Convex Sets

Yuanming Shi

ShanghaiTech University

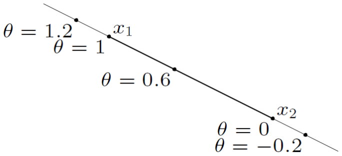
Outline

1. Affine and Convex Sets
2. Some Important Examples
3. Operations that Preserve Convexity
4. Generalized Inequalities
5. Separating and Supporting Hyperplanes

Definition of Affine Set

- **Line:** through $\boldsymbol{x}_1, \boldsymbol{x}_2$: all points

$$\boldsymbol{x} = \theta \boldsymbol{x}_1 + (1 - \theta)\boldsymbol{x}_2 \quad (\theta \in \mathbb{R})$$



- **Affine set:** contains the line through any two distinct points in the set
- **Example:** solution set of linear equations $\{\boldsymbol{x} | \boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}\}$
  (conversely, every affine set can be expressed as solution set of system of linear equations)

$x_1 \in S$, $x_2 \in S$, $\theta \in R$

$x = \theta x_1 + (1-\theta) x_2 \in S$

$Ax = \theta \underbrace{A x_1}_{b} + (1-\theta) \underbrace{A x_2}_{b} = b$

Definition of Convex Set

- **Line segment:** between $\boldsymbol{x}_1$ and $\boldsymbol{x}_2$: all points

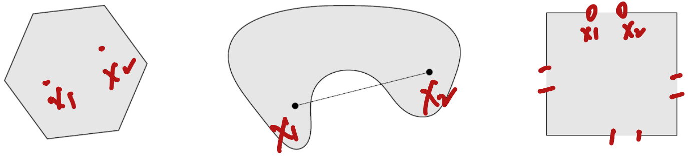
$$\boldsymbol{x} = \theta \boldsymbol{x}_1 + (1 - \theta)\boldsymbol{x}_2$$

with $0 \leq \theta \leq 1$

- **Convex set:** contains line segment between any two points in the set C

$$\boldsymbol{x}_1, \boldsymbol{x}_2 \in C, 0 \leq \theta \leq 1 \quad \Longrightarrow \quad \theta \boldsymbol{x}_1 + (1 - \theta)\boldsymbol{x}_2 \in C$$

- **Examples** (one convex, two nonconvex sets)



Outline

1 Affine and Convex Sets

2 Some Important Examples

3 Operations that Preserve Convexity

4 Generalized Inequalities

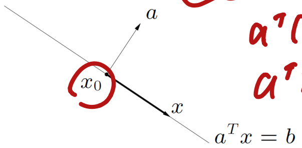
5 Separating and Supporting Hyperplanes

Examples: Hyperplanes and Halfspaces

- **Hyperplane:** set of the form $\{\boldsymbol{x} | \boldsymbol{a}^T\boldsymbol{x} = b\} (\boldsymbol{a} \neq \mathbf{0})$

$a^T x_0 = b$

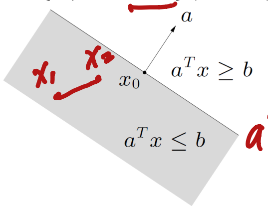
$a \perp (x - x_0)$



$a^T(x - x_0) = 0$

$a^T x = a^T x_0 = b$

- **Halfspace:** set of the form $\{\boldsymbol{x} | \boldsymbol{a}^T\boldsymbol{x} \leq b\} (\boldsymbol{a} \neq \mathbf{0})$



$a^T x = b$

- $\boldsymbol{a}$ is the normal vector
- hyperplanes are affine and convex; halfspaces are convex

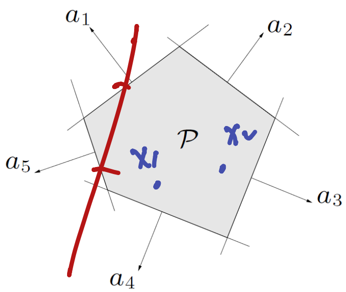
Example: Polyhedra

Solution set of finitely many linear inequalities and equalities

$$\boldsymbol{A}\boldsymbol{x} \preceq \boldsymbol{b}, \qquad \boldsymbol{C}\boldsymbol{x} = \boldsymbol{d}$$

