Volume 31, N. 1, pp. 205–218, 2012

ISSN 0101-8205 / ISSN 1807-0302 (Online)
www.scielo.br/cam

A sensitivity result for quadratic semidefinite programs with an application to a sequential quadratic semidefinite programming algorithm

RODRIGO GARCÉS[1], WALTER GÓMEZ[2]* and FLORIAN JARRE[3]

[1]EuroAmerica S.A., Av. Apoquindo 3885, Las Condes, Santiago, Chile,
[2]Department of Mathematical Engineering, Universidad de La Frontera, Av. Francisco Salazar, 01145, Temuco, Chile
[3]Institut für Mathematik, Universität Düsseldorf, Universitätsstraße 1, D–40225 Düsseldorf, Germany

E-mails: rgarces@euroamerica.cl / wgomez@ufro.cl / jarre@opt.uni-duesseldorf.de

Abstract. In this short note a sensitivity result for quadratic semidefinite programming is presented under a weak form of second order sufficient condition. Based on this result, also the local convergence of a sequential quadratic semidefinite programming algorithm extends to this weak second order sufficient condition.

Mathematical subject classification: 90C22, 90C30, 90C31, 90C55.

Key words: semidefinite programming, second order sufficient condition, sequential quadratic programming, quadratic semidefinite programming, sensitivity, convergence.

1 Introduction

A sequential semidefinite programming (SSDP) algorithm for solving nonlinear semidefinite programs was proposed in [5, 7]. It is a generalization of the well-known sequential quadratic programming (SQP) method and considers linear semidefinite subproblems that can be solved using standard interior point packages. Note that linear SDP relaxations with a convex quadratic objective

#CAM-362/11. Received: 27/IV/11. Accepted: 24/V/11.
*Corresponding author

function can be transformed to equivalent linear SDP subproblems. However, as shown in [4], under standard assumptions, local superlinear convergence is possible only when the iterates are defined by SDP relaxations with a nonconvex quadratic objective function. Since this class of problems is no longer equivalent to the linear semidefinite programming case we refer to the algorithm in this note as Sequential Quadratic Semidefinite Programming (SQSDP) method.

In the papers [5, 7] a proof is given showing local quadratic convergence of the SSDP algorithm to a local minimizer assuming a strong second order sufficient condition. This condition ensures, in particular, that the quadratic SDP subproblems close to the local minimizers are convex, and therefore reducible to the linear SDP case. However, as pointed out in [4], there are examples of perfectly well-conditioned nonlinear SDP problems that do not satisfy the strong second order sufficient condition used in [5, 7].

These examples satisfy a weaker second order condition [10], that considers explicitly the curvature of the semidefinite cone.

In this short note we study the sensitivity of quadratic semidefinite problems (the subproblems of SQSDP), using the weaker second order condition. Based on this sensitivity result, the fast local convergence of the SQSDP method can also be established under the weaker assumption in [10]; in this case the quadratic SDP subproblems may be nonconvex.

The sensitivity results presented in this paper were used in [8] for the study of a local self-concordance property for certain nonconvex quadratic semidefinite programming problems.

2 Notation and preliminaries

By $\mathbb{S}^m$ we denote the linear space of $m \times m$ real symmetric matrices. The space $\mathbb{R}^{m\times n}$ is equipped with the inner product

$$A \bullet B := \text{trace}(A^T B) = \sum_{i=1}^{m} \sum_{j=1}^{n} A_{ij} B_{ij}.$$

The corresponding norm is the Frobenius norm defined by $\|A\|_F = \sqrt{A \bullet A}$. The negative semidefinite order $\preceq$ for $A, B \in \mathbb{S}^m$ is defined in the standard form, that is, $A \preceq B$ iff $A - B$ is a negative semidefinite matrix. The order relations $\prec$, $\succeq$ and $\succ$ are defined similarly. By $\mathbb{S}^m_+$ we denote the set of positive semidefinite matrices.

The following simple Lemma is used in the sequel.

