Journal of Convex Analysis
Volume 15 (2008), No. 4, 767–801

# Set-Semidefinite Optimization

**Gabriele Eichfelder**
*Department Mathematik, Universität Erlangen-Nürnberg,*
*Martensstr. 3, 91058 Erlangen, Germany*
*Gabriele.Eichfelder@am.uni-erlangen.de*

**Johannes Jahn**
*Department Mathematik, Universität Erlangen-Nürnberg,*
*Martensstr. 3, 91058 Erlangen, Germany*
*jahn@am.uni-erlangen.de*

Received: July 6, 2007

In this paper set-semidefinite optimization is introduced as a new field of vector optimization in infinite dimensions covering semidefinite and copositive programming. This unified approach is based on a special ordering cone, the so-called $K$-semidefinite cone for which properties are given in detail. Optimality conditions in the KKT form and duality results including the linear case are presented for $K$-semidefinite optimization problems. A penalty approach is developed for the treatment of the special constraint arising in $K$-semidefinite optimization problems.

*Keywords:* Vector optimization, convex analysis, semidefinite programming, copositive programming

*2000 Mathematics Subject Classification:* 90C29, 90C48, 90C22, 90C46

## 1. Introduction

Semidefinite programming is a rapidly growing field of optimization (compare [48] and [49] for an overview). Various important practical problems lead to a semidefinite programming formulation (e.g., see [48] and [28]). It is characteristic for these finite dimensional optimization problems that the image of a certain matrix-valued function is positive semidefinite. During the last years a further problem class (where the copositivity of the image of this matrix-valued function is required) has been added. These problems are called copositive programming problems (e.g., see [6], [40], [5], [17], [16] and [41]). It turns out that both problem classes can be covered by a general class of optimization problems with cone constraint.

Copositive programming problems arise in different fields of optimization. For instance, continuous quadratic optimization problems ([6], [18], [11] and the references in [21]), the maximum stable set problem ([6], [18], [11]), the quadratic assignment problem ([18], [42], [11]) and a minimum-cut graph tri-partitioning problem ([11], [43]) can be written as copositive programs. It is interesting to note that copositivity also has its application in elasticity theory (see [46]), differential equations e.g. appearing in population genetics (see [21]), linear quadratic control (see [26]) and game theory in biology (see [7], [32]).

In this paper we investigate a general partial ordering being useful for both problem classes, we explore these problems with vector-valued objective function, and we develop

ISSN 0944-6532 / $ 2.50 

the whole theory in an infinite-dimensional setting. Since the partial ordering is defined by a so-called $K$-semidefinite ordering cone for an arbitrary nonempty set $K$, we speak of set-semidefinite optimization. For the definition of an ordering cone see Def. 1.19, [27].

To be more specific we have the following standard assumption in this paper.

**Assumption 1.1.** Let $X, Y$ and $Z$ be topological linear spaces; let $Z$ be partially ordered by a pointed convex cone $C_Z$; and let $f : X \to Z$ and $G : X \to L(Y, Y^*)$ (here $L(Y, Y^*)$ denotes the linear space of continuous linear maps from $Y$ to its topological dual space $Y^*$) be given maps.

In the following we write linear forms $\ell$ of a topological dual space as $\langle \ell, \cdot \rangle$, and for the scalar product in $\mathbb{R}^n$ we use the notation $a^\top b$ for $a, b \in \mathbb{R}^n$. Under Assumption 1.1 we investigate the vector optimization problem

$$\min_{x \in S} \; f(x) \tag{1}$$

with the constraint set

$$S := \{x \in X \mid G(x) \preccurlyeq 0_{L(Y,Y^*)}\}. \tag{2}$$

Here $\preccurlyeq$ denotes a special partial ordering defined later. Recall that a *minimal solution* $\bar{x} \in S$ of the vector optimization problem (1) is defined as the preimage of a minimal element $f(\bar{x})$ of the image set $f(S)$, i.e. $\bar{x}$ is a minimal solution if

$$(\{f(\bar{x})\} - C_Z) \cap f(S) = \{f(\bar{x})\}$$

(compare [27]). In connection with optimality conditions we are also interested in weakly minimal solutions being defined under the additional assumption that $C_Z$ has a nonempty interior $\mathrm{int}(C_Z)$. The element $\bar{x} \in S$ is called a *weakly minimal solution* of problem (1), if $f(\bar{x})$ is a weakly minimal element of the image set $f(S)$, i.e.

$$(\{f(\bar{x})\} - \mathrm{int}(C_Z)) \cap f(S) = \emptyset.$$

Special problems with scalar-valued objective function and matrix-valued function $G$ belong to the general problem class defined by (1). For instance, if $\mathcal{M}^n$ denotes the linear space of real $(n, n)$ matrices with the scalar product $\langle \cdot, \cdot \rangle$ defined by $\langle A, B \rangle = \mathrm{trace}(A \cdot B^\top)$ for all $A, B \in \mathcal{M}^n$, and the partial ordering $\preccurlyeq$ induced by the Löwner ordering cone

$$\mathcal{M}^n_+ := \{A \in \mathcal{M}^n \mid A \text{ is positive semidefinite}\}, \tag{3}$$

then the optimization problem (1) describes a semidefinite programming problem (Löwner [34] has introduced this partial ordering in 1934). One often considers the linear subspace $\mathcal{S}^n$ of $\mathcal{M}^n$ consisting of symmetric matrices. In this special case the scalar product is defined by $\langle A, B \rangle = \mathrm{trace}(A \cdot B)$ for all $A, B \in \mathcal{S}^n$. The ordering cone in this subspace is denoted by $\mathcal{S}^n_+ := \mathcal{M}^n_+ \cap \mathcal{S}^n$.

Recall that a matrix $A \in \mathcal{M}^n$ is called *copositive*, if $y^\top A y \geq 0$ for all $y \in \mathbb{R}^n$ with $y_1, \ldots, y_n \geq 0$. This notion has been introduced by Motzkin [38] in 1952. If the partial ordering $\preccurlyeq$ is induced by

$$\mathcal{M}^n_{++} := \{A \in \mathcal{M}^n \mid A \text{ is copositive}\}, \tag{4}$$

then the optimization problem (1) describes a copositive programming problem. We have again $\mathcal{S}^n_{++} := \mathcal{M}^n_{++} \cap \mathcal{S}^n$.

Based on this special ordering structure for matrices we use a general partial ordering $\preccurlyeq$.

**Definition 1.2.** For an arbitrary nonempty set $K \subset Y$ the set

$$C^K_{L(Y,Y^*)} := \{A \in L(Y,Y^*) \mid \langle Ay, y\rangle \geq 0 \text{ for all } y \in K\}$$

is called *$K$-semidefinite cone* (for simplicity we write $C^K_L$). Every $A \in C^K_L$ is called *$K$-semidefinite*.

In this definition we require that a map $A \in C^K_L$ is positive semidefinite on a set $K$. This set $K$ plays the role of a parameter for the ordering cone $C^K_L$. Notice that $C^K_L$ is a convex cone for every parameter $K$. Therefore, it is an ordering cone inducing a partial ordering $\preccurlyeq$ in $L(Y,Y^*)$. This partial ordering is used in the definition of the constraint set (2). $K$-semidefinite quadratic forms for a closed convex cone $K$ and a real Hilbert space $(Y, \langle\cdot,\cdot\rangle)$ have already been studied by Martin [35] in 1980. An application in control theory is also discussed in [35].

**Example 1.3.** In this example we consider the matrix space $\mathcal{M}^n$ and special sets $K$:

(a) For $K = \{0_{\mathbb{R}^n}\}$ we obtain $C^K_{\mathcal{M}^n} = \mathcal{M}^n$.

(b) With $K = \mathbb{R}^n_+$ we conclude $C^K_{\mathcal{M}^n} = \mathcal{M}^n_{++}$ (defined by (4)). In this case the ordering cone consists of all copositive matrices.

(c) If we choose $K = \mathbb{R}^n$, we get $C^K_{\mathcal{M}^n} = \mathcal{M}^n_+$ (defined by (3)). So, this ordering cone describes all positive semidefinite matrices.

**Remark 1.4.** In nonlinear optimization a necessary optimality condition of second order for constrained problems in Banach spaces says that the quadratic form of the second Fréchet derivative of the Lagrangian at the optimal solution is nonnegative on a certain contingent cone (for details see [37], Thm. 3.3, [8], [53], Thm. 5.5.2, and compare also [14]). If the map defining the inequality constraint is convex, then this contingent cone is even convex. This necessary optimality condition of second order means that the second Fréchet derivative of the Lagrangian is set-semidefinite. So, the set-semidefinite cone is an important concept in nonlinear optimization.

In continuous finite dimensional optimization $K$-semidefinite maps have been investigated in connection with a conjugate decomposition by Han and Mangasarian [23] (see also [22]) and in connection with optimality conditions (see also [14]), and further in [15], [44] and [47] in view of quadratic programming. This notion has been used by Gowda [20] for the investigation of complementarity problems. Blum and Oettli [3] have utilized this concept for the investigation of equilibrium problems. In copositive programming this notion has been already mentioned in [6]. In the literature $K$-semidefinite matrices are also called copositive with respect to the set $K$ (see e.g. [26], [20], [47], [2]), $K$-copositive ([7], [4], [6], [14]), cone-positive (if $K$ is a cone, see [44]), positive semidefinite ([22], [24]) or nonnegative definite ([51], [52]) on the set $K$, or $M$-conditionally positive definite if $K = \{y \in \mathbb{R}^n \mid My \in \mathbb{R}^n_+\}$ for a matrix $M$ (see e.g. [36]).

In the next sections we investigate problem (1). This is done using a convex cone $C^K_L$ inducing the partial ordering $\preccurlyeq$ in (2). Properties of this ordering cone are given in Section

2. Optimality conditions of the Lagrange type reducing to extended KKT conditions are formulated in Section 3. The fourth section presents duality results for problem (1) including the linear case. Finally, we introduce a penalty approach for the treatment of the constraint.

## 2. Properties of $C_L^K$

In this section we study the $K$-semidefinite cone $C_L^K$ for an arbitrary nonempty set $K \subset Y$ and in addition to that especially for the case, that the set $K$ is a cone. We present various properties of this ordering cone and we also examine special variants of this cone in a finite dimensional setting. We begin our investigations with simple calculation rules.

As $C_L^K$ is a convex cone we have for all $A^1, A^2 \in C_L^K$ and $\lambda \geq 0$

$$\lambda(A^1 + A^2) \in C_L^K.$$

**Lemma 2.1.** *Let $K_1, K_2 \subset Y$ be given nonempty sets. Then it is*

$$C_L^{K_1 \cup K_2} = C_L^{K_1} \cap C_L^{K_2}.$$

**Proof.** We first show that $C_L^{K_1 \cup K_2} \subset C_L^{K_1} \cap C_L^{K_2}$. For that let $A \in C_L^{K_1 \cup K_2}$ be arbitrarily chosen, i.e.

$$\langle Ay, y \rangle \geq 0 \quad \text{for all } y \in K_1 \cup K_2.$$

We conclude $\langle Ay, y \rangle \geq 0$ for all $y \in K_1$ and hence $A \in C_L^{K_1}$ and also $\langle Ay, y \rangle \geq 0$ for all $y \in K_2$ and hence $A \in C_L^{K_2}$. Summarizing this we get $A \in C_L^{K_1} \cap C_L^{K_2}$.

It remains to show $C_L^{K_1} \cap C_L^{K_2} \subset C_L^{K_1 \cup K_2}$. For $A \in C_L^{K_1} \cap C_L^{K_2}$ it is $\langle Ay, y \rangle \geq 0$ for all $y \in K_1$ and for all $y \in K_2$. Thus we have $\langle Ay, y \rangle \geq 0$ for all $y \in K_1 \cup K_2$ and therefore $A \in C_L^{K_1 \cup K_2}$. □

**Lemma 2.2.** *Let $K_1, K_2 \subset Y$ be given nonempty sets with $K_1 \subset K_2$. Then it is*

$$C_L^{K_2} \subset C_L^{K_1}.$$

**Proof.** For $A \in C_L^{K_2}$ it is $\langle Ay, y \rangle \geq 0$ for all $y \in K_2$ and due to $K_1 \subset K_2$ it is a fortiori $\langle Ay, y \rangle \geq 0$ for all $y \in K_1$, i.e. $A \in C_L^{K_1}$. □

Using this lemma on $K_1 \cap K_2 \subset K_1 \cup K_2$ together with Lemma 2.1 we conclude:

**Corollary 2.3.** *Let $K_1, K_2 \subset Y$ be given nonempty sets. Then it is*

$$C_L^{K_1 \cap K_2} \supset C_L^{K_1} \cap C_L^{K_2}.$$

The converse inclusion is generally not true as it is demonstrated in the following example.

**Example 2.4.** Let $Y = \mathbb{R}^2$ and let the sets $K_1$, $K_2 \subset \mathbb{R}^2$ be given by

$$K_1 = \{y \in \mathbb{R}^2 \mid y_1 \geq 0,\ y_2 = 0\}$$

and $K_2 = \mathbb{R}^2$. Then the map $A \colon \mathbb{R}^2 \to \mathbb{R}^2$,

$$y \mapsto \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} y$$

is $K_1 \cap K_2$-semidefinite, as $y^\top Ay = y_1^2 - y_2^2 \geq 0$ for all $y \in K_1 \cap K_2 = K_1$. But it is $y^\top Ay < 0$ for $y = (0, 2)^\top \in K_2$ and therefore it is $A \notin C_L^{K_1} \cap C_L^{K_2}$.

In the following lemma it is shown that $K$-semidefiniteness already implies $(-K)$-semidefiniteness.

**Lemma 2.5.** *Let $K \subset Y$ be a given nonempty set. Then it is*

$$C_L^K = C_L^{-K}.$$

**Proof.** Let $A \in C_L^K$ be arbitrarily given. Because of the linearity of $A \in L(Y, Y^*)$ and $Ay \in Y^*$ we have for all $y \in K$

$$\begin{aligned} 0 &\le \langle Ay, y \rangle \\ &= \langle -Ay, -y \rangle \\ &= \langle A(-y), -y \rangle. \end{aligned}$$

This is equivalent to

$$0 \le \langle Ay, y \rangle \text{ for all } y \in -K$$

and thus to $A \in C_L^{-K}$. □

As a direct consequence of Lemma 2.1 and Lemma 2.5 it follows

$$C_L^{K \cup (-K)} = C_L^K. \tag{5}$$

In some special cases $K$-semidefinite maps are already $Y$-semidefinite.

