In this section, we are interested in the so-called order of optimizers. The order of maximizers provides information about the sharpness of the optimal solution: the smaller the order, the sharper is the maximizer. Geometrically, the order is related to the “curvature” of the feasible set around the optimal solution $\bar{x}$.

**Definition 4.1.** A feasible solution $\bar{x}$ (or $\bar{X}$) of $(P)$ is called a maximizer of order $p > 0$ if there exist $\gamma > 0$ and $\varepsilon > 0$ such that

$$c^T \bar{x} - c^T x \geq \gamma \|x - \bar{x}\|^p \quad \text{for all } x \in \mathcal{F}_P \text{ with } \|x - \bar{x}\| < \varepsilon.$$

The following figures illustrate first and second order maximizers for the problem

$$\max \quad c^T x$$

$$\text{s.t. } x \in \mathcal{F}_P$$

---

*Submitted as M. Dür, B. Jargalsaikhan, and G. Still. First order solutions in conic programming. Preprint, 2014*
From the definition of the order of a maximizer, we directly obtain the following

- If $\bar{x}$ is a maximizer of order $p > 0$, then $\bar{x}$ is a unique maximizer.
- A maximizer of order $p$ is also a maximizer of any order $p' > p$.

In linear programming, since there is no curvature, any unique maximizer is a first order maximizer, for the proof see e.g. [GLT95, Theorem 3.1]. Let us directly show the above statement.

**Proposition 4.2.** Suppose that the following LP has a unique optimal solution $\bar{x}$.

\[
(LP) \quad \max \ c^T x \quad \text{s.t.} \quad B - A^T x \geq 0
\]

where $A$ is an $n \times m$ matrix and $c \in \mathbb{R}^n$, $B \in \mathbb{R}^m$. Then $\bar{x}$ is a first order maximizer.

**Proof.** Let us denote the active index set

\[ I(\bar{x}) = \{ i \in \{1, \ldots, m\} \mid A_i^T \bar{x} = B_i \} \]

where $A_i$ is the $i$th column of the matrix $A$. Locally $\bar{x}$ is given uniquely by the equations $A_i^T \bar{x} = B_i$ with $i \in I(\bar{x})$. As $\bar{x}$ is unique, we can assume that $|I(\bar{x})| = n$ and that the $A_i$ with $i \in I(\bar{x})$ are linearly independent. By the KKT conditions, we have $c = \sum_{i \in I(\bar{x})} y_i A_i$. Uniqueness is equivalent to the condition $c \in \text{int} \, \mathcal{M}(\bar{x}) = \text{int} \, \text{cone}\{A_i \mid i \in I(\bar{x})\}$, for the proof see e.g. [GLT95, Theorem 3.1]. Thus, the KKT condition holds with $y_i > 0$ for all $i \in I(\bar{x})$ and we can rewrite the KKT condition as $c = A_{I(\bar{x})} y$ where $y > 0$ and $A_{I(\bar{x})}$ is the $n \times n$ nonsingular matrix with columns $A_i$ with $i \in I(\bar{x})$. Thus,

\[ c^T \bar{x} - c^T x = c^T (\bar{x} - x) = y^T A_{I(\bar{x})}^T (\bar{x} - x). \]

Let us denote $\varepsilon := \min_j y_j > 0$. Then we have

\[ y^T A_{I(\bar{x})}^T (\bar{x} - x) \geq \varepsilon \| A_{I(\bar{x})}^T (\bar{x} - x) \|_1 \geq \varepsilon \frac{1}{\| (A_{I(\bar{x})}^{-1})^T \|_1} \| (\bar{x} - x) \|. \]
where we used with appropriate matrix norm $\|A\|_1$ for any $d \in \mathbb{R}^n$

$$\|d\|_1 = \|(A^T I)^{-1}A^T d\|_1 \leq \|(A^T I)^{-1}\|_1 \|A^T d\|_1.$$

Therefore, $c^T \bar{x} - c^T x \geq \gamma \|\bar{x} - x\|$ holds with $\gamma = \varepsilon/\|(A^T I)^{-1}\|_1$. $\square$

First order maximizers have been well studied in semi-infinite programming, see e.g. [Fis91], [Nür85], [HT98]. It is a classical result, for example, that the optimal solution of the problem “Tchebycheff approximation by polynomials” always has a first order optimal solution, see e.g. [Che66]. In semi-infinite programming, one of the numerical approaches of solving SIP is the uniform discretization method. The rate of convergence of these uniform discretization methods depends partially on the order of maximizer, see [Sti01]. In copositive programming, the discretization method has been applied and analysed in [ADS13, Yil12].

In general, the order of an optimal solution does not have to be an integer. It is shown in [GLT95, Theorem 3.1] that if an optimal solution is of order $0 < p < 1$, then the feasible set is a singleton. Also, there are examples whose optimal solutions have any order $p \geq 1$ (not necessarily an integer), see [GLT95, Example 3.2]. Furthermore, it is possible that a problem has a unique optimal solution but no order. Let us illustrate this example with no order.

**Example 4.3.** [GLT95, Example 3.3] Consider the following problem in $\mathbb{R}^2$:

$$\min x_2 \quad \text{s.t. } x_2 \leq e^{-\frac{3}{2}} \text{ and } x_2 \geq f(x_1)$$

where

$$f(x_1) = \begin{cases} 
  e^{-\frac{1}{x_1}}, & x_1 \neq 0 \\
  0, & x_1 = 0.
\end{cases}$$

We can check that the problem has the unique optimal solution $\bar{x}^T = (\bar{x}_1, \bar{x}_2)^T = (0, 0)$. The shaded region in Figure 4.2 is the feasible set of the problem.

We show that there is no $p > 0$ exists such that $\bar{x}$ has order $p$. Suppose by contradiction that $\bar{x}$ was of some order $p \geq 1$. Then there exist $\gamma > 0$
and $\rho > 0$ such that

$$x_2 \geq \gamma ||x||^p \text{ for all } x \in \mathcal{F} \cap \mathcal{B}_\rho(0)$$

where $\mathcal{F}$ is the feasible set of the problem. Let us consider solutions on the boundary of the feasible set $\mathcal{F}$. For $\varepsilon > 0$ small enough, we have for $0 < x_1 < \varepsilon$ that

$$f(x_1) \geq \gamma ||x||^p > \gamma x_1^p.$$

Equivalently,

$$\frac{f(x_1)}{x_1^p} > \gamma \text{ for all } 0 < x_1 < \varepsilon.$$

However, this is a contradiction, as $\lim_{x_1 \to 0^+} f(x_1)/x_1^p = 0$.

In conic programming, the order of maximizers may differ depending on the geometry of the cone $\mathcal{K}$. In the following, we discuss the order of maximizers for some specific cones.

- As we have seen, in linear programming unique maximizers are of first order.

- In SDP or COP, maximizers of arbitrarily high order are possible. Let us give an example from [ADS13] for a 4th order maximizer:

$$\max x_1$$

s.t. $X = \begin{pmatrix} -x_1 & x_2 & 0 & 0 \\ x_2 & 1 & 0 & 0 \\ 0 & 0 & -x_2 & x_3 \\ 0 & 0 & x_3 & 1 \end{pmatrix} \in S^4_+.$$

Figure 4.2: A problem with no order
The optimal solution is \( \bar{x} = (0, 0, 0) \). Note that a feasible \( x \) must fulfill
\[
x_1 \leq -x_3^4, \quad x_1 \leq 0.
\]
These imply \( c^T \bar{x} - c^T x \geq x_3^4 = ||x||_\infty^4 \), so \( \bar{x} \) is a 4th order maximizer.

