Part II Sets, spaces and matrices

Wenxun Xing

Department of Mathematical Sciences
Tsinghua University
Tel. 62787945
Office hour: 4:00-5:00 pm, Thursday
Email: wxing@tsinghua.edu.cn

Sept., 2019

Preliminaries

Content

- Vectors, linear spaces and matrices
- Semi-definite positive matrices
- Convex sets and cones
- Dual sets
- Linear Systems

Sets, spaces and matrices

- Real numbers: $\mathbb{R}$, $\mathbb{R}_+$, $\mathbb{R}_{++}$
- Euclidean space: $\mathbb{R}^n$
- First orthant: $\mathbb{R}^n_+$
- n-dimensional (column) vector:

$$x = (x_1, x_2, \ldots, x_n)^T$$

- Matrices space: $\mathbb{R}^{m\times n}$
- Matrix: $M \in \mathbb{R}^{m\times n}$, ith row $M_{i\bullet}$, jth column $M_{\bullet j}$, ijth entry M_{ij}
- Symmetric square matrices space ($n(n+1)/2$-dimensional space):

$$\mathcal{S}^n = \{M \in \mathbb{R}^{n\times n} \mid M = M^T\}.$$

Vectors, spaces and matrices

Given $M \in \mathbb{R}^{m\times n}$, $N \in \mathbb{R}^{n\times m}$, $S \in \mathbb{R}^{n\times n}$

- Determinant: $\det(S)$
- Trace: $\mathrm{tr}(S)$=$\sum_{i=1}^{n} s_{ii}$

$$\mathrm{tr}(MN) = \mathrm{tr}(NM)$$

- Null space: $\mathcal{N}(M)$= $\{x \in \mathbb{R}^n | Mx = 0\}$.
- Range space: $\mathcal{R}(M)$= $\{y \in \mathbb{R}^m | y = Mx \text{ for some } x \in \mathbb{R}^n\}$.
- Positive semidefinite matrix:

$$S \succeq 0 \iff z^T S z \geq 0,\ \forall\ z \in \mathbb{R}^n$$

- Positive definite matrix:

$$S \succ 0 \iff z^T S z > 0,\ \forall\ z \in \mathbb{R}^n,\ z \neq 0$$

Linear Systems

Given $x^1, \cdots, x^m \in \mathbb{R}^n$

- Linear combination:
$$\sum_{i=1}^{m} \lambda_i x^i,$$
where $\lambda_i \in \mathbb{R}$, $i = 1, \ldots, m$.
- Linearly independent
$$\sum_{i=1}^{m} \lambda_i x^i = 0 \Rightarrow \lambda_1 = \cdots = \lambda_m = 0$$
- Affine combination: a linear combination with
$$\sum_{i=1}^{m} \lambda_i = 1$$
- Affinely independent: if $x^2 - x^1, \cdots, x^m - x^1$ are linearly independent.

Linear Systems

- Convex combination: a linear combination with

$$\sum_{i=1}^{m} \lambda_i = 1 \text{ and } \lambda_i \geq 0, i = 1, \ldots, m$$

- Hyperplane:

$$\mathcal{X} = \{x \in \mathbb{R}^n | a^T x = \sum_{i=1}^{n} a_i x_i = b\}$$

- Affine space: affine combination of any two points in the space is still in the space. (An intersection of finitely many hyperplanes.)
- Linear subspace: an affine space containing the origin.

  We can always transform an affine space $\mathcal{Y} \subset \mathbb{R}^n$ into a linear subspace $\mathcal{X} \subset \mathbb{R}^n$ by choosing $x^0 \in \mathcal{Y}$ such that

$$\mathcal{X} = \{x - x^0 | x \in \mathcal{Y}\}$$

Linear Systems

- Half space:

$$\mathcal{X} = \{x \in \mathbb{R}^n | a^T x = \sum_{i=1}^{n} a_i x_i \leq b\}$$

- Polyhedron: an intersection of finitely many half spaces.
- Dimension of a linear subspace: the maximum number of linearly independent vectors in the subspace.
- Dimension of an affine space: the dimension of the transformed linear subspace.
- Dimension of a polyhedron: the dimension of the smallest affine space containing it.

Linear Systems

- Linear equations

$$\left.\begin{array}{ccc} a^1 \bullet x & = & b_1 \\ a^2 \bullet x & = & b_2 \\ \dots & \dots & \dots \\ a^m \bullet x & = & b_m \end{array}\right. \Rightarrow Ax = b,$$

where $a^1, \cdots, a^m$ and x are all in $\mathbb{R}^n$.

$$\left.\begin{array}{ccc} A_1 \bullet X & = & b_1 \\ A_2 \bullet X & = & b_2 \\ \dots & \dots & \dots \\ A_m \bullet X & = & b_m \end{array}\right. \Rightarrow \mathcal{A}X = b,$$

where $A_1, \cdots, A_m$ and X are all in $\mathcal{S}^n$.

- For convenience, $\mathcal{A}^* y = \sum_{i=1}^m y_i A_i$.

Properties of Trace

- $\mathrm{tr}(A) = \mathrm{tr}(A^T)$, where $A \in \mathcal{S}^n$.
- $\mathrm{tr}(AB^T) = \mathrm{tr}(B^T A)$, where A and B are the same size.
- $\mathrm{tr}(A(\sum_{i=1}^k B_i)^T) = \sum_{i=1}^k \mathrm{tr}(AB_i^T)$, where A and B_i are the same size.
- $\mathrm{tr}(kAB^T) = k \cdot \mathrm{tr}(AB^T)$, where $k \in \mathbb{R}$, A and B are the same size.
- $\mathrm{tr}(A^T A) \geq 0$ and $\mathrm{tr}(A^T A) = 0$ if and only if $A = 0$.
- $\mathrm{tr}(Dxx^T) = x^T Dx$, where $D \in \mathcal{S}^n$ and $x \in \mathbb{R}^n$.