**Lemma 2.6.** *Let $(Y, \langle \cdot, \cdot \rangle)$ be a real Hilbert space, $K \subset Y$ a set with int$(K) \neq \emptyset$ and let $A \in L(Y, Y)$ be a self adjoint map. If $A$ is $K$-semidefinite and there exists a $\bar{y} \in$ int$(K)$ with $A\bar{y} = 0_Y$, then $A$ is already $Y$-semidefinite.*

**Proof.** Let $\bar{y} \in \text{int}(K)$ with $A\bar{y} = 0_Y$ and thus with $\langle A\bar{y}, \bar{y} \rangle = 0$. For all $y \in Y$ there exists a small scalar $\varepsilon > 0$ such that it holds $\bar{y} + \varepsilon\, y \in K$ and thus due to $A$ self adjoint

$$\begin{aligned} 0 &\le \langle A(\bar{y} + \varepsilon\, y), (\bar{y} + \varepsilon\, y) \rangle \\ &= \underbrace{\langle A\bar{y}, \bar{y} \rangle}_{=0} + 2\varepsilon \langle \underbrace{A\bar{y}}_{=0_Y}, y \rangle + \varepsilon^2 \langle Ay, y \rangle \\ &= \underbrace{\varepsilon^2}_{>0} \langle Ay, y \rangle. \end{aligned}$$

Therefore it is $\langle Ay, y \rangle \ge 0$ for all $y \in Y$. □

For symmetric copositive matrices this result is shown in [13]. Cottle et al. discuss in [13] also the notion of $K$-flatness and the connection to $K$-semidefiniteness for symmetric $(n, n)$ matrices. This result can be generalized to continuous linear maps $A \in L(Y, Y^*)$ with an arbitrary real Hilbert space $(Y, \langle \cdot, \cdot \rangle)$.

**Definition 2.7.** Let $(Y, \langle \cdot, \cdot \rangle)$ be a real Hilbert space and let $K \subset Y$ be a nonempty set. Then $A \in L(Y, Y)$ is called *$K$-flat* if for all $y \in K$ the equation $\langle Ay, y \rangle = 0$ implies $(A + A^*)y = 0_Y$ with $A^* \colon Y \to Y$ the adjoint map to $A$.

For $K = \mathbb{R}^n_+$ a matrix $A \in \mathcal{S}^n$ which is copositive and for which it also holds that $\langle Ay, y \rangle = 0$ and $y \in \mathbb{R}^n_+$ implies $Ay = 0_n$, is also called copositive plus, see e.g. [26], Remark on p. 14.

**Theorem 2.8.** *Let $(Y, \langle\cdot,\cdot\rangle)$ be a real Hilbert space, $K \subset Y$ a nonempty convex set and $A \in L(Y,Y)$ a map. If $A$ is $K$-flat then $A$ or $-A$ is $K$-semidefinite.*

**Proof.** If $\langle Ay, y\rangle \geq 0$ or $\langle(-A)y, y\rangle = -\langle Ay, y\rangle \geq 0$ for all $y \in K$ the conclusion is shown. Therefore, we assume there exist $y^1,\ y^2 \in K$ with

$$q_A(y^1) > 0 > q_A(y^2) \tag{6}$$

for $q_A\colon Y \to \mathbb{R},\ q_A(y) = \langle Ay, y\rangle$ for all $y \in Y$. As $K$ is convex the line segment with endpoints $y^1,\ y^2$ belongs to $K$. As the map $q_A$ is continuous, there exists an interior point $y^0$ of this line segment with

$$q_A(y^0) = \langle Ay^0, y^0\rangle = 0.$$

Because $A$ is $K$-flat it follows $(A + A^*)y^0 = 0_Y$. Any point of the line segment can be written as $y = y^0 + \lambda \cdot v$ for some direction $v \in Y$ and a scalar $\lambda \in \mathbb{R}$. We have for all $y$ of the line segment

$$\begin{aligned} q_A(y) &= \langle A(y^0 + \lambda v), y^0 + \lambda v\rangle \\ &= \underbrace{\langle Ay^0, y^0\rangle}_{=0} + \lambda\langle\underbrace{(A + A^*)y^0}_{=0_Y}, v\rangle + \lambda^2\langle Av, v\rangle \\ &= \lambda^2\langle Av, v\rangle = \lambda^2 q_A(v). \end{aligned}$$

Consequently $q_A$ is equal to zero or has the sign of $q_A(v)$ along the entire line contradicting the assumption (6). This completes the proof. □

The converse implication is generally not true as the following example demonstrates.

**Example 2.9.** Let $Y = \mathbb{R}^2$ and $K = \{y \in \mathbb{R}^2 \mid y_1^2 - y_2^2 \geq 0,\ y_1 \geq 0\}$. Then the map $A\colon \mathbb{R}^2 \to \mathbb{R}^2$,

$$y \mapsto \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} y$$

is $K$-semidefinite, but not $K$-flat, as for the point $(3,3)^\top \in K$ it holds

$$y^\top Ay = 0 \ \text{ but } (A + A^\top)y = (6, -6)^\top \neq 0_2.$$

A weak converse implication under the additional assumption that $K$ is a convex cone with $\text{int}(K) \neq \emptyset$ is given in Lemma 5.6. There it is shown that for a $K$-semidefinite map $A$, $y \in K$ and $\langle Ay, y\rangle = 0$ implies $(A + A^*)y \in K^*$.

In this paper $K^*$ denotes the topological *dual set* to a set $K$, i.e.

$$K^* = \{y \in Y^* \mid \langle y, k\rangle \geq 0 \text{ for all } k \in K\}$$

(in the case of a convex cone $K$ the set $K^*$ equals the dual cone).

The following criteria for $K$-semidefiniteness is similar to the criteria given in Theorem 2.8.

**Theorem 2.10.** *Let $(Y, \langle\cdot,\cdot\rangle)$ be a real Hilbert space, $K \subset Y$ a nonempty convex set and let $A \in L(Y,Y)$ be given. If $y \in K$ and $\langle Ay, y\rangle = 0$ implies $(A + A^*)y \in K^*$ and there exists a $\hat{y} \in K$ with $\langle A\hat{y}, \hat{y}\rangle > 0$, then $A$ is $K$-semidefinite.*

**Proof.** We assume $A$ is not $K$-semidefinite and thus there exists a $\bar{y} \in K$ with $\langle A\bar{y}, \bar{y}\rangle < 0$. Then due to the continuity of the map $y \mapsto \langle Ay, y\rangle$ and as $K$ is a convex set there is a $\lambda \in ]0,1[$ such that for $y^\lambda = (1-\lambda)\bar{y} + \lambda\hat{y} \in K$ it is

$$\langle Ay^\lambda, y^\lambda\rangle = 0.$$

Then we have $(A + A^*)y^\lambda \in K^*$ and we conclude

$$\begin{aligned}
\langle Ay^\lambda, y^\lambda\rangle &= (1-\lambda)^2\langle A\bar{y}, \bar{y}\rangle + \lambda(1-\lambda)\langle (A+A^*)\bar{y}, \hat{y}\rangle + \lambda^2\langle A\hat{y}, \hat{y}\rangle \\
&= (1-\lambda)\,\langle (A+A^*)\bar{y}, y^\lambda\rangle - (1-\lambda)^2\langle A^*\bar{y}, \bar{y}\rangle + \lambda^2\underbrace{\langle A\hat{y}, \hat{y}\rangle}_{>0} \\
&> \underbrace{(1-\lambda)}_{>0}\langle \underbrace{(A+A^*)y^\lambda}_{\in K^*}, \underbrace{\bar{y}}_{\in K}\rangle - \underbrace{(1-\lambda)^2}_{>0}\underbrace{\langle A\bar{y}, \bar{y}\rangle}_{<0} \\
&> 0
\end{aligned}$$

in contradiction to $\langle Ay^\lambda, y^\lambda\rangle = 0$. □

For the case of copositive symmetric matrices this result can also be found in [50]. Next we come to some results concerning the eigenvalues and eigenvectors of a $K$-semidefinite linear map.

**Lemma 2.11.** *Let $(Y, \langle\cdot,\cdot\rangle)$ be a real Hilbert space. Let $K \subset Y$ be a given nonempty set and $A \in L(Y,Y)$ be $K$-semidefinite. Then for every eigenvector $y \in K$ of $A$ the correspondent eigenvalue $\lambda$ is nonnegative.*

**Proof.** Due to $y \in K$ and $y \neq 0_Y$ we have for the associated eigenvalue $\lambda$

$$0 \le \langle Ay, y\rangle = \langle \lambda y, y\rangle = \lambda\underbrace{\langle y, y\rangle}_{>0}$$

and thus $\lambda \ge 0$. □

Kaplan shows in [30] (see also [31]) that if a matrix $A \in \mathcal{S}^n$ is copositive then $A$ has no eigenvector $y \in \operatorname{int}(\mathbb{R}^n_+)$ with associated eigenvalue $\lambda < 0$.

**Lemma 2.12.** *Let $(Y, \langle\cdot,\cdot\rangle)$ be a real Hilbert space and $A \in L(Y,Y)$ arbitrarily chosen. Let $y^1, \ldots, y^k \in Y$ ($k \in \mathbb{N}$) be $k$ eigenvectors of $A$ with $\langle y^i, y^j\rangle \ge 0$ for all $i, j \in \{1, \ldots, k\}$, $i \neq j$, and with eigenvalues $\lambda_i \ge 0$. Let $K \subset Y$ be a nonempty set with*

$$K \subset \mathit{cone}(\mathit{convex\ hull}\{y^1, \ldots, y^k\}).$$

*Then $A$ is $K$-semidefinite.*

**Proof.** Let $y \in K$ be arbitrarily chosen. Then there exist $\beta, \mu_1, \ldots, \mu_k \ge 0$ with

$$\sum_{i=1}^k \mu_i = 1 \quad \text{and} \quad y = \beta\sum_{i=1}^k \mu_i y^i.$$

Because of the linearity of $A$ we conclude

$$\begin{aligned}\langle Ay, y\rangle &= \left\langle A\left(\beta \sum_{i=1}^{k} \mu_i y^i\right), \beta \sum_{i=1}^{k} \mu_i y^i \right\rangle \\ &= \beta^2 \sum_{i=1}^{k} \sum_{j=1}^{k} \mu_i \mu_j \langle A y^i, y^j \rangle \\ &= \beta^2 \sum_{i=1}^{k} \sum_{j=1}^{k} \underbrace{\mu_i \mu_j}_{\geq 0} \underbrace{\lambda_i}_{\geq 0} \underbrace{\langle y^i, y^j \rangle}_{\geq 0} \\ &\geq 0.\end{aligned}$$

Thus $A$ is $K$-semidefinite. □

**Example 2.13.** Let $Y = \mathbb{R}^3$ and consider the matrix

$$A = \begin{pmatrix} 2 & 0 & 4 \\ 0 & 6 & 0 \\ 4 & 0 & 2 \end{pmatrix} \in \mathcal{S}^3.$$

Let the nonempty set $K \subset \mathbb{R}^3$ be given by

$$K \subset \text{cone}\left(\text{convex hull}\left\{\begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}\right\}\right).$$

It is simple to see that the eigenvalues and eigenvectors of $A$ are $\lambda_1 = \lambda_2 = 6$ with $y = \alpha\,(1,0,1)^\top + \beta\,(0,1,0)^\top$ $((\alpha,\beta) \neq (0,0))$ and $\lambda_3 = -2$ with $y = \gamma\,(-1,0,1)^\top$ $(\gamma \neq 0)$. Then the matrix $A$ is $K$-semidefinite.

Now we consider the special case that the set $K$ is a cone. For the case of a total ordering introduced by the cone $K$ in $Y$ the following result can be of interest.

**Theorem 2.14.** *Let $\leq$ define a total ordering in $Y$, i.e. for all $y^1, y^2 \in Y$ it holds $y^1 \leq y^2$ or $y^2 \leq y^1$, and let the ordering cone $K$ be defined by*

$$K = \{y \in Y \mid y \geq 0_Y\}.$$

*Then it is*

$$C_L^K = C_L^Y,$$

*i.e. $K$-semidefiniteness equals $Y$-semidefiniteness.*

**Proof.** As the ordering $\leq$ is total we have for the cone $K$ the property $Y = K \cup (-K)$. With (5) we conclude

$$C_L^Y = C_L^{K\cup(-K)} = C_L^K.$$

□

As an application of this result to the matrix space $\mathcal{M}^n$ we get a characterization of positive semidefinite matrices.

**Example 2.15.** Consider the matrix space $\mathcal{M}^n$ and the lexicographical ordering in $\mathbb{R}^n$, which is total, defined by the cone

$$K = \{y \in \mathbb{R}^n \mid \exists k \in \{1, \dots, n\} \text{ with } y_i = 0 \text{ for } i < k \text{ and } y_k > 0\} \cup \{0_n\}.$$

Then, a matrix $A \in \mathcal{M}^n$ is positive semidefinite (compare Example 1.3 (c)) if and only if it is $K$-semidefinite.

For non-pointed cones $K$ we can show that semidefiniteness w. r. t. a special smaller cone already implies $K$-semidefiniteness. Recall that a cone $K \subset Y$ is called *pointed* if $K \cap (-K) = \{0_Y\}$. We need the following decomposition of a cone $K$.

**Lemma 2.16.** *Let $K \subset Y$ be a cone and $b \in Y^* \setminus \{0_{Y^*}\}$ be arbitrarily given. Then it is*

$$K = K_1 \cup K_2 \cup (-K_1)$$

*with cones*

$$K_1 := \{y \in K \cap (-K) \mid \langle b, y \rangle \geq 0\}$$

*and*

$$K_2 := \{y \in K \mid y \notin -K\} \cup \{0_Y\}.$$

**Proof.** Of course it is $K_1$, $K_2 \subset K$. For an arbitrary $y \in -K_1$ it is $-y \in K_1$ and therefore $-y \in K \cap (-K)$, thus $-y \in -K$ and hence $y \in K$. Then we have also $-K_1 \subset K$ and we conclude $K_1 \cup K_2 \cup (-K_1) \subset K$.

