Primal and dual affine scaling Newton's methods for linear semidefinite programming problems

V. G. Zhadan*

*Dorodnicyn Computing Centre of RAS, zhadan@ccas.ru

We consider the standard form semidefinite programming (SDP) problem:

$$\begin{array}{rcll} \min \ C \bullet X, & & & \\ A_i \bullet X & = & b^i, & i = 1, \ldots, m, \\ X & \succeq & 0, & \end{array} \qquad (1)$$

and its dual

$$\begin{array}{rcl} \max \ b^T u, & & \\ \sum_{i=1}^m u^i A_i + V & = & C, \\ V & \succeq & 0, \end{array} \qquad (2)$$

where C, X and A_i, $1 \le i \le m$ belong to the space $\mathcal{S}^n$ of $n \times n$ real symmetric matrices, the operator $\bullet$ denotes the standard inner product in $\mathcal{S}^n$, i.e., $C \bullet X := \mathrm{tr}(CX) = \sum_{i,j} C_{ij} X_{ij}$, and the inequality $X \succeq 0$ means that X is positive semidefinite. We write also $X \succ 0$ to indicate that X is positive definite.

In what follows we assume that the matrices A_i, $1 \le i \le m$, are linear independent. We suppose also that the Slater constraint qualification is fulfilled for both problems (1) and (2), i.e. there are feasible matrices X and V such that $X \succ 0$, $V \succ 0$. In this case the strong duality holds and both problems (1), (2) have nonempty compact sets of solutions [1].

If X_* and V_* are optimal solutions of problems (1) and (2), respectively, then $X_* \bullet V_* = 0$, and the matrices X_* and V_* must commute. Hence, there exists an orthogonal matrix Q such that

$$X_* = Q \mathrm{Diag}(\eta_*) Q^T, \quad V_* = Q \mathrm{Diag}(\theta_*) Q^T,$$

where $\eta_* = [\eta_*^1, \ldots, \eta_*^n]$ and $\theta_* = [\theta_*^1, \ldots, \theta_*^n]$ are the eigenvalues of X_* and V_* respectively. The eigenvalues η_*^i and θ_*^i satisfy the complementarity conditions $\eta_*^i \theta_*^i = 0$, $1 \le i \le n$. The *strict complementarity condition* means that, for each $1 \le i \le n$, one of the values η_*^i or θ_*^i is strictly positive.

Denote by $X * V$ the symmetrized product of square matrices X and V defined by the formula $X * V = (XV + V^T X^T)/2$. The well-known optimality conditions for (1) and (2), having the form

$$\begin{array}{rcl} X \bullet V & = & 0, \\ A_i \bullet X & = & b^i, \quad 1 \le i \le m, \\ V & = & C - \sum_{i=1}^m u^i A_i, \\ X \succeq 0, & & V \succeq 0, \end{array}$$

can be written in the following way

$$\begin{array}{rcl} X * V & = & 0_{nn}, \\ A_i \bullet X & = & b^i, \quad 1 \le i \le m, \\ V & = & C - \sum_{i=1}^m u^i A_i, \\ X \succeq 0, & & V \succeq 0. \end{array} \qquad (3)$$

Let the symbol $\mathrm{vec} X$ denote the direct sum of the columns of $X \in \mathcal{S}^n$, 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 $\mathrm{vech} X$ of dimension $k_\triangle(n) = n(n+1)/2$. 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 $\mathrm{vecs} X$ is defined similarly. It differs from the preceding operation $\mathrm{vech} X$ only in that the off-diagonal entries of X are multiplied by two before placing into $\mathrm{vecs} X$.

Let also $\mathcal{L}_n$ and $\mathcal{D}_n$ are the *elimination* and *duplicated* matrices respectively [2]. The matrix $\mathcal{L}_n$

for an arbitrary square matrix X effects the transformation $\mathcal{L}_n \mathrm{vec} X = \mathrm{vech} X$. By contrast, the matrix $\mathcal{D}_n$ acts on an arbitrary symmetric matrix X so that $\mathcal{D}_n \mathrm{vech} X = \mathrm{vec} X$.

