2 GEOMETRY OF SEMIDEFINITE PROGRAMMING

Gabor Pataki

My work has always tried to unite the true with the beautiful and when I had to choose one or the other, I usually chose the beautiful.

Hermann Weyl (1885 - 1955) quoted in an obituary by Freeman J. Dyson in Nature, March 10, 1956

[Inscription above Plato's Academy:]
Let no one ignorant of geometry enter here.

Plato

2.1 INTRODUCTION

Consider the primal-dual pair of optimization problems

$$
(P) \quad \begin{array}{rl} Min & \langle c, x\rangle \\ s.t. & x \in K \\ & Ax = b \end{array} \qquad\qquad \begin{array}{rl} Max & \langle b, y\rangle \\ s.t. & z \in K^* \\ & A^*y + z = c \end{array} \quad (D)
$$

where

- X and Y are Euclidean spaces with $\dim X \geq \dim Y$.
- $A : X \to Y$ is a linear operator, assumed to be onto.

- $A^* : Y \to X$ is its adjoint.
- K is a closed, convex, facially exposed cone in X.
- $K^* := \{ z \mid \langle z, x \rangle \geq 0 \ \forall x \in K \}$ is the dual of K, also a closed, convex, facially exposed cone.

The problems (P) and (D) are called a primal-dual pair of conic linear programs, cone programs, or cone-LP's. With the appropriate choice of X, Y and K, they include: ordinary LP's; semidefinite programs (SDP's); programs over p-cones, in particular over second order cones.

Besides the wide applicability of cone-LP's, their main attraction is their elegance: both their duality theory, and the algorithmic approaches to solve them are natural extensions of their counterparts in linear programming. As we shall see in this chapter, the situation is similar regarding their geometry. By the "geometry" of a cone-LP we mean the characterization of the

(1) Set of optimal solutions, in particular, of whether this set is a singleton (the question of uniqueness).

(2) Tangent cone and tangent space of the feasible set at an optimal solution (thus through polarity, also of the normal cone).

(3) Analogous sets in the dual problem.

In this chapter we give the overview of a theory that describes the geometry of cone-LP's. It is reminiscent of how the geometry of LP is usually described: through the *facial structure* of the feasible set. Since the solution set of a cone-LP is always a face of the feasible set, regardless of what the underlying cone is, this approach is quite natural. It generalizes the notions, and corresponding theorems known in LP about the facial structure of the feasible set, on nondegeneracy, and strict complementarity.

Its essence: given a feasible solution to a cone program, the minimal face of the *feasible set* containing it is the intersection of the minimal face of the *cone* containing it with the affine constraints. Whether this solution is an extreme point of the feasible set can be characterized using these two latter sets. (E.g. when the cone program is an LP, the extremity of a feasible solution depends only on the *position* of the nonzeros in it; in other words on the minimal face of the nonnegative orthant that contains it). Moreover, its nondegeneracy is defined by imposing the extremity condition on the *dual* with the *complementary* face of the cone; and strict complementarity of a solution-pair is imposed by requiring them to be in the relative interior of complementary faces of K and K^*. The other objects we want to study (the tangent cone and tangent space of the feasible set at a solution) will have similar descriptions.

This theory can be specialized to various classes of cone-LP's by using the description of the faces of the underlying cones. In all the interesting cases this description is quite handy; for the nonnegative orthant and p-cones, it is trivial; for the semidefinite cone it is given by a classical result (see e.g. [71]).

The chapter is structured as follows: section 2.2 collects the notation, and necessary basic results that will be used later on. Section 2.3 presents the theory on the geometry of cone programs. In subsection 2.3.1 we describe their facial structure, the notions of nondegeneracy and strict complementarity, and prove the generalizations of the results connecting them in LP. We also show how several previous results on the geometry of SDP are subsumed by this framework. Subsection 2.3.2 describes the tangent spaces of the feasible sets. In subsection 2.3.3 we derive the family of *boundary structure inequalities* that relate the dimensions of

- Minimal faces in the primal and dual *cones* that contain a given optimal solution.
- Minimal faces in the primal and dual *feasible sets* that contain a given optimal solution.
- Tangent spaces at the primal and dual feasible sets at a given optimal solution.

These inequalities provide a surprising amount of information about the boundary structure of (P) and (D). For example: at a strictly complementary optimal solution pair in an SDP, it is impossible to have full-dimensional normal cones at both the primal and dual optima; i.e. both the primal and dual optimal solutions cannot be "kinky".

Subsection 2.3.4 translates the previous results for equivalent cone programs formulated with different variables (e.g. when the dual slack is eliminated), and subsection 2.3.5 gives a detailed example. In section 2.4 we present several examples of "semidefinite combinatorics", ie. apply the results on the geometry of cone programs to deduce some instructive structural results about problems that can be formulated as an SDP.

Finally, in section 2.5 we study two algorithmic aspects; converting a feasible solution of a cone-lp into one, which is also an extreme point of the feasible set, and performing sensitivity analysis.

2.2 PRELIMINARIES

Spaces and cones of interest. The space of n by n symmetric, and the cone of n by n symmetric, positive semidefinite matrices are denoted by $\mathcal{S}^n$, and $\mathcal{S}^n_+$, respectively. The space $\mathcal{S}^n$ is equipped with the inner product

$$\langle x, z \rangle \quad := \quad \sum_{i,j=1}^{n} x_{ij} z_{ij},$$

and it is a well-known fact, that $\mathcal{S}^n_+$ is self-dual with respect to it.

If $1 < p < +\infty$, then the p-cone is defined as

$$K_p \quad = \quad \{\, (x_0, x) \mid x_0 \geq ||x||_p \,\}.$$

If $K = K_p$, then $K^* = K_q$, where $\frac{1}{p} + \frac{1}{q} = 1$.

Operators and matrices. Linear operators are denoted by capital letters; when a matrix is considered to be an element of a euclidean space, and not a linear operator, it is usually denoted by a small letter. The i^{th} row of matrix a is denoted by $a_{i\,.}$ and the j^{th} column by $a_{.\,j}$.
The range space of an operator A [of a matrix x] is denoted by $\mathcal{R}(A)$ [$\mathcal{R}(x)$]. If $x \in \mathcal{S}^n$, then $\lambda_i(x)$ denotes its i^{th} largest eigenvalue, and $\lambda(x)$ the vector $(\lambda_1(x), \ldots, \lambda_n(x))^T$.
The identity linear operator, and the identity matrix are denoted by I, and the vector of all ones by e.
The positive part of the vector $x \in \mathcal{R}^n$ is denoted by x_+, i.e. $(x_+)_i = x_i$, if $x_i > 0$, and $(x_+)_i = 0$, otherwise.
The inner product of $x^1, x^2 \in X$ is denoted by $\langle x^1, x^2 \rangle$. Even if the inner products in X and Y are different (say if $X = \mathcal{S}^n$ and $Y = \mathcal{R}^m$), we still use the notation $\langle , \rangle$ for both; the context should make it clear, which one is meant. The matrix product of matrices x^1 and x^2 is denoted by $x^1 x^2$. The block diagonal matrix obtained by placing x^1 and x^2 on the main diagonal is denoted by $x^1 \oplus x^2$.

The dimension of the intersection of subspaces. The following simple proposition will be used many times, hence we state, and prove it for convenience.

Proposition 2.2.1 *Suppose that L_1 and L_2 are subspaces. Then*

$$\dim\,[\,L_1 + L_2\,] \;=\; \dim\,L_1 + \dim\,L_2 - \dim\,[\,L_1 \cap L_2\,].$$

Proof Case 1 If $L_1 \cap L_2 = \{0\}$, then the claim is obvious.
Case 2 $L_1 \cap L_2 \neq \{0\}$. Let L_3 be a subspace that satisfies

$$L_3 \subseteq L_1, \quad L_3 \cap L_2 = \{0\}, \quad L_1 = L_3 + [\,L_1 \cap L_2\,].$$

Then using Case 1 with L_3 in place of L_1, and $L_1 \cap L_2$ in place of L_2, we obtain

$$\dim\,L_1 \;=\; \dim\,L_3 + \dim\,[\,L_1 \cap L_2\,]. \tag{2.2.1}$$

Also,

$$\begin{aligned} L_1 + L_2 &= [\,L_3 + [L_1 \cap L_2]\,] + L_2 \\ &= L_3 + L_2. \end{aligned}$$

Again using Case 1, we get

$$\dim\,[\,L_1 + L_2\,] \;=\; \dim\,L_3 + \dim\,L_2. \tag{2.2.2}$$

Subtracting (2.2.2) from (2.2.1) and rearranging yields the required formula.

■

Faces, feasible directions, tangent cones and tangent spaces in convex sets. For vectors y and z, we denote the open line-segment between y and z by

$$(y, z) \quad = \quad \{\, \mu y + (1-\mu) z \mid 0 < \mu < 1 \,\}.$$

Let C be a closed convex set. A convex subset F of C is called a *face* of C, and this fact is denoted by $F \lhd C$, if

$$x \in F,\ y, z \in C,\ x \in (y, z) \text{ implies } y, z \in F. \tag{2.2.3}$$

An *extreme point* of C is a face consisting of a single element. If S is a subset of C, then we denote by face (S, C) the *minimal face of* C containing S, and if $x \in C$, then we write face (x, C) for face $(\{x\}, C)$.
A *supporting hyperplane* of C is a set of the form

$$H \quad = \quad \{\, x \mid \langle a, x \rangle = \alpha \,\},$$

where

$$\alpha \quad = \quad \max \{\, \langle a, x \rangle \mid x \in C \,\}.$$

The set $C \cap H$ is called an *exposed face* of C. An exposed face is always a face, whereas the opposite may not be true. For a discussion of the difference between faces and exposed faces, see e.g. Section 5 in [139].

Proposition 2.2.2 *Let C be a convex set, $C' \lhd C$, and D a convex subset of C.*

(i) If $\mathrm{ri}\, \mathrm{D} \cap \mathrm{C}' \neq \emptyset$*, then* $D \subseteq C'$.

(ii) If $D \subseteq C'$*, and* $D \cap \mathrm{ri}\, \mathrm{C}' \neq \emptyset$*, then* $C' = \mathrm{face}\,(D, C)$.

(iii) $C' = \mathrm{face}\,(D, C)$*, iff* $\mathrm{ri}\, \mathrm{D} \cap \mathrm{ri}\, \mathrm{C}' \neq \emptyset$.

(iv) $C' = C \cap \mathrm{aff}\, C'$.

Proof The statement (i) is [662, Theorem 18.1]; (ii) follows from [662, Theorem 18.2], and (iii) by putting (i) and (ii) together. Statement (iv) is Exercise 5.4 in [139]. ■

For $x \in C$, the cone of feasible directions, the tangent cone and the tangent space at x in C are defined as

$$\begin{aligned}
\mathrm{dir}\,(x, C) &= \{\, y \mid x + ty \in C \text{ for some } t > 0 \,\}, \\
\mathrm{tcone}\,(x, C) &= \mathrm{Cl}\ \mathrm{dir}\,(x, C) \;=\; \{\, y \mid \mathrm{dist}(x + ty, C) = o(t) \,\}, \\
\tan(x, C) &= \mathrm{tcone}\,(x, C) \cap -\mathrm{tcone}\,(x, C) \\
&= \{\, y \mid \mathrm{dist}(x \pm ty, C) = o(t) \,\}.
\end{aligned}$$

The equivalence of the alternative expressions for $\text{tcone}\,(x, C)$ follows e.g. from [339], page 135.
An important fact ([339, Proposition 5.3.1]), that we state for the ease of reference is

Proposition 2.2.3 *If C_1 and C_2 are nonempty closed convex sets, $x \in C_1 \cap C_2$, then*

$$\text{tcone}\,(x, C_1 \cap C_2) \quad \subseteq \quad \text{tcone}\,(x, C_1) \cap \text{tcone}\,(x, C_2),$$

with equality holding, if $\text{ri}\, \mathrm{C}_1 \cap \text{ri}\, \mathrm{C}_2 \neq \emptyset$.

■

Faces, complementary faces, feasible directions and tangent spaces in cones. A convex set K is a *cone*, if $\mu K \subseteq K$ holds for all $\mu \geq 0$. If K is a cone, then a simple argument shows that (2.2.3) is equivalent to

$$x \in F, \; y, z \in K, \; x = y + z \;\text{ implies }\; y, z \in F. \tag{2.2.4}$$

The *dual* of the cone K is

$$K^* \quad = \quad \{\, z \,|\, \langle z, x \rangle \geq 0 \text{ for all } x \in K \,\}.$$

If K, K_1 and K_2 are convex cones, then

$$\begin{aligned} K^{**} &= \text{Cl}\, K, \\ (K_1 + K_2)^* &= K_1^* \cap K_2^*, \\ (K_1 \cap K_2)^* &= \text{Cl}\,(K_1^* + K_2^*). \end{aligned}$$

If $F \lhd K$, and $\bar{x} \in \text{ri}\, \mathrm{F}$ is fixed, then the *complementary* (or *conjugate*) face of F is defined alternatively as

$$\begin{aligned} F^{\triangle} &= \{\, z \in K^* \,|\, \langle z, x \rangle = 0 \;\text{ for all } x \in F \,\} \\ &= \{\, z \in K^* \,|\, \langle z, \bar{x} \rangle = 0 \,\}. \end{aligned}$$

The equivalence of the two definitions is straightforward. The complementary face of $G \lhd K^*$ is defined analogously, and is denoted by $G^{\triangle}$. K is facially exposed, i.e. all faces of K arise as the intersection of K with a supporting hyperplane, iff for all $F \lhd K$, $F^{\triangle\triangle} = F$, see [139, Theorem 6.7]. For brevity, if $F \lhd K$, then we write $F^{\triangle *}$ for $(F^{\triangle})^*$, and $F^{\triangle \perp}$ for $(F^{\triangle})^{\perp}$.
It is well known that if K is a polyhedral cone, then for all $F \lhd K$, then $\text{lin}\, F + \text{lin}\, F^{\triangle}$ is the whole space. The *residual subspace* of $F \lhd K$ is meant to measure, "to what extent F is nonpolyhedral". It is defined as

$$\text{res}\; F \quad = \quad (\text{lin}\, F + \text{lin}\, F^{\triangle})^{\perp}.$$

We say, that K is *nice*, if one of the following equivalent statements holds

$$\begin{array}{rll} K^* + F^\perp & \text{is closed} & \forall\ F \lhd K, \\ \operatorname{Proj}_{\operatorname{lin} F}(K^*) & \text{is closed} & \forall\ F \lhd K. \end{array} \tag{2.2.5}$$

Next, we list several examples of cones, along with the description of their faces. The corresponding complementary faces, and residual subspaces can be found in Table 2.2.