It remains to show $K \subset K_1 \cup K_2 \cup (-K_1)$. For that let $y \in K$ be arbitrarily chosen. If $y \notin -K$ we have immediately $y \in K_2$. Else, for $y \in K \cap (-K)$ we have to consider the case $\langle b, y \rangle \geq 0$, which induces $y \in K_1$, and the case $\langle b, y \rangle < 0$. In the latter case we conclude $\langle b, -y \rangle > 0$. Besides, because of $y \in K \cap (-K)$, we have $-y \in K \cap (-K)$, too. This leads to $-y \in K_1$ and thus $y \in -K_1$, which completes the proof. □

Using this decomposition we get the following result for a $K$-semidefinite map:

**Theorem 2.17.** *Let $K \subset Y$ be a cone. Then we get, using the decomposition of Lemma 2.16,*

$$C_L^K = C_L^{K_1 \cup K_2}.$$

**Proof.** Using the decomposition $K = K_1 \cup K_2 \cup (-K_1)$ of the cone $K$ and applying Lemma 2.1 and Lemma 2.5 we get

$$\begin{aligned} C_L^K &= C_L^{K_1} \cap C_L^{K_2} \cap \underbrace{C_L^{-K_1}}_{=C_L^{K_1}} \\ &= C_L^{K_1} \cap C_L^{K_2} \\ &= C_L^{K_1 \cup K_2} \end{aligned}$$

□

We illustrate this result with three examples:

**Example 2.18.** (a) Let $K \subset Y$ be a pointed cone. Then it is $K_1 = \{0_Y\}$ and $K_2 = K$. Thus the result of Theorem 2.17 is trivial.

(b) Let $K = Y$. Thus it is $K_2 = \{0_Y\}$ and $K_1 = \{y \in Y \mid \langle b, y\rangle \geq 0\}$ for an arbitrary $b \in Y^* \setminus \{0_{Y^*}\}$. Then the map $A \in L(Y, Y^*)$ is $K$-semidefinite if and only if it is $K_1$-semidefinite.

(c) Let $Y = \mathbb{R}^2$ and $K = \{y \in \mathbb{R}^2 \mid y_1 \geq 0 \ \vee \ y_2 \geq 0\}$. Then it is for $b = (1, -1)^\top$

$$K_1 = \{y \in \mathbb{R}^2 \mid y_1 \geq 0 \ \wedge \ y_2 \leq 0\}$$

and

$$K_2 = \{y \in \mathbb{R}^2 \mid y_1 > 0 \ \wedge \ y_2 > 0\} \cup \{0_2\}$$

(see Figure 2.1). Therefore, if $A \in \mathcal{M}^2$ is semidefinite w. r. t. the cone $K_1 \cup K_2$ then it is also $K$-semidefinite.

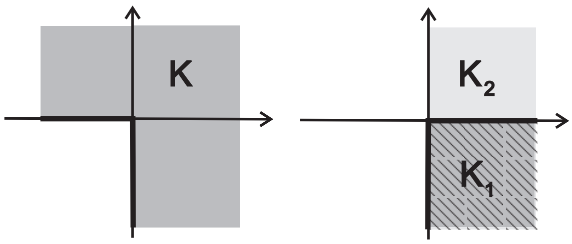


Figure 2.1: Cones $K$, $K_1$, and $K_2$ of Example 2.18 (c).

To show that a map is $K$-semidefinite, sometimes it is simpler to work with a suitable subset $B \subset K$ instead of the whole cone $K$. This result is given in the following theorem.

**Theorem 2.19.** *Let $K \subset Y$ be a cone and let $B$ be a subset of $K$ such that for any $y \in K$ there exists a $\lambda \in \mathbb{R}$ and a $b \in B$ with*

$$y = \lambda b. \tag{7}$$

*Then $A \in C_L^K$ if and only if*

$$\langle Ay, y\rangle \geq 0 \ \ \text{for all } y \in B. \tag{8}$$

**Proof.** Because of $B \subset K$ for any $A \in C_L^K$ (8) follows immediately (compare Lemma 2.2). For showing the converse implication we assume that (8) is satisfied. Let $y \in K$ be arbitrarily chosen. Then there is a $\lambda \in \mathbb{R}$ and a $b \in B$ with (7) and we get

$$\begin{aligned} \langle Ay, y\rangle &= \langle A(\lambda b), \lambda b\rangle \\ &= \underbrace{\lambda^2}_{\geq 0} \underbrace{\langle Ab, b\rangle}_{\geq 0} \geq 0. \end{aligned}$$

Thus $A \in C_L^K$. □

This result can of course be generalized to arbitrary nonempty sets $K \subset Y$. To give an example for Theorem 2.19 it is demonstrated in [52] that for the case $Y = \mathbb{R}^n$, $A \in \mathcal{S}^n$ it suffices to show for a $k \in \{1, \dots, n\}$

$$y^\top Ay \geq 0 \text{ for all } y \in \mathbb{R}^n_+ \text{ with } y_k = 1$$

to prove the copositivity, i.e. the $\mathbb{R}^n_+$-semidefiniteness, of $A$.

Recall that a nonempty convex subset $B$ of a convex cone $K \neq \{0_Y\}$ is called a *base* for $K$, if for every $y \in K \setminus \{0_Y\}$ there exists a $\lambda > 0$ and a $b \in B$ such that the representation in (7) is unique. Since a base $B$ fulfills the assumptions of Theorem 2.19, the result of this lemma also holds for a base. For example the set

$$B = \{y \in \mathbb{R}^n_+ \mid \|y\|_1 = 1\}$$

is a base for the cone $K = \mathbb{R}^n_+$. This set is used in [10] and already in [1] to check a symmetric matrix for $\mathbb{R}^n_+$-semidefiniteness. Each nontrivial convex cone with a base is pointed (see [27], Lemma 1.14). If the cone $K$ is not pointed but the linear space $Y$ is normed, we can use in Theorem 2.19 for example the set $B = \{y \in K \mid \|y\|_Y = 1\}$.

If the cone $K$ has a special structure the following result is of interest.

**Theorem 2.20.** *Let $Y$, $W$ be real separated locally convex linear spaces, with $W$ partially ordered by a convex cone $K_W$. Let the convex cone $K \subset Y$ be given by*

$$K = \{y \in Y \mid y = \overline{K}w,\ w \in K_W\}$$

*with a continuous linear map $\overline{K} : W \to Y$. Then $A \in L(Y, Y^*)$ is $K$-semidefinite if and only if $\overline{K}^* A\overline{K}$ is $K_W$-semidefinite, with $\overline{K}^* : Y^* \to W^*$ the adjoint map to $\overline{K}$.*

**Proof.** The map $A$ is $K$-semidefinite if and only if

$$\begin{aligned} & \langle Ay, y\rangle \geq 0 \quad \text{for all } y \in K \\ \Leftrightarrow \quad & \langle A\overline{K}w, \overline{K}w\rangle \geq 0 \quad \text{for all } w \in K_W \\ \Leftrightarrow \quad & \langle \overline{K}^* A\overline{K}w, w\rangle \geq 0 \quad \text{for all } w \in K_W \end{aligned}$$

being equivalent to $\overline{K}^* A\overline{K}$ $K_W$-semidefinite. □

This result is especially interesting in the finite dimensional case.

**Corollary 2.21.** *Let $Y = \mathbb{R}^n$, $A \in \mathcal{M}^n$ and $K \subset Y$ a polyhedral convex cone. Then there is a matrix $\overline{K} \in \mathbb{R}^{n\times s}$ with*

$$K = \{y \in \mathbb{R}^n \mid y = \overline{K}x,\ x \in \mathbb{R}^s_+\} \tag{9}$$

*and $A$ is $K$-semidefinite if and only if $\overline{K}^\top A\overline{K} \in \mathcal{M}^s_{++}$.*

**Proof.** According to [45], Prop 2.1.12 a convex cone is polyhedral if and only if it is finitely generated. Thus, there exists a matrix $\overline{K} \in \mathbb{R}^{n\times s}$ with (9). Then the assumptions of Theorem 2.20 are fulfilled and $K$-semidefiniteness of $A$ is equivalent to $\mathbb{R}^s_+$-semidefiniteness of $\overline{K}^\top A\overline{K}$. According to Example 1.3 (b) $\mathbb{R}^s_+$-semidefiniteness is equivalent to copositivity. □

Thus testing for $K$-semidefiniteness w. r. t. a polyhedral cone can be reduced to test for copositivity. For general non-polyhedral cones this is generally not true, but Loewy and Schneider show in [33] that semidefiniteness of a matrix $A \in \mathcal{S}^n$ w. r. t. the non-polyhedral ice-cream cone

$$K = \left\{ y \in \mathbb{R}^n \;\middle|\; \sqrt{y_1^2 + \ldots + y_{n-1}^2} \leq y_n \right\}$$

is equivalent to the existence of a scalar $\mu \geq 0$ such that the matrix

$$A - \mu \operatorname{diag}(-1, \ldots, -1, 1)$$

is positive semidefinite, i.e. $\mathbb{R}^n$-semidefinite.

The following theorem shows a certain invariance of set-semidefinite maps under transformation with invertible maps. It can easily be concluded from Theorem 2.20 but it can also be proved directly very easily and thus the proof is omitted here.

**Theorem 2.22.** *Let $Y$ be a real separated locally convex linear space, $B\colon Y \to Y$ a continuous linear invertible map and let $K$ be a nonempty subset of $Y$. Then a map $A \in L(Y, Y^*)$ is $K$-semidefinite if and only if $B^*AB$ is $B^{-1}K$-semidefinite. Here $B^*\colon Y^* \to Y^*$ denotes the adjoint of $B$ and we set $B^{-1}K := \{B^{-1}k \mid k \in K\}$.*

**Remark 2.23.** If $Y = \mathbb{R}^n$ and $B$ is an orthogonal matrix, the classical Jacobi method (for the determination of all eigenvalues of a matrix) working with orthogonal transformations, produces set-semidefinite matrices if the original matrix is set-semidefinite. After a sufficiently large number of iterations one then obtains nearly a diagonal matrix which is, by Theorem 2.22, semidefinite with respect to the transformed cones.

The next corollary shows for special polyhedral cones $K$ in $\mathbb{R}^n$ that the $K$-semidefiniteness reduces to copositivity. It is very similar to the result in Corollary 2.21 but works with a different representation of the polyhedral cone.

**Corollary 2.24.** *Let $Y = \mathbb{R}^n$ and an invertible matrix $F \in \mathcal{M}^n$ be given. If*

$$K = \{y \in \mathbb{R}^n \mid Fy \geq 0_{\mathbb{R}^n}\}$$

*where $\geq$ has to be understood in a componentwise sense, then a matrix $A \in \mathcal{M}^n$ is $K$-semidefinite if and only if $(F^\top)^{-1}AF^{-1}$ is copositive, i.e. $\mathbb{R}^n_+$-semidefinite.*

**Proof.** This result follows from Theorem 2.22, if we set $B := F^{-1}$ and if we notice that

$$B^{-1}K = FK = \{Fy \mid y \in \mathbb{R}^n,\ Fy \geq 0_{\mathbb{R}^n}\} = \mathbb{R}^n_+.$$

The result can also be concluded from Corollary 2.21 for $\overline{K} = F^{-1}$ and $s = n$. □

In the finite dimensional case $Y = \mathbb{R}^n$ we present some further special results on $K$-semidefinite matrices.

**Lemma 2.25.** *Let $Y = \mathbb{R}^n$ and $K = \mathbb{R}^n_+$. If $A = (a_{ij})_{i,j\in\{1,\ldots,n\}} \in \mathcal{M}^n$ is $K$-semidefinite, then it is*

$$a_{ii} \geq 0 \quad \textit{for all } i = 1 \ldots, n.$$

**Proof.** Let $i \in \{1, \ldots, n\}$ be arbitrarily chosen. Then we have for the unit vector $e^i \in K$

$$0 \leq e^{i\top} A e^i = a_{ii}.$$

□

So, every copositive matrix in $\mathcal{M}^n$ has nonnegative diagonal elements. This is of course also true for arbitrary $K$-semidefinite matrices with $e^i \in K$ for all $i = 1, \ldots, n$, and therefore also for positive semidefinite matrices (see also [41]). The result of Lemma 2.25 can also be found in [4], [12], [25], [26] and [50] for symmetric matrices. Bomze presents in [4] even an algorithm with which one can test whether a tridiagonal matrix is copositive or not only by checking the signs of the entries. We illustrate now the results of Corollary 2.24 and Lemma 2.25 with an example:

**Example 2.26.** Let $Y = \mathbb{R}^n$ and we choose the matrices

$$A = \begin{pmatrix} -23 & 7 & 34 \\ 7 & 5 & 11 \\ 34 & 11 & 13 \end{pmatrix} \quad \text{and} \quad F = \begin{pmatrix} 2 & 0 & -1 \\ 5 & 1 & 0 \\ 0 & 1 & 3 \end{pmatrix}.$$

The cone $K$ is defined as in Corollary 2.24. Then the matrix $A$ is not $K$-semidefinite because the matrix

$$(F^\top)^{-1} A F^{-1} = \begin{pmatrix} -17 & 6 & -5 \\ 6 & -3 & 4 \\ -5 & 4 & 0 \end{pmatrix}$$

has a negative diagonal element and is thus, according to Lemma 2.25, not copositive.

The result of Lemma 2.25 also implies the following:

**Lemma 2.27.** *Let $A = (a_{ij})_{i,j\in\{1,\ldots,n\}} \in \mathcal{S}^n$ be a diagonal matrix. Then $A$ is copositive if and only if $A$ is positive semidefinite.*

**Proof.** If $A$ is copositive then by Example 1.3 (b) $A$ is $\mathbb{R}^n_+$-semidefinite and we have according to Lemma 2.25

$$a_{ii} \geq 0 \quad \text{for all } i = 1 \ldots, n.$$

Since $A$ is a diagonal matrix, all eigenvalues of $A$ are diagonal entries and thus nonnegative. Therefore, $A$ is positive semidefinite.

Now let $A$ be positive semidefinite. Then of course (compare Lemma 2.2) $A$ is $\mathbb{R}^n_+$-semidefinite, too. □

In view of the optimality conditions in the following section the dual cone and the interior of the cone $C_L^K$ are of interest. For the finite dimensional case $Y = \mathbb{R}^n$ and the matrix space $\mathcal{S}^n$ we summarize some results on dual cones. Proofs can be found in [28] for a closed convex cone $K$, in [47] for a given nonempty set $K \subset \mathbb{R}^n$ and in [44] for a polyhedral cone $K \subset \mathbb{R}^n$.