($\boldsymbol{A} \in \mathbb{R}^{m\times n}$, $\boldsymbol{C} \in \mathbb{R}^{p\times n}$, $\preceq$ is componentwise inequality)

$$Ax = \begin{pmatrix} a_1^T x \\ a_2^T x \\ \vdots \\ a_n^T x \end{pmatrix} \qquad b = \begin{pmatrix} b_1 \\ \vdots \\ b_m \end{pmatrix}$$



$a_i^T x \le b_i$

$i = 1, \dots, m$

polyhedron is intersection of finite number of halfspaces and hyperplanes

Examples: Euclidean Balls and Ellipsoids

$\Rightarrow \|ru\|_2 \le r \Rightarrow \|u\|_2 \le 1$

- **(Euclidean) Ball** with center $\boldsymbol{x}_c$ and radius r:

$$B(\boldsymbol{x}_c, r) = \{\boldsymbol{x} \,|\, \|\underbrace{\boldsymbol{x} - \boldsymbol{x}_c}_{=ru}\|_2 \le r\} = \{\boldsymbol{x}_c + r\boldsymbol{u} \,|\, \|\boldsymbol{u}\|_2 \le 1\}$$

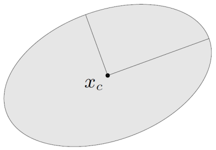
- **Ellipsoid:** set of the form

$$\begin{aligned} E(\boldsymbol{x}_c, \boldsymbol{P}) &= \{\boldsymbol{x} \,|\, (\boldsymbol{x} - \boldsymbol{x}_c)^T \boldsymbol{P}^{-1} (\boldsymbol{x} - \boldsymbol{x}_c) \le 1\} \\ &= \{\boldsymbol{x}_c + \boldsymbol{A}\boldsymbol{u} \,|\, \|\boldsymbol{u}\|_2 \le 1\} \end{aligned}$$

with $\boldsymbol{P} \in \mathbb{S}^n_{++}$ (i.e., $\boldsymbol{P}$ symmetric positive definite), $\boldsymbol{A}$ square and nonsingular

$$\frac{x_1^2}{r_1} + \frac{x_2^2}{r_2} = 1$$

$$P = \begin{pmatrix} r_1 & \\ & r_2 \end{pmatrix}$$



E

the length of the semi-axes of E

are given by $\sqrt{\lambda_i}$:

λ_i eigenvalues of φ .

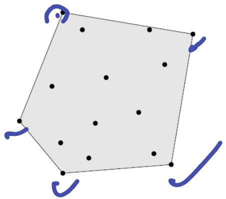
Convex Combination and Convex Hull

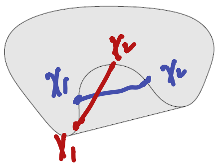
- **Convex combination** of $\boldsymbol{x}_1, \cdots, \boldsymbol{x}_k$: any point $\boldsymbol{x}$ of the form

$$\boldsymbol{x} = \theta_1 \boldsymbol{x}_1 + \theta_2 \boldsymbol{x}_2 + \cdots + \theta_k \boldsymbol{x}_k$$

  with $\theta_1 + \cdots + \theta_k = 1, \theta_i \geq 0$

- **Convex hull** conv S: set of all convex combinations of points in S



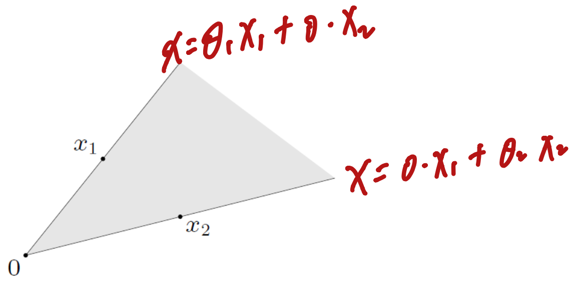


Conic Combination and Convex Cone

- **Conic (nonnegative) combination** of $\boldsymbol{x}_1$ and $\boldsymbol{x}_2$: any point of the form