Using the vectorization operators *vech* and *vecs*, the optimality conditions (3) can be rewritten as

$$\begin{array}{rcl} \mathcal{L}_n X^{\otimes} \mathcal{D}_n \mathrm{vech} V & = & 0_{k_\triangle(n)}, \\ \mathcal{A}_{vecs} \mathrm{vech} X & = & b, \\ \mathrm{vech} V & = & \mathrm{vech} C - \mathcal{A}_{vech}^T u, \\ X \succeq 0, & & V \succeq 0, \end{array} \quad (4)$$

where $X^{\otimes} = [X \otimes I_n + I_n \otimes X]/2$ is the Kronecker sum of a matrix X, and I_n is the identity matrix of order n. By $\mathcal{A}_{vech}$ and $\mathcal{A}_{vecs}$ we denote the $m \times n^2$ matrices with $\mathrm{vech} A_i$ and $\mathrm{vecs} A_i$ respectively as their rows, $1 \le i \le m$.

We consider at first *the primal affine scaling Newton's method*. For its deriving from (4) we substitute the expression for $\mathrm{vech} V$ from the third equality to the first one and multiply both sides of that resulting equality by the matrix $\mathcal{A}_{vecs}$ from the left. Adding also to this equality the second equality from the (4) multiplied by some parameter $\tau > 0$, we obtain the following system of linear algebraic equations with respect to the vector u:

$$\Gamma(X) u = \mathcal{A}_{vecs} \tilde{X}^{\otimes} \mathrm{vech} C + \tau \left(b - \mathcal{A}_{vecs} \mathrm{vech} X \right), \quad (5)$$

where

$$\Gamma(X) = \mathcal{A}_{vecs} \tilde{X}^{\otimes} (\mathcal{A}_{vech})^T, \quad \tilde{X}^{\otimes} = \mathcal{L}_n X^{\otimes} \mathcal{D}_n.$$

If the matrix $\Gamma(X)$ is nonsingular, then solving system (5), we obtain

$$\begin{array}{rcl} u(X) & = & \Gamma^{-1}(X) \left[\mathcal{A}_{vecs} \tilde{X}^{\otimes} \mathrm{vech} C + \right. \\ & & \left. + \, \tau \left(b - \mathcal{A}_{vecs} \mathrm{vech} X \right) \right]. \end{array}$$

Denote $V(u) = C - \sum_{i=1}^m u^i A_i$, $V(X) = V(u(X))$. Substituting $V(X)$ into the first equality from (3), we obtain the nonlinear system of equations with respect to X:

$$X * V(X) = 0_{nn}. \quad (6)$$

Let $F(X) = X * V(X)$. Since the matrix function $F(X)$ is symmetric, the system (6) can be written in the form

$$\mathrm{vech} F(X) = 0_{k_\triangle(n)}.$$

Now we apply the Newton method for solving this system

$$\mathrm{vech} X_{k+1} = \mathrm{vech} X_k - \left[\mathcal{L}_n F_X(X_k) \mathcal{D}_n \right]^{-1} \mathrm{vech} F(X_k). \quad (7)$$

Here $F_X(X)$ is the Jacobian matrix of a symmetric matrix function $F(X)$.

Lemma 1 *The matrix $F_X(X)$ has the form:*

$$F_X(X) = V^{\otimes}(X) + X^{\otimes} V_X(X).$$

Therefore, to calculate $F_X(X)$ we need to know the Jacobian matrix $V_X(X)$ of the matrix function $V(X)$. Since $V_X(X) = -\mathcal{A}_{vec}^T u_X(X)$, the calculation of $V_X(X)$ is reduced to calculation of the Jacobian matrix $u_X(X)$. In the case, where the matrix $\Gamma(X)$ is nonsingular, we obtain

$$u_X(X) = \Gamma^{-1}(X) \left(\mathcal{A}_{vec} V^{\otimes}(X) - \tau \mathcal{A}_{vec} \right),$$

where $V^{\otimes} = \left[V \otimes I_n + I_n \otimes V \right]/2$ is a Kronecker sum of V.

