1

DUALITY, OPTIMALITY CONDITIONS AND PERTURBATION ANALYSIS

Alexander Shapiro [1]

School of Industrial and Systems Engineering,
Georgia Institute of Technology,
Atlanta, Georgia 30332-0205, USA,
E-mail: ashapiro@isye.gatech.edu

1 INTRODUCTION

Consider the optimization problem

$$\operatorname*{Min}_{x \in C} f(x) \ \text{ subject to } \ G(x) \preceq 0, \tag{1.1}$$

where C is a convex closed cone in the Euclidean space $\mathbb{R}^n$, $f : \mathbb{R}^n \to \mathbb{R}$ and $G : \mathbb{R}^n \to \mathcal{Y}$ is a mapping from $\mathbb{R}^n$ into the space $\mathcal{Y} := \mathcal{S}^m$ of $m \times m$ symmetric matrices. We refer to the above problem as a nonlinear semidefinite programming problem. In particular, if $C = \mathbb{R}^n$, the objective function is linear, i.e. $f(x) := \sum_{i=1}^n b_i x_i$, and the constraint mapping is affine, i.e. $G(x) := A_0 + \sum_{i=1}^n x_i A_i$ where $A_0, A_1, ..., A_n \in \mathcal{S}^m$ are given matrices, problem (1.1) becomes a linear semidefinite programming problem

$$\operatorname*{Min}_{x \in \mathbb{R}^n} \sum_{i=1}^n b_i x_i \ \text{ subject to } \ A_0 + \sum_{i=1}^n x_i A_i \preceq 0. \tag{1.2}$$

In this article we discuss duality, optimality conditions and perturbation analysis of such nonlinear semidefinite programming problems.

Let us observe at this point that problem (1.1) can be formulated in the form

$$\operatorname*{Min}_{x \in C} f(x) \ \text{ subject to } \ G(x) \in K, \tag{1.3}$$

where $K := \mathcal{S}^m_-$ is the cone of negative semidefinite $m \times m$ symmetric matrices. That is, the feasible set of problem (1.1) can be defined by the "cone constraints" $\{x \in C : G(x) \in K\}$. Some of the results presented in this article can be formulated in the general framework of such "cone constrained" problems, while the others use a particular structure of the considered cone $\mathcal{S}^m_-$. Therefore we start our analysis by considering an optimization problem in the form (1.3) with $\mathcal{Y}$ being a finite dimensional vector space and $K \subset \mathcal{Y}$ being a convex closed cone. Note that if $C = \mathbb{R}^n$ and the cone K is given by the negative orthant $\mathbb{R}^p_- := \{y \in \mathbb{R}^p : y_i \le 0,\ i = 1, ..., p\}$, then problem (1.3) becomes a standard (nonlinear) programming problem. As we shall see there are certain similarities between such nonlinear programming problems and semidefinite programming problems. There are also, however, some essential differences.

We assume that spaces $\mathbb{R}^n$ and $\mathcal{Y}$ are equipped with respective scalar products, denoted by "$\cdot$". In particular, in the Euclidean space $\mathbb{R}^n$ we use the standard scalar product $x \cdot z := x^T z$, and in the space $\mathcal{S}^m$ the scalar product $A \bullet B := \operatorname{trace}(AB)$. With the cone K is associated its polar (negative dual) cone K^-,

$$K^- := \{y \in \mathcal{Y} : y \cdot w \le 0,\ \forall\, w \in K\}.$$

Since the cone K is convex and closed we have the following duality relation $(K^-)^- = K$ (and similarly for the cone C). This is a classical result which can be easily derived from the separation theorem. It can be noted that the polar of the cone $\mathcal{S}^m_-$ is the cone $\mathcal{S}^m_+ := \{A \in \mathcal{S}^m : A \succeq 0\}$ of positive semidefinite matrices.

2 DUALITY

The Lagrangian function, associated with the problem (1.3), can be written in the form

$$L(x, \lambda) := f(x) + \lambda \cdot G(x), \quad (x, \lambda) \in \mathbb{R}^n \times \mathcal{Y}.$$

It follows from the duality relation $(K^-)^- = K$ that

$$\sup_{\lambda \in K^-} \lambda \cdot G(x) = \begin{cases} 0, & if\ G(x) \in K, \\ +\infty, & otherwise. \end{cases}$$

Therefore problem (1.3) can be also written in the form

$$\operatorname*{Min}_{x \in C} \left\{ \sup_{\lambda \in K^-} L(x, \lambda) \right\}. \tag{2.1}$$

By formally interchanging the "min" and "max" operators in (2.1) we obtain the following problem

$$\operatorname*{Max}_{\lambda \in K^-} \left\{ \inf_{x \in C} L(x, \lambda) \right\}. \tag{2.2}$$

We refer to (1.3) as the *primal* (P) and to (2.2) as its *dual* (D) problems, and denote their optimal values by $\mathrm{val}(P)$ and $\mathrm{val}(D)$ and their sets of optimal solutions by $\mathrm{Sol}(P)$ and $\mathrm{Sol}(D)$, respectively. In particular, the dual of the semidefinite problem (1.1) can be written in the form

$$\operatorname*{Max}_{\Omega \succeq 0} \left\{ \inf_{x \in C} L(x, \Omega) \right\}, \tag{2.3}$$

where $\Omega \in \mathcal{S}^m$ denotes the dual variable. Note that the optimal value of a minimization (maximization) problem with an empty feasible set is defined to be $+\infty$ $(-\infty)$.

In case the objective function is linear and the constraint mapping is affine the dual problem can be written explicitly. For example, in the case of the linear semidefinite programming problem (1.2) we have

$$\inf_{x \in \mathbb{R}^n} L(x, \Omega) = \begin{cases} \Omega \bullet A_0, & \text{if } b_i + \Omega \bullet A_i = 0,\ i = 1, ..., n, \\ -\infty, & \text{otherwise.} \end{cases}$$

Therefore the dual of (1.2) is

$$\operatorname*{Max}_{\Omega \succeq 0} \Omega \bullet A_0 \ \text{ subject to } \ \Omega \bullet A_i + b_i = 0,\ i = 1, ..., n. \tag{2.4}$$

Note that we can consider the above problem as a particular case of the cone constrained problem (1.3) by defining $f(\Omega) := \Omega \bullet A_0$, $C := \mathcal{S}^m_+$, $K := \{0\} \subset \mathbb{R}^n$ and $G(\Omega) := (\Omega \bullet A_1 + b_1, ..., \Omega \bullet A_n + b_n)$. Its dual then coincides with the primal problem (1.2). Therefore there is a complete symmetry between the dual pair (1.2) and (2.4), and which one is called primal and which is dual is somewhat arbitrary.

We say that $(\bar{x}, \bar{\lambda})$ is a *saddle point* of the Lagrangian $L(x, \lambda)$ if

$$\bar{x} = \arg\min_{x \in C} L(x, \bar{\lambda}) \ \textit{ and } \ \bar{\lambda} = \arg\max_{\lambda \in K^-} L(\bar{x}, \lambda). \tag{2.5}$$

Recall that the supremum of $\lambda \cdot G(\bar{x})$ over all $\lambda \in K^-$ equals 0 if $G(\bar{x}) \in K$ and is $+\infty$ otherwise. Therefore the second condition in (2.5) means that $G(\bar{x}) \in K$, $\bar{\lambda} \in K^-$ and the following, so-called *complementarity*, condition holds $\bar{\lambda} \cdot G(\bar{x}) = 0$. It follows that conditions (2.5) are equivalent to:

$$\bar{x} = \arg\min_{x \in C} L(x, \bar{\lambda}), \ \ \bar{\lambda} \cdot G(\bar{x}) = 0, \ \ G(\bar{x}) \in K, \ \ \bar{\lambda} \in K^-. \tag{2.6}$$

In particular, in the case of semidefinite programming problem (1.1) these conditions become

$$\bar{x} = \arg\min_{x \in C} L(x, \bar{\Omega}), \ \ \bar{\Omega} \bullet G(\bar{x}) = 0, \ \ G(\bar{x}) \preceq 0, \ \ \bar{\Omega} \succeq 0. \tag{2.7}$$

The following proposition is an easy consequence of the min-max representations of the primal and dual problems. It can be applied to the primal (1.1) and its dual (2.3) semidefinite programming problems in a straightforward way.

Proposition 2.1 *Let (P) and (D) be the primal and dual problems (1.3) and (2.2), respectively. Then* $\mathrm{val}(D) \le \mathrm{val}(P)$. *Moreover,* $\mathrm{val}(P) = \mathrm{val}(D)$ *and $\bar{x}$ and $\bar{\lambda}$ are optimal solutions of (P) and (D), respectively, if and only if $(\bar{x}, \bar{\lambda})$ is a saddle point of the Lagrangian $L(x, \lambda)$.*

Proof. For any $(x', \lambda') \in C \times K^-$ we have

$$\inf_{x \in C} L(x, \lambda') \le L(x', \lambda') \le \sup_{\lambda \in K^-} L(x', \lambda), \tag{2.8}$$

and hence

$$\sup_{\lambda \in K^-} \inf_{x \in C} L(x, \lambda) \le \inf_{x \in C} \sup_{\lambda \in K^-} L(x, \lambda). \tag{2.9}$$

It follows then from the min-max representations (2.1) and (2.2), of the primal and dual problems respectively, that $\mathrm{val}(D) \le \mathrm{val}(P)$.

Now if $(\bar{x}, \bar{\lambda})$ is a saddle point, then it follows from (2.5) that

$$\inf_{x \in C} L(x, \bar{\lambda}) = L(\bar{x}, \bar{\lambda}) = \sup_{\lambda \in K^-} L(\bar{x}, \lambda), \tag{2.10}$$

and hence

$$\sup_{\lambda \in K^-} \inf_{x \in C} L(x, \lambda) \ge L(\bar{x}, \bar{\lambda}) \ge \inf_{x \in C} \sup_{\lambda \in K^-} L(x, \lambda). \tag{2.11}$$

Inequalities (2.9) and (2.11) imply that $\mathrm{val}(P) = \mathrm{val}(D)$ and $\bar{x}$ and $\bar{\lambda}$ are optimal solutions of problems (P) and (D), respectively.

Conversely, if $\mathrm{val}(P) = \mathrm{val}(D)$ and $\bar{x}$ and $\bar{\lambda}$ are optimal solutions of problems (P) and (D), respectively, then (2.10) follows from (2.8), and hence $(\bar{x}, \bar{\lambda})$ is a saddle point. ∎

By the above proposition the optimal value of the primal problem is always greater than or equal to the optimal value of the dual problem. The difference $\mathrm{val}(P) - \mathrm{val}(D)$ is called the *duality gap* between the primal and dual problems. It is said that there is no duality gap between the primal and dual problems, if val$(P) = $ val(D). It is well known that in the case of linear programming (i.e. if $f(x)$ is linear, $G(x)$ is affine, $C = \mathbb{R}^n$ and K is a convex polyhedral cone), there is no duality gap between the primal and dual problems, provided the feasible set of the primal or dual problem is nonempty. Also a linear programming problem always possesses an optimal solution provided its optimal value is finite. As the following examples show these properties do not always hold for linear semidefinite programming problems.

