Chapter II Conic optimization

<u>Goal</u>: Develop general theory of convex optimization in the framework of conic programs, i.e. linear programs over convex cones.

§1 Convex cones

In the following we shall work with subsets of an n-dimensional Euclidean space E with inner product (x,y). For concreteness we identify E with $\mathbb{R}^n$ and $(x,y) = x^T y = \sum_{i=1}^{n} x_i y_i$.

<u>Def. 1</u> (a) $K \neq \emptyset$, $K \subseteq \mathbb{R}^n$ is called <u>(convex) cone</u> if $\forall \alpha, \beta \in \mathbb{R}_+$, $x, y \in K$: $\alpha x + \beta y \in K$.

(b) K is <u>pointed</u> if $K \cap (-K) = \{0\}$

(c) K is <u>proper</u> („ordentlich") if K is closed, pointed & full dimensional (i.e. $\operatorname{span} K = \mathbb{R}^n$, or $\operatorname{int} K \neq \emptyset$).

Proper cones are <u>positivity domain</u> in $\mathbb{R}^n$; they define a <u>partial order</u> by

$$x \succeq_K y \iff x - y \in K.$$

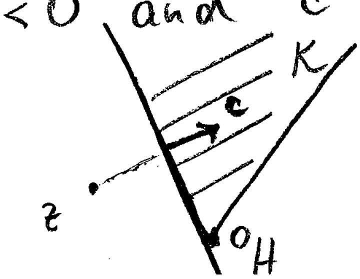
<u>Remarks</u>
- $K = [0, \infty)$, then $x \succeq_K y \iff x \geq y$
- can happen: neither $x \succeq_K y$ nor $y \succeq_K x$.
  ($\rightsquigarrow$ partial order)
- $x \succ_K y \iff x - y \in \operatorname{int} K$; strict inequality.

<u>Def. 2</u> Let $K \subseteq \mathbb{R}^n$ be a convex cone. Its <u>dual cone</u> is defined as

$$K^* = \{ y \in \mathbb{R}^n : x^T y \geq 0 \quad \forall x \in K \}.$$

<u>Recall from OR</u>: Separation theorem

<u>Theorem 3</u> Let $K \subseteq \mathbb{R}^n$ be a closed convex cone. Suppose $z \notin K$. Then there exist $c \in \mathbb{R}^n$ so that $c^T z < 0$ and $c^T x \geq 0$ for all $x \in K$. (i.e. $c \in K^*$)



H = separating hyperplane of z and K.

<u>Proposition 4</u> Let $K \subseteq \mathbb{R}^n$ be a convex cone. Then,

(i) $K \subseteq (K^*)^*$

(ii) if K is closed, $K = (K^*)^*$.

<u>Proof</u> (i) clear

(ii) to show $K \supseteq (K^*)^*$.

Use the complement. Let $z \in \mathbb{R}^n \setminus K$. Then by Thm. 3 there is $c \in \mathbb{R}^n$ s.t. $c^T z < 0$ & $c^T x \geq 0$ $\forall x \in K$. So $c \in K^*$ and $z \notin (K^*)^*$. $\boxtimes$

§2 Examples

(i) Polyhedral cones

<u>Def. 1</u> For $X \subseteq \mathbb{R}^n$ define the <u>conic hull</u> of X

$$\operatorname{cone} X = \left\{ \sum_{i=1}^{N} \alpha_i x_i : N \in \mathbb{N},\ x_1, \dots, x_N,\ \alpha_1, \dots, \alpha_N \in \mathbb{R}_+ \right\}.$$

$$= \bigcap_{\substack{C \text{ cone} \\ X \subseteq C}} C$$

cone X is the cone <u>generated</u> by X. If $|X| < \infty$, then cone X is called <u>finitely generated</u>.

A cone C is called <u>polyhedral</u> if there is a matrix $A \in \mathbb{R}^{m \times n}$ such that

$$C = \{x \in \mathbb{R}^n : Ax \leq 0\}.$$

<u>Theorem 2</u> (Minkowski-Weyl, ⇝ OR)

A convex cone is polyhedral if and only if it is finitely generated.

<u>Theorem 3</u> (Carathéodory)

Let $C = \operatorname{cone} X$ be a cone. For every $c \in C$ there are linearly independent vectors $x_1, \dots, x_k \in X$ s.t. $c \in \operatorname{cone}\{x_1, \dots, x_k\}$.

<u>Corollary 4</u> (of Thm. 2)

Let

$$C = \{x \in \mathbb{R}^n : Ax \leq 0\} = \operatorname{cone}\{x_1, \dots, x_N\}$$

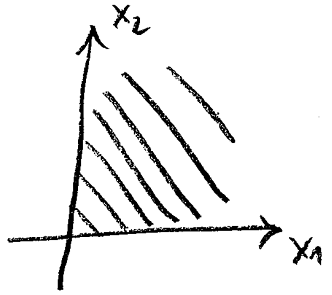
be a polyhedral cone. Then its dual cone is

$$C^* = \operatorname{cone}\{-a_1, \dots, -a_m\} = \{x \in \mathbb{R}^n : -x_1^T x \leq 0, \dots, -x_N^T x \leq 0\},$$

where $a_1^T, \dots, a_m^T$ rows of A.

