Optimization for the information and data sciences

Mahdi Soltanolkotabi

Convex Sets

Ming Hsieh Department of Electrical Engineering

USC University of Southern California

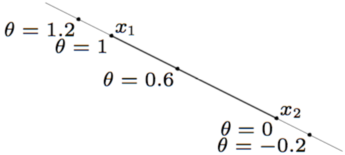
Convex sets

- affine and convex sets
- some important examples
- operations that preserve convexity
- generalized inequalities
- separating and supporting hyperplanes
- dual cones and generalized inequalities

Affine sets

line through x_1, x_2: all points

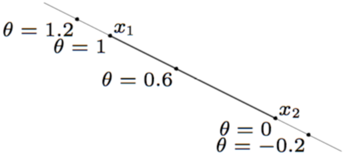
$$x = \theta x_1 + (1 - \theta) x_2 \qquad (\theta \in \mathbb{R})$$



Affine sets

line through x_1, x_2: all points

$$x = \theta x_1 + (1-\theta)x_2 \qquad (\theta \in \mathbb{R})$$

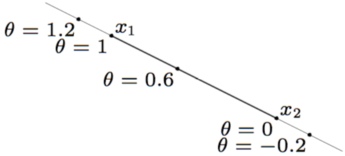


affine set: contains the line through any two distinct points in the set

Affine sets

line through x_1, x_2: all points

$$x = \theta x_1 + (1 - \theta) x_2 \qquad (\theta \in \mathbb{R})$$



affine set: contains the line through any two distinct points in the set

example: solution set of linear equations $\{x \mid Ax = b\}$
(conversely, every affine set can be expressed as solution set of system of linear equations)

Convex set

line segment between x_1 and x_2: all points

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

with $0 \leq \theta \leq 1$

Convex set

line segment between x_1 and x_2: all points

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

with $0 \leq \theta \leq 1$

convex set: contains line segment between any two points in the set

$$x_1, x_2 \in C, \quad 0 \leq \theta \leq 1 \quad \Longrightarrow \quad \theta x_1 + (1 - \theta) x_2 \in C$$

Convex set

line segment between x_1 and x_2: all points

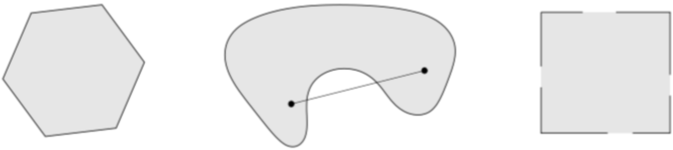
$$x = \theta x_1 + (1-\theta)x_2$$

with $0 \leq \theta \leq 1$

convex set: contains line segment between any two points in the set

$$x_1, x_2 \in C, \quad 0 \leq \theta \leq 1 \quad \Longrightarrow \quad \theta x_1 + (1-\theta)x_2 \in C$$

examples (one convex, two nonconvex sets)



Convex combination and convex hull

convex combination of x_1,..., x_k: any point x of the form

$$x = \theta_1 x_1 + \theta_2 x_2 + \cdots + \theta_k x_k$$

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

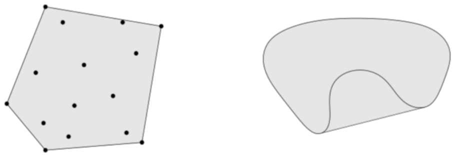
Convex combination and convex hull

convex combination of x_1,..., x_k: any point x of the form

$$x = \theta_1 x_1 + \theta_2 x_2 + \cdots + \theta_k x_k$$

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

convex hull conv S: set of all convex combinations of points in S

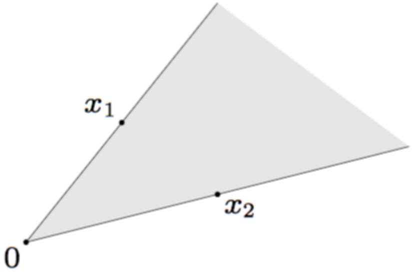


Convex cone

conic (nonnegative) combination of x_1 and x_2: any point of the form

$$x = \theta_1 x_1 + \theta_2 x_2$$

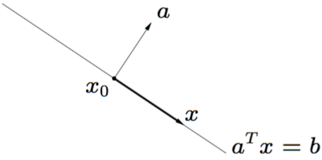
with $\theta_1 \geq 0$, $\theta_2 \geq 0$