For the case that the cone \( \mathcal{K} \) is semi-algebraic (such as the semidefinite or the copositive cone), a genericity result with respect to the parameter \( c \) is given in [BDL11]:

**Proposition 4.4.** Let \( \mathcal{K} \) be a semi-algebraic cone and \( A, B \) be given such that \( \mathcal{F}_P \) is compact. Then there exists a generic set \( S \subset \mathbb{R}^n \) such that for all \( c \in S \) the corresponding program \( (P) \) has a second order maximizer.

In SDP, it is shown that uniqueness of an optimal solution is equivalent to second order maximizer under generic properties.

**Theorem 4.5.** [SS00, Theorem 4.2.1] Consider an SDP with optimal solution \( \bar{x} \). Assume that Slater’s condition and strict complementarity hold. Then \( \bar{x} \) is the unique optimal solution if and only if \( \bar{x} \) is of second order.

As we have seen in Proposition 3.20, generically Slater’s condition, strict complementarity and uniqueness are stable properties of SDP. Thus, the following can be obtained:

**Corollary 4.6.** Let parameter \( Q = (A, B, c) \in \mathcal{P} \) describe a generic instance of SDP such that unique primal and dual nondegenerate optimal solutions exist and satisfy strict complementarity. Then the optimal solution \( \bar{X} \) of \( (P) \) is a stable second order maximizer at \( Q \in \mathcal{P} \).

Combining the above corollary and Corollary 3.17, we have that second order maximizer is a generic property in the parameter space \( \mathcal{D}(\mathcal{L}) \) for SDP. Let us illustrate a simple example with stable second order (non-first order) maximizer in SDP.
**Example 4.7.** [SDP, stable second order (non-first order) maximizer] Consider the program

\[
(P) \quad \max (0,1)x \quad \text{s.t.} \quad X := \begin{pmatrix} 1-x_1 & -x_2 \\ -x_2 & 1+x_1 \end{pmatrix} \in S_+^2.
\]

Our problem data is given by

\[
\bar{c} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \quad \bar{B} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad \bar{A}_1 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \quad \bar{A}_2 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}
\]

It is not difficult to see that this program is equivalent to

\[
\max x_2 \quad \text{s.t.} \quad x_1^2 + x_2^2 \leq 1, \quad -1 \leq x_1 \leq 1.
\]

The solution \( \bar{x} = (0,1) \) or \( \bar{X} = \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix} \) is not of first order, but of second order. It is not difficult to verify that this second (non-first) order maximizer is stable.

### 4.1 First order solutions in conic programming

We have seen that second order maximizers are weakly generic in conic programming. In this section, we focus on first order maximizers (which are also of second order) and their characterizations. The next lemma is frequently used in the following discussion.

**Lemma 4.8.** Let \( \bar{X} \in \mathcal{F}_P, \bar{Y} \in \mathcal{F}_D \) be complementary solutions of \((P)\) and \((D)\). Let the KKT condition \( c = \sum_{j=1}^l y_j a(Y_j) \) hold with \( y_j > 0, Y_j \in J^\Delta(\bar{X}) \) and let \( \bar{Y} := \sum_{j=1}^l y_j Y_j \). Suppose that

\[
\text{span}\{Y_1, \ldots, Y_l\} = \text{span} J^\Delta(\bar{X}).
\]

Then

\[
\text{span}\{Y_1, \ldots, Y_l\} = \text{span} G(\bar{Y}).
\]

In particular, \( \bar{X} \) and \( \bar{Y} \) are strictly complementary.
4.1. First order solutions in conic programming

Proof. Observe that for the minimal face $G(\overline{Y})$ we have

$$G(\overline{Y}) = \{ Y \in \mathcal{K}^* \mid \overline{Y} \pm \lambda Y \in \mathcal{K}^* \text{ for some } \lambda > 0 \}.$$ 

Since the coefficients in the representation of $Y$ satisfy $y_j > 0$, we must have $Y_j \in G(\overline{Y})$ for all $j$. Moreover, $G(\overline{Y}) \subseteq J^\Delta(\overline{X})$ holds, for the proof see [PT01, p. 452]. These imply

$$\text{span}\{Y_1, \ldots, Y_l\} = \text{span} J^\Delta(\overline{X}) \subseteq \text{span} G(\overline{Y}) \subseteq \text{span} J^\Delta(\overline{X}),$$

so equality must hold for all these sets. We also have $\overline{Y} \in \text{ri} J^\Delta(\overline{X})$ since $\text{span}\{Y_1, \ldots, Y_l\} = \text{span} J^\Delta(\overline{X})$. By definition of strict complementarity (2.17), the complementary solutions $\overline{X}$ and $\overline{Y}$ are strictly complementary. \qed

Next, let us translate well-known characterizations of first order solutions from semi-infinite programming to the special case of conic programs. Consider the general SIP problem (SIP$_P$) and denote its feasible set as $\mathcal{F}_{SIP}$ and the active index set as

$$I(\overline{x}) := \{ Y \in \mathcal{Z} \mid b(Y) - a(Y)^T \overline{x} = 0 \}.$$ 

Recall that we identify $\overline{x}$ with $\overline{X} = b(Y) - a(Y)^T \overline{x}$. The following necessary and sufficient conditions for first order maximizers are well-known, see e.g. [GLT95, Theorem 4.1] or [Fis91, Theorem 3.1].

**Theorem 4.9.** [SIP-result] Let $\overline{x} \in \mathcal{F}_{SIP}$ and consider the moment cone $\mathcal{M}(\overline{x}) := \text{cone}\{ a(Y_j) \mid Y_j \in I(\overline{x}) \}$. We have the following

1. If $c \in \text{int} \mathcal{M}(\overline{x})$, then $\overline{x}$ is a first order maximizer. Conversely, if the Slater condition holds for (SIP$_P$) and $\overline{x}$ is a first order maximizer, then $c \in \text{int} \mathcal{M}(\overline{x})$.

2. The following conditions are equivalent.

   (a) $c \in \text{int} \mathcal{M}(\overline{x})$
Chapter 4. Order of maximizers

\[(b) \quad c = \sum_{j=1}^{l} y_j a(Y_j) \text{ with } y_j > 0, \quad Y_j \in I(\bar{x}), \quad \text{and} \quad \text{span}\{a(Y_1), \ldots, a(Y_l)\} = \mathbb{R}^n.\]

Note that in the conic formulation, the active index set is precisely the complementary face \(J^\Delta(X)\). In order to formulate these conditions in terms of conic programs, we need an auxiliary lemma.