Example 2.2.1 (The nonnegative orthant) If $\bar{x} \in K = \mathcal{R}^n_+$, then

$$\operatorname{face}(\bar{x}, \mathcal{R}^n_+) \quad = \quad \{\, x \in \mathcal{R}^n_+ \mid x_i = 0 \ \ \forall\, i \ s.t.\ \bar{x}_i = 0 \,\}.$$

This face, (after permuting components) can be brought to the form

$$\operatorname{face}((e, 0)^T, \mathcal{R}^n_+),$$

for an e of appropriate size. ■

Example 2.2.2 (The semidefinite cone) If $\bar{x} \in K = \mathcal{S}^n_+$, then

$$\operatorname{face}(\bar{x}, \mathcal{S}^n_+) \quad = \quad \{\, x \in \mathcal{S}^n_+ \mid \mathcal{R}(x) \subseteq \mathcal{R}(\bar{x}) \,\}, \tag{2.2.6}$$

$$\operatorname{face}(\bar{x}, \mathcal{S}^n_+)^\triangle \quad = \quad \{\, x \in \mathcal{S}^n_+ \mid \mathcal{R}(x) \subseteq \mathcal{R}(\bar{x})^\perp \,\}, \tag{2.2.7}$$

([71], for a simple proof, see Appendix 2.7.1.) Let q be an orthonormal matrix such that

$$\bar{x} \quad = \quad q \left(\begin{smallmatrix} \Lambda & 0 \\ 0 & 0 \end{smallmatrix}\right) q^T,$$

where Λ is a diagonal matrix with positive diagonal. All transformations $v^T(.)v$, where v is an invertible matrix, are one-to-one mappings of $\mathcal{S}^n_+$ to itself. Therefore $\operatorname{face}(\bar{x}, \mathcal{S}^n_+)$ can be brought to the form

$$q^T(\operatorname{face}(\bar{x}, \mathcal{S}^n_+))q \ = \ \operatorname{face}(q^T \bar{x} q, \mathcal{S}^n_+) \ = \ \operatorname{face}(\left(\begin{smallmatrix} \Lambda & 0 \\ 0 & 0 \end{smallmatrix}\right), \mathcal{S}^n_+) \ = \ \operatorname{face}(\left(\begin{smallmatrix} I & 0 \\ 0 & 0 \end{smallmatrix}\right), \mathcal{S}^n_+).$$

If the rank of $\bar{x}$ is r, then

$$\dim(\operatorname{face}(\bar{x}, \mathcal{S}^n_+)) \quad = \quad t(r) \quad := \quad r(r+1)/2,$$

where $t(r)$ denotes the r^{th} triangular number. ■

Example 2.2.3 (The p-cones) Let $1 < p < +\infty$. Since K_p is obtained by "lifting" the unit ball of the norm $||\,.\,||_p$, all of its nontrivial faces (i.e. apart from the origin and itself) are of the form

$$\operatorname{cone}\{\, (||x||_p, x)^T \,\}$$

K	A typical F	$F^{\triangle}$	res F
$\mathcal{R}^n_+$	$\mathrm{face}\,((e,0)^T, \mathcal{R}^n_+)$	$\mathrm{face}\,((0,e)^T, \mathcal{R}^n_+)$	$\{0\}$
$\mathcal{S}^n_+$	$\mathrm{face}\,(\left(\begin{smallmatrix} I & 0 \\ 0 & 0 \end{smallmatrix}\right), \mathcal{S}^n_+)$	$\mathrm{face}\,(\left(\begin{smallmatrix} 0 & 0 \\ 0 & I \end{smallmatrix}\right), \mathcal{S}^n_+)$	$\{\, y \in \mathcal{S}^n \mid y = \left(\begin{smallmatrix} 0 & v \\ v^T & 0 \end{smallmatrix}\right) \,\}$
K_p	$\mathrm{cone}\,\{\, (\lvert\lvert x \rvert\rvert_p, x)^T \,\}$	$\mathrm{cone}\,\{\, (\lvert\lvert x \rvert\rvert_q, -x)^T \,\}$	$\{\, (0, y)^T \mid \langle y, x \rangle = 0 \,\}$

Table 2.1 The faces, complementary faces, and residual subspaces in $\mathcal{R}^n_+$, $\mathcal{S}^n_+$ and K_p

for some x. ■

It is not hard to see, that all these cones are facially exposed. They are also nice, by using the second criterion in (2.2.5). In the case of $\mathcal{R}^n_+$ and $\mathcal{S}^n_+$, the projection in question is just a smaller copy of the original cone. In the case of K_p the linear span of any nontrivial face is a line, and all cones contained in a line are closed.
Next, we show that the set of feasible directions, and several related sets for an $x \in K$ can be conveniently described in terms of $\mathrm{face}\,(x, K)$.

Lemma 2.2.1 *Let $x \in K$, and write $F = \mathrm{face}\,(x, K)$. Then the following relations hold.*

$$\begin{aligned} \mathrm{dir}\ (x,K) &= K + \mathrm{lin}\, F, && (2.2.8) \\ \mathrm{dir}\ (x,K)^* &= K^* \cap F^{\perp} = F^{\triangle} = K^* \cap \mathrm{lin}\, F^{\triangle}, && (2.2.9) \\ \mathrm{Cl}\ \mathrm{dir}\ (x,K) &= \mathrm{Cl}\ (K + \mathrm{lin}\, F) = F^{\triangle *} = \mathrm{Cl}\ (K + \mathrm{lin}\, F + \mathrm{res}\ F). && (2.2.10) \end{aligned}$$

Furthermore, if K is nice, then

$$\tan(x,K) = \mathrm{lin}\, F + \mathrm{res}\ F. \quad (2.2.11)$$

Proof of (2.2.8) "$\supseteq$:" Let $v \in \mathrm{lin}\, F$, $z \in K$. Then $x + \alpha v \in K$ for some $\alpha > 0$ $x + \alpha v \in K$, hence $x + \alpha(v + z) \in K$.

"$\subseteq$:" Let $y \in \mathrm{dir}\ (x,K)$, $\alpha > 0$, $x' := x + \alpha y \in K$. Then $y = \frac{1}{\alpha}(x' - x) \in K + \mathrm{lin}\, F$.
Proof of (2.2.9) The first equality follows from (2.2.8) by taking the dual, the second by the definition of $F^{\triangle}$, and the third by Proposition 2.2.2 (iii), since $F^{\triangle}$ is a face.
Proof of (2.2.10) The first and third equalities follow from (2.2.9).
Proof of (2.2.11) From the definition of $\tan(x,K)$ and since K is nice,

$$\tan(x,K) = (K + \mathrm{lin}\, F + \mathrm{res}\ F) \cap -(K + \mathrm{lin}\, F + \mathrm{res}\ F).$$

Therefore "$\supseteq$" in (2.2.11) is obvious. For "$\subseteq$", let

$$\begin{array}{l} x^1 \in K,\ y^1 \in \operatorname{lin} F + \operatorname{res}\ F, \text{such that} \\ x^1 + y^1 \in (K + \operatorname{lin} F + \operatorname{res}\ F) \cap -(K + \operatorname{lin} F + \operatorname{res}\ F). \end{array}$$

That is, for some $x^2 \in K, y^2 \in \operatorname{lin} F + \operatorname{res}\ F$,

$$\begin{array}{lclclcll} x^1 + y^1 & = & -(x^2 + y^2) & \Rightarrow & x^1 + x^2 & = & y^1 + y^2 & \Rightarrow \\ x^1 + x^2 & \in & K \cap (\operatorname{lin} F + \operatorname{res}\ F) = F & \Rightarrow & x^1, x^2 & \in & F & \Rightarrow \\ x^1 + y^1 & \in & \operatorname{lin} F + \operatorname{res}\ F. & & & & & \end{array}$$

■

2.3 THE GEOMETRY OF CONE-LP'S: MAIN RESULTS

2.3.1 *Facial structure, nondegeneracy and strict complementarity*

We say that (P) satisfies the Slater condition, if there is an $\bar{x} \in \operatorname{ri} \mathrm{K}$ feasible for (P), and that (D) satisfies the Slater condition, if there is a $(\bar{y}, \bar{z}) \in Y \times \operatorname{ri} \mathrm{K}^*$ feasible for (D). We assume that the optimal values of (P) and (D) are equal, both attained, (this is ensured if both satisfy the Slater condition) and denote:

- their feasible sets by *Feas(P)* and *Feas(D)*, resp.
- the set of their optimal solutions by *Opt(P)* and *Opt(D)*, resp.

Theorem 2.3.1 (Primal Faces) *Let $\bar{x} \in Feas(P)$, and let*

$$\begin{array}{rcl} \mathcal{F} & = & \operatorname{face}(\bar{x}, Feas(P)), \\ F & = & \operatorname{face}(\bar{x}, K). \end{array}$$

Then

(1) $F = \operatorname{face}(\mathcal{F}, K)$.

(2) *(a)* $\begin{array}{rcl} \operatorname{aff} \mathcal{F} & = & \operatorname{lin} F \cap \{ x \mid Ax = b \} \\ & = & \bar{x} + [\operatorname{lin} F \cap \mathcal{N}(A)], \end{array}$

(b) $\mathcal{F} = F \cap \{ x \mid Ax = b \}$.

(3) $\dim \mathcal{F} = \dim [\operatorname{lin} F \cap \mathcal{N}(A)] = \dim F - \dim Y + \dim [F^{\perp} \cap \mathcal{R}(A^*)]$.

(4) *$\mathcal{F}$ is a singleton set, i.e. $\mathcal{F}$ (or equivalently $\bar{x}$) is an extreme point of Feas(P), if and only if*

$$\operatorname{lin} F \cap \mathcal{N}(A) = \{0\}.$$

Proof
(1): Follows from (iii) in Proposition 2.2.2.

(2)(a) "$\subseteq$" : Follows from (1).
(2)(a) "$\supseteq$" : Let $v \in \operatorname{lin} F \cap \{ x \mid Ax = b \}$. As $\bar{x} \in \mathcal{F} \cap \operatorname{ri} \mathrm{F}$, there exists $\epsilon > 0$ such that

$$\begin{array}{rcl} x^1 & = & \bar{x} + \epsilon(\bar{x} - v), \\ x^2 & = & \bar{x} - \epsilon(\bar{x} - v), \\ x^1, x^2 & \in & F. \end{array}$$

Clearly, x^1 and x^2 also satisfy the affine constraint, so they are in *Feas(P)*. As $\bar{x} \in \mathcal{F} \lhd Feas(P)$, and $\bar{x} \in (x^1, x^2)$, we get

$$\begin{array}{rcl} x^1, x^2 & \in & \mathcal{F} \implies \\ v & \in & \operatorname{aff} \{ x^1, x^2 \} \subseteq \operatorname{aff} \mathcal{F}, \end{array}$$

as required.
(2)(b): We have

$$\begin{array}{rcl} \mathcal{F} & = & Feas(P) \cap \operatorname{aff} \mathcal{F} \\ & = & (K \cap \{ x \mid Ax = b \}) \cap (\operatorname{lin} F \cap \{ x \mid Ax = b \}) \\ & = & (K \cap \operatorname{lin} F) \cap \{ x \mid Ax = b \} \\ & = & F \cap \{ x \mid Ax = b \}, \end{array}$$

where the first equality follows by $\mathcal{F} \lhd$ *Feas(P)*, and the last by $F \lhd K$.
(3): By (2)(a),

$$\begin{array}{rcll} \operatorname{aff} \mathcal{F} - \bar{x} & = & \mathcal{N}(A) \cap \operatorname{lin} F & \Rightarrow \\ \dim \mathcal{F} & = & \dim [\operatorname{lin} F \cap \mathcal{N}(A)] & \\ & = & \dim F + \dim \mathcal{N}(A) - \dim [\mathcal{N}(A) + \operatorname{lin} F] & \\ & = & \dim F + \dim X - \dim Y - \dim [\mathcal{N}(A) + \operatorname{lin} F] & \\ & = & \dim F - \dim Y + \dim [\mathcal{R}(A^*) \cap F^{\perp}], & \end{array}$$

with the third equality following from Proposition 2.2.1.
(4): $\mathcal{F}$ is a singleton set, iff $\operatorname{aff} \mathcal{F}$ is, so the equivalence follows from (2)(a).
■