**Lemma 2.28.**

(*a*) *Let $K \subset \mathbb{R}^n$ be a nonempty given set. Then it is*

$$(C^K_{\mathcal{S}^n})^* = cl\,cone\big(convex\ hull\{\, xx^\top \mid x \in K\}\big).$$

(*b*) *Let $K \subset \mathbb{R}^n$ be a closed convex cone. Then it is*

$$(C^K_{\mathcal{S}^n})^* = \textit{convex hull}\{\, xx^\top \mid x \in K\}$$

*and $(C^K_{\mathcal{S}^n})^*$ is closed.*

(*c*) *Let $K = \mathbb{R}^n$, then $C^K_{\mathcal{S}^n} = \mathcal{S}^n_+$ and it is $(C^K_{\mathcal{S}^n})^* = C^K_{\mathcal{S}^n}$, i.e. $\mathcal{S}^n_+$ is self-dual.*

In the literature elements of the dual cone

$$\left(C^{\mathbb{R}^n_+}_{\mathcal{S}^n}\right)^* = \left(\mathcal{S}^n_{++}\right)^* = \text{convex hull}\{\, xx^\top \mid x \in \mathbb{R}^n_+\}$$

are called completely positive matrices.

For determining the interior of the cone $C^K_L$ let $\hat{C}^K_L$ denote the set of strict $K$-semidefinite maps, i.e. let

$$\hat{C}^K_L := \{A \in L(Y, Y^*) \mid \langle Ay, y\rangle > 0 \text{ for all } y \in K \setminus \{0_Y\}\},$$

for an arbitrary nonempty set $K \subset Y$. Of course it is $\hat{C}^K_L \subset C^K_L$. If the set $\hat{C}^K_L$ is nonempty for a reflexive Banach space $Y$, is under some additional assumptions the interior of the cone $C^K_L$.

**Theorem 2.29.** *Let $(Y, \|\cdot\|_Y)$ be a real reflexive Banach space, let $K$ be a closed convex cone, and let the set $\hat{C}^K_L$ be nonempty. If there are linear functionals $\ell_1, \ldots, \ell_k \in Y^*$ and real numbers $\alpha_1 \neq 0, \ldots, \alpha_k \neq 0$ for some $k \in \mathbb{N}$ so that*

$$K = \bigcup_{i=1}^{k} \textit{cone}\,(K_i)$$

*with*

$$K_i := \{y \in K \cap B \mid \ell_i(y) = \alpha_i\} \quad \textit{for all } i \in \{1, \ldots, k\}$$

*($B$ denotes the closed unit ball) and for every $A \in \hat{C}^K_L$ the quadratic form $\langle A\cdot, \cdot\rangle$ is weakly lower semicontinuous, then*

$$\textit{int}(C^K_L) = \hat{C}^K_L.$$

**Proof.** For the proof of the inclusion $\hat{C}^K_L \subset \text{int}(C^K_L)$ we fix an arbitrary map $A \in \hat{C}^K_L$. Under the assumptions of the theorem the cone $K$ is weakly closed and the closed unit ball $B$ is weakly compact (see [27], Lemma 1.41) and thus the set $K \cap B$ is also weakly compact. Consequently, for every $i \in \{1, \ldots, k\}$ the set $K_i$ is weakly compact as well. Further it is $K_i \subset K \setminus \{0_Y\}$ for $i = 1, \ldots, k$. Since the quadratic form $\langle A\cdot, \cdot\rangle$ is weakly lower semicontinuous, by Thm. 2.3 in [28] there exists a scalar $\varepsilon > 0$ with

$$\varepsilon \leq \min_{y \in K_i} \langle Ay, y\rangle \quad \text{for all } i \in \{1, \ldots, k\}.$$

Let $\mathcal{N}_\varepsilon(A) := \{D \in L(Y, Y^*) \mid \|D - A\|_{L(Y,Y^*)} < \varepsilon\}$ denote the $\varepsilon$-neighborhood of $A$ and let a map $D \in \mathcal{N}_\varepsilon(A)$ be arbitrarily chosen. For an arbitrary $y \in K\setminus\{0_Y\}$ we have $y \in \text{cone}(K_i)\setminus\{0_Y\}$ for some $i \in \{1, \ldots, k\}$ or

$$y = \lambda_i k_i \quad \text{for some } \lambda_i > 0 \text{ and some } k_i \in K_i.$$

Then we obtain because of $K_i \subset B$

$$\begin{aligned}
\frac{1}{\lambda_i^2}\langle Dy, y\rangle &= \langle Dk_i, k_i\rangle \\
&= \underbrace{\langle Ak_i, k_i\rangle}_{\geq\varepsilon} + \langle (D-A)k_i, k_i\rangle \\
&\geq \varepsilon - |\,\langle\,(D-A)k_i, k_i\,\rangle\,| \\
&\geq \varepsilon - \|\,(D-A)k_i\,\|_{Y^*}\cdot\underbrace{\|k_i\|_Y}_{\leq 1} \\
&\geq \varepsilon - \underbrace{\|\,D-A\,\|_{L(Y,Y^*)}}_{<\varepsilon}\cdot\underbrace{\|k_i\|_Y}_{\leq 1} \\
&\geq 0
\end{aligned}$$

and thus $D \in C_L^K$. Therefore we have shown $\mathcal{N}_\varepsilon(A) \subset C_L^K$ and thus $A \in \operatorname{int}(C_L^K)$. As $A \in \hat{C}_L^K$ is arbitrarily chosen we have $\hat{C}_L^K \subset \operatorname{int}(C_L^K)$.

It remains to show $\operatorname{int}(C_L^K) \subset \hat{C}_L^K$. Let $A \in \operatorname{int}(C_L^K)$ be arbitrarily chosen. Then there exists a $\varepsilon$-neighborhood $\mathcal{N}_\varepsilon(A)$ of $A$ with $\mathcal{N}_\varepsilon(A) \subset C_L^K$. For arbitrarily chosen $D \in \hat{C}_L^K$ there exists a $\lambda > 0$ with

$$D^\lambda := A + \lambda\,(A - D) \in \mathcal{N}_\varepsilon(A).$$

Consequently, we have

$$A = \frac{1}{1+\lambda}D^\lambda + \frac{\lambda}{1+\lambda}D$$

and we obtain for all $y \in K \setminus \{0_Y\}$

$$\langle Ay, y\rangle = \frac{1}{1+\lambda}\underbrace{\langle D^\lambda y, y\rangle}_{\geq 0} + \frac{\lambda}{1+\lambda}\underbrace{\langle Dy, y\rangle}_{>0} > 0,$$

i.e. $A \in \hat{C}_L^K$. □

If the cone $K$ is a subset of a finite-dimensional real Banach space the quadratic form $\langle A\cdot, \cdot\rangle$ for $A \in \hat{C}_L^K$ is weakly lower semicontinuous. For a real Hilbert space $(Y, \langle\cdot,\cdot\rangle)$ any quadratic form $\langle A\cdot,\cdot\rangle$ for which $A$ is $Y$-semidefinite is weakly lower semicontinuous (see [19], Theorem 2.1). Besides, if $(Y, \langle\cdot,\cdot\rangle)$ is a real Hilbert space then $\hat{C}_L^K \neq \emptyset$ as the identity map is always an element of $\hat{C}_L^K$ for any set $K \subset Y$. Notice that in the case of $Y = \mathbb{R}^n$ the convex cone $K = \mathbb{R}^n_+$ fulfills the assumptions of Theorem 2.29 because of the equality

$$K = \operatorname{cone}\underbrace{\left\{y \in \mathbb{R}^n_+ \;\middle|\; \sum_{i=1}^n y_i = 1\right\}}_{=:K_1}.$$

Using various subsets $K_1, \ldots, K_k$ any closed convex nontrivial cone $K$ in $\mathbb{R}^n$ fulfills the assumptions of the previous theorem.

**Corollary 2.30.** *Let $Y = \mathbb{R}^n$ and $K \subset \mathbb{R}^n$ be a closed convex nontrivial cone. Then*

$$int(C_L^K) = \hat{C}_L^K.$$

**Proof.** As $Y$ is finite dimensional the set $\hat{C}_L^K$ is nonempty and the quadratic forms $\langle A\cdot, \cdot\rangle$ are weakly lower semicontinuous for all $A \in \hat{C}_L^K$. Consider the $l_1$ norm in $\mathbb{R}^n$, i.e. the closed unit ball $B$ equals $\{y \in \mathbb{R}^n \mid \sum_{i=1}^n |y_i| \leq 1\}$. For $l_1, \ldots, l_{2^n} \in \mathbb{R}^n$ denoting all vectors of the form $(\pm 1, \ldots, \pm 1) \in \mathbb{R}^n$ we define

$$K_i := \{y \in K \cap B \mid l_i(y) = 1\} \text{ for all } i \in \{1, \ldots, 2^n\}$$

(notice that the cone generated by the set $K_i$ equals the intersection of the $i$-th orthant in $\mathbb{R}^n$ with the cone $K$). If we set $I := \{i \in \{1, \ldots, 2^n\} \mid K_i \neq \emptyset\}$, we then obtain

$$K = \bigcup_{i \in I} \text{cone}\,(K_i)$$

and the assertion follows from Theorem 2.29. □

The result of Theorem 2.29 is a generalization of a result in [9]. There it is shown that in the finite dimensional case $Y = \mathbb{R}^n$ for $K = \mathbb{R}^n_+$ it holds

$$\text{int}(C_{\mathcal{S}^n}^K) = \{A \in \mathcal{S}^n \mid y^\top Ay > 0 \text{ for all } y \in \mathbb{R}^n_+ \setminus \{0_{\mathbb{R}^n}\}\},$$

i.e. the interior elements of the cone of copositive matrices $\mathcal{S}^n_{++}$ are exactly the strict copositive matrices.

As a consequence of Corollary 2.30 we have the following results for $Y = \mathbb{R}^n$ and the matrix space $\mathcal{S}^n$, but direct proofs can be found in [28].

**Corollary 2.31.** *Let $K \subset \mathbb{R}^n$ be a convex cone. Then*

$$\{A \in \mathcal{S}^n \mid A \textit{ is positive definite}\} \subset int(C_{\mathcal{S}^n}^K).$$

*For $K = \mathbb{R}^n$ equality holds, i.e.*

$$int(C_{\mathcal{S}^n}^K) = \{A \in \mathcal{S}^n \mid A \textit{ is positive definite}\}.$$

## 3. Optimality Conditions

In this section we investigate the set-semidefinite optimization problem (1) and formulate necessary and sufficient optimality conditions for this problem. For an arbitrary nonempty set $K \subset Y$ we consider the $K$-semidefinite cone $C_L^K$ inducing the partial ordering $\preccurlyeq$ in the definition of the constraint set $S$ in (2). To be more concrete, under Assumption 1.1 we examine the set-semidefinite optimization problem

$$\begin{array}{c} \min\ f(x) \quad \text{subject to the constraint} \\ -G(x) \in C_L^K, \quad x \in X \end{array} \tag{10}$$

with the constraint set $S = \{x \in X \mid G(x) \in -C_L^K\}$. The following theorem gives a necessary condition for a minimal solution of problem (10).

**Theorem 3.1.** *Let the set-semidefinite optimization problem* (10) *be given under Assumption* 1.1*, and let $X$ be a real Banach space and $Y$ and $Z$ real normed spaces. Let the ordering cones $C_Z$ and $C_L^K$ have a nonempty interior. Let $\bar{x} \in S$ be a weakly minimal solution of problem* (10)*. Moreover, let $f$ and $G$ be Fréchet differentiable at $\bar{x}$. Then there are continuous linear functionals $t \in C_Z^*$, $U \in (C_L^K)^*$ with $(t, U) \neq 0_{Z^* \times L(Y,Y^*)^*}$ so that*

$$t \circ f'(\bar{x}) + U \circ G'(\bar{x}) = 0_{X^*} \tag{11}$$

*and*

$$(U \circ G)(\bar{x}) = 0. \tag{12}$$

*If, in addition to the given assumptions*

$$G'(\bar{x})(X) + cone\left(C_L^K + \{G(\bar{x})\}\right) = L(Y, Y^*) \tag{13}$$

*then $t \neq 0_{Z^*}$.*

**Proof.** The first part of this theorem follows from a general Lagrange multiplier rule given in [27], Thm. 7.4, if we notice that the superset used in [27] equals the whole space $X$ in our case. For the proof of the second part of this theorem let the condition (13) be satisfied. Now assume that $t = 0_{Z^*}$. Then for an arbitrary element $A \in L(Y, Y^*)$ there is a nonnegative number $\alpha$, a vector $x \in X$ and a map $D \in C_L^K$ with

$$A = G'(\bar{x})(x) + \alpha(D + G(\bar{x})).$$

Then we obtain with the equations (11) and (12) and the positivity of $U$

$$\langle U, A\rangle = \underbrace{(U \circ G'(\bar{x}))\,(x)}_{=0} + \underbrace{\alpha\langle U, D\rangle}_{\geq 0} + \underbrace{\alpha(U \circ G)(\bar{x})}_{=0} \geq 0.$$

Consequently, we have $U = 0_{L(Y,Y^*)^*}$. But this contradicts the assertion that $(t, U) \neq 0_{Z^* \times L(Y,Y^*)^*}$. □

The necessary optimality condition given in Theorem 3.1 generalizes the well-known Lagrange multiplier rule. The regularity assumption (13) is the known Kurcyusz-Robinson-Zowe regularity condition (see [28], p. 114). It does not use the interior of the ordering cone and, therefore, it is more general than the known Slater condition.

**Corollary 3.2.** *Let the assumptions of Theorem* 3.1 *be satisfied and let $C_Z \neq Z$. If $\bar{x} \in S$ is a minimal solution of problem* (10)*, then we obtain the necessary condition in Theorem* 3.1*.*

**Proof.** Under the additional assumption $C_Z \neq Z$ every minimal solution of problem (10) is also a weakly minimal solution of (10) (see [27], Lemma 4.14). Hence, we get the same assertion as in Theorem 3.1. □

Next we specialize the previous results to the finite dimensional case and we present KKT-type conditions.