**Lemma 4.10.** Let \(A_1, \ldots, A_n \in S^k\) be linearly independent matrices and let \(Y_1, \ldots, Y_l \in S^k\). Define \(\mathcal{L} := \text{span}\{A_1, \ldots, A_n\}, \mathcal{R} := \text{span}\{Y_1, \ldots, Y_l\}\), and \(T := (\langle A_i, Y_j \rangle)_{i,j} \in \mathbb{R}^{n \times l}\). Then we have:

\[
\text{rank } T = n \iff \mathcal{L} \cap \mathcal{R}^\perp = \{0\} \quad \text{and} \quad \dim \mathcal{R} \geq n + \dim(\mathcal{L}^\perp \cap \mathcal{R}).
\]

**Proof.** Denote \(i_1 := \dim(\mathcal{L} \cap \mathcal{R}^\perp)\) and \(j_1 := \dim(\mathcal{L}^\perp \cap \mathcal{R})\). Without loss of generality, we can assume that the matrices \(Y_j\) are linearly independent, and that they are ordered according to

\[
\text{span}\{Y_1, \ldots, Y_{l-j_1}\} \cap \mathcal{L}^\perp = \{0\} \quad \text{and} \quad \text{span}\{Y_{l-j_1+1}, \ldots, Y_l\} \subseteq \mathcal{L}^\perp.
\]

Then \(\mathcal{R}\) is decomposed as

\[
\mathcal{R} = \text{span}\{Y_1, \ldots, Y_{l-j_1}\} \oplus \text{span}\{Y_{l-j_1+1}, \ldots, Y_l\}.
\]

Similarly, we may assume that

\[
\text{span}\{A_1, \ldots, A_{n-i_1}\} \cap \mathcal{R}^\perp = \{0\} \quad \text{and} \quad \text{span}\{A_{n-i_1+1}, \ldots, A_n\} \subseteq \mathcal{R}^\perp.
\]

Then \(\mathcal{L}\) is decomposed as

\[
\mathcal{L} = \text{span}\{A_1, \ldots, A_{n-i_1}\} \oplus \text{span}\{A_{n-i_1+1}, \ldots, A_n\}.
\]

By removing all zero rows and columns from \(T\), we obtain a matrix \(\tilde{T} := (\langle A_i, Y_j \rangle)_{i,j}\) for \(i = 1, \ldots, n-i_1\) and \(j = 1, \ldots, l-j_1\) which obviously has \(\text{rank } \tilde{T} = \text{rank } T\).

\((\Rightarrow): \) If \(n = \text{rank } T = \text{rank } \tilde{T}\), then we immediately get that \(n - i_1 \geq n\).
and \( l - j_1 \geq n \). Hence \( i_1 = 0 \), which implies that \( \mathcal{L} \cap \mathcal{R}^\perp = \{0\} \), and from \( l - j_1 \geq n \) we conclude that \( \dim \mathcal{R} \geq n + \dim(\mathcal{L}^\perp \cap \mathcal{R}) \).

(\( \Leftarrow \)): Let \( \mathcal{L} \cap \mathcal{R}^\perp = \{0\} \) and \( \dim \mathcal{R} \geq n + \dim(\mathcal{L}^\perp \cap \mathcal{R}) \). To prove that \( \text{rank } T = n \), it is sufficient to show that \( d^T T = 0 \) implies \( d = 0 \). By contradiction, assume that \( d \neq 0 \) solves \( d^T T = 0 \). Then

\[
0 \neq D := \sum_{i=1}^{n} d_i A_i \in \mathcal{L}.
\]

Since \( \mathcal{L} \cap \mathcal{R}^\perp = \{0\} \), we have \( D \notin \mathcal{R}^\perp \). Hence there exists an index \( j_0 \) such that for the corresponding unit basis vector \( e_{j_0} \) we have

\[
0 \neq \langle D, Y_{j_0} \rangle = d^T T e_{j_0}.
\]

Therefore, we have \( d^T T \neq 0 \), a contradiction. \( \square \)

We can now restate the SIP-condition from Theorem 4.9 (2) in terms of conic programming.

**Theorem 4.11.** The following conditions are equivalent for \( \overline{X} \) in \( \mathcal{F}_P \).

(a) \( c \in \text{int } \mathcal{M}(\overline{X}) \).

(b) There exist \( Y_j \in J^\Delta(\overline{X}) \), \( j = 1, \ldots, l \), and multipliers \( y_j > 0 \) such that \( \overline{Y} := \sum_{j=1}^{l} y_j Y_j \) is an optimal solution of (D), and we have

\[
\mathcal{L} \cap \mathcal{R}^\perp = \{0\} \quad \text{and} \quad \dim \mathcal{R} \geq n + \dim(\mathcal{L}^\perp \cap \mathcal{R}). \quad (4.1)
\]

Here again \( \mathcal{L} = \text{span}\{A_1, \ldots, A_n\} \) and \( \mathcal{R} = \text{span}\{Y_1, \ldots, Y_l\} \).

**Proof.** The proof follows directly from Theorem 4.9 (2) and Lemma 4.10 by noticing that \( a(Y_1), \ldots, a(Y_l) \) are the columns of the matrix \( T \). \( \square \)

Suppose now that \( \overline{X} \in \mathcal{F}_P \) and \( \overline{Y} \in \mathcal{F}_D \) satisfy the conditions of Theorem 4.11 (b). By the arguments in the proof of Lemma 4.8 the relations

\[
\mathcal{R} \subseteq \text{span } G(\overline{Y}) \subseteq \text{span } J^\Delta(\overline{X}) \quad (4.2)
\]

hold, and the inequality in (4.1) implies that \( n \leq \dim \mathcal{R} \). In view of Theorem 4.9, this immediately yields the following
Chapter 4. Order of maximizers

**Corollary 4.12.** Consider problem $(P)$ satisfying the Slater condition. If $\overline{X} = B - \sum_{i=1}^{n} \overline{x}_i A_i$ is a first order maximizer of $(P)$, then there exists an optimal solution $\overline{Y}$ of $(D)$ with $\dim G(\overline{Y}) \geq n$.

Moreover, for a nondegenerate optimal solution, the first order characterization can be stated as follows

**Corollary 4.13.** Let $\overline{X} = B - \sum_{i=1}^{n} \overline{x}_i A_i \in \mathcal{F}_P$ be a nondegenerate maximizer of $(P)$ and let $\overline{Y}$ be the optimal solution of $(D)$ which is unique by Lemma 2.25(b). Then $\overline{X}$ is a first order maximizer if and only if $\dim G(\overline{Y}) = n$ and $\overline{X}, \overline{Y}$ are strictly complementary.

**Proof.** $(\Rightarrow)$ From Proposition 2.21, we have that if there exists a nondegenerate $\overline{X} \in \mathcal{F}_P$, then the primal Slater condition holds. By Theorem 4.9 (1), we have $c \in \text{int} \mathcal{M}(\overline{X})$ and thus the conditions of Theorem 4.11 (b) hold. As $\overline{X}$ is nondegenerate, we have by definition that $\mathcal{L}^\perp \cap \text{span} J^\Delta(\overline{X}) = \{0\}$. Since $\dim \mathcal{L}^\perp = \frac{1}{2}k(k + 1) - n$, we obtain $\dim J^\Delta(\overline{X}) \leq n$. Combining this with (4.2) and Corollary 4.12, we conclude that if $\overline{X}$ is nondegenerate, then

$$n \leq \dim \mathcal{R} \leq \dim G(\overline{Y}) \leq \dim J^\Delta(\overline{X}) \leq n.$$  

This implies $\mathcal{R} = \text{span} G(\overline{Y}) = \text{span} J^\Delta(\overline{X})$, and by Lemma 4.8, the matrices $\overline{X}$ and $\overline{Y}$ are strictly complementary.