Example 2.1 Consider the following linear semidefinite programming problem

$$\text{Min } x_1 \text{ subject to } \begin{bmatrix} -x_1 & 1 \\ 1 & -x_2 \end{bmatrix} \preceq 0. \tag{2.12}$$

The feasible set of this problem is $\{(x_1, x_2) : x_1 \geq 0,\ x_2 \geq 0,\ x_1 x_2 \geq 1\}$, and hence its optimal value is 0 and this problem does not have an optimal solution. The dual of this problem is

$$\text{Max } 2\omega_{12} \text{ subject to } \begin{bmatrix} 1 & \omega_{12} \\ \omega_{21} & 0 \end{bmatrix} \succeq 0. \tag{2.13}$$

Its feasible set contains one point with $\omega_{12} = \omega_{21} = 0$, which is also its optimal solution, and hence its optimal value is 0. Therefore in this example there is no duality gap between the primal and dual problems, although the primal problem does not have an optimal solution.

Example 2.2 Consider the linear semidefinite programming problem

$$\text{Min } -x_2 \text{ subject to } \begin{bmatrix} x_2 - a & 0 & 0 \\ 0 & x_1 & x_2 \\ 0 & x_2 & 0 \end{bmatrix} \preceq 0, \tag{2.14}$$

where $a > 0$ is a given number. The dual of this problem is

$$\text{Max } -a\omega_{11} \text{ subject to } \Omega \succeq 0,\ \omega_{22} = 0,\ \omega_{11} + 2\omega_{23} = 1. \tag{2.15}$$

The feasible set of the primal problem is $\{(x_1, x_2) : x_1 \leq 0,\ x_2 = 0\}$ and hence its optimal value is 0. On the other hand any feasible Ω of the dual problem has $\omega_{11} = 1$, and hence the optimal value of the dual problem is $-a$. Therefore the duality gap in this example is a.

In case of linear semidefinite programming problem (1.2) and its dual (2.4), a pair $(\bar{x}, \bar{\Omega})$ is a saddle point iff $\bar{x}$ and $\bar{\Omega}$ are feasible points of (1.2) and (2.4), respectively, and the following complementarity condition

$$\bar{\Omega}\left(A_0+\sum_{i=1}^{n}\bar{x}_iA_i\right)=0 \tag{2.16}$$

holds. The above condition corresponds to the complementarity condition of (2.6), while feasibility of $\bar{\Omega}$ is equivalent to the first and forth conditions of (2.6) and feasibility of $\bar{x}$ is the third condition of (2.6). Note that since $\bar{\Omega}\succeq 0$ and $A_0+\sum_{i=1}^{n}\bar{x}_iA_i\preceq 0$, the complementarity condition (2.16) is equivalent to $\bar{\Omega}\bullet(A_0+\sum_{i=1}^{n}\bar{x}_iA_i)=0$.

It is not always easy to verify existence of a saddle point, and moreover "no gap" property can hold even if the primal or/and dual problems do not have optimal solutions. We approach now the "no duality gap" problem from a somewhat different point of view. With the primal problem we associate the following parametric problem

$$\operatorname*{Min}_{x\in C} f(x)\ \text{ subject to }\ G(x)+y\in K, \tag{2.17}$$

depending on the parameter vector $y\in\mathcal{Y}$. We denote this problem by (P_y) and by $v(y)$ we denote its optimal value, i.e. $v(y):=\operatorname{val}(P_y)$. Clearly problem (P_0) coincides with the primal problem (P) and $\operatorname{val}(P)=v(0)$. The conjugate of the function $v(y)$ is defined as

$$v^*(y^*):=\sup_{y\in\mathcal{Y}}\{y^*\cdot y-v(y)\}.$$

We have then that

$$\begin{array}{rcl} v^*(y^*) & = & \sup\{y^*\cdot y-f(x):(x,y)\in C\times\mathcal{Y},\ G(x)+y\in K\}\\ & = & \sup_{x\in C}\sup_{y\in -G(x)+K}\{y^*\cdot y-f(x)\}. \end{array}$$

Now

$$\sup_{y\in -G(x)+K} y^*\cdot y=-y^*\cdot G(x)+\sup_{y\in K} y^*\cdot y,$$

and hence the above supremum equals $-y^*\cdot G(x)$ if $y^*\in K^-$ and is $+\infty$ otherwise.

We obtain that

$$-v^*(y^*)=\left\{\begin{array}{cl} \inf_{x\in C}L(x,y^*), & if\ y^*\in K^-,\\ -\infty, & otherwise. \end{array}\right. \tag{2.18}$$

Therefore the dual problem (2.2) can be formulated as

$$\operatorname*{Max}_{\lambda\in\mathcal{Y}}\{-v^*(\lambda)\}. \tag{2.19}$$

Consequently we have that $\mathrm{val}(D) = v^{**}(0)$, where

$$v^{**}(y) := \sup_{y^* \in \mathcal{Y}} \{y^* \cdot y - v^*(y^*)\}$$

denotes the conjugate of the conjugate function $v^*(\cdot)$. The relations $\mathrm{val}(P) = v(0)$ and $\mathrm{val}(D) = v^{**}(0)$, between the primal and dual problems and the optimal value function $v(\cdot)$, allow to apply powerful tools of convex analysis.

In the subsequent analysis we deal with convex cases where the optimal value function $v(y)$ is convex. By the Fenchel-Moreau duality theorem we have that if $v(y)$ is *convex*, then

$$v^{**}(\cdot) = \begin{cases} \mathrm{lsc}\, v(\cdot), & if \;\; \mathrm{lsc}\, v(y) > -\infty \; for \; all \; y \in \mathcal{Y}, \\ -\infty, & otherwise, \end{cases} \tag{2.20}$$

where

$$\mathrm{lsc}\, v(y) := \min \left\{ v(y), \liminf_{y' \to y} v(y') \right\}$$

denotes the lower semicontinuous hull of v, i.e. the largest lower semicontinuous function majorized by v.

The quantity

$$\mathrm{lsc}\, v(0) = \min \left\{ v(0), \liminf_{y \to 0} v(y) \right\}$$

is called the *subvalue* of the problem (P). The problem (P) is said to be *subconsistent* if its subvalue is less than $+\infty$, i.e. either if the feasible set of (P) is nonempty, and hence its optimal value is less than $+\infty$, or if by making "small" perturbations in the feasible set of (P) the corresponding optimal value function $v(y)$ becomes bounded from above. Note that if the lower semicontinuous hull of a convex function has value $-\infty$ in at least one point, then it can take only two possible values $+\infty$ or $-\infty$. Therefore if $\mathrm{lsc}\, v(0)$ is finite, then $v^{**}(0) = \mathrm{lsc}\, v(0)$. By the above discussion we obtain the following result.

Theorem 2.3 *Suppose that the optimal value function $v(y)$ is convex and that the primal problem (P) is subconsistent. Then*

$$\mathrm{val}(D) = \min \left\{ \mathrm{val}(P), \liminf_{y \to 0} v(y) \right\}, \tag{2.21}$$

and hence $\mathrm{val}(P) = \mathrm{val}(D)$ *if and only if* $v(y)$ *is lower semicontinuous at* $y = 0$.

We discuss now conditions ensuring convexity of $v(y)$. Let $Q \subset Y$ be a closed convex cone. We say that the mapping $G : \mathbb{R}^n \to Y$ is *convex* with respect to Q if for any $x_1, x_2 \in \mathbb{R}^n$ and $t \in [0,1]$ the following holds

$$tG(x_1) + (1-t)G(x_2) \succeq_Q G(tx_1 + (1-t)x_2). \tag{2.22}$$

Here " $\succeq_Q$ " denotes the partial ordering associated with the cone Q, i.e. $a \succeq_Q b$ if $a - b \in Q$. We say that the problem (P) is convex, if the function $f(x)$ is convex and the mapping $G(x)$ is convex with respect to the cone $Q := -K$. Any affine mapping $G(x)$ is convex with respect to any cone Q, and hence any linear problem is convex.

Let us observe that if (P) is convex and $\lambda \in K^-$, then the function $L(\cdot, \lambda)$ is convex. Indeed, Condition (2.22) means that

$$tG(x_1) + (1-t)G(x_2) - G(tx_1 + (1-t)x_2) \in -K,$$

and hence for any $\lambda \in K^-$,

$$t\lambda \cdot G(x_1) + (1-t)\lambda \cdot G(x_2) - \lambda \cdot G(tx_1 + (1-t)x_2) \geq 0.$$

It follows that $\lambda \cdot G(\cdot)$ is convex, and since $f(x)$ is convex, we obtain that $L(\cdot, \lambda)$ is convex. Convexity of (P) also implies convexity of $v(\cdot)$.

Proposition 2.2 *Suppose that the primal problem (P) is convex. Then the optimal value function $v(y)$ is convex.*

Proof. Consider the following extended real valued function

$$\varphi(x,y) := \begin{cases} f(x), & \textit{if } x \in C \textit{ and } G(x) + y \in K, \\ +\infty, & \textit{otherwise}. \end{cases}$$

Note that $v(y) = \inf_{x \in \mathbb{R}^n} \varphi(x,y)$. Let us show that the function $\varphi(\cdot,\cdot)$ is convex. Since $f(x)$ is convex, it suffices to show that the domain $\operatorname{dom}\varphi := \{(x,y) \in C \times \mathcal{Y} : \varphi(x,y) < +\infty\}$ of φ is a convex set. Consider $(x_1,y_1),(x_2,y_2) \in \operatorname{dom}\varphi$ and $t \in [0,1]$. By convexity of C, $tx_1 + (1-t)x_2 \in C$. By convexity of G, with respect to $-K$, we have that

$$k := G(tx_1 + (1-t)x_2) - [tG(x_1) + (1-t)G(x_2)] \in K.$$

Therefore, since K is a convex cone,

$$G(tx_1 + (1-t)x_2) + ty_1 + (1-t)y_2 = k + t[G(x_1) + y_1] + (1-t)[G(x_2) + y_2] \in K.$$

It follows that $t(x_1, y_1) + (1-t)(x_2, y_2) \in \operatorname{dom}\varphi$, and hence $\operatorname{dom}\varphi$ is convex.