An inner product: $X \bullet Y = traceXY^T$, where $X, Y \in \mathcal{M}(m, n)$.

Let $X, Y_1, Y_2 \in \mathcal{M}(m, n)$, $k_1, k_2 \in \mathbb{R}$.

- Linearity. $X \bullet (k_1 Y_1 + k_2 Y_2) = k_1 X \bullet Y_1 + k_2 X \bullet Y_2$.
- Symmetry. $X \bullet Y = Y \bullet X$.
- Nonnegativity. $X \bullet X \geq 0$ and $X \bullet X = 0$ if and only if $X = 0$.

An Example: QCQP

Quadratically constrained quadratic programming problem

$$
\begin{array}{ll}
\min & \frac{1}{2}x^T Q_0 x + q_0^T x + c_0 \\
s.t. & \frac{1}{2}x^T Q_i x + q_i^T x + c_i \leq 0, i = 1, 2, \ldots, m \\
& x \in \mathbb{R}^n
\end{array}
$$

where $Q_i \in \mathcal{S}^n$, $q_i \in \mathbb{R}^n$, $c_i \in \mathbb{R}$, $i = 0, 1, 2, \ldots, m$ are given coefficients, x is a decision variable.

$$
\begin{array}{ll}
\min & \frac{1}{2}Q_0 \bullet X + q_0^T x + c_0 \\
s.t. & \frac{1}{2}Q_i \bullet X + q_i^T x + c_i \leq 0, i = 1, 2, \ldots, m \\
& X = xx^T \\
& x \in \mathbb{R}^n
\end{array}
$$

An Example: SDP Relaxation

Formulation 1

$$
\begin{array}{ll}
\min & \frac{1}{2}Q_0 \bullet X + q_0^T x + c_0 \\
s.t. & \frac{1}{2}Q_i \bullet X + q_i^T x + c_i \leq 0, i = 1, 2, \ldots, m \\
 & x \in \mathbb{R}^n, X \in \mathcal{S}_+^n.
\end{array}
$$

松弛！

Formulation 2

$$
\begin{array}{ll}
\min & \frac{1}{2}Q_0 \bullet X + q_0^T x + c_0 \\
s.t. & \frac{1}{2}Q_i \bullet X + q_i^T x + c_i \leq 0, i = 1, 2, \ldots, m \\
 & \begin{pmatrix} 1 & x^T \\ x & X \end{pmatrix} \in \mathcal{S}_+^{n+1}.
\end{array}
$$

Inner Products and Norms

- Inner products:

$$x \bullet y = x^T y = \sum_i x_i y_i$$

$$X \bullet Y = \mathrm{tr}(X^T Y) = \sum_{i,j} X_{ij} Y_{ij}$$

- Norms:
  - Euclidean norm: $\|x\|_2 = \sqrt{x \bullet x}$
  - p-norm: $\|x\|_p = (\sum_{i=1}^n |x_i|^p)^{1/p}$ for $p \geq 1$.
  - Infinity-norm: $\|x\|_\infty = \max\{|x_1|, \ldots, |x_n|\}$
  - Frobenius norm:

$$\|X\|_F = \sqrt{X \bullet X} = \sqrt{\mathrm{tr}(X^T X)}$$

- Note that: $x^T A x = A \bullet x x^T$

Properties of semi-definite positive matrices

Theorem

(i) Given $A \in \mathcal{S}^n$, there exists an orthogonal matrix Q, i.e., $Q^TQ = QQ^T = I$, such that

$$Q^TAQ = \mathrm{diag}(\lambda_1, \lambda_2, \ldots, \lambda_n),$$

$$\mathrm{tr}(A) = \sum_{i=1}^{n} \lambda_i, \quad \det(A) = \lambda_1\lambda_2 \cdots \lambda_n,$$

where $\lambda_1, \lambda_2, \ldots, \lambda_n$ are eigenvalues of A.
(ii) $A \in \mathcal{S}_+^n$ if and only if $\lambda_i \geq 0, i = 1, 2, \ldots, n$.
(iii) If $A = (a_{ij}) \in \mathcal{S}_+^n$, then $a_{ii} \geq 0, i = 1, 2, \ldots, n$.

Theorem: (Schur complementary theorem)

证明见讲义P32

$$A \succ 0, X = \begin{bmatrix} A & B \\ B^T & C \end{bmatrix}, S = C - B^T A^{-1} B$$

Then

$$X \succeq (\succ) 0 \Leftrightarrow S \succeq (\succ) 0$$

Examples: Nonlinear to linear representable equations

将非线性的转化为线性的

线性矩阵不等式

LMI

$$\sqrt{x_1^2 + x_2^2 + \cdots + x_{n-1}^2} \leq x_n \Leftrightarrow \begin{pmatrix} x_n & x_{1:n-1}^T \\ x_{1:n-1} & x_n I_{n-1} \end{pmatrix} \in \mathcal{S}_+^n.$$

$$X - X^T (I + X)^{-1} X \in \mathcal{S}_+^n, X \in \mathcal{S}_+^n \Leftrightarrow \begin{pmatrix} I + X & X \\ X^T & X \end{pmatrix} \in \mathcal{S}_+^{2n}.$$

Theorem

(Congruent diagonalization) Given $A \in \mathcal{S}^n_{++}$ *and* $B \in \mathcal{S}^n$*, there exists an insertable matrix* P *such that* $P^T AP$ *and* $P^T BP$ *are diagonal.*