Lemma 1 (See [7]). *Let $Y, S \in \mathbb{S}^m$.*

(a) *If $Y, S \succeq 0$ then*

$$YS + SY = 0 \iff YS = 0. \tag{1}$$

(b) *If $Y + S \succ 0$ and $YS + SY = 0$ then $Y, S \succeq 0$.*

(c) *If $Y + S \succ 0$ and $YS + SY = 0$ then for any $\dot{Y}, \dot{S} \in \mathbb{S}^m$,*

$$Y\dot{S} + \dot{Y}S = 0 \iff Y\dot{S} + \dot{Y}S + \dot{S}Y + S\dot{Y} = 0. \tag{2}$$

Moreover, Y, S have representations of the form

$$Y = U \begin{bmatrix} Y_1 & 0 \\ 0 & 0 \end{bmatrix} U^T, \quad S = U \begin{bmatrix} 0 & 0 \\ 0 & S_2 \end{bmatrix} U^T,$$

where U is an $m \times m$ orthogonal matrix, $Y_1 \succ 0$ is a $(m-r) \times (m-r)$ diagonal matrix and $S_2 \succ 0$ is a $r \times r$ diagonal matrix, and any matrices $\dot{Y}, \dot{S} \in \mathbb{S}^m$ satisfying (2) are of the form

$$\dot{Y} = U \begin{bmatrix} \dot{Y}_1 & \dot{Y}_3 \\ \dot{Y}_3^T & 0 \end{bmatrix} U^T, \quad \dot{S} = U \begin{bmatrix} 0 & \dot{S}_3 \\ \dot{S}_3^T & \dot{S}_2 \end{bmatrix} U^T,$$

where

$$Y_1 \dot{S}_3 + \dot{Y}_3 S_2 = 0. \tag{3}$$

Proof. For (a), (c) see [7].

(b) By contradiction we assume that λ is a negative eigenvalue of S and u a corresponding eigenvector. The equality $YS + SY = 0$ implies that

$$\begin{aligned} u^T Y(Su) + (Su)^T Yu &= 0, \\ u^T Y(\lambda u) + (\lambda u)^T Yu &= 0, \\ \lambda(u^T Yu + u^T Yu) &= 0, \\ \Rightarrow u^T Yu &= 0, \text{ since } \quad \lambda < 0. \end{aligned}$$

Now using the fact that $Y + S \succ 0$, we have

$$0 < u^T(Y + S)u = u^T Yu + u^T Su = \lambda u^T u = \lambda \parallel u \parallel^2 < 0$$

which is a contradiction. Hence, $S \succeq 0$. The same arguments give us $Y \succeq 0$. □

Remark 1. Due to (3) and the positive definiteness of the diagonal matrices Y_1 and S_2, it follows that $(\dot{Y}_3)_{ij}(\dot{S}_3)_{ij} < 0$ whenever $(\dot{Y}_3)_{ij} \neq 0$. Hence, if, in addition to (3), also $\langle \dot{Y}_3, \dot{S}_3 \rangle = 0$ holds true, then $\dot{Y}_3 = \dot{S}_3 = 0$.

In the sequel we refer to the set of symmetric and strict complementary matrices

$$C = \left\{(Y, S) \in \mathbb{S}^m \times \mathbb{S}^m \mid YS + SY = 0,\ Y + S \succ 0\right\}. \tag{4}$$

As a consequence of Lemma 1(b), the set C is (not connected, in general, but) contained in $\mathbb{S}^m_+ \times \mathbb{S}^m_+$. Moreover, Lemma 1(c) implies that the rank of the matrices Y and S is locally constant on C.

2.1 Nonlinear semidefinite programs

Given a vector $b \in \mathbb{R}^n$ and a matrix-valued function $G : \mathbb{R}^n \to \mathbb{S}^m$, we consider problems of the following form:

$$\underset{x \in \mathbb{R}^n}{\text{minimize}} \quad b^T x \quad \text{subject to} \quad G(x) \preceq 0. \tag{5}$$

Here, the function G is at least C^3-differentiable.

For simplicity of presentation, we have chosen a simple form of problem (5). All statements about (5) in this paper can be modified so that they apply to additional nonlinear equality and inequality constraints and to nonlinear objective functions. The notation and assumptions in this subsection are similar to the ones used in [8].

The Lagrangian $\mathcal{L} : \mathbb{R}^n \times \mathbb{S}^m \to \mathbb{R}$ of (5) is defined as follows:

$$\mathcal{L}(x, Y) := b^T x + G(x) \bullet Y. \tag{6}$$

Its gradient with respect to x is given by

$$g(x, Y) := \nabla_x \mathcal{L}(x, Y) = b + \nabla_x \left(G(x) \bullet Y\right) \tag{7}$$

and its Hessian by

$$H(x, Y) := \nabla^2_x \mathcal{L}(x, Y) = \nabla^2_x \left(G(x) \bullet Y\right). \tag{8}$$

Assumptions.