Now, since φ is convex, we have for any $(x_1, y_1), (x_2, y_2) \in \mathbb{R}^n \times \mathcal{Y}$ and $t \in [0, 1]$,

$$t\varphi(x_1, y_1) + (1-t)\varphi(x_2, y_2) \geq \varphi(tx_1 + (1-t)x_2, ty_1 + (1-t)y_2) \geq v(ty_1 + (1-t)y_2).$$

By minimizing the left hand side of the above inequality over x_1 and x_2, we obtain

$$tv(y_1) + (1-t)v(y_2) \geq v(ty_1 + (1-t)y_2),$$

which shows that $v(y)$ is convex. ∎

In the case of semidefinite programming, i.e. when $K := \mathcal{S}^m_-$ and hence $Q := -K = \mathcal{S}^m_+$, convexity of $G(x)$ means that it is convex with respect to the Löwner partial order in the space $\mathcal{S}^m$. The inequality (2.22) is then equivalent to

$$tz^T G(x_1)z + (1-t)z^T G(x_2)z \geq z^T G(tx_1 + (1-t)x_2)z, \quad \forall\, z \in \Re^m. \tag{2.23}$$

This, in turn, means that the function $\psi(x) := z^T G(x)z$ is convex for any $z \in \Re^m$.

By theorem 2.3 we have that in the convex case, and if (P) is subconsistent, there is no duality gap between the primal and dual problems iff $v(y)$ is lower semicontinuous at $y = 0$. This is a topological condition which may be not easy to verify in a particular situation. We derive now more directly verifiable conditions ensuring the no duality gap property. Recall that a vector $z \in \mathcal{Y}$ is said to be a *subgradient* of $v(\cdot)$, at a point y such that $v(y)$ is finite, if

$$v(y') - v(y) \geq z \cdot (y' - y), \quad \forall\, y' \in \mathcal{Y}.$$

The set of all subgradients of $v(\cdot)$, at y, is called the *subdifferential* and denoted $\partial v(y)$. The function $v(\cdot)$ is said to be *subdifferentiable* at y if $v(y)$ is finite and the subdifferential $\partial v(y)$ is nonempty. Results of the following proposition are easy consequences of the definitions.

Proposition 2.3 *Suppose that the function $v(\cdot)$ has a finite value at a point $y \in \mathcal{Y}$. Then:*
(i) If $v(\cdot)$ is subdifferentiable at y, then $v(\cdot)$ is lower semicontinuous at y.
(ii)

$$\partial v^{**}(y) = \arg\max_{y^* \in Y} \{y^* \cdot y - v^*(y^*)\}. \tag{2.24}$$

(iii) If v is subdifferentiable at y, then $v^{**}(y) = v(y)$.
(iv) If $v^{**}(y) = v(y)$, then $\partial v^{**}(y) = \partial v(y)$.

Together with representation (2.19) of the dual problem, the above proposition implies the following results. Recall that $\mathrm{Sol}(D)$ denotes the set of optimal solutions of the dual problem (D).

Theorem 2.4 *Suppose that the primal problem (P) is convex. Then:*
(i) If $\mathrm{val}(D)$ is finite, then $\mathrm{Sol}(D) = \partial v^{**}(0)$.
(ii) If $v(y)$ is subdifferentiable at $y = 0$, then there is no duality gap between the primal and dual problems and $\mathrm{Sol}(D) = \partial v(0)$.
(iii) If $\mathrm{val}(P) = \mathrm{val}(D)$ and is finite, then the (possibly empty) set $\mathrm{Sol}(D)$, of optimal solutions of the dual problem, coincides with $\partial v(0)$.

Proof. Assertion (i) follows from representation (2.19) of the dual problem and formula (2.24) applied at $y = 0$.

If $v(y)$ is subdifferentiable at $y = 0$, then $v^{**}(0) = v(0)$ and $\partial v^{**}(y) = \partial v(y)$. Consequently, by theorem 2.3, we obtain that $\mathrm{val}(P) = \mathrm{val}(D)$. Together with (i) this proves assertion (ii).

If $\mathrm{val}(P) = \mathrm{val}(D)$ and is finite, then $v^{**}(y) = v(y)$ and hence $\partial v^{**}(y) = \partial v(y)$. Together with (i) this proves (iii). ∎

Suppose now that $v(\cdot)$ is convex and that $v(y) < +\infty$ for all y in a neighborhood of zero. By convex analysis this implies that $v(y)$ is continuous at $y = 0$ (it still can happen that $v(y) = -\infty$ for all y in a neighborhood of zero). Moreover, if in addition $v(0)$ is finite, then $v(y)$ is subdifferentiable at $y = 0$ and $\partial v(0)$ is bounded. Note that $v(y) < +\infty$ iff the feasible set of (P_y) is nonempty, which means that there exists $x \in C$ such that $G(x)+y \in K$. Therefore the condition that $v(y) < +\infty$ for all y in a neighborhood of zero, can be written in the form

$$0 \in \mathrm{int}\{G(C) - K\}, \tag{2.25}$$

where $G(C)$ denotes the set of points $G(x)$, $x \in C$, in the space $\mathcal{Y}$, and "int" stands for the interior of the corresponding set. We obtain the following result.

Theorem 2.5 *Suppose that the primal problem (P) is convex and that the regularity condition (2.25) holds. Then there is no duality gap between the primal and dual problems and, moreover, if their common optimal value is finite, then the set $\mathrm{Sol}(D)$, of optimal solutions of the dual problem, is nonempty and bounded.*

It is interesting that in a sense a converse of the above result also holds. By theorem 2.2(i) we have that if the set of optimal solutions of the dual problem is nonempty (and hence $\mathrm{val}(D)$ is finite) and bounded, then $\partial v^{**}(0)$ is nonempty and bounded. By convexity of $v(\cdot)$ this implies that $v(y)$ is continuous at $y = 0$, and hence (2.25) follows. Therefore we obtain the following result.

Theorem 2.6 *Suppose that the primal problem (P) is convex and that the dual problem has a nonempty and bounded set of optimal solutions. Then condition (2.25) holds and there is no duality gap between the primal and dual problems.*

We say that the *Slater condition* holds, for the primal problem, if there exists a point $\bar{x} \in C$ such that $G(\bar{x}) \in \mathrm{int}(K)$. In the case of semidefinite programming, i.e. when $K = \mathcal{S}^m_-$, this means that $G(\bar{x}) \prec 0$, i.e. that the matrix $G(\bar{x})$ is negative definite. It is clear that the Slater condition implies the regularity condition (2.25). Converse of that is also true if the cone K has a nonempty interior, and hence in particular in the case of convex semidefinite programming problems.

Proposition 2.4 *Suppose that the mapping $G(x)$ is convex with respect to the cone $Q := -K$ and that K has a nonempty interior. Then (2.25) is equivalent to the Slater condition.*

Proof. It is clear that Slater condition implies (2.25). We proof now that the converse implication also holds. Suppose that (2.25) holds. Let x be a feasible point of (P), i.e. $x \in C$ and $G(x) \in K$. It is clear that existence of such a point follows from (2.25). Let $\bar{y}$ be an interior point of K, i.e. for some neighborhood $N \subset \mathcal{Y}$ of zero the inclusion $\bar{y} + N \subset K$ holds. It follows from (2.25) that for sufficiently small $\alpha > 0$, there is a point $x' \in C$ such that $\alpha(\bar{y} - G(x)) \in G(x') - K$. Consider $k_1 := G(x') - \alpha(\bar{y} - G(x)) \in K$ and let $\bar{x} := (x + x')/2$. By convexity of C, $\bar{x} \in C$, and by convexity of G we have

$$k_2 := -\tfrac{1}{2}G(x) - \tfrac{1}{2}G(x') + G(\bar{x}) \in K.$$

It follows that

$$G(\bar{x}) = \tfrac{1}{2}k_1 + k_2 + \tfrac{1}{2}(1-\alpha)G(x) + \tfrac{1}{2}\alpha\bar{y} = k + \tfrac{1}{2}\alpha\bar{y},$$

where $k := \frac{1}{2}k_1 + k_2 + \frac{1}{2}(1-\alpha)G(x) \in K$ for $\alpha < 1$. Consequently

$$G(\bar{x}) + \tfrac{1}{2}\alpha N = k + \tfrac{1}{2}\alpha(\bar{y} + N) \subset K,$$

and hence $G(\bar{x}) \in \text{int}(K)$, which completes the proof. ∎

All the above results can be applied to convex semidefinite programming problems, and in particular to linear semidefinite programming problems. Consider, for instance, a linear semidefinite programming problem in the form (1.2) and its dual (2.4). If the matrix A_0 is negative definite, then the Slater condition holds (with $\bar{x} = 0$ for example) and hence in that case there is no duality gap, and moreover the dual problem has a nonempty and bounded set of optimal solutions. Suppose now that A_0 is positive definite and the optimal value of the dual problem is finite. Consider the set $\{\Omega \in \mathcal{S}^m_+ : \Omega \bullet A_0 = \text{val}(D)\}$. Since A_0 is positive definite this set is bounded and hence the set of optimal solutions of (D) is nonempty and bounded. We obtain then that the Slater condition, for the primal problem, holds and there is no duality gap between the primal and dual problems.

We can also apply such results to the dual linear semidefinite programming problem (2.4) as well. The regularity condition (2.25), for the problem (2.4), takes the form

$$0 \in \text{int}\{x \in \mathbb{R}^n : x_i = \Omega \bullet A_i + b_i,\ i = 1, ..., n,\ \Omega \succeq 0\}. \tag{2.26}$$

If the above condition holds, then there is no duality gap between the primal and dual problems and the primal problem (1.2) has a nonempty and bounded set of optimal solutions. Conversely, if primal problem (1.2) has a nonempty and bounded set of optimal solutions, then condition (2.26) holds and there is no duality gap between the primal and dual problems. In particular, condition (2.26) holds if there exists a positive definite matrix Ω such that $\Omega \bullet A_i + b_i = 0$, $i = 1, ..., n$.

3 OPTIMALITY CONDITIONS

In this section we discuss first and second order optimality conditions for the semidefinite programming problem (1.1). Again we adopt here an approach of considering a general "cone constrained" problem in the form (1.3), and then specifying the obtained results to the semidefinite programming setting. In order to simplify the presentation we assume in the remainder of this article that the cone C *coincides* with the space $\mathbb{R}^n$, that is $C = \mathbb{R}^n$. We also assume from now on that the function $f(x)$ and mapping $G(x)$ are sufficiently smooth, at least are continuously differentiable.