$(\Leftarrow)$ By strict complementarity (2.20), we have $G(\overline{Y}) = J^\Delta(\overline{X})$. Using $\overline{Y} \in \text{ri} J^\Delta(\overline{X})$, there exist $Y_j \in J^\Delta(\overline{X})$, $y_j > 0$ ($j = 1, \ldots, l$) such that $\overline{Y} = \sum_{j=1}^{l} y_j Y_j$ and

$$\mathcal{R} := \text{span}\{Y_1, \ldots, Y_l\} = \text{span} J^\Delta(\overline{X}) \quad (4.3)$$

with $\dim \mathcal{R} = \dim G(\overline{Y}) = n$. By Theorems 4.9 and 4.11, it suffices to show that (4.1) is satisfied. Nondegeneracy of $\overline{X}$ together with (4.3) implies that $\{0\} = \mathcal{L}^\perp \cap \text{span} J^\Delta(\overline{X}) = \mathcal{L}^\perp \cap \mathcal{R}$. Using $\dim \mathcal{R} = n$, we see that the second condition in (4.1) is valid. Since $\dim \mathcal{L}^\perp = \frac{1}{2}k(k + 1) - n$ and $\dim \mathcal{R} = n$, we get from $\{0\} = \mathcal{L}^\perp \cap \mathcal{R}$ that $\mathcal{L}^\perp + \mathcal{R} = S^k$, and thus $\mathcal{L} \cap \mathcal{R}^\perp = \{0\}$. □
4.1. First order solutions in conic programming

Remark 4.14. Note that the conditions for first order maximizers in Theorem 4.11 (b) allow that \( \dim \mathcal{R} > n \), in which case \( \overline{X} \) is degenerate.

Remark 4.15. From Corollary 4.13, we conclude that if \( \overline{X} \) is a nondegenerate first order maximizer of \((P)\), then for the (unique) solution \( \overline{Y} \) of \((D)\) the strict complementary holds. For non-first order maximizers this need not be the case, even if both \( \overline{X} \) and \( \overline{Y} \) are nondegenerate, see [AHO97, p.117] for a counterexample in SDP.

We now present some examples of first order maximizers in SDP and COP. Note that it is geometrically clear that for the case \( n = 1 \), i.e. \( \dim \mathcal{L} = 1 \), under the conditions \( \emptyset \neq \mathcal{F}_P \not\subset \text{bd} \mathcal{K} \) and \( c \neq 0 \), any maximizer is of first order. So we give examples with \( n \geq 2 \).

**Example 4.16.** [SDP with \( n = k = 2 \)] Consider

\[
(P) \quad \text{max} \, (1, 1)x \quad \text{s.t.} \quad X := \begin{pmatrix} 1-x_1 & 0 \\ 0 & 1-x_2 \end{pmatrix} \in \mathcal{S}_+^2.
\]

\[
(D) \quad \text{min} \left\langle \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, Y \right\rangle \quad \text{s.t.} \quad Y_{11} = 1, \ Y_{22} = 1, \ Y \in \mathcal{S}_+^2.
\]

The solution of \((P)\) is \( \overline{x} = (1, 1) \ or \ \overline{X} = 0 \), and all matrices of the form

\[
Y = \begin{pmatrix} 1 & Y_{12} \\ Y_{12} & 1 \end{pmatrix}
\]

with \(-1 \leq Y_{12} \leq 1\) are optimal solutions of \((D)\). With \( J^\Delta(\overline{X}) = \mathcal{S}_+^2 \), we can see that the assumptions of Theorem 4.11 (b) are satisfied for

\[
\overline{Y} := \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = 1 \cdot Y_1 + 1 \cdot Y_2
\]

with

\[
Y_1 = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \ Y_2 = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \in J^\Delta(\overline{X}).
\]

So \( \overline{X} = 0 \) is a first order maximizer. However, for the solution \( \overline{Y} \) of \((D)\) we have here \( \dim \mathcal{G}(\overline{Y}) = 3 \), so \( \overline{X} \) must be degenerate by Corollary 4.13. We will see later that the first order maximizer \( \overline{X} \) is not stable in this example, see Example 4.22.
Example 4.17. \textit{[COP with } \( n = k = 2 \text{]} \quad \text{It is known that}
\[
\text{COP}^2 = \mathcal{N} \mathcal{N}^2 + \mathcal{S}^2_+ \quad \text{and} \quad (\text{COP}^2)^* = \mathcal{C} \mathcal{P}^2 = \mathcal{N} \mathcal{N}^2 \cap \mathcal{S}^2_+.
\]

Consider
\[
(P) \quad \text{max}(1, 1)x \quad \text{s.t.} \quad \begin{pmatrix} 1 - x_1 & 1 \\ 1 & 1 - x_2 \end{pmatrix} \in \text{COP}^2,
\]
\[
(D) \quad \text{min} \left\langle \begin{pmatrix} 1 \\ 1 \end{pmatrix}, Y \right\rangle \quad \text{s.t.} \quad Y_{11} = 1, Y_{22} = 1, Y \in \mathcal{C} \mathcal{P}^2.
\]

The dual can be rewritten as:
\[
(D) \quad \text{min}(2 + 2Y_{12}) \quad \text{s.t.} \quad Y = \begin{pmatrix} 1 & Y_{12} \\ Y_{12} & 1 \end{pmatrix} \in \mathcal{N} \mathcal{N}^2 \cap \mathcal{S}^2_+.
\]

The feasibility condition for \((D)\) is given by \(0 \leq Y_{12} \leq 1\) and the feasibility condition for \((P)\) by \(x_i \leq 1\) for \(i = 1, 2\). So \(\bar{x} = (1, 1)\) or \(\bar{X} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}\) is the optimal solution of \((P)\) with
\[
J^\Delta(\bar{X}) = \{ \alpha \cdot Y_1 + \beta \cdot Y_2 \mid \alpha, \beta \geq 0 \}
\]

where
\[
Y_1 := \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \quad Y_2 := \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \in J^\Delta(\bar{X}).
\]

The unique solution of \((D)\) is
\[
\bar{Y} := \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = 1 \cdot Y_1 + 1 \cdot Y_2,
\]
so again the conditions of Thereom 4.11 (b) are satisfied and \(\bar{X}\) is a first order maximizer. It is not difficult to see that \(\bar{Y}\) is also a first order minimizer and that both first order solutions are stable.

The next example is similar to Example 4.16 but with a stable first order maximizer.
Example 4.18. [SDP with $n = 3, k = 2$] Consider

\[
(P) \quad \max (1, 1, 1)x \quad \text{s.t.} \quad \begin{pmatrix} 1 - x_1 & 1 - x_3 \\ 1 - x_3 & 1 - x_2 \end{pmatrix} \in \mathcal{S}^2_+,
\]

\[
(D) \quad \min \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}, Y \quad \text{s.t.} \quad Y_{11} = 1, Y_{22} = 1, Y_{12} = \frac{1}{2}, Y \in \mathcal{S}^2_+.
\]

The maximizer of $(P)$ is $\bar{x} = (1, 1, 1)$ (or $X = 0$) and the minimizer of $(D)$ is

\[
Y = \begin{pmatrix} 1 & 1/2 \\ 1/2 & 1 \end{pmatrix} \in \text{int} \mathcal{S}^2_+
\]

with $\dim G(Y) = 3$. It is not difficult to see that the optimal solutions $\bar{X}$ of $(P)$ and $\bar{Y}$ of $(D)$ are both of first order, as they are given by the unique solutions of the linear systems

\[
B - x_1 A_1 - x_2 A_2 - x_3 A_3 = 0 \quad \text{and} \quad \langle A_i, Y \rangle = c_i, \ i = 1, 2, 3.
\]

Also here, the first order solutions $\bar{X}$ and $\bar{Y}$ are stable.

Semidefinite programming: For the SDP case, the result in Corollary 4.13 can be further specified as one can determine the dimension of a particular face.

