REMARKS ON NONDEGENERACY IN MIXED SEMIDEFINITE–QUADRATIC PROGRAMMING

JEAN–PIERRE HAEBERLY

ABSTRACT. We consider the definitions of nondegeneracy and strict complementarity given in [5] for semidefinite programming (SDP) and their obvious extensions to mixed semidefinite–quadratic programming (SDQP). We show that a solution to SDQP satisfies strict complementarity and primal and dual nondegeneracy if and only if the Jacobian of the Newton system determined by the optimality conditions is nonsingular at the solution.

1. MIXED SEMIDEFINITE–QUADRATIC PROGRAMS

We begin by introducing some notation, based on [3]. Let $N = (N_1, \dots, N_s)$ and $n = (n_1, \dots, n_q)$ denote two vectors of positive integers. We define the following spaces.

- Let $\mathcal{S}^N \equiv \mathcal{S}^{N_1} \times \cdots \times \mathcal{S}^{N_s}$ denote the space of real, symmetric, block diagonal matrices with block sizes $N_1, \dots, N_s$. Let $\mathcal{S}_+^{N_i}$ denote the semidefinite cone of positive semidefinite $N_i \times N_i$ matrices and let $\mathcal{S}_+^N \equiv \mathcal{S}_+^{N_1} \times \cdots \times \mathcal{S}_+^{N_s}$. For a matrix $X_S \in \mathcal{S}^N$ we write $X_S \succeq 0$ to denote that $X_S \in \mathcal{S}_+^N$.
- Let $\mathcal{Q}^n \equiv \mathbb{R}^{n_1+1} \times \cdots \times \mathbb{R}^{n_q+1}$ denote the space of real block vectors with block sizes $n_1+1, \dots, n_q+1$. Let $\mathcal{Q}_+^{n_j} \subset \mathbb{R}^{n_j+1}$ denote the quadratic cone of vectors $(x_0, \dots, x_{n_j})^T$ satisfying
$$x_0 \geq \|\overline{x}\|,$$
where $\overline{x} \equiv (0, x_1, \dots, x_{n_j})^T$. Let $\mathcal{Q}_+^n \equiv \mathcal{Q}_+^{n_1} \times \cdots \times \mathcal{Q}_+^{n_q}$. For a vector $x_q \in \mathcal{Q}^n$ we write $x_Q \geq_Q 0$ to denote that $x_Q \in \mathcal{Q}_+^n$.

Both cones $\mathcal{S}_+^N$ and $\mathcal{Q}_+^n$ are closed, convex and self–dual. For a matrix $T \in \mathcal{S}^N$ (likewise a block vector in $\mathcal{Q}^n$), we refer to the i^{th} block of T as $T(i)$, and write $T = (T(1), \dots, T(s))$. Let

$$d_S = \sum_{i=1}^{s} \frac{N_i(N_i+1)}{2} \quad \text{and} \quad d_Q = \sum_{j=1}^{q} (n_j + 1) \tag{1.1}$$

denote the dimensions of $\mathcal{S}^N$ and $\mathcal{Q}^n$ respectively, and let $d = d_S + d_Q$, so that

$$\mathcal{S}^N \times \mathcal{Q}^n \cong \mathbb{R}^d.$$

A mixed semidefinite–quadratic program (SDQP) is an optimization problem of the form

$$\begin{array}{ll} \min & C_S \bullet X_S + c_Q \bullet x_Q \\ \text{s.t.} & (A_S)_k \bullet X_S + (a_Q)_k \bullet x_Q = b_k, \quad k = 1, \dots, m, \\ & X_S \succeq 0, \quad x_Q \geq_Q 0, \end{array} \tag{1.2}$$

Date: September 25, 1998.

where

$$C_S, (A_S)_k, X_S \in \mathcal{S}^N, \quad c_Q, (a_Q)_k, x_Q \in \mathcal{Q}^n, \quad \text{and} \quad b = (b_1, \ldots, b_m)^T \in \mathbb{R}^m.$$

The symbol $\bullet$ is used to denote the inner product on both $\mathcal{S}^N$ and $\mathcal{Q}^n$, so that for $U, V \in \mathcal{S}^N$, and $u, v \in \mathcal{Q}^n$, we have

$$U \bullet V \equiv \text{trace}\,(UV), \quad u \bullet v \equiv u^T v.$$

The dual of (1.2) is

$$\begin{aligned} \max \quad & b \bullet y \\ \text{s.t.} \quad & \sum_{k=1}^{m} y_k (A_S)_k + Z_S = C_S, \\ & \sum_{k=1}^{m} y_k (a_Q)_k + z_Q = c_Q, \\ & Z_S \succeq 0, \quad z_Q \geq_Q 0. \end{aligned} \tag{1.3}$$

Let us write $\mathcal{S}^N_{++}$ and $\mathcal{Q}^n_{++}$ for the interiors of $\mathcal{S}^N_+$ and $\mathcal{Q}^n_+$ respectively. We assume that the following conditions hold:

- *Assumption 1:* (Slater condition) There exist strictly feasible primal and dual points, *i.e.* there exist $X_S, Z_S \in \mathcal{S}^N_{++}$, $x_Q, z_Q \in \mathcal{Q}^n_{++}$, and $y \in \mathbb{R}^m$ satisfying the equality constraints in (1.2) and (1.3);
- *Assumption 2:* The $((A_S)_k, (a_Q)_k)$, $k = 1, \ldots, m$, are linearly independent in $\mathcal{S}^N \times \mathcal{Q}^n$.