Theorem 2.3.2 (Dual Faces) *Let $(\bar{y}, \bar{z}) \in Feas(D)$, and let*

$$\begin{array}{rcl} \mathcal{G} & = & \operatorname{face}((\bar{y}, \bar{z}), Feas(D)), \\ Y \times G & = & \operatorname{face}((\bar{y}, \bar{z}), (Y \times K^*)) \quad (\Leftrightarrow\ G = \operatorname{face}(\bar{z}, K^*)). \end{array}$$

Then

(1) $Y \times G = \operatorname{face}(\mathcal{G}, (Y \times K^*))$

(2) *(a)* $$\begin{array}{rcl} \operatorname{aff} \mathcal{G} & = & (Y \times \operatorname{lin} G) \cap \{ (y, z) \mid A^* y + z = c \} \\ & = & (\bar{y}, \mathbf{z}) + [(Y \times \operatorname{lin} G) \cap \mathcal{N}(A^*, I)] \end{array}$$

(b) $\mathcal{G} = (Y \times G) \cap \{ (y, z) \mid A^* y + z = c \}$

(3) $$\begin{array}{rcl} \dim \mathcal{G} & = & \dim [\operatorname{lin} G \cap I\!R(A^*)] \\ & = & \dim G - (\dim X - \dim Y) + \dim [G^{\perp} \cap \mathcal{N}(A)] \end{array}$$

(4) $\mathcal{G}$ is a singleton set, i.e. $\mathcal{G}$ (or equivalently $(\bar{y}, \bar{z})$) is an extreme point of Feas(D), if and only if

$$\operatorname{lin} G \cap \mathcal{R}(A^*) \quad = \quad \{0\}.$$

Proof (1) and (2) follow along the same lines as their counterparts in the Primal Faces Theorem, by noting that the relative interior [affine hull], of $Y \times G$ is the the direct product of Y with the relative interior [affine hull] of G.
(3): From (2)(a)

$$\begin{aligned}
\operatorname{aff} \mathcal{G} - (\bar{y}, \bar{z}) &= (Y \times \operatorname{lin} G) \cap \mathcal{N}(A^*, I), \\
\dim \mathcal{G} &= \dim [\operatorname{lin} G \cap \mathcal{R}(A^*)] \\
&= \dim \mathcal{R}(A^*) + \dim G - \dim [\operatorname{lin} G + \mathcal{R}(A^*)] \\
&= \dim G - (\dim X - \dim Y) + \dim [G^{\perp} \cap \mathcal{N}(A)].
\end{aligned}$$

(4): $\mathcal{G}$ is a singleton set, iff $\operatorname{aff} \mathcal{G}$ is, so the equivalence follows from (2)(a). ■

Remark 2.3.1 If A is not onto, then $\dim Y$ in all the previous results should be replaced by $\operatorname{rank} A$.

Given a feasible solution of a cone-LP, and the minimal face of the *cone* containing it, its nondegeneracy is defined by imposing the extremity condition in the dual problem, with the complementary face in the dual cone; and strict complementarity of a solution pair by requiring them to be in the relative interior of complementary faces.

Definition 2.3.1 *Let $\bar{x}$ be feasible for (P), $(\bar{y}, \bar{z})$ feasible for (D), $F =$ face $(\bar{x}, K)$, and $G =$ face $(\bar{z}, K^*)$. We say that*

- $\bar{x}$ *is* (primal) nondegenerate *if*

$$\operatorname{lin} F^{\triangle} \cap \mathcal{R}(A^*) \quad = \quad \{0\},$$

- $(\bar{y}, \bar{z})$ *is* (dual) nondegenerate *if*

$$\operatorname{lin} G^{\triangle} \cap \mathcal{N}(A) \quad = \quad \{0\}.$$

Furthermore, if $\bar{x}$ and $(\bar{y}, \bar{z})$ are optimal solutions, then we say that

- *They are* strictly complementary *if $F^{\triangle} = G$.*

It turns out, that all results well known from LP that connect nondegeneracy in either (P) or (D), strict complementarity, and uniqueness of the optimal solution in the "opposite" problem carry over word by word to our more general framework.

Theorem 2.3.3 *Let $\bar{x}$, $(\bar{y}, \bar{z})$, F and G be as in Definition 2.3.1. Then the following hold.*

(1) *(a)* *If $(\bar{y}, \bar{z})$ is nondegenerate, then $\bar{x}$ is a unique primal optimal solution.*

(b) *The converse of (1)(a) holds, assuming that they are strictly complementary.*

(c) *If $(\bar{y}, \bar{z})$ is nondegenerate, then*

$$\dim G^{\triangle\perp} \;\geq\; \dim X - \dim Y.$$

(2) *(a)* *If $\bar{x}$ is nondegenerate, then $(\bar{y}, \bar{z})$ is a unique dual optimal solution.*

(b) *The converse holds, assuming that they are strictly complementary.*

(c) *If $\bar{x}$ is nondegenerate, then*

$$\dim F^{\triangle\perp} \;\geq\; \dim Y.$$

Proof of (1) The duality gap between arbitrary primal and dual feasible solutions x and (y, z) is $\langle x, z\rangle$. Hence

$$\begin{aligned} Opt(P) &= Feas(P) \cap \{\, x \in K^* \mid \langle x, \bar{z}\rangle = 0 \,\} \\ &= Feas(P) \cap G^{\triangle} \implies \\ E := \text{face}\,(Opt(P), K) &\subseteq G^{\triangle}. \end{aligned}$$

By the Primal Faces Theorem (3), *Opt(P)* is a singleton, iff

$$\operatorname{lin} E \cap \mathcal{N}(A) \;=\; \{\, 0 \,\}.$$

From this (1)(a) is immediate. If strict complementarity holds, i.e. $\bar{x} \in Opt(P) \cap \operatorname{ri} \mathrm{G}^{\triangle}$, then by Proposition 2.2.2, (ii)

$$E \;=\; G^{\triangle}.$$

So, (1)(b) also follows. The proof of (1)(c) is immediate from the definition.
Proof of (2) Analogous. ■

Remark 2.3.2 With the sole exception of the (b) parts of Theorem 2.3.3, all results derived so far hold true, even in the case when K is not facially exposed. The only difficulty arising in this case is that (using the notation there), strict complementarity could be stated in 2 different ways:

$$\text{(i) } F^{\triangle} = G, \quad \text{or} \quad \text{(ii) } F = G^{\triangle}.$$

Neither of (i) or (ii) implies the other, unless K is facially exposed, in which case $F = F^{\triangle\triangle}$, and $G = G^{\triangle\triangle}$, hence (1) and (2) are equivalent.
However, the proof of the (b) parts in Theorem 2.3.3 implies, that if (1) (which may be called "primal strict complementarity") holds, and $(\bar{y}, \bar{z})$ is unique, then $\bar{x}$ is (primal) nondegenerate. Similarly, if (2) (which may be called "dual strict complementarity") holds, and $\bar{x}$ is unique, then $(\bar{y}, \bar{z})$ is (dual) nondegenerate.

Since all cones known so far that occur in practice are facially exposed, and the results for the non-exposed case are simple generalizations of the exposed case, we restrict ourselves to the latter. ■

Remark 2.3.3 Note that nondegeneracy simply requires the *extremity condition* given in parts (4) of the Primal and Dual Faces Theorems to hold for a *strictly complementary* solution in the "opposite" program (even though a strictly complementary solution pair may not always exist; for SDP, see the discussion at the end of this subsection). ■

Remark 2.3.4 Our definition of nondegeneracy is a generalization of the one used by Alizadeh, Haeberly and Overton for SDP in [24], and of the one by Alizadeh and Schmieta in [26]. In Definition 2.3.1 primal nondegeneracy is defined by

$$\begin{aligned} \mathcal{R}(A^*) \cap \operatorname{lin} F^{\triangle} &= \{0\} \;\Leftrightarrow \\ \mathcal{N}(A) + F^{\triangle\perp} &= X, \end{aligned} \tag{2.3.12}$$

and in [26] by

$$\mathcal{N}(A) + \tan(\bar{x}, K) \;=\; X, \tag{2.3.13}$$

"assuming that $\tan(\bar{x}, K)$ exists". If we define $\tan(\bar{x}, K)$ as in Section 2.2, and assume that K (such as $\mathcal{R}^n_+$, $\mathcal{S}^n_+$ and the p-cone) is nice, then by Lemma 2.2.1 $F^{\triangle\perp} = \tan(\bar{x}, K)$. Therefore (2.3.13) and (2.3.12) are equivalent. The case of dual nondegeneracy is similar, assuming that K^* is nice.
On the other hand, our theory covers not only nondegeneracy and strict complementarity, but also the characterization of the faces of the feasible sets (and tangent spaces and other sets, see the discussions below), which is of independent interest. ■

In the remainder of Subsection 2.3.1 we specialize these results for linear and semidefinite programs.

Example 2.3.1 (Linear programming) If $X = \mathcal{R}^n, Y = \mathcal{R}^m, K = K^* = \mathcal{R}^n_+$, then (P) and (D) are a standard pair of primal and dual linear programs. For $\bar{x} \in K$,

$$\operatorname{lin}\operatorname{face}(\bar{x}, K) \;=\; \{\, y \,|\, y_i = 0 \;\; \forall i \; s.t. \; \bar{x}_i = 0 \,\}.$$

In particular, (3) in Theorem 2.3.1 specializes to: $\bar{x}$ is an extreme point iff the columns of A corresponding to its nonzero components are linearly independent. Similarly, $\bar{x}$ is nondegenerate, iff the *rows* of this submatrix are linearly independent.

If $\bar{x}$ and $(\bar{y}, \bar{z})$ are a strictly complementary pair of solutions, then

$$\begin{aligned}\dim \mathcal{F} &= \dim F - \dim Y + \dim [\, F^{\perp} \cap \mathcal{R}(A^*)\,] \\ &= \dim F - m + \dim [\, \mathrm{lin}\, G \cap \mathcal{R}(A^*)\,] \\ &= \dim F - m + \dim \mathcal{G}, \end{aligned} \tag{2.3.14}$$

where the first equality follows from (3) in the Primal Faces Theorem, the second by $F^{\perp} = \mathrm{lin}\, G$, and the third by (3) in the Dual Faces Theorem. The formula (2.3.14) reduces to a result of Tijssen and Sierksma [763], as follows. They define

$$\sigma(\mathcal{F}) = \dim \mathcal{F} + \mathrm{bnd}\, \mathcal{F} - n,$$

as the *degree of degeneracy* of face $\mathcal{F}$, with $\mathrm{bnd}\, \mathcal{F}$ being the number of hyperplanes in the description of the polyhedron *Feas(P)*, which are tight at $\bar{x}$ (equivalently at every point in $\mathcal{F}$). (The number $\sigma(\mathcal{F})$ can depend on the representation of *Feas(P)*.) In words, the degeneracy degree is the number of "superfluous" hyperplanes at $\mathcal{F}$, i.e. of hyperplanes which are tight at $\mathcal{F}$, but not necessary to define its affine hull. They prove

$$\dim \mathcal{G} = \sigma(\mathcal{F}). \tag{2.3.15}$$

Rewriting (2.3.14) yields

$$\dim \mathcal{G} = \dim \mathcal{F} + m - \dim F,$$

and it is not hard to see, that

$$m - \dim F = \mathrm{bnd}\, \mathcal{F} - n.$$

Therefore (2.3.14) and (2.3.15) are indeed equivalent. ■

Example 2.3.2 (Semidefinite programming) If $X = \mathcal{S}^n, Y = \mathcal{R}^m, K = K^* = \mathcal{S}^n_+$, A and A^* are defined via $a^1, \ldots, a^m \in \mathcal{S}^n$ as

$$Ax = \begin{pmatrix} \langle a^1, x\rangle \\ \vdots \\ \langle a^m, x\rangle \end{pmatrix} \quad , \quad \mathcal{A}^* y = \sum_{i=1}^{m} y_i a^i,$$

then (P) and (D) are a pair of semidefinite programs. By the characterization of the faces of $\mathcal{S}^n_+$ given in Example 2.2.2,

$$\mathrm{lin}\,(\mathrm{face}\,(\bar{x}, \mathcal{S}^n_+)) = \{\, x \in \mathcal{S}^n \mid \mathcal{R}(x) \subseteq \mathcal{R}(\bar{x})\,\}.$$

Therefore specializing the results in the Primal and Dual Faces Theorems is straightforward. We obtain

Corollary 2.3.1 *Suppose that* $\bar{x} \in Feas(P), (\bar{y}, \bar{z}) \in Feas(D)$, *where (P) and (D) are a pair of SDP's defined by the operator* A *above, and* $b \in \mathcal{R}^m$, $c \in \mathcal{S}^n$. *Let*

$$\begin{array}{rclrcl} \mathcal{F} & = & \text{face}\,(\bar{x}, Feas(P)), & r & = & \text{rank}\,\bar{x}, \\ \mathcal{G} & = & \text{face}\,((\bar{y}, \bar{z}), Feas(D)), & s & = & \text{rank}\,\bar{z}. \end{array}$$

Then the following hold.

(1) $t(r) \leq m + \dim \mathcal{F}$.

(2) If $\bar{x}$ *is nondegenerate, then* $t(n-r) \leq t(n) - m$.

(3) $t(s) \leq (t(n) - m) + \dim \mathcal{G}$.

(4) If $(\bar{y}, \bar{z})$ *is nondegenerate, then* $t(n-s) \leq m$.