(A1) We assume that $\bar{x}$ is a local minimizer of (5) that satisfies the Mangasarian-Fromovitz constraint qualification, i.e., there exists a vector $\Delta x \neq 0$ such that $G(\bar{x}) + DG(\bar{x})[\Delta x] \prec 0$, where by definition $DG(x)[s] = \sum_{i=1}^{n} s_i D_{x_i} G(x)$.

Assumption (**A1**) implies that the first-order optimality condition is satisfied, i.e., there exist matrices $\bar{Y},\ \bar{S} \in \mathbb{S}^m$ such that

$$\begin{aligned} G(\bar{x}) + \bar{S} &= 0, \\ g(\bar{x}, \bar{Y}) &= 0, \\ \bar{Y}\bar{S} &= 0, \\ \bar{Y},\ \bar{S} &\succeq 0. \end{aligned} \tag{9}$$

A triple $(\bar{x}, \bar{Y}, \bar{S})$ satisfying (9), will be called a stationary point of (5).

Due to Lemma 1(a) the third equation in (9) can be substituted by $\bar{Y}\bar{S} + \bar{S}\bar{Y} = 0$. This reformulation does not change the set of stationary points, but it reduces the underlying system of equations (via a symmetrization of YS) in the variables (x, Y, S), such that it has now the same number of equations and variables. This is a useful step in order to apply the implicit function theorem.

(A2) We also assume that $\bar{Y}$ is unique and that $\bar{S}, \bar{Y}$ are strictly complementary, i.e. $(\bar{Y}, \bar{S}) \in C$.

According to Lemma 1(c), there exists a unitary matrix $U = [U_1, U_2]$ that simultaneously diagonalizes $\bar{Y}$ and $\bar{S}$. Here, U_2 has $r := \operatorname{rank}(\bar{S})$ columns and U_1 has $m - r$ columns. Moreover the first $m - r$ diagonal entries of $U^T \bar{S} U$ are zero, and the last r diagonal entries of $U^T \bar{Y} U$ are zero. In particular, we obtain

$$U_1^T G(\bar{x}) U_1 = 0 \quad \text{and} \quad U_2^T \bar{Y} U_2 = 0. \tag{10}$$

A vector $h \in \mathbb{R}^n$ is called a critical direction at $\bar{x}$ if $b^T h = 0$ and it is the limit of feasible directions of (5), i.e. if there exist $h^k \in \mathbb{R}^n$ and $\epsilon_k > 0$ with $\lim_{k\to\infty} h^k = h$, $\lim_{k\to\infty} \epsilon_k = 0$, and $G(\bar{x} + \epsilon_k h^k) \preceq 0$ for all k. As shown in [1] the cone of critical directions at a strictly complementary local solution $\bar{x}$ is given by

$$C(\bar{x}) := \{h \mid U_1^T DG(\bar{x})[h]U_1 = 0\}. \tag{11}$$

In the following we state second order sufficient conditions due to [10] that are weaker than the ones used in [5, 7].

(A3) We further assume that $\bar{x}$, $\bar{Y}$ satisfies the second order sufficient condition:

$$h^T(\nabla_x^2 \mathcal{L}(\bar{x}, \bar{Y}) + \mathcal{H}(\bar{x}, \bar{Y}))h > 0 \qquad \forall\, h \in C(\bar{x}) \setminus \{0\} \tag{12}$$

Here $\mathcal{H}$ is a nonnegative matrix related to the curvature of the semidefinite cone in $G(\bar{x})$ along direction $\bar{Y}$ (see [10]) and is given by its matrix entries

$$\mathcal{H}_{i,j} := -2\bar{Y} \bullet G_i(\bar{x})G(\bar{x})^\dagger G_j(\bar{x}),$$

where $G_i(\bar{x}) := DG(\bar{x})[e_i]$ with e_i denoting the i-th unit vector. Furthermore, $G(\bar{x})^\dagger$ denotes the Moore-Penrose pseudo-inverse of $G(\bar{x})$, i.e.

$$G(\bar{x})^\dagger = \sum \lambda_i^{-1} u_i u_i^T,$$

where λ_i are the nonzero eigenvalues of $G(\bar{x})$ and u_i corresponding orthonormal eigenvectors.