Theorem

(Cholesky decomposition) Suppose $A \in \mathcal{S}^n_{++}$*, we have a lower triangular matrix with positive diagonal elements* L *such that* $A = LL^T$*.*

Corollary

(i) Denote the eigenvalues of $A \in \mathcal{S}^n$ *as* $\lambda_1, \lambda_2, \ldots, \lambda_n$*. We have* $\mathrm{tr}(A^T A) = \sum_{i=1}^n \lambda_i^2$*. When* $x \neq 0$*,*

$$\min_{1 \le i \le n} \{\lambda_i\} \le \frac{x^T Ax}{x^T x} \le \max_{1 \le i \le n} \{\lambda_i\}.$$

(ii) Given $A = (a_{ij}), B = (b_{ij}) \in \mathcal{S}^n$*, we have* $\sum_{i=1}^k a_{ii} \ge \sum_{i=1}^k b_{ii}$ *for any* $1 \le k \le n$ *when* $A - B \in \mathcal{S}^n_+$*.*

Corollary

Given $A \in \mathcal{M}(m,n)$ *and* $B \in \mathcal{M}(n,p)$, *we have*(i) $\|A\|_2 \le \|A\|_F$; (ii) *for any* $x \in \mathbb{R}^n$, *and* $\|Ax\|_2 \le \|A\|_2 \|x\|_2$; (iii) $\|AB\|_F \le \|A\|_2 \|B\|_F$.

Theorem

Suppose $A \in \mathcal{S}_+^n$ *and* $\operatorname{rank}(A) = r$. *There exist* $p^i \in \mathbb{R}^n,\ i = 1, 2, \ldots, r$, *such that* $A = \sum_{i=1}^r p^i (p^i)^T$.

Theorem

For a given $X \in \mathcal{S}_+^n$ *of rank* r *and any* $G \in \mathcal{S}^n$, $G \bullet X \ge 0$ *if and only if there exist* $p^i \in \mathbb{R}^n$, $i = 1, 2, \cdots, r$ *such that*

$$X = \sum_{i=1}^{r} p^i (p^i)^T \ \ \textit{and} \ \ (p^i)^T G p^i \ge 0.$$

In case of $G \bullet X = 0$, *there exist* $p^i \in \mathbb{R}^n$, $i = 1, 2, \cdots, r$ *such that*

$$X = \sum_{i=1}^{r} p^i (p^i)^T \ \ \textit{and} \ \ (p^i)^T G p^i = 0.$$

Proof of the decomposition theorem

$' \Leftarrow '$

$$G \bullet X = \mathrm{tr}(GX^T) = \sum_{i=1}^{r} G \bullet \left(p^i (p^i)^T\right) = \sum_{i=1}^{r} (p^i)^T G p^i \geq 0,$$

$' \Rightarrow '$

- Input: $X \in \mathcal{S}_+^n$ and G.
- Output: a vector y such that $0 \leq y^T G y \leq G \bullet X$, $X - yy^T \in \mathcal{S}_+^n$ and the rank of $X - yy^T$ is $r-1$.

Step 0 Calculate $p^1, p^2, \cdots, p^r$ such that $X = \sum_{i=1}^r p^i (p^i)^T$.

Step 1 If $[(p^1)^T G p^1][(p^i)^T G p^i] \geq 0$ for all $i = 2, 3, \cdots, r$, output $y = p^1$. Otherwise select one j such that $[(p^1)^T G p^1][(p^j)^T G p^j] < 0$.

Step 2 Calculate the α such that $(p^1 + \alpha p^j)^T G (p^1 + \alpha p^j) = 0$. out put $y = (p^1 + \alpha p^j)/\sqrt{1+\alpha^2}$.

Open, Closed, Interior and Boundary Sets

- Neighborhood: $N(x^0;\epsilon) = \{x \in \mathbb{R}^n \mid \|x - x^0\| < \epsilon\}$.
- Open: $\mathcal{X} \subset \mathbb{R}^n$ is open if for any $x \in \mathcal{X}$, there exists $\epsilon > 0$ such that $N(x;\epsilon) \subset \mathcal{X}$.
- Closed: $\mathcal{X} \subset \mathbb{R}^n$ is closed, if $\mathbb{R}^n \backslash \mathcal{X} = \{x \in \mathbb{R}^n | x \notin \mathcal{X}\}$ is open.
- Closed: An equivalent statement: any accumulation point of $\mathcal{X}$ is in $\mathcal{X}$.
- Closure of a set $\mathcal{X} \subset \mathbb{R}^n$ is the smallest closed set containing $\mathcal{X}$ and is denoted as $\mathrm{cl}(\mathcal{X})$.

Open, Closed, Interior and Boundary Sets

- Interior: the interior of a given set $\mathcal{X} \subset \mathbb{R}^n$ is

$$\mathrm{int}(\mathcal{X}) = \{x \in \mathcal{X} | \exists\ \epsilon_x > 0 \text{ such that } N(x; \epsilon_x) \subset \mathcal{X}\}$$

- Boundary of a set $\mathcal{X} \subset \mathbb{R}^n$:

$$\mathrm{bdry}(\mathcal{X}) = \mathrm{cl}(\mathcal{X}) \backslash \mathrm{int}(\mathcal{X}) = \{x \in \mathrm{cl}(\mathcal{X}) | x \notin \mathrm{int}(\mathcal{X})\}$$