Under these assumptions, (1.2) and (1.3) admit a unique solution satisfying the optimality conditions for SDQP [9], namely that, in addition to the constraints in (1.2) and (1.3), the duality gap be zero, *i.e.* $X_S \bullet Z_S + x_Q \bullet z_Q = 0$. This condition is equivalent to

$$X_S(i) \bullet Z_S(i) = 0, \ 1 \leq i \leq s, \quad \text{and} \quad x_Q(j) \bullet z_Q(j) = 0, \ 1 \leq j \leq q. \tag{1.4}$$

Remark 1.1. Observe that SDQP is equivalent to mixed semidefinite–quadratic–linear programming (see [3, 6]). Indeed, a linear program is exactly the same as an SDP with diagonal blocks of size 1. It is also possible to view quadratic programs as SDPs (see [8] for example). But such an embedding of a quadratic cone into a cone of positive semidefinite matrices is not useful from our point of view because it does not preserve duality, in the sense that the SDP dual of an embedded quadratic program is not the same as the embedding of the dual quadratic program. In particular, nondegeneracy of quadratic programs (see Section 2) is in general not preserved under such an embedding.

In order to describe the Newton system resulting from these optimality conditions we need to introduce some more notation. We write

$$X = (X_S, x_Q) \in \mathcal{S}^N \times \mathcal{Q}^n,$$

for the primal variable of SDQP and similarly $Z = (Z_S, z_Q)$ for the dual slack variable. We write $X \bullet Z \equiv X_S \bullet Z_S + x_Q \bullet z_Q$.

Following [1], given a block vector $x_Q \in \mathcal{Q}^n$, we define the "arrow matrix" $\mathbf{Arw}(x_Q(j))$ to be

$$\mathbf{Arw}(x_Q(j)) = \begin{bmatrix} x_0 & x_1 & x_2 & \cdots & x_{n_j} \\ x_1 & x_0 & 0 & \cdots & 0 \\ x_2 & 0 & x_0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ x_{n_j} & 0 & 0 & 0 & x_0 \end{bmatrix}$$

where $(x_0, x_1, \ldots, x_{n_j})^T$ is the vector $x_Q(j)$. We let $\mathbf{Arw}(x_Q)$ denote the block diagonal matrix whose j^{th} block is $\mathbf{Arw}(x_Q(j))$. We have $x_Q \geq_Q 0$ if and only if $\mathbf{Arw}(x_Q) \succeq 0$, and $x_Q >_Q 0$ if and only if $\mathbf{Arw}(x_Q) \succ 0$. In particular, for $x_Q \geq_Q 0$, the matrix $\mathbf{Arw}(x_Q)$ is nonsingular if and only if $x_Q >_Q 0$. Observe that for $v \in \mathbb{R}^{n_j+1}$ we have

$$\mathbf{Arw}(x_Q(j))v = (x_Q(j) \bullet v)e_Q(j) + x_0\overline{v} + v_0\overline{x} \tag{1.5}$$

where $e_Q(j) = (1, 0, \ldots, 0)^T \in \mathbb{R}^{n_j+1}$.

Now for $X_S, Z_S \succeq 0$, and $x_Q, z_Q \geq_Q 0$, the complementarity condition (1.4) reduces to the following system of nonlinear equations:

$$X_S Z_S + Z_S X_S = 0, \tag{1.6}$$

$$\mathbf{Arw}(x_Q)\ \mathbf{Arw}(z_Q)\ e_Q = 0. \tag{1.7}$$

Here, $e_Q = (e_Q(1), \ldots, e_Q(q)) \in \mathcal{Q}^n$. The complementarity condition (1.7) for the quadratic cone is discussed in [1]. The complementarity condition (1.6) for the semidefinite cone can be found in [2]. The equivalent symmetrized condition $X_S Z_S + Z_S X_S = 0$ results in the XZ+ZX direction [4] (also called the AHO direction in the SDP literature).