Remark 2. The Moore-Penrose inverse $M^\dagger$ is a continuous function of M, when the perturbations of M do not change its rank, see [3].

The curvature term can be rewritten as follows:

$$\begin{aligned} h^T\mathcal{H}(\bar{x}, \bar{Y})h &= \sum_{i,j} h_i h_j(-2\bar{Y} \bullet G_i(\bar{x})G(\bar{x})^\dagger G_j(\bar{x})), \\ &= -2\bar{Y} \bullet \left(\sum_{i,j} h_i h_j G_i(\bar{x})G(\bar{x})^\dagger G_j(\bar{x}) \right), \qquad (13) \\ &= -2\bar{Y} \bullet \left(\sum_{i=1}^{n} h_i G_i(\bar{x})G(\bar{x})^\dagger \sum_{j=1}^{n} h_j G_j(\bar{x}) \right), \\ &= -2\bar{Y} \bullet DG(\bar{x})[h]G(\bar{x})^\dagger DG(\bar{x})[h]. \end{aligned}$$

Note that in the particular case where G is affine (i.e. $G(x) = \mathcal{A}(x) + C$, with a linear map $\mathcal{A}$ and $C \in \mathbb{S}^m$), the curvature term is given by

$$h^T \mathcal{H}(\bar{x}, \bar{Y})h := -2\bar{Y} \bullet (\mathcal{A}(h)(\mathcal{A}(\bar{x}) + C)^\dagger \mathcal{A}(h)). \tag{14}$$

The following very simple example of [4] shows that the classical second order sufficient condition is generally too strong in the case of semidefinite constraints, since it does not exploit curvature of the non-polyhedral semidefinite cone.

$$\begin{aligned} \min_{x\in\mathbb{R}^2} \quad & -x_1 - (x_2 - 1)^2 \\ \text{s.t.} \quad & \\ & \begin{bmatrix} -1 & 0 & -x_1 \\ 0 & -1 & -x_2 \\ -x_1 & -x_2 & -1 \end{bmatrix} \preceq 0 \end{aligned} \tag{15}$$

It is a trivial task to check that the constraint $G(x) \preceq 0$ is equivalent to the inequality $x_1^2 + x_2^2 \leq 1$, such that $\bar{x} = (0, -1)^T$ is the global minimizer of the problem.

The first order optimality conditions (9) are satisfied at $\bar{x}$ with associated multiplier

$$\bar{Y} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 2 & 2 \\ 0 & 2 & 2 \end{bmatrix}.$$

The strict complementarity condition also holds true, since

$$\bar{Y} - G(\bar{x}) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 3 & 1 \\ 0 & 1 & 3 \end{bmatrix} \succ 0$$

The Hessian of the Lagrangian at $(\bar{x}, \bar{Y})$ for this problem can be calculated as

$$\nabla^2_{xx} \mathcal{L}(\bar{x}, \bar{Y}) = \begin{bmatrix} -2 & 0 \\ 0 & -2 \end{bmatrix}.$$

It is negative definite, and the stronger second order condition is not satisfied.

In order to calculate the curvature term in (12) let us consider the orthogonal matrix U, which simultaneously diagonalizes $\bar{Y}$, $G(\bar{x})$

$$U = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ 0 & -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix}.$$

The Moore-Penrose pseudoinverse matrix at $\bar{x}$ is then given by

$$G(\bar{x})^{\dagger} = \frac{-1}{4} \begin{bmatrix} 4 & 0 & 0 \\ 0 & 1 & -1 \\ 0 & -1 & 1 \end{bmatrix}, \tag{16}$$

and the matrix associated to the curvature becomes

$$\mathcal{H}(\bar{x}, \bar{Y}) = \begin{bmatrix} 4 & 0 \\ 0 & 0 \end{bmatrix}.$$

Finally, every $h \in \mathbb{R}^2$ such that $\nabla f(\bar{x})^T \cdot h = 0$, has the form $h = (h_1, 0)^T$ with $h_1 \in \mathbb{R}$. Therefore, the weaker second order sufficient condition holds, i.e.,

$$h^t(\nabla^2_{xx}\mathcal{L}(\bar{x}, \bar{Y}) + \mathcal{H}(\bar{x}, \bar{Y}))h = 2h_1^2 > 0 \qquad \forall h \in C(\bar{x}) \setminus \{0\}.$$