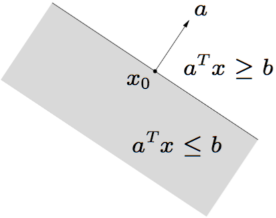
convex cone: set that contains all conic combinations of points in the set

Hyperplanes and halfspaces

hyperplane: set of the form $\{x \mid a^T x = b\}$ $(a \neq 0)$



halfspace: set of the form $\{x \mid a^T x \leq b\}$ $(a \neq 0)$



- a is the normal vector
- hyperplanes are affine and convex; halfspaces are convex

Euclidean balls and ellipsoids

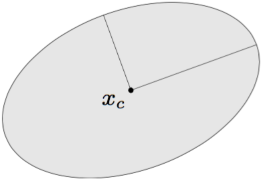
(Euclidean) ball with center x_c and radius r:

$$B(x_c, r) = \{x \mid \|x - x_c\|_2 \leq r\} = \{x_c + ru \mid \|u\|_2 \leq 1\}$$

ellipsoid: set of the form

$$\{x \mid (x - x_c)^T P^{-1}(x - x_c) \leq 1\}$$

with $P \in \mathcal{S}^n_{++}$ (i.e., P symmetric positive definite)



other representation: $\{x_c + Au \mid \|u\|_2 \leq 1\}$ with A square and nonsingular

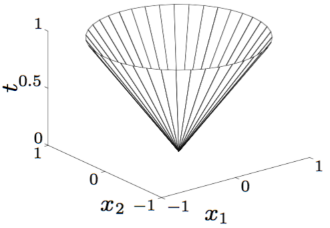
Norm balls and norm cones

norm: a function $\|\cdot\|$ that satisfies

- $\|x\| \geq 0$; $\|x\| = 0$ if and only if $x = 0$
- $\|tx\| = |t| \, \|x\|$ for $t \in \mathbb{R}$
- $\|x + y\| \leq \|x\| + \|y\|$

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

norm ball with center x_c and radius r: $\{x \mid \|x - x_c\| \leq r\}$



norm cone: $\{(x, t) \mid \|x\| \leq t\}$

Euclidean norm cone is called second-order cone

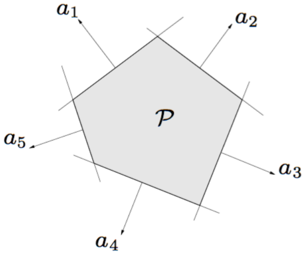
norm balls and cones are convex

Polyhedra

solution set of finitely many linear inequalities and equalities

$$Ax \preceq b, \qquad Cx = d$$

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



polyhedron is intersection of finite number of halfspaces and hyperplanes

Positive semidefinite cone

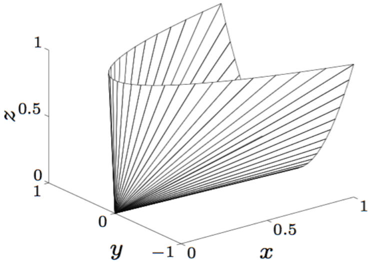
notation:

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

$$X \in \mathcal{S}^n_+ \quad \Longleftrightarrow \quad z^T X z \geq 0 \text{ for all } z$$

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

example: $\left[\begin{array}{cc} x & y \\ y & z \end{array}\right] \in \mathcal{S}^2_+$



Operations that preserve convexity

practical methods for establishing convexity of a set C

1. apply definition

$$x_1, x_2 \in C, \quad 0 \leq \theta \leq 1 \quad \Longrightarrow \quad \theta x_1 + (1-\theta)x_2 \in C$$

2. show that C is obtained from simple convex sets (hyperplanes, halfspaces, norm balls, ...) by operations that preserve convexity
   - intersection
   - affine functions
   - perspective function
   - linear-fractional functions

Intersection

the intersection of (any number of) convex sets is convex

- If S_1 and S_2 are convex then $S_1 \bigcap S_2$ is convex.

