CSED700H: Convex Optimization

Convex sets[1]

Namhoon Lee

POSTECH

Fall 2023

[1]slides credits to Prof. Lieven Vandenberghe

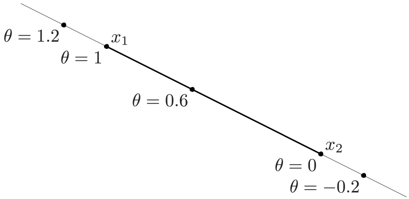
Contents

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

Affine set

Line through points x_1, x_2: all points

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



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 linear equations

Convex set

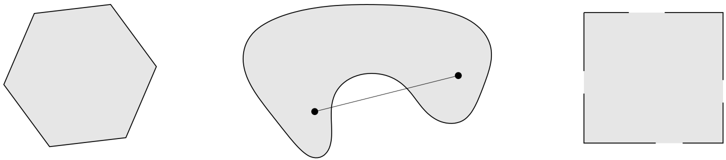
Line segment between points $x1, x2$: all points

$$x = \theta x_1 + (1-\theta)x_2 \qquad \text{with } 0 \le \theta \le 1$$

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

$$x1, x2 \in \mathbb{C}, \quad 0 \le \theta \le 1 \quad \Longrightarrow \quad \theta x_1 + (1-\theta)x_2 \in \mathbb{C}$$

Examples (one convex, two nonconvex sets)



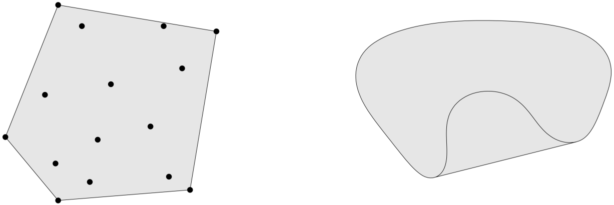
Convex combination and convex hull

Convex combination of $x_1, \ldots, x_k$: any point x of the form

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

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

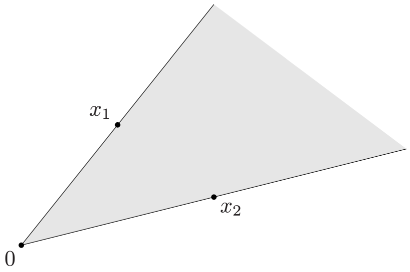
Convex hull: $\mathrm{conv}\,\mathbb{S}$ is set of all convex combinations of points in $\mathbb{S}$



Convex cone

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

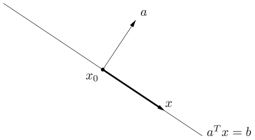
$$x = \theta_1 x_1 + \theta_2 x_2 \qquad \text{with } \theta_1 \geq 0, \theta_2 \geq 0$$



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

Hyperplanes and halfspaces

Hyperplane: set of the form $\{x \mid a^\top x = b\}$ where $a \neq 0$

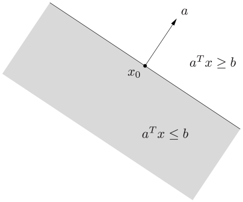


x_0 is a particular element, *e.g.*,

$$x_0 = \frac{b}{a^\top a} a$$

$a^\top x = b$ if and only if $a \perp (x - x_0)$

Halfspace: set of the form $\{x \mid a^\top x \leq b\}$ where $a \neq 0$



hyperplanes are affine and convex; halfspaces are convex

Euclidean balls and ellipsoids

(Euclidean) ball with center x_c and radius r:

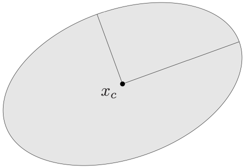
$$B(x_c, r) = \{x \mid \|x - x_c\| \leq r\} = \{x_c + ru \mid \|u\|_2 \leq 1\}$$

$\|\cdot\|_2$ denotes the Euclidean norm

Ellipsoid: set of the form

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

with P symmetric positive definite



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

Principal axes

$$\mathcal{E} = \{x \mid (x - x_c)^\top P^{-1}(x - x_c) \le 1\}$$

Eigendecomposition: $P = Q\Lambda Q^\top = \sum_{i=1}^n \lambda_i q_i q_i^\top$

- Q is orthogonal ($Q^\top = Q^{-1}$) with columns q_i
- Λ is diagonal with diagonal elements $\lambda_1 \ge \lambda_2 \ge \cdots \ge \lambda_n > 0$

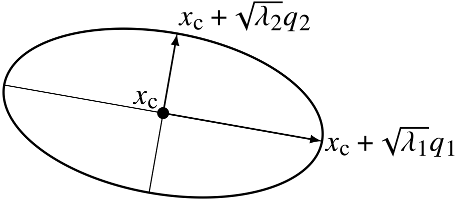
Change of variables: $y = Q^\top(x - x_c), x = x_c + Qy$

