Vl. D. Mazurov, and N. N. Astaf'ev, *Improper Problems of Linear and Convex Programming* (Nauka, Moscow, 1983) [in Russian].
11. M. Frank and P. Wolfe, "An algorithm for quadratic programming," Naval Res. Logist. Quart. **3**, 95–110 (1956).
12. A. Albert, *Regression and the Moore–Penrose Pseudo-Inverse* (Academic, New York, 1972; Nauka, Moscow, 1977).
13. O. L. Mangasarian, "A Newton method for linear programming," J. Optim. Theory Appl. **121** (1), 1–18 (2004).
14. C. Kanzow, H. Qi, and L. Qi, "On the minimum norm solution of linear programs," J. Optim. Theory Appl. **116**, 333–345 (2003).
15. V. A. Garanzha, A. I. Golikov, Yu. G. Evtushenko, and M. Kh. Nguen, "Parallel implementation of Newton's method for solving large-scale linear programs," Comp. Math. Math. Phys. **49** (8), 1303–1317 (2009).
16. L. D. Popov, "Quadratic approximation of penalty functions for solving large-scale linear programs," Comp. Math. Math. Phys. **47** (2), 200–214 (2007).

Translated by I. Tselishcheva