(5) $\bar{x}$ *and* $(\bar{y}, \bar{z})$ *are strictly complementary, iff* $r + s = n$.

Proof Note that $\dim X = t(n)$, $\dim Y = m$, and by Example 2.2.2 if $x \in \mathcal{S}^n_+$ then

$$\text{rank}\, x \leq r \quad \Leftrightarrow \quad \dim\ \text{face}\,(x, \mathcal{S}^n_+) \leq t(r).$$

Then, (1) follows by (3) in the Primal Faces Theorem, (2) by (1)(c) in Theorem 2.3.3, (3) by (3) in the Dual Faces Theorem, and (4) by (2)(c) in Theorem 2.3.3. ■

Moreover, just as in the LP case, one can obtain and elegant characterization of extreme and nondegenerate solutions. Let $\bar{x}$ be a feasible solution of (P), and q an orthonormal matrix such that

$$\bar{x} \quad = \quad q\left(\begin{smallmatrix} \Lambda & 0 \\ 0 & 0 \end{smallmatrix}\right)q^{\mathrm{T}},$$

where Λ is a diagonal matrix with positive diagonal. Partition q as $q = [q^1, q^2]$, where $\mathcal{R}(q^1) = \mathcal{R}(\bar{x})$ and $\mathcal{R}(q^2)$ is its orthogonal complement. Since

$$A\bar{x} \quad = \quad \begin{pmatrix} \langle a^1, \bar{x} \rangle \\ \vdots \\ \langle a^m, \bar{x} \rangle \end{pmatrix} \quad = \quad \begin{pmatrix} \langle q^T a^1 q, \left(\begin{smallmatrix} \Lambda & 0 \\ 0 & 0 \end{smallmatrix}\right) \rangle \\ \vdots \\ \langle q^T a^m q, \left(\begin{smallmatrix} \Lambda & 0 \\ 0 & 0 \end{smallmatrix}\right) \rangle \end{pmatrix},$$

we obtain that $\bar{x}$ is an extreme [nondegenerate] point in *Feas(P)* iff $\left(\begin{smallmatrix} \Lambda & 0 \\ 0 & 0 \end{smallmatrix}\right)$ is such a point in the primal SDP defined by rhs b, and linear operator A_q, where

$$A_q x \quad = \quad \begin{pmatrix} \langle q^T a^1 q, x \rangle \\ \vdots \\ \langle q^T a^m q, x \rangle \end{pmatrix}.$$

Therefore,

Corollary 2.3.2 *Using the notation of Corollary 2.3.1, the following hold.*

(1) $\bar{x}$ *is an extreme point of Feas(P)* $\Leftrightarrow$ *the matrices*

$$(q^1)^T a^1 q^1, \ldots, (q^1)^T a^m q^1$$

span $\mathcal{S}^r$.

(2) $\bar{x}$ *is a nondegenerate point of Feas(P)* $\Leftrightarrow$ *the matrices*

$$\begin{pmatrix} (q^1)^T a^1 q^1 & (q^1)^T a^1 q^2 \\ (q^2)^T a^1 q^1 & 0 \end{pmatrix}, \ldots, \begin{pmatrix} (q^1)^T a^m q^1 & (q^1)^T a^m q^2 \\ (q^2)^T a^m q^1 & 0 \end{pmatrix}$$

are linearly independent.

■

To characterize the extremity, and nondegeneracy of a *dual* feasible solution $(\bar{y}, \bar{z})$, one does not need to repeat the calculations. Recall from Remark 2.3.3 that nondegeneracy requires the extremity condition to hold for a strictly complementary solution in the "opposite" problem. Suppose that $(\bar{y}, \bar{z})$ is a feasible solution of (D), and write

$$\bar{x} = \tilde{q} \left(\begin{smallmatrix} 0 & 0 \\ 0 & \Omega \end{smallmatrix}\right) \tilde{q}^{\mathrm{T}},$$

where Ω is a diagonal matrix with positive diagonal, $s = \operatorname{rank} \bar{z}$, $\tilde{q} = [\tilde{q}^1, \tilde{q}^2]$, where $\mathcal{R}(\tilde{q}^2) = \mathcal{R}(\bar{z})$ and $\mathcal{R}(\tilde{q}^1)$ is its orthogonal complement. Then we obtain

Corollary 2.3.3 *Using the notation of Corollary 2.3.1, the following hold.*

(1') $(\bar{y}, \bar{z})$ *is an extreme point of Feas(D) iff the matrices in (2) of Corollary 2.3.2 with* $\tilde{q}$ *in place of* q *are linearly independent.*

(2') $(\bar{y}, \bar{z})$ *is a nondegenerate point of Feas(D) iff the matrices in (1) of Corollary 2.3.2 with* $\tilde{q}$ *in place of* q *span* $\mathcal{S}^{n-s}$.

■

In ordinary linear programs, a strictly complementary solution-pair always exists, as was shown by Goldman and Tucker [282]. This is not the case for SDP. An example was given in [24], which we reproduce here.
Let $n = m = 3$, $b = (1\,0\,0)^{\mathrm{T}}$

$$c = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad a^1 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad a^2 = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix}, \quad a^3 = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Then it is straightforward to see that $\bar{x}$ and $(\bar{y}, \bar{z})$ given by

$$\bar{x} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad \bar{y} = \begin{pmatrix} 0 & 0 & 0 \end{pmatrix}^{\mathrm{T}}, \quad \bar{z} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

are both nondegenerate, therefore both unique optimal solutions, which do not satisfy strict complementarity.
An instructive family of SDP examples where strict complementarity fails can be created from (convex) quadratically constrained quadratic programs (QCQP's). All such problems can be written in the form

$$\begin{array}{lll} \max & \langle b, y\rangle & \\ s.t. & \|d^i y - f^i\|^2 \leq & \gamma_i \;\; (i=1,\ldots,m) \end{array} \tag{QCQP}$$

with the d^i's being appropriate symmetric matrices and the f^i's vectors. The problem $(QCQP)$ then has an SDP representation

$$\begin{array}{ll} \max & \langle b, y\rangle \\ s.t. & \begin{pmatrix} I & d^i y - f^i \\ (d^i y - f^i)^T & \gamma_i \end{pmatrix} \succeq 0 \;\; (i=1,\ldots,m) \end{array}$$

which is equivalent to

$$\begin{array}{lll} \max & & \langle b, y\rangle \\ s.t. & & z^i \succeq 0 \\ & \displaystyle\sum_{j=1}^m \begin{pmatrix} 0 & -d^i_{\cdot j} \\ -d^i_{\cdot j} & 0 \end{pmatrix} y_j & +z^i = \begin{pmatrix} I & -f^i \\ -(f^i)^T & \gamma_i \end{pmatrix} \;\; (i=1,\ldots,m) \end{array} \tag{SDP-QCQP}$$

The proof of the following theorem is straightforward, therefore omitted.

Theorem 2.3.4 *Suppose that*

- *$\bar{y}$ is a unique optimal solution of $(QCQP)$ (therefore also of (SDP-QCQP) with the appropriate $\bar{z}^1, \ldots, \bar{z}^m$ slacks), and*
- *in $(QCQP)$ constraints 1 through k are the tight ones.*

Then (1) and (2) below are equivalent.

(1) b is a strictly positive combination of the gradients of the tight constraints of $(QCQP)$.

(2) The dual of (SDP-QCQP) has an optimal solution, which is strictly complementary with $(\bar{y}, \bar{z}^1, \ldots, \bar{z}^m)$.

■

Finally, since the dimension formulas that follow from parts (3) in the Primal and Dual Faces Theorems and Theorem 2.3.3 look quite elegant for *extreme and nondegenerate* solutions, we state them separately in Corollary 2.3.4. More refined and tighter formulas will follow in subsection 2.3.3.

Corollary 2.3.4 *Suppose that $\bar{x}$ is an extreme and nondegenerate point in Feas(P) and $(\bar{y}, \bar{z})$ in Feas(D),*

$$F = \text{face}(\bar{x}, K), \quad G = \text{face}(\bar{z}, K^*).$$

Then

$$\begin{array}{rcccl} \dim F & \leq & \dim Y & \leq & \dim F^{\triangle\perp}, \\ \dim G & \leq & \dim X - \dim Y & \leq & \dim G^{\triangle\perp}. \end{array} \tag{2.3.18}$$

In particular, if (P) and (D) are ordinary linear programs, with data defined as in Example 2.3.1,

$$r = \# \text{ of nonzeros in } \bar{x}, \quad s = \# \text{ of nonzeros in } \bar{z},$$

then

$$\begin{array}{rcl} r & = & m, \\ s & = & n - m. \end{array} \tag{2.3.19}$$

If (P) and (D) are semidefinite programs, with data defined as in Example 2.3.2,

$$r = \text{rank}\, \bar{x}, \quad s = \text{rank}\, \bar{z},$$

then

$$\begin{array}{rcccl} t(r) & \leq & m & \leq & t(n) - t(n-r), \\ t(s) & \leq & t(n) - m & \leq & t(n) - t(n-s). \end{array} \tag{2.3.20}$$

■

2.3.2 Tangent spaces

In this subsection we characterize the *tangent spaces* of the feasible sets of (P) and (D) at given feasible solutions. It turns out, that they can described in a manner similar to how the faces were described. The main result is

Theorem 2.3.5 (Tangent Spaces) *Suppose that K and K^* are nice, and both (P) and (D) satisfy the Slater condition. Let $\bar{x}$ and $(\bar{y}, \bar{z})$ be primal and dual feasible solutions, respectively,*

$$\begin{array}{rclrcl} F & = & \text{face}(\bar{x}, K), & \mathcal{T} & = & \tan(\bar{x}, Feas(P)), \\ G & = & \text{face}(\bar{z}, K^*), & \mathcal{U} & = & \tan((\bar{y}, \bar{z}), Feas(D)). \end{array}$$

Then

(1) $\mathcal{T} = F^{\triangle\perp} \cap \mathcal{N}(A)$.

(2) $$\begin{array}{rcl} \dim \mathcal{T} & = & \dim\, [F^{\triangle\perp} \cap \mathcal{N}(A)] \\ & = & \dim F^{\triangle\perp} - \dim Y + \dim\, [\text{lin}\, F^{\triangle} \cap I\!R(A^*)] \end{array}.$$

(3) $\mathcal{U} = (Y \times G^{\triangle\perp}) \cap \mathcal{N}(A^*, I)$.

$$\text{(4)} \quad \dim \mathcal{U} = \dim [\, G^{\triangle\perp} \cap I\!R(A^*)\,] = \dim G^{\triangle\perp} - (\dim X - \dim Y) + \dim [\,\mathrm{lin}\, G^{\triangle} \cap \mathcal{N}(A)\,] \, .$$

Proof First, from (2.2.11) in Lemma 2.2.1 recall

$$\tan(\bar{x}, K) = F^{\triangle\perp}, \quad \tan(\bar{z}, K^*) = G^{\triangle\perp}.$$

(1): Since (P) satisfies the Slater condition, by Proposition 2.2.3

$$\begin{array}{rcll} \mathrm{tcone}\,(\bar{x}, Feas(P)) & = & \mathrm{tcone}\,(\bar{x}, K) \cap \mathcal{N}(A) & \Rightarrow \\ {}_s \quad \mathcal{T} = \tan(\bar{x}, Feas(P)) & = & \tan(\bar{x}, K) \cap \mathcal{N}(A) & = \; F^{\triangle\perp} \cap \mathcal{N}(A). \end{array}$$

(2): Analogous to the proof of (1).
(3): Since

$$\begin{array}{rcl} \mathrm{aff}\, \mathcal{F} - \bar{x} & = & \mathrm{lin}\, F \cap \mathcal{N}(A), \\ \mathcal{T} & = & F^{\triangle\perp} \cap \mathcal{N}(A), \end{array}$$

we can use a similar calculation (with $F^{\triangle\perp}$ in place of $\mathrm{lin}\, F$) as in computing $\dim \mathcal{F}$ in the proof of the Primal Faces Theorem.
(4): Since

$$\begin{array}{rcl} \mathrm{aff}\, \mathcal{G} - (\bar{y}, \bar{z}) & = & (Y \times \mathrm{lin}\, G) \cap \mathcal{N}(A^*, I), \\ \mathcal{U} & = & (Y \times G^{\triangle\perp}) \cap \mathcal{N}(A^*, I), \end{array}$$

we can use a similar calculation (with $G^{\triangle\perp}$ in place of $\mathrm{lin}\, G$) as in computing $\dim \mathcal{G}$ in the proof of the Primal Faces Theorem. ■

Remark 2.3.5 If we can compute $\dim \mathcal{T}$ and $\dim F^{\triangle\perp}$, then equation (2) makes it possible to check whether $\bar{x}$ is nondegenerate (with an analogous statement relating dual nondegeneracy to equation (4)).

To keep the presentation relatively short, we do not write out the specializations for LP and SDP; they are quite straightforward.

2.3.3 The boundary structure inequalities

Putting together the results of the previous subsections, we shall now derive several instructive inequalities that relate the dimensions of

- Minimal faces in the primal and dual *cones* that contain a given optimal solution.
- Minimal faces in the primal and dual *feasible sets* that contain a given optimal solution.
- Tangent spaces at the primal and dual feasible sets at a given optimal solution.

We will call these inequalities the *boundary structure inequalities*. Theorem 2.3.6 below contains their statement, and an intuitive explanation follows.