Suppose that the primal problem (P) is convex. Then, as we know, the function $L(\cdot,\lambda)$ is convex for any $\lambda \in K^-$, and hence $\bar{x}$ is a minimizer of $L(\cdot,\lambda)$, over $\mathbb{R}^n$, iff $\nabla_x L(\bar{x},\lambda) = 0$. Therefore, for convex problems, conditions (2.6) can be written as

$$\nabla_x L(\bar{x},\bar{\lambda}) = 0, \ \ \bar{\lambda}\cdot G(\bar{x}) = 0, \ \ G(\bar{x}) \in K, \ \ \bar{\lambda} \in K^-. \tag{3.1}$$

In particular, for the semidefinite programming problem (1.1) the above conditions take the form

$$\nabla_x L(\bar{x},\bar{\Omega}) = 0, \ \ \bar{\Omega} \bullet G(\bar{x}) = 0, \ \ G(\bar{x}) \preceq 0, \ \ \bar{\Omega} \succeq 0. \tag{3.2}$$

Moreover, for the linear semidefinite programming problem (1.2) these conditions are reduced to feasibility of $A_0 + \sum_{i=1}^n \bar{x}_i A_i$ and $\bar{\Omega}$, considered as points of the respective primal (1.2) and dual (2.4) problems, and the complementarity condition (2.16).

Conditions (3.1) can be viewed as *first order* optimality conditions for the problem (1.3). A point $\bar{x} \in X$ is said to be a *stationary* point of the problem (P) if there exists $\bar{\lambda} \in Y$ such that conditions (3.1) hold. We refer to $\bar{\lambda}$, satisfying (3.1), as a *Lagrange multiplier* vector (in the case of semidefinite programming we refer to $\bar{\Omega}$ as a Lagrange multiplier matrix), and denote by $\Lambda(\bar{x})$ the set of all Lagrange multiplier vectors. In the convex case conditions (3.1) ensure that the point $(\bar{x},\bar{\lambda})$ is a saddle point of the Lagrangian, and hence that $\bar{x}$ is an optimal solution of (P). That is, in the convex case conditions (3.1) are *sufficient* for optimality. However, even in the case of linear semidefinite programming, these conditions are not necessary and in order to ensure their necessity a *constraint qualification* is required.

Theorem 3.1 *Suppose that the primal problem (P) is convex and let $\bar{x}$ be its optimal solution. If the constraint qualification (2.25) holds, then the set $\Lambda(\bar{x})$, of Lagrange multiplier vectors, is nonempty and bounded and is the same for any optimal solution of (P). Conversely, if $\Lambda(\bar{x})$ is nonempty and bounded, then (2.25) holds.*

Proof. If condition (2.25) holds, then by theorem 2.5 we have that there is no duality gap between the primal and dual problems, and moreover $\Lambda(\bar{x})$ coincides with the set of optimal solutions of the dual problem and is nonempty and bounded. The converse assertion follows from theorem 2.6. ∎

Recall that by proposition 2.4, in the convex case, constraint qualification (2.25) is equivalent to the Slater condition provided cone K has a nonempty interior.

In particular, in the case of convex semidefinite programming we have that the set of Lagrange multiplier matrices is nonempty and bounded iff the Slater condition holds.

Let us consider now a possibly nonconvex optimization problem (P) of the form (1.3) (with $C = \mathbb{R}^n$), and let $\bar{x}$ be a feasible point of (P). We can derive then first order optimality conditions by considering the following linearization of (P) at $\bar{x}$:

$$\operatorname*{Min}_{h \in \mathbb{R}^n} Df(\bar{x})h \ \text{ subject to } \ DG(\bar{x})h \in T_K(G(\bar{x})). \tag{3.3}$$

We denote here by $Df(\bar{x})$ the differential of $f(\cdot)$ at $\bar{x}$, i.e. $Df(\bar{x})h$ is a linear function of h given by $Df(\bar{x})h = h \cdot \nabla f(\bar{x})$, and similarly for the differential of the mapping $G(\cdot)$, $DG(\bar{x})h = \sum_{i=1}^n h_i G_i(\bar{x})$ where $G_i(\bar{x}) := \partial G(\bar{x})/\partial x_i$. By $T_K(y)$ we denote the tangent cone to K at a point $y \in K$, that is

$$T_K(y) := \{z \in \mathcal{Y} : \operatorname{dist}(y + tz, K) = o(t),\ t \geq 0\}.$$

It is also possible to linearize the constraint qualification (2.25), at the point $\bar{x}$ and for $C = \mathbb{R}^n$, as follows

$$0 \in \operatorname{int}\{G(\bar{x}) + DG(\bar{x})\mathbb{R}^n - K\}. \tag{3.4}$$

This constraint qualification was introduced by Robinson, and we refer to it as *Robinson constraint qualification.* Since the affine mapping $h \to G(\bar{x}) + DG(\bar{x})h$ is convex, with respect to any cone, we have by proposition 2.4 that if the cone K has a nonempty interior, then (3.4) is equivalent to existence of a vector $\bar{h} \in \mathbb{R}^n$ such that

$$G(\bar{x}) + DG(\bar{x})\bar{h} \in \operatorname{int}(K). \tag{3.5}$$

The above condition can be viewed as an extended Mangasarian-Fromovitz constraint qualification. Recall that if $K := \mathcal{S}^m_-$ is the cone of negative semidefinite matrices, then its interior is formed by negative definite matrices. Therefore in the case of semidefinite programming problem (1.1) the above constraint qualification takes the form: there exists a vector $\bar{h} \in \mathbb{R}^n$ such that

$$G(\bar{x}) + DG(\bar{x})\bar{h} \prec 0. \tag{3.6}$$

It is possible to show that if $\bar{x}$ is a locally optimal solution of (P) and Robinson constraint qualification holds, then $h = 0$ is an optimal solution of the linearized problem (3.3). Proof of that is based on the following stability result. Suppose that the constraint qualification (3.4) holds. Then there exists $\kappa > 0$ such that for for all x in a neighborhood of $\bar{x}$ one has

$$\text{dist}(x, G^{-1}(K)) \le \kappa \, \text{dist}(G(x), K), \tag{3.7}$$

where $G^{-1}(K) := \{x : G(x) \in K\}$ is the feasible set of the problem (P). It follows from the Taylor expansion $G(\bar{x}+th) = G(\bar{x}) + tDG(\bar{x})h + o(t)$ and (3.7) that

$$T_{G^{-1}(K)}(\bar{x}) = \{h : DG(\bar{x})h \in T_K(G(\bar{x}))\},$$

i.e. the feasible set of the linearized problem (3.3) coincides with the tangent cone to the feasible set of the problem (P) at the point $\bar{x}$. By the linearization of the objective function $f(\cdot)$ at $\bar{x}$, we obtain then that it follows from local optimality of $\bar{x}$ that $Df(\bar{x})h$ should be nonnegative for any $h \in T_{G^{-1}(K)}(\bar{x})$, and hence $h = 0$ should be an optimal solution of (3.3).

Since the linearized problem (3.3) is convex, we can apply to it the corresponding first order optimality conditions, at $h = 0$. Note that the polar of the tangent cone $T_K(G(\bar{x}))$ is given by

$$[T_K(G(\bar{x}))]^- = \{\lambda \in K^- : \lambda \cdot G(\bar{x}) = 0\}.$$

Therefore the first order optimality conditions for the linearized problem are equivalent to conditions (3.1) for the problem (P) at the point $\bar{x}$. By theorem 3.1 we have then the following result.

Theorem 3.2 *Let $\bar{x}$ be a locally optimal solution of the primal problem (P), and suppose that Robinson constraint qualification (3.4) holds. Then the set $\Lambda(\bar{x})$, of Lagrange multiplier vectors, is nonempty and bounded. Conversely, if $\bar{x}$ is a feasible point of (P) and $\Lambda(\bar{x})$ is nonempty and bounded, then (3.4) holds.*

We discuss now *second order* optimality conditions. Since second order conditions are intimately related to a particular structure of the cone $\mathcal{S}^m_-$, we consider here only the case of semidefinite programming problems in the form (1.1), and with $C = \mathbb{R}^n$. In the remainder of this section we assume that $f(x)$ and $G(x)$ are twice continuously differentiable. In order to derive second order optimality conditions the following construction will be useful. Let $A \in \mathcal{S}^m_-$ be a matrix of rank $r < m$. Then there exists a neighborhood $\mathcal{N} \subset \mathcal{S}^m$ of A and a mapping $\Xi : \mathcal{N} \to \mathcal{S}^{m-r}$ such that: (i) $\Xi(\cdot)$ is twice continuously differentiable, (ii) $\Xi(A) = 0$, (iii) $\mathcal{S}^m_- \cap \mathcal{N} = \{X \in \mathcal{N} : \Xi(X) \in \mathcal{S}^{m-r}_-\}$, and (iv)

$D\Xi(A)\mathcal{S}^m = \mathcal{S}^{m-r}$, i.e. the differential $D\Xi(A)$ maps the space $\mathcal{S}^m$ onto the space $\mathcal{S}^{m-r}$.

In order to construct such a mapping we proceed as follows. Denote by $\mathcal{E}(X)$ the eigenspace of $X \in \mathcal{S}^m$ corresponding to its $m-r$ largest eigenvalues, and let $P(X)$ be the orthogonal projection matrix onto $\mathcal{E}(X)$. Let E_0 be a (fixed) $m \times (m-r)$ matrix whose columns are orthonormal and span the space $\mathcal{E}(A)$. Consider $F(X) := P(X)E_0$ and let $U(X)$ be the $m \times (m-r)$ matrix whose columns are obtained by applying the Gram-Schmidt orthonormal procedure to the columns of $F(X)$. It is known that $P(X)$ is twice continuously differentiable (in fact even analytic) function of X in a sufficiently small neighborhood of A. Consequently $F(\cdot)$ and hence $U(\cdot)$ are twice continuously differentiable near A. Also by the above construction, $U(A) = E_0$, the column space of $U(X)$ coincides with $\mathcal{E}(X)$ and $U(X)^T U(X) = I_{m-r}$ for all X near A. Finally define $\Xi(X) := U(X)^T X U(X)$. It is straightforward to verify that this mapping Ξ satisfies the properties (i)-(iv). Note that since $D\Xi(A)$ is onto, Ξ maps a neighborhood of A *onto* a neighborhood of the null matrix in the space $\mathcal{S}^{m-r}$.