To simplify our formulas we use the notation $\mathcal{P}(X) = X^{\otimes} \mathcal{A}_{vec}^T \Gamma^{-1}(X) \mathcal{A}_{vec}$. Then after substituting $u_X(X)$ into $V_X(X)$, we have

$$F_X(X) = \left[I_{n^2} - \mathcal{P}(X) \right] V^{\otimes}(X) + \tau \mathcal{P}(X).$$

Thus, the iterative process (7) can be written in more detailed form as

$$\begin{array}{rcl} \mathrm{vech} X_{k+1} & = & \mathrm{vech} X_k - \\ & - & \left[\left(I_{k_\triangle(n)} - \tilde{\mathcal{P}}_k \right) \tilde{V}_k^{\otimes} + \tau \tilde{\mathcal{P}}_k \right]^{-1} \cdot \\ & \cdot & \mathrm{vech} F(X_k), \end{array} \quad (8)$$

where $\tilde{\mathcal{P}}_k = \mathcal{L}_n \mathcal{P}(X_k) \mathcal{D}_n$, $\tilde{V}_k^{\otimes} = \mathcal{L}_n V^{\otimes}(X_k) \mathcal{D}_n$.

Let $\mathcal{T}(X)$ denote the tangent space of $\mathcal{S}^n$ at the point X. We denote also by $\mathcal{R}_A$ the subspace of $\mathcal{S}_n$ generated by matrices A_i, $1 \le i \le m$. Let $\mathcal{R}_A^{\perp}$ be the orthogonal complement of $\mathcal{R}_A$. Following [3], we give definitions of nondegenerate points in primal and dual problems (1), (2).

Definition 1 *A feasible point X of primal problem (1) is nondegerate if $\mathcal{T}(X) + \mathcal{R}_A^{\perp} = \mathcal{S}^n$. Similarly, a feasible point V of dual problem (2) is nondegenerate if $\mathcal{T}(V) + \mathcal{R}_A = \mathcal{S}^n$.*

Lemma 2 *Let X be a nondegenerate feasible point of primal problem (1). Then the matrix $\Gamma(X)$ is nonsingular.*

We assume that the problem (1) is *nondegenerate*, i.e. all feasible points X are nondegenerate. Then, due to continuity, there exists a neighborhood of the feasible set such that the iterative process (7) is completely determined in this neighborhood.

Theorem 1 *Assume that solutions X_* and V_* of primal and dual SDP problems (1) and (2) are strictly complementary. Let also the points X_* and V_* be nondegenerate. Then the method (7) locally converges to X_* at a superlinear rate.*

Now let us consider the dual analogue of the primal Newton method (7). With this aim we rewrite optimality conditions (3) as

$$\begin{array}{rcl} \mathcal{L}_n V^{\otimes}\mathcal{D}_n \mathrm{vech} X & = & 0_{k_\triangle(n)}, \\ \mathcal{A}_{vecs}\mathrm{vech} X & = & b, \\ \mathrm{vech} V & = & \mathrm{vech} C - \mathcal{A}^T_{vech} u, \\ X \succeq 0, & & V \succeq 0. \end{array} \qquad (9)$$

Multiplying both sides of the second equality in (9) by the matrix $\mathcal{A}^T_{vec}$ and adding this equality to the first one we comes to the following equation

$$\Phi(V)\mathrm{vech} X = \mathcal{A}^T_{vech} b, \qquad (10)$$

where

$$\Phi(V) = \mathcal{A}^T_{vech}\mathcal{A}_{vecs} + \mathcal{L}_n V^{\otimes}\mathcal{D}_n.$$

If the matrix $\Phi(V)$ is invertible, then, solving equation (10), we obtain

$$\mathrm{vech} X = \Phi^{-1}(V)\mathcal{A}^T_{vech} b. \qquad (11)$$

Thus $X = X(V)$.