Corollary 4.19. Let $\bar{X} = B - \sum_{i=1}^n \pi_i A_i \in \mathcal{S}^k_+$ be a nondegenerate maximizer of $(SDP_P)$ with rank $\bar{X} = r$ and let $\bar{Y} \in \mathcal{S}^k_+$ be the minimizer of $(SDP_D)$ with rank $\bar{Y} = s$. Then $\bar{X}$ is of first order if and only if

\[
\frac{1}{2} s(s + 1) = \frac{1}{2} (k - r)(k - r + 1) = n.
\]

Proof. From Proposition 2.31, we know that strict complementarity is equivalent to $k = r + s$, and $\dim G(Y) = \frac{1}{2} s(s + 1)$ holds by (2.26). By applying Corollary 4.13, the result follows directly. \qed

By this formula, nondegenerate first order maximizers are excluded for many choices of $n$. The above formulation provides a necessary condition:
Corollary 4.20. Let $\overline{X} = B - \sum_{i=1}^{n} \overline{x}_i A_i \in S^k_+$ be a nondegenerate first order maximizer of $(SDP_P)$. Then there exists $s \in \mathbb{N}$ such that

$$\frac{1}{2}s(s + 1) = n. \quad (4.5)$$

Table 4.1 gives a list of pairs $(s, n)$ such that $\frac{1}{2}s(s + 1) = n.$

<table>
<thead>
<tr>
<th>$s$</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
</tr>
</thead>
<tbody>
<tr>
<td>$n$</td>
<td>1</td>
<td>3</td>
<td>6</td>
<td>10</td>
<td>15</td>
<td>21</td>
</tr>
</tbody>
</table>

Table 4.1: Pairs $(s, n)$ such that (4.5) holds.

Note that for the numbers $n$ in Table 4.1, stable first order maximizers are possible for all $k \geq s$. Recall that nondegeneracy and strict complementarity are generically fulfilled in SDP and hence first order maximizers of ($P$) are generically excluded for numbers $n$ not satisfying the condition of (4.5).

The cones $COP^k$ and $CP^k$ have much richer structure than $S^k_+$. Therefore, in contrast to SDP, we expect that in COP stable first order maximizers occur for any $n$. For instance, the maximizer in Example 4.17 with $n = k = 2$ is stable.

We close this section with some remarks on related properties. The tangent space given in (2.21) is the subspace of directions where the boundary of $K$ at $\overline{X}$ is smooth. The smaller the dimension of $\text{tan} (\overline{X}, K)$, the higher the non-smoothness (“kinkiness”) of $K$ at $\overline{X}$. Let us introduce the order of kinkiness. Defining $m := \frac{1}{2}k(k + 1)$, we say that the cone $K$ has a kink at $\overline{X} \in K$ of order $l = m - p$ (i.e. a kink of co-order $p$), if $\dim(\text{tan}(\overline{X}, K)) = p$. Clearly, there is a relation between the order of a kink at $\overline{X} \in K$ and the fact that $\overline{X}$ is a first order maximizer. Assume $K$ is a nice cone (cf. Definition 2.26). Let ($P$) satisfy the Slater condition and let $\overline{X}$ be a first order maximizer with unique strict complementary solution $\overline{Y} \in F_D$. Then by Corollary 4.12, we obtain that

$$\dim J^\Delta(\overline{X}) = \dim G(\overline{Y}) \geq n.$$
4.2 Stability of first order maximizers

Using the fact that $\mathcal{K}$ is nice and equation (2.22), this implies
\[
\dim(\tan(\overline{X}, \mathcal{K})) = \dim(J^\Delta(\overline{X})^\perp) \leq m - n.
\]

If moreover $\overline{X}$ is nondegenerate, we deduce from Corollary 4.13 that $\dim(G(\overline{Y})) = n$ and thus
\[
\dim(\tan(\overline{X}, \mathcal{K})) = \dim(J^\Delta(\overline{X})^\perp) = m - n.
\]

4.2 Stability of first order maximizers

In this section, we investigate the stability of first order maximizers of conic programs. As mentioned before, in SDP generically the optimal solutions are at least of order two. However, stable first order maximizers do occur, see Example 4.18. In this respect, first order maximizers are especially “nice” sharp maximizers which also may occur in the generic situation.

In semi-infinite optimization (SIP), starting with the paper [Nür85], the stability of first order maximizers of SIP was studied in several papers, see e.g. [HT98, GTVdS12]. In [Nür85, HT98], the set
\[
\Pi := \{(a, b, c) \in C(\mathcal{Z})^n \times C(\mathcal{Z}) \times \mathbb{R}^n\}
\]
has been considered as the set of input data for SIP of the general form $(\text{SIP}_\rho)$, where $\mathcal{Z} \subset \mathbb{R}^l$ and $C(\mathcal{Z})$ denotes the set of continuous functionals on $\mathcal{Z}$. In these papers, the set $\Pi$ is endowed with the topology given by
\[
d((a, b, c), (\tilde{a}, \tilde{b}, \tilde{c})) := \max_{Y \in \mathcal{Z}} ||a(Y) - \tilde{a}(Y)|| + \max_{Y \in \mathcal{Z}} |b(Y) - \tilde{b}(Y)| + ||c - \tilde{c}||.
\]

Consider the general SIP problem $(\text{SIP}_\rho)$ and let $\overline{x}$ be the maximizer of $(\text{SIP}_\rho)$. Before discussing genericity results for general SIP, let us define the subsets $\mathcal{U}, \mathcal{U}_1 \subset \Pi$ as
\[
\mathcal{U} := \{(a, b, c) \in \Pi \mid \overline{x} \text{ of the corresponding SIP is unique}\},
\]
\[
\mathcal{U}_1 := \{(a, b, c) \in \Pi \mid \overline{x} \text{ of the corresponding SIP is of first order}\}.
\]

73
As before, denote its feasible set of \((\text{SIP}_p)\) as \(\mathcal{F}_{\text{SIP}}\) and the active index set for \(\bar{x} \in \mathcal{F}_{\text{SIP}}\) as \(I(\bar{x}) = \{Y \in \mathcal{Z} \mid b(Y) - a(Y)^T \bar{x} = 0\}\). For general linear SIP, the following stability results have been proven in [N"ur85, HT98]

**Theorem 4.21.** Let \(\pi = (a, b, c) \in \Pi\) satisfy the Slater condition and let \(0 \neq c \in \mathbb{R}^n\). Then the following statements are equivalent:

(a) \(\pi \in \text{int} \mathcal{U}_1\),

(b) \(\pi \in \text{int} \mathcal{U}\),

(c) There exist \(\bar{x} \in \mathcal{F}_{\text{SIP}}\) and \(Y_j \in I(\bar{x}), y_j > 0 (j = 1, \ldots, n)\) such that \(c = \sum_{j=1}^{n} y_j a(Y_j)\), and for every selection \(\tilde{Y}_j \in I(\bar{x}) (j = 1, \ldots, n)\) with \(c = \sum_{j=1}^{n} \tilde{y}_j a(\tilde{Y}_j), \tilde{y}_j \geq 0\), any subset of \(n\) vectors from

\[
c, a(\tilde{Y}_j), (j = 1, \ldots, n) \quad \text{are linearly independent.} \quad (4.6)
\]

As we mentioned before in Section 3.7, the parameter space \(\mathcal{P}\) as defined in (3.1) of conic programs given by the linear functions \(a(Y), b(Y)\) in (2.1) is only a small subset of the parameter set \(\Pi\) of general SIP problems. It is clear that the topology in \(\Pi\) allows more (also nonsmooth) perturbations and the stability conditions in SIP are too strong for conic programming.