Thus the Newton system for SDQP based on the XZ+ZX direction is the nonlinear system of equations (1.6), (1.7), together with the equality constraints in (1.2) and (1.3). Let $\mathcal{L}$ denote the linear span of the $((A_S)_k, (a_Q)_k)$, $k = 1, \ldots, m$, in $\mathcal{S}^N \times \mathcal{Q}^n$. Note that Assumption 2 implies that $\mathcal{L}$ has dimension m. Given $X, Z \in \mathcal{S}^N \times \mathcal{Q}^n$, let $\mathcal{W}$ denote the linear subspace of $(\mathcal{S}^N \times \mathcal{Q}^n) \times (\mathcal{S}^N \times \mathcal{Q}^n)$ consisting of those $U = (U_S, u_Q)$ and $V = (V_S, v_Q)$ satisfying the following equations:

$$X_S V_S + V_S X_S + Z_S U_S + U_S Z_S = 0, \tag{1.8}$$

$$\mathbf{Arw}(x_Q)v_Q + \mathbf{Arw}(z_Q)u_Q = 0. \tag{1.9}$$

Thus $\mathcal{W}$ is the kernel of the linear map

$$(\mathcal{S}^N \times \mathcal{Q}^n) \times (\mathcal{S}^N \times \mathcal{Q}^n) \longrightarrow \mathcal{S}^N \times \mathcal{Q}^n$$

sending (U_S, u_Q, V_S, v_Q) to

$$(X_S V_S + V_S X_S + Z_S U_S + U_S Z_S, \mathbf{Arw}(x_Q)v_Q + \mathbf{Arw}(z_Q)u_Q).$$

In particular $\dim \mathcal{W} \geq d$. Now let (X, y, Z) be the solution to SDQP. It is easy to see that the Jacobian of the Newton system at (X, y, Z) is nonsingular if and only if

$$(\mathcal{L}^\perp \times \mathcal{L}) \cap \mathcal{W} = 0. \tag{1.10}$$

We conclude this section by recalling a few properties of the cones $\mathcal{S}^N_+$ and $\mathcal{Q}^n_+$. For $X_S \in \mathcal{S}^N_+$, let $0 \leq r_i \leq N_i$ denote the rank of the matrix $X_S(i)$. We let $\mathcal{T}_{X_S(i)} \subset \mathcal{S}^{N_i}$ denote the tangent space at $X_S(i)$ to the smooth manifold of symmetric matrices of rank r_i. If $Q(i)$ is an orthogonal matrix that diagonalizes $X_S(i)$ with

$$X_S(i) = Q(i) \begin{bmatrix} \Lambda & 0 \\ 0 & 0 \end{bmatrix} Q(i)^T \tag{1.11}$$

and $\Lambda = \mathbf{Diag}(\lambda_1, \dots, \lambda_{r_i})$, then we have the following representation for $\mathcal{T}_{X_S(i)}$ and its orthogonal complement [7]:

$$\mathcal{T}_{X_S(i)} = \left\{ Q(i) \begin{bmatrix} W_{11} & W_{12} \\ W_{12}^T & 0 \end{bmatrix} Q(i)^T \;\middle|\; W_{11} \in \mathcal{S}^{r_i},\ W_{12} \in \mathbb{R}^{r_i \times (N_i - r_i)} \right\}, \tag{1.12}$$

$$\mathcal{T}^{\perp}_{X_S(i)} = \left\{ Q(i) \begin{bmatrix} 0 & 0 \\ 0 & W_{22} \end{bmatrix} Q(i)^T \;\middle|\; W_{22} \in \mathcal{S}^{(N_i - r_i)} \right\}. \tag{1.13}$$

In particular, we have

$$V \in \mathcal{T}^{\perp}_{X_S(i)} \Longrightarrow X_S(i)\, V = 0. \tag{1.14}$$

We write

$$\mathcal{T}_{X_S} \equiv \prod_{i=1}^{s} \mathcal{T}_{X_S(i)} \subset \mathcal{S}^N.$$

Note that if X_S and Z_S satisfy the complementarity condition (1.6), then we may choose the matrices $Q(i)$ to diagonalize $X_S(i)$ and $Z_S(i)$ simultaneously with

$$Z_S(i) = Q(i) \begin{bmatrix} 0 & 0 \\ 0 & \Omega \end{bmatrix} Q(i)^T \tag{1.15}$$

and $\Omega = \mathbf{Diag}(\omega_1, \dots, \omega_{t_i})$, $t_i = \operatorname{rank} Z_S(i)$. In this case, we have the following representation for $\mathcal{T}_{Z_S(i)}$ and its orthogonal complement:

$$\mathcal{T}_{Z_S(i)} = \left\{ Q(i) \begin{bmatrix} 0 & W_{12} \\ W_{12}^T & W_{22} \end{bmatrix} Q(i)^T \;\middle|\; W_{22} \in \mathcal{S}^{t_i},\ W_{12} \in \mathbb{R}^{(N_i - t_i) \times t_i} \right\}, \tag{1.16}$$

$$\mathcal{T}^{\perp}_{Z_S(i)} = \left\{ Q(i) \begin{bmatrix} W_{11} & 0 \\ 0 & 0 \end{bmatrix} Q(i)^T \;\middle|\; W_{11} \in \mathcal{S}^{(N_i - t_i)} \right\}. \tag{1.17}$$