After substituting the expression (11) for $\mathrm{vech}X$ at the second equality from (9) and taking $V = V(u)$ we derive the system of m nonlinear equations with respect of m variables $u_1, \ldots, m$:

$$\left[\mathcal{A}_{vecs}\Phi^{-1}(V(u))\mathcal{A}_{vech} - I_m\right] b = 0_m.$$

We apply also Newton's method for solving this system

$$u_{k+1} = u_k + \Lambda^{-1}(u_k)\left(b - \mathcal{A}_{vecs}\mathrm{vech} X_k\right), \qquad (12)$$

where $u_0 \in \mathbb{R}^m$, $X_k = X(u_k)$, $X(u) = X(V(u))$, and the matrix $\Lambda(u)$ is as follows

$$\Lambda(u) = \frac{d}{du}\mathcal{A}_{vecs}\Phi^{-1}(V(u))\mathcal{A}^T_{vech} b.$$

It is possible to derive that

$$\Lambda(u) = \mathcal{A}_{vecs}\Phi^{-1}(V(u))\mathcal{L}_n X^{\otimes}(u)\mathcal{D}_n\mathcal{A}^T_{vech}.$$

Therefore, the iterative process (12) can be written as

$$\begin{array}{rcl} u_{k+1} & = & u_k + \left\{\mathcal{A}_{vecs}\Phi^{-1}(V_k)\mathcal{L}_n X^{\otimes}_k\mathcal{D}_n\mathcal{A}^T_{vech}\right\}^{-1} \cdot \\ & \cdot & \left(b - \mathcal{A}_{vecs}\mathrm{vech} X_k\right), \end{array} \qquad (13)$$

where $V_k = V(u_k)$.

Lemma 3 *Let $V = V(u)$ be a nondegenerate feasible point of dual problem (2). Then the matrix $\Phi(V)$ is nonsingular.*

Similarly to previous case we assume that the dual problem is nondegenerate, i.e. all feasible points $V = V(u)$ of the dual problem (2) are nondegenerate. In this case, due to continuity, the iterative process (13) is fully determined in some neighborhood of the feasible set.

Theorem 2 *Assume that solutions X_* and $V_* = V(u_*)$ of primal and dual SDP problems (1), (2) are strictly complementary. Let also the points X_* and V_* be nondegenerate. Then the method (13) locally converges to u_* at a superlinear rate.*

Methods (7) and (13) can be regarded as extensions onto semidefinite programming the barrier-Newton methods, which was previously proposed for solving linear programming problems [4], [5]. The properties of these methods are given in [6], [7].

ACKNOWLEDGEMENTS. This study was supported by Russian Foundation for Basic Research (project no. 11-01-00786), by the Program for Support of Leading Scientific Scholls (project no NSh-4096.2010.1), and by the Program of Presidium of RAS (project No. Pi-14).

References

[1] E. De Klerk *Aspects of semidefinite programming. Interior point algorithms and selected applications.* Dordrecht etc.: Kluwer Acad. Publs, 2004.

[2] J. R. Magnus, N. Neudecker *Matrix Differential Calculas with Applications in Statistics and Econometrics.* Chichester, New York, Toronto: John Willey & Sons, 1988.

[3] *F. Alizadeh, J.-P. F. Haeberly, M. L. Overton Complementarity and nondegeneracy in semidefinite programming.* Mathematical Programming., *Series B.* **77**, *129–162 (1997).*

[4] Yu.G. Evtushenko, V.G. Zhadan Stable Barrier-Projection and Barrier-Newton Methods in Linear Programming. Comput. Optimaz. and Appl., **3**, *289-304 (1994).*

[5] Yu.G. Evtushenko, V.G. Zhadan Dual Barrier-*Projection and Barrier-Newton Methods for Linear Programming.* Comp. Maths Math. *Phys.,* **36**, *847-859 (1996).*

[6] V.G. Zhadan *Primal Newton's method for a linear problem of semidefinite programming* Proceedings of Steklov Institute of Mathemat*ics,* **263, Nu. 2 suppl.**, *(2008).*

[7] V.G. Zhadan, A.A. Orlov On convergence of the dual Newton method for linear semidefinite programming problem. Proceedings of Irkutsk state University, Mathematica, **4, Nu. 2**, *75-90 (2011), in Russian.*