ISSN 0081-5438, Proceedings of the Steklov Institute of Mathematics, 2017, Vol. 299, Suppl. 1, pp. S246–S256.

published in Trudy Instituta Matematiki i Mekhaniki UrO RAN, 2016, Vol. 22, No. 3.

A Variant of the Dual Simplex Method for a Linear Semidefinite Programming Problem

V. G. Zhadan[1]

Received May 16, 2016

Abstract—A linear semidefinite programming problem in the standard statement is considered, and a variant of the dual simplex method is proposed for its solution. This variant generalizes the corresponding method used for linear programming problems. The transfer from an extreme point of the feasible set to another extreme point is described. The convergence of the method is proved.

Keywords: linear semidefinite programming problem, dual problem, extreme points, dual simplex method.

DOI: 10.1134/S0081543817090267

INTRODUCTION

Linear semidefinite programming problems are an important generalization of linear programming problems [1,2]. In turn, they are a particular case of conic programming problems, where the variables must belong to a closed convex cone (in semidefinite programming, the cone of symmetric positive semidefinite matrices is taken for this purpose). Many convex nonlinear problems of mathematical programming, as well as problems of discrete and combinatorial optimization, are reduced to such statements [3,4], which explains the interest in numerical methods for solving these problems.

The interior point methods, mainly, of affine scaling type [4], are ones of the most developed among such methods at present. In addition, generalizations of the primal simplex method were proposed both for semidefinite programming problems and for conic programming problems [5–7].

One of the difficulties in extending the simplex method to semidefinite programming problems in the standard statement is that the number of equality type constraints, as a rule, is not a triangular number, i.e., the number of elements of a symmetric matrix located on its diagonal and under the diagonal. This requires a special passage from one extreme point of the admissible set to another extreme point. A possible scheme of this passage in the primal simplex method was described in [8]. Here we use a similar technique to generalize the dual simplex method.

The paper consists of three sections. In Section 1 we give the problem statement and optimality conditions. In Sections 2 and 3 we consider the passage from one extreme point of the admissible set of the dual problem to another extreme point in two cases depending on whether the inequality that connects the rank of the dual residual matrix with the number of constraints in the problem turns into an equality. In the end of Section 3, we prove the local convergence of the method.

[1]Dorodnitsyn Computing Center of the Russian Academy of Sciences, Moscow, 119333 Russia
e-mail: zhadan@ccas.ru

1. PROBLEM STATEMENT AND OPTIMALITY CONDITIONS

Let $\mathbb{S}^n$ denote the space of symmetric matrices of order n, and let $\mathbb{S}^n_+$ be the subset of $\mathbb{S}^n$ consisting of positive semidefinite matrices. The set $\mathbb{S}^n_+$ is a cone in $\mathbb{S}^n$. To specify that a matrix $M \in \mathbb{S}^n$ is positive semidefinite, we will also use the inequality $M \succeq 0$. The cone $\mathbb{S}^n_+$ is not polyhedral; its dimension is the triangular number $n_\triangle = n(n+1)/2$.

The scalar (inner) product of two matrices M_1 and M_2, denoted by $M_1 \bullet M_2$, is defined as follows: $M_1 \bullet M_2 = \operatorname{tr}(M_1^T M_2)$; i.e., it is the sum of products of the elements of M_1 and M_2 located at the same positions. If M_1 and M_2 are two positive semidefinite matrices from $\mathbb{S}^n$, then $M_1 \bullet M_2 \geq 0$ and $M_1 \bullet M_2 = 0$ if and only if $M_1 M_2 = M_2 M_1 = 0_{nn}$. The cone $\mathbb{S}^n_+$ is self-dual [4].

Consider the semidefinite programming problem

$$\min\ C \bullet X, \quad A_i \bullet X = b^i, \quad 1 \leq i \leq m, \quad X \succeq 0. \tag{1.1}$$

Here the matrices C, X, and A_i, $1 \leq i \leq m$, belong to the space $\mathbb{S}^n$. We assume that the matrices A_i, $1 \leq i \leq m$, are linearly independent.

The dual problem to (1.1) is the problem

$$\max\ \langle b, u\rangle, \quad V = V(u) = C - \sum_{i=1}^m u^i A_i \succeq 0, \tag{1.2}$$

where $b = (b^1, \ldots, b^m)^T \in \mathbb{R}^m$ and the angular brackets stand for the usual Euclidean scalar product in $\mathbb{R}^m$. We assume that both problems (1.1) and (1.2) have solutions and $b \neq 0_m$.

Let $\mathcal{F}_D$ be the admissible set in the dual problem (1.2); i.e.,

$$\mathcal{F}_D = \left\{[u, V] \in \mathbb{R}^m \times \mathbb{S}^n_+ \colon V = V(u)\right\}.$$

The projection of $\mathcal{F}_D$ to the space $\mathbb{R}^m$ is the set

$$\mathcal{F}_{D,u} = \left\{u \in \mathbb{R}^m \colon [u, V] \in \mathcal{F}_D \text{ for some } V \in \mathbb{S}^n_+\right\}.$$

The optimality conditions for the pair of problems (1.1) and (1.2) consist in the existence of $X \succeq 0$ and $V \succeq 0$ satisfying the system of equalities

$$X \bullet V = 0, \quad A_i \bullet X = b^i, \quad 1 \leq i \leq m, \quad V = C - \sum_{i=1}^m u^i A_i. \tag{1.3}$$

Let us write these equalities in a slightly different form using the operation of vectorization of matrices.

For a square matrix M of order n, denote by $\operatorname{vec} M$ the column vector of length n^2 that is the direct sum of the columns M. If the matrix M is symmetric, then, instead of $\operatorname{vec} M$, it is reasonable to use the column vector $\operatorname{svec} M$ of the smaller dimension $n_\triangle$. This vector contains the lower parts of the columns of M starting from the diagonal element, and each off-diagonal elements is multiplied by $\sqrt{2}$. Then the scalar product $M_1 \bullet M_2$ of two matrices $M_1 \in \mathbb{S}^n$ and $M_2 \in \mathbb{S}^n$ is written as the usual scalar product in the space $\mathbb{R}^{n_\triangle}$; i.e., $M_1 \bullet M_2 = \langle \operatorname{svec} M_1, \operatorname{svec} M_2\rangle$. Thus, equality (1.3) takes the following form in terms of the vector representations of matrices:

$$\langle \operatorname{svec} X, \operatorname{svec} V\rangle = 0, \quad \mathcal{A}_{\rm svec} \operatorname{svec} X = b, \quad \operatorname{svec} V = \operatorname{svec} C - \mathcal{A}_{\rm svec}^T u. \tag{1.4}$$

Here $\mathcal{A}_{\rm svec}$ is the matrix of size $m \times n_\triangle$ with the rows $\operatorname{svec} A_i$, $1 \leq i \leq m$.