**Theorem 3.3.** *Let $f\colon \mathbb{R}^m \to \mathbb{R}^k$ and $G\colon \mathbb{R}^m \to \mathcal{M}^n$ be given functions, let $C_{\mathbb{R}^k}$ and $C^K_{\mathcal{M}^n}$ be ordering cones with a nonempty interior, and let $C_{\mathbb{R}^k}$ be pointed. Let $\bar{x} \in S$ be a weakly minimal solution of (10). Let $f$ be differentiable at $\bar{x}$ and let $G$ be elementwise differentiable at $\bar{x}$. Then there is a vector $t \in C^*_{\mathbb{R}^k}$ and a matrix $U \in (C^K_{\mathcal{M}^n})^*$ with $(t, U) \neq (0_{\mathbb{R}^k}, 0_{\mathcal{M}^n})$ so that*

$$\sum_{i=1}^{k} t_i \nabla f_i(\bar{x}) + \begin{pmatrix} \langle U, G_{x_1}(\bar{x})\rangle \\ \vdots \\ \langle U, G_{x_m}(\bar{x})\rangle \end{pmatrix} = 0_{\mathbb{R}^m} \tag{14}$$

*and*

$$\langle U, G(\bar{x})\rangle = 0. \tag{15}$$

*If, in addition to the above assumptions*

$$G'(\bar{x})(\mathbb{R}^m) + cone\left(C^K_{\mathcal{M}^n} + \{G(\bar{x})\}\right) = \mathcal{M}^n, \tag{16}$$

*then $t \neq 0_{\mathbb{R}^k}$. Here we use the notation*

$$G_{x_i} := \begin{pmatrix} \frac{\partial}{\partial x_i} G_{11} & \dots & \frac{\partial}{\partial x_i} G_{1n} \\ \vdots & & \vdots \\ \frac{\partial}{\partial x_i} G_{n1} & \dots & \frac{\partial}{\partial x_i} G_{nn} \end{pmatrix} \quad \textit{for all } i \in \{1, \dots, m\}$$

*and*

$$G'(\bar{x})(h) = \sum_{i=1}^{m} G_{x_i}(\bar{x}) h_i \quad \textit{for all } h \in \mathbb{R}^m.$$

**Proof.** This proof follows the proof of Theorem 7.8 in [28]. In this special case the Fréchet derivative of $G$ at $\bar{x}$ is given by

$$G'(\bar{x})(h) = \sum_{i=1}^{m} G_{x_i}(\bar{x}) h_i \quad \text{for all } h \in \mathbb{R}^m$$

(see [28], Lemma 7.7) and, therefore, we obtain for $U \in (C^K_{\mathcal{M}^n})^*$

$$(U \circ G'(\bar{x}))(h) = \sum_{i=1}^{m} \langle U, G_{x_i}(\bar{x})\rangle h_i \quad \text{for all } h \in \mathbb{R}^m.$$

Moreover, for the Fréchet derivative of $f$ at $\bar{x}$ and $t \in C^*_{\mathbb{R}^k}$ we get

$$t \circ f'(\bar{x}) = \sum_{i=1}^{k} t_i \nabla f_i(\bar{x}).$$

So, the equation (14) follows from equation (11) and the equation (15) is equivalent to the equation (12) in this special case. The regularity assumption (16) is equivalent to the Kurcyusz-Robinson-Zowe regularity condition (13). □

Under generalized convexity assumptions the necessary optimality condition in Theorem 3.1 is also a sufficient optimality condition. Here we recall the concept of $\tilde{C}$-quasiconvexity (see also [28], Def. 5.12).

**Definition 3.4.** Let $T$ be a nonempty subset of a real linear space $X$, and let $\tilde{C}$ be a nonempty subset of a real normed space $V$. Let $h\colon T \to V$ be a given map having a directional derivative at some $\bar{x} \in T$ in every direction $x - \bar{x}$ with arbitrary $x \in T$. The map $h$ is called *$\tilde{C}$-quasiconvex* at $\bar{x}$, if for all $x \in T$

$$h(x) - h(\bar{x}) \in \tilde{C} \;\Rightarrow\; h'(\bar{x})(x - \bar{x}) \in \tilde{C}.$$

Now we present the sufficient optimality condition.

**Theorem 3.5.** *Let the set-semidefinite optimization problem* (10) *be given under Assumption* 1.1*, let $Y$ and $Z$ be real normed spaces, and let the ordering cone $C_Z$ be nontrivial. Let $f$ and $G$ have a directional derivative at some $\bar{x} \in S$. Moreover, let the map $(f, G)\colon X \to Z \times L(Y, Y^*)$ be $\tilde{C}$-quasiconvex at $\bar{x}$ with*

$$\tilde{C} := (-C_Z \setminus \{0_Z\}) \times \left(-C_L^K + lin(G(\bar{x}))\right).$$

*If there are continuous linear functionals $t \in C_Z^{\#}$, $U \in (C_L^K)^*$ so that*

$$t \circ f'(\bar{x}) + U \circ G'(\bar{x}) = 0_{X^*} \tag{17}$$

*and*

$$(U \circ G)(\bar{x}) = 0, \tag{18}$$

*then $\bar{x}$ is a minimal solution of problem* (10).

*Here*

$$C_Z^{\#} := \{t \in Z^* \mid \langle t, c\rangle > 0 \text{ for all } c \in C_Z \setminus \{0_Z\}\} \tag{19}$$

*denotes the so-called quasi-interior of the dual cone $C_Z^*$ and $lin(G(\bar{x}))$ denotes the one-dimensional linear space spanned by $G(\bar{x})$.*

**Proof.** Assume that there is a vector $x \in X$ with

$$(f'(\bar{x})(x - \bar{x}) \,,\, G'(\bar{x})(x - \bar{x})) \in \tilde{C}.$$

Then we have

$$f'(\bar{x})(x - \bar{x}) \in -C_Z \setminus \{0_Z\}$$

and

$$G'(\bar{x})(x - \bar{x}) \in -C_L^K + \mathrm{lin}(G(\bar{x})),$$

and we obtain with the definition of the quasi-interior $C_Z^{\#}$ and with the equation (18) for some $\alpha \in \mathbb{R}$

$$(t \circ f'(\bar{x}) + U \circ G'(\bar{x}))(x - \bar{x}) < \alpha\langle U, G(\bar{x})\rangle = 0.$$

This contradicts the equation (17). Hence, we have for all $x \in X$

$$(f'(\bar{x})(x - \bar{x}) \,,\, G'(\bar{x})(x - \bar{x})) \notin \tilde{C}$$

and because $(f,G)$ is $\tilde{C}$-quasiconvex this implies

$$(f(x)-f(\bar{x})\,,\,G(x)-G(\bar{x}))\notin\tilde{C}\quad\text{for all } x\in X.$$

This means that there is no $x\in X$ with

$$f(x)-f(\bar{x})\in -C_Z\setminus\{0_Z\}$$

and

$$\begin{aligned} G(x)-G(\bar{x}) &\in\; -C_L^K+\mathrm{lin}(G(\bar{x}))\\ \Leftrightarrow\qquad G(x) &\in \{G(\bar{x})\}-C_L^K+\mathrm{lin}(G(\bar{x}))\\ &\subset\; -C_L^K-C_L^K+\mathrm{lin}(G(\bar{x}))\\ &=\; -C_L^K+\mathrm{lin}(G(\bar{x}))\end{aligned}$$

and thus in particular with $G(x)\in -C_L^K$. Consequently, there is no feasible vector $x\in S$ with

$$f(x)\in\{f(\bar{x})\}-C_Z\setminus\{0_Z\}$$

or

$$(f(\bar{x})-C_Z)\cap f(S)=\{f(\bar{x})\}.$$

This means that $\bar{x}$ is a minimal solution of problem (10). □

A similar result can also be shown for the weak minimality notion (see [27], Thm. 7.20 and Cor. 7.21). Notice that for the sufficient optimality condition we use a stronger assumption on $t$. Here $t$ should be in the quasi-interior of the dual cone whereas in Theorem 3.1 the functional $t$ is an element of the dual cone. This theoretical gap results from the fact that there is no complete characterization of minimal solutions, if one works with linear scalarization (for instance, see Thm. 5.18 and the following discussion in [27]).

Finally, we specialize the preceding sufficient optimality condition to the finite dimensional case. Recall, that a differentiable function $h\colon T\to\mathbb{R}$ with $T\subset\mathbb{R}^m$ is called *pseudoconvex* at some $\bar{x}\in T$ if for all $x\in T$

$$\nabla h(\bar{x})^\top(x-\bar{x})\geq 0\;\Rightarrow\; h(x)-h(\bar{x})\geq 0.$$

**Theorem 3.6.** *Let $f\colon\mathbb{R}^m\to\mathbb{R}^k$ and $G\colon\mathbb{R}^m\to\mathcal{M}^n$ be given functions, let $C_{\mathbb{R}^k}\neq\{0_{\mathbb{R}^k}\}$ and $C^K_{\mathcal{M}^n}$ be ordering cones, and let $C_{\mathbb{R}^k}$ be pointed. Let $f$ be differentiable at some $\bar{x}\in S$ and let $G$ be elementwise differentiable at $\bar{x}$. Assume that $f_1,\ldots,f_k$ are pseudoconvex at $\bar{x}$ and the map $G$ is $\big(-C^K_{\mathcal{M}^n}+lin(G(\bar{x}))\big)$-quasiconvex at $\bar{x}$. If there is a vector $t\in C^{\#}_{\mathbb{R}^k}$ and a matrix $U\in(C^K_{\mathcal{M}^n})^*$ so that*

$$\sum_{i=1}^k t_i\nabla f_i(\bar{x})+\begin{pmatrix}\langle U\,,\,G_{x_1}(\bar{x})\rangle\\ \vdots\\ \langle U\,,\,G_{x_m}(\bar{x})\rangle\end{pmatrix}=0_{\mathbb{R}^m}\tag{20}$$

*and*

$$\langle U\,,\,G(\bar{x})\rangle=0,\tag{21}$$

*then $\bar{x}$ is a minimal solution of problem* (10).

**Proof.** By Lemma 7.7 in [28] the equation (20) equals

$$\nabla(t^\top f(\bar{x})) + U \circ G'(\bar{x}) = 0_{\mathbb{R}^m}$$

for the Fréchet derivative $G'(\bar{x})$ of $G$ at $\bar{x}$. With Lemma 5.16 and Cor. 5.15 in [28] we conclude that $\bar{x}$ then is a minimal solution of the problem $\min_{x\in S}\ t^\top f(x)$. By a scalarization result (Thm. 5.18 (b) in [27]) $\bar{x}$ is a minimal solution of the set-semidefinite optimization problem (10). □

For the special ordering cone $C_{\mathbb{R}^k} = \mathbb{R}^k_+$ the quasi-interior of the dual cone is given as

$$C^{\#}_{\mathbb{R}^k} = \left\{ t \in \mathbb{R}^k \mid t_i > 0 \text{ for all } i \in \{1,\ldots,k\} \right\}.$$

In this case one can set $t_1 = \ldots = t_k = 1$ in the equation (20) and then one can try to determine a matrix $U \in (C^K_{\mathcal{M}^n})^*$ and a vector $\bar{x} \in S$ satisfying the KKT-system (20) and (21).

We complete our investigations with a simple example.

**Example 3.7.** We investigate the set-semidefinite optimization problem (10) with $X = Z = \mathbb{R}^2$, $Y = \mathbb{R}^3$, and the ordering cones $C_Z = \mathbb{R}^2_+$ and $C^K_{\mathcal{S}^3} = \mathcal{S}^3_+$. We consider the objective function $f\colon \mathbb{R}^2 \to \mathbb{R}^2$ with

$$f(x_1,x_2) = \begin{pmatrix} (x_1+1)^2 + (x_2-2)^2 \\ x_1 + 4x_2 \end{pmatrix} \quad \text{for all } (x_1,x_2) \in \mathbb{R}^2$$

and the constraint function $G\colon \mathbb{R}^2 \to \mathcal{S}^3$ with

$$G(x_1,x_2) = -\begin{pmatrix} x_1 & 1 & 0 \\ 1 & x_2 & 0 \\ 0 & 0 & x_1+1 \end{pmatrix} \quad \text{for all } (x_1,x_2) \in \mathbb{R}^2.$$

Then the constraint set is given by

$$S = \{x \in \mathbb{R}^2 \mid -G(x) \text{ positive semidefinite }\}.$$

The matrix $-G(x)$ has the eigenvalues

$$\lambda_{1/2} = \frac{x_1+x_2}{2} \pm \sqrt{\frac{(x_1+x_2)^2}{4} - x_1x_2 + 1}$$

and $\lambda_3 = x_1 + 1$ being nonnegative if and only if

$$x_1x_2 \geq 1, \quad x_1 > 0, \quad x_2 > 0.$$

Therefore, we obtain the constraint set

$$S = \{x \in \mathbb{R}^2_+ \mid x_1x_2 \geq 1\}.$$

We want to determine a minimal solution $\bar{x}$ of the optimization problem using Theorem 3.6. With the derivatives $\nabla f_1(\bar{x}) = 2(\bar{x}_1+1, \bar{x}_2-2)$, $\nabla f_2(\bar{x}) = (1,4)$,

$$G_{x_1}(\bar{x}) = -\begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad \text{and} \quad G_{x_2}(\bar{x}) = -\begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

we get the KKT-conditions (20) and (21) as

$$2\,t_1 \begin{pmatrix} \bar{x}_1 + 1 \\ \bar{x}_2 - 2 \end{pmatrix} + t_2 \begin{pmatrix} 1 \\ 4 \end{pmatrix} + \begin{pmatrix} -U_{11} - U_{33} \\ -U_{22} \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \tag{22}$$

and

$$U_{11}\,\bar{x}_1 + 2\,U_{21} + U_{22}\,\bar{x}_2 + U_{33}\,(\bar{x}_1 + 1) = 0 \tag{23}$$

where $t_1, t_2$ are positive real numbers and $U_{ij}$ are the coefficients of a symmetric positive semidefinite $(3,3)$ matrix $U$. Remember, that according to Lemma 2.28 (c) the cone $\mathcal{S}^3_+$ is self-dual. If we choose the parameters $t_1 = 1$, $t_2 = 2$, and the positive semidefinite matrix

$$U = \begin{pmatrix} 6 & -6 & 0 \\ -6 & 6 & 0 \\ 0 & 0 & 0 \end{pmatrix},$$

then the equation (22) is equivalent to $\bar{x}_1 = \bar{x}_2 = 1$ and the equation (23) is also satisfied for these values. Since the matrix

$$-G(\bar{x}_1, \bar{x}_2) = \begin{pmatrix} 1 & 1 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & 2 \end{pmatrix}$$

is positive semidefinite, the pair $(\bar{x}_1, \bar{x}_2) = (1,1)$ is a feasible point of the set-semidefinite optimization problem. Finally notice that the objective functions $f_1$ and $f_2$ are convex and the map $G$ is linear. Then all assumptions of Theorem 3.6 are fulfilled. Hence, we conclude that $(\bar{x}_1, \bar{x}_2) = (1,1)$ is a minimal solution of the set-semidefinite optimization problem.