- Bounded: a set $\mathcal{X} \subset \mathbb{R}^n$ is bounded if there exists an $r > 0$ such that

$$\|x\| < r, \forall x \in \mathcal{X}$$

An example: open, closed, boundary

(i) $\mathcal{S}^n_{++}$: open, (ii) $\mathcal{S}^n_{+}$: closed, (iii) $\mathrm{int}(\mathcal{S}^n_{+}) = \mathcal{S}^n_{++}$, (iv) $\mathrm{cl}(\mathcal{S}^n_{+}) = \mathrm{cl}(\mathcal{S}^n_{++}) = \mathcal{S}^n_{+}$, (v)

$$\mathrm{bdry}(\mathcal{S}^n_{+}) = \{A \in \mathcal{S}^n_{+} \mid \exists x \in \mathbb{R}^n \text{ and } x \neq 0 \text{ such that } x^T A x = 0\}.$$

Proof. (i) Let $\lambda_{max}(A)$ and $\lambda_{min}(A)$ be the maximal and minimal eigenvalue of A. For any $A \in \mathcal{S}^n_{++}$, we have $\lambda_{min}(A) > 0$. Let $\epsilon = \frac{\lambda_{min}(A)}{2}$. For any $B \in N(A;\epsilon) = \{B \in \mathcal{S}^n \mid \ \|B - A\| < \epsilon\}$, we know the absolute eigenvalue of $B - A$ is less than ϵ. For any $x \neq 0$,

$$x^T B x = x^T A x + x^T (B - A) x > (\lambda_{min}(A) - \epsilon) x^T x = \frac{\lambda_{min}}{2} x^T x > 0,$$

so $B \succ 0$, and then $\mathcal{S}^n_{++}$ is open.

(ii) We prove $\mathcal{S}^n \setminus \mathcal{S}^n_+$ is open. $\forall A \in \mathcal{S}^n \setminus \mathcal{S}^n_+$, we get $\lambda_{min}(A) < 0$. Let $\epsilon = \frac{|\lambda_{min}(A)|}{2}$. With almost the same arguments as the above, we get the result.

By the definition of the accumulation point, for any accumulation point B of $\mathcal{S}^n_+$, there exist $\{A_i \mid i = 1, 2, \dots\} \subseteq \mathcal{S}^n_+$ such that $A_i \to B, i \to +\infty$. Then $x \in \mathbb{R}^n$, $x^T A_i x \geq 0, \forall i \geq 1$, which imply

$$\lim_{i \to +\infty} x^T A_i x = x^T B x \geq 0.$$

Hence $B \in \mathcal{S}^n_+$ and $\mathcal{S}^n_+$ is closed.

(iii) Obviously, $\mathrm{int}(\mathcal{S}^n_+) \supseteq \mathcal{S}^n_{++}$. For $\forall A \in \mathrm{int}(\mathcal{S}^n_+)$, if A is not positive definite, suppose $A = Q\mathrm{diag}(\lambda_1, \lambda_2, \dots, \lambda_n)Q^T$, and $\lambda_1 \leq \lambda_2 \leq \dots \leq \lambda_n$. We have $\lambda_1 = 0$. Then for any $\epsilon > 0$, let

$$B = Q\mathrm{diag}(-\epsilon/2, \lambda_2, \dots, \lambda_n)Q^T.$$

We have $\|B - A\| = \epsilon/2 < \epsilon$ but $B \notin \mathcal{S}^n_+$.

Convex Sets and Properties

- A set $\mathcal{X} \subset \mathbb{R}^n$ is convex if for any $x^1 \in \mathcal{X}$ and $x^2 \in \mathcal{X}$, we have $\lambda x^1 + (1-\lambda)x^2 \in \mathcal{X}$, for all $0 \le \lambda \le 1$.
- Convex hull: the smallest convex set containing a given set

$$\begin{aligned}\mathrm{conv}(\mathcal{X}) = \{x \in \mathbb{R}^n | x = \textstyle\sum_{i=1}^m \lambda_i y^i \text{ for some } m \in \mathbb{N}_+, \\ \lambda_i \ge 0, \textstyle\sum_{i=1}^m \lambda_i = 1, \text{ and } y^i \in \mathcal{X}, i = 1, \ldots, m\}\end{aligned}$$

- Dimension of a convex set: the dimension of the smallest affine space containing it.
- Relative interior of a convex set $\mathcal{X} \subset \mathbb{R}^n$: suppose $\mathcal{H}$ is the smallest affine space containing $\mathcal{X}$,

$$\mathrm{ri}(\mathcal{X}) = \{x \in \mathbb{R}^n | \exists \text{ open set } \mathcal{Y} \subseteq \mathbb{R}^n \text{ such that } x \in \mathcal{Y} \cap \mathcal{H} \subset \mathcal{X}\}$$

- Supporting hyperplane $\mathcal{H} = \{x \in \mathbb{R}^n | a^T x = b\}$ of a convex set $\mathcal{X}$:

$$a^T y \ge b, \forall\, y \in \mathcal{X} \text{ and } \mathrm{cl}(\mathcal{X}) \cap \mathcal{H} \ne \emptyset.$$

Relative Interior–An Example

$$\mathcal{X} = \{x_1 \in \mathbb{R} | 0 \le x_1 \le 2\}.$$

A linear programming standard reformulation

$$\mathcal{Y} = \{(x_1, x_2) \in \mathbb{R}^2_+ | x_1 + x_2 = 2\}.$$

Relative interior

$$\mathrm{ri}(\mathcal{X}) = \mathrm{int}(\mathcal{X}) = \{x_1 \in \mathbb{R} | 0 < x_1 < 2\}.$$

$$\mathrm{ri}(\mathcal{Y}) = \{(x_1, x_2) \in \mathbb{R}^2 | x_1 + x_2 = 2, x_1 > 0, x_2 > 0\},$$

where the small affine space is $\mathcal{H} = \{(x_1, x_2) \in \mathbb{R}^2 | x_1 + x_2 = 2\}$, and the open set is defined as in $\mathbb{R}^2$.