In what follows, we consider a numerical method for solving problem (1.2) and, consequently, problem (1.1), based on solving system (1.4) in a special way. This method can be interpreted as a generalization of the dual simplex method for linear programming problems. All the points of the iterative process in the space $\mathbb{R}^m$ belong to the boundary of the set $\mathcal{F}_{D,u}$ and are extreme points of this set.

Take a point $u \in \mathcal{F}_{D,u}$ and consider the corresponding dual residual $V = V(u)$. Let $\operatorname{rank} V = s$ and define the sth triangular number $s_\triangle = s(s+1)/2$. If u is an *extreme* point of the set $\mathcal{F}_{D,u}$, then the rank of the matrix V satisfies the inequality $s_\triangle \le n_\triangle - m$ (see [4]). Since there are gaps between neighboring triangular numbers, it may happen that, for specific n and m from the statement of problem (1.1), this inequality is strict. Let us call an extreme point u *regular* if $s_\triangle = n_\triangle - m$. Otherwise, when $s_\triangle < n_\triangle - m$, an extreme point u will be called *irregular*.

2. ITERATION AT A REGULAR POINT

Given an initial extreme point $u_0 \in \mathcal{F}_{D,u}$, we construct a sequence of extreme points $\{u_k\}$ so that the corresponding values of the objective function in problem (1.2) monotonically increase from iteration to iteration.

Assume that $u \in \mathcal{F}_{D,u}$ is a current extreme point and the dual residual matrix $V = V(u)$ at this point has rank $s < n$. For the matrix V, we have a decomposition $V = HD(\theta)H^T$, where H is an orthogonal matrix and θ is the vector of eigenvalues of V. Since V is rank deficient, we can divide the matrix H and the vector θ into two parts in accordance with the zero and positive eigenvalues of V. We assume for definiteness that this decomposition has the form

$$H = [H_B, H_N], \quad \theta = [\theta_B, \theta_N], \quad \theta_B = 0_r, \quad \theta_N > 0_s, \quad r = n - s. \tag{2.1}$$

In accordance with (2.1), we decompose the space $\mathbb{S}^n$ into linear subspaces $\mathbb{S}^n_B$ and $\mathbb{S}^n_N$. The second subspace $\mathbb{S}^n_N$ consists of matrices $M \in \mathbb{S}^n$ in which only the lower right block of order s may contain nonzero elements. On the contrary, the first subspace $\mathbb{S}^n_B$ consists of matrices $M \in \mathbb{S}^n$ in which the lower right block contains only zeros. These two subspaces are orthogonal to each other, and any matrix $M \in \mathbb{S}^n$ can be represented as $M = M_1 + M_2$, where $M_1 \in \mathbb{S}^n_B$ and $M_2 \in \mathbb{S}^n_N$.

If we pass from the matrix V to the matrix $V^H = H^T V H$, i.e., to the representation of V in the basis given by the columns of the orthogonal matrix H, we come to the relation $V^H = V^H_B + V^H_N$, where

$$V^H_B = H^T V_B H = 0_{nn}, \quad V^H_N = H^T V_N H = \begin{bmatrix} 0_{ss} & 0_{sr} \\ 0_{rs} & D(\theta_N) \end{bmatrix}. \tag{2.2}$$

We will treat the matrices X, C, and A_i, $1 \le i \le m$, in a similar manner; i.e., we pass to the representations $X^H = H^T X H$, $C^H = H^T C H$, and $A^H_i = H^T A_i H$ and decompose each of them into two matrices. For example,

$$X^H = X^H_B + X^H_N, \quad X^H_B = \begin{bmatrix} H^T_B X H_B & H^T_B X H_N \\ H^T_N X H_B & 0_{rr} \end{bmatrix}, \quad X^H_N = \begin{bmatrix} 0_{ss} & 0_{sr} \\ 0_{rs} & H^T_N X H_N \end{bmatrix}.$$

The matrix X^H_B is a bordering matrix if its off-diagonal blocks are nonzero. A point $u \in \mathcal{F}_{D,u}$ is extreme if and only if the matrices $A^H_{i,B}$, $1 \le i \le m$, are linearly independent [4].

Since $X \bullet V = \operatorname{tr} V^H X^H = V^H \bullet X^H$, the first equality in (1.4) can be written in the form $\langle \operatorname{svec} X, \operatorname{svec} V\rangle = \langle \operatorname{svec} X^H, \operatorname{svec} V^H\rangle = 0$. In addition, this equality is equivalent to the following:

$$\langle \operatorname{svec} X^H, \operatorname{svec} V^H\rangle = \langle \operatorname{svec} X^H_B, \operatorname{svec} V^H_B\rangle + \langle \operatorname{svec} X^H_N, \operatorname{svec} V^H_N\rangle = 0. \tag{2.3}$$

Note that the corresponding matrices V_B^H and V_N^H at the point u have the form (2.2). Hence, $\operatorname{svec} V_B^H$ is the zero vector, and the vector $\operatorname{svec} V_N^H$ is $\operatorname{svec} D(\theta)$; i.e., its first $n_\triangle - s_\triangle$ components are also zeros. Therefore, equality (2.3) holds if the matrix X^H is such that the corresponding matrix X_N^H is zero. The number $n_\triangle - s_\triangle$ plays an important role and is hereinafter denoted by l.

Assume that u is not an optimal solution of the dual problem (1.2), and it is desirable to pass to a new extreme point $\bar{u}$ with a larger value of the objective function. Let us first assume for simplicity that u is a regular extreme point. In this case, $m = l$.