Now let $\bar{x}$ be a stationary point of the semidefinite problem (1.1), i.e. the corresponding set $\Lambda(\bar{x})$ of Lagrange multiplier matrices is nonempty. This, of course, implies that $\bar{x}$ is a feasible point of (1.1). Since it is assumed that $C = \mathbb{R}^n$, feasibility of $\bar{x}$ means that $G(\bar{x}) \in \mathcal{S}^m_-$. Let Ξ be a corresponding mapping, from a neighborhood of $G(\bar{x})$ into $\mathcal{S}^{m-r}$ (where r is the rank of $G(\bar{x})$), satisfying the corresponding properties (i)-(iv). Consider the composite mapping $\mathcal{G}(x) := \Xi(G(x))$, from a neighborhood of $\bar{x}$ into $\mathcal{S}^{m-r}$. Note that by the property (ii) of Ξ, we have that $\mathcal{G}(\bar{x}) = 0$. Since Ξ maps a neighborhood of $G(\bar{x})$ *onto* a neighborhood of the null matrix in the space $\mathcal{S}^{m-r}$, we have that for all x near $\bar{x}$ the feasible set of (1.1) can be defined by the constraint $\mathcal{G}(x) \preceq 0$. Therefore problem (1.1) is locally equivalent to the following, so called reduced, problem

$$\operatorname*{Min}_{x \in \mathbb{R}^n} f(x) \text{ subject to } \mathcal{G}(x) \preceq 0. \tag{3.8}$$

It is relatively easy to write second order optimality conditions for the reduced problem (3.8) at the point $\bar{x}$. This is mainly because the "constraint cone" $\mathcal{S}^{m-r}_-$ coincides with its tangent cone at the point $\mathcal{G}(\bar{x}) = 0$.

Let

$$\mathcal{L}(x, \Psi) := f(x) + \Psi \bullet \mathcal{G}(x), \quad \Psi \in \mathcal{S}^{m-r},$$

be the Lagrangian of the problem (3.8). Since $D\Xi(\bar{y})$ is onto and (by the chain rule of differentiation) $D\mathcal{G}(\bar{x}) = D\Xi(\bar{y})DG(\bar{x})$, where $\bar{y} := G(\bar{x})$, we have that $\bar{x}$ is also a stationary point of the problem (3.8), and that Ψ is a Lagrange multiplier matrix of the problem (3.8) iff $\Omega := [D\Xi(\bar{y})]^*\Psi$ is a Lagrange multiplier matrix of the problem (1.1). We say that $h \in \mathbb{R}^n$ is a *critical direction* at the point $\bar{x}$, if $D\mathcal{G}(\bar{x})h \preceq 0$ and $Df(\bar{x})h = 0$. Note that since $\mathcal{G}(\bar{x}) = 0$, the tangent cone to $\mathcal{S}_-^{m-r}$ at $\mathcal{G}(\bar{x})$ coincides with $\mathcal{S}_-^{m-r}$, and hence $T_K(\bar{y}) = [D\Xi(\bar{y})]^{-1}\mathcal{S}_-^{m-r}$. Therefore the set of all critical directions can be written as

$$C(\bar{x}) = \{h \in \mathbb{R}^n : DG(\bar{x})h \in T_K(\bar{y}),\ Df(\bar{x})h = 0\}, \tag{3.9}$$

or, equivalently, for any $\bar{\Omega} \in \Lambda(\bar{x})$,

$$C(\bar{x}) = \{h \in \mathbb{R}^n : DG(\bar{x})h \in T_K(\bar{y}),\ \bar{\Omega} \bullet DG(\bar{x})h = 0\}. \tag{3.10}$$

Note that $C(\bar{x})$ is a closed convex cone, referred to as the *critical cone*. Note also that since $\Lambda(\bar{x})$ is nonempty, the optimal value of the linearized problem (3.3) is 0, and hence $C(\bar{x})$ coincides with the set of optimal solutions of the linearized problem (3.3). Cone $C(\bar{x})$ represents those directions for which the first order linearization of (P) does not provide an information about local optimality of $\bar{x}$.

Consider a critical direction $h \in C(\bar{x})$ and a curve $x(t) := \bar{x} + th + \frac{1}{2}t^2 w + \varepsilon(t)$, $t \geq 0$, where the remainder $\varepsilon(t)$ is of order $o(t^2)$. Suppose that Robinson constraint qualification (3.4) holds. We have then that Robinson constraint qualification for the reduced problem (3.8) holds as well. It follows from the stability result (3.7) and the Taylor expansion

$$\mathcal{G}(x(t)) = \mathcal{G}(\bar{x}) + tD\mathcal{G}(\bar{x})h + \tfrac{1}{2}t^2\left[D\mathcal{G}(\bar{x})w + D^2\mathcal{G}(\bar{x})(h,h)\right] + o(t^2),$$

that $x(t)$ can be feasible, i.e. the remainder term $\varepsilon(t) = o(t^2)$ can be chosen in such a way that $\mathcal{G}(x(t)) \in S_-^{m-r}$ for $t > 0$ small enough, iff

$$\operatorname{dist}\left(D\mathcal{G}(\bar{x})h + t\left[D\mathcal{G}(\bar{x})w + D^2\mathcal{G}(\bar{x})(h,h)\right], S_-^{m-r}\right) = o(t).$$

That is, iff

$$D\mathcal{G}(\bar{x})w + D^2\mathcal{G}(\bar{x})(h,h) \in T_{S_-^{m-r}}(D\mathcal{G}(\bar{x})h). \tag{3.11}$$

Note that $D\mathcal{G}(\bar{x})h \in S_-^{m-r}$ since it is assumed that h is a critical direction.

We have that if $\bar{x}$ is a locally optimal solution of the problem (3.8) and $x(t)$ is feasible, then $f(x(t)) \geq f(\bar{x})$ for $t > 0$ small enough. By using the corresponding second order Taylor expansion of $f(x(t))$ and since $Df(\bar{x})h = 0$ we obtain that the optimization problem

$$\begin{array}{ll} \operatorname{Min}_{w\in\mathbb{R}^n} & Df(\bar{x})w + D^2 f(\bar{x})(h,h) \\ \text{subject to} & D\mathcal{G}(\bar{x})w + D^2\mathcal{G}(\bar{x})(h,h) \in T_{S_-^{m-r}}(D\mathcal{G}(\bar{x})h), \end{array} \tag{3.12}$$

has a nonnegative optimal value. The above problem is a linear problem subject to "cone constraint". The dual of that problem consists in maximization of $h^T \nabla^2_{xx} \mathcal{L}(\bar{x}, \Psi) h$ over $\Psi \in M(\bar{x})$, where $M(\bar{x})$ denotes the set of all Lagrange multiplier matrices of the problem (3.8). Therefore the following second order necessary conditions, for $\bar{x}$ to be a locally optimal solution of the problem (3.8), hold

$$\sup_{\Psi \in M(\bar{x})} h^T \nabla^2_{xx} \mathcal{L}(\bar{x}, \Psi) h \geq 0, \quad \forall\, h \in C(\bar{x}). \tag{3.13}$$

The above second order necessary conditions can be formulated in terms of the original problem (1.1). By the chain rule we have

$$h^T \nabla^2_{xx} \mathcal{L}(\bar{x}, \Psi) h = h^T \nabla^2_{xx} L(\bar{x}, \Omega) h + \varsigma(\Omega, h), \tag{3.14}$$

where $\Omega := [D\Xi(\bar{y})]^* \Psi$ is a Lagrange multiplier matrix of the problem (1.1) and

$$\varsigma(\Omega, h) := \Omega \bullet D^2 \Xi(\bar{y})(DG(\bar{x})h, DG(\bar{x})h). \tag{3.15}$$

The additional term $\varsigma(\Omega, h)$, which appears here, is a quadratic function of h and in a sense represents the curvature of the cone $\mathcal{S}^m_-$ at the point $\bar{y} := G(\bar{x})$. This term can be calculated explicitly by various techniques, and can be written as follows (unfortunately the involved derivations are not trivial and will be not given here)

$$\varsigma(\Omega, h) = h^T H(\bar{x}, \Omega) h, \tag{3.16}$$

where $H(\bar{x}, \Omega)$ is an $n \times n$ symmetric matrix with typical elements

$$[H(\bar{x}, \Omega)]_{ij} := -2\Omega \bullet \left(G_i(\bar{x}) [G(\bar{x})]^{\dagger} G_j(\bar{x}) \right), \quad i, j = 1, ..., n. \tag{3.17}$$

Here $G_i(\bar{x}) := \partial G(\bar{x}) / \partial x_i$ is $m \times m$ matrix of partial derivatives and $[G(\bar{x})]^{\dagger}$ denotes the Moore-Penrose pseudo-inverse of $G(\bar{x})$, i.e. $[G(\bar{x})]^{\dagger} = \sum_{i=1}^r \alpha_i^{-1} e_i e_i^T$ where α_i are nonzero eigenvalues of $G(\bar{x})$ and e_i are corresponding orthonormal eigenvectors. Note that in the case of linear semidefinite problem (1.2), partial derivatives matrices $G_i(\bar{x})$ do not depend on $\bar{x}$, and $G_i(\bar{x}) = A_i$.

Together with (3.13) this implies the following second order necessary conditions.

Theorem 3.3 *Let $\bar{x}$ be a locally optimal solution of the semidefinite problem (1.1), and suppose that Robinson constraint qualification (3.4) holds. Then*

$$\sup_{\Omega \in \Lambda(\bar{x})} h^T \left(\nabla^2_{xx} L(\bar{x}, \Omega) + H(\bar{x}, \Omega) \right) h \geq 0, \quad \forall\, h \in C(\bar{x}). \tag{3.18}$$

The matrix $H(\bar{x},\Omega)$ can be written in the following equivalent form

$$H(\bar{x},\Omega) = -2\left(\frac{\partial G(\bar{x})}{\partial x}\right)^T \left(\Omega \otimes [G(\bar{x})]^\dagger\right)\left(\frac{\partial G(\bar{x})}{\partial x}\right), \tag{3.19}$$

where $\partial G(\bar{x})/\partial x$ denotes the $m^2 \times n$ Jacobian matrix

$$\partial G(\bar{x})/\partial x := [\operatorname{vec} G_1(\bar{x}), ..., \operatorname{vec} G_n(\bar{x})],$$

and "$\otimes$" stands for the Kronecker product of matrices. Since $\Omega \succeq 0$ and $G(\bar{x}) \preceq 0$ we have that $\Omega \otimes [G(\bar{x})]^\dagger$ is a negative semidefinite matrix of rank rp, where $r = \operatorname{rank} G(\bar{x})$ and $p = \operatorname{rank} \Omega$. Therefore the matrix $H(\bar{x},\Omega)$ is positive semidefinite of rank less than or equal to rp. It follows that the additional term $h^T H(\bar{x},\Omega)h$ is always nonnegative. Of course if the semidefinite problem is linear, then $\nabla^2_{xx} L(\bar{x},\Omega) = 0$. Nevertheless, even in the linear case the additional term can be strictly positive.

It is said that the *quadratic growth* condition holds, at a feasible point $\bar{x}$ of the problem (P), if there exists $c > 0$ such that for any feasible point x in a neighborhood of $\bar{x}$ the following inequality holds

$$f(x) \geq f(\bar{x}) + c\|x - \bar{x}\|^2. \tag{3.20}$$

Clearly this quadratic growth condition implies that $\bar{x}$ is a locally optimal solution of (P).