(ii) Nonnegative orthant

$$\mathbb{R}^n_+ = \{x \in \mathbb{R}^n : x_1 \geq 0, \dots, x_n \geq 0\} = \mathbb{R}_+ \times \dots \times \mathbb{R}_+$$
$$= \text{cone}\{e_1, \dots, e_n\}$$
$$= (\mathbb{R}^n_+)^*$$

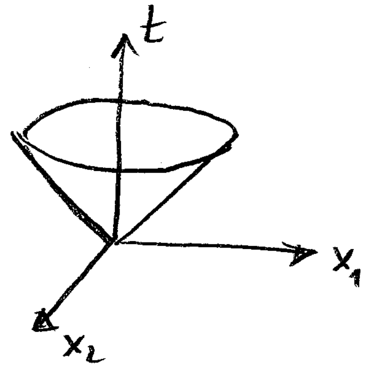


Here $K \times K' = \{(k, k') \in \mathbb{R}^{n+m} : k \in K, k' \in K'\}$

direct product of two convex cones $K \subseteq \mathbb{R}^n$, $K' \subseteq \mathbb{R}^m$.

(iii) Lorentz cone $\hat{=}$ light cone $\hat{=}$ ice cream cone

$$\mathcal{L}^{n+1} = \{(x, t) \in \mathbb{R}^{n+1} : \|x\| = \sqrt{x_1^2 + \dots + x_n^2} \leq t\}.$$



Can show $(\mathcal{L}^{n+1})^* = \mathcal{L}^{n+1}$.

(iii) Cone of positive semidefinite matrices

$\hat{=}$ semidefinite cone $\hat{=}$ PSD cone.

$$S^n = \{X \in \mathbb{R}^{n \times n} : X \text{ symmetric}, X_{ij} = X_{ji}\}$$

space of symmetric matrices, $\dim S^n = \frac{n(n+1)}{2}$

The <u>standard basis</u> of S^n is

$$E_{ij} = \tfrac{1}{2}(e_i e_j^T + e_j e_i^T) \qquad 1 \le i \le j \le n.$$

S^n is a Euclidean vector space with inner product

$$\langle A, B \rangle = \mathrm{Tr}(AB) = \sum_{i=1}^n \sum_{j=1}^n A_{ij} B_{ij},$$

where $\mathrm{Tr}(A) = \sum_{i=1}^n A_{ii}$ denotes the <u>trace</u> of A.

[Recall: $\mathrm{Tr}(AB) = \mathrm{Tr}(BA)$.]

$(S^n, \langle \cdot, \cdot \rangle)$ is isometric to $\left(\mathbb{R}^{\frac{n(n+1)}{2}}, (x,y) \mapsto x^T y\right)$

via $T : S^n \to \mathbb{R}^{\frac{n(n+1)}{2}}$

$$A \mapsto \begin{pmatrix} A_{11}, & \sqrt{2} A_{12}, & \sqrt{2} A_{13}, & \dots, & \sqrt{2} A_{1n} \\ & A_{22}, & \sqrt{2} A_{23}, & \dots, & \sqrt{2} A_{2n} \\ & & & \ddots & \\ & & & & A_{nn} \end{pmatrix}$$

T takes essentially the upper triangular part of A.

Theorem 5 (Spectral decomposition)

Every symmetric matrix $X \in S^n$ has a spectral decomposition, i.e. there is an orthonormal basis $u_1, \dots, u_n \in \mathbb{R}^n$ and real numbers $\lambda_1, \dots, \lambda_n$ such that

$$X = \sum_{i=1}^{n} \lambda_i u_i u_i^T$$

holds; $\lambda_1, \dots, \lambda_n$ are the eigenvalues of X, $u_1, \dots, u_n$ corresponding eigenvectors.

In matrix form: $X = P^T D P$, where

$P \in O(n)$ matrix in orthogonal group

$$O(n) = \{ P \in \mathbb{R}^{n \times n} : PP^T = P^T P = I_n \}, \quad I_n \text{ identity matrix},$$

and where $D = \mathrm{Diag}(\lambda_1, \dots, \lambda_n) = \begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{bmatrix}$ diagonal matrix.

Proof $\rightarrow$ Linear Algebra.

Def. 6 $X \in S^n$ is called positive semidefinite if for all $v \in \mathbb{R}^n$: $v^T X v \geq 0$ holds; it is called positive definite if for all $v \neq 0$: $v^T X v > 0$.

<u>Notation</u>: $X \succeq 0$ or $X \succ 0$

$\uparrow$ `$\succeq 0$`

<u>Note</u>: $v^T X v = \langle X, vv^T \rangle$ (direct check)

<u>Proposition 7</u> For $X \in S^n$ the following assertions are equivalent:

(i) $X \succeq 0$

(ii) all eigenvalues of X are nonnegative

(iii) $\exists\, L \in \mathbb{R}^{n \times k}$: $X = LL^T$, a <u>Cholesky factorisation</u> of X. (L can be chosen as lower triangular matrix $L = \left(\triangle \right)$)

(iv) $\exists\, v_1, \dots, v_n \in \mathbb{R}^k$: $X_{ij} = v_i^T v_j$, $i, j \in [n]$, a <u>Gram representation</u> of X

(v) All principal minors of X are nonnegative.

<u>Proof</u> $\rightarrow$ Linear Algebra

<u>Remarks</u>
- Corresponding characterisations for $X \succ 0$
- smallest possible k in (iii) and (iv) is equal to rank X.

Def 8 $S^n_+ = \{X \in S^n : X \succeq 0\}$

is the cone of positive semidefinite matrices (PSD cone);

$S^n_{++} = \{X \in S^n : X \succ 0\}$

Proposition 9 $S^n_+ = \text{cone}\{xx^T : x \in \mathbb{R}^n\}$ is a proper convex cone, its interior is S^n_{++} and it is self dual $(S^n_+)^* = S^n_+$.