Along with the decomposition of the matrix $M \in \mathbb{S}^n$ into components $M_B \in \mathbb{S}_B^n$ and $M_N \in \mathbb{S}_N^n$, we will need a decomposition of the vector $\operatorname{svec} M$ into two subvectors; specifically, $\operatorname{svec} M = [\operatorname{svec}_B M, \operatorname{svec}_N M]^T$, where the dimensions of the components $\operatorname{svec}_B M$ and $\operatorname{svec}_N M$ are l and $s_\triangle$, respectively. In particular, $\operatorname{svec} V^H = \left[\operatorname{svec}_B V^H, \operatorname{svec}_N V^H\right]^T$ and the following relations hold in view of the regularity of the point u:

$$\operatorname{svec}_B V^H = \operatorname{svec}_B V_B^H = 0_m, \quad \operatorname{svec}_N V^H = \operatorname{svec}_N V_N^H = \operatorname{svec} D(\theta_N) \in \mathbb{R}^{s_\triangle}.$$

Let $\mathcal{A}_{\text{svec}}^H$ be the $(m \times n_\triangle)$-matrix whose rows are the vectors $\operatorname{svec} A_i^H$, $1 \leq i \leq m$, and let $\mathcal{A}_{\text{svec}_B}^H$ be its submatrix consisting of the first m columns, i.e., of the rows $\operatorname{svec}_B A_i^H$. The second equality in the optimality conditions (1.4) can be written with the use of the introduced notation in an equivalent vector form as $\mathcal{A}_{\text{svec}}^H \operatorname{svec} X^H = b$. If we now require that $\operatorname{svec}_N X^H = 0_{s_\triangle}$, then this equality is reduced to a system of linear equations with respect to the vector $\operatorname{svec}_B X^H$:

$$\mathcal{A}_{\text{svec}_B}^H \operatorname{svec}_B X^H = b. \tag{2.4}$$

Since u is an extreme point, the matrix of this system is nonsingular. Hence, solving system (2.4), we obtain

$$\operatorname{svec}_B X^H = \left(\mathcal{A}_{\text{svec}_B}^H\right)^{-1} b. \tag{2.5}$$

For the whole matrix $X^H \in \mathbb{S}^n$, we have the vector representation

$$\operatorname{svec} X^H = \operatorname{svec} X_B^H = \begin{bmatrix} \left(\mathcal{A}_{\text{svec}_B}^H\right)^{-1} b \\ 0_{s_\triangle} \end{bmatrix}.$$

Hence, in particular, the lower right block of order s in the matrix X^H consists of zeros. Therefore, if X^H is a bordering matrix, there are negative numbers among its eigenvalues.

If the matrix X^H is positive semidefinite, then the matrix X, which is similar to it, is also positive semidefinite. In this case, the point u and the corresponding weak dual variable $V(u)$ form a solution of the dual problem (1.2), because the optimality conditions (1.3) are fulfilled. The point X will be a solution of the original problem (1.1).

Further, assume that X is not a positive semidefinite matrix. Consider its decomposition $X = QD(\eta)Q^T$, where Q is an orthogonal matrix and η is the vector of eigenvalues of X, which coincide with the eigenvalues of the matrix $X^H = X_B^H$. Vectorizing the matrix X using the known formula

$$\operatorname{vec} M_1 M_2 M_3 = (M_3^T \otimes M_1) \operatorname{vec} M_2, \tag{2.6}$$

where $\otimes$ denotes the Kronecker product of matrices, we obtain $\operatorname{vec} X = (Q \otimes Q) \operatorname{vec} D(\eta)$.

We pass in this formula from $\operatorname{vec} X$ and $\operatorname{vec} D(\eta)$ to $\operatorname{svec} X$ and $\operatorname{svec} D(\eta)$. For this, we will need special *elimination* and *duplication* matrices (see [9]), denoted by $\widetilde{\mathcal{L}}_n$ and $\widetilde{\mathcal{D}}_n$. They are full rank matrices of size $n_\triangle \times n^2$ and $n^2 \times n_\triangle$, respectively. If M is a symmetric matrix of order n, then $\mathrm{svec}\, M = \widetilde{\mathcal{L}}_n \mathrm{vec}\, M$ and $\mathrm{vec}\, M = \widetilde{\mathcal{D}}_n \mathrm{svec}\, M$. Using the introduced matrices, we get

$$\mathrm{svec}\, X = \widetilde{\mathcal{L}}_n (Q \otimes Q) \mathrm{vec}\, D(\eta) = \widetilde{\mathcal{L}}_n (Q \otimes Q) \widetilde{\mathcal{D}}_n \,\mathrm{svec}\, D(\eta).$$

Note that $\widetilde{\mathcal{L}}_n$ and $\widetilde{\mathcal{D}}_n$ are somewhat different from the elimination and duplication matrices $\mathcal{L}_n$ and $\mathcal{D}_n$ from [9]; specifically, $\widetilde{\mathcal{L}}_n = \mathrm{Diag}\,(\mathrm{svec}\, E_n)\, \mathcal{L}_n$ and $\widetilde{\mathcal{D}}_n = \mathcal{D}_n \mathrm{Diag}^{-1}(\mathrm{svec}\, E_n)$. Here E_n is the square matrix of order n all of whose elements are equal to one, and $\mathrm{Diag}\,(a)$ is the diagonal matrix with vector a on its diagonal.

Let q_i, $1 \le i \le n$, be the columns of the orthogonal matrix Q (i.e., the eigenvectors of the matrix X). Then X can also be written in the following matrix and vector forms:

$$X = \sum_{i=1}^{n} \eta^i q_i q_i^T, \quad \mathrm{vec}\, X = \sum_{i=1}^{n} \eta^i \left(q_i \otimes q_i\right). \tag{2.7}$$

Assume that η^k is a negative eigenvalue of X and q_k is the corresponding eigenvector. We pass to a new point $\bar{u}$, setting

$$\bar{u} = u - \alpha \Delta u, \tag{2.8}$$

where $\alpha > 0$. We require that the vector $\Delta u \in \mathbb{R}^m$ satisfy the system of linear equations

$$\left(\mathcal{A}^H_{\mathrm{svec}_B}\right)^T \Delta u = \mathrm{svec}_B\, Q_k^H. \tag{2.9}$$

Here and below, $Q_k^H = H^T Q_k H$ and $Q_k = q_k q_k^T$. The symmetric matrix Q_k is positive semidefinite and has unit rank.