Lemma

Suppose $\mathcal{X} \subseteq \mathbb{R}^n$ has an interior point. Then for any hyperplane $\mathcal{H} = \{x \in \mathbb{R}^n \mid a^T x = b\}$, there exists an $\bar{x} \in \mathcal{X}$ such that $a^T \bar{x} \neq b$.

Lemma

Suppose $\mathcal{X} \subseteq \mathbb{R}^n$ be nonempty, $\mathcal{A}$is the minimal affine space containing $\mathcal{X}$, and $\mathrm{ri}(\mathcal{X}) \neq \emptyset$. For any hyperplane $\mathcal{H} = \{x \in \mathbb{R}^n \mid a^T x = b\}$ such that $\dim(\mathcal{H} \cap \mathcal{A}) \leq \dim(\mathcal{A}) - 1$, there exists an $\bar{x} \in \mathcal{X}$ such that $a^T \bar{x} \neq b$.

Lemma

Suppose $\mathcal{X}_1, \mathcal{X}_2$ be convex. Then $\mathcal{X}_1 + \mathcal{X}_2$ and $\mathcal{X}_1 \times \mathcal{X}_2$ are convex.
Suppose $\mathcal{X}_i$ be convex for $i = 1, 2, \cdots$. Then $\bigcap_{i=1}^{\infty} \mathcal{X}_i$ is convex.
Suppose $\mathcal{X}_i$ be closed for $i = 1, 2, \cdots$. Then $\bigcap_{i=1}^{\infty} \mathcal{X}_i$ is closed.
Suppose $\mathcal{X}_i$ be open for $i = 1, 2, \cdots$. Then $\bigcup_{i=1}^{\infty} \mathcal{X}_i$ is open.

Lemma

Suppose $\mathcal{X} \subseteq \mathbb{R}^n$ be nonempty, convex and closed. For any $z \in \mathbb{R}^n$, there exists an unique $\bar{x} \in \mathcal{X}$ such that

$$\mathrm{dist}(z, \mathcal{X}) = \|z - \bar{x}\| = \min\left\{\|z - x\| \mid x \in \mathcal{X}\right\},$$

and

$$(z - \bar{x})^T (x - \bar{x}) \leq 0, \ \forall x \in \mathcal{X}.$$

Lemma

Suppose $\mathcal{X} \subseteq \mathbb{R}^n$ be nonempty and convex. For any $z \notin \mathrm{cl}(\mathcal{X})$, there exist $a \neq 0, a \in \mathbb{R}^n, b \in \mathbb{R}$ such that

$$a^T x \geq b > a^T z \textit{ for any } x \in \mathcal{X}.$$

Theorem

Suppose $\mathcal{X} \subseteq \mathbb{R}^n$ be a nonempty convex set. For any $y \in \mathrm{bdry}(\mathcal{X})$, there exist $a \in \mathbb{R}^n, b \in \mathbb{R}$ such that $a^T y = b$ and $a^T x \geq b$ for any $x \in \mathcal{X}$.

Theorem

Suppose two nonempty convex sets $\mathcal{X}_1 \cap \mathcal{X}_2 = \emptyset$. There exist $a \in \mathbb{R}^n, b \in \mathbb{R}$ such that the hyperplane $\{x \in \mathbb{R}^n \mid a^T x = b\}$ separate $\mathcal{X}_1$ and $\mathcal{X}_2$.

Theorem

Two nonempty sets $\mathcal{X}_1 \subseteq \mathbb{R}^n$ and $\mathcal{X}_2 \subseteq \mathbb{R}^n$ is properly separated by a hyperplane if and only if there exist an $a \in \mathbb{R}^n$ such that

$$\begin{array}{ll} \text{(i)} & \inf_{x \in \mathcal{X}_1} a^T x \geq \sup_{x \in \mathcal{X}_2} a^T x, \\ \text{(ii)} & \sup_{x \in \mathcal{X}_1} a^T x > \inf_{x \in \mathcal{X}_2} a^T x. \end{array}$$

Lemma

Suppose $\mathcal{X}$ be convex, $r = \dim(\mathcal{X}) \geq 1$, and $\{x^1, x^2 \ldots, x^{r+1}\} \subseteq \mathcal{X}$ be $r+1$ linearly independent affine points. For any $\lambda_i > 0, i = 1, 2, \ldots, r+1$ and $\sum_{i=1}^{r+1} \lambda_i = 1$, $y = \sum_{i=1}^{r+1} \lambda_i x^i$ is a relatively interior point of $\mathcal{X}$. Conversely, for any $y \in \mathrm{ri}(\mathcal{X})$, there exists $r+1$ linearly independent points $\{x^1, x^2 \ldots, x^{r+1}\} \subseteq \mathcal{X}$, $\lambda_i > 0, i = 1, 2, \ldots, r+1$ with $\sum_{i=1}^{r+1} \lambda_i = 1$, such that $y = \sum_{i=1}^{r+1} \lambda_i x^i$.

Corollary

Suppose $\mathcal{X} \subseteq \mathbb{R}^n$ *be a nonempty convex set. Then* $\mathrm{ri}(\mathcal{X}) \neq \emptyset$.