Proof

- $S^n_+ = \text{cone}\{xx^T : x \in \mathbb{R}^n\}$: follows from spectral decomposition.

- S^n_+ closed : $S^n_+ = \bigcap_{v \in \mathbb{R}^n} \{X : \langle X, vv^T \rangle \geq 0\}$
  is intersection of closed halfspace

- S^n_+ pointed : Let $X \in S^n$ and denote the smallest (largest) eigenvalue of X by $\lambda_{min}(X)$ $(\lambda_{max}(X))$.
  Suppose $X, -X \in S^n_+$. Then,

$$0 \leq \lambda_{min}(-X) = -\lambda_{max}(X) \leq 0.$$

Hence, $X = 0$.

- $\operatorname{int} S^n_+ = S^n_{++}$: $\to$ exercise

- $(S^n_+)^* = S^n_+$: Recall $(S^n_+)^* = \{ Y \in S^n : \langle X, Y \rangle \geq 0 \ \forall X \in S^n_+ \}$.

„$\subseteq$": For $x \in \mathbb{R}^n$ consider rank-1-matrix $xx^T \in S^n_+$.

For $Y \in (S^n_+)^*$ we have

$$0 \leq \langle Y, xx^T \rangle = x^T Y x \implies Y \in S^n_+$$

„$\supseteq$": Consider $X, Y \in S^n_+$. Y has a spectral decomposition with $Y = \sum_{i=1}^n \lambda_i u_i u_i^T$ with $\lambda_i \geq 0$.

Then

$$\langle X, Y \rangle = \sum_i \lambda_i \langle X, u_i u_i^T \rangle = \sum_i \underbrace{\lambda_i}_{\geq 0} \overbrace{u_i^T X u_i}^{\geq 0} \geq 0,$$

hence $X \in (S^n_+)^*$.

Some properties and constructions of psd matrices

Proposition 10 (Schur complement)

Let $X \in S^n$ be a matrix in the following block form

$$X = \begin{bmatrix} A & B \\ B^T & C \end{bmatrix} \quad \text{with } A \in S^p, C \in S^{n-p}, B \in \mathbb{R}^{p\times(n-p)}.$$

Suppose A is invertible. Then $X \in S_+^n$ if and only if

$$A \in S_{++}^p \quad \text{and} \quad \underbrace{C - B^T A^{-1} B}_{\text{\underline{Schur complement} of } A \text{ in } X} \in S_+^{n-p}.$$

<u>Proof</u> Verify:

$$X = P^T \begin{bmatrix} A & 0 \\ 0 & C - B^T A^{-1} B \end{bmatrix} P \quad \text{with } P = \begin{bmatrix} I & A^{-1}B \\ 0 & I \end{bmatrix}.$$

$\boxtimes$

Def. 11 Define the <u>Schur-Hadamard product</u> of $X, Y \in S^n$ by $(X \circ Y)_{ij} = X_{ij} Y_{ij}$.

Prop. 12 Suppose $X, Y \in S^n_+$, then $X \circ Y \in S^n_+$.

Proof Express X, Y by spectral decomposition and verify $x^T (X \circ Y) x \geq 0$ directly. $\boxtimes$

Def 13 The <u>Kronecker product</u> ($\rightsquigarrow$ <u>tensor product</u>) of $A \in \mathbb{R}^{m \times n}$, $B \in \mathbb{R}^{p \times q}$ is $A \otimes B \in \mathbb{R}^{mp \times nq}$ defined by

$$(A \otimes B)_{((i,h),(j,k))} = A_{i,j} B_{h,k}.$$

$A \otimes B$ is a $m \times n$ block matrix with blocks of size $p \times q$ (and vice versa)

$$A \otimes B = \begin{bmatrix} A_{11} B & A_{12} B & \dots & A_{1n} B \\ & \dots & & \\ A_{m1} B & A_{m2} B & \dots & A_{mn} B \end{bmatrix}.$$

Proposition 14

(i) $(A \otimes B)(C \otimes D) = AC \otimes BD$

(ii) If $A \in S^m$ has eigenvalues $\alpha_1, \dots, \alpha_m$ and if $B \in S^n$ has eigenvalues $\beta_1, \dots, \beta_n$. Then, $A \otimes B$ has eigenvalues $\alpha_i \beta_j$, $i \in [m]$, $j \in [n]$.

(iii) $A \in S^m_+$, $B \in S^n_+ \Longrightarrow A \otimes B \in S^{mn}_+$.

Proof direct verification. $\boxtimes$

(iv) Cone of completely positive matrices

Def 15 $CP_n = \text{cone}\{xx^T : x \in \mathbb{R}^n_+\} \subseteq S^n_+ \subseteq S^n$ is called the cone of completely positive matrices.

Facts CP_n is a proper convex cone. Its dual cone is

$$(CP_n)^* = COP_n = \{X \in S^n : x^T X x \geq 0 \ \forall x \in \mathbb{R}^n_+\},$$

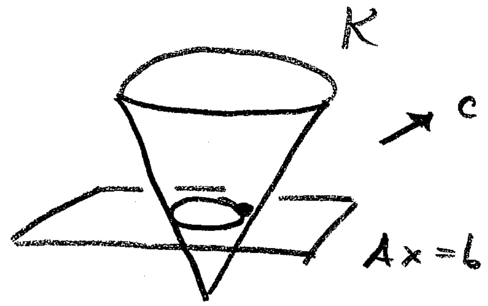
the cone of copositive matrices.