Theorem 2.3.6 (Boundary Structure Inequalities) *Suppose that K and K^* are nice, and both (P) and (D) satisfy the Slater condition. Let $\bar{x}$ and $(\bar{y}, \bar{z})$ be complementary solutions of (P) and (D), respectively,*

$$\begin{array}{lll} F = \text{face}\,(\bar{x}, K), & \mathcal{F} = \text{face}\,(\bar{x}, Feas(P)), & \mathcal{T} = \tan(\bar{x}, Feas(P)), \\ G = \text{face}\,(\bar{z}, K^*), & \mathcal{G} = \text{face}\,((\bar{y}, \bar{z}), Feas(D)), & \mathcal{U} = \tan((\bar{y}, \bar{z}), Feas(D)). \end{array}$$

Then

(1) $\dim F - \dim Y + \dim \mathcal{U} \;\le\; \dim \mathcal{F} \;\le\; \dim G^{\triangle} - \dim Y + \dim \mathcal{U}$,

(2) $\dim \mathcal{F} + (\dim X - \dim \mathcal{T}) \;\ge\; \dim F + \dim F^{\triangle} + \dim\,[\,\text{res } F \cap I\!R(A^*)\,]$.

Suppose that strict complementarity holds. Then

(3) $\dim \mathcal{U} + \dim \mathcal{T} \;=\; \dim \mathcal{F} + \dim \mathcal{G} + \dim \text{res } F$,

(4) $\dim F^{\triangle\perp} - \dim \mathcal{T} \;\le\; \dim Y \;\le\; \dim F + \dim \mathcal{U}$,

with the left inequality at equality iff the primal solution is unique, and the right inequality at equality iff the dual solution is unique.

Proof First, from the Primal and Dual Faces Theorems, and the Tangent Spaces Theorem we recall

$$\begin{aligned} \dim \mathcal{F} &= \dim F - \dim Y + \dim\,[\,F^{\perp} \cap \mathcal{R}(A^*)\,], && (2.3.21) \\ \dim \mathcal{T} &= \dim F^{\triangle\perp} - \dim Y + \dim\,[\,\text{lin}\, F^{\triangle} \cap \mathcal{R}(A^*)\,], && (2.3.22) \\ \dim \mathcal{G} &= \dim\,[\,\text{lin}\, G \cap \mathcal{R}(A^*)\,], && (2.3.23) \\ \dim \mathcal{U} &= \dim\,[\,G^{\triangle\perp} \cap \mathcal{R}(A^*)\,]. && (2.3.24) \end{aligned}$$

Moreover,

$$F \;\subseteq\; G^{\triangle}, \quad F^{\perp} \;\supseteq\; G^{\triangle\perp}.$$

(1): The first inequality:

$$\begin{aligned} \dim \mathcal{F} &= \dim F - \dim Y + \dim\,[\,F^{\perp} \cap \mathcal{R}(A^*)\,] \\ &\ge \dim F - \dim Y + \dim\,[\,G^{\triangle\perp} \cap \mathcal{R}(A^*)\,] \\ &= \dim F - \dim Y + \dim \mathcal{U}. \end{aligned}$$

The second:

$$\begin{aligned} \dim \mathcal{F} &= \dim F - \dim Y + \dim\,[\,F^{\perp} \cap \mathcal{R}(A^*)\,] \\ &\le \dim G^{\triangle} - \dim Y + \dim\,[\,G^{\triangle\perp} \cap \mathcal{R}(A^*)\,] \\ &= \dim G^{\triangle} - \dim Y + \dim \mathcal{U}. \end{aligned}$$

as the number of linearly independent vectors in $\text{lin}\, G^{\triangle} \setminus \text{lin}\, F$ is at least as large as in $[\,F^{\perp} \cap \mathcal{R}(A^*)\,] \setminus [\,G^{\triangle\perp} \cap \mathcal{R}(A^*)\,]$. Also, this inequality is an equality if strict complementarity holds.

(2): Taking (2.3.21) $-$ (2.3.22) yields

$$\begin{aligned}
\dim \mathcal{F} - \dim \mathcal{T} &= \dim F - \dim F^{\triangle\perp} + \dim [\, F^{\perp} \cap \mathcal{R}(A^*)\,] \\
&\qquad - \dim [\operatorname{lin} F^{\triangle} \cap \mathcal{R}(A^*)\,] \quad \Leftrightarrow \\
\dim \mathcal{F} + (\dim X - \dim \mathcal{T}) &= \dim F + \dim F^{\triangle} + \dim [\, F^{\perp} \cap \mathcal{R}(A^*)\,] \\
&\qquad - \dim [\operatorname{lin} F^{\triangle} \cap \mathcal{R}(A^*)\,] \\
&\geq \dim F + \dim F^{\triangle} + \dim [\operatorname{res}\ F \cap \mathcal{R}(A^*)\,].
\end{aligned}$$

(3): By symmetry from (1), or from (2.3.23) we obtain

$$\begin{aligned}
\dim G - (\dim X - \dim Y) + \dim \mathcal{T} &\leq \dim \mathcal{G} \leq \dim F^{\triangle} \\
&\qquad -(\dim X - \dim Y) + \dim \mathcal{T} \quad \Leftrightarrow \\
\dim Y - \dim G^{\perp} + \dim \mathcal{T} &\leq \dim \mathcal{G} \leq \dim Y \\
&\qquad - \dim F^{\triangle\perp} + \dim \mathcal{T}.
\end{aligned}$$

Adding the second of these inequality chains to (1) yields

$$\begin{aligned}
\dim F - \dim G^{\perp} + (\dim \mathcal{U} + \dim \mathcal{T}) &\leq \dim \mathcal{F} + \dim \mathcal{G} \\
&\leq \dim G^{\triangle} - \dim F^{\triangle\perp} + (\dim \mathcal{U} + \dim \mathcal{T}).
\end{aligned}$$

therefore, if strict complementarity holds, then

$$\dim \mathcal{U} + \dim \mathcal{T} \quad = \quad \dim \mathcal{F} + \dim \mathcal{G} + \dim \operatorname{res}\ F.$$

(4): From (1) we have

$$\dim \mathcal{F} \quad = \quad \dim F - \dim Y + \dim \mathcal{U} \tag{2.3.25}$$

By symmetry, and strict complementarity

$$\begin{aligned}
\dim \mathcal{G} &= \dim F^{\triangle} - (\dim X - \dim Y) + \dim \mathcal{T} \\
&= \dim Y - \dim F^{\triangle\perp} + \dim \mathcal{T}. \qquad (2.3.26)
\end{aligned}$$

Putting (2.3.25) and (2.3.26) together we obtain

$$\begin{aligned}
\dim Y &= \dim F - \dim \mathcal{F} + \dim \mathcal{U} \\
&= \dim \mathcal{G} + \dim F^{\triangle\perp} - \dim \mathcal{T},
\end{aligned}$$

from which the two inequalities in (4) readily follow. ■

The geometry behind the inequalities Inequality (1) in the Boundary Structure Inequalities Theorem is a generalization of the Tijssen-Sierksma equality (2.3.15). Inequality (2) proves that

"the primal *feasible set* at $\bar{x}$ is at least as nonsmooth as the primal *cone* at $\bar{x}$".

To see what this means, for a convex set $C \subseteq \mathcal{R}^n$, and $x \in C$, the normal cone to C at x is defined as

$$\operatorname{ncone}(x, C) \quad = \quad \{\, v \mid \langle v, x \rangle \; \geq \; \langle v, y \rangle \, \text{for all}\, y \in C \,\}.$$

It is well known, that

$$\begin{aligned}
\text{ncone}\,(x,C)^* &= -\text{tcone}\,(x,C) \;\Rightarrow \\
\text{ncone}\,(x,C)^\perp &= \tan(x,C) \;\Rightarrow \\
\dim\ \text{ncone}\,(x,C) + \dim\ \tan(x,C) &= n. \qquad (2.3.27)
\end{aligned}$$

Also, if $\bar{x}$ and F are as in the Boundary Structure Inequalities Theorem, then

$$\text{ncone}\,(x,K) \;=\; -F^\triangle.$$

The quantity

$$\dim\ \text{face}\,(x,C) + \dim\ \text{ncone}\,(x,C)$$

is an intuitive measure of the nonsmoothness of C at x. Since face $(x,C) \subseteq \tan(x,C)$, by (2.3.27) this number is always less than or equal than n. If C is a polyhedron, then it is equal to n; and if C is a sphere (that is "as smooth as possible"), and x is on its boundary, then it is 1.
Now, inequality (2) can be rewritten as

$$\begin{aligned}
\dim \mathcal{F} + \dim\ \text{ncone}\,(\bar{x}, Feas(P)) \quad &\geq \dim F + \dim F^\triangle \qquad (2.3.28) \\
&\quad + \dim\ [\mathcal{R}(A^*) \cap \text{res}\ F]. \quad (2.3.29)
\end{aligned}$$

Since

- $\dim \mathcal{F}$ + dim ncone $(\bar{x}, Feas(P))$ is the measure of nonsmoothness of *Feas(P)* at $\bar{x}$, and
- $\dim F + \dim F^\triangle$ is the measure of nonsmoothness of K at $\bar{x}$,

inequality (2.3.28) indeed shows that "*Feas(P)* at $\bar{x}$ is at least as nonsmooth as K at $\bar{x}$". Confer this with the case, when $K = \mathcal{R}^n_+$; in this case, res $F = \{0\}$, and inequality (2.3.28) proves that *Feas(P)*, which is a polyhedron, is exactly as nonsmooth as K.
Finally, consider equality (3) in the theorem. This proves, that if strict complementarity holds, then

$$\dim \mathcal{T} + \dim \mathcal{U} \;\geq\; \dim \text{res}\ F. \qquad (2.3.30)$$

That is, unless the face F is polyhedral, ie. res $F = \{0\}$, then at least one of the primal and dual feasible sets will have a nontrivial tangent space at the given optimal solutions.
Specializing the Boundary Structure Inequalities Theorem for SDP, and using the substitution for the dimension of the normal cone, as described above, we obtain

Corollary 2.3.5 *If (P) and (D) are semidefinite programs, with data defined as in Example 2.3.2, $\bar{x}$ and $(\bar{y},\bar{z})$ are complementary solutions of (P) and (D), respectively,*

$$\begin{array}{lll}
r = \text{rank}\,\bar{x}, & \mathcal{F} = \text{face}\,(\bar{x}, Feas(P)), & \mathcal{T} = \tan(\bar{x}, Feas(P)), \\
s = \text{rank}\,\bar{z}, & \mathcal{G} = \text{face}\,((\bar{y},\bar{z}), Feas(D)), & \mathcal{U} = \tan((\bar{y},\bar{z}), Feas(D)).
\end{array}$$

Then

(1) $t(r) - m + \dim \mathcal{U} \ \leq\ \dim \mathcal{F} \ \leq\ t(n-s) - m + \dim\ \mathcal{U}$.

(2) $\dim \mathcal{F} + \dim \operatorname{ncone}(\bar{x}, Feas(P)) \ \geq\ t(r) + t(s) + \dim\,[\operatorname{res}\ F \cap \mathcal{R}(A^*)\,]$.

Suppose that strict complementarity holds. Then

(3) $\dim\ \mathcal{U} + \dim\ \mathcal{T} \ =\ \dim\ \mathcal{F} + \dim \mathcal{G} + r(n-r)$,

(4) $t(n) - t(n-r) - \dim\ \mathcal{T} \ \leq\ m \ \leq\ t(r) + \dim\ \mathcal{U}$,
with the left inequality at equality iff the primal solution is unique, and the right inequality at equality iff the dual solution is unique.

■

2.3.4 *The geometry of the feasible sets expressed with different variables*

Frequently, the dual problem is given in a form without the slack variable z, as

$$\begin{array}{lll} Max & \langle b, y\rangle & \\ s.t. & A^*y \preceq_{K^*} c & \qquad (D_y) \end{array}$$

The points in *Feas(D)* and *Feas(D)* are obviously in one-to-one correspondence: $y \in Feas(D_y) \Leftrightarrow (y, \phi(y)) \in Feas(D)$, with $\phi(y) = c - A^*y$. It is easy, if somewhat tedious to translate all results proved previously on the geometry of *Feas(D)* to describe the geometry of *Feas(D_y)*. The proof of Lemma 2.3.1 is deferred to Appendix B.

Lemma 2.3.1 *The following hold.*

(1) (a) *F is a face of Feas(D), if and only if* $\operatorname{Proj}_y(F)$ *is a face of Feas(D_y).*
(b) $\dim F = \dim \operatorname{Proj}_y(F)$.

(2) *If $(\bar{y}, \bar{z}) \in$ Feas(D), then*

(a) $\tan(\bar{y}, Feas(D_y)) = \operatorname{Proj}_y[\tan((\bar{y}, \bar{z}), Feas(D))]$,
(b) $\dim[\tan(\bar{y}, Feas(D_y))] = \dim[\tan((\bar{y}, \bar{z}), Feas(D))]$.