Lemma

Suppose $\mathcal{X}$ *be a nonempty convex set. For* $y \in \mathrm{cl}(\mathcal{X})$ *and* $z \in \mathrm{ri}(\mathcal{X})$*, the point* $x = \alpha y + (1-\alpha)z \in \mathrm{ri}(\mathcal{X}), \forall 0 \leq \alpha < 1$*.*

Theorem

Suppose $\mathcal{X}$ *be nonempty and convex. We have* $\mathrm{cl}(\mathrm{ri}(\mathcal{X})) = \mathrm{cl}(\mathcal{X})$ *and* $\mathrm{ri}(\mathrm{cl}(\mathcal{X})) = \mathrm{ri}(\mathcal{X})$*.*

Theorem

Suppose $\mathcal{X} \subseteq \mathbb{R}^n$ *be nonempty and convex,* A *be given* $m \times n$ *matrix. We have* $\mathrm{ri}(A\mathcal{X}) = A(\mathrm{ri}(\mathcal{X}))$*.*

Theorem

Suppose $\mathcal{X}_1$ *and* $\mathcal{X}_2$ *be nonempty and convex. We have* $\mathrm{ri}(\mathcal{X}_1 \times \mathcal{X}_2) = \mathrm{ri}(\mathcal{X}_1) \times \mathrm{ri}(\mathcal{X}_2)$ *and* $\mathrm{ri}(\mathcal{X}_1 + \mathcal{X}_2) = \mathrm{ri}(\mathcal{X}_1) + \mathrm{ri}(\mathcal{X}_2)$.

Theorem

Given a nonempty convex set $\mathcal{X}$, for any $x^0 \notin \mathrm{ri}(\mathcal{X})$, there exists a hyperplane such that $a^T x \geq b, \forall x \in \mathcal{X}$, $a^T x^0 = b$ and $a^T x > b, \forall x \in \mathrm{ri}(\mathcal{X})$.

Theorem

Given two nonempty convex sets $\mathcal{X}_1$ and $\mathcal{X}_2$, they are properly separated if and only if $\mathrm{ri}(\mathcal{X}_1) \cap \mathrm{ri}(\mathcal{X}_2) = \emptyset$.

Theorem

Suppose $\mathcal{C}$ and $\mathcal{D}$ be polytopes in $\mathbb{R}^n$. The following sets are polytopes

(i) $\mathcal{C} \cap \mathcal{D}$,

(ii) $\mathcal{C} \times \mathcal{D} = \left\{ \begin{pmatrix} x \\ y \end{pmatrix} \in \mathbb{R}^{2n} \mid x \in \mathcal{C}, y \in \mathcal{D} \right\}$,

(iii) $\mathcal{C} + \mathcal{D}$.

Cones and Properties

- A set $K \subset \mathbb{R}^n$ is a cone if

$$\forall x \in K \text{ and } \lambda > 0 \Rightarrow \lambda x \in K;$$

- A cone $K \subset \mathbb{R}^n$ is pointed if

$$K \cap -K = \{0\};$$

- A cone $K \subset \mathbb{R}^n$ is solid if

$$\text{int} K \neq \emptyset;$$

- A cone $K \subset \mathbb{R}^n$ is proper if it is pointed, solid, closed and convex.

Dual Cones

- Conic combination: a linear combination $\sum_{i=1}^m \lambda_i x^i$ with $\lambda_i \geq 0$, $x^i \in \mathbb{R}^n$ for all $i = 1, \ldots, m$.
- The conic hull of a set $\mathcal{X} \subset \mathbb{R}^n$ is

$$\text{cone}(\mathcal{X}) = \{x \in \mathbb{R}^n | x = \sum_{i=1}^m \lambda_i x^i, \text{ for some } m \in \mathbb{N}_+ \\ \text{and } x^i \in \mathcal{X}, \lambda_i \geq 0, i = 1, \ldots, m.\}$$

- The dual cone $K^* \subset \mathbb{R}^n$ of a cone $K \subset \mathbb{R}^n$ is

$$K^* = \{y \in \mathbb{R}^n | y \bullet x \geq 0, \forall\, x \in K\}$$

  K^* is a *closed, convex* cone.
- If $K^* = K$, then K is a self-dual cone.

Properties

Theorem

If $\mathcal{K}_1, \mathcal{K}_2, \ldots, \mathcal{K}_m$ are simultaneously pointed (solid, closed or convex) cones, then their Cartesian product is a cone and keeps the same property of pointed (solid, closed or convex). Their intersection is a cone and keeps the property of pointed (closed or convex).

Theorem

Given $\mathcal{X}_1$ and $\mathcal{X}_2$ in $\mathbb{R}^n$, $\mathrm{(i)}$ If $\mathcal{X}_1 \subseteq \mathcal{X}_2$, then $\mathcal{X}_1^* \supseteq \mathcal{X}_2^*$; $\mathrm{(ii)}$ If $0 \in \mathcal{X}_1 \cap \mathcal{X}_2$, then $(\mathcal{X}_1 + \mathcal{X}_2)^* = \mathcal{X}_1^* \cap \mathcal{X}_2^*$.

Theorem

Suppose $\mathcal{X}$ be a closed, pointed and convex cone containing at least one nonzero point, and $\operatorname{int}(\mathcal{X}^*) \neq \emptyset$. Then $y \in \operatorname{int}(\mathcal{X}^*)$ if and only if $y^T x > 0$ for any $x \in \mathcal{X}$ and $x \neq 0$.