§ 2 Conic programs

<u>Def. 1</u> Let $K \subseteq \mathbb{R}^n$ be a proper convex cone. Given $c, a_1, \dots, a_m \in \mathbb{R}^n$, $b_1, \dots, b_m \in \mathbb{R}$, a <u>primal conic program</u> (in standard form) is the following maximization problem

$$(P) \quad p^* = \sup \{ c^T x : x \in K, \ a_1^T x = b_1, \dots, a_m^T x = b_m \}$$
$$= \sup \{ c^T x : x \in K, \ Ax = b \},$$

with $A = \begin{bmatrix} - a_1^T - \\ \vdots \\ - a_m^T - \end{bmatrix} \in \mathbb{R}^{m \times n}$.



The corresponding <u>dual conic program</u> is the following minimization problem

$$(D) \quad d^* = \inf \{ b^T y : y \in \mathbb{R}^m, \ \sum_{j=1}^m y_j a_j - c \in K^* \}$$
$$= \inf \{ b^T y : y \in \mathbb{R}^m, \ A^T y - c \in K^* \}.$$

<u>Remark</u> We say $x \in \mathbb{R}^n$ is <u>feasible</u> for (P) if $x \in K$ and $Ax = b$. It is <u>strictly feasible</u> if additionally $x \in \operatorname{int} K$

Similarly, $y \in \mathbb{R}^m$ is <u>feasible</u> for (D) if $A^T y - c \in K^*$, and <u>strictly feasible</u> if $A^T y - c \in \operatorname{int} K^*$.

<u>Geometric interpretation</u>

Define linear subspaces

$$L = \{ x \in \mathbb{R}^n : Ax = 0 \} = \ker A$$

$$L^\perp = \{ z \in \mathbb{R}^n : x^T z = 0 \text{ for all } x \in L \}$$

$$= \{ y^T A : y \in \mathbb{R}^m \}$$

Suppose there is $x_0 \in \mathbb{R}^n$ with $Ax_0 = b$. Then,

$$b^T y = (Ax_0)^T y = x_0^T A^T y = x_0^T (A^T y - c) + x_0^T c,$$

and

$$p^* = \sup \{ c^T x : x \in K \cap (x_0 + L) \}$$

$$d^* = c^T x_0 + \inf \{ x_0^T z : z \in K^* \cap (-c + L^\perp) \}.$$

Rewrite the last equation in primal form:

$$d^* = c^T x_0 - \sup \{ -x_0^T z : z \in K^* \cap (-c + L^\perp) \}$$

and take the dual.

Then one gets back the primal because of bipolarity $K = (K^*)^*$ (Prop. 1.4).

Examples

Def. 2 Let $K_1 \subseteq \mathbb{R}^{n_1}$, $K_2 \subseteq \mathbb{R}^{n_2}$ be convex cones. The direct product of K_1 and K_2 is defined as

$$K_1 \times K_2 = \{ (x_1, x_2) \in \mathbb{R}^{n_1+n_2} : x_1 \in K_1,\ x_2 \in K_2 \}.$$

Prop. 3 (a) If K_1, K_2 are both proper convex cones, then so is $K_1 \times K_2$.

(b) $(K_1 \times K_2)^* = K_1^* \times K_2^*$.

Proof easy

(i) Linear programming (LP)

$$\boxed{K = \mathbb{R}^n_+}$$

$$p^* = \sup\{c^T x : \overbrace{x \in \mathbb{R}^n_+}^{x \geq 0 \text{ (componentwise)}}, Ax = b\}$$

$$d^* = \inf\{b^T y : y \in \mathbb{R}^m, A^T y - c \in \mathbb{R}^n_+\}.$$

(ii) Conic quadratic programming (CQP)

$$\boxed{K = \mathcal{L}^{n_1+1} \times \mathcal{L}^{n_2+1} \times \dots \times \mathcal{L}^{n_r+1}}$$

$$\begin{aligned} p^* = \sup\Big\{ & ((c_1,\gamma_1),\dots,(c_r,\gamma_r))^T((x_1,t_1),\dots,(x_r,t_r)) : \\ & ((x_1,t_1),\dots,(x_r,t_r)) \in \mathcal{L}^{n_1+1} \times \dots \times \mathcal{L}^{n_r+1} \\ & ((a_{j1},\alpha_{j1}),\dots,(a_{jr},\alpha_{jr}))^T((x_1,t_1),\dots,(x_r,t_r)) = b_j \\ & \text{for } j \in [m] \Big\} \end{aligned}$$

$$d^* = \inf \Big\{ b^T y : y \in \mathbb{R}^m,$$

$$(*) \left\{ \sum_{j=1}^{m} y_j \big((a_{j1}, \alpha_1), \dots, (a_{jr}, \alpha_r) \big) - \big((c_1, \gamma_1), \dots, (c_r, \gamma_r) \big) \in \mathcal{L}^{n_1+1} \times \dots \times \mathcal{L}^{n_r+1} \right\}$$

Better notation: Define matrices

$$A_i = \left[a_{1i} \dots a_{mi} \right] \in \mathbb{R}^{n_i \times m}, \quad i \in [r]$$

and vectors

$$d_i = (\alpha_{1i}, \dots, \alpha_{mi})^T, \quad i \in [r].$$

<u>Then</u>: $(*) \iff \| A_i y - c_i \| \leq d_i^T y - \gamma_i$ for $i \in [r]$.

<u>In particular</u>: LP is special case of CQP: set $A_i = 0$ and $c_i = 0$.