$$\boldsymbol{x} = \theta_1 \boldsymbol{x}_1 + \theta_2 \boldsymbol{x}_2$$

with $\theta_1 \geq 0, \theta_2 \geq 0$



- **Convex cone:** set that contains all conic combinations of points in the set

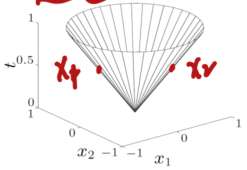
Convex Cones: Norm Balls and Norm Cones

- **Norm:** a function $\|\cdot\|$ that satisfies
  - $\|\boldsymbol{x}\| \geq 0$; $\|\boldsymbol{x}\| = 0$ if and only if $\boldsymbol{x} = \boldsymbol{0}$
  - $\|t\boldsymbol{x}\| = |t|\|\boldsymbol{x}\|$ for $t \in \mathbb{R}$
  - $\|\boldsymbol{x} + \boldsymbol{y}\| \leq \|\boldsymbol{x}\| + \|\boldsymbol{y}\|$

  notation: $\|\cdot\|$ general (unspecified) norm; $\|\cdot\|_{\text{symb}}$ a particular norm

- **Norm ball** with center $\boldsymbol{x}_c$ and radius r: $\{\boldsymbol{x} \mid \|\boldsymbol{x} - \boldsymbol{x}_c\| \leq r\}$

- **Norm cone:** $\{(\boldsymbol{x}, t) \in \mathbb{R}^{n+1} \mid \|\boldsymbol{x}\| \leq t\}$



Euclidean norm cone or second-order cone (aka ice-cream cone)

Positive Semidefinite Cone

Notation

- $\mathbb{S}^n$ is set of symmetric $n \times n$ matrices
- $\mathbb{S}^n_+ = \{\boldsymbol{X} \in \mathbb{S}^n | \boldsymbol{X} \succeq 0\}$: positive semidefinite $n \times n$ matrices

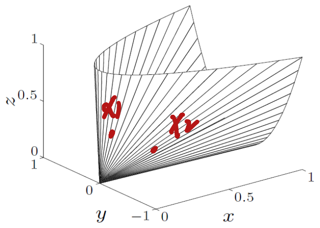
$$\boldsymbol{X} \in \mathbb{S}^n_+ \quad \Longleftrightarrow \quad \boldsymbol{z}^\top \boldsymbol{X} \boldsymbol{z} \geq 0 \text{ for all } \boldsymbol{z}$$

  $\mathbb{S}^n_+$ is a convex cone
- $\mathbb{S}^n_{++} = \{\boldsymbol{X} \in \mathbb{S}^n | \boldsymbol{X} \succ 0\}$: positive definite $n \times n$ matrices

Example: $\begin{bmatrix} x & y \\ y & z \end{bmatrix} \in \mathbb{S}^2_+$

$xz - y^2 \geq 0$

$x \geq 0, \ z \geq 0$



$M_1, \; M_2 \in S_+^n$

$M = \theta_1 M_1 + \theta_2 M_2 \in S_+^n, \quad \theta_1, \theta_2 \geq 0$

$$z^T M z = \theta_1 \underbrace{z^T M_1 z}_{\geq 0} + \theta_2 \underbrace{z^T M_2 z}_{\geq 0}$$

≥ 0

Outline

1 Affine and Convex Sets

2 Some Important Examples

3 Operations that Preserve Convexity

4 Generalized Inequalities

5 Separating and Supporting Hyperplanes

Operations that Preserve Convexity

How to establish the convexity of a given set C

- Apply the definition (can be cumbersome)

$$\boldsymbol{x}_1, \boldsymbol{x}_2 \in C, 0 \leq \theta \leq 1 \quad \Longrightarrow \quad \theta \boldsymbol{x}_1 + (1-\theta)\boldsymbol{x}_2 \in C$$

- Show that C is obtained from simple convex sets(hyperplanes, halfspaces, norm balls,$\cdots$) by operations that preserve convexity
  - intersection
  - affine functions
  - perspective function
  - linear-fractional functions