Since the matrix $\mathcal{A}^H_{\mathrm{svec}_B}$ is nonsingular, we can solve this system and obtain

$$\Delta u = \left(\mathcal{A}^H_{\mathrm{svec}_B}\right)^{-T} \mathrm{svec}_B\, Q_k^H, \tag{2.10}$$

where we use the conventional notation $M^{-T} = (M^T)^{-1}$.

Assertion 1. *The vector q_k does not belong to the subspace $\mathcal{R}(H_N)$ generated by the columns of the matrix H_N.*

Proof. Indeed, if assume that $q_k = H_N z$ for some nonzero vector $z \in \mathbb{R}^s$, then we would have $X q_k = X H_N z = \eta^k H_N z$. Multiplying this equality on the left by the matrix H_N^T, we get $H_N^T X H_N z = \eta^k z$, which is impossible because $H_N^T X H_N$ is the zero matrix.

The assertion is proved.

Remark. Since the nonzero vector q_k does not belong to the subspace $\mathcal{R}(H_N)$, it can be represented in the form

$$q_k = H_B q_k^{H,B} + H_N q_k^{H,N}, \tag{2.11}$$

where $q_k^{H,B} = H_B^T q_k \neq 0_r$. If $q_k^{H,N} = H_N^T q_k = 0_s$, then $q_k = H_B q_k^{H,B}$. Hence, the matrix Q_k belongs to the face $\mathcal{G}^*_{\min}(V; \mathbb{S}^n_+)$, which is dual to the minimum face $\mathcal{G}_{\min}(V; \mathbb{S}^n_+)$ of the cone $\mathbb{S}^n_+$ containing the point $V = V(u)$.

Let us consider the change of the value of the objective function in the dual problem (1.2) when we pass to the new point $\bar{u}$.

Assertion 2. *The increment of the value of the objective function in the dual problem satisfies the formula*

$$\langle b, \bar{u}\rangle = \langle b, u\rangle - \alpha \eta^k > \langle b, u\rangle. \tag{2.12}$$

Proof. According to (2.8), we have $\langle b, \bar{u}\rangle = \langle b, u\rangle - \alpha\langle b, \Delta u\rangle$. However,

$$\langle b, \Delta u\rangle = \langle b, \left(\mathcal{A}^H_{\mathrm{svec}_B}\right)^{-T} \mathrm{svec}_B\, Q^H_k\rangle = \langle \left(\mathcal{A}^H_{\mathrm{svec}_B}\right)^{-1} b, \mathrm{svec}_B\, Q^H_k\rangle$$
$$= \langle \mathrm{svec}_B X^H_B,\ \mathrm{svec}_B\, Q^H_k\rangle = \langle \mathrm{svec}\, X^H_B, \mathrm{svec}\, Q^H_k\rangle = \langle \mathrm{vec}\, X^H_B, \mathrm{vec}\, Q^H_k\rangle.$$

Here we used the fact that the last $s_\triangle$ components of the vector $\mathrm{svec}\, X^H_B$ are zeros.

Since $X^H_B = X^H$, it follows from (2.7) that $X^H = \sum\limits_{i=1}^m \eta^i Q^H_i$, where $Q^H_i = H^T q_i q_i^T H$. Then, using (2.6), we obtain $\mathrm{vec}\, X^H_B = (H^T \otimes H^T)\sum\limits_{i=1}^m \eta^i \mathrm{vec}\,(q_i q_i^T) = (H \otimes H)^T \sum\limits_{i=1}^m \eta^i (q_i \otimes q_i)$. Since $\mathrm{vec}\, Q^H_k = (H \otimes H)^T(q_k \otimes q_k)$, we come to the relation $\langle \mathrm{vec}\, X^H_B, \mathrm{vec}\, Q^H_k\rangle = \eta^k$. Thus, (2.12) is true.
The assertion is proved.

Consider the matrix $\Delta V^H = \sum_{i=1}^m (\Delta u)^i A^H_i$ and decompose it into two matrices: $\Delta V^H = \Delta V^H_B + \Delta V^H_N$. For the first matrix ΔV^H_B, we obtain

$$\mathrm{svec}_B\, \Delta V^H_B = \left(\mathcal{A}^H_{\mathrm{svec}_B}\right)^T \Delta u = \mathrm{svec}_B\, Q^H_k. \tag{2.13}$$

Since the lower right block of the matrices $A_{i,B}$, $1 \le i \le m$, is zero, it follows from (2.13) that $\Delta V^H_B = (Q^H_k)_B$. Note that the lower right block of the matrix $(Q^H_k)_B$ is also zero. Let us calculate ΔV^H_N. Vectorizing, we get

$$\mathrm{svec}_N\, \Delta V^H_N = \left(\mathcal{A}^H_{\mathrm{svec}_N}\right)^T \Delta u = \left(\mathcal{A}^H_{\mathrm{svec}_N}\right)^T \left(\mathcal{A}^H_{\mathrm{svec}_B}\right)^{-T} \mathrm{svec}_B\, Q^H_k. \tag{2.14}$$

Formula (2.8) corresponds to the formula for calculating the weak dual variable

$$\bar{V}^H(\alpha) = V^H(\bar{u}) = V^H(u) + \alpha\Delta V^H. \tag{2.15}$$

Assertion 3. *There exists $\bar{\alpha} > 0$ such that $\bar{V}^H(\alpha) \succeq 0$ for any $0 < \alpha \le \bar{\alpha}$.*