Theorem 3.4 *Let $\bar{x}$ be a stationary point of the semidefinite problem (1.1), and suppose that Robinson constraint qualification (3.4) holds. Then the quadratic growth condition (3.20) holds if and only if the following conditions are satisfied*

$$\sup_{\Omega \in \Lambda(\bar{x})} h^T \left(\nabla^2_{xx} L(\bar{x},\Omega) + H(\bar{x},\Omega)\right) h > 0, \quad \forall\, h \in C(\bar{x}) \setminus \{0\}. \tag{3.21}$$

The above second order sufficient conditions can be proved in two steps. First, the corresponding second order sufficient conditions (without the additional term) can be derived for the reduced problem (3.8). That is, it is possible to show that second order conditions (3.13) become sufficient if the weak inequality sign is replaced by the strict inequality sign. Proof of that is based on the simple observation that for any $\Psi \succeq 0$, and in particular for any $\Psi \in M(\bar{x})$, and any feasible x, i.e. such that $\mathcal{G}(x) \preceq 0$, the inequality $f(x) \geq \mathcal{L}(x,\Psi)$ holds. Second, the obtained second order sufficient conditions for the reduced problem are translated into conditions (3.21) in exactly the same way as it was done for the corresponding second order necessary conditions.

It can be noted that the only difference between the second order necessary conditions (3.18) and the second order sufficient conditions (3.21) is the change of the weak inequality sign in (3.18) into the strict inequality in (3.21). That is, (3.18) and (3.21) give a pair of "no gap" second order optimality conditions for the semidefinite programming problem (1.1). As we mentioned earlier, even in the case of linear semidefinite programming problems the additional term can be positive and hence the corresponding quadratic growth condition can hold.

Note that it can happen that, for a stationary point $\bar{x}$, the critical cone $C(\bar{x})$ contains only one point 0. Of course, in that case the second order conditions (3.18) and (3.21) trivially hold. Since $\bar{x}$ is stationary, we have that $h = 0$ is an optimal solution of the linearized problem (3.3). Therefore condition $C(\bar{x}) = \{0\}$ implies that $Df(\bar{x})h > 0$ for any $h \neq 0$ such that $DG(\bar{x})h \in T_K(G(\bar{x}))$. This in turn implies that for some $\alpha > 0$ and all feasible x sufficiently close to $\bar{x}$, the inequality

$$f(x) \geq f(\bar{x}) + \alpha \|x - \bar{x}\| \tag{3.22}$$

holds. That is, if $\bar{x}$ is a stationary point and $C(\bar{x}) = \{0\}$, then $\bar{x}$ is a *sharp* local minimizer of the problem (P).

Let us discuss now calculation of the critical cone $C(\bar{x})$. The tangent cone to $\mathcal{S}^m_-$ at $G(\bar{x})$ can be written as follows

$$T_{\mathcal{S}^m_-}(G(\bar{x})) = \{Z \in \mathcal{S}^m : E^T Z E \preceq 0\}, \tag{3.23}$$

where $r = \text{rank}(G(\bar{x}))$ and E is an $m \times (m-r)$ matrix complement of $G(\bar{x})$, i.e. E is of rank $m - r$ and such that $G(\bar{x})E = 0$. By (3.10) we obtain then that for any $\Omega \in \Lambda(\bar{x})$,

$$C(\bar{x}) = \left\{ h \in \mathbb{R}^n : \sum_{i=1}^n h_i E^T G_i(\bar{x}) E \preceq 0,\ \sum_{i=1}^n h_i \Omega \bullet G_i(\bar{x}) = 0 \right\}. \tag{3.24}$$

Let us observe that it follows from the first order optimality conditions (3.2) (i.e. since $\Omega \succeq 0$, $G(\bar{x}) \preceq 0$ and because of the complementarity condition $\Omega \bullet G(\bar{x}) = 0$) that for any $\Omega \in \Lambda(\bar{x})$, the inequality

$$\text{rank}(\Omega) + \text{rank}(G(\bar{x})) \leq m$$

holds. We say that the *strict complementarity* condition holds at $\bar{x}$ if there exists $\Omega \in \Lambda(\bar{x})$ such that

$$\operatorname{rank}(\Omega) + \operatorname{rank}(G(\bar{x})) = m. \tag{3.25}$$

Consider a Lagrange multipliers matrix Ω. By the complementarity condition $\Omega \bullet G(\bar{x}) = 0$ we have that $\Omega = E_1 E_1^T$, where E_1 is an $m \times s$ matrix of rank $s = \operatorname{rank}(\Omega)$ and such that $G(\bar{x})E_1 = 0$. Let E_2 be a matrix such that $E_2^T E_1 = 0$ and the matrix $E := [E_1, E_2]$ forms a complement of $G(\bar{x})$. (If the strict complementarity condition (3.25) holds, then $E = E_1$.) Then the critical cone can be written in the form

$$C(\bar{x}) = \left\{ h : \begin{array}{l} \sum_{i=1}^n h_i E_1^T G_i(\bar{x}) E_1 = 0,\ \sum_{i=1}^n h_i E_1^T G_i(\bar{x}) E_2 = 0, \\ \sum_{i=1}^n h_i E_2^T G_i(\bar{x}) E_2 \preceq 0 \end{array} \right\}. \tag{3.26}$$

In particular if the strict complementarity condition (3.25) holds, and hence $\Omega = EE^T$ for some complement E of $G(\bar{x})$, then

$$C(\bar{x}) = \left\{ h \in \mathbb{R}^n : \sum_{i=1}^n h_i E^T G_i(\bar{x}) E = 0 \right\}, \tag{3.27}$$

and hence in that case $C(\bar{x})$ is a linear space.

Let us note that Robinson constraint qualification (3.4) is equivalent to the condition

$$DG(\bar{x})\mathbb{R}^n + T_K(G(\bar{x})) = \mathcal{S}^m. \tag{3.28}$$

Consider the linear space

$$\mathcal{L}(G(\bar{x})) := \{Z \in \mathcal{S}^m : E^T Z E = 0\}, \tag{3.29}$$

where E is a complement matrix of $G(\bar{x})$. (It is not difficult to see that this linear space does not depend on a particular choice of the complement matrix E.) This linear space represents the tangent space to the (smooth) manifold of matrices of rank $r = G(\bar{x})$ in the space $\mathcal{S}^m$ at the point $G(\bar{x})$. It follows from (3.23) that $\mathcal{L}(G(\bar{x})) \subset T_{\mathcal{S}^m_-}(G(\bar{x}))$. In fact $\mathcal{L}(G(\bar{x}))$ is the largest linear subspace of $T_{\mathcal{S}^m_-}(G(\bar{x}))$, i.e. it is the lineality space of the tangent cone $T_{\mathcal{S}^m_-}(G(\bar{x}))$. We have then that condition

$$DG(\bar{x})\mathbb{R}^n + \mathcal{L}(G(\bar{x})) = \mathcal{S}^m \tag{3.30}$$

is stronger than the corresponding Robinson constraint qualification, i.e. (3.30) implies (3.28). It is said that the *nondegeneracy condition* holds at $\bar{x}$ if condition

(3.30) is satisfied. It is not difficult to show that the nondegeneracy condition (3.27) holds iff the n-dimensional vectors $f_{ij} := (e_i^T G_1(\bar{x}) e_j, ..., e_i^T G_n(\bar{x}) e_j)$, $1 \le i \le j \le m - r$, are linearly independent. Here $e_1, ..., e_{m-r}$ are column vectors of the complement matrix E. Since there are $(m-r)(m-r+1)/2$ such vectors f_{ij}, the nondegeneracy condition can hold only if $n \ge (m-r)(m-r+1)/2$.

If the nondegeneracy condition holds, then $\Lambda(\bar{x})$ is a singleton. Moreover, if the strict complementarity condition holds, then $\Lambda(\bar{x})$ is a singleton iff the nondegeneracy condition is satisfied. If the nondegeneracy and strict complementarity conditions hold, then the critical cone $C(\bar{x})$ is defined by a system of $(m-r)(m-r+1)/2$ linearly independent linear equations and hence is a linear space of dimension $n - (m-r)(m-r+1)/2$.

4 STABILITY AND SENSITIVITY ANALYSIS

In this section we consider a parameterized semidefinite programming problem in the form

$$\operatorname*{Min}_{x \in \mathbb{R}^n} f(x, u) \text{ subject to } G(x, u) \preceq 0, \tag{4.1}$$

depending on the parameter vector u varying in a finite dimensional vector space $\mathcal{U}$. Here $f : \mathbb{R}^n \times \mathcal{U} \to \Re$, $G : \mathbb{R}^n \times \mathcal{U} \to \mathcal{S}^m$, and we assume that for a given value u_0 of the parameter vector, problem (4.1) coincides with the "unperturbed" problem (P), i.e. $f(\cdot, u_0) = f(\cdot)$ and $G(\cdot, u_0) = G(\cdot)$. We study continuity and differentiability properties of the optimal value $v(u)$ and an optimal solution $\hat{x}(u)$ of the parameterized problem (4.1) as functions of u near the point u_0. We assume throughout this section that the optimal value of the problem (P) is *finite*. Since the analysis is somewhat involved we only state some main results referring to the literature for proofs and extensions.

Let us consider first the parameterized problem (P_y), $y \in \mathcal{Y}$, in the form (2.17) and the corresponding optimal value function $v(y)$. We have then that if the problem (P) is convex and the regularity condition (2.25) holds, then $v(y)$ is continuous at $y = 0$ and $\partial v(0)$ is nonempty and bounded and coincides with the set of optimal solutions of the dual problem (D). By convex analysis we have that if a convex function $v(\cdot)$ is finite valued and continuous at a point

$y \in \mathcal{Y}$, then for any $d \in \mathcal{Y}$,

$$v'(y, d) = \sup_{y^* \in \partial v(y)} y^* \cdot d, \tag{4.2}$$

where

$$v'(y, d) := \lim_{t \downarrow 0} \frac{v(y + td) - v(y)}{t}$$

denotes the directional derivative of $v(\cdot)$ at y in the direction d.

In particular, we can consider the following parameterization of the semidefinite problem (1.1)

$$\operatorname*{Min}_{x \in C} f(x) \text{ subject to } G(x) + Y \preceq 0, \tag{4.3}$$

parameterized by $Y \in \mathcal{S}^m$, and the corresponding optimal value function $v(Y)$. By the above discussion we have the following result.

Theorem 4.1 *Suppose that the semidefinite problem (1.1) is convex and that the Slater condition holds. Then the set* $\mathrm{Sol}(D)$, *of optimal solutions of the dual problem* (D), *is nonempty and bounded, the optimal value function* $v(Y)$ *of the problem (4.3) is convex, continuous at* $Y = 0$ *and*

$$v'(0, A) = \sup_{\Omega \in \mathrm{Sol}(D)} \Omega \bullet A. \tag{4.4}$$

In particular, we have that, under the assumptions of the above theorem, if the dual problem (D) has a unique optimal solution $\bar{\Omega}$, then the optimal value function $v(\cdot)$ is differentiable at $Y = 0$ and $\nabla v(0) = \bar{\Omega}$.