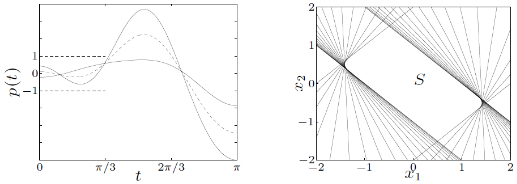
Intersection

$S :=$

- **Intersection:** if $S_1, S_2, \ldots, S_k$ are convex, then $S_1 \cap S_2 \cap \cdots \cap S_k$ is convex (k can be any positive integer)

- Example 1: a polyhedron is the intersection of halfspaces and hyperplanes

- Example 2:
$$S = \{\boldsymbol{x} \in \mathbb{R}^m \,|\, |p(t)| \leq 1 \text{ for } |t| \leq \pi/3\}$$
where $p(t) = x_1 \cos t + x_2 \cos 2t + \cdots + x_m \cos mt$



for $m = 2$

S_α is convex for any $\alpha \in A$

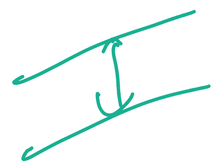
$\bigcap_{\alpha \in A} S_\alpha$ is convex

Examples: infinite number of $\underset{\text{convex}}{\underline{\text{slabs}}}$

given t,

$$S_t = \{ x \in \mathbb{R}^n \mid -1 \le \underbrace{(\cos t, \dots, \cos mt)}_{a(t)}^T x \le 1 \}$$

$$S = \bigcap_{|t| \le \frac{\pi}{3}} S_t \text{ . convex}$$



Affine Function

suppose $f : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ is affine ($f(\boldsymbol{x}) = \boldsymbol{A}\boldsymbol{x} + \boldsymbol{b}$ with $\boldsymbol{A} \in \mathbb{R}^{m\times n}$, $\boldsymbol{b} \in \mathbb{R}^m$)

- the image of a convex set under f is convex

$$S \subseteq \mathbb{R}^n \text{ convex} \quad \Longrightarrow \quad f(S) = \{f(\boldsymbol{x}) | \boldsymbol{x} \in S\} \text{ convex}$$

- the inverse image $f^{-1}(C)$ a convex set under f is convex

$$C \subseteq \mathbb{R}^m \text{ convex} \quad \Longrightarrow \quad f^{-1}(C) = \{\boldsymbol{x} \in \mathbb{R}^n | f(\boldsymbol{x}) \in C\} \text{ convex}$$

Examples

- scaling, translation, projection
- inverse solution set of linear matrix inequality $\{\boldsymbol{x} | x_1 \boldsymbol{A}_1 + \cdots + x_m \boldsymbol{A}_m \preceq \boldsymbol{B}\}$ (with $\boldsymbol{A}_i, \boldsymbol{B} \in \mathbb{S}^p$)

  inverse img.

  $A(x) :=$

  $f(x) = B - A(x) \succeq 0$

  PSD cone

- $\{(\boldsymbol{x}, t) \in \mathbb{R}^{n+1} | \|\boldsymbol{x}\| \le t\}$ is convex, so is

$$\{\boldsymbol{x} \in \mathbb{R}^n | \|\boldsymbol{A}\boldsymbol{x} + \boldsymbol{b}\| \le \boldsymbol{c}^T\boldsymbol{x} + d\}$$

$f(x) = \begin{pmatrix} Ax+b \\ c^Tx+d \end{pmatrix} \begin{matrix} \to y \\ \to s \end{matrix}$

$\Rightarrow \|y\| \le s$, $f(x) \in$ second-order cone

Perspective and Linear-fractional Function I

(x_1, x_2, x_3)

$x_3 = 0$

- **Perspective function** $P : \mathbb{R}^{n+1} \to \mathbb{R}^n$

$$P(\boldsymbol{x}, t) = \boldsymbol{x}/t, \quad \mathrm{dom}P = \{(\boldsymbol{x}, t) | t > 0\}$$

images and inverse images of convex sets under perspective are convex

$x_3 = -1$

$-\left(\frac{x_1}{x_3}, \frac{x_2}{x_3}, 1\right)$

- **Linear-fractional function** $f : \mathbb{R}^n \to \mathbb{R}^m$

$$f(\boldsymbol{x}) = \frac{\boldsymbol{A}\boldsymbol{x} + \boldsymbol{b}}{\boldsymbol{c}^T\boldsymbol{x} + d}, \quad \mathrm{dom}f = \{\boldsymbol{x} | \boldsymbol{c}^T\boldsymbol{x} + d > 0\}$$

images and inverse images of convex sets under linear-fractional functions are convex