■

2.3.5 *A detailed example*

In this subsection we give a detailed example that illustrates the formulas

$$\begin{aligned} \dim \mathcal{G} &= \dim\,[\operatorname{lin} G \cap \mathcal{R}(A^*)\,] \\ &= \dim G - (\dim X - \dim Y) + \dim\,[\,G^{\perp} \cap \mathcal{N}(A)\,], \end{aligned} \tag{2.3.31}$$

$$\begin{aligned}
\dim \mathcal{U} &= \dim\,[\,G^{\triangle\perp} \cap \mathcal{R}(A^*)\,] \\
&= \dim G^{\triangle\perp} - (\dim X - \dim Y) + \dim\,[\,\mathrm{lin}\, G^{\triangle} \cap \mathcal{N}(A)\,] \qquad (2.3.32)\\
\dim \mathcal{F} &= \dim\,[\,\mathrm{lin}\, F \cap \mathcal{N}(A)\,], \qquad (2.3.33)\\
\dim \mathcal{T} &= \dim\,[\,F^{\triangle\perp} \cap \mathcal{N}(A)\,], \qquad (2.3.34)
\end{aligned}$$

from the Dual Faces and Tangent Spaces Theorems, by showing a primal-dual pair of cone program, with 3 different objective functions for the dual. For each objective, we will

- Find a complementary, hence optimal primal-dual solution pair.
- Compute $\dim G$ and $\dim G^{\triangle\perp}$ algebraically.
- Find $\dim \mathcal{U}$ and $\dim \mathcal{G}$ by inspecting the graph of the dual feasible set projected onto the y space.
- Deduce, whether the dual solution is nondegenerate from (2.3.32).
- If strict complementary holds, compute $\dim \mathcal{F}$ from (2.3.33) and $\dim \mathcal{T}$ from (2.3.34).

For each objective, the reader may easily check, that both (P) and (D) satisfy the Slater condition, hence the results of the Tangent Spaces Theorem will hold. The pair of cone programs are defined by $X = S^2 \times S^2$, $Y = \mathcal{R}^2$, $K = K^* = S^2_+ \times S^2_+$ as

$$(P) \quad \begin{array}{rll}
Min & \langle \begin{pmatrix} 0 & -1 \\ -1 & 1 \end{pmatrix}, x^1 \rangle & + \langle \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix}, x^2 \rangle \\
st. & x^1 \in S^2_+, & x^2 \in S^2_+ \\
 & \langle \begin{pmatrix} 0 & -1 \\ -1 & 0 \end{pmatrix}, x^1 \rangle & + \langle \begin{pmatrix} 0 & -1 \\ -1 & 0 \end{pmatrix}, x^2 \rangle = b_1 \\
 & \langle \begin{pmatrix} -1 & 0 \\ 0 & 0 \end{pmatrix}, x^1 \rangle & + \langle \begin{pmatrix} -1 & 0 \\ 0 & 0 \end{pmatrix}, x^2 \rangle = b_2
\end{array}$$

$$(D) \quad \begin{array}{rllll}
Max & b_1 y_1 + b_2 y_2 & & & \\
st. & z^1 \in S^2_+, & z^2 \in S^2_+ & & \\
 & y_1 \begin{pmatrix} 0 & -1 \\ -1 & 0 \end{pmatrix} & + y_2 \begin{pmatrix} -1 & 0 \\ 0 & 0 \end{pmatrix} & + z^1 & = \begin{pmatrix} 0 & -1 \\ -1 & 1 \end{pmatrix} \\
 & y_1 \begin{pmatrix} 0 & -1 \\ -1 & 0 \end{pmatrix} & + y_2 \begin{pmatrix} -1 & 0 \\ 0 & 0 \end{pmatrix} & + z^2 & = \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix}
\end{array}$$

and we will have 3 different choices for b. A simple calculation shows

$$Feas(D_y) = \{\,(y_1, y_2) \,|\, y_2 \geq (y_1 - 1)^2,\; y_2 \geq (y_1 + 1)^2\,\}$$

and $Feas(D_y)$ is shown on Figure 2.1. First, let $(b_1, b_2) = (0, -1)$. An optimal solution pair of (D) and (P) is given by

$$\begin{aligned}
\bar{y} = (0,1)^T, \quad \bar{z}^1 &= \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}, \quad \bar{z}^2 = \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix}, \\
\bar{x}^1 &= \tfrac{1}{2}\begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix}, \quad \bar{x}^2 = \tfrac{1}{2}\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}.
\end{aligned}$$

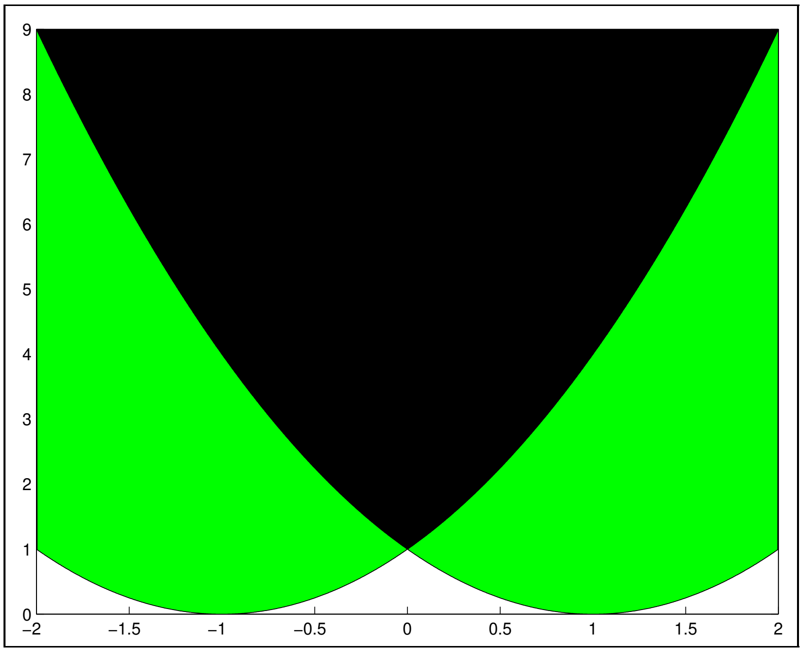


Figure 2.1 $\mathrm{Feas}(D_y)$ for the example of subsection 2.3.5

We find

$$\dim \mathcal{U} \;=\; 0, \qquad \dim \mathcal{G} \;=\; 0$$

from Figure 2.1, since the dimensions of the optimal face, and tangent space are the same in the (y, z)-space as in the y-space by the results of the previous subsection. Also,

$$\begin{aligned} \dim G &= 2, \qquad \dim G^{\triangle\perp} \;=\; 4, \\ \dim X - \dim Y &= 4. \end{aligned}$$

Plugging these numbers into (2.3.32) and (2.3.31) yields

$$\begin{aligned} \dim\,[\,\mathrm{lin}\, G^{\triangle} \cap \mathcal{N}(A)\,] &= 0, \\ \dim\,[\,G^{\perp} \cap \mathcal{N}(A)\,] &= 2. \end{aligned}$$

Therefore, $(\bar{y}, \bar{z}^1 \oplus \bar{z}^2)$ is nondegenerate, hence $\bar{x}^1 \oplus \bar{x}^2$ is a unique primal optimal solution. Also, as they are strictly complementary,

$$\dim \mathcal{T} \;=\; \dim\,[\,G^{\perp} \cap \mathcal{N}(A)\,] \;=\; 2.$$

Next, let $(b_1, b_2) = (-2, -1)$. An optimal solution pair of (D) and (P) is given by the same $(\bar{y}, \bar{z}^1 \oplus \bar{z}^2)$ as before and

$$\bar{x}^1 \;=\; 0, \quad \bar{x}^2 \;=\; \left(\begin{smallmatrix} 1 & 1 \\ 1 & 1 \end{smallmatrix}\right).$$

Note that now strict complementary does not hold; indeed, since $(-2,-1)$ is the gradient of the constraint function $-y_2+(y_1-1)^2$ at $\bar{y}$, so by Theorem 2.3.4 a strictly complementary solution pair cannot exist. Still, by nondegeneracy, the primal optimal solution is unique, and

$$\dim \mathcal{T} \;\leq\; \dim\,[\,G^{\perp} \cap \mathcal{N}(A)\,] \;=\; 2,$$

although the exact dimension of $\mathcal{T}$ now can only be computed algebraically (by finding $\dim\,[\,F^{\triangle\perp} \cap \mathcal{N}(A)\,]$).
Finally, let $(b_1, b_2) = (-4, -1)$. An optimal solution pair of (D) and (P) is now

$$\begin{array}{ll} \bar{y} = (-1,4)^T, \quad \bar{z}^1 = \begin{pmatrix} 4 & -2 \\ -2 & 1 \end{pmatrix}, & \bar{z}^2 = \begin{pmatrix} 4 & 0 \\ 0 & 1 \end{pmatrix}, \\ \qquad\qquad\qquad\; \bar{x}^1 = \begin{pmatrix} 1 & 2 \\ 2 & 4 \end{pmatrix}, & \bar{x}^2 = 0. \end{array}$$

We have

$$\begin{array}{rclrcl} \dim \mathcal{U} & = & 1, & \dim \mathcal{G} & = & 0, \\ \dim G & = & 4, & \dim G^{\triangle\perp} & = & 5, \\ \dim X - \dim Y & = & 4. & & & \end{array}$$

hence

$$\begin{array}{rcl} \dim\,[\,\mathrm{lin}\, G^{\triangle} \cap \mathcal{N}(A)\,] & = & 0, \\ \dim\,[\,G^{\perp} \cap \mathcal{N}(A)\,] & = & 0. \end{array}$$

therefore $(\bar{y}, \bar{z}^1 \oplus \bar{z}^2)$ is nondegenerate, so $\bar{x}$ is a unique primal optimal solution, and by strict complementarity

$$\dim \mathcal{T} \;=\; \dim\,[\,G^{\perp} \cap \mathcal{N}(A)\,] \;=\; 0.$$

2.4 SEMIDEFINITE COMBINATORICS

The subject of polyhedral combinatorics is the study of optimization problems via the polyhedral structure of their linear programming formulations. Combining knowledge on the extremal structure of polyhedra with the specifics of the LP-formulation can give valuable insights about the given optimization problem.
In this section we give three examples to illustrate that similar studies are possible, and worthwhile using convex programming, in particular SDP formulations. In a convex programming formulation one may want to characterize and study

(1) The extreme point optimal solutions in the primal and dual problems.

(2) The nondegenerate optimal solutions in the primal and dual problems.

(3) The solution pairs which lack strict complementarity.

2.4.1 *The Multiplicity of Optimal Eigenvalues*

Let A be an affine function from $\mathcal{R}^m$ to $\mathcal{S}^n$. The *eigenvalue-optimization problem* is

$$Min\ \{\ f_k(A(x)) : x \in \mathcal{R}^m\ \}, \qquad (EV_k)$$

where $f_k(A(x))$ is the sum of the k largest eigenvalues of $A(x)$.
The problem (EV_k) can be formulated as a semidefinite program, as it was shown by Alizadeh [21] and Nesterov and Nemirovskii [576]. In fact, it is the earliest instance of SDP that has been the subject of a computational study; see [168, 186]. In these studies, and in many more recent papers dealing with eigenvalue-optimization, the following phenomenon was observed. At optimal solutions of (EV_k) the eigenvalues of the optimal matrix tend to coalesce; if x^* achieves the minimum, then frequently $\lambda_k(A(x^*)) = \lambda_{k+1}(A(x^*))$, and $\lambda_k(A(x^*))$ often has multiplicity even larger than two.
The clustering phenomenon plays a central role in eigenvalue-optimization. As proven by Cullum, Donath, and Wolfe [168] the function f_k is differentiable at a fixed symmetric matrix a if and only if $\lambda_k(a) > \lambda_{k+1}(a)$. If this condition fails to hold, then the dimension of the subgradient of f_k at a grows quadratically with the multiplicity of $\lambda_k(a)$. Furthermore, if f_k is nonsmooth at $A(x^*)$ then generally the composite function $f_k \circ A$ is also nonsmooth at x^*.
Therefore, clustering tends to cause the nondifferentiability of the objective function $f_k \circ A$ at a solution point, making (EV_k) a "model problem" in nonsmooth optimization. We remark, that $f_k \circ A$ *may* be differentiable at x^* even if f_k is *not* differentiable at $A(x^*)$ (e.g. when $A(x) \equiv I$); this, however is usually not the case.

Example 2.4.1 Consider the problem with $k = 1$, and

$$A(x) \quad = \quad x_1 \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} + x_2 \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} + \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

The graph of the function $f_1 \circ A : \mathcal{R}^2 \longrightarrow \mathcal{R}$ is pictured in Figure 2. Clearly, at the unique optimal solution $x^* = (0, 0)$ $f_1 \circ A$ is not differentiable.
Somewhat surprisingly, in the case $n = m = 2$ every choice of the coefficient matrices in the definition of $A(x)$ gives rise to at least one point where the objective function is nonsmooth. Another example is given in Figure 3. Here the optimal solution is not unique, still, there is a nonsmooth optimum. ∎

In the rest of the section we outline, how the rankbounds on extreme matrices in SDP's can be used to explain eigenvalue-clustering in extreme point optimal solutions of (EV_k). This material is taken from [615], and proofs are omitted, or only sketched.