3 Sensitivity result

Let us now consider the following quadratic semidefinite programming problem

$$\begin{array}{ll} \min\limits_{x \in \mathbb{R}^n} & b^T x + \frac{1}{2} x^T H x \\ \text{s.t.} & \\ & \mathcal{A}(x) + C \preceq 0. \end{array} \tag{17}$$

Here, $\mathcal{A} : \mathbb{R}^n \longrightarrow \mathbb{S}^m$ is a linear function, $b \in \mathbb{R}^n$, and $C, H \in \mathbb{S}^m$. The data to this problem is

$$\mathcal{D} := [\mathcal{A}, b, C, H]. \tag{18}$$

In the next theorem, we present a sensitivity result for the solutions of (17), when the data $\mathcal{D}$ is changed to $\mathcal{D} + \triangle\mathcal{D}$ where

$$\triangle\mathcal{D} := [\triangle\mathcal{A}, \triangle b, \triangle C, \triangle H] \tag{19}$$

is a sufficiently small perturbation.

The triple $(\bar{x}, \bar{Y}, \bar{S}) \in \mathbb{R}^n \times \mathbb{S}^m \times \mathbb{S}^m$ is a stationary point for (17), if

$$\begin{aligned} \mathcal{A}(\bar{x}) + C + \bar{S} &= 0, \\ b + H\bar{x} + \mathcal{A}^*(\bar{Y}) &= 0, \\ \bar{Y}\bar{S} + \bar{S}\bar{Y} &= 0, \\ \bar{Y}, \bar{S} &\succeq 0. \end{aligned} \tag{20}$$

Remark 3. Below, we consider tiny perturbations $\triangle\mathscr{D}$ such that there is an associated strictly complementary solution $(x, Y, S)(\triangle\mathscr{D})$ of (20). For such x there exists $U_1 = U_1(x)$ and an associated cone of critical directions $C(x)$. The basis $U_1(x)$ generally is not continuous with respect to x. However, the above characterization (11) of $C(\bar{x})$ under strict complementarity can be stated using any basis of the orthogonal space of $G(\bar{x})$. Since such basis can be locally parameterized in a smooth way over the set C in (4) it follows that locally, the set $C(x)$ forms a closed point to set mapping.

The following is a slight generalization of Theorem 1 in [7].

Theorem 1. *Let the point* $(\bar{x}, \bar{Y}, \bar{S})$ *be a stationary point satisfying the assumptions (**A1**)-(**A3**) for the problem* (17) *with data* $\mathscr{D}$*. Then, for all sufficiently small perturbations* $\triangle\mathscr{D}$ *as in* (19)*, there exists a locally unique stationary point* $(\bar{x}(\mathscr{D} + \triangle\mathscr{D}), \bar{Y}(\mathscr{D} + \triangle\mathscr{D}), \bar{S}(\mathscr{D} + \triangle\mathscr{D}))$ *of the perturbed program* (17) *with data* $\mathscr{D} + \triangle\mathscr{D}$*. Moreover, the point* $(\bar{x}(\mathscr{D}), \bar{Y}(\mathscr{D}), \bar{S}(\mathscr{D}))$ *is a differentiable function of the perturbation* (19)*, and for* $\triangle\mathscr{D} = 0$*, we have* $(\bar{x}(\mathscr{D}), \bar{Y}(\mathscr{D}), \bar{S}(\mathscr{D})) = (\bar{x}, \bar{Y}, \bar{S})$*. The derivative* $D_{\mathscr{D}}(\bar{x}(\mathscr{D}), \bar{Y}(\mathscr{D}), \bar{S}(\mathscr{D}))$ *of* $(\bar{x}(\mathscr{D}), \bar{Y}(\mathscr{D}), \bar{S}(\mathscr{D}))$ *with respect to* $\mathscr{D}$ *evaluated at* $(\bar{x}, \bar{Y}, \bar{S})$ *is characterized by the directional derivatives*

$$(\dot{x}, \dot{Y}, \dot{S}) := D_{\mathscr{D}}(\bar{x}(\mathscr{D}), \bar{Y}(\mathscr{D}), \bar{S}(\mathscr{D}))[\triangle\mathscr{D}]$$

for any $\triangle\mathscr{D}$*. Here,* $(\dot{x}, \dot{Y}, \dot{S})$ *is the unique solution of the system of linear equations,*