Proof. Let us apply the representation (2.11) of the vector q_k. Then the vector $q^H_k = H^T q_k$ is decomposed into two subvectors: $q^H_k = \left[q^{H,B}_k, q^{H,N}_k\right]^T$, where $q^{H,B}_k \in \mathbb{R}^r$, $q^{H,N}_k \in \mathbb{R}^s$, and $q^{H,B}_k \ne 0_r$. Let us represent the increment matrix ΔV^H in the block form:

$$\Delta V^H = \begin{bmatrix} \Omega_{BB} & \Omega_{BN} \\ \Omega_{NB} & \Omega_{NN} \end{bmatrix},$$

where the diagonal blocks Ω_{BB} and Ω_{NN} have orders r and s, respectively. Since $\Delta V^H_B = (Q^H_k)_B$, we have $\Omega_{BB} = q^{H,B}_k (q^{H,B}_k)^T$ and $\Omega_{BN} = (\Omega_{NB})^T = q^{H,B}_k (q^{H,N}_k)^T$. The matrix Ω_{NN} has vector representation (2.14). Therefore, according to (2.2) and (2.15),

$$\bar{V}^H(\alpha) = \begin{bmatrix} \alpha q^{H,B}_k \left(q^{H,B}_k\right)^T & \alpha q^{H,B}_k \left(q^{H,N}_k\right)^T \\ \alpha q^{H,N}_k \left(q^{H,B}_k\right)^T & D(\theta_N) + \alpha\Omega_{NN} \end{bmatrix}. \tag{2.16}$$

Adding and subtracting the matrix $\alpha q^{H,N}_k (q^{H,N}_k)^T$ in the lower right block and setting $\widetilde{\Omega}_{NN} = \Omega_{NN} - q^{H,N}_k (q^{H,N}_k)^T$, we obtain one more representation of $\bar{V}^H(\alpha)$:

$$\bar{V}^H(\alpha) = \alpha Q^H_k + \begin{bmatrix} 0_{rr} & 0_{rs} \\ 0_{sr} & D(\theta_N) + \alpha\widetilde{\Omega}_{NN} \end{bmatrix}. \tag{2.17}$$

The matrix Q_k^H has unit rank and is positive semidefinite. In addition, since $\theta_N > 0_s$, we conclude that the lower right submatrix $\widetilde{Y}_{NN}(\alpha) = D(\theta_N) + \alpha\widetilde{\Omega}_{NN}$ in the second matrix in (2.17) will be positive definite for sufficiently small α. Consequently, we can specify $\bar{\alpha} > 0$ such that $\bar{V}^H(\alpha) \succeq 0$ for all $0 < \alpha \leq \bar{\alpha}$.

The assertion is proved.

Remark. If the matrix Ω_{NN} is such that $\Omega_{NN} \succeq q_k^{H,N}(q_k^{H,N})^T$, then the matrix $\widetilde{\Omega}_{NN}$ and, hence, the matrix $\widetilde{Y}_{NN}(\alpha)$ are positive semidefinite for all $\alpha > 0$. In this case, the dual problem (1.2) has no solution.

Further, assume that the matrix $\widetilde{\Omega}_{NN}$ has negative eigenvalues. Therefore, the inequality $\Omega_{NN} \succeq q_k^{H,N}(q_k^{H,N})^T$ does not hold. In this case, an upper bound for the greatest $\bar{\alpha}$ that preserves the positive semidefiniteness of $\bar{V}^H(\alpha)$ is found as the smallest α for which the matrix $\widetilde{Y}_{NN}(\alpha)$ has zero eigenvalue.

The bound for the greatest possible $\bar{\alpha}$ can be refined. Indeed, since $\theta_N > 0_s$, the lower right matrix $Y_{NN}(\alpha) = D(\theta_N) + \alpha\Omega_{NN}$ remains positive definite for sufficiently small α. Hence, as seen from (2.16), the whole matrix $\bar{V}^H(\alpha)$ is positive semidefinite if the Schur complement of the matrix $Y_{NN}(\alpha)$, i.e., the matrix

$$\bar{Y}_{NN}(\alpha) = \alpha \left\{ q_k^{H,B} \left(q_k^{H,B}\right)^T - \alpha q_k^{H,B} \left(q_k^{H,N}\right)^T [D(\theta_N) + \alpha\Omega_{NN}]^{-1} q_k^{H,N} \left(q_k^{H,B}\right)^T \right\},$$

is also positive semidefinite.

It is clear that this condition holds if $q_k^{H,N}$ is the zero vector. The matrix $\bar{V}^H$ in this case becomes block-diagonal. Further, we assume that the vector $q_k^{H,N}$ is nonzero. Defining $p(\alpha) = \left(q_k^{H,N}\right)^T [D(\theta_N) + \alpha\Omega_{NN}]^{-1} q_k^{H,N}$, we find that

$$\bar{Y}_{NN}(\alpha) = \alpha \left[1 - \alpha p(\alpha)\right] q_k^{H,B} \left(q_k^{H,B}\right)^T.$$

The matrix $\bar{Y}_{NN}$ also remains positive semidefinite for sufficiently small α. Thus, $\bar{\alpha}$ is found as follows. First, one should find the smallest α (denote it by $\bar{\alpha}_1$) for which the matrix $Y_{NN}(\alpha)$ has zero eigenvalue. Second, one should make sure that the inequality $\alpha \leq p(\alpha)^{-1}$ is fulfilled. If this inequality is violated for the first time for some $\bar{\alpha}_2 < \bar{\alpha}_1$, then the upper bound for $\bar{\alpha}$ is $\bar{\alpha} = \bar{\alpha}_2$. Otherwise, $\bar{\alpha} = \bar{\alpha}_1$.

3. ITERATION AT AN IRREGULAR POINT

Assume now that the point $u \in \mathcal{F}_{D,u}$ is irregular, i.e., that the rank s of the matrix $V(u)$ satisfies the strict inequality $s_\triangle < n_\triangle - m$. In this case, system (2.4) becomes underdetermined. Then, we take for $\mathrm{svec}_B\, X^H$ a solution of (2.4) with minimum norm:

$$\mathrm{svec}_B\, X^H = \left(\mathcal{A}^H_{\mathrm{svec}_B}\right)^T \left[\mathcal{A}^H_{\mathrm{svec}_B} \left(\mathcal{A}^H_{\mathrm{svec}_B}\right)^T\right]^{-1} b. \quad (3.1)$$

It belongs to the row space of the matrix $\mathcal{A}^H_{\mathrm{svec}_B}$. The general solution of system (2.4) has the form $\mathrm{svec}_B\, X^H = \left(\mathcal{A}^H_{\mathrm{svec}_B}\right)^T \left[\mathcal{A}^H_{\mathrm{svec}_B} \left(\mathcal{A}^H_{\mathrm{svec}_B}\right)^T\right]^{-1} b + g$, where g is an arbitrary vector belonging to the zero space of the matrix $\mathcal{A}^H_{\mathrm{svec}_B}$. If $m + p = l$, then the dimension of the zero space of $\mathcal{A}^H_{\mathrm{svec}_B}$

is p. Below, we assume that $p < s$. The irregular points u at which the rank s of the matrix $V(u)$ satisfies the inequality $s_\triangle + s > n_\triangle - m$ will be called *quasi-regular*.