For example, the conditions in Theorem 4.21 (c) can never hold at first order maximizers of conic problems if \(n > 1\): Note that the condition (4.6) states in particular that the KKT relations \(c = \sum_{j=1}^{l} y_j a(Y_j)\) cannot be fulfilled with \(l < n\). However, given any first order optimizer \(\varpi\) (or \(\overline{X}\)) of a conic program, the linearity of \(b(Y) - a(Y)^T x = \langle (B - \sum_i x_i A_i), Y \rangle\) in \(Y\) implies that if \(c = \sum_{j=1}^{l} y_j a(Y_j)\) with \(Y_j \in I(\overline{X}), y_j > 0\) for \(j = 1, \ldots, l\), then with the solution \(\overline{Y} := \sum_{j=1}^{l} y_j Y_j\) of (D) we have (assuming \(\overline{Y} \neq 0\)):

\[
\tilde{Y} := \frac{1}{\|\overline{Y}\|} \overline{Y} \in I(\overline{X}) \quad \text{and} \quad c = \|\overline{Y}\| \cdot a(\tilde{Y}).
\]
4.2. Stability of first order maximizers

Consequently, the KKT condition holds with \( l = 1 \) and for \( n > 1 \) the condition (4.6) in Theorem 4.21 (c) always fails. Moreover, the equivalence of conditions (a) and (b) of Theorem 4.21 does not hold as there are stable second order (not first) maximizers in conic programming, see Example 4.7.

Next, we discuss how the conditions for stability of first order maximizers in Theorem 4.21 (c) have to be modified in conic programming in order to obtain necessary and sufficient conditions for the stability of first order maximizers in conic programming. As an illustrative example, we go back to the SDP problem of Example 4.16.

**Example 4.22.** [Example 4.16 continued] The SDP instance with \( n = k = 2 \) is given by \( \mathcal{Q} = (\overline{c}, \overline{B}, \overline{A}_1, \overline{A}_2) \) with

\[
\overline{c} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \quad \overline{B} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad \overline{A}_1 = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \quad \overline{A}_2 = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}.
\]

Recall that the primal maximizer wrt. \( \overline{Q} \) is given by \( \overline{X} = 0, \; \overline{x} = (1, 1) \) and the corresponding set of dual optimal solutions is

\[
\mathcal{F}_D^*(\overline{X}) := \left\{ \overline{Y} = \begin{pmatrix} 1 & Y_{12} \\ Y_{12} & 1 \end{pmatrix} \mid -1 \leq Y_{12} \leq 1 \right\}.
\]

For \( Y_{12} = \pm 1 \), we obtain solutions on the boundary of \( S_2^+ \):

\[
\overline{Y}_+ := \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}, \quad \overline{Y}_- := \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix}.
\]

Both are extreme rays satisfying \( \dim G(\overline{Y}_{\pm}) = 1 < n \). Choose one of them, say \( \overline{Y}_+ \), and consider for small \( \varepsilon > 0 \) the perturbed instance

\[
\mathcal{Q}_\varepsilon = (\overline{c}, B_\varepsilon, \overline{A}_1, \overline{A}_2) \quad \text{with} \quad B_\varepsilon = \overline{B} + \varepsilon \overline{Y}_+ = \begin{pmatrix} 1 + \varepsilon & \varepsilon \\ \varepsilon & 1 + \varepsilon \end{pmatrix}.
\]

The primal feasibility condition now changes from \( x_1 \leq 1, \; x_2 \leq 1 \) (for \( \overline{Q} \) which corresponds to \( \varepsilon = 0 \)) to

\[
B_\varepsilon - x_1 \overline{A}_1 - x_2 \overline{A}_2 = \begin{pmatrix} 1 + \varepsilon - x_1 & \varepsilon \\ \varepsilon & 1 + \varepsilon - x_2 \end{pmatrix} \in S_2^+
\]
for $Q_\varepsilon$, i.e.

\[
\begin{align*}
  x_1 &\leq 1 + \varepsilon \\
  x_2 &\leq 1 + \varepsilon \\
  (1 + \varepsilon - x_1)(1 + \varepsilon - x_2) &\geq \varepsilon^2.
\end{align*}
\]

The primal maximizer wrt. $Q_\varepsilon$ is now given by

\[
\bar{x}_\varepsilon = (1, 1), \quad \bar{X}_\varepsilon = \begin{pmatrix} \varepsilon & \varepsilon \\ \varepsilon & \varepsilon \end{pmatrix}
\text{ with } J^\Delta(\bar{X}_\varepsilon) = \left\{ \gamma \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix} | \gamma \geq 0 \right\}.
\]

The corresponding dual problem has the unique minimizer

\[
\bar{Y}_\varepsilon = \bar{Y}_- = \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix}
\text{ with } \dim G(\bar{Y}_\varepsilon) = 1 < n.
\]

Note that the primal program satisfies the Slater condition, so by Corollary 4.12 the solution $\bar{x}_\varepsilon$ is not a first order minimizer for $\varepsilon > 0$. The maximizer $\bar{x}_\varepsilon$ is of second order for any $\varepsilon > 0$ and the first order maximizer $\bar{x}$ is unstable at $\bar{Q}$, as displayed in Figure 4.3.

![Figure 4.3: Perturbed and unperturbed feasible set](image)

Let us now consider a general instance $\bar{Q} \in \mathcal{P}$. If $\bar{Q}$ satisfies the primal Slater condition and a solution $\bar{X}$ of $(P)$ exists, then by strong duality a complementary dual solution $\bar{Y}$ exists and thus the set

\[
\mathcal{F}_D^*(\bar{X}) := \{ \bar{Y} \in \mathcal{F}_D | \langle \bar{Y}, \bar{X} \rangle = 0 \}
\]

of dual optimal solutions is nonempty and compact, see e.g. [GL98a, Theorem 9.8]). The analysis in Example 4.22 now suggests to replace the condition in Theorem 4.21 (c) by the following condition.
4.2. Stability of first order maximizers

**C1.** There is no dual optimal solution \( \bar{Y} \in \mathcal{F}_D^*(\bar{X}) \) with \( \dim F < n \), where \( F \subset \mathcal{K}^* \) is the minimal exposed face containing \( G(\bar{Y}) \).

Note that since the set \( J^\Delta(\bar{X}) \) is an exposed face of \( \mathcal{K}^* \) and \( G(\bar{Y}) \subset J^\Delta(\bar{X}) \), this minimal face \( F \) must satisfy \( F \subset J^\Delta(\bar{X}) \). It turns out that condition **C1** is necessary for the stability of the first order maximizer.

**Theorem 4.23.** Assume that \( \bar{Q} = (\bar{c}, \bar{B}, \bar{A}_1, \ldots, \bar{A}_n) \in \mathcal{P} \) satisfies the primal Slater condition, and let \( \bar{X} \) be the first order maximizer of the corresponding primal program. If the first order maximizer is stable, then condition **C1** holds.