$$\begin{aligned} \mathcal{A}(\dot{x}) + \dot{S} &= -\triangle C - \triangle\mathcal{A}(\bar{x}), \\ H\dot{x} + \mathcal{A}^*(\dot{Y}) &= -\triangle b - \triangle H\bar{x} - \triangle\mathcal{A}^*(\bar{Y}), \\ \bar{Y}\dot{S} + \dot{Y}\bar{S} + \dot{S}\bar{Y} + \bar{S}\dot{Y} &= 0, \end{aligned} \tag{21}$$

for the unknowns $\dot{x} \in \mathbb{R}^n$, $\dot{Y}, \dot{S} \in \mathbb{S}^m$. *Finally, the second-order sufficient condition holds at* $(\bar{x}(\mathscr{D}), \bar{Y}(\mathscr{D}))$ *whenever* $\triangle\mathscr{D}$ *is sufficiently small.*

This theorem is related to other sensitivity results for semidefinite programming problems (see, for instance, [2, 6, 11]). Local Lipschitz properties under strict complementarity can be found in [9]. In [10] the directional derivative $\dot{x}$ is given as solution of a quadratic problem.

Proof. Following the outline in [7] this proof is based on the application of the implicit function theorem to the system of equations (20). In order to apply this result we show that the matrix of partial derivatives of system (20) with respect to the variables (x, Y, S) is regular. To this end it suffices to prove that the system

$$\begin{aligned} \mathcal{A}(\dot{x}) + \dot{S} &= 0, \\ H\dot{x} + \mathcal{A}^*(\dot{Y}) &= 0, \\ \bar{Y}\dot{S} + \dot{Y}\bar{S} + \dot{S}\bar{Y} + \bar{S}\dot{Y} &= 0, \end{aligned} \tag{22}$$

only has the trivial solution $\dot{x} = 0$, $\dot{Y} = \dot{S} = 0$.

Let $(\dot{x}, \dot{Y}, \dot{S}) \in \mathbb{R}^n \times \mathbb{S}^m \times \mathbb{S}^m$ be a solution of (22). Since $\bar{Y}$ and $\bar{S}$ are strictly complementary, it follows from part (c) of Lemma 1, the existence of an orthonormal matrix U such that:

$$\bar{Y} = U\tilde{Y}U^T, \quad \bar{S} = U\tilde{S}U^T \tag{23}$$

where

$$\tilde{Y} = \begin{bmatrix} \bar{Y}_1 & 0 \\ 0 & 0 \end{bmatrix}, \quad \tilde{S} = \begin{bmatrix} 0 & 0 \\ 0 & \bar{S}_2 \end{bmatrix}, \tag{24}$$

with $\bar{Y}_1, \bar{S}_2$ diagonal and positive definite. Furthermore, the matrices $\dot{Y}, \dot{S} \in \mathbb{S}^m$ satisfying (22) fulfill the relations

$$\dot{Y} = U\check{Y}U^T, \quad \dot{S} = U\check{S}U^T \tag{25}$$

where

$$\check{Y} = \begin{bmatrix} \dot{Y}_1 & \dot{Y}_3 \\ \dot{Y}_3^T & 0 \end{bmatrix}, \quad \check{S} = \begin{bmatrix} 0 & \dot{S}_3 \\ \dot{S}_3^T & \dot{S}_2 \end{bmatrix}, \quad \text{and} \quad \dot{Y}_3\bar{S}_2 + \bar{Y}_1\dot{S}_3 = 0. \tag{26}$$

Using the decomposition given in (10), the first equation of (22), and (25) we have

$$U_1^T \mathcal{A}(\dot{x}) U_1^T = -U_1^T U \check{S} U^T U_1 = 0.$$

It follows that $\dot{x} \in C(\bar{x})$. Now using (14), (23-26), the first equation in (20), and the first equation in (22), we obtain

$$\begin{aligned} \dot{x}^T \mathcal{H}(\bar{x}, \bar{Y})\dot{x} &= -2\bar{Y} \bullet \mathcal{A}(\dot{x})(\mathcal{A}(\bar{x}) + C)^\dagger \mathcal{A}(\dot{x}), \\ &= -2\bar{Y} \bullet \dot{S}(-\bar{S})^\dagger \dot{S}, \\ &= -2\tilde{Y} \bullet \check{S}(-\tilde{S})^\dagger \check{S}, \quad \text{since} \quad \bar{S}_2 \succ 0, \quad \tilde{S}^\dagger = \begin{bmatrix} 0 & 0 \\ 0 & \bar{S}_2^{-1} \end{bmatrix}. \\ &= -2\dot{Y}_3 \bullet \dot{S}_3. \end{aligned}$$