For $x_Q \in \mathcal{Q}^n$ with $x_Q(j) = (x_0, x_1, \dots, x_{n_j})^T$, $1 \le j \le q$, we write

$$\mathcal{T}_{x_Q(j)} = \begin{cases} \mathbb{R}^{n_j + 1} & \text{if } x_0 > \|\overline{x}\|, \\ \langle \nabla f(x_Q(j)) \rangle^{\perp} & \text{if } x_0 = \|\overline{x}\| > 0, \\ 0 & \text{if } x_0 = 0, \end{cases} \tag{1.18}$$

where $f(x_Q(j)) = x_0^2 - \sum_{k=1}^{n_j} x_k^2$ and $\langle \nabla f(x_Q(j)) \rangle$ denotes the linear span of the gradient of f at $x_Q(j)$. Its orthogonal complement, $\langle \nabla f(x_Q(j)) \rangle^{\perp} \subset \mathcal{Q}^{n_j}$, is then the tangent space to the boundary of the quadratic cone at $x_Q(j) \neq 0$. In particular, we have

$$v \in \mathcal{T}^{\perp}_{x_Q(j)} \Longrightarrow \mathbf{Arw}(x_q) v = 0. \tag{1.19}$$

We write

$$\mathcal{T}_{x_Q} \equiv \prod_{j=1}^{q} \mathcal{T}_{x_Q(j)} \subset \mathcal{Q}^n.$$

Finally, for $X = (X_S, x_Q) \in \mathcal{S}^N \times \mathcal{Q}^n$, we let

$$\mathcal{T}_X \equiv \mathcal{T}_{X_S} \times \mathcal{T}_{x_Q}.$$

Lemma 1.2. *Let $x \in \mathbb{R}^{p+1}$ satisfy $x_0 \geq \|\overline{x}\|$ and $x_0 > 0$. Then the set of solutions $z \in \mathbb{R}^{p+1}$ to the equation $x \bullet z = 0$ that satisfy $z_0 \geq \|\overline{z}\|$ is*

$$\begin{cases} \{0\}, & \text{if} \quad x_0 > \|\overline{x}\|, \\ \{\frac{z_0}{x_0}(x_0, -x_1, \ldots, -x_p) \mid z_0 \geq 0\}, & \text{if} \quad x_0 = \|\overline{x}\|. \end{cases}$$

Thus if $x \neq 0$ lies on the boundary of the quadratic cone then the set of solutions z to $x \bullet z = 0$ that lie in the quadratic cone consists of all the nonnegative multiples of $\nabla f(x)$.

Proof. (See [1]) For any x and z satisfying $x_0 \geq \|\overline{x}\|$, $z_0 \geq \|\overline{z}\|$ we have

$$-\overline{x} \bullet \overline{z} \leq \|\overline{x}\|\|\overline{z}\| \leq x_0 z_0$$

where the first inequality is Cauchy–Schwarz. Thus $x \bullet z = x_0 z_0 + \overline{x} \bullet \overline{z} = 0$ if and only if both inequalities are equalities. But $\|\overline{x}\|\|\overline{z}\| = x_0 z_0$ implies $z = 0$ if $x_0 > \|\overline{x}\|$, and $z_0 = \|\overline{z}\|$ if $x_0 = \|\overline{x}\|$.

If $x_0 = \|\overline{x}\| > 0$, then equality in Cauchy–Schwarz implies that

$$\overline{z} = \frac{\overline{x} \bullet \overline{z}}{\|\overline{x}\|^2}\overline{x}.$$

But $\overline{x} \bullet \overline{z} = -x_0 z_0$ and $\|\overline{x}\|^2 = x_0^2$, so the result follows. □

A corolloray to Lemma 1.2 is that if $x_Q(j), z_Q(j) \in \mathcal{Q}_+^{n_j}$ satisfy the complementarity condition

$$\mathbf{Arw}(x_Q(j))\, z_Q(j) = \mathbf{Arw}(x_Q(j))\, \mathbf{Arw}(z_Q(j))\, e_Q(j) = 0$$

then $x_Q(j)$, $z_Q(j)$ must satisfy one of the following five conditions:

$$x_Q(j) = 0 \quad \text{and} \quad z_Q(j) > \|\overline{z}\|, \tag{1.20}$$

$$z_Q(j) = 0 \quad \text{and} \quad x_Q(j) > \|\overline{x}\|, \tag{1.21}$$

$$x_Q(j) = \|\overline{x}\| > 0 \quad \text{and} \quad z_Q(j) = \|\overline{z}\| > 0, \tag{1.22}$$

$$x_Q(j) = 0 \quad \text{and} \quad z_Q(j) = \|\overline{z}\|, \tag{1.23}$$

$$z_Q(j) = 0 \quad \text{and} \quad x_Q(j) = \|\overline{x}\|. \tag{1.24}$$

Lemma 1.3. *Let $x \in \mathbb{R}^{p+1}$ satisfy $x_0 = \|\overline{x}\|$. Then there exists an orthogonal matrix Q such that $Qx = (x_0, -x_0, 0, \ldots, 0)^T$ and $\mathbf{Arw}(Qx) = Q\mathbf{Arw}(x)Q^T$. Moreover, if $x \neq 0$ then $\nabla f(Qx) = Q\nabla f(x)$. In particular, if z satisfies $x \bullet z = 0$ and $z_0 = \|\overline{z}\|$, then $Qz = (z_0, z_0, 0, \ldots, 0)^T$.*