$g(x) = \begin{pmatrix} A \\ c^T \end{pmatrix} x + \begin{pmatrix} b \\ d \end{pmatrix}$, affine function

$f(x) = P \circ g$

if C is convex, then

$P^{-1}(C) = \{(x,t) \mid \frac{x}{t} \in C,\ t>0\}$ is convex

proof: $(x,t) \in P^{-1}(C)$, $(y,s) \in P^{-1}(C)$, $0 \le \theta \le 1$

$\theta(x,t) + (1-\theta)(y,s) \in P^{-1}(C)$

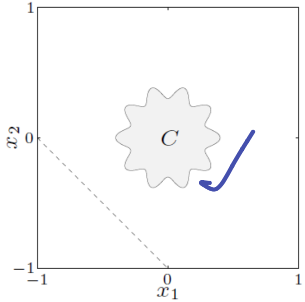
$\frac{\theta x + (1-\theta) y}{\theta t + (1-\theta) s} \in C \iff u\left(\frac{x}{t}\right) + (1-u)\left(\frac{y}{s}\right) \in C$

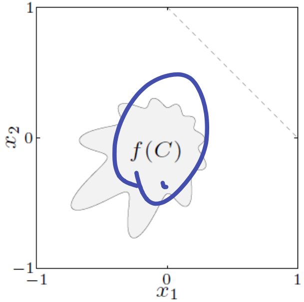
$u = \frac{\theta t}{\theta t + (1-\theta)s} \in [0,1]$

Perspective and Linear-fractional Function II

❧ **Examples** of a linear-fractional function

$$f(\boldsymbol{x}) = \frac{1}{x_1 + x_2 + 1}\boldsymbol{x}$$





Outline

1 Affine and Convex Sets

2 Some Important Examples

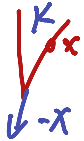
3 Operations that Preserve Convexity

4 Generalized Inequalities

5 Separating and Supporting Hyperplanes

Generalized Inequalities I

- A convex cone $K \subseteq \mathbb{R}^n$ is a **proper cone** if



  - K is closed (contains its boundary)
  - K is solid (has nonempty interior)
  - K is pointed (contains no line)

$x \in K,\ -x \in K \Rightarrow x = 0$

- **Examples**
  - nonnegative orthant

$$K = \mathbb{R}^n_+ = \{\boldsymbol{x} \in \mathbb{R}^n | x_i \geq 0, i = 1, \cdots, n\}$$

  - positive semidefinite cone

(PSD)

$$K = \mathbb{S}^n_+ = \{\boldsymbol{X} \in \mathbb{R}^{n\times n} | \boldsymbol{X} = \boldsymbol{X}^T \succeq \boldsymbol{0}\}$$

$\lambda_i \geq 0,\ i = 1, \cdots, n$

  - nonnegative polynomials on $[0, 1]$:

$$K = \{\boldsymbol{x} \in \mathbb{R}^n | x_1 + x_2 t + x_3 t^2 + \cdots + x_n t^{n-1} \geq 0 \text{ for } t \in [0, 1]\}$$

Generalized Inequalities II

Generalized inequality defined by a proper cone K:

$$\boldsymbol{y} \succeq_K \boldsymbol{x} \iff \boldsymbol{y} - \boldsymbol{x} \succeq_K \boldsymbol{0} \text{ or } \boldsymbol{y} - \boldsymbol{x} \in K$$

Examples

- Componentwise inequality ($K = \mathbb{R}^n_+$) $y_i - x_i \geq 0, \; i = 1, \ldots, n$

$$\boldsymbol{y} \succeq_{\mathbb{R}^n_+} \boldsymbol{x} \iff y_i \geq x_i, \quad i = 1, \cdots, n$$

- Matrix inequality ($K = \mathbb{S}^n_+$)

$$\boldsymbol{Y} \succeq_{\mathbb{S}^n_+} \boldsymbol{X} \iff \boldsymbol{Y} - \boldsymbol{X} \text{ positive semidefinite}$$