Consider the matrix $X = HX^HH^T$. It similar to the matrix $X^H = X^H_B$, whose vector $\mathrm{svec}_B\, X^H$ is defined by (3.1). Let η^k be a negative eigenvalue of X, and let q_k be the corresponding eigenvector. Just as at a regular point u, the vector q_k does not belong to the linear subspace $\mathcal{R}(H_N)$.

System (2.9) for the direction Δu becomes overdetermined in this case. That is why we consider another, more general, method of finding Δu. We pass from Δu to the direction ΔV in the V-space, and search for the latter in the form

$$\Delta V = [q_k\ H_N]\begin{bmatrix} 1 & w^T \\ w & \Delta Z \end{bmatrix}\begin{bmatrix} q_k^T \\ H_N^T \end{bmatrix} = Q_k + q_k w^T H_N^T + H_N w q_k^T + H_N \Delta Z H_N^T. \tag{3.2}$$

Here, $\Delta Z \in \mathbb{S}^s$ and $w \in \mathbb{R}^s$.

We require the vector w to be chosen as follows: $w = Wy$, where all columns $w_j \in \mathbb{R}^s$, $1 \le j \le p$, of the matrix W are linearly independent and $y \in \mathbb{R}^p$. In addition, we require the vectors $h_{w_j} = H_N w_j$, $1 \le j \le p$, to be orthogonal to the vector q_k:

$$\langle q_k, h_{w_j}\rangle = \langle H_N^T q_k, w_j\rangle = 0, \quad 1 \le j \le p. \tag{3.3}$$

All vectors h_{w_j}, $1 \le j \le p$, belong to the subspace $\mathcal{R}(H_N)$.

Along with (3.2), there is a relation between ΔV and Δu; specifically, $\Delta V = \sum_{i=1}^m \Delta u^i A_i$. Equating this representation for ΔV to (3.2), we come to the equality

$$\sum_{i=1}^m \Delta u^i A_i = Q_k + q_k y^T W^T H_N^T + H_N W y q_k^T + H_N \Delta Z H_N^T.$$

Let us vectorize it. Preliminarily, we define for brevity

$$W_N = H_N W, \quad U_{W_N,q_k} = W_N \otimes q_k + q_k \otimes W_N, \quad \mathcal{H}_N = H_N \otimes H_N.$$

Then, in view of the equality $\mathrm{vec}\, y^T = \mathrm{vec}\, y = y$, we obtain

$${\mathcal{A}_{\mathrm{vec}}}^T \Delta u - U_{W_N,q_k} y - \mathcal{H}_N \mathrm{vec}\, \Delta Z = \mathrm{vec}\, Q_k. \tag{3.4}$$

Rewrite this equality in the basis given by the orthogonal matrix H. For this, we multiply the equality on the left by the matrix $(H \otimes H)^T = H^T \otimes H^T$. Since

$$(H^T \otimes H^T)(H_N \otimes H_N) = \begin{bmatrix} 0_{rs} \\ I_s \end{bmatrix} \otimes \begin{bmatrix} 0_{rs} \\ I_s \end{bmatrix} = \begin{bmatrix} 0_{(rn)s^2} \\ \mathrm{Diag}\left(\begin{bmatrix} 0_{rs} \\ I_s \end{bmatrix}, \ldots, \begin{bmatrix} 0_{rs} \\ I_s \end{bmatrix}\right) \end{bmatrix} \tag{3.5}$$

and $U^H_{W_N,q_k} = (H^T \otimes H^T)U_{W_N,q_k} = W_N^H \otimes q_k^H + q_k^H \otimes W_N^H$, where $q_k^H = H^T q_k$, $W_N^H = H^T W_N$, we have

$$({\mathcal{A}_{\mathrm{vec}}}^H)^T \Delta u - \left\{U^H_{W_N,q_k} y + \Gamma^H \mathrm{vec}\, \Delta Z\right\} = \mathrm{vec}\, Q_k^H, \tag{3.6}$$

where Γ^H is the matrix from the right-hand side of (3.5).

The columns of the matrices $U^H_{W_N,q_k}$ and Γ^H correspond to symmetric matrices (in the case of the matrix Γ^H, the only unit elements in a column are at the positions of diagonal elements). Therefore, system (3.6), can be written as follows:

$$(\mathcal{A}^H_{\mathrm{svec}})^T \Delta u - \widetilde{U}^H_{W_N,q_k} y - \widetilde{\Gamma}^H \mathrm{svec}\, \Delta Z = \mathrm{svec}\, Q_k^H, \tag{3.7}$$

where $\widetilde{U}^H_{W_N,q_k} = \widetilde{\mathcal{L}}_n U^H_{W_N,q_k}$ and $\widetilde{\Gamma}^H = \widetilde{\mathcal{L}}_n \Gamma^H \widetilde{\mathcal{D}}_s$. Note that the matrix $\widetilde{\Gamma}^H$ of size $n_\triangle \times s_\triangle$ is such that its upper submatrix of size $l \times s_\triangle$ is zero. System (3.7) is a system of $n_\triangle$ linear equations with respect to $n_\triangle$ variables: Δu, y, and $\operatorname{svec} \Delta Z$. If its matrix

$$\mathcal{M} = \left[(\mathcal{A}^H_{\text{svec}})^T \vdots \widetilde{U}^H_{W_N,q_k} \vdots \widetilde{\Gamma}^H\right]$$

is nonsingular, then the system has a unique solution.

Assume in what follows that the matrix $\mathcal{M}$ is nonsingular. Since the matrices A_i, $1 \le i \le m$, are linearly independent, the matrix $\mathcal{A}^H_{\text{svec}}$ has full rank equal to m. Moreover, since u is an extreme point of the set $\mathcal{F}_{D,u}$, the matrix $\mathcal{A}^H_{\text{svec}_B}$ of size $m \times l$ also has rank m, i.e., full row rank.