(iii) Semidefinite programming (SDP)

$$\boxed{K = S^n_+}$$

$$p^* = \sup\{ \langle C, X\rangle : X \in S^n_+,$$
$$\langle A_j, X\rangle = b_j,\ j \in [m]\}.$$

$$d^* = \inf\{ b^T y : y \in \mathbb{R}^m,\ \sum_{j=1}^m y_j A_j - C \in S^n_+\}$$

Inequalities of the form $\sum_{j=1}^m y_j A_j - C \succeq 0$ are called linear matrix inequalities (LMIs)

Remarks • LP is a special case of SDP: restrict X to diagonal matrix

• CQP is a special case of SDP:

Lemma 4 $\mathcal{L}^{n+1} = \{(x,t) \in \mathbb{R}^{n+1} : \|x\| \le t\}$

$$= \{(x,t) \in \mathbb{R}^{n+1} : \begin{bmatrix} tI_n & x \\ x & t \end{bmatrix} \succeq 0\}.$$

Proof If $t = 0$, then $x = 0$.

If $t > 0$, then $t I_n \succ 0$.

Consider the Schur complement of tI_n in $\begin{bmatrix} tI_n & x \\ x & t \end{bmatrix}$.

Then $\begin{bmatrix} tI_n & x \\ x & t \end{bmatrix} \succeq 0 \iff t - x^T \frac{1}{t} I_n x \geq 0$

$$\iff x^T x \leq t^2.$$

$$\iff \|x\| \leq t.$$

(iv) <u>Copositive programming</u>

$$p^* = \sup\{\langle C, X\rangle : X \in CP_n,\ \langle A_j, X\rangle = b_j,\ j \in [m]\}$$

$$d^* = \inf\{b^T y : y \in \mathbb{R}^m,\ \sum_{j=1}^m y_j A_j - C \in COP_n\}.$$

<u>Interesting fact</u> ($\rightarrow$ Sheet 3): One can model NP-hard problems as copositive programs; so convex optimisation is not necessarily computationally easy.

back to page 14 : $S^n_{++} = \operatorname{int} S^n_+$.

$S^n_{++} \subseteq \operatorname{int} S^n_+$:

To show: For $A \in S^n_{++}$ there is $\varepsilon > 0$ so that for all $B \in S^n$ with $\|A - B\| < \varepsilon$ we have $B \in S^n_+$.

Set $C = A - B$. For $x \in \mathbb{R}^n$ with $\|x\| = 1$ we have

$$x^T B x = x^T (A - C) x = x^T A x - x^T C x.$$

Since the unit sphere ($\|x\| = 1$) is compact and $A \in S^n_{++}$:

$$\delta = \min \{ x^T C x : \|x\| = 1 \} > 0,$$

[In fact δ is smallest eigenvalue of A]

and by Cauchy-Schwarz

$$|x^T C x| = |\langle C, x x^T \rangle| = \|C\| \, \|x x^T\| < \varepsilon \cdot 1.$$

Hence

$$x^T B x \geq \delta - \varepsilon \geq 0 \quad \text{if we choose } \varepsilon \in [0, \delta].$$

Then for arbitrary $x \in \mathbb{R}^n$:

$$x^T B x = \underbrace{\|x\|^2}_{\geq 0} \underbrace{\frac{x^T}{\|x\|} B \frac{x}{\|x\|}}_{\geq 0}.$$

$\operatorname{int} S^n_+ \subseteq S^n_{++}$:

To show: $S^n \setminus \operatorname{int} S^n_+ \supseteq S^n \setminus S^n_{++}$.

For $A \in S^n \setminus S^n_{++}$ spectral decomposition $A = \sum_{i=1}^n \lambda_i u_i u_i^T$ gives a $j \in [n]$ so that $\lambda_j \leq 0$. For every $\varepsilon > 0$ the matrix $B = -\varepsilon u_j u_j^T$ has norm $\|B\| = \varepsilon$, and $A - B \notin S^n_+$. Hence, $A \notin \operatorname{int} S^n_{++}$.

---

§3 Theorem of alternatives

Recall: Farkas Lemma (OR)

Lemma 1 For $A \in \mathbb{R}^{m \times n}$, $b \in \mathbb{R}^m$ exactly one of the following two alternatives holds:

(i) $\exists x \geq 0 : Ax = b$

(ii) $\exists y \in \mathbb{R}^m : A^T y \geq 0$ and $b^T y < 0$.

Goal Generalise to cones: Replace "≥ 0" by "$\geq_K 0$".

__First try__ Let $K \subseteq \mathbb{R}^n$ be a proper convex cone. For $A \in \mathbb{R}^{m \times n}$, $b \in \mathbb{R}^m$ exactly one of the following two alternatives holds:

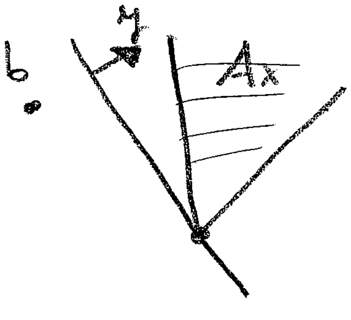
(i) $\exists\, x \in K : Ax = b$

(ii) $\exists\, y \in \mathbb{R}^m : A^T y \in K^*$ and $b^T y < 0$.

__"Proof"__ __(i) $\Rightarrow \neg$(ii)__ Suppose x is feasible for (i) and y is feasible for (ii). Then,

$$0 \le (A^T y)^T x = y^T A x = y^T b < 0.$$

__$\neg$(i) $\Rightarrow$ (ii)__ Suppose (i) has no solution. Then $b \notin \{Ax : x \in K\}$ and $\{Ax : x \in K\}$ is a convex cone. By the separation theorem there is a vector $y \in \mathbb{R}^m$ with $y^T b < 0$ and $y^T A x \ge 0$ for all $x \in K$. Hence $(y^T A)^T = A^T y \in K^*$ and (ii) follows.