By the same way, using the first two relations in (25) and the first two equations of (22), one readily verifies that

$$\dot{x}^T H \dot{x} = \langle H\dot{x}, \dot{x}\rangle = -\langle \mathcal{A}^*(\dot{Y}), \dot{x}\rangle = -\dot{Y} \bullet \mathcal{A}(\dot{x}) = \dot{Y} \bullet \dot{S} = 2\dot{Y}_3 \bullet \dot{S}_3.$$

Consequently

$$\dot{x}^T (H + \mathcal{H}(\bar{x}, \bar{Y}))\dot{x} = 0. \tag{27}$$

This implies that $\dot{x} = 0$, since $\dot{x} \in C(\bar{x})$. Using Remark 1 it follows also that $\dot{Y}_3 = \dot{S}_3 = 0$.

By the first equation of (22), we obtain

$$\dot{S} = -\mathcal{A}(\dot{x}) = -\mathcal{A}(0) = 0. \tag{28}$$

Thus, it only remains to show that $\dot{Y} = 0$. In view of (26) we have

$$\check{Y} = \begin{bmatrix} \dot{Y}_1 & 0 \\ 0 & 0 \end{bmatrix}. \tag{29}$$

Now suppose that $\dot{Y}_1 \neq 0$. Since $\bar{Y}_1 \succ 0$, it is clear that there exists some $\bar{\tau} > 0$ such that $\bar{Y}_1 + \tau\dot{Y}_1 \succ 0 \quad \forall \tau \in (0, \bar{\tau}]$. If we define $\bar{Y}_\tau := \bar{Y} + \tau\dot{Y}$ it follows that

$$\bar{Y}_\tau = U(\tilde{Y} + \tau\check{Y})U^T \succeq 0 \wedge \bar{Y}_\tau \neq \bar{Y}, \quad \forall \tau \in (0, \bar{\tau}].$$

Moreover, using (20), (22), and (28), one readily verifies that $(\bar{x}, \bar{Y}_\tau, \bar{S})$ also satisfies (20) for all $\tau \in (0, \bar{\tau}]$. This contradicts the assumption that $(\bar{x}, \bar{Y}, \bar{S})$ is a locally unique stationary point. Hence $\dot{Y}_1 = 0$ and by (29), $\dot{Y} = 0$.

We can now apply the implicit function theorem to the system

$$\begin{aligned} \mathcal{A}(x) + C + S &= 0, \\ Hx + b + \mathcal{A}^*(Y) &= 0, \\ YS + SY &= 0. \end{aligned} \tag{30}$$

As we have just seen, the linearization of (30) at the point $(\bar{x}, \bar{Y}, \bar{S})$ is nonsingular. Therefore the system (30) has a differentiable and locally unique solution $(\bar{x}(\triangle\mathscr{D}), \bar{Y}(\triangle\mathscr{D}), \bar{S}(\triangle\mathscr{D}))$. By the continuity of $\bar{Y}(\triangle\mathscr{D}), \bar{S}(\triangle\mathscr{D})$ with respect to $\triangle\mathscr{D}$ it follows that for $\| \triangle\mathscr{D} \|$ sufficiently small, $\bar{Y}(\triangle\mathscr{D}) + \bar{S}(\triangle\mathscr{D}) \succ 0$, i.e. $(\bar{Y}(\triangle\mathscr{D}) + \bar{S}(\triangle\mathscr{D})) \in C$.

Consequently, by part (b) of Lemma 1 we have $\bar{Y}(\triangle\mathscr{D}), \bar{S}(\triangle\mathscr{D}) \succeq 0$. This implies that the local solutions of the system (30) are actually stationary points.

Note that the dimension of the image space of $\bar{S}(\triangle\mathscr{D})$ is constant for all $\| \triangle\mathscr{D} \|$ sufficiently small. According to Remark 2 it holds that $\bar{S}(\triangle\mathscr{D})^\dagger \to \bar{S}^\dagger$ when $\triangle\mathscr{D} \to 0$.