Let $\mathcal{K}$ be an arbitrary matrix of size $p \times l$ whose rows are linearly independent vectors from the zero space of the matrix $\mathcal{A}^H_{\text{svec}_B}$. Using $\mathcal{A}^H_{\text{svec}_B}$ and $\mathcal{K}$, we compose the square matrix

$$\mathcal{Q} = \begin{bmatrix} \bar{\mathcal{A}}^H_{\text{svec}_B} & 0_{ls_\triangle} \\ 0_{s_\triangle l} & I_{s_\triangle} \end{bmatrix}, \quad \bar{\mathcal{A}}^H_{\text{svec}_B} = \begin{bmatrix} \mathcal{A}^H_{\text{svec}_B} \\ \mathcal{K} \end{bmatrix}.$$

Note that the matrix $\bar{\mathcal{A}}^H_{\text{svec}_B}$ is nonsingular by its definition. If we multiply system (3.7) on the left by the nonsingular matrix $\mathcal{Q}$, its solution does not change.

Denote by $\widetilde{U}^{H,B}_{W_N,q_k}$ the upper submatrix of $\widetilde{U}^H_{W_N,q_k}$ of size $l \times p$. Multiplying (3.7) by the first row of $\mathcal{Q}$, we get

$$\mathcal{W}\Delta u = \mathcal{A}^H_{\text{svec}_B}\left[\widetilde{U}^{H,B}_{W_N,q_k} y + \operatorname{svec}_B Q^H_k\right], \quad \mathcal{K}\left[\widetilde{U}^{H,B}_{W_N,q_k} y + \operatorname{svec}_B Q^H_k\right] = 0_p, \tag{3.8}$$

where $\mathcal{W} = \mathcal{A}^H_{\text{svec}_B}(\mathcal{A}^H_{\text{svec}_B})^T$.

Since the square matrix $\mathcal{W}$ of order m is nonsingular, we have

$$\Delta u = \mathcal{W}^{-1}\mathcal{A}^H_{\text{svec}_B}\left[\widetilde{U}^{H,B}_{W_N,q_k} y + \operatorname{svec}_B Q^H_k\right]. \tag{3.9}$$

If the square matrix $\mathcal{K}\widetilde{U}^{H,B}_{W_N,q_k}$ of order p is nonsingular, then, solving the second system in (3.8), we obtain $y = -\left(\mathcal{K}\widetilde{U}^{H,B}_{W_N,q_k}\right)^{-1}\mathcal{K}\operatorname{svec}_B Q^H_k$. Substituting the found y into the expression (3.9) for Δu, we get

$$\Delta u = \mathcal{W}^{-1}\mathcal{A}^H_{\text{svec}_B}\left[I_l - \mathcal{P}\right]\operatorname{svec}_B Q^H_k, \quad \mathcal{P} = \widetilde{U}^{H,B}_{W_N,q_k}\left(\mathcal{K}\widetilde{U}^{H,B}_{W_N,q_k}\right)^{-1}\mathcal{K}. \tag{3.10}$$

Let us now calculate the change of the value of the objective function in the dual problem (1.2) along the direction Δu.

Assertion 4. *Suppose that the columns w_j of the matrix W, $1 \le j \le p$, satisfy equality* (3.3). *Then the dual objective function takes at the point $\bar{u}$ value* (2.12).

Proof. Substituting Δu from (3.10), we get

$$\begin{aligned} \langle b, \Delta u\rangle &= \langle b, \mathcal{W}^{-1}\mathcal{A}^H_{\text{svec}_B}\left[I_l - \mathcal{P}\right]\operatorname{svec}_B Q^H_k\rangle = \langle(\mathcal{A}^H_{\text{svec}_B})^T\mathcal{W}^{-T}b, \operatorname{svec}_B Q^H_k - \mathcal{P}\operatorname{svec}_B Q^H_k\rangle \\ &= \langle\operatorname{svec}_B X^H, \operatorname{svec}_B Q^H_k\rangle - \langle\operatorname{svec}_B X^H, \mathcal{P}\operatorname{svec}_B Q^H_k\rangle \\ &= \langle\operatorname{vec} X, \operatorname{vec} Q_k\rangle - \langle\operatorname{svec}_B X^H, \mathcal{P}\operatorname{svec}_B Q^H_k\rangle = \eta^k - \langle\operatorname{svec}_B X^H, \mathcal{P}\operatorname{svec}_B Q^H_k\rangle. \end{aligned} \tag{3.11}$$

Consider the matrix $\widetilde{U}^{H,B}_{W_N,q_k}$ in more detail, finding preliminarily the form of the matrix $U^H_{W_N,q_k}$. Since $W^H_N = [0_{rs} \vdots W]^T$, the ith column of $W^H_N \otimes q^H_k$ is $\operatorname{vec}\left[0_{nr} \vdots w_i \otimes q^H_k\right]$, $1 \le i \le p$. Therefore, the upper $(rn) \times p$-submatrix in the matrix $W_N^H \otimes q_k^H$ is zero. Hence, the upper $l \times p$-submatrix in the matrix $\widetilde{\mathcal{L}}_n(W_N^H \otimes q_k^H)$ is also zero. Thus, $\widetilde{U}^{H,B}_{W_N,q_k}$ coincides with the upper $(l \times p)$-submatrix of $\widetilde{\mathcal{L}}_n(q_k^H \otimes W_N^H)$.