## 4. Duality

As in the previous section we examine the vector optimization problem (1) under Assumption 1.1, i.e. we consider the problem

$$\begin{array}{c} \min\ f(x) \quad \text{subject to the constraint} \\ -G(x) \in C^K_L, \quad x \in X \end{array} \tag{24}$$

for an arbitrary nonempty set $K \subset Y$. Here we use the additional assumptions that $f$ and $G$ are convex, and that the quasi-interior $C^{\#}_Z$ of the dual cone is nonempty. So, problem (24) is a convex vector optimization problem.

In this section problem (24) is called *primal problem.* We will associate a so-called dual problem to this primal problem and we will present weak, strong and converse duality theorems. Finally, we specialize our theory to the linear case and obtain duality results for linear set-semidefinite vector optimization problems.

### 4.1. Convex Optimization Problems

The standard assumption of this subsection is summarized in

**Assumption 4.1.** Let Assumptions 1.1 be satisfied, let $f$ and $G$ be convex, let $K \subset Y$ be an arbitrary nonempty set and let the quasi-interior $C^{\#}_Z$ (see (19)) be nonempty.

Under this assumption we assign a *dual problem* to the primal problem (24)

$$\begin{aligned} &\max\ z \quad \text{subject to the constraints} \\ &\langle t, z\rangle \le \inf_{x\in X}(t\circ f + U\circ G)(x), \quad z\in Z,\ t\in C_Z^{\#},\ U\in (C_L^K)^*. \end{aligned} \tag{25}$$

A triple $(z, t, U)$ is called a *maximal solution* of this problem, if it is a minimal solution w.r.t. the ordering cone $-C_Z$. A first relationship between the primal and dual problem is given in

**Theorem 4.2 (weak duality theorem).** *Let Assumption* 4.1 *be satisfied. For every feasible pair* $(\bar z, \bar t, \bar U)$ *of the dual problem* (25)

$$\langle \bar t, \bar z\rangle \le (\bar t\circ f)(\bar x) \quad \textit{for all } \bar x\in S.$$

**Proof.** We have for an arbitrary $\bar x \in S$

$$\langle \bar t, \bar z\rangle \le \inf_{x\in X}(\bar t\circ f + \bar U\circ G)(x) \le (\bar t\circ f)(\bar x) + \underbrace{(\bar U\circ G)(\bar x)}_{\le 0} \le (\bar t\circ f)(\bar x).$$

□

**Theorem 4.3 (strong duality theorem).** *Let Assumption* 4.1 *be satisfied, and let* $\bar x \in S$ *be a minimal solution of the primal problem* (24) *with the additional property that for some* $t \in C_Z^{\#}$

$$(t\circ f)(\bar x) \le (t\circ f)(x) \quad \textit{for all } x\in S.$$

*Let the scalar optimization problem* $\min\limits_{x\in S}(t\circ f)(x)$ *be stable, i.e.*

$$\inf_{x\in S}(t\circ f)(x) = \sup_{U\in(C_L^K)^*}\inf_{x\in X}(t\circ f + U\circ G)(x)$$

*and the problem*

$$\sup_{U\in(C_L^K)^*}\inf_{x\in X}(t\circ f + U\circ G)(x)$$

*has at least one solution. Then there is a maximal solution* $(\bar z, \bar t, \bar U)$ *of the dual problem* (25) *with* $f(\bar x) = \bar z$.

**Proof.** This theorem is an equivalent formulation of Theorem 8.7 (b) in [27] in the case of set-semidefinite optimization. □

If for the problem $\min_{x\in S} f(x)$ the generalized Slater condition is satisfied, i.e. there exists a vector $x \in X$ with $G(x) \in -\mathrm{int}(C_L^K)$, then the problem $\min_{x\in S}(t\circ f)(x)$ is stable (compare [27], p. 197).

Notice that it is also possible to work with the weak minimality notion according to [27], p. 196. But here we restrict ourselves to the minimality notion which is of interest in practice.

**Theorem 4.4 (strong converse duality theorem).** *Let Assumption* 4.1 *be satisfied and, in addition, let $Z$ be locally convex, let the set $f(S)+C_Z$ be closed, and let for arbitrary $t \in C_Z^{\#}$*

$$\inf_{x\in S}(t\circ f)(x) = \sup_{U\in (C_L^K)^*} \inf_{x\in X}(t\circ f + U\circ G)(x).$$

*Then for every maximal solution $(\bar z, \bar t, \bar U)$ of the dual problem* (25) *there exists a minimal solution $\bar x$ of the primal problem* (24) *with $\bar z = f(\bar x)$.*

**Proof.** This theorem is an equivalent formulation of Theorem 8.9 (b) in [27] in the case of set-semidefinite optimization. □

## 4.2. Linear Optimization Problems

In this subsection we specialize the convex problem to a linear vector optimization problem. First, we modify our standard assumption.

**Assumption 4.5.** Let $X$ and $Z$ be real separated locally convex topological linear spaces. Let $Y$ be a real normed space and let the linear space $L(Y,Y^*)$ be partially ordered with an ordering cone $C_L^K$ for an arbitrary nonempty set $K \subset Y$. Let the ordering cone $C_Z$ be nontrivial, i.e. $C_Z \neq \{0_Z\}$, and let the quasi-interior $C_Z^{\#}$ be nonempty. Let $C\colon X \to Z$ and $A\colon X \to L(Y,Y^*)$ be continuous linear maps, and let $B \in L(Y,Y^*)$ be fixed. Moreover, let the constraint set

$$S := \{x \in X \mid A(x) - B \in C_L^K\}$$

be nonempty.

Under this assumption we consider the *primal linear problem*

$$\begin{array}{c} \min\ C(x) \quad \text{subject to the constraint} \\ A(x) - B \in C_L^K, \quad x \in X. \end{array} \tag{26}$$

Then the dual linear problem can be written as

$$\begin{array}{c} \max\ z \quad \text{subject to the constraints} \\ \langle t, z\rangle \le \langle t, C(x)\rangle + \langle U, B - A(x)\rangle \ \text{ for all } x \in X; \quad z \in Z,\ t \in C_Z^{\#},\ U \in (C_L^K)^*. \end{array}$$

If we denote

$$\begin{array}{rl} \hat S := \{(z,t,U) \in Z \times C_Z^{\#} \times (C_L^K)^* \mid & \langle t, z\rangle \le \langle t, C(x)\rangle + \langle U, B - A(x)\rangle \\ & \text{for all } x \in X\}, \end{array} \tag{27}$$

we can rewrite this set using the following lemma.

**Lemma 4.6.** *Let Assumption* 4.5 *be satisfied, and let the set $\hat S$ be given by* (27)*. Then*

$$\hat S = \{(z,t,U) \in Z \times C_Z^{\#} \times (C_L^K)^* \mid A^*(U) = C^*(t) \ \textit{and}\ \langle t, z\rangle \le \langle U, B\rangle\}$$

*($A^*$ and $C^*$ denote adjoint maps).*

**Proof.** (a) Let $(z,t,U) \in \hat{S}$ be arbitrarily given. Then we get

$$(t \circ C - U \circ A)(x) \geq \langle t, z\rangle - \langle U, B\rangle \;\text{ for all } x \in X \tag{28}$$

implying

$$t \circ C - U \circ A = 0_{X^*}$$

or

$$A^*(U) = C^*(t).$$

For $x = 0_X$ it follows from (28)

$$\langle t, z\rangle \leq \langle U, B\rangle.$$

(b) Let $(z,t,U) \in Z \times C_Z^{\#} \times (C_L^K)^*$ with $A^*(U) = C^*(t)$ and $\langle t, z\rangle \leq \langle U, B\rangle$ be arbitrarily given. Then we conclude

$$(t \circ C - U \circ A)(x) = 0 \geq \langle t, z\rangle - \langle U, B\rangle \;\text{ for all } x \in X$$

implying $(z,t,U) \in \hat{S}$. □

Hence, the dual linear problem can be written as

$$\begin{gathered}\max\ z \quad \text{subject to the constraints} \\ A^*(U) = C^*(t), \quad \langle t, z\rangle \leq \langle U, B\rangle, \quad z \in Z,\ t \in C_Z^{\#},\ U \in (C_L^K)^*.\end{gathered} \tag{29}$$

The inequality constraint $\langle t, z\rangle \leq \langle U, B\rangle$ can be treated as an equality. This result is proved in the next lemma.

**Lemma 4.7.** *Let Assumption* 4.5 *be satisfied. The triple* $(\bar{z}, \bar{t}, \bar{U})$ *is a maximal solution of the dual problem* (29) *if and only if it is a maximal solution of the problem*

$$\begin{gathered}\max\ z \quad \textit{subject to the constraints} \\ A^*(U) = C^*(t), \quad \langle t, z\rangle = \langle U, B\rangle, \quad z \in Z,\ t \in C_Z^{\#},\ U \in (C_L^K)^*.\end{gathered} \tag{30}$$

**Proof.** (a) Let $(\bar{z}, \bar{t}, \bar{U})$ be a maximal solution of problem (29). Assume that $\langle \bar{t}, \bar{z}\rangle < \langle \bar{U}, B\rangle$, then there is an element $c \in C_Z \setminus \{0_Z\}$ with $\langle \bar{t}, c\rangle = \langle \bar{U}, B\rangle - \langle \bar{t}, \bar{z}\rangle > 0$. So, we obtain

$$\langle \bar{t}, \bar{z} + c\rangle = \langle \bar{t}, \bar{z}\rangle + \langle \bar{t}, c\rangle = \langle \bar{U}, B\rangle.$$

Since the element $(\bar{z} + c, \bar{t}, \bar{U})$ satisfies the constraints of problem (29), we obtain a contradiction to the assumption that $(\bar{z}, \bar{t}, \bar{U})$ is a maximal solution of problem (29). Consequently, we have $\langle \bar{t}, \bar{z}\rangle = \langle \bar{U}, B\rangle$ and $(\bar{z}, \bar{t}, \bar{U})$ is a maximal solution of problem (30).

(b) Let $(\bar{z}, \bar{t}, \bar{U})$ be a maximal solution of problem (30) and assume that it is not a maximal solution of problem (29). Then there is an element $c \in C_Z \setminus \{0_Z\}$ so that $(\bar{z} + c, t, U)$ is a feasible point of problem (29). Hence, we have $\langle t, \bar{z} + c\rangle \leq \langle U, B\rangle$. In analogy to part (a) the case $\langle t, \bar{z} + c\rangle < \langle U, B\rangle$ means that there is some $\tilde{c} \in C_Z \setminus \{0_Z\}$ with $\langle t, \bar{z} + c + \tilde{c}\rangle = \langle U, B\rangle$. So, $(\bar{z} + c + \tilde{c}, t, U)$ is a feasible point of (30), a contradiction to the maximality of $(\bar{z}, \bar{t}, \bar{U})$. □

Although problem (30) is only an equivalent formulation of the dual problem (29), we call it the *dual linear problem* being used in the following.

It may be of interest to note that based on the formulation (30) and using the approach in [27], Section 8.3, under Assumption 4.1 and the additional assumption $B \neq 0_{L(Y,Y^*)}$ the dual problem is even equivalent to the problem

$$\begin{gathered}\max\ T(B) \quad \text{subject to the constraints} \\ (C - TA)^*(t) = 0_{X^*}, \quad T^*(t) \in (C_L^K)^*, \quad t \in C_Z^{\#}, \quad T \in L(L(Y,Y^*), Z).\end{gathered}$$

It is known (compare the proof of Theorem 2.3 in [27]) that $T$ has a special structure. It maps the linear space $L(Y,Y^*)$ to a two-dimensional subspace of $Z$. This subspace is even one-dimensional in many cases.

**Remark 4.8.** We now discuss the special case that the primal linear problem is a scalar optimization problem, i.e. $Z = \mathbb{R}$. Then we have the ordering cone $C_Z = \mathbb{R}_+$ with the quasi-interior $C_Z^{\#} = \mathbb{R}_+ \setminus \{0\}$ of the dual cone. This means that $t \in C_Z^{\#}$ is equivalent to $t > 0$. And the equation $\langle t, z\rangle = \langle U, B\rangle$ is equivalent to $z = \frac{1}{t}\langle U, B\rangle$. So, the dual problem (30) can be written as

$$\begin{gathered}\max\ \tfrac{1}{t}\langle U, B\rangle \quad \text{subject to the constraints} \\ A^*(\tfrac{1}{t}U) = C^*, \quad t > 0,\ U \in (C_L^K)^*.\end{gathered}$$

being equivalent to the problem

$$\begin{gathered}\max\ \langle U, B\rangle \quad \text{subject to the constraints} \\ A^*(U) = C^*, \quad U \in (C_L^K)^*.\end{gathered}$$

**Remark 4.9.** We consider the linear vector optimization problem in finite dimensions. Now we assume that $X = \mathbb{R}^m$, $Y = \mathbb{R}^n$, and $Z = \mathbb{R}^k$. Then the linear map $A\colon \mathbb{R}^m \to \mathcal{M}^n$ can be written as

$$A(x) = A^{(1)}x_1 + \ldots + A^{(m)}x_m \ \text{ for all } x \in \mathbb{R}^m$$

with matrices $A^{(1)}, \ldots, A^{(m)} \in \mathcal{M}^n$. For given matrices $C \in \mathbb{R}^{k\times m}$ and $B \in \mathcal{M}^n$ the primal problem (26) is

$$\begin{gathered}\min\ Cx \quad \text{subject to the constraint} \\ A^{(1)}x_1 + \ldots + A^{(m)}x_m - B \in C_{\mathcal{M}^n}^K, \quad x \in \mathbb{R}^m.\end{gathered}$$

For the formulation of the dual problem we need the adjoint map $A^*$ given by

$$\begin{aligned}\langle A^*U, x\rangle &= \langle U\,,\, A(x)\rangle \\ &= \langle U\,,\, A^{(1)}\rangle x_1 + \ldots + \langle U\,,\, A^{(m)}\rangle x_m \\ &= \big(\langle U\,,\, A^{(1)}\rangle, \ldots, \langle U\,,\, A^{(m)}\rangle\big)\cdot x \ \text{ for all } x \in \mathbb{R}^m \text{ and all } U \in \mathcal{M}^n.\end{aligned}$$

From the general formulation (30) we then obtain the special dual problem

$$\begin{gathered}\max\ z \quad \text{subject to the constraints} \\ \langle U\,,\, A^{(1)}\rangle = c_1^\top t \\ \vdots \\ \langle U\,,\, A^{(m)}\rangle = c_m^\top t \\ t^\top z = \langle U\,,\, B\rangle, \quad z \in \mathbb{R}^k,\ t \in C_{\mathbb{R}^k}^{\#},\ U \in (C_{\mathcal{M}^n}^K)^*\end{gathered} \tag{31}$$

where $c_i$ $(1 \leq i \leq m)$ denotes the $i$-th column of the matrix $C$. For instance, if we choose the subspace $\mathcal{S}^n$, the cones $K = \mathbb{R}^n_+$ (i.e., we consider the copositive case) and $C_{\mathbb{R}^k} = \mathbb{R}^k_+$, then the dual problem (31) can be specialized to

$$\begin{array}{c} \max\ z \quad \text{subject to the constraints} \\ \langle U\,,\,A^{(1)}\rangle = c_1^\top t \\ \vdots \\ \langle U\,,\,A^{(m)}\rangle = c_m^\top t \\ t^\top z = \langle U\,,\,B\rangle, \quad t_1,\ldots,t_k > 0, \quad z \in \mathbb{R}^k,\ U \in \mathcal{S}^n \text{ completely positive.} \end{array}$$

Remember that according to Lemma 2.28 (b) the dual cone of the cone of the copositive matrices is given by the cone of the completely positive matrices.
Notice that this duality approach in finite dimensions extends the approach in [29] given for a scalar-valued linear objective and positive semidefinite matrices.