Outline

1 Affine and Convex Sets

2 Some Important Examples

3 Operations that Preserve Convexity

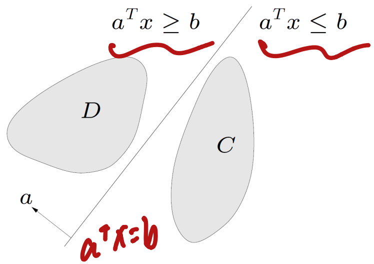
4 Generalized Inequalities

5 Separating and Supporting Hyperplanes

Separating Hyperplane Theorem

If C and D are nonempty disjoint convex sets, there exist $\boldsymbol{a} \neq \mathbf{0}$ and b, such that

$$\boldsymbol{a}^T\boldsymbol{x} \leq b \text{ for } \boldsymbol{x} \in C, \quad \boldsymbol{a}^T\boldsymbol{x} \geq b \text{ for } \boldsymbol{x} \in D$$



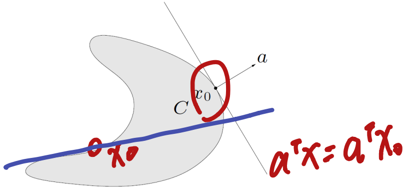
the hyperplane $\{\boldsymbol{x} | \boldsymbol{a}^T\boldsymbol{x} = b\}$ separates C and D

Supporting Hyperplane Theorem

Supporting hyperplane to set C at boundary point $\boldsymbol{x}_0$:

$$\{\boldsymbol{x} | \boldsymbol{a}^T \boldsymbol{x} = \boldsymbol{a}^T \boldsymbol{x}_0\}$$

where $\boldsymbol{a} \neq \boldsymbol{0}$ and $\boldsymbol{a}^T \boldsymbol{x} \leq \boldsymbol{a}^T \boldsymbol{x}_0$ for all $\boldsymbol{x} \in C$

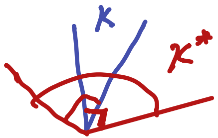


Supporting hyperplane theorem: if C is convex, then there exists a supporting hyperplane at every boundary point of C

Dual Cones and Generalized Inequalities

- **Dual cone** of a cone K:

$\langle y, x \rangle \geq$



$$K^* = \{\boldsymbol{y} | \boldsymbol{y}^T \boldsymbol{x} \geq 0 \text{ for all } \boldsymbol{x} \in K\}$$

- Examples
  - $K = \mathbb{R}^n_+$: $K^* = \mathbb{R}^n_+$
  - $K = \mathbb{S}^n_+$: $K^* = \mathbb{S}^n_+$
  - $K = \{(\boldsymbol{x}, t) | \|\boldsymbol{x}\|_2 \leq t\}$: $K^* = \{(\boldsymbol{x}, t) | \|\boldsymbol{x}\|_2 \leq t\}$
  - $K = \{(\boldsymbol{x}, t) | \|\boldsymbol{x}\|_1 \leq t\}$: $K^* = \{(\boldsymbol{x}, t) | \|\boldsymbol{x}\|_\infty \leq t\}$

  First three examples are **self-dual** cones

- Dual cones of proper cones are proper, hence define generalized inequalities:

$$\boldsymbol{y} \succeq_{K^*} \boldsymbol{0} \quad \Longleftrightarrow \quad \boldsymbol{y}^T \boldsymbol{x} \geq 0 \text{ for all } \boldsymbol{x} \succeq_K \boldsymbol{0}$$

① nonnegative orthant

$y^T x \geq 0$, $\forall x \geq 0 \iff y \geq 0$

② positive semidefinite cone:

$\langle X, Y \rangle = \text{Trace}(XY^T) = \sum_{i,j=1}^{n} X_{ij} Y_{ij}$

The PSD cone is self-dual, i.e., for

$X, Y \in S^n$, $\underline{\text{Trace}(XY^T) \geq 0, \forall X \geq 0} \iff \underline{Y \geq 0}$