Example 2 $K = S_+^2$

(i) $\langle E_{11}, X \rangle = 0, \ \langle E_{12}, X \rangle = 1, \quad X \in S_+^2$

(ii) $y_1 E_{11} + y_2 E_{12} \in S_+^2, \quad y_2 < 0$

Neither (i) nor (ii) has a feasible solution!

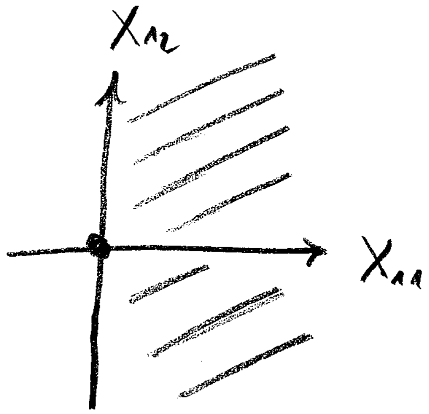
What? Why?

The convex cone

$$A S_+^2 = \left\{ \begin{bmatrix} \langle E_{11}, X \rangle \\ \langle E_{12}, X \rangle \end{bmatrix} \in \mathbb{R}^2 : X \in S_+^2 \right\}$$

$$= \left\{ \begin{bmatrix} X_{11} \\ X_{12} \end{bmatrix} : X_{11} > 0, X_{12} \in \mathbb{R} \right\} \cup \left\{ \begin{bmatrix} 0 \\ 0 \end{bmatrix} \right\}$$

is <u>not closed</u>



<u>Def. 3</u> For $A \in \mathbb{R}^{m \times n}$, $b \in \mathbb{R}^m$ the system

$$x \in K, \ Ax = b$$

is called <u>weakly feasible</u> if for all

$\varepsilon > 0$ there exists $x \in K$ with $\|Ax - b\| \leq \varepsilon$.

In other words, there is a sequence $(x_i)_{i \in \mathbb{N}}$ with $x_i \in K$ and $\lim_{i \to \infty} A x_i = b$.

(weakly feasible $\hat{=}$ <u>limit feasible</u>)

<u>Theorem 4</u> Let $K \subseteq \mathbb{R}^n$ be a proper convex cone. For $A \in \mathbb{R}^{m \times n}$, $b \in \mathbb{R}^m$ exactly one of the following two alternatives holds:

(i) The system $x \in K$, $Ax = b$ is weakly feasible

(ii) $\exists y \in \mathbb{R}^m : A^T y \in K^*, \; b^T y < 0$.

<u>Proof</u> Same argument as above but use the topological closure $\overline{\{Ax : x \in K\}}$. $\boxtimes$

In Example 2, system (i)

$$X \in S_+^2, \quad \langle E_{11}, X \rangle = 0, \quad \langle E_{12}, X \rangle = 1$$

is weakly feasible: Choose sequence

$$X_i = \begin{bmatrix} \frac{1}{i} & 1 \\ 1 & i \end{bmatrix}, \quad i \in \mathbb{N}.$$

Def 5 For $A \in \mathbb{R}^{m \times n}$, $c \in \mathbb{R}^n$ the system

$$A^T y - c \in K^*$$

is called weakly feasible if for all $\varepsilon > 0$ there exists $y \in \mathbb{R}^m$ and $z \in K$ s.t. $\| A^T y - c - z \| \leq \varepsilon$.

Theorem 4'

Let $K \subseteq \mathbb{R}^n$ be a proper convex cone. For $A \in \mathbb{R}^{m \times n}$, $c \in \mathbb{R}^n$ exactly one of the following two alternatives holds:

(i) $\exists x \in K$: $Ax = 0$ and $c^T x > 0$

(ii) The system $A^T y - c \in K^*$ is weakly feasible.

Proof → Exercise.

Theorem 6 Suppose that the linear system $Ax = b$ has a solution x_0. Then, exactly one of the following two alternatives holds:

(i) $\exists x \in \operatorname{int} K$: $Ax = b$. (x is strictly feasible)

(ii) $\exists y \in \mathbb{R}^m$: $A^T y \in K^* \setminus \{0\}$ and $b^T y \leq 0$

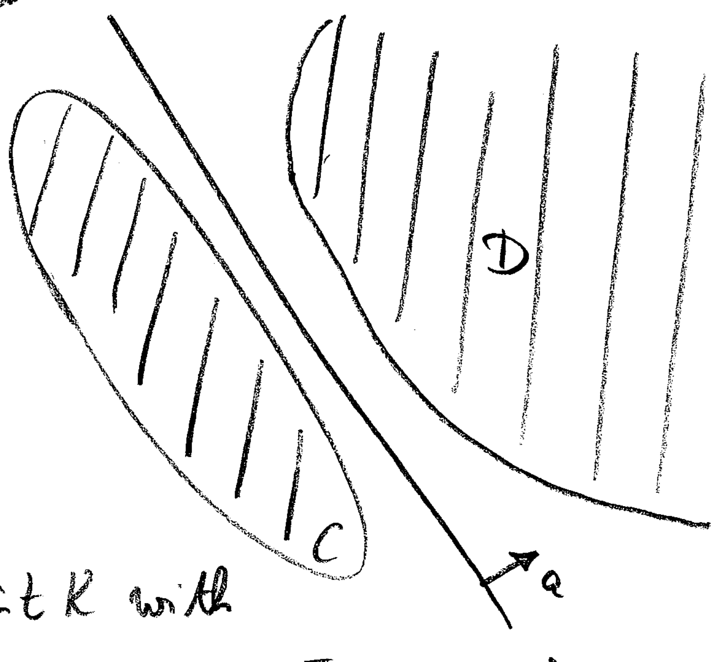
In Example 2 ($\langle E_{11}, X\rangle = 0$, $\langle E_{12}, X\rangle = 1$, $X \in S^2_+$)
condition (ii) is satisfied:

$$y_1 E_{11} + y_2 E_{12} \succeq 0 \qquad \text{with} \quad y_1 = 1,\ y_2 = 0.$$
$$y_2 \le 0$$

For the proof of Thm. 7 recall the following version of the separation theorem (from OR).