Lemma 2.4.1 *Fix $a \in \mathcal{S}^n$, and suppose that*

$$a \quad = \quad q(\mathrm{Diag}\,(\lambda(a)))q^T.$$

Figure 2.2 The graph of $f_1 \circ A$

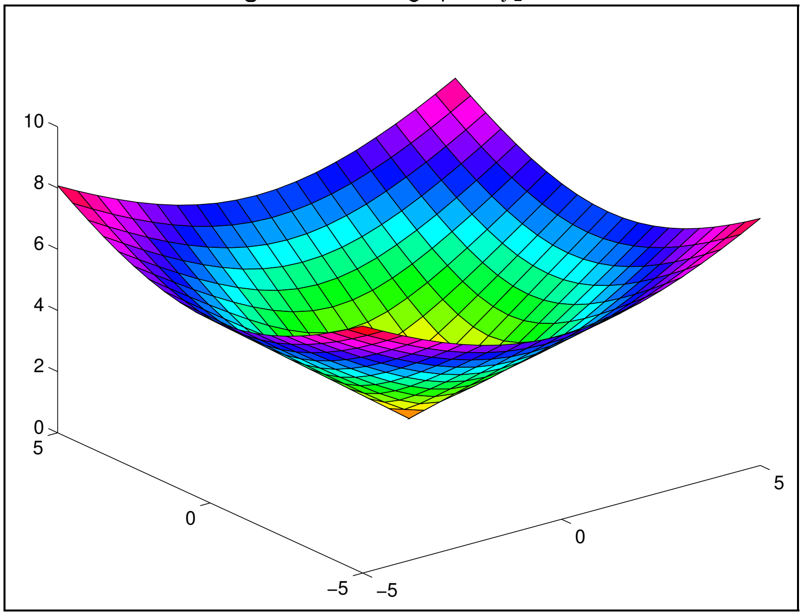
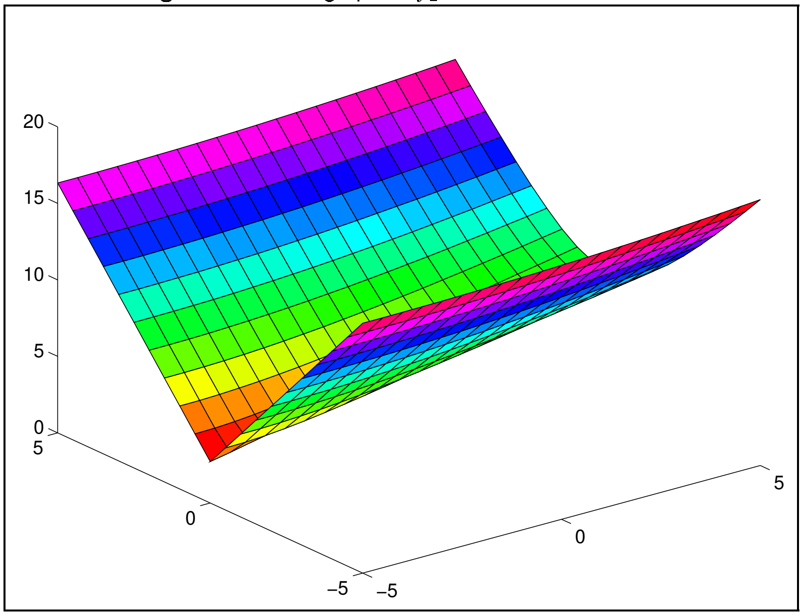


Figure 2.3 The graph of $f_1 \circ A$ for different data



The optimal value of the problem

$$\begin{array}{lll} Min_{z,v,w} & kz + \langle I, v\rangle & \\ s.t. & v \in \mathcal{S}^n_+,\ w \in \mathcal{S}^n_+ & \quad (2.4.36) \\ & zI + v - w = a & \end{array}$$

is $f_k(a)$*. The triple* (z^*, v^*, w^*) *is optimal to (2.4.36) if and only if*

$$\begin{array}{rcl} z^* & \in & [\,\lambda_k(a),\ \lambda_{k+1}(a)\,], \\ v^* & = & q\,(\,\mathrm{Diag}\,(\lambda(a) - z^*e\,)_+\,)\,q^T, \\ w^* & = & q\,(\,\mathrm{Diag}\,(\,z^*e - \lambda(a)\,)_+\,)\,q^T. \end{array}$$

■

Remark 2.4.1 The optimal solutions of (2.4.36) do not depend on the choice of a's eigenvectors. Suppose that the distinct eigenvalues of a are

$$\lambda_{i_1}(a), \ldots, \lambda_{i_r}(a)$$

in descending order and $\lambda_k(a) = \lambda_{i_s}(a)$. Then the distinct eigenvalues of v^* and w^* in Lemma 2.4.1 are

$$\lambda_{i_1}(a) - z^*, \ldots, \lambda_{i_{s-1}}(a) - z^* \quad \text{and} \quad z^* - \lambda_{i_{s+1}}(a), \ldots, z^* - \lambda_{i_{s-1}}(a),$$

respectively. Therefore, choosing different eigenvectors of a to represent the eigenspace corresponding to $\lambda_{i_j}(a)$ $(j = 1, \ldots, r)$ does not change v^* and w^*.

It follows from Lemma 2.4.1, that (by plugging $A(x)$ into the place of a in (2.4.36)) the problem (EV_k) can be formulated as the SDP

$$\begin{array}{lll} Min_{x,z,v,w} & kz + \langle I, v\rangle & \\ s.t. & v \in \mathcal{S}^n_+,\ w \in \mathcal{S}^n_+ & \quad (SDP_k) \\ & zI + v - w = A(x) & \end{array}$$

In the following we assume the mapping A to be fixed, and denote the set of optimal solutions of

- (2.4.36) (in the (z, v, w)-space) by $\Omega_k(a)$ for a given a.
- (EV_k) (in the x-space) by *Opt*(EV_k), and
- (SDP_k) (in the (x, z, v, w)-space) by *Opt*(SDP_k).

Theorem 2.4.1 *The following hold.*

(1) The point x^* *is an extreme point of Opt*(EV_k) *if and only if* $f_k \circ A$ *is strictly convex at* x^*.

(2) If nonempty, the set Opt(EV_k) has at least one extreme point.

(3) Let x^* be an extreme point of Opt(EV_k). If $m > k(n-k)$, then

$$\lambda_k(A(x^*)) \quad = \quad \lambda_{k+1}(A(x^*)). \tag{2.4.38}$$

Sketch of Proof of (3) If x^* is an extreme point of *Opt(EV_k)*, then

$$F^* \quad = \quad \{\, x^* \,\} \times \Omega_k(A(x^*))$$

is a face of *Opt(SDP_k)*, therefore also of *Feas(SDP_k)*. Since the only possible degree of freedom in F^* is in choosing z^*, we have $\dim F^* \le 1$. Now, let $(x^*, z^*, v^*, w^*) \in F^*$, $i = \operatorname{rank} v^*$, $j = \operatorname{rank} w^*$ and apply the Primal Faces Theorem with

$$X \quad = \quad \mathcal{R} \times \mathcal{R}^m \times \mathcal{S}^n \times \mathcal{S}^n, \quad Y \quad = \quad \mathcal{S}^n, \quad K \quad = \quad \mathcal{R} \times \mathcal{R}^m \times \mathcal{S}^n_+ \times \mathcal{S}^n_+$$

to conclude

$$1 + m + t(i) + t(j) \quad \le \quad t(n) + 1. \tag{2.4.39}$$

By (2.4.39) if $m > k(n-k)$, then $i = k$, $j = n - k$ cannot hold simultaneously. Since $(z^*, v^*, w^*) \in \Omega_k(A(x^*))$ was arbitrary, we conclude that (2.4.38) must hold. ■

2.4.2 *The geometry of a max-cut relaxation*

Consider the set

$$\mathcal{E}_n \quad = \quad \{\, x \in \mathcal{S}^n_+ \mid x_{ii} = 1 \, (i = 1, \ldots, n) \,\}.$$

Optimizing over $\mathcal{E}_n$ provides a strong relaxation of the maximum cut problem ([174], [275]); having this motivation, it was termed the *elliptope* and its geometry studied in [462] and [463].
In Theorem 2.4.2 we show that several results in these papers can be easily derived from our general framework. We consider $\mathcal{E}_n$ as a feasible set of a primal type SDP, ie. let $X = \mathcal{S}^n, Y = \mathcal{R}^n, K = K^* = \mathcal{S}^n_+$ and define A and A^* via $a^1, \ldots, a^n \in \mathcal{S}^n$, with a^i having 1 in the i^{th} position on the main diagonal, and 0 everywhere else.

Theorem 2.4.2 *Let $\bar{x} \in \mathcal{E}_n$. Then*

(1) $\bar{x}$ is a nondegenerate point.

(2) $\dim \tan(\bar{x}, \mathcal{E}_n) \; = \; 0$ *if and only if* $\operatorname{rank} \bar{x} = 1$.

Proof Let

$$\mathcal{T} \quad = \quad \tan(\bar{x}, \mathcal{E}_n), \qquad F \quad = \quad \operatorname{face}(\bar{x}, \mathcal{S}^n_+), \qquad r \quad = \quad \operatorname{rank} \bar{x}.$$

(1): After exchanging rows and columns, we may assume that the elements of $\mathcal{R}(A^*)$ are of the form

$$\begin{pmatrix} I & 0 & 0 \\ 0 & -I & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

Suppose that such a matrix is in $\operatorname{lin} F^{\triangle}$, with the size of the nonzero block equal to some $s \geq 0$. Then

$$\begin{pmatrix} I & 0 & 0 \\ 0 & -I & 0 \\ 0 & 0 & 0 \end{pmatrix} \in \operatorname{lin} F^{\triangle} \Rightarrow \begin{pmatrix} I & 0 & 0 \\ 0 & I & 0 \\ 0 & 0 & 0 \end{pmatrix} \in F^{\triangle} \Rightarrow \left\langle \begin{pmatrix} I & 0 & 0 \\ 0 & I & 0 \\ 0 & 0 & 0 \end{pmatrix}, \bar{x} \right\rangle = 0.$$

Since $\bar{x} \succeq 0$, its lower $(n-s) \times (n-s)$ principal submatrix must be zero. But $\bar{x}_{ii} = 1$ $(i = 1, \ldots, n)$, hence $s = 0$, proving

$$\operatorname{lin} F^{\triangle} \cap \mathcal{R}(A^*) \quad = \quad \{0\},$$

as required.

(2): From (2) in the Tangent Spaces Theorem,

$$\begin{aligned} \dim \mathcal{T} &= \dim F^{\triangle\perp} - \dim Y + \dim [\operatorname{lin} F^{\triangle} \cap \mathcal{R}(A^*)] \\ &= (t(n) - t(n-r)) - n + 0 \\ &\geq 0, \end{aligned} \tag{2.4.40}$$

with equality holding iff $r = 0$. ■

Remark 2.4.2 By the nondegeneracy of all points in $\mathcal{E}_n$ it also follows that for any $c \in \mathcal{S}^n$, the dual of

$$\begin{array}{ll} \min & \langle c, x \rangle \\ st. & x \in \mathcal{E}_n \end{array}$$

has a unique optimal solution, a result proved independently in [692] and [174]. Also, the formula for $\dim \mathcal{T}$ given in (2.4.40) is equivalent to the formula for $\dim \operatorname{ncone}(\bar{x}, \mathcal{E}_n)$ in [463] since

$$\dim \operatorname{ncone}(\bar{x}, \mathcal{E}_n) \quad = \quad t(n) - \dim \mathcal{T}.$$

2.4.3 The embeddability of graphs

As one more example of semidefinite combinatorics, we give a simple proof of a theorem of Barvinok [73].

Definition 2.4.1 *Let $G = (V, E; w)$ be a graph, with $\{w_{ij} : (i,j) \in E\}$ a nonnegative weighting on the edges. We say that G is* realizable *in $\mathcal{R}^d$, if there exists vectors $x^1, \ldots, x^n \in \mathcal{R}^d$, such that*

$$||x^i - x^j|| = w_{ij} \;\; \forall (i,j) \in E.$$

Theorem 2.4.3 *Suppose that $G = (V, E; w)$ is realizable in some dimension. Then it is realizable in $\mathcal{R}^d$, with d satisfying*

$$t(d) \quad \leq \quad |E|.$$

Proof Define the matrices $m^{ij} \in S^n$ for $(i,j) \in E$ by

$$m^{ij}_{ii} = m^{ij}_{jj} = 1, \;\; m^{ij}_{ij} = m^{ij}_{ji} = -1,$$

and with all other components zero. Then G is realizable in $\mathcal{R}^d$, iff

$$\begin{array}{rll} \exists\, x^1, \ldots, x^n \in \mathcal{R}^d & s.t.\ \langle x^i, x^i \rangle - 2\langle x^i, x^j \rangle + \langle x^j, x^j \rangle = w^2_{ij}\ \forall (i,j) \in E & \Leftrightarrow \\ \exists\, x \in \mathcal{R}^{n \times d} & s.t.\ \langle m^{ij}, x x^{\mathrm{T}} \rangle = w^2_{ij}\ \ \forall (i,j) \in E & \Leftrightarrow \\ \exists\, y \in \mathcal{S}^n_+,\ \mathrm{rank}\, y = d & s.t.\ \langle m^{ij}, y \rangle = w^2_{ij}\ \ \forall (i,j) \in E. & \end{array}$$

Therefore, G is representable in *some* dimension, iff the SDP in the last line above has a feasible solution, and it is representable in dimension $\leq d^*$, iff it has a feasible solution of rank $\leq d^*$. However, the first of these statements implies the second with d^* satisfying

$$t(d^*) \quad \leq \quad |E|,$$

by taking a solution which is an extreme point of the feasible set.

2.5 TWO ALGORITHMIC ASPECTS

In this section we describe, how two algorithmic aspects of cone programming can be handled using our framework: transforming a feasible solution into one, which is an extreme point of the feasible set; and determining by how much one can perturb the objective function in a given direction, while keeping the current solution optimal.

2.5.1 *Finding an extreme point solution*

This section is devoted to the following question: Given a feasible solution $\bar{x}$ to the problem

$$\begin{array}{ll} Min & \langle c, x \rangle \\ s.t. & x \in K \\ & Ax = b \end{array} \qquad (P)$$

we want to find an extreme point feasible solution, called $\bar{\bar{x}}$ with no worse objective value.