Theorem

(i) *Given* $\mathcal{K}_1$ *and* $\mathcal{K}_2$ *are two convex cones, then* $\mathcal{K}_1 \cap \mathcal{K}_2$ *and* $\mathcal{K}_1 + \mathcal{K}_2$ *are convex cones.*
(ii) *Given* $\mathcal{K}_1$ *and* $\mathcal{K}_2$ *are solid cones, then* $\mathcal{K}_1 + \mathcal{K}_2$ *is a solid cone.*

Theorem

Given a nonempty set $\mathcal{X} \subseteq \mathbb{R}^n$*, we have the follows.*
(i) $\mathcal{X}^*$ *is a closed convex cone,*
(ii) $\mathcal{X} \subseteq (\mathcal{X}^*)^*$*,*
(iii) $(\mathcal{X}^*)^* = \mathcal{X}$ *if* $\mathcal{X}$ *is a closed convex cone,*
(iv) $\mathcal{X}^*$ *is a pointed cone if* $\operatorname{int}(\mathcal{X}) \neq \emptyset$*,*
(v) $\operatorname{int}(\mathcal{X}^*) \neq \emptyset$ *if* $\mathcal{X}$ *is a closed, convex and pointed cone.*

Partial Order and Ordered Vector Space

- A relation "$\geq$" is a partial order on a set $\mathcal{X}$ if it has:

  1. *reflexivity*: $a \geq a$ for all $a \in \mathcal{X}$;
  2. *antisymmetry*: $a \geq b$ and $b \geq a$ imply $a = b$;
  3. *transitivity*: $a \geq b$ and $b \geq c$ imply $a \geq c$.

- An ordered vector space $\mathcal{X}$ is equipped with a partial order "$\geq$" which also satisfies:

  - *homogeneity*: $a \geq b$ and $\lambda \in \mathbb{R}_+$ imply $\lambda a \geq \lambda b$;
  - *additivity*: $a \geq b$ and $c \geq d$ imply $a + c \geq b + d$.

Partial Order and Ordered Vector Space

- A *proper* cone K in a vector space can induce a partial order "$\geq_K$"

$$a \geq_K b \Leftrightarrow a - b \in K$$

  which leads to an ordered vector space.
- Similarly, we can define "$\leq_K$"

$$a \leq_K b \Leftrightarrow b \geq_K a,$$

- *Closeness* of K allows passing limits in $\geq_K$:

$$a^i \geq_K b^i,\ a^i \to a,\ b^i \to b \text{ as } i \to \infty \ \Rightarrow \ a \geq_K b.$$

- *Solidness* of K allows us to define a strict inequality:

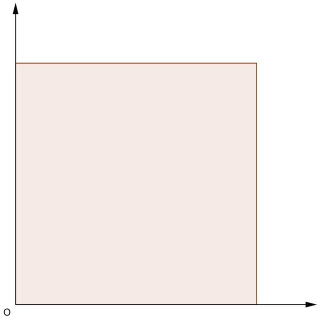
$$a >_K b \Leftrightarrow a - b \in \text{int}K,$$

  and

$$a <_K b \Leftrightarrow b >_K a.$$

Examples: $\mathbb{R}^n_+$

- $\mathbb{R}^n_+$ is a proper cone;
- Inner product: $x \bullet y = x^T y$;
- $(\mathbb{R}^n_+)^* = \mathbb{R}^n_+$ (self-dual);
- Partial order: "$\geq_{\mathbb{R}^n_+}$"

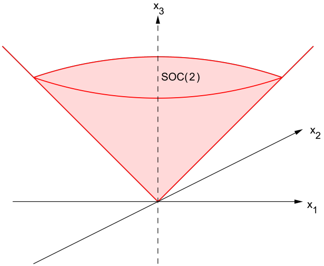


Examples: $\mathcal{L}^n$

- $\mathcal{L}^n$ / $SOC(n-1)$Lorentz cone (second order cone)

$$\mathcal{L}^n = \{x \in \mathbb{R}^n | x_n \geq \sqrt{x_1^2 + \cdots + x_{n-1}^2}\}$$

- $\mathcal{L}^n$ is a proper cone;
- Inner product: $x \bullet y = x^T y$;
- $(\mathcal{L}^n)^* = \mathcal{L}^n$ (self-dual);
- Partial order: "$\geq_{\mathcal{L}^n}$"



Examples: $\mathcal{S}_+^n$

- $\mathcal{S}_+^n \subset \mathcal{S}^n$: the set of symmetric positive semidefinite matrices
- $\mathcal{S}_+^n$ is a proper cone;
- Inner product:

$$X \bullet Y = \mathrm{tr}(X^T Y)$$

- *Another view*:

$$\mathrm{vec}(X) = [X_{11}, \sqrt{2}X_{12}, X_{22}, \sqrt{2}X_{13}, \sqrt{2}X_{23}, X_{33}, \cdots, X_{nn}]^T \in \mathbb{R}^{\frac{n(n+1)}{2}}$$

Then

$$X \bullet Y = \mathrm{vec}(X) \bullet \mathrm{vec}(Y) = \sum_{i,j} X_{ij} Y_{ij}$$

- Partial order: "$\geq_{\mathcal{S}_+^n}$" or "$\succeq$"

Examples: $\mathcal{S}_+^n$

Lemma

$(\mathcal{S}_+^n)^* = \mathcal{S}_+^n$ (self-dual)

Proof.

"$\subseteq$": If $X \in (\mathcal{S}_+^n)^*$, then $z^T X z = X \bullet zz^T \geq 0$, for all $z \in \mathbb{R}^n$.
Therefore, $X \in \mathcal{S}_+^n$.

"$\supseteq$": For any $Y \in \mathcal{S}_+^n$,

$$Y = \sum_{i=1}^n \lambda_i z^i (z^i)^T,$$

with $\lambda_i \geq 0$.
If $X \in \mathcal{S}_+^n$, then

$$X \bullet Y = \sum_{i=1}^n \lambda_i X \bullet z^i (z^i)^T = \sum_{i=1}^n \lambda_i (z^i)^T X z^i \geq 0.$$

Therefore, $X \in (\mathcal{S}_+^n)^*$.