**Proof.** Assume by contradiction that there exist \( Y \in \mathcal{F}_D^*(\bar{X}) \) and a minimal exposed face \( F \) such that \( G(\bar{Y}) \subset F \subset J^\Delta(\bar{X}) \) and \( \dim F < n \). Since \( F \subset \mathcal{K}^* \) is an exposed face, there exists a supporting hyperplane \( H = \{ Y \in S^k \mid \langle S, Y \rangle = 0 \} \) with normal vector \( 0 \neq S \in \mathcal{K} \) such that

\[
\langle S, Y \rangle = 0 \quad \text{for all } Y \in F,
\]

\[
\langle S, Y \rangle > 0 \quad \text{for all } 0 \neq Y \in \mathcal{K}^* \setminus F.
\]  

(4.7)

For small \( \varepsilon > 0 \) consider the perturbed instance

\[
Q_\varepsilon = (\bar{c}, B_\varepsilon, \bar{A}_1, \ldots, \bar{A}_n) \quad \text{with} \quad B_\varepsilon = \bar{B} + \varepsilon S.
\]

Let \( \bar{X}_\varepsilon := \bar{X} + \varepsilon S \in \mathcal{K} \). Using (4.7) and \( F \subset J^\Delta(\bar{X}) \), we have the following for any \( Y \in \mathcal{K}^* \):

\[
\langle \bar{X}_\varepsilon, Y \rangle = \langle \bar{X}, Y \rangle + \varepsilon \langle S, Y \rangle = 0 \quad \Leftrightarrow \quad Y \in F.
\]

So wrt. \( Q_\varepsilon \) the matrix \( \bar{X}_\varepsilon \) is a maximizer with dual complementary solution \( \bar{Y} \) such that \( G(\bar{Y}) \subset F = J^\Delta(\bar{X}_\varepsilon) \). Furthermore, note that any dual solution \( \tilde{Y} \) with respect to \( Q_\varepsilon \) must be contained in \( F = J^\Delta(\bar{X}_\varepsilon) \) and thus satisfies

\[
\dim G(\tilde{Y}) \leq \dim F < n.
\]

Since the primal Slater condition holds at \( \bar{Q} \), it also holds at \( Q_\varepsilon \) for \( \varepsilon \) small enough. So by Corollary 4.12, the maximizer \( \bar{X}_\varepsilon \) cannot be of first order, i.e. first order stability fails. \( \square \)
Chapter 4. Order of maximizers

We now turn to sufficient conditions for stability of a first order maximizer $\overline{X}$ of a conic problem. We look for natural assumptions such as nondegeneracy of $\overline{X}$ or conditions on $G(\overline{Y})$. Let again $\overline{Q} \in \mathcal{P}$ be fixed with a first order maximizer $\overline{X}$ and a unique complementary minimizer $\overline{Y}$ with $\dim G(\overline{Y}) = n$. Let $\overline{Q}$ satisfy the Slater condition and assume that uniqueness of the primal and dual optimizers are stable at $\overline{Q}$, i.e. for $Q$ in an neighbourhood of $\overline{Q}$ there are unique complementary solutions $X = X(Q), Y = Y(Q)$. Since the Slater condition is stable, we infer from Corollary 4.12 that there exists an $\varepsilon > 0$ such that for any first order maximizer $X = X(Q)$ we have

$$\dim G(Y) \geq n = \dim G(\overline{Y}) \quad \text{for} \quad Y = Y(Q), \|Q - \overline{Q}\| < \varepsilon.$$  

This means that the function $\dim G(Y)$ is lower semicontinuous at $\overline{Y}$. So we consider the lower semicontinuity of the set-valued minimal face mapping

$$G : \mathcal{K}^* \rightrightarrows \mathcal{K}^*, \quad G(Y) = \text{face}(Y, \mathcal{K}^*).$$

**Definition 4.24.** The set-valued mapping $G : \mathcal{K}^* \rightrightarrows \mathcal{K}^*$ is called lower semicontinuous (lsc) at $\overline{Y} \in \mathcal{K}^*$, if for any open set $V \subset \mathcal{S}^k$ there exists $\delta > 0$ such that

$$G(\overline{Y}) \cap V \neq \emptyset \quad \Rightarrow \quad G(Y) \cap V \neq \emptyset \quad \text{for all} \quad Y \in \mathcal{K}^* \text{ with } \|Y - \overline{Y}\| < \delta.$$  

It is easy to see that lower semicontinuity of $G$ implies lower semicontinuity of $\dim G(Y)$: if $G$ is lsc at $\overline{Y}$, then there exists $\delta > 0$ such that $\dim G(Y) \geq n = \dim G(\overline{Y})$ for all $Y \in \mathcal{K}^*$ with $\|Y - \overline{Y}\| < \delta$. Moreover, it is shown in [Pap77, Theorem 2.3] that lower semi-continuity of $G$ and lower semi-continuity of $\dim G(Y)$ are equivalent. By these arguments it is clear that the lower semicontinuity of $G$ is a natural condition for stable first order maximizers at $\overline{Q}$.

We shortly discuss the lower semicontinuity of the face mapping $G$ before turning to sufficient conditions of first order maximizers. This condition depends on the structure of the cone $\mathcal{K}^*$. 

78
4.2. Stability of first order maximizers

**Definition 4.25.** [Pap77, Definition 3.1] The closed convex cone $\mathcal{K}^*$ is called stable at $Y_0 \in \mathcal{K}^*$, if the mapping $h : \mathcal{K}^* \times \mathcal{K}^* \to \mathcal{K}^*$,

$$h(V, W) = \frac{1}{2}(V + W)$$

is open at $Y_0$.

It has been shown in [Pap77, Proposition 3.3] that if a closed convex set $\mathcal{O}$ of a topological vector space is stable at $Z \in \mathcal{O}$, then the corresponding minimal face map $G : \mathcal{O} \rightrightarrows \mathcal{O}$ is lower semi-continuous at $Z$. Also, it is shown that lower semi-continuity of the minimal face mapping implies the closedness of the so-called $q$-skeletons, the set of all points $x \in \mathcal{O}$ such that the minimal face $F$ containing $x$ has dimension less than or equal to $q$ (see [Pap77]). For compact convex sets in the Euclidean space, the stability of $\mathcal{O}$, the lower semicontinuity of $G$, and the closedness of all $q$-skeletons are equivalent conditions (see [Pap77]). Let us look at an example of a convex compact set where the minimal face map is not lower semi-continuous.

**Example 4.26.** [Pap77, p.193] Consider the set

$$\mathcal{O} := \text{conv}(\{(x, y, 0) \mid (x - 1)^2 + y^2 \leq 1\} \cup \{(0, 0, 1), (0, 0, -1)\})$$

Extreme points of $\mathcal{O}$ include $\{(x, y, 0) \mid (x - 1)^2 + y^2 = 1\}$ except the point $(0, 0, 0)$. In other words, a sequence of extreme points of the set $\mathcal{O}$ does not necessarily converge to an extreme point, see Figure 4.4. This means that the so-called 0-skeleton (the set of extreme points of $\mathcal{O}$) is not closed and the minimal face map $G$ is not lower semi-continuous.

![Figure 4.4](image-url)
We now give a sufficient condition for the stability.

**Theorem 4.27.** Let $\bar{Q} \in \mathcal{P}$ and let $\bar{X}$ be a corresponding nondegenerate primal first order maximizer. (Then there is a unique dual optimal solution $\bar{Y}$ with $\dim G(\bar{Y}) = n$.) Assume in addition that primal uniqueness and nondegeneracy are stable at $\bar{Q}$ and that the minimal face mapping $G$ is lsc at $\bar{Y}$. Then the first order maximizer $\bar{X}$ is stable at $\bar{Q}$.

**Proof.** Stability of nondegeneracy of the maximizer at $\bar{Q}$ implies stability of the primal Slater condition by Proposition 2.21. Since $\bar{X}$ is a maximizer of order $p > 0$ and the Slater condition holds, standard results in parametric SIP (see e.g. [GL98a, Theorem 10.4] or [BS00, Proposition 4.41]) yield the continuity condition (upper semicontinuity) for the unique solutions $X_\nu$ of any sequence $Q_\nu \to \bar{Q}$:

$$X_\nu \to \bar{X}.$$  

The stability of the Slater condition implies that for $Q \approx \bar{Q}$ the solution set of the dual is nonempty and compact (see e.g. [GL98a, Theorem 9.8]), and the stability of primal nondegeneracy assures that there is a unique dual solution $Y = Y(Q)$ by Lemma 2.25. Since there is no duality gap for the problems $Q_\nu$ and the primal optimal solutions $X_\nu$ converge to $ar{X}$, the corresponding unique dual solutions $Y_\nu$ are bounded. Therefore, $Y_\nu$ must converge to $Y$ as well.