Let us consider now the parametric problem (P_u) in the general form (4.1). We denote by $L(x, \Omega, u)$ the Lagrangian associated with (P_u), that is

$$L(x, \Omega, u) := f(x) + \Omega \bullet G(x, u).$$

We assume that the functions $f(x, u)$ and $G(x, u)$ are sufficiently smooth, at least are continuously differentiable, and that the following, so-called *inf-compactness*, condition holds: there exist a number $\alpha > v(u_0)$ and a compact set $S \subset \mathbb{R}^n$ such that

$$\{x \in \mathbb{R}^n : f(x, u) \leq \alpha,\ G(x, u) \preceq 0\} \subset S \tag{4.5}$$

for all u in a neighborhood of u_0. We have that under the above inf-compactness condition, for all u near u_0, the optimization is actually performed in the compact set S, and hence the set of optimal solutions of (P_u) is nonempty and bounded. In particular, it follows that the set $\mathrm{Sol}(P)$, of optimal solutions of the "unperturbed" problem (1.1), is nonempty and bounded.

In a sense the following result is an extension of formula (4.4).

Theorem 4.2 *Suppose that the semidefinite problem (1.1) is convex and that the Slater and inf-compactness conditions hold. Then the optimal value function $v(u)$ is directionally differentiable at u_0 and for any $d \in \mathcal{U}$,*

$$v'(u_0, d) = \inf_{x \in \mathrm{Sol}(P)} \sup_{\Omega \in \mathrm{Sol}(D)} d \cdot \nabla_u L(x, \Omega, u_0). \tag{4.6}$$

It can be noted that in the case of parametric problem (4.3), the corresponding Lagrangian $L(x, \Omega, Y)$ takes the form $L(x, \Omega) + \Omega \bullet Y$. Therefore its gradient, with respect to Y, coincides with Ω and does not depend on x. Consequently in that case formula (4.6) reduces to formula (4.4). Note, however, that formula (4.4) may hold even if the set $\mathrm{Sol}(P)$ is empty. Note also that in the formulation of the above theorem only the "unperturbed" problem is assumed to be convex while perturbed problems (P_u), $u \neq u_0$, can be nonconvex.

In a nonconvex case the analysis is more delicate and an analogue of formula (4.6) may not hold. Somewhat surprisingly, in nonconvex cases, the question of first order differentiability of the optimal value function may require a second order analysis of the considered optimization problem. Yet the following result holds.

Theorem 4.3 *Suppose that the inf-compactness condition holds and that for any $x \in \mathrm{Sol}(P)$ there exists a unique Lagrange multipliers matrix $\bar{\Omega}(x)$, i.e. $\Lambda(x) = \{\bar{\Omega}(x)\}$. Then the optimal value function is directionally differentiable at u_0 and*

$$v'(u_0, d) = \inf_{x \in \mathrm{Sol}(P)} d \cdot \nabla_u L(x, \bar{\Omega}(x), u_0). \tag{4.7}$$

Note that existence and uniqueness of the Lagrange multipliers matrix can be considered as a constraint qualification. Since then clearly $\Lambda(x)$ is nonempty and bounded, we have that this condition implies Robinson constraint qualification.

Let us discuss now continuity and differentiability properties of an optimal solution $\hat{x}(u)$ of the problem (P_u). We have that, under the inf-compactness condition, the distance from $\hat{x}(u)$ to the set $\mathrm{Sol}(P)$, of optimal solutions of the unperturbed problem, tends to zero as $u \to u_0$. In particular, if $\mathrm{Sol}(P) = \{\bar{x}\}$ is a singleton, i.e. (P) has unique optimal solution $\bar{x}$, then $\hat{x}(u) \to \bar{x}$ as $u \to u_0$. However, the rate at which $\hat{x}(u)$ converges to $\bar{x}$ can be slower than $O(\|u - u_0\|)$, i.e. $\|\hat{x}(u) - \bar{x}\| / \|u - u_0\|$ can tend to ∞ as $u \to u_0$, even if the quadratic growth condition (3.20) holds.

We assume in the remainder of this section that $\mathrm{Sol}(P) = \{\bar{x}\}$ is a singleton and that $f(x,u)$ and $G(x,u)$ are twice continuously differentiable. For $K := \mathcal{S}^m_-$, a given direction $d \in \mathcal{U}$ and $t \geq 0$ consider the following linearization of the problem (P_{u_0+td}),

$$\operatorname*{Min}_{h \in \mathbb{R}^n} Df(\bar{x}, u_0)(h, d) \text{ subject to } DG(\bar{x}, u_0)(h, d) \in T_K(G(\bar{x}, u_0)), \tag{4.8}$$

where

$$Df(\bar{x}, u_0)(h, d) = D_x f(\bar{x}, u_0)h + D_u f(\bar{x}, u_0)d,$$

and similarly for $DG(\bar{x}, u_0)(h, d)$. This is a linear problem subject to "cone constraint", and its dual is given by

$$\operatorname*{Max}_{\Omega \in \Lambda(\bar{x})} d \cdot \nabla_u L(x, \Omega, u_0). \tag{4.9}$$

We refer to the above problems as (PL_d) and (DL_d), respectively. If Robinson constraint qualification holds, then there is no duality gap between problems (PL_d) and (DL_d), the set $\Lambda(\bar{x})$ is nonempty and bounded, and hence the set $\mathrm{Sol}(DL_d)$, of the problem (4.9), is also nonempty and bounded.

Consider the following, so-called strong form, of second order conditions

$$\sup_{\Omega \in \mathrm{Sol}(DL_d)} h^T \left(\nabla^2_{xx} L(\bar{x}, \Omega) + H(\bar{x}, \Omega) \right) h > 0, \quad \forall\, h \in C(\bar{x}) \setminus \{0\}. \tag{4.10}$$

Of course, if $\mathrm{Sol}(DL_d) = \Lambda(\bar{x})$, then the above second order conditions coincide with the second order conditions (3.21). In particular, this happens if $\Lambda(\bar{x})$ is a singleton or if $G(x,u)$ does not depend on u.

We can formulate now the basic sensitivity theorem for semidefinite programming problems.

Theorem 4.4 *Let $\hat{x}(t) := \hat{x}(u_0 + td)$ be an optimal solution of the problem (4.1), for $u = u_0 + td$ and $t \geq 0$, converging to $\bar{x}$ as $t \downarrow 0$. Suppose that Robinson constraint qualification holds at $\bar{x}$, that the strong second order sufficient conditions (4.10) are satisfied and that the set $\mathrm{Sol}(PL_d)$, of optimal solutions of the linearized problem (4.8), is nonempty. Then:*
(i) $\hat{x}(u)$ is Lipschitz stable at $\bar{x}$ in the direction d, i.e.

$$\|\hat{x}(t) - \bar{x}\| = O(t), \quad t \geq 0. \tag{4.11}$$

(ii) For $t \geq 0$, the optimal value function has the following second order expansion along the direction d,

$$v(u_0 + td) = v(u_0) + t\,\mathrm{val}(DL_d) + \tfrac{1}{2}t^2\,\mathrm{val}(Q_d) + o(t^2), \tag{4.12}$$

where $\mathrm{val}(DL_d)$ is the optimal value of the problem (4.9) and $\mathrm{val}(Q_d)$ is the optimal value of the following min-max problem:

$$\operatorname*{Min}_{h \in \mathrm{Sol}(PL_d)} \operatorname*{Max}_{\Omega \in \mathrm{Sol}(DL_d)} \left\{ D^2 L(\bar{x}, \Omega, u_0)((h,d),(h,d)) + \zeta(\Omega, h, d) \right\}, \tag{4.13}$$

referred to as (Q_d), with $D^2 L(\bar{x}, \Omega, u_0)((h,d),(h,d))$ given by

$$h^T \nabla^2_{xx} L(\bar{x}, \Omega, u_0) h + 2h^T \nabla^2_{xu} L(\bar{x}, \Omega, u_0) d + d^T \nabla^2_{uu} L(\bar{x}, \Omega, u_0) d,$$

and

$$\zeta(\Omega, h, d) := h^T H_{xx}(\Omega) h + 2h^T H_{xu}(\Omega) d + d^T H_{uu}(\Omega) d,$$

$$H_{xx}(\Omega) := -2 \left(\frac{\partial G(\bar{x}, u_0)}{\partial x} \right)^T \left(\Omega \otimes [G(\bar{x}, u_0)]^\dagger \right) \left(\frac{\partial G(\bar{x}, u_0)}{\partial x} \right),$$

$$H_{xu}(\Omega) := -2 \left(\frac{\partial G(\bar{x}, u_0)}{\partial x} \right)^T \left(\Omega \otimes [G(\bar{x}, u_0)]^\dagger \right) \left(\frac{\partial G(\bar{x}, u_0)}{\partial u} \right),$$

$$H_{uu}(\Omega) := -2 \left(\frac{\partial G(\bar{x}, u_0)}{\partial u} \right)^T \left(\Omega \otimes [G(\bar{x}, u_0)]^\dagger \right) \left(\frac{\partial G(\bar{x}, u_0)}{\partial u} \right).$$

(iii) Every accumulation point of $(\hat{x}(t) - \bar{x})/t$, as $t \downarrow 0$, is an optimal solution of the min-max problem (Q_d), given in (4.13). In particular, if (Q_d) has a unique optimal solution $\bar{h}$, then

$$\hat{x}(t) = \bar{x} + t\bar{h} + o(t), \quad t \geq 0. \tag{4.14}$$

Let us make the following remarks. If the constraint mapping $G(x, u)$ does not depend on u, then the set $\mathrm{Sol}(PL_d)$, of optimal solutions of the linearized problem, coincides with the critical cone $C(\bar{x})$. In any case $C(\bar{x})$ forms the recession cone of $\mathrm{Sol}(PL_d)$ provided $\mathrm{Sol}(PL_d)$ is nonempty. The above matrix $H_{xx}(\Omega)$ is exactly the same as the matrix $H(\bar{x}, \Omega)$ defined in (3.17) and (3.19), and used in the second order optimality conditions. Strong second order conditions (4.10) ensure that the min-max problem (Q_d) has a finite optimal value and at least one optimal solution. It is possible to show that strong second order conditions (4.10) are "almost necessary" for the directional Lipschitzian stability of $\hat{x}(u)$. That is, if the left hand side of (4.10) is less than 0 for some $h \in C(\bar{x})$, then (4.11) cannot hold.