Intersection

the intersection of (any number of) convex sets is convex

- If S_1 and S_2 are convex then $S_1 \bigcap S_2$ is convex.
- If S_α is convex for every $\alpha \in \mathcal{A}$ then $\bigcap_{\alpha \in \mathcal{A}} S_\alpha$ is convex

Intersection

the intersection of (any number of) convex sets is convex

- If S_1 and S_2 are convex then $S_1 \bigcap S_2$ is convex.
- If S_α is convex for every $\alpha \in \mathcal{A}$ then $\bigcap_{\alpha \in \mathcal{A}} S_\alpha$ is convex
- Example: PSD cone $\mathcal{S}_+^n$ can be expressed as

$$\bigcap_{z \neq 0} \left\{ X \in \mathcal{S}^n \mid z^T X z \geq 0 \right\}$$

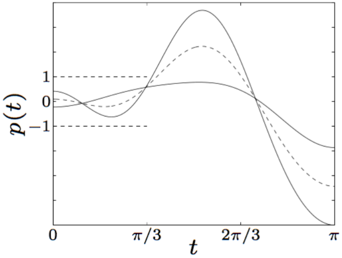
Intersection

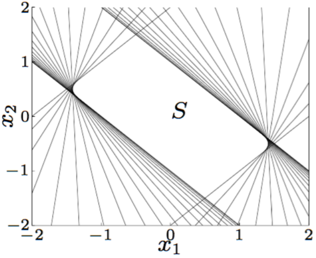
the intersection of (any number of) convex sets is convex

example:

$$S = \{x \in \mathbb{R}^m \mid |p(t)| \leq 1 \text{ for } |t| \leq \pi/3\}$$

where $p(t) = x_1 \cos t + x_2 \cos 2t + \cdots + x_m \cos mt$
for $m = 2$:





Affine function

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

- the image of a convex set under f is convex

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

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

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

examples

- scaling, translation, projection
- solution set of linear matrix inequality $\{x \mid x_1 A_1 + \cdots + x_m A_m \preceq B\}$ (with $A_i, B \in \mathcal{S}^p$)
- hyperbolic cone $\{x \mid x^T P x \leq (c^T x)^2,\ c^T x \geq 0\}$ (with $P \in \mathcal{S}^n_+$)

Perspective and linear-fractional function

perspective function $P : \mathbb{R}^{n+1} \to \mathbb{R}^n$:

$$P(x,t) = x/t, \qquad \textbf{dom} P = \{(x,t) \mid t > 0\}$$

images and inverse images of convex sets under perspective are convex

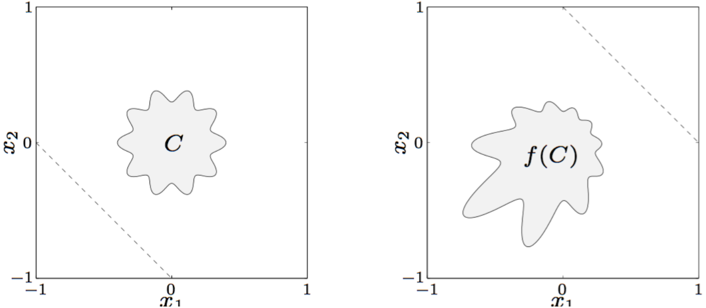
linear-fractional function $f : \mathbb{R}^n \to \mathbb{R}^m$:

$$f(x) = \frac{Ax+b}{c^T x + d}, \qquad \textbf{dom} f = \{x \mid c^T x + d > 0\}$$

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

example of a linear-fractional function

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



Generalized inequalities

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)

examples

- nonnegative orthant $K = \mathbb{R}^n_+ = \{x \in \mathbb{R}^n \mid x_i \geq 0, i = 1, \ldots, n\}$
- positive semidefinite cone $K = \mathcal{S}^n_+$
- coefficients of nonnegative polynomials on $[0, 1]$:

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

generalized inequality defined by a proper cone K:

$$x \preceq_K y \quad \Longleftrightarrow \quad y - x \in K, \qquad\qquad x \prec_K y \quad \Longleftrightarrow \quad y - x \in \mathbf{int} K$$

examples

- componentwise inequality ($K = \mathbb{R}^n_+$)

$$x \preceq_{\mathbf{R}^n_+} y \quad \Longleftrightarrow \quad x_i \leq y_i, \quad i = 1, \ldots, n$$