Now assume that the first order maximizer is not stable at $\bar{Q}$. Then there exists a sequence $Q_\nu = (c_\nu, B_\nu, A_{1\nu}, \ldots, A_{n\nu})$ with $Q_\nu \to \bar{Q}$ such that

the maximizers $X_\nu$ wrt. $Q_\nu$ are not of first order.  \hspace{2cm} (4.8)

Since $\bar{X}$ is a first order maximizer, we have $\dim G(\bar{Y}) = n$ and strict complementarity holds for $\bar{X}, \bar{Y}$ by Corollary 4.13. Consider now the unique dual solutions $Y_\nu$ wrt. $Q_\nu$. Since $G$ is lsc at $\bar{Y}$, it follows that $\dim G(Y_\nu) \geq n$. Since $Y_\nu \in \text{ri} G(Y_\nu)$, for any fixed $\nu$ there exist linearly independent matrices $V_j^\nu \in G(Y_\nu)$ and scalars $v_j^\nu > 0$ (for $j = 1, \ldots, k_\nu$ with $k_\nu \geq n$) such that

$$Y_\nu = \sum_{j=1}^{k_\nu} v_j^\nu V_j^\nu.$$  

80
4.3. First order minimizers of (D)

Let us put $\mathcal{L}_{\nu} := \text{span}\{A_{1\nu}^\nu, \ldots, A_{n\nu}^\nu\}$ and $\mathcal{R}_{\nu} := \text{span}\{V_{1\nu}^\nu, \ldots, V_{k_{\nu}}^\nu\}$ with $\dim \mathcal{R}_{\nu} = k_{\nu} \geq n$. We will show that (4.1) is satisfied for $X_{\nu}, Y_{\nu}$, so that by Theorems 4.9 and 4.11 the maximizers $X_{\nu}$ are of first order in contradiction to (4.8). To do so, note that $\mathcal{R}_{\nu} \subseteq G(Y_{\nu}) \subseteq J^\Delta(X_{\nu})$, and the stable nondegeneracy of $X_{\nu}$ gives

$$\mathcal{L}_{\nu}^\perp \cap \text{span} J^\Delta(X_{\nu}) = \{0\} \quad \Rightarrow \quad \mathcal{L}_{\nu}^\perp \cap \mathcal{R}_{\nu} = \{0\}.$$  

This immediately gives that $\dim \mathcal{R}_{\nu} \geq n = n + \dim(\mathcal{L}_{\nu}^\perp \cap \mathcal{R}_{\nu})$.

Finally, as in the proof of Corollary 4.13, using $\dim \mathcal{L}_{\nu}^\perp = \frac{1}{2}k(k+1) - n$ and $\mathcal{L}_{\nu}^\perp \cap \mathcal{R}_{\nu} = \{0\}$, we find that $\dim \mathcal{R}_{\nu} \leq n$, and thus $\dim \mathcal{R}_{\nu} = n$. With $\{0\} = \mathcal{L}_{\nu}^\perp \cap \mathcal{R}_{\nu}$, we obtain $\mathcal{L}_{\nu}^\perp + \mathcal{R}_{\nu} = S^k$, and thus $\mathcal{L}_{\nu} \cap \mathcal{R}_{\nu}^\perp = \{0\}$. Hence, the conditions in (4.1) are satisfied. $\square$

Let us specify further sufficient conditions for the stability of first order maximizers for the SDP case. By Proposition 3.20, there is a generic subset $\mathcal{P}_g \subset \mathcal{P}$ such that for any $\overline{Q} \in \mathcal{P}_g$, properties such as nondegeneracy and uniqueness are satisfied and $\text{rank} \overline{X} = r$ and $\text{rank} \overline{Y} = s$ with $k = r + s$ are stable. In other words, the condition $\frac{1}{2}s(s + 1) = \frac{1}{2}(k - r)(k - r + 1) = n$ in Corollary 4.19 is stable for $\overline{Q} \in \mathcal{P}_g$. Thus, we obtain the following for SDP case:

**Corollary 4.28.** Consider an SDP instance given by a parameter $\overline{Q} \in \mathcal{P}_g$. Suppose the unique maximizer $\overline{X}$ wrt. $\overline{Q}$ is of first order. Then the first order maximizer is a stable at $\overline{Q}$.

4.3 First order minimizers of (D)

Finally, in this short section, we finish with some remarks on first order solutions for the dual problem $(D)$. We can apply the results for the first order maximizers $\overline{X}$ of $(P)$ to the dual. Let us consider the self-dual formulation of the dual problem: denote as usual $m = \frac{1}{2}k(k + 1)$ and choose (under Assumption 2.6) a basis $\{A_1^\perp, \ldots, A_{m-n}^\perp\}$ of $\mathcal{L}^\perp$. Then any feasible solution $Y$ has the form $Y = C + \sum_{j=1}^{m-n} y_j A_j^\perp$. By defining
\( b := (\langle B, A_1^\perp \rangle, \ldots, \langle B, A_{m-n}^\perp \rangle) \), the dual \((D)\) can be equivalently written in form of a “primal”:

\[
(D_0) \quad \min b^T y \quad \text{s.t.} \quad Y = C + \sum_{j=1}^{m-n} y_j A_j^\perp \in \mathcal{K}^*,
\]

with corresponding “dual” problem \((P_0)\). We can now apply all results of the previous sections to \((D_0)\), we only have to make the obvious changes, e.g., we have to replace \(G(Y)\) by \(J(X)\) and \(n\) by \(m - n\). As an example, we formulate Corollary 4.13 in terms of the dual.

**Corollary 4.29.** Let \(Y \in \mathcal{F}_{D_0}\) be a nondegenerate minimizer of \((D_0)\) and let \(X\) be the optimal solution of \((P_0)\) which is unique by Lemma 2.25 (b). Then \(Y\) is a first order minimizer if and only if \(\dim J(X) = m - n\) and \(X, Y\) satisfy the dual strict complementarity condition \(J(X) = G^\Delta(Y)\).

**Remark 4.30.** *Semidefinite programming:* Again we can be more specific in the case of \(\mathcal{K} = \mathcal{K}^* = \mathcal{S}_k^+\). As before, let us denote the rank of the optimal solutions \(X\) of \((P_0)\) and \(Y\) of \((D_0)\) by \(r := \text{rank } X\) and \(s := \text{rank } Y\), and recall that strict complementarity is equivalent to \(k = r + s\). We obtain: if \(Y\) is a nondegenerate minimizer of \((D_0)\), then

\[
\text{\(Y\) is of first order} \quad \iff \quad \frac{1}{2}(k - s)(k - s + 1) = \frac{1}{2}k(k + 1) - n.
\]

Moreover, under primal and dual nondegeneracy and strict complementarity, the optimal solutions \(X, Y\) are both first order if and only if

\[
\frac{1}{2}s(s + 1) = n \quad \text{and} \quad \frac{1}{2}r(r + 1) = \frac{1}{2}k(k + 1) - n.
\]

Note that this is only possible for \(r = k\) and \(s = n = 0\) or for \(s = k\) and \(r = 0\) (see Example 4.18).