Proof. Let $\overline{Q} \in \mathbb{R}^{p \times p}$ be an orthogonal matrix satisfying

$$\overline{Q}\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_p \end{bmatrix} = \begin{bmatrix} -x_0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

which is possible since $x_0 = \|(x_1, \ldots, x_p)\|$, and let

$$Q = \begin{bmatrix} 1 & 0 \\ 0 & \overline{Q} \end{bmatrix}.$$

It is easy to verify that Q satisfies the conclusions of the Lemma. □

2. Nondegeneracy and Strict Complementarity in SDQP

The concepts of primal and dual nondegeneracy for SDPs were introduced in [5] (see also [11, 10]) and were extended to mixed semidefinite–quadratic programs in [6, 3]. We briefly recall their definitions as well as the definition of strict complementarity discussed in [2, 3, 6]. We also prove a few Lemmas that will be needed in the proof of the main Theorem.

Definition 2.1. Let $X, Z \in \mathcal{S}_+^N \times \mathcal{Q}_+^n$. We say that X is primal nondegenerate if

$$\mathcal{T}_X + \mathcal{L}^\perp = \mathbb{R}^d$$

where the linear space $\mathcal{L}$ was defined in Section 1. We say that Z is dual nondegenerate if

$$\mathcal{T}_Z + \mathcal{L} = \mathbb{R}^d.$$

Definition 2.2. Let $X, Z \in \mathcal{S}_+^N \times \mathcal{Q}_+^n$ satisfy the complementarity conditions (1.6) and (1.7). We say that a block pair $(X_S(i), Z_S(i))$, $1 \le i \le s$, satisfies strict complementarity if $\operatorname{rank} X_S(i) + \operatorname{rank} Z_S(i) = N_i$. We say that a block pair $(x_Q(j), z_Q(j))$, $1 \le j \le q$, satisfies strict complementarity if it satisfies one of the conditions (1.20), (1.21), or (1.22). We say that (X_S, Z_S) (respectively (x_Q, z_Q)) satisfies strict complementarity if all block pairs $(X_S(i), Z_S(i))$, $1 \le i \le s$ (respectively $(x_Q(j), z_Q(j))$, $1 \le j \le q$) satisfy strict complementarity. Finally, we say that (X, Z) satisfies strict complementarity if both (X_S, Z_S) and (x_Q, z_Q) do.

Lemma 2.3. *Let $(X_S, Z_S) \in \mathcal{S}_+^N \times \mathcal{S}_+^N$ satisfy strict complementarity. Then a pair $(U, V) \in \mathcal{S}^N \times \mathcal{S}^N$ satisfies*

$$X_S V + V X_S + Z_S U + U Z_S = 0 \tag{2.1}$$

if and only if

$$X_S V + U Z_S = 0. \tag{2.2}$$

Proof. Clearly it is sufficient to show the result for each block. So fix i, $1 \le i \le s$, and let $Q(i)$ be an orthogonal matrix satisfying (1.11) and (1.15). Conjugating the i^{th} block of equation (2.1) by $Q(i)$ we get

$$\begin{bmatrix} \Lambda & 0 \\ 0 & 0 \end{bmatrix} \widehat{V} + \widehat{V} \begin{bmatrix} \Lambda & 0 \\ 0 & 0 \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ 0 & \Omega \end{bmatrix} \widehat{U} + \widehat{U} \begin{bmatrix} 0 & 0 \\ 0 & \Omega \end{bmatrix} = 0,$$

where

$$\widehat{U} \equiv Q(i)^T U Q(i) = \begin{bmatrix} \widehat{U}_{11} & \widehat{U}_{12} \\ \widehat{U}_{12}^T & \widehat{U}_{22} \end{bmatrix}, \quad \widehat{V} \equiv Q(i)^T V Q(i) = \begin{bmatrix} \widehat{V}_{11} & \widehat{V}_{12} \\ \widehat{V}_{12}^T & \widehat{V}_{22} \end{bmatrix}.$$

Let us write

$$M = \begin{bmatrix} \Lambda & 0 \\ 0 & 0 \end{bmatrix} \widehat{V} + \widehat{U} \begin{bmatrix} 0 & 0 \\ 0 & \Omega \end{bmatrix} = \begin{bmatrix} \Lambda \widehat{V}_{11} & \Lambda \widehat{V}_{12} + \widehat{U}_{12}\Omega \\ 0 & \widehat{U}_{22}\Omega \end{bmatrix}.$$