Proof: ① Suppose $\underline{Y \notin S_+^n}$, there exists $q \in R^n$

$q^T Y q = \langle q, Yq \rangle = \text{Trace}(qq^T Y^T) < 0$

$X := $ PSD matrix

Have the PSD matrix $X = qq^T$ satisfies

$\text{Trace}(XY) < 0$,

base on the definition of K^*

it follows that $Y \notin (S_+^n)^*$

② Now suppose $X, \underline{Y \in S_+^n}$

we can express X in terms of eigenvalue decomposition as

$X = \sum_{i=1}^{n} \lambda_i q_i q_i^T$, $\lambda_i \geq 0$, $i=1, \dots, n$

$\text{Trace}(YX) = \underline{\text{Trace}(Y \sum_{i=1}^{n} \lambda_i q_i q_i^T)}$

$= \underline{\sum_{i=1}^{n} \lambda_i q_i^T Y q_i} \geq 0$

this shows that $Y \in (S_+^n)^*$

(definition of K^*)

second-order cone is self-dual

$\underbrace{\text{second-order cone}}_{\mathcal{L}^d}$

proof: the dual cone of $\mathcal{L}^d$ is

$$C = \left\{ \begin{pmatrix} y \\ \delta \end{pmatrix} \;\middle|\; 0 \le \underbrace{y^T x + \delta t}_{\langle \begin{pmatrix} y \\ \delta \end{pmatrix}, \begin{pmatrix} x \\ t \end{pmatrix} \rangle \ge 0}, \ \forall \begin{pmatrix} x \\ t \end{pmatrix} \in \mathcal{L}^d \right\}$$

$= \mathcal{L}^d$? $\quad \|y\| \le t$?

inner product

$$= \left\{ \begin{pmatrix} y \\ \delta \end{pmatrix} \;\middle|\; 0 \le \inf_{t \ge 0} \inf_{\|x\| \le t} (y^T x + \delta t) \right\}$$

$$= \left\{ \begin{pmatrix} y \\ \delta \end{pmatrix} \;\middle|\; 0 \le \inf_{t \ge 0} \inf_{\|x\| \le t} (-\|y\|\|x\| + \delta t) \right\}$$

$$= \left\{ \begin{pmatrix} y \\ \delta \end{pmatrix} \;\middle|\; 0 \le \inf_{t \ge 0} (\delta - \|y\|) t \right\}$$

since $t \ge 0$, one has

$\begin{pmatrix} y \\ \delta \end{pmatrix} \in C$ iff $\|y\| \le \delta \Rightarrow C = \mathcal{L}^d$

second-order cone

Dual norm:

Let $\|\cdot\|$ be a norm on $\mathbb{R}^n$. The associated dual norm, $\|\cdot\|_*$, is defined as

$$\|z\|_* = \sup \{ z^T x \mid \|x\| \le 1 \}$$

fact: $\|x\|_{**} = \|x\|$, $\forall x$

Example 1: The dual of Euclidean norm is the Euclidean norm, i.e.,

$$\sup\{z^T x \mid \|x\|_2 \le 1\} = \|z\|_2$$

This follows from Cauchy-Schwarz inequality

for nonzero z, the value of x that (given z)

$$\left.\begin{array}{ll} \underset{x \in \mathbb{R}^n}{\text{maximize}} & z^T x \\ \text{subject to} & \|x\|_2 \le 1 \end{array}\right\} \Rightarrow x^* = \frac{z}{\|z\|_2}$$

optimal objective value: $z^T x^* = \|z\|_2$

Example 2: The dual of the l_∞-norm is the l_1-norm

$$\sup\{z^T x \mid \underbrace{\|x\|_\infty}_{\max_i |x_i|} \le 1\} = \sum_{i=1}^n |z_i| = \|z\|_1$$

$$\left.\begin{array}{ll} \text{maximize} & z^T x \\ \text{subject to} & \|x\|_\infty \le 1 \end{array}\right\} \Rightarrow x_i^* = \frac{z_i}{|z_i|}$$

General results: the dual of the l_p-norm is the l_q-norm, q satisfies

$$\frac{1}{p} + \frac{1}{q} = 1$$

Reference

Chapter 2 of:

- Stephen Boyd and Lieven Vandenberghe, *Convex Optimization*. Cambridge, U.K.: Cambridge University Press, 2004.