(1) Let $F = \mathrm{face}\,(\bar{x}, K)$.

(2) Find a nonzero $\Delta x \in \mathrm{lin}\, F \cap \mathcal{N}(A)$. If no such vector exists, set $\bar{\bar{x}} = \bar{x}$ and STOP.

(3) If $\langle c, \Delta x \rangle > 0$, set $\Delta x = -\Delta x$. Determine $\alpha^* = \max\{\, \alpha \,|\, \bar{x} + \alpha \Delta x \in F^{\triangle} \,\}$.

If $\alpha^* = +\infty$, STOP; (P) is unbounded.

(4) Set $\bar{x} = \bar{x} + \alpha^* \Delta x$, and go to (1).

The current $\bar{x}$ is not an extreme point, iff a nonzero Δx can be found in Step (2), by (4) in the Primal Faces Theorem. Also, by elementary convex analysis, if α^* is found in Step (3), then

$$\text{face}\,(\bar{x} + \alpha^* \Delta x, K) \quad \subset \quad \text{face}\,(\bar{x}, K)$$

therefore the algorithm is correct, and finite.

2.5.2 Sensitivity Analysis

Consider the primal-dual pair of cone-lp's parametrized by the scalar $t \geq 0$.

$$(P_t) \qquad \begin{array}{rl} Min & \langle c + t\Delta c, x\rangle \\ s.t. & x \in K \\ & Ax = b \end{array} \qquad\qquad \begin{array}{rl} Max & \langle b, y\rangle \\ s.t. & z \in K^* \\ & A^* y + z = c + t\Delta c \end{array} \qquad (D_t)$$

Let $\bar{x}$ be an optimal solution of (P_0), and

$$t^* \quad = \quad \sup\{\, t \mid \bar{x} \text{ is an optimal solution of } (P_t)\,\}.$$

Theorem 2.5.1 *Suppose (P_0) satisfies the Slater condition, and $\bar{x}$ and $(\bar{y}, \bar{z})$ is an optimal solution-pair of (P_0) and (D_0). Then*

(1) t^* is the optimal value of

$$\begin{array}{lrcl} Max & & t & \\ st. & \bar{z} + t\Delta z & \in & F^{\triangle} \\ & \Delta z & \in & \operatorname{lin} F^{\triangle} \\ & A^*\Delta y + \Delta z & = & \Delta c \end{array} \qquad (D_t')$$

(if the last two constraints of (D_t') are infeasible then the optimal value of (D_t') is understood to be 0).

(2) If $\bar{x}$ and $(\bar{y}, \bar{z})$ are strictly complementary, then $t^* > 0$.

(3) If $\bar{x}$ is nondegenerate, and there is $\Delta x \in \operatorname{lin} F^{\triangle}$ s.t. $A^*\Delta y + \Delta z = \Delta c$, then

$$t^* \quad = \quad \max\{\, t \mid \bar{z} + t\Delta z \in F^{\triangle}\,\}$$

for some such fixed Δz.

Proof

(1): Since (P_0), equivalently (P_t) for an arbitrary $t \geq 0$ satisfies the Slater condition, (D_t) attains its optimal value. Therefore t^* is the optimal value of

$$\begin{array}{rrcl} Max & t & & \\ st. & z & \in & F^\triangle \\ & A^*y + z & = & c + t\Delta c \end{array} \qquad (D_t'')$$

Then

$$\begin{array}{rl} t^* > 0 & \Leftrightarrow \\ \exists\, t > 0,\ (y(t), z(t)) \quad \text{feasible for } (D_t'') & \Leftrightarrow \\ \exists\, t > 0,\ (\Delta y, \Delta z) \quad \text{feasible for } (D_t') & \end{array}$$

where the second equivalence follows by taking

$$(\Delta y, \Delta z) \quad = \quad \frac{1}{t}\left((y(t) - \bar{y}), (z(t) - \bar{z})\right)$$

and this also proves that (D_t') and (D_t'') have the same optimal value.
(2): Straightforward.
(3): The primal optimal solution $\bar{x}$ is nondegenerate, iff

$$\operatorname{lin} F^\triangle \cap \mathcal{R}(A^*) \quad = \quad \{0\}.$$

In this case, the system

$$\begin{array}{rcl} \Delta z & \in & \operatorname{lin} F^\triangle \\ A^*\Delta y + \Delta z & = & \Delta c, \end{array}$$

which is part of (D_t') has a unique solution. The claim follows. ■

2.6 LITERATURE

For Section 2.2 Several important papers on the facial structure of convex cones and their duals were written by Barker, see [66, 67, 68], and by Tam, see [749]. Nice cones, (although without this name) were introduced by Borwein and Wolkowicz [128, 129], and used in the theory of *facial reduction*, see also Chapter 6.

For Section 2.3 A general theorem stating that all faces of the intersection of two convex sets arise as the intersection of two corresponding faces, was proven by Dubins in [189]; parts of the Primal and Dual faces Theorems already follow from it. The bounds on the rank of extreme matrices in SDP's were proved by Pataki in [612], and in [615]. The existence of a solution with a small rank (a rank that satisfies the bound stated for an extreme point) was proved independently by Barvinok [73]. A recent interesting result obtained by him (see [74]) shows that for semidefinite programs, if one of the equality constraints is of the form $\langle I, x\rangle = 1$, then this bound can be improved by 1.

The faces of "spectrahedra" ie. of feasible sets of SDP's were characterized by Ramana and Goldman [646]. Nondegeneracy and strict complementarity

for SDP were introduced and studied by Shapiro and Fan [701], Alizadeh, Haeberly and Overton [24], and for symmetric cones (ie. for the semidefinite, and second order cones) by Faybusovich [208]. For the second order cone, see also Alizadeh and Schmieta [26]. In [24] also the *genericity* of the property of strict complementarity was proved, ie. they showed that the instances of SDP which do not have a strictly complementary solution pair form a set of measure zero.
The general framework on the facial structure, nondegeneracy and strict complementarity for general cone programs was described by Pataki in [613], and in the dissertation [614]. Strict complementarity was also introduced independently by Luo, Sturm and Zhang [504].
For Section 2.4 The multiplicity of the critical eigenvalue in eigenvalue-optimization was studied in [615] by Pataki.
For Section 2.5 The algorithm to find an extreme point feasible solution of a cone-lp was given in [613] and [614] by Pataki. The method for sensitivity analysis is a generalization of the one given for SDP by Goldfarb and Scheinberg in [279]. The treatment given here is rather restrictive of course; it handles only a perturbation in a given direction, however, it is algorithmic, rather than purely structural. An extensive study on the structural properties of how the solution changes under perturbations is given by Bonnans and Shapiro [120]. For structural results on the sensitivity of a central solution of an SDP, see Sturm and Zhang [743]. A simple treatment on by how much the solution of an SDP can change, when the problem data is perturbed is described by Nayakkankuppam and Overton [556].

2.7 APPENDICES

2.7.1 A: The faces of the semidefinite cone

We will use some facts from linear algebra stated in

Proposition 2.7.1 *Suppose that x and y are in $\mathcal{S}^n_+$. Then*

(1)
$$\mathcal{R}(x+y) \;=\; \mathcal{R}(x) + \mathcal{R}(y)$$

(2) If $\mathcal{R}(y) \subseteq \mathcal{R}(x)$, then there exists $z \in \mathcal{S}^n_+$ such that
$$x \;\in\; (y, z)$$
In other words, the line-segment from y to x can be extended past x within $\mathcal{S}^n_+$.

(3) $\langle x, y\rangle = 0$ if and only if $\mathcal{R}(x) \perp \mathcal{R}(y)$.

To prove (2.2.6) in Example 2.2.2, it is enough to verify the PSD Faces Theorem below. Formula (2.2.7) for the conjugate face will then follow by using (3) in Proposition 2.7.1.

Theorem 2.7.1 (PSD Faces) *A set is a face of $\mathcal{S}^n_+$ if and only if it is of the form*

$$\{\, x \mid x \in \mathcal{S}^n_+,\ \mathcal{R}(x) \subseteq L \,\}$$

for some subspace L of $\mathcal{R}^n$.

Proof Denote the set in the statement of the theorem by $F(L)$.
(If) Proposition 2.7.1 (1) implies both the convexity of $F(L)$, and

$$\text{if } x + y \in F(L) \quad \text{then} \quad x \in F(L), \text{ and } y \in F(L)$$

(Only if) Let F be a face of $\mathcal{S}^n_+$. Define L as the subspace spanned by the range spaces of all matrices in F. We claim

$$F = F(L)$$

with the inclusion $\subseteq$ being obvious. To show the reverse, we first construct a matrix $\hat{x} \in F$ with $\mathcal{R}(\hat{x}) = L$. As there are matrices $x^1, \ldots, x^k$ in F such that

$$L = \sum_{i=1}^{k} \mathcal{R}(x^i) = \mathcal{R}(\sum_{i=1}^{k} x^i)$$

therefore

$$\hat{x} = \sum_{i=1}^{k} x^i$$

will do. Now, pick any $y \in F(L)$. By (2) in Proposition 2.7.1 there exists $z \in \mathcal{S}^n_+$ such that

$$\hat{x} \in (y, z)$$

Since F is a face of $\mathcal{S}^n_+$ containing $\hat{x}$, we conclude that y (and z) must be in F.
■

2.7.2 B: Proof of Lemma 2.3.1

Recall that $\phi(y) = c - A^*y$.
Proof of (1) We have

$$\begin{array}{c} F \lhd Feas(D) \qquad \Leftrightarrow \\ (y^1, z^1),\ (y^2, z^2) \in Feas(D) \text{ and } \tfrac{1}{2}[(y^1, z^1) + (y^2, z^2)] \in F \\ \text{imply} \\ (y^1, z^1), (y^2, z^2) \in F \end{array} \qquad \begin{array}{r} \\ (2.7.44) \\ \\ (2.7.45) \end{array}$$

But, (2.7.44) is equivalent to

$$(y^1, \phi(y^1)), (y^2, \phi(y^2)) \in Feas(D) \text{ and } (\tfrac{1}{2}(y^1 + y^2), \phi(\tfrac{1}{2}(y^1 + y^2))) \in F \Leftrightarrow$$
$$y^1, y^2 \in Feas(D_y) \text{ and } \tfrac{1}{2}(y^1 + y^2) \in \text{Proj}_{\,y}(F) \qquad (2.7.46)$$

and (2.7.45) to

$$\begin{aligned} (y^1, \phi(y^1)),\ (y^2, \phi(y^2)) \in F \quad \Leftrightarrow \\ y^1,\ y^2 \in \text{Proj}_{\,y}(F) \end{aligned} \tag{2.7.47}$$

Therefore,

$$F \triangleleft Feas(D) \quad \Leftrightarrow \quad (2.7.46) \text{ implies } (2.7.47) \quad \Leftrightarrow \quad \text{Proj}_{\,y}(F) \triangleleft Feas(D_y)$$

This proves (1)(a) in Lemma 2.3.1. To see (1)(b), one only needs to note

$$F \;=\; \{\, (y, \phi(y)) \mid y \in \text{Proj}_{\,y}(F) \,\}$$

Proof of (2) By the proof of the Tangent Spaces Theorem

$$\text{tcone}\,((\bar{y}, \bar{z}), Feas(D)) \;=\; \{\, (y, z) \mid A^* y + z = 0,\ z \in \text{tcone}\,(\bar{z}, K^*) \,\} \tag{2.7.48}$$

therefore

$$\text{Proj}_{\,y}\,[\,\text{tcone}\,((\bar{y}, \bar{z}), Feas(D))\,] \;=\; \{\, y \mid -A^* y \in \text{tcone}\,(c - A^* \bar{y}, K^*) \,\} \tag{2.7.49}$$

A straightforward calculation shows

$$\begin{aligned} \text{dir}\ (\bar{y}, Feas(D_y)) &= (-A^*)^{-1}\,[\,\text{dir}\ (c - A^* \bar{y}, K^*)\,] \;\Rightarrow \\ \text{Cl}\,\text{dir}\ (\bar{y}, Feas(D_y)) &= (-A^*)^{-1}\,[\,\text{Cl}\,\text{dir}\ (c - A^* \bar{y}, K^*)\,] \;\Leftrightarrow \\ \text{tcone}\,(\bar{y}, Feas(D_y)) &= (-A^*)^{-1}\,[\,\text{tcone}\,(c - A^* \bar{y}, K^*)\,] \end{aligned} \tag{2.7.50}$$

where the first implication follows by Theorem 6.7 in [662]. Putting (2.7.50) and (2.7.49) together yields

$$\begin{aligned} \text{tcone}\,(\bar{y}, Feas(D_y)) &= \text{Proj}_{\,y}\,[\,\text{tcone}\,((\bar{y}, \bar{z}), Feas(D))\,] \;\Rightarrow \\ \tan(\bar{y}, Feas(D_y)) &= \text{Proj}_{\,y}\,[\,\tan((\bar{y}, \bar{z}), Feas(D))\,] \end{aligned}$$

which proves (2)(a). Also, by (2.7.48)

$$\tan((\bar{y}, \bar{z}), Feas(D)) \;=\; \{\, (y, -A^* y) \mid -A^* y \in \tan(c - A^* \bar{y}, K^*) \,\}$$

which shows that projecting $\tan((\bar{y}, \bar{z}), Feas(D))$ onto the y-space preserves its dimension, proving (2)(b).