It follows from (4.12) that the directional derivative $v'(u_0, d)$ exists and is equal to the optimal value of the problem (4.9).

Existence of an optimal solution of the linearized problem (PL_d) is a *necessary* condition for the Lipschitzian stability (4.11) of an optimal solution. The following example shows that it can happen in semidefinite programming that the corresponding linearized problem does not possess an optimal solution.

Example 4.5 Consider the linear space $\mathcal{Y} := \mathcal{S}^2$, the cone $K := \mathcal{S}^2_+$, the mapping $G : \Re^2 \times \Re \to \mathcal{S}^2$ defined as $G(x_1, x_2, t) := \mathrm{diag}(x_1, x_2) + tA$, where $\mathrm{diag}(x_1, x_2)$ denotes the diagonal matrix with diagonal elements x_1 and x_2 and $A = (a_{ij})$ is the 2×2 symmetric matrix with zero diagonal elements and $a_{12} = a_{21} = 1$, and the parameterized problem

$$\operatorname*{Min}_{x \in \Re^2} x_1 + \tfrac{1}{2}x_1^2 + \tfrac{1}{2}x_2^2 \text{ subject to } G(x_1, x_2, t) \in \mathcal{S}^2_+. \tag{4.15}$$

It is not difficult to see that the feasible set of this problem is $\{x : x_1 x_2 \geq t^2,\ x_1 \geq 0, x_2 \geq 0\}$, and that the unperturbed problem (for $t = 0$) has unique optimal solution $\bar{x} = (0, 0)$ and the unique corresponding Lagrange matrix $\bar{\Omega} = \mathrm{diag}(-1, 0)$. Moreover, the above problem (4.15) is convex, the Slater condition holds and the strongest form of second order sufficient conditions is satisfied at the point $\bar{x}$. Yet it is not difficult to verify that the second coordinate of $\hat{x}(t)$ is of order $t^{2/3}$, as $t \downarrow 0$, and hence $\hat{x}(t)$ is not Lipschitz stable. The reason for such behavior of $\hat{x}(t)$ is that the corresponding linearized problem, at the direction $d = 1$,

$$\operatorname*{Min}_{x \in \Re^2} x_1 \text{ subject to } x_1 x_2 \geq 1,\ x_1 \geq 0,\ x_2 \geq 0, \tag{4.16}$$

does not possess an optimal solution.

Let Ω be an optimal solution of (DL_d), and let E_1 and E_2 be such matrices that $\Omega = E_1 E_1^T$, $E_2^T E_1 = 0$ and $E := [E_1, E_2]$ is a complement of $G(\bar{x}, u_0)$. Then similar to (3.26) we have

$$\mathrm{Sol}(PL_d) = \left\{ h : \begin{array}{c} \sum_i h_i E_1^T \frac{\partial G(\bar{x},u_0)}{\partial x_i} E_1 + \sum_j d_j E_1^T \frac{\partial G(\bar{x},u_0)}{\partial u_j} E_1 = 0 \\ \sum_i h_i E_1^T \frac{\partial G(\bar{x},u_0)}{\partial x_i} E_2 + \sum_j d_j E_1^T \frac{\partial G(\bar{x},u_0)}{\partial u_j} E_2 = 0 \\ \sum_i h_i E_2^T \frac{\partial G(\bar{x},u_0)}{\partial x_i} E_2 + \sum_j d_j E_2^T \frac{\partial G(\bar{x},u_0)}{\partial u_j} E_2 \preceq 0 \end{array} \right\}. \quad (4.17)$$

In particular, if the strict complementarity and nondegeneracy conditions hold, then $\Lambda(\bar{x}) = \{\bar{\Omega}\}$ is a singleton, $\mathrm{Sol}(DL_d) = \{\bar{\Omega}\}$ and

$$\mathrm{Sol}(PL_d) = \left\{ h : \sum_i h_i E^T \frac{\partial G(\bar{x}, u_0)}{\partial x_i} E + \sum_j d_j E^T \frac{\partial G(\bar{x}, u_0)}{\partial u_j} E = 0 \right\}. \quad (4.18)$$

In that case the corresponding problem (Q_d), defined in (4.13), becomes a problem of minimization of a quadratic function subject to linear constraints, and hence can be solved in a closed form. It follows then, under the second order sufficient conditions, that: (i) $\mathrm{val}(Q_d)$ is quadratic in d, and hence the optimal value function $v(u)$ is twice differentiable at u_0, and (ii) (Q_d) has a unique optimal solution $\bar{h} = \bar{h}(d)$ which is a linear function of d, and hence $v(u)$ is differentiable at u_0.

5 NOTES

Lagrangian duality is a well developed concept in mathematical programming. Its origins go back to von Neumann's game theory. In the context of semidefinite programming particular examples of duality schemes were considered, for example, in [1, 19, 25]. Example 2.2, of a linear semidefinite program with a duality gap, is taken from [24]. The parametric approach to duality, by applying convex analysis to the parametric problem (2.17), was developed in Rockafellar [17, 18]. A proof of the Fenchel-Moreau duality theorem can be found in [17]. Convex semidefinite programming problems were discussed in [20].

First order necessary conditions of the form (3.1), for optimization problems subject to "cone constraints", were discussed in [10, 16, 26]. The constraint qualification (3.4) was introduced by Robinson in [15]. The stability result (3.8) is called metric regularity and is based on the Robinson-Ursescu stability theorem [14, 23]. Constraint qualifications (3.5) and (3.7) can be considered as extensions of the Mangasarian-Fromovitz [13] constraint qualification used in nonlinear programming. The result of theorem 3.2 is essentially due to Zowe and Kurcyusz [26].

The reduction approach to semidefinite programming (of considering the reduced problem (3.8)) is due to Bonnans and Shapiro [4]. A general formula for the additional term in second order optimality conditions under cone constraints is given in Cominetti [7], in a form of the support function of a certain second order tangent set, see also Kawasaki [9]. In the case of semidefinite programming this additional term is explicitly calculated and the second order necessary conditions (3.18) are given in Shapiro [20]. Sufficiency of second order conditions (3.21) is proved in [5]. The nondegeneracy condition in the form (3.30) was introduced in Shapiro and Fan [22] as a transversality condition (see also [2, 20] for a discussion of the nondegeneracy concept).

Theorem 4.2, giving the min-max formula (4.6) for the directional derivatives of the optimal value function in the convex case, is essentially due to Gol'shtein [8]. The result of theorem 4.3 is due to Levitin [12] and Lempio and Maurer [11]. Uniqueness of Lagrange multipliers in cone constrained, and in particular in semidefinite programming, problems is discussed in [21].

An extensive discussion of sensitivity type results of theorem 4.4 can be found in the review paper [3]. In the semidefinite programming first results of that type were obtained by an application of the Implicit Function Theorem in [20]. A proof of theorem 4.4 is given in [6]. Example 4.5 is taken from [3].

REFERENCES

[1] Alizadeh, F., Interior point methods in semidefinite programming with applications to combinatorial optimization. SIAM J. Optim. 1995; 5:13-51.

[2] Alizadeh, F., Haeberly, J. A., Overton, M. L., Complementarity and nondegeneracy in semidefinite programming. Mathematical Programming 1997; 77:111-128.

[3] Bonnans, J.F. and Shapiro, A., Optimization problems with perturbations, a guided tour. SIAM Review, to appear.

[4] Bonnans, J.F., and Shapiro A., Nondegeneracy and quantitative stability of parameterized optimization problems with multiple solutions. SIAM J. Optimization, to appear.

[5] Bonnans,J.F., Cominetti, R. and Shapiro, A., Second order optimality conditions based on parabolic second order tangent sets. SIAM J. Optimization, to appear.

[6] Bonnans,J.F., Cominetti, R. and Shapiro, A., Sensitivity analysis of optimization problems under second order regular constraints. Mathematics of Operations Research, to appear.

[7] Cominetti, R., Metric regularity, tangent sets and second order optimality conditions. Applied Mathematics and Optimization 1990; 21:265-287.

[8] Gol'shtein, E.G., *Theory of Convex Programming*, Translations of Mathematical Monographs 36, American Mathematical Society, Providence, RI, 1972.

[9] Kawasaki, H., An envelope-like effect of infinitely many inequality constraints on second-order necessary conditions for minimization problems. Mathematical Programming 1988; 41:73-96.

[10] Kurcyusz, S., On the existence and nonexistence of Lagrange multipliers in Banach spaces. Journal of Optimization Theory and Applications 1976; 20:81-110.

[11] Lempio, F. and Maurer, H., Differential stability in infinite-dimensional nonlinear programming. Applied Mathematics and Optimization 1980; 6:139-152.

[12] Levitin, E.S., On differential properties of the optimal value of parametric problems of mathematical programming. Dokl. Akad. Nauk SSSR 1975; 224:1354-1358.

[13] Mangasarian, O.L., Fromovitz, S., The Fritz John necessary optimality conditions in the presence of equality and inequality constraints. J. Math. Anal. Appl. 1967;17:37-47.

[14] Robinson, S.M., Regularity and stability for convex multivalued functions. Mathematics of Operations Research 1976; 1:130-143.

[15] Robinson, S.M., Stability theorems for systems of inequalities, Part II: differentiable nonlinear systems. SIAM J. Numerical Analysis 1976; 13:497-513.

[16] Robinson, S.M., First order conditions for general nonlinear optimization. SIAM Journal on Applied Mathematics 1976; 30:597-607.

[17] Rockafellar, R.T., *Convex Analysis*, Princeton University Press, Princeton, NJ, 1970.

[18] Rockafellar, R.T., *Conjugate Duality and Optimization*, Regional Conference Series in Applied Mathematics, SIAM, Philadelphia, 1974.

[19] Shapiro, A., Extremal problems on the set of nonnegative definite matrices. Linear Algebra and Applications 1985; 67:7-18.

[20] Shapiro, A., First and second order analysis of nonlinear semidefinite programs. Mathematical Programming 1997; 77: 301-320.

[21] Shapiro, A., On uniqueness of Lagrange multipliers in optimization problems subject to cone constraints. SIAM J. Optimization 1997; 7:508-518.

[22] Shapiro, A., and and Fan, M.K.H., On eigenvalue optimization. SIAM J. on Optimization 1995; 5:552-569.

[23] Ursescu, C., Multifunctions with convex closed graph. Czechoslovak Mathematical Journal 1975; 25:438-441.

[24] Vandenberghe, L., and Boyd, S., Semidefinite programming. SIAM Review 1996; 38:49-95.

[25] Wolkowicz, H., Some applications of optimization in matrix theory. Linear Algebra and Its Applications 1981; 40:101-118.

[26] Zowe, J. and Kurcyusz, S., Regularity and stability for the mathematical programming problem in Banach spaces. Applied Mathematics and Optimization 1979; 5:49-62.