Examples: $\mathcal{C}_n$ and $\mathcal{C}_n^*$

- Copositive cone:

$$\mathcal{C}_n = \{X \in \mathcal{S}^n | z^T X z \geq 0, \forall\ z \geq_{\mathbb{R}^n_+} 0\}$$

- Completely positive(nonnegative) cone:

$$\mathcal{C}_n^* = \left\{ X \in \mathcal{S}^n \middle| \begin{array}{l} X = \sum_{i=1}^m z^i (z^i)^T, \text{ for some } m \in \mathbb{N}_+ \\ \text{and } z^i \geq_{\mathbb{R}^n_+} 0, i = 1, \ldots, m \end{array} \right\}$$

- $(\mathcal{C}_n)^* = \mathcal{C}_n^*$ and $\mathcal{C}_n = (\mathcal{C}_n^*)^*$
- $\mathcal{C}_n^* \subset \mathcal{S}_+^n \subset \mathcal{C}_n$

Examples: Cones of Nonnegative Quadratic Functions — Homogeneous

- $\mathcal{F} \subset \mathbb{R}^n$
- Nonnegative homogeneous quadratic functions over $\mathcal{F}$

$$f(x) = x^T Ax \geq 0, \forall x \in \mathcal{F}$$

$$f \Leftrightarrow A$$

- $\mathcal{HD}_{\mathcal{F}} = \{A \in \mathcal{S}^n | x^T Ax \geq 0, \forall x \in \mathcal{F}\}$ is a closed, convex cone.
  (i) Closeness:

$$x^T A_i x \geq 0 \text{ and } A_i \to A \Rightarrow x^T Ax \geq 0$$

  (ii) Convexity:

$$x^T A_i X \geq 0, i = 1, 2 \Rightarrow x^T (\lambda A_1 + (1 - \lambda) A_2) x \geq 0, \forall\, 0 \leq \lambda \leq 1$$

Examples: Cones of Nonnegative Quadratic Functions — Homogeneous

- $\mathcal{HD}^*_{\mathcal{F}} = \mathrm{cl}(\mathrm{cone}\{xx^T | x \in \mathcal{F}\})$
- $(\mathcal{HD}_{\mathcal{F}})^* = \mathcal{HD}^*_{\mathcal{F}}$ and $(\mathcal{HD}^*_{\mathcal{F}})^* = \mathcal{HD}_{\mathcal{F}}$
- Examples:
  - $\mathcal{F} = \mathbb{R}^n$
    $\mathcal{HD}_{\mathcal{F}} = \mathcal{HD}^*_{\mathcal{F}} = \mathcal{S}^n_+$
  - $\mathcal{F} = \mathbb{R}^n_+$
    $\mathcal{HD}_{\mathcal{F}} = \mathcal{C}_n$ and $\mathcal{HD}^*_{\mathcal{F}} = \mathcal{C}^*_n$
  - $\mathcal{F} = \{x | e^T x = 1, x \in \mathbb{R}^n_+\}$
    $\mathcal{HD}_{\mathcal{F}} = \mathcal{C}_n$ and $\mathcal{HD}^*_{\mathcal{F}} = \mathcal{C}^*_n$

Examples: Cones of Nonnegative Quadratic Functions — Nonhomogeneous

- Nonnegative quadratic functions over $\mathcal{F} \subset \mathbb{R}^n$

$$f(x) = x^T A x + 2b^T x + c \geq 0, \forall x \in \mathcal{F}$$

$$f \Leftrightarrow \begin{bmatrix} c & b^T \\ b & A \end{bmatrix}$$

- $\mathcal{D}_{\mathcal{F}} = \{ \begin{bmatrix} c & b^T \\ b & A \end{bmatrix} \in \mathcal{S}^{n+1} | \begin{bmatrix} 1 \\ x \end{bmatrix}^T \begin{bmatrix} c & b^T \\ b & A \end{bmatrix} \begin{bmatrix} 1 \\ x \end{bmatrix} \geq 0, \forall x \in \mathcal{F} \}$ is a closed, convex cone.
- $\mathcal{D}_{\mathcal{F}}^* = \mathrm{cl}(\mathrm{cone}\{ \begin{bmatrix} 1 & x^T \\ x & xx^T \end{bmatrix} | x \in \mathcal{F} \})$
- $(\mathcal{D}_{\mathcal{F}}^*)^* = \mathcal{D}_{\mathcal{F}}$ and $(\mathcal{D}_{\mathcal{F}})^* = \mathcal{D}_{\mathcal{F}}^*$

Examples: Cones of Nonnegative Quadratic Functions — Nonhomogeneous

- Examples:
  - $\mathcal{F} = \mathbb{R}^n$, $\mathcal{D}_{\mathcal{F}} = \mathcal{D}_{\mathcal{F}}^* = \mathcal{S}_+^{n+1}$
  - $\mathcal{F} = \mathbb{R}_+^n$, $\mathcal{D}_{\mathcal{F}} = \mathcal{C}_{n+1}$ and $\mathcal{D}_{\mathcal{F}}^* = \mathcal{C}_{n+1}^*$
- Not a self-dual cone.

$$\mathcal{D}_{\mathcal{F}} = \left\{ U \in S^{n+1} \;\middle|\; \begin{pmatrix} 1 \\ x \end{pmatrix}^T U \begin{pmatrix} 1 \\ x \end{pmatrix} \geq 0 \text{ for any } x \in [0,1]^n. \right\}$$

Obviously, $\mathcal{S}_+^{n+1} \subseteq \mathcal{D}_{\mathcal{F}}$. Then $\mathcal{S}_+^{n+1} \supseteq \mathcal{D}_{\mathcal{F}}^*$. Any matrix $U \in S^{n+1}$ with each element nonnegative is in $\mathcal{D}_{\mathcal{F}}$ which may not be in S_+^{n+1}. For example, $\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \in \mathcal{D}_{\mathcal{F}}$ when $n = 1$, but $\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \notin \mathcal{D}_{\mathcal{F}}^*$.