<u>Theorem 7</u> Let $C, D \subseteq \mathbb{R}^n$ be nonempty convex sets which do not intersect. Then there exist $a \in \mathbb{R}^n \setminus \{0\}$ so that

$$\sup_{x\in C} a^T x \le \inf_{x \in D} a^T x$$



<u>Proof</u> (of Thm. 6)

<u>(i) $\Rightarrow \neg$(ii)</u>: Suppose $\exists x \in \operatorname{int} K$ with $Ax = b$ and $\exists y \in \mathbb{R}^m$: $A^T y \in K^* \setminus \{0\}$ with $b^T y \le 0$. Then,

$$0 \le (A^T y)^T x = y^T A x = y^T b \le 0.$$

Hence, $(A^T y)^T x = 0$. Contradiction to Ex. 2.1.
($x \in \operatorname{int} K \Leftrightarrow \forall y \in K^* \setminus \{0\}: x^T y > 0$).

$\neg(i) \Rightarrow (ii)$: Suppose $Ax = b$ does not have a solution $x \in \operatorname{int} K$. Consider the linear subspace $L = \{x \in \mathbb{R}^n : Ax = 0\}$. We have

$$x_0 + L \cap \operatorname{int} K = \emptyset.$$

By the separation theorem there is a vector $a \in \mathbb{R}^n \setminus \{0\}$ and $\alpha \in \mathbb{R}$ with

$$a^T x \geq \alpha \quad \forall x \in K \quad \text{and} \quad a^T x \leq \alpha \quad \forall x \in x_0 + L.$$

- Since $0 \in K$, we have $0 \geq \alpha$.
- Furthermore, $a \in K^*$ because: For suppose not. Then $\exists x' \in K$ s.t. $a^T x' < 0$. For large enough $t > 0$: $a^T(tx') < \alpha$. $\lightning$
- Also, $a \in L^\perp$ because: For suppose not. Then $\exists x' \in L$ s.t. $a^T x' \neq 0$. We may assume that $a^T x' > 0$. For large enough $t > 0$: $a^T(x_0 + tx') > \alpha$. $\lightning$ So there exists $y \in \mathbb{R}^m$ with $a = A^T y$ because $L^\perp = \{A^T y : y \in \mathbb{R}^m\}$

Together: $\exists y \in \mathbb{R}^m$: $A^T y \in K^* \setminus \{0\}$ and

$$y^T b = y^T (A x_0) = a^T x_0 \leq \alpha \leq 0. \qquad \boxtimes$$

<u>Theorem 6'</u> Let $K \subseteq \mathbb{R}^n$ be a proper convex cone, let $A \in \mathbb{R}^{m \times n}$, $c \in \mathbb{R}^n$. Then, exactly one of the following two alternatives holds:

(i) $\exists\, x \in K \setminus \{0\} : Ax = 0$ and $c^T x \geq 0$

(ii) $\exists\, y \in \mathbb{R}^m : A^T y - c \in \operatorname{int} K^*$.

<u>Proof</u> $\rightarrow$ Exercise.

§4 Duality theory

(P) $p^* = \sup\{c^T x : x \in K,\ Ax = b\}$

(D) $d^* = \inf\{b^T y : y \in \mathbb{R}^m,\ A^T y - c \in K^*\}$.

<u>Def. 1</u> The difference $d^* - p^*$ (which turns out to be ≥ 0) is called <u>duality gap</u>.

Examples

<u>Example 2</u> d^* not attained.

$p^* = \sup\{\langle E_{12}, X\rangle : X \in S^2_+,\ \langle X, E_{11}\rangle = 1,\ \langle X, E_{22}\rangle = 0\}$

$$d^* = \inf \{ y_1 : \underbrace{y_1 E_{11} + y_2 E_{22} - E_{12}}_{\begin{bmatrix} y_1 & -\frac{1}{2} \\ -\frac{1}{2} & y_2 \end{bmatrix} \succeq 0} \in S_+^2 \}$$

We have $p^* = d^* = 0$, but d^* is not attained.

Problem: (P) is not strictly feasible.

Example 3 positive duality gap.

$$p^* = \sup \{ \langle -E_{11} - E_{22}, X \rangle : X \in S_+^3, \langle E_{11}, X \rangle = 0, \langle E_{22} + 2E_{13}, X \rangle = 1 \}$$

$$d^* = \inf \{ y_2 : \underbrace{y_1 E_{11} + y_2 (E_{22} + 2E_{13}) + E_{11} + E_{22}}_{\begin{bmatrix} y_1+1 & 0 & y_2 \\ 0 & y_2+1 & 0 \\ y_2 & 0 & 0 \end{bmatrix} \succeq 0} \in S_+^3 \}$$

Every feasible solution of primal satisfies

$$X_{11} = 0 = X_{13}, \ X_{22} = 1 \implies p^* = -1$$

Every feasible solution of dual satisfies

$$y_2 = 0 \implies d^* = 0.$$

Theorem 4 Let $K \subseteq \mathbb{R}^n$ be a proper convex cone, and $A \in \mathbb{R}^{m \times n}$, $b \in \mathbb{R}^m$, $c \in \mathbb{R}^n$. Consider the primal-dual pair of conic programs

$$(P) \quad p^* = \sup\{c^T x : x \in K,\ Ax = b\}$$
$$(D) \quad d^* = \inf\{b^T y : y \in \mathbb{R}^m,\ A^T y - c \in K^*\}.$$

(a) weak duality

Suppose x is feasible for (P) and y is feasible for (D). Then $c^T x \leq b^T y$. In particular, $p^* \leq d^*$, so that the duality gap is always ≥ 0.