Finally we prove that the second-order sufficient condition is invariant under small perturbations $\triangle\mathscr{D}$ of the problem data $\mathscr{D}$. We just need to show that there exists $\bar{\varepsilon} > 0$ such that for all $\triangle\mathscr{D}$ with $\| \triangle\mathscr{D} \| \leq \bar{\varepsilon}$ it holds:

$$h^T((H+\triangle H)+\mathcal{H}(\bar{x}(\triangle\mathscr{D}), \bar{Y}(\triangle\mathscr{D})))h > 0 \qquad \forall h \in C(\bar{x}(\triangle\mathscr{D}))\backslash\{0\}. \tag{31}$$

Since $C(\bar{x}(\triangle\mathscr{D}))/\{0\}$ is a cone, it suffices to consider unitary vectors, i.e. $\| h \| = 1$. We assume by contradiction that there exists $\varepsilon_k \to 0$, $\{\triangle\mathscr{D}_k\}$ with $\| \triangle\mathscr{D}_k \| \leq \varepsilon_k$, and $\{h_k\}$ with $h_k \in C(\bar{x}(\triangle\mathscr{D}_k)) \setminus \{0\}$ such that

$$h_k^T((H + \triangle H_k) + \mathcal{H}(\bar{x}(\triangle\mathscr{D}), \bar{Y}(\triangle\mathscr{D}_k)))h_k \leq 0. \tag{32}$$

We may assume that h_k converges to h with $\| h \| = 1$, when $k \to \infty$. Since $\triangle\mathscr{D}_k \to 0$, we obtain from the already mentioned convergence $\bar{S}(\triangle\mathscr{D})^\dagger \to \bar{S}^\dagger$ and simple continuity arguments that:

$$0 < h^T Hh + h^T \mathcal{H}(\bar{x}, \bar{Y})h \leq 0. \tag{33}$$

The left inequality of this contradiction follows from the second order sufficient condition since $h \in C(\bar{x}(0)) \setminus \{0\}$ due to Remark 3. $\square$

4 Conclusions

The sensitivity result of Theorem 1 was used in [7] to establish local quadratic convergence of the SSP method. By extending this result to the weaker form of second order sufficient condition, the analysis in [7] can be applied in a straightforward way to this more general class of nonlinear semidefinite programs. In fact, the analysis in [7] only used local Lipschitz continuity of the solution with respect to small changes of the data, which is obviously implied by the differentiability established in Theorem 1.

Acknowledgements. The authors thank to an anonymous referee for the helpful and constructive suggestions. This work was supported by Conicyt Chile under the grants Fondecyt Nr. 1050696 and Nr. 7070326.

REFERENCES

[1] F. Bonnans and H.C Ramírez, *Strong regularity of semidefinite programming problems.* Informe Técnico, DIM-CMM, Universidad de Chile, Santiago, Departamento de Ingeniería Matemática, No CMM-B-05/06-137 (2005).

[2] F. Bonnans and A. Shapiro, *Perturbation Analysis of Optimization Problems*. Springer Series in Operations Research (2000).

[3] A. Bjork, *Numerical methods for least squares problems.* SIAM, Society for Industrial and Applied Mathematics, Philadelphia (1996).

[4] M. Diehl, F. Jarre and C.H. Vogelbusch, *Loss of Superlinear Convergence for an SQP-type method with Conic Constraints.* SIAM Journal on Optimization, (2006), 1201–1210.

[5] B. Fares, D. Noll and P. Apkarian, *Robust control via sequential semidefinite programming.* SIAM J. Control Optim., **40** (2002), 1791–1820.

[6] R.W. Freund and F. Jarre, *A sensitivity result for semidefinite programs.* Oper. Res. Lett., **32**(2004), 126–132.

[7] R.W. Freund, F. Jarre and C. Vogelbusch, *Nonlinear semidefinite programming: sensitivity, convergence, and an application in passive reduced-order modeling.* Math. Program., 109(2-3) (2007), 581–611.

[8] R. Garcés, W. Gómez and F. Jarre, *A self-concordance property for nonconvex semidefinite programming.* Math. Meth. Oper. Res., **74** (2011), 77–92.

[9] M.V. Nayakkankuppam and M.L. Overton, *Conditioning of semidefinite programs.* Math. Program., **85** (1999), 525–540.

[10] A. Shapiro, *First and second order analysis of nonlinear semidefinite programs.* Math. Program., **77** (1997), 301–320.

[11] J.F. Sturm and S. Zhang, *On sensitivity of central solutions in semidefinite programming.* Math. Program., **90** (2001), 205–227.