- matrix inequality ($K = \mathcal{S}^n_+$)

$$X \preceq_{\mathbf{S}^n_+} Y \quad \Longleftrightarrow \quad Y - X \text{ positive semidefinite}$$

these two types are so common that we drop the subscript in $\preceq_K$

properties: many properties of $\preceq_K$ are similar to $\leq$ on $\mathbb{R}$, e.g.,

$$x \preceq_K y, \quad u \preceq_K v \quad \Longrightarrow \quad x + u \preceq_K y + v$$

Minimum and minimal elements

$\preceq_K$ is not in general a *linear ordering*: we can have $x \not\preceq_K y$ and $y \not\preceq_K x$

$x \in S$ is **the minimum element** of S with respect to $\preceq_K$ if

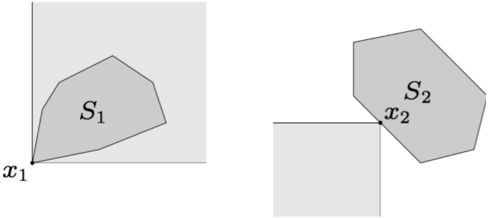
$$y \in S \quad \Longrightarrow \quad x \preceq_K y$$

$x \in S$ is **a minimal element** of S with respect to $\preceq_K$ if

$$y \in S, \quad y \preceq_K x \quad \Longrightarrow \quad y = x$$

example $(K = \mathbb{R}^2_+)$

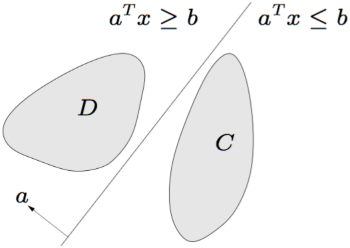
x_1 is the minimum element of S_1
x_2 is a minimal element of S_2



Separating hyperplane theorem

if C and D are disjoint convex sets, then there exists $a \neq 0$, b such that

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



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

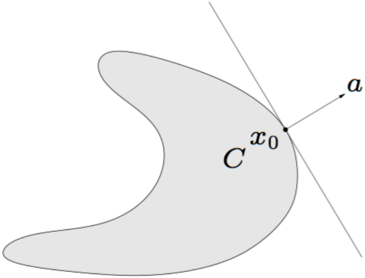
strict separation requires additional assumptions (e.g., C is closed, D is a singleton)

Supporting hyperplane theorem

supporting hyperplane to set C at boundary point x_0:

$$\{x \mid a^T x = a^T x_0\}$$

where $a \neq 0$ and $a^T x \leq a^T x_0$ for all $x \in C$



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:

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

examples

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

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

dual cones of proper cones are proper, hence define generalized inequalities:

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

Dual cones and generalized inequalities

dual cone of a cone K:

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

examples

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

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

dual cones of proper cones are proper, hence define generalized inequalities:

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

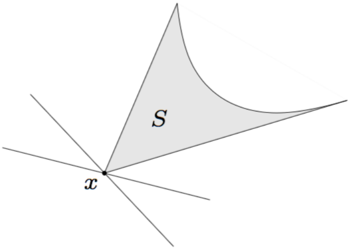
important properties relating a generalized inequality and its dual

- $x \preceq_K y$ if and only if $\lambda^T x \leq \lambda^T y$ for all $\lambda \succeq_{K^*} 0$
- $x \prec_K y$ if and only if $\lambda^T x < \lambda^T y$ for all $\lambda \succeq_{K^*} 0$, $\lambda \neq 1$

Minimum and minimal elements via dual inequalities

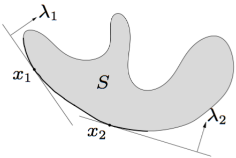
minimum element w.r.t. $\preceq_K$

x is minimum element of S iff for all $\lambda \succ_{K^*} 0$, x is the unique minimizer of $\lambda^T z$ over S



minimal element w.r.t. $\preceq_K$

- if x minimizes $\lambda^T z$ over S for some $\lambda \succ_{K^*} 0$, then x is minimal



- if x is a minimal element of a *convex* set S, then there exists a nonzero $\lambda \succ_{K^*} 0$ such that x minimizes $\lambda^T z$ over S

optimal production frontier

- different production methods use different amounts of resources $x \in \mathbb{R}^n$
- production set P: resource vectors x for all possible production methods
- efficient (Pareto optimal) methods correspond to resource vectors x that are minimal w.r.t. $\mathbb{R}^n_+$

example $(n = 2)$

x_1, x_2, x_3 are efficient; x_4, x_5 are not