We end our investigations again with an example.

**Example 4.10.** Similar to Example 3.7 we investigate the following vector optimization problem with $X = Z = \mathbb{R}^2$, $Y = \mathbb{R}^3$ and the ordering cones $C_Z = \mathbb{R}^2_+$, $C^K_{\mathcal{S}^3} = \mathcal{S}^3_+$:

$$\begin{array}{c} \min\ f(x_1,x_2) = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \quad \text{subject to the constraint} \\ -G(x_1,x_2) = \begin{pmatrix} x_1 & 1 & 0 \\ 1 & x_2 & 0 \\ 0 & 0 & x_1+1 \end{pmatrix} \in \mathcal{S}^3_+, \quad x = (x_1,x_2) \in \mathbb{R}^2. \end{array} \tag{32}$$

The constraint set is according to Example 3.7 given by

$$\begin{aligned} S &= \{x \in \mathbb{R}^2 \mid -G(x) \text{ positive semidefinite }\} \\ &= \{x \in \mathbb{R}^2_+ \mid x_1 x_2 \geq 1\}. \end{aligned}$$

The set of minimal solutions of problem (32) is thus

$$S^0 := \{x \in \mathbb{R}^2_+ \mid x_1 x_2 = 1\},$$

see Figure 4.1.

The objective function $f$ is linear, i.e. $f(x) = Cx$ with

$$C = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix},$$

and the constraint function $G$ is also linear. It is $G(x) = B - A(x)$ with

$$B = \begin{pmatrix} 0 & -1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & -1 \end{pmatrix}$$

and $A(x) = A^{(1)}x_1 + A^{(2)}x_2$ with

$$A^{(1)} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad \text{and} \quad A^{(2)} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

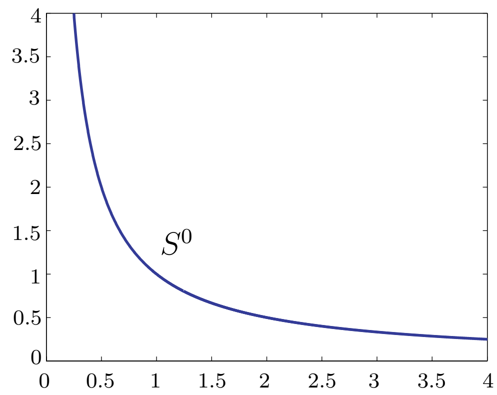


Figure 4.1: Minimal solution set of Example 4.10.

By Remark 4.9 the dual problem to (32) is then given by problem (31), i.e.

$$\max\ z = \begin{pmatrix} z_1 \\ z_2 \end{pmatrix} \quad \text{subject to the constraints}$$
$$\langle U\,,\,A^{(1)}\rangle = c_1^\top t, \quad \langle U\,,\,A^{(2)}\rangle = c_2^\top t, \quad t^\top z = \langle U\,,\,B\rangle,$$
$$z \in \mathbb{R}^2,\ t \in C^{\#}_{\mathbb{R}^2}, U \in \mathcal{S}^3_+,\ \text{i.e. } U \text{ is positive semidefinite.}$$

Remember that $(\mathcal{S}^3_+)^* = \mathcal{S}^3_+$ according to Lemma 2.28 (c) and that $\langle A\,,\,B\rangle = \text{trace}(A\cdot B)$ for arbitrary $A, B \in \mathcal{S}^3$. Further it is $C^{\#}_{\mathbb{R}^2} = \{y \in \mathbb{R}^2 \mid y_1, y_2 > 0\}$. Thus the dual problem is equivalent to

$$\begin{gathered} \max\ z \quad \text{subject to the constraints} \\ U_{11} + U_{33} = t_1, \quad U_{22} = t_2, \quad t_1 z_1 + t_2 z_2 = -2U_{12} - U_{33}, \\ t_1,\ t_2 > 0, \quad z \in \mathbb{R}^2, U \in \mathcal{S}^3_+ \end{gathered} \tag{33}$$

with $U_{ij}$ being the coefficients of a symmetric positive semidefinite matrix $U \in \mathcal{S}^3_+$. The point $\bar{x} = (1,1) \in S^0$ is a minimal solution of problem (32) and for $t = (1,1) \in C^{\#}_{\mathbb{R}^2}$ it is

$$(t \circ f)(x) = x_1 + x_2 \geq x_1 + \frac{1}{x_1} = \underbrace{\frac{(x_1-1)^2}{x_1}}_{\geq 0} + 2 \geq 2 = (t \circ f)(\bar{x}) \ \text{ for all } x \in S.$$

For $\hat{x} = (2,2)$ the eigenvalues of $-G(\hat{x})$ are $\lambda_1 = 1$, $\lambda_{2/3} = 3$ and thus (with Corollary 2.31)

$$-G(\hat{x}) \in \text{int}(\mathcal{S}^3_+) = \{A \in \mathcal{S}^3 \mid A \text{ is positive definite }\}.$$

Then the generalized Slater condition is satisfied for the problem $\min_{x\in S} f(x)$ and thus the problem $\min_{x\in S}(t \circ f)(x)$ is stable. Therefore the assumptions of Theorem 4.3 are satisfied for $\bar{x} = (1,1)$ and the dual problem (33) has at least one solution $(\bar{z}, \bar{t}, \bar{U})$ with $f(\bar{x}) = (1,1) = \bar{z}$. Here this is the case for $\bar{t} = (1,1)$ and

$$\bar{U} = \begin{pmatrix} 1 & -1 & 0 \\ -1 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

## 5. Penalty Approach

For treating the constraint $G(x) \preccurlyeq 0_{L(Y,Y^*)}$ in the optimization problem (1) we develop a penalty approach. In this section we assume the set $K \subset Y$ to be a cone. First we need some results about a minimization of the map $y \mapsto \langle Ay, y\rangle$ for $A \in L(Y, Y^*)$ over the cone $K$.

**Lemma 5.1.** *Let $K \subset Y$ be a cone and $A \in L(Y, Y^*)$. Then it is*

$$\inf_{y\in K} \langle Ay, y\rangle \in \{-\infty, 0\}.$$

**Proof.** Due to $0_Y \in K$ it is

$$\inf_{y\in K} \langle Ay, y\rangle \leq \langle A0_Y, 0_Y\rangle = 0.$$

If $\inf_{y\in K}\langle Ay, y\rangle < 0$, then there exists a $\bar{y} \in K$ with $\langle A\bar{y}, \bar{y}\rangle < 0$. For arbitrary $\lambda > 0$ it is $\lambda\bar{y} \in K$, too, because $K$ is a cone. Thus we have

$$\langle A(\lambda\bar{y}), \lambda\bar{y}\rangle = \lambda^2 \underbrace{\langle A\bar{y}, \bar{y}\rangle}_{<0}.$$

Because we can choose $\lambda > 0$ arbitrarily we conclude

$$\inf_{y\in K} \langle Ay, y\rangle = -\infty.$$

□

**Corollary 5.2.** *Under the assumptions of Lemma* 5.1 *it is*

$$\inf_{y\in K} \langle Ay, y\rangle = \begin{cases} 0 & \text{for } A \text{ } K\text{-semidefinite,} \\ -\infty & \text{else.} \end{cases}$$

**Proof.** If $A$ is $K$-semidefinite we have $\langle Ay, y\rangle \geq 0$ for all $y \in K$ and thus $\inf_{y\in K}\langle Ay, y\rangle = 0$. Else there exists a $\bar{y} \in K$ with $\langle A\bar{y}, \bar{y}\rangle < 0$ and therefore $\inf_{y\in K}\langle Ay, y\rangle = -\infty$. □

If there exists a set $B \subset K$ as described in Theorem 2.19 it can be sufficient to consider $B$ instead of $K$.

**Theorem 5.3.** *Let $K \subset Y$ be a cone and $B \subset Y$ as in Theorem* 2.19. *Then $A$ is $K$-semidefinite if and only if*

$$0 \leq \inf_{y\in B} \langle Ay, y\rangle.$$

**Proof.** According to Theorem 2.19 $A$ is $K$-semidefinite if and only if

$$\langle Ay, y\rangle \geq 0 \quad \text{for all } y \in B,$$

and thus if and only if $\inf_{y\in B}\langle Ay, y\rangle \geq 0$. □

For example, for $Y = \mathbb{R}^n$ this result can be used for checking whether a matrix $A \in \mathcal{M}^n$ is $K$-semidefinite. For $K = \mathbb{R}^n_+$ the compact set $B = \{y \in \mathbb{R}^n_+ \mid \|y\| = 1\}$ (with $\|\cdot\|$ an arbitrary norm in $\mathbb{R}^n$) fulfills the assumptions of Theorem 2.19 and can be used as a base on which the problem

$$\min_{y \in B} y^\top A y \tag{34}$$

is solved for testing whether the matrix $A$ is copositive. For solving (34) numerical methods of global optimization are needed, but it is well known that the determination of the copositivity of a given symmetric matrix is a NP-hard problem ([39]).

According to Lemma 5.1 if $\bar{y}$ is a minimal solution of

$$\min_{y \in K} \langle Ay, y \rangle$$

it is $\langle A\bar{y}, \bar{y} \rangle = 0$. Using the generalized Lagrange multiplier rule we can show that we then also have $A\bar{y} \in K^*$. For that we need the Fréchet derivative of the map $q_A : Y \to \mathbb{R}, y \mapsto \langle Ay, y \rangle$.

**Lemma 5.4.** *Let $(Y, \langle \cdot, \cdot \rangle)$ be a real Hilbert space and let $A \in L(Y, Y)$. For $\bar{y} \in Y$ the Fréchet derivative of the map $q_A : Y \to \mathbb{R},\ q_A(y) = \langle Ay, y \rangle$ is given by*

$$q'_A(\bar{y})(y) = \langle (A + A^*)\bar{y}, y \rangle$$

*for all $y \in Y$ with $A^* \colon Y \to Y$ the adjoint map to $A$.*

**Proof.** As it is

$$\begin{aligned}
& \lim_{\|h\|_Y \to 0} \frac{|q_A(\bar{y} + h) - q_A(\bar{y}) - q'_A(\bar{y})(h)|}{\|h\|_Y} \\
= & \lim_{\|h\|_Y \to 0} \frac{|\langle A(\bar{y} + h), \bar{y} + h \rangle - \langle A\bar{y}, \bar{y} \rangle - \langle (A + A^*)\bar{y}, h \rangle|}{\|h\|_Y} \\
= & \lim_{\|h\|_Y \to 0} \frac{|\langle Ah, h \rangle|}{\|h\|_Y} \\
\leq & \lim_{\|h\|_Y \to 0} \frac{\|Ah\|_Y \|h\|_Y}{\|h\|_Y} \\
= & \lim_{\|h\|_Y \to 0} \|Ah\|_Y = 0
\end{aligned}$$

we have

$$\lim_{\|h\|_Y \to 0} \frac{|q_A(\bar{y} + h) - q_A(\bar{y}) - q'_A(\bar{y})(h)|}{\|h\|_Y} = 0$$

and thus $q_A$ is Fréchet differentiable with Fréchet derivative $q'_A(\bar{y})(y) = \langle (A + A^*)\bar{y}, y \rangle$ for all $y \in Y$. □

If the map $A$ is self-adjoint then we conclude that the Fréchet derivative is given by $q'_A(\bar{y})(y) = 2\langle A\bar{y}, y \rangle$ for all $y \in Y$.