Let us calculate the p-dimensional vector $z = \left(\widetilde{U}^{H,B}_{W_N,q_k}\right)^T \operatorname{svec}_B X^H$. Since the last $s_\triangle$ elements of the vector svec X^H are zeros, the jth element of z is defined as follows:

$$\begin{aligned} z^j &= \langle \widetilde{\mathcal{L}}_n \left(q_k^H \otimes (H^T H_N w_j)\right), \operatorname{svec} X^H \rangle = \langle q_k^H \otimes (H^T H_N w_j), \operatorname{vec} X^H \rangle \\ &= \langle (H \otimes H) \left((H^T q_k) \otimes (H^T H_N w_j)\right), \operatorname{vec} X \rangle = \langle q_k \otimes (H_N w_j), \operatorname{vec} X \rangle \\ &= (q_k \otimes (H_N w_j))^T \operatorname{vec} X = \left(q_k^T \otimes (H_N w_j)^T\right) \operatorname{vec} X. \end{aligned}$$

Hence, in view of condition (3.3) and the equality $\operatorname{vec} X = \sum_{i=1}^{n} \eta^i (q_i \otimes q_i)$, we obtain

$$z^j = \sum_{i=1}^{n} \eta^i \left(q_k^T \otimes (w_j^T H_N^T)\right) (q_i \otimes q_i) = \eta^k w_j^T H_N^T q_k = 0.$$

Therefore, $z = 0_p$, which leads to the equality

$$\langle \operatorname{svec}_B X_B^H, \mathcal{P} \operatorname{svec}_B Q_k^H \rangle = \left\langle \left(\widetilde{U}^{H,B}_{W_N,q_k}\right)^T \operatorname{svec}_B X_B^H, \left(\mathcal{K} \widetilde{U}^{H,B}_{W_N,q_k}\right)^{-1} \mathcal{K} \operatorname{svec}_B Q_k^H \right\rangle = 0.$$

It remains to use equality (3.11).

The assertion is proved.

In conclusion, we consider the issue of the convergence of the method. We say that problem (1.2) is *quasi-regular* if all extreme points from $\mathcal{F}_{D,u}$ are regular or quasi-regular. We assume additionally that the negative eigenvalue η^k taken at each step is the value with the greatest absolute value.

Theorem. *Suppose that problem* (1.2) *is quasi-regular and the initial point* $u_0 \in \mathcal{F}_{D,u}$ *is such that the set* $\mathcal{F}_{D,u}(u_0) = \{u \in \mathcal{F}_{D,u} \colon \langle b, u \rangle \geq \langle b, u_0 \rangle\}$ *is bounded. Then the dual simplex method generates a sequence of points* $\{u_k\} \subset \mathcal{F}_{D,u}(u_0)$. *If this sequence is finite, then its last point is a solution of* (1.2). *Otherwise, any of its limit points is a solution of* (1.2).

Proof. We restrict ourselves to considering the case when the sequence $\{u_k\}$ is infinite. Since it is bounded, there are limit points. Let $u_{k_s} \to \bar{u}$. The point $\bar{u}$ is extreme.

The sequence $\{u_k\}$ corresponds to the sequence of matrices $\{V_k\}$, where $V_k = V(u_k)$. The rank of such matrices is bounded at extreme points. Therefore, the corresponding matrices H_N are also bounded in the Frobenius norm, i.e., belong to a compact set. Hence, $\{V_{k_s}\}$ contains a subsequence for which the matrices H_N converge to some matrix $\bar{H}_N$. Without loss of generality, we assume that the sequence $\{V_{k_s}\}$ itself has this property and the rank of all matrices H_N is the same. Denote by $\bar{H}$ the orthogonal matrix whose second component is $\bar{H}_N$.

Consider the matrix $\mathcal{A}^{\bar{H}}_{\operatorname{svec}_B}$, which enters system (2.4) for finding the vector $\operatorname{svec}_B X^{\bar{H}}$ at the point $\bar{u}$. Since $\bar{u}$ is an extreme point, this matrix has full rank coinciding with the row rank. Hence, the solutions of system (2.4), specifically, the vectors $\operatorname{svec}_B X^{H_{k_s}}$ defined either by (2.5) or by (3.1), converge to $\operatorname{svec}_B \bar{X}^{\bar{H}}$.

The matrix $\bar{X}$ must be positive semidefinite; otherwise, $\bar{X}$ would have a negative eigenvalue. However, eigenvalues of matrices are Lipschitz continuous. Therefore, matrices X_{k_s} sufficiently close to $\bar{X}$ also have negative eigenvalues. Hence, these iterations correspond to the passage from the points u_{k_s} to the next points u_{k_s+1} with step α_{k_s} and with the value of the objective function increasing by $-\alpha_{k_s}\bar{\eta}_{k_s}$, where $\bar{\eta}_{k_s}$ is the negative component of η_{k_s} with maximum absolute value. However, since the vectors Δu_k are bounded, the steps α_{k_s} cannot tend to zero. Therefore, we will have $\langle b, u_{k_s+1}\rangle > \langle b, \bar{u}\rangle$ at a certain k_sth iteration. In view of the monotone increase of the values of the objective function along the trajectory, this contradicts the convergence of $\{u_{k_s}\}$ to $\bar{u}$.

The theorem is proved.

ACKNOWLEDGMENTS

This work was supported by the Russian Foundation for Basic Research (project no. 15-01-08259), by the Program for State Support of Leading Scientific Schools of the Russian Federation (project no. NSh-8860.2016.1), and by Program I.33 P of the Russian Academy of Sciences.

REFERENCES

1. I. I. Eremin, *Theory of Linear Optimization* (Izd. Ekaterinburg, Yekaterinburg, 1999) [in Russian].
2. F. P. Vasil'ev and A. Yu. Ivanitskii, *Linear Programming* (Faktorial, Moscow, 2008) [in Russian].
3. L. Vandenberghe and S. Boyd, "Semidefinite programming," SIAM Rev. **38** (1), 49–95 (1996).
4. *Handbook of Semidefinite Programming*, Ed. by H. Wolkowicz, R. Saigal, and L. Vandenberghe (Kluwer Acad., Dordrecht, 2000).
5. J. B. Lasserre, "Linear programming with positive semi-definite matrices," Math. Probl. Engineering **2**, 499–522 (1996).
6. G. Pataki, "Cone-LP's and semidefinite programs: Geometry and a simplex-type method," in *Proceedings of the 5th International Conference on Integer Programming and Combinatorial Optimization, Vancouver, Canada, 1996*, pp. 162–174.
7. A. I. Kosolap, "The simplex method for semidefinite programming problems," Vestn. Donetsk. Nats. Univ., Ser. Estestv. Nauki, No. 2, pp. 365–369 (2009).
8. V. G. Zhadan, "On a variant of the simplex method for a linear semidefinite programming problem," Trudy Inst. Mat. Mekh. UrO RAN **21** (3), 117–127 (2015).
9. J. R. Magnus and H. Neudecker, *Matrix Differential Calculus with Applications in Statistics and Econometrics* (Wiley, Chichester, 1999; Fizmatlit, Moscow, 2002).

Translated by E. Vasil'eva