(b) complementary slackness

Suppose $p^* = d^*$ and suppose x is optimal for (P) and y is optimal for (D). Then,

$$(A^T y - c)^T x = 0.$$

(c) optimality criterion

Suppose x is feasible for (P) and y is feasible for (D). Then, x and y are both optimal and $p^* = d^*$ if and only if $(A^T y - c)^T x = 0$.

(d) <u>strong duality</u>

Suppose $d^* > -\infty$ and (D) is strictly feasible. Then there is an optimal solution of (P), i.e. $\exists x^*$ feasible for (P) with $p^* = c^T x^*$. Furthermore, $p^* = d^*$.

<u>"Dually"</u>: Suppose $p^* < \infty$ and (P) is strictly feasible. Then $\exists y^*$ feasible for (D) with $d^* = b^T y^*$. Furthermore, $p^* = d^*$.

<u>Remarks</u>
- (c) strongly related to the <u>KKT-condition</u> which are famous in "classical" convex optimization / analysis (KKT = Karush-Kuhn-Tucker)
- (d): If (P) and (D) are both strictly feasible [then $p^* = d^*$ and (P), (D) attain optimal solution], then we say that <u>Slater's condition</u> is fulfilled.

Proof

(a) – (c): follow from

$$0 \le (A^T y - c)^T x = y^T A x - c^T x = y^T b - c^T x .$$

(d): Suppose $d^* > -\infty$ and (D) is strictly feasible.

to show: $\exists x^* \in K : Ax^* = b$ and $c^T x^* \ge d^* (\ge p^*)$

1st case $b = 0$

Then $d^* = 0$. Set $x^* = 0$.

2nd case $b \neq 0$

Consider

$$M = \{ A^T y - c : y \in \mathbb{R}^m, \ b^T y \le d^* \}.$$

Since $d^* > -\infty$, this is a nonempty polyhedron. (in particular convex)

We have

$$M \cap \operatorname{int} K^* = \emptyset .$$

Because: For suppose not. Then there would exist $y \in \mathbb{R}^m$ with $b^T y \leq d^*$ and $A^T y - c \in \operatorname{int} K^*$.
Consider $y' = y - \varepsilon \frac{b}{\|b\|}$ for $\varepsilon > 0$. Then

$$b^T y' = b^T \left(y - \varepsilon \frac{b}{\|b\|} \right) \leq d^* - \varepsilon$$

and

$$A^T y' - c \in K^*$$

for small enough ε since A^T is continuous.
Contradiction to the definition of d^*.

Now by the separation theorem there exist $a \in \mathbb{R}^n \setminus \{0\}$ such that

$$\sup_{z \in M} a^T z \leq \inf_{z \in \operatorname{int} K^*} a^T z = \inf_{z \in K^*} a^T z. \qquad (*)$$

Claims 1) $a \in K$

2) $\exists \mu > 0$: $Aa = \mu b$ and $c^T a \geq \mu d^*$

3) $x^* = \frac{a}{\mu}$ is optimal for (P).

Proofs 3) Follows from 1) & 2) + weak duality.

1) <u>We show:</u> $\inf_{z \in K^*} a^T z \geq 0$.

Then $a \in (K^*)^* = K$ because K is proper.

Suppose $\exists z \in K^*$: $a^T z < 0$. Then

$a^T(tz) \to -\infty$ for $t \to +\infty$.

By $(*)$: $\sup_{z \in M} a^T z = -\infty \Rightarrow M = \emptyset$ $\lightning$.

2) Claim 1 implies $\inf_{z \in K^*} a^T z = 0$ because $0 \in K^*$.

Hence, $\sup_{z \in M} a^T z \leq 0$. So, for $y \in \mathbb{R}^m$:

$$b^T y \leq d^* \Rightarrow a^T(A^T y - c) \leq 0 \qquad (**)$$

So, the halfspace $\{y : b^T y \leq d^*\}$ is contained in $\{y : (Aa)^T y \leq a^T c\}$ $\Rightarrow$ b and Aa have to be linearly dependent. If $Aa \neq 0$, then $\{y : (Aa)^T y \leq a^T c\}$ is a halfspace and

$\exists \mu > 0$: $Aa = \mu b$ and $\mu d^* \leq a^T c$

To show $\mu > 0$, For suppose not; $\mu = 0$.
i.e. $Aa = 0$.

Since (D) is strictly feasible $\exists y' : A^T y' - c \in \operatorname{int} \mathbb{R}^n_+$.

By Ex. 2.1

$$0 < (A^T y' - c)^T a = (y')^T A a - c^T a = (y')^T \mu b - c^T a$$

$$\overset{\mu = 0}{=} -c^T a$$

Thus, $c^T a < 0$, but from (xx) it follows

$$a^T(A^T y - c) \le 0 \iff 0 \le c^T a \qquad \lightning.$$

$$\boxtimes$$

The second statement of (d) follows from the first by taking the dual.