We have to show that $M = 0$ if and only if $M + M^T = 0$. But

$$M + M^T = 0 \iff \begin{cases} \Lambda \widehat{V}_{11} + \widehat{V}_{11}\Lambda = 0 \\ \Omega \widehat{U}_{22} + \widehat{U}_{22}\Omega = 0 \\ \Lambda \widehat{V}_{12} + \widehat{U}_{12}\Omega = 0 \end{cases}$$

Since $\Lambda \succ 0$ and $\Omega \succ 0$, we conclude that $M = 0$. □

Remark 2.4. Observe that the proof of Lemma 2.3 shows that if (X, Z) satisfies strict complementarity, then $(U, V) \in \mathcal{S}^N \times \mathcal{S}^N$ is a solution to (2.2) if and only if the blocks of U and V are of the following form

$$U(i) = Q(i) \begin{bmatrix} \widehat{U}_{11} & \widehat{U}_{12} \\ \widehat{U}_{12}^T & 0 \end{bmatrix} Q(i)^T, \quad V(i) = Q(i) \begin{bmatrix} 0 & -\Lambda^{-1}\widehat{U}_{12}\Omega \\ -\Omega\widehat{U}_{12}^T\Lambda^{-1} & \widehat{V}_{22} \end{bmatrix} Q(i)^T. \tag{2.3}$$

3. Main Result

We are now ready to state and prove our main result.

Theorem 3.1. *Let (X, y, Z) be a solution of SDPQ. Then the Jacobian of the Newton system at (X, y, Z) is nonsingular, i.e.* (1.10) *holds, if and only if X is primal nondegenerate, Z is dual nondegenerate, and (X, Z) satisfies strict complementarity.*

Proof. Let us assume that the Jacobian of the Newton system is nonsingular. We begin by showing that (X, Z) satisfies strict complementarity. Suppose that this is not the case, *i.e.* that strict complementarity is violated either by a pair $(X_S(i), Z_S(i))$, $1 \le i \le s$, or else by a pair $(x_Q(j), z_Q(j))$, $1 \le j \le q$. We derive a contradiction by showing that the dimension of $\mathcal{W}$ is greater than d so that $\mathcal{W}$ and $\mathcal{L}^\perp \times \mathcal{L}$ must intersect nontrivially. If $(x_Q(j), z_Q(j))$ violates strict complementarity then it satisfies conditions (1.23) or (1.24). It follows that one of $\mathbf{Arw}(x_Q(j))$ and $\mathbf{Arw}(z_Q(j))$ is zero and the other is singular. Hence the linear map from $\mathcal{Q}^{n_j} \times \mathcal{Q}^{n_j}$ to $\mathcal{Q}^{n_j}$ defined by

$$(u, v) \mapsto \mathbf{Arw}(x_Q(j)v + \mathbf{Arw}(z_Q(j))u$$

is not surjective, so its kernel has dimension greater than n_j and $\dim \mathcal{W} > d$. If $(X_S(i), Z_S(i))$ violates strict complementarity, then $\operatorname{rank} X_S(i) + \operatorname{rank} Z_S(i) < N_i$. Let $r_i = \operatorname{rank} X_S(i)$, $t_i = \operatorname{rank} Z_S(i)$, and $f_i = N_i - r_i - t_i > 0$. Let $Q(i)$ diagonalize $X_S(i)$ and $Z_S(i)$ with

$$X_S(i) = Q(i) \begin{bmatrix} \Lambda & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} Q(i)^T, \quad Z_S(i) = Q(i) \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & \Omega \end{bmatrix} Q(i)^T,$$

where $\Lambda = \mathbf{Diag}(\lambda_1, \ldots, \lambda_{r_i})$, $\Omega = \mathbf{Diag}(\omega_1, \ldots, \omega_{t_i})$. Consider the linear subspace of $\mathcal{S}^{N_i} \times \mathcal{S}^{N_i}$ consisting of all pairs of matrices (U, V) of the form

$$U = \begin{bmatrix} U_{11} & U_{12} & U_{13} \\ U_{12}^T & U_{22} & 0 \\ U_{13}^T & 0 & 0 \end{bmatrix}, \quad V = \begin{bmatrix} 0 & 0 & -\Lambda^{-1}U_{13}\Omega \\ 0 & V_{22} & V_{23} \\ -\Omega U_{13}^T\Lambda^{-1} & V_{23}^T & V_{33} \end{bmatrix},$$

where $U_{11} \in \mathcal{S}^{r_i}$, $U_{22}, V_{22} \in \mathcal{S}^{f_i}$, $V_{33} \in \mathcal{S}^{t_i}$, $U_{12} \in \mathbb{R}^{r_i \times f_i}$, $U_{13} \in \mathbb{R}^{r_i \times t_i}$, and $V_{23} \in \mathbb{R}^{f_i \times t_i}$. The dimension of this subspace is greater than N_i since U_{22} and V_{22} can be chosen independently of one another. It is easy to verify that this subspace lies in the kernel of the linear map from $\mathcal{S}^{N_i} \times \mathcal{S}^{N_i}$ to $\mathcal{S}^{N_i}$ defined by

$$(U, V) \mapsto X_S(i)V + VX_S(i) + Z_S(i)U + UZ_S(i).$$

We may conclude that $\dim \mathcal{W} > d$.