**Theorem 5.5.** *Let $(Y, \langle\cdot,\cdot\rangle)$ be a real Hilbert space, $K \subset Y$ a convex cone with $int(K) \neq \emptyset$ and $A \in L(Y,Y)$. If $\bar{y}$ is a minimal solution of*

$$\min_{y\in K} \langle Ay, y\rangle, \tag{35}$$

*it is $(A + A^*)\bar{y} \in K^*$ and $\langle (A + A^*)\bar{y}, \bar{y}\rangle = 0$ with the adjoint map $A^* \colon Y \to Y$ to $A$.*

**Proof.** The constraint $y \in K$ of the optimization problem (35) is equivalent to $g(y) \in -K$ with the convex function $g\colon Y \to Y,\ g(y) = -y$ for all $y \in Y$. As $\text{int}(K) \neq \emptyset$ there exists a $\hat{y} \in \text{int}(K)$ and thus

$$g(\bar{y}) + g'(\bar{y})(\hat{y} - \bar{y}) = -\bar{y} - (\hat{y} - \bar{y}) = -\hat{y} \in -\text{int}(K). \tag{36}$$

Therefore, applying Theorem 5.6 in [28] the Kurcyusz-Robinson-Zowe regularity assumption is satisfied. Note that here the condition (36) is equivalent to the generalized Slater condition, i.e. to the existence of an element $\hat{y} \in Y$ with $g(\hat{y}) \in -\text{int}(K)$ ([27], p. 197). Using the generalized Lagrange multiplier rule (see [28], Theorem 5.3) it follows that there exists a $u \in K^*$ with $\langle u, \bar{y}\rangle = 0$ and

$$q_A'(\bar{y}) + u \circ g'(\bar{y}) = 0_Y$$

with $q_A$ as in Lemma 5.4. We conclude $(A + A^*)\bar{y} - u = 0_Y$ and thus

$$(A + A^*)\bar{y} = u \in K^*.$$

Then the equation $\langle u, \bar{y}\rangle = 0$ can be written as $\langle (A + A^*)\bar{y}, \bar{y}\rangle = 0$. □

Thus, for $A$ self-adjoint, we conclude under the assumptions of Theorem 5.5

$$A\bar{y} \in K^* \quad \text{and} \quad \langle A\bar{y}, \bar{y}\rangle = 0.$$

Han and Mangasarian discuss in [23], p. 2-3, the optimization problem (35) for the finite dimensional case $Y = \mathbb{R}^n$ and deduce a result as in Theorem 5.5.

Now we are able to formulate a weak converse implication of Theorem 2.8.

**Lemma 5.6.** *Let $(Y, \langle\cdot,\cdot\rangle)$ be a real Hilbert space, $K \subset Y$ a convex cone with $int(K) \neq \emptyset$ and let $A \in L(Y,Y)$ be a $K$-semidefinite map. Then $y \in K$ with $\langle Ay, y\rangle = 0$ implies $(A + A^*)y \in K^*$ with $A^* \colon Y \to Y$ the adjoint map to $A$.*

**Proof.** As $A$ is $K$-semidefinite there exists according to Corollary 5.2 a minimal solution of

$$\min_{y\in K} \langle Ay, y\rangle \tag{37}$$

with minimal value 0. Thus any $y \in K$ with $\langle Ay, y\rangle = 0$ is a minimal solution of (37) and Theorem 5.5 then implies $(A + A^*)y \in K^*$. □

Note, that with additionally assuming that $A$ is self-adjoint we get under the assumptions of Lemma 5.6 that $y \in K$ with $\langle Ay, y\rangle = 0$ implies $Ay \in K^*$. For $A \in \mathcal{S}^n$ and $K = \mathbb{R}^n_+$ this result is stated in [26], Lemma 4.4.

We now come to the announced penalty formulation of the constraint $G(x) \preccurlyeq 0_{L(Y,Y^*)}$.

**Theorem 5.7.** *Let $K \subset Y$ be a cone and let $B$ be given as in Theorem 2.19. Let $S$ be defined as in (2), i.e.*

$$S = \{x \in X \mid -G(x) \text{ is } K\text{-semidefinite}\}.$$

*If we define a function $P : X \to \mathbb{R}_+$ by*

$$P(x) = \max\{\,0\,, \sup_{y \in B}\langle G(x)y, y\rangle\} \quad \text{for all } x \in X$$

*it follows*

$$P(x) = 0 \;\Leftrightarrow\; x \in S.$$

**Proof.** Let $x \in X$ be given. By Theorem 5.3 $-G(x)$ is $K$-semidefinite, i.e. $x \in S$, if and only if

$$\inf_{y \in B}\langle -G(x)y, y\rangle \geq 0$$

or

$$\langle G(x)y, y\rangle \leq 0 \quad \text{for all } y \in B$$

being equivalent to

$$P(x) = \max\{\,0\,, \underbrace{\sup_{y \in B}\langle G(x)y, y\rangle}_{\leq 0}\,\} = 0.$$

□

This result of Theorem 5.7 can also be formulated for arbitrary nonempty sets $K$. The function $P$ in the preceding theorem is of Penalty type, i.e. $x \in S$ if and only if $P(x) = 0$ and $P(x) > 0$ for $x \in X \setminus S$. But note that $P$ may not be a continuous function. We can apply Theorem 5.7 for example if $B$ is a base of $K$. We can use this, for instance, for $X = \mathbb{R}^m$, $Y = \mathbb{R}^n$, $K \subset \mathbb{R}^n$ a pointed convex cone with a base $B \subset Y$ and $Z = \mathbb{R}$ for solving problem (1):

$$\begin{array}{l} \min f(x) \quad \text{subject to the constraint} \\ -G(x) \;\; K\text{-semidefinite}, \quad x \in \mathbb{R}^m \end{array}$$

by the following Penalty approach ($\mu > 0$)

$$\min_{x \in \mathbb{R}^m} f(x) + \mu P(x) = \min_{x \in \mathbb{R}^m} f(x) + \mu \cdot \max\{\,0\,, \sup_{y \in B} y^\top G(x) y\,\}.$$

Here, the subproblem

$$\sup_{y \in B} y^\top G(x) y$$

for $x \in \mathbb{R}^m$ has to be solved by a numerical method of global optimization.

## 6. Conclusions

A new field of vector optimization in infinite dimensions called set-semidefinite optimization is presented which covers the known semidefinite and copositive programming in finite dimensions. The basic principle of this generalization is the introduction of a special ordering cone, the $K$-semidefinite cone, which is the convex cone of all quadratic forms being positive semidefinite on a set $K$. Various results for this cone are proved such as calculation rules, specifications, connections to eigenvalues and a characterization of the interior.

Necessary as well as sufficient optimality conditions are given for the general case and specialized to finite dimensions extending the known results in semidefinite optimization to problems with vector-valued objective function. The theory is completed by several duality results with a special focus on linear problems. Moreover, a penalty approach is developed for the solution of set-semidefinite optimization problems using an appropriate global optimization solver. Numerical examinations of this approach are subject to further research.

The advantage of this generalization is that the gained results can be applied both to semidefinite and copositive optimization. Thus these problem classes can be extended to vector-valued objective functions based on the presented theory. Further this paper establishes a basis for the consideration of more general problems of semidefinite type in the future.

## References

[1] L.-F. Andersson, G. Chang, T. Elfving: Criteria for copositive matrices using simplices and barycentric coordinates, Linear Algebra Appl. 220 (1995) 9–30.

[2] A. Berman, R. J. Plemmons: Nonnegative Matrices in the Mathematical Sciences, Academic Press, New York (1979).

[3] E. Blum, W. Oettli: From optimization and variational inequalities to equilibrium problems, Math. Stud. 63 (1994) 123–145.

[4] I. M. Bomze: Linear-time copositivity detection for tridiagonal matrices and extension to block-tridiagonality, SIAM J. Matrix Anal. Appl. 21 (2000) 840–848.

[5] I. M. Bomze, E. de Klerk: Solving standard quadratic optimization problems via linear, semidefinite and copositive programming, J. Glob. Optim. 24 (2002) 163–185.

[6] I. M. Bomze, M. Dür, E. de Klerk, C. Roos, A. J. Quist, T. Terlaky: On copositive programming and standard quadratic optimization problems, J. Glob. Optim. 18 (2000) 301–320.

[7] I. M. Bomze, B. M. Pötscher: Game Theoretic Foundations of Evolutionary Stability, Springer, Berlin (1989).

[8] J. M. Borwein: Necessary and sufficient conditions for quadratic minimality, Numer. Funct. Anal. Optimization 5 (1982) 127–140.

[9] S. Bundfuss, M. Dür: Criteria for copositivity and approximations of the copositive cone, Preprint, Department of Mathematics, Darmstadt University of Technology, Darmstadt (2006).

[10] S. Bundfuss, M. Dür: Algorithmic copositivity detection by simplicial partition, Linear Algebra Appl. 428 (2008) 1511–1523.

[11] S. Burer: On the copositive representation of binary and continuous nonconvex quadratic programs, Working paper, Department of Management Sciences, University of Iowa, Iowa City (2007).

[12] R. W. Cottle, G. J. Habetler, C. E. Lemke: On classes of copositive matrices, Linear Algebra Appl. 3 (1970) 295–310.

[13] R. W. Cottle, G. J. Habetler, C. E. Lemke: Quadratic forms semi-definite over convex cones, in: Proc. Princeton Sympos. Math. Programming (Princeton, 1967), H. W. Kuhn (ed.), Princeton University Press, Princeton (1970) 551–565.

[14] G. Danninger: Role of copositivity in optimality criteria for nonconvex optimization problems, J. Optimization Theory Appl. 75 (1992) 535–558.

[15] G. Danninger, I. M. Bomze: Using copositivity for global optimality criteria in concave quadratic programming problems, Math. Program. 62 (1993) 575–580.

[16] E. de Klerk, M. Laurent, P. Parrillo: On the equivalence of algebraic approaches to the minimization of forms on the simplex, in: Positive Polynomials in Control, D. Henrion, A. Garulli (eds.), Lecture Notes in Control and Information Sciences 312, Springer, Berlin (2005) 121–132.

[17] E. de Klerk, D. V. Pasechnik: Approximation of the stability number of a graph via copositive programming, SIAM J. Optim. 12 (2002) 875–892.

[18] M. Fukuda, M. Yamashita, M. Kojima: Computational prospects on copositive programming, RIMS Kokyuroku 1526 (2006) 207–213.

[19] M. S. Gowda: A characterization of positive semidefinite operators on a Hilbert space, J. Optimization Theory Appl. 48 (1986) 419–425.

[20] M. S. Gowda: Complementarity problems over locally compact cones, SIAM J. Control Optimization 27 (1989) 836–841.

[21] K. P. Hadeler: On copositive matrices, Linear Algebra Appl. 49 (1983) 79–89.

[22] S.-P. Han, O. L. Mangasarian: A conjugate decomposition of the Euclidean space, Proc. Natl. Acad. Sci. USA 80 (1983) 5156–5157.

[23] S.-P. Han, O. L. Mangasarian: Conjugate cone characterization of positive definite and semidefinite matrices, Linear Algebra Appl. 56 (1984) 89–103.

[24] T. L. Hayden, J. Wells: Approximation by matrices positive semidefinite on a subspace, Linear Algebra Appl. 109 (1988) 115–130.

[25] L. Hogben, C. R. Johnson, R. Reams: The copositive completion problem, Linear Algebra Appl. 408 (2005) 207–211.

[26] K. D. Ikramov, N. V. Savel'eva: Conditionally definite matrices, J. Math. Sci., New York 98 (2000) 1–50.

[27] J. Jahn: Vector Optimization. Theory, Applications, and Extensions, Springer, Berlin (2004).

[28] J. Jahn: Introduction to the Theory of Nonlinear Optimization, 3rd Ed., Springer, Berlin (2007).

[29] F. Jarre, J. Stoer: Optimierung, Springer, Berlin (2004).

[30] W. Kaplan: A test for copositive matrices, Linear Algebra Appl. 313 (2000) 203–206.

[31] W. Kaplan: A copositivity probe, Linear Algebra Appl. 337 (2001) 237–251.

[32] H. Länger: Strictly copositive matrices and ESS's, Arch. Math. 55 (1990) 516–520.

[33] R. Loewy, H. Schneider: Positive operators on the n-dimensional ice cream cone, J. Math. Anal. Appl. 49 (1975) 375–392.

[34] K. Löwner: Über monotone Matrixfunktionen, Math. Z. 38 (1934) 177–216.

[35] D. H. Martin: Conditional positivity of quadratic forms in Hilbert space, SIAM J. Math. Anal. 11 (1980) 1047–1057.

[36] D. H. Martin: Finite criteria for conditional definiteness of quadratic forms, Linear Algebra Appl. 39 (1981) 9–21.

[37] H. Maurer, J. Zowe: First and second-order necessary and sufficient optimality conditions for infinite-dimensional programming problems, Math. Program. 16 (1979) 98–110.

[38] T. S. Motzkin: Copositive quadratic forms, National Bureau of Standards Report 1818 (1952) 11–22.

[39] K. G. Murty, S. N. Kabadi: Some NP-complete problems in quadratic and linear programming, Math. Program. 39 (1987) 117–129.

[40] P. A. Parrillo: Structured Semidefinite Programs and Semi-Algebraic Geometry Methods in Robustness and Optimization, PhD Thesis, California Institute of Technology, Pasadena (2000).

[41] J. Povh: Application of Semidefinite and Copositive Programming in Combinatorial Optimization, PhD Thesis, Faculty of Mathematics and Physics, University of Ljubljana, Ljubljana (2006).

[42] J. Povh, F. Rendl: Copositive and semidefinite relaxations of the quadratic assignment problem, Manuscript, Department of Mathematics, University of Klagenfurt (2006).

[43] J. Povh, F. Rendl: A copositive programming approach to graph partitioning, SIAM J. Optim. 18 (2007) 223–241.

[44] A. J. Quist, E. de Klerk, C. Roos, T. Terlaky: Copositive relaxation for general quadratic programming, Optim. Methods Softw. 9 (1998) 185–208.

[45] Y. Sawaragi, H. Nakayama, T. Tanino: Theory of Multiobjective Optimization, Academic Press, London (1985).

[46] H. C. Simpson, S. J. Spector: On copositive matrices and strong ellipticity for isotropic elastic materials, Arch. Ration. Mech. Anal. 84 (1983) 55–68.

[47] J. F. Sturm, S. Zhang: On cones of nonnegative quadratic functions, Math. Oper. Res. 28 (2003) 246–267.

[48] L. Vandenberghe, S. Boyd: Semidefinite programming, SIAM Rev. 38 (1996) 49–95.

[49] H. Wolkowicz, R. Saigal, L. Vandenberghe (eds.): Handbook on Semidefinite Programming, Kluwer, Dordrecht (2000).

[50] H. Väliaho: Criteria for copositive matrices, Linear Algebra Appl. 81 (1986) 19–34.

[51] H. Väliaho: Testing the definiteness of matrices on polyhedral cones, Linear Algebra Appl. 101 (1988) 136–165.

[52] H. Väliaho: Quadratic-programming criteria for copositive matrices, Linear Algebra Appl. 119 (1989) 163–182.

[53] J. Werner: Optimization, Theory and Applications, Vieweg, Braunschweig (1984).