Next we show that X is primal nondegenerate and Z is dual nondegenerate, *i.e.* that $\mathcal{T}_X^\perp \cap \mathcal{L} = 0$ and $\mathcal{T}_Z^\perp \cap \mathcal{L}^\perp = 0$. So let $V \in \mathcal{T}_X^\perp \cap \mathcal{L}$ and $U \in \mathcal{T}_Z^\perp \cap \mathcal{L}^\perp$. By (1.14) and (1.19) we have

$$X_S V_S = 0, \quad Z_S U_S = 0, \quad \mathbf{Arw}(x_Q) v_Q = 0, \quad \text{and} \quad \mathbf{Arw}(z_Q) u_Q = 0,$$

showing that $(U, V) \in \mathcal{W}$. Since $(\mathcal{L}^\perp \times \mathcal{L}) \cap \mathcal{W} = 0$ by (1.10), we conclude that $U = 0$ and $V = 0$.

Now, for the converse, we assume that (X, Z) satisfies strict complementarity and that X is primal nondegenerate and Z is dual nondegenerate. We let $(U, V) \in (\mathcal{L}^\perp \times \mathcal{L}) \cap \mathcal{W}$ and we show that $U = 0$, $V = 0$, by showing that $U \in \mathcal{T}_Z^\perp \cap \mathcal{L}^\perp$ and $V \in \mathcal{T}_X^\perp \cap \mathcal{L}$. It is enough to show that for each $1 \leq i \leq s$ and $1 \leq j \leq q$

$$\begin{aligned} &X_S(i)V_S(i) + V_S(i)X_S(i) + Z_S(i)U_S(i) + U_S(i)Z_S(i) = 0 \\ &\qquad \Longrightarrow U_S(i) \in \mathcal{T}_{Z_S(i)}^\perp, \quad V_S(i) \in \mathcal{T}_{X_S(i)}^\perp \end{aligned} \tag{3.1}$$

$$\begin{aligned} &\mathbf{Arw}(x_Q(j))v_Q(j) + \mathbf{Arw}(z_Q(j))u_Q(j) = 0 \\ &\qquad \Longrightarrow u_Q(j) \in \mathcal{T}_{z_Q(j)}^\perp, \quad v_Q(j) \in \mathcal{T}_{x_Q(j)}^\perp. \end{aligned} \tag{3.2}$$

To prove (3.1), we use the description for $U_S(i)$ and $V_S(i)$ given in (2.3) and the fact that $U_S(i) \bullet V_S(i) = 0$ to get

$$\begin{aligned} 0 &= \left(Q(i) \begin{bmatrix} \widehat{U}_{11} & \widehat{U}_{12} \\ \widehat{U}_{12}^T & 0 \end{bmatrix} Q(i)^T \right) \bullet \left(Q(i) \begin{bmatrix} 0 & -\Lambda^{-1}\widehat{U}_{12}\Omega \\ -\Omega \widehat{U}_{12}^T \Lambda^{-1} & \widehat{V}_{22} \end{bmatrix} Q(i)^T \right) \\ &= \operatorname{trace} \left(\begin{bmatrix} \widehat{U}_{11} & \widehat{U}_{12} \\ \widehat{U}_{12}^T & 0 \end{bmatrix} \begin{bmatrix} 0 & -\Lambda^{-1}\widehat{U}_{12}\Omega \\ -\Omega \widehat{U}_{12}^T \Lambda^{-1} & \widehat{V}_{22} \end{bmatrix} \right) \\ &= \operatorname{trace} \begin{bmatrix} -\widehat{U}_{12}\Omega\widehat{U}_{12}^T\Lambda^{-1} & -\widehat{U}_{11}\Lambda^{-1}\widehat{U}_{12}\Omega + \widehat{U}_{12}\widehat{V}_{22} \\ 0 & -\widehat{U}_{12}^T\Lambda^{-1}\widehat{U}_{12}\Omega \end{bmatrix} \\ &= -\left[\operatorname{trace}(\widehat{U}_{12}\Omega\widehat{U}_{12}^T\Lambda^{-1}) + \operatorname{trace}(\widehat{U}_{12}^T\Lambda^{-1}\widehat{U}_{12}\Omega) \right] \\ &= -2\operatorname{trace}(\widehat{U}_{12}\Omega\widehat{U}_{12}^T\Lambda^{-1}). \end{aligned}$$

Since $\Lambda \succ 0$ and $\Omega \succ 0$, $\operatorname{trace}(\widehat{U}_{12}\Omega\widehat{U}_{12}^T\Lambda^{-1}) = 0$ if and only if $\widehat{U}_{12} = 0$. It follows that

$$U_S(i) = Q(i) \begin{bmatrix} \widehat{U}_{11} & 0 \\ 0 & 0 \end{bmatrix} Q(i)^T \in \mathcal{T}_{Z_S(i)}^\perp \cap \mathcal{L}^\perp$$

and

$$V_S(i) = Q(i) \begin{bmatrix} 0 & 0 \\ 0 & \widehat{V}_{22} \end{bmatrix} Q(i)^T \in \mathcal{T}_{X_S(i)}^\perp \cap \mathcal{L}.$$

To prove (3.2) we first observe that the result is obvious in cases (1.20) and (1.21). So we may assume that $x_Q(j) = (x_0, \ldots, x_{n_j})$ and $z_Q(j) = (z_0, \ldots, z_{n_j})$ satisfy $x_0 = \|\overline{x}\| > 0$ and $z_0 = \|\overline{z}\| > 0$. By Lemma 1.3 it is enough to consider the special case $x_Q(j) = (x_0, -x_0, 0, \ldots, 0)$ and $z_Q(j) = (z_0, z_0, 0, \ldots, 0)$. Then, with

$u_Q(j) = (u_0, \ldots, u_{n_j})$ and $v_Q(j) = (v_0, \ldots, v_{n_j})$, we have

$$\mathbf{Arw}(x_Q(j))\, v_Q(j) + \mathbf{Arw}(z_Q(j))\, u_Q(j) = \begin{bmatrix} x_0(v_0 - v_1) + z_0(u_0 + u_1) \\ x_0(v_1 - v_0) + z_0(u_0 + u_1) \\ x_0 v_2 + z_0 u_2 \\ \vdots \\ x_0 v_{n_j} + z_0 u_{n_j} \end{bmatrix}.$$

So we have

$$\begin{aligned} z_0(u_0 + u_1) + x_0(v_0 - v_1) &= 0, \\ z_0(u_0 + u_1) - x_0(v_0 - v_1) &= 0, \\ z_0 u_k + x_0 v_k, \quad 2 \le k &\le n_j. \end{aligned}$$

It follows that $u_1 = -u_0$, $v_1 = -v_0$, and it only remains to show that $u_k = 0$, $v_k = 0$, for $2 \le k \le n_j$. But

$$0 = u_Q(j) \bullet v_q(j) = u_0 v_0 + u_1 v_1 + \sum_{k=2}^{n_j} u_k v_k = \sum_{k=2}^{n_j} u_k v_k.$$

The conclusion follows since $v_k = -\frac{z_0}{x_0} u_k$, $2 \le k \le n_j$. □

References

1. I. Adler and F.Alizadeh, *Primal–dual interior point algorithms for convex quadratically constrained and semidefinite optimization problems*, Tech. Report RRR 46–95, RUTCOR. Rutgers University, Piscataway, NJ, December 1995.
2. F. Alizadeh, *Interior–point methods in semidefinite programming with applications to combinatorial optimization*, SIAM Journal on Optimization **5** (1995), 13–51.
3. F. Alizadeh, J.-P. Haeberly, M.V. Nayakkankuppam, M.L. Overton, and S. Schmieta, *SDPpack User Guide (Version 0.9 Beta)*, Tech. Report 737, Department of Computer Science, New York University, New York, NY, June 1997.
4. F. Alizadeh, J.-P. Haeberly, and M. L. Overton, *Primal–dual interior–point methods for semidefinite programming: convergence rates, stability, and numerical methods*, SIAM Journal on Optimization (1998), to appear.
5. F. Alizadeh, J.-P. Haeberly, and M.L. Overton, *Complementarity and nondegeneracy in semidefinite programming*, Mathematical Programming **77** (1997), 111–128.
6. F. Alizadeh and S. Schmieta, *Optimization with semidefinite, quadratic and linear constraints*, Tech. Report RRR23-97, RUTCOR. Rutgers University, Piscataway, NJ, November 1997.
7. V.I. Arnold, *On matrices depending on parameters*, Russian Mathematical Surveys **26** (1971), 29–43.
8. M. Lobo, L. Vandenberghe, S. Boyd, and H. Lebret, *Second–order cone programming*, Linear Algebra and its Applications (1997), to appear.
9. Yu. E. Nesterov and A. Nemirovskii, *Interior–point polynomial algorithms in convex programming*, Studies in applied mathematics, vol. 13, SIAM, Philadelphia, PA, 1994.
10. G. Pataki, *Cone lp's and semidefinite programs: Geometry and a simplex–type method*, Proceedings of the Fifth Integer Programming and Combinatorial Optimization Conference (Waterloo), 1996.
11. A. Shapiro, *First and seconf order analysis of nonlinear semidefinite programs*, Mathematical Programming **77** (1997), 301–320.

Department of Mathematics, Fordham University
E-mail address: `haeberly@murray.fordham.edu`