- after the change of variables the ellipsoid is described by

$$y^\top \Lambda^{-1} y = y_1^2/\lambda_1 + \cdots + y_n^2/\lambda_n \le 1$$

  an ellipsoid centered at the origin, and aligned with the coordinate axes
- eigenvectors q_i of P give the principal axes of $\mathcal{E}$
- the width of $\mathcal{E}$ along the principal axis determined by q_i is $2\sqrt{\lambda_i}$

Example in $\mathbb{R}^2$



Exercise: give an interpretation of $\mathrm{tr}(P)$ as a measure of the size of

$$\mathcal{E} = \{x \mid (x - x_c)^\top P^{-1}(x - x_c) \leq 1\}$$

Norms

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

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

Notation

- $\|\cdot\|$ is a general (unspecified) norm
- $\|\cdot\|_{\mathsf{symb}}$ is a particular norm

Frequently used norms

Vector norm ($x \in \mathbb{R}^n$)

- Euclidean norm $\|x\|_2 = (x_1^2 + \cdots + x_n^2)^{1/2}$
- p-norm ($p \geq 1$) and ∞-norm (Chebyshev norm)

$$\|x\|_p = (|x_1|^p + \cdots + |x_n|^p)^{1/p}, \qquad \|x\|_\infty = \max_{k=1,\ldots,n} |x_k|$$

- quadratic norm: $\|x\|_A = (x^\top A x)^{1/2}$, with A symmetric positive definite

Matrix norms ($X \in \mathbb{R}^{m \times n}$)

- Frobenius norm: $\|X\|_F = (\sum_{i=1}^m \sum_{j=1}^n X_{ij}^2)^{1/2}$
- 2-norm (spectral norm, operator norm)

$$\|X\|_2 = \sup_{y \neq 0} \frac{\|Xy\|_2}{\|y\|_2} = \sigma_{\text{max}}(X)$$

$\sigma_{\text{max}}(X) = (\lambda_{\text{max}}(X^\top X))^{1/2}$ is largest singular value of X

Norm balls and norm cones

Norm ball with center x_c and radius r:

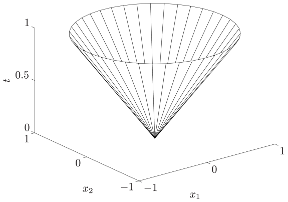
$$\{x \mid \|x - x_c\| \leq r\}$$

norm balls are convex

Norm cone:

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

- norm cones are convex
- example: second order cone (norm cone for Euclidean norm)

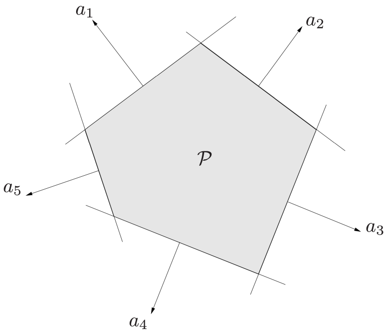


Polyhedra

Polyhedron: solution set of *finitely many* linear inequalities and equalities

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

$\preceq$ denotes componentwise inequality between vectors



polyhedron is intersection of finite number of halfspaces and hyperplanes

Positive semidefinite cone

Notation

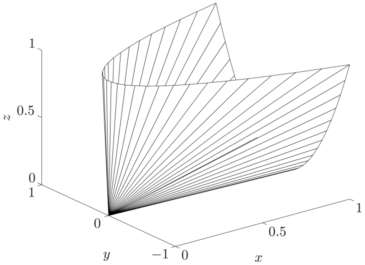
- $\mathbb{S}^n$ is set of symmetric $n \times n$ matrices
- $\mathbb{S}^n_+ = \{X \in \mathbb{S}^n \mid X \succeq 0\}$: positive semidefinite $n \times n$ matrices

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

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

Example

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



Operations that preserve convexity

methods for establishing convexity of a set $\mathbb{C}$

1. apply definition

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

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

Intersection

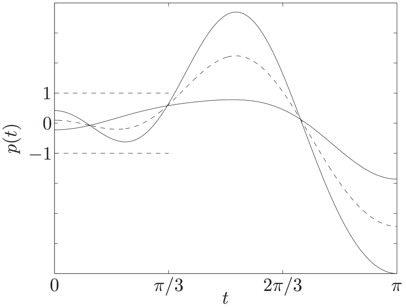
the intersection of (any number of) convex sets is convex

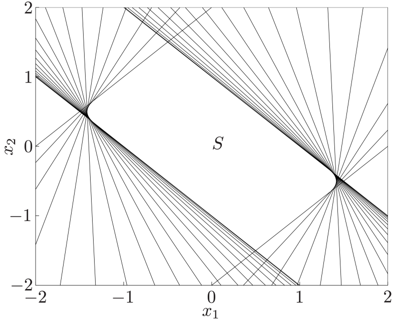
Example

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

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 an affine function:

$$f(x) = Ax + b$$

with $A \in \mathbb{R}^{m \times n}$, $b \in \mathbb{R}^m$

- the image of a convex set under f is convex

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

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

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

Examples

- scaling, translation, projection
- image and inverse image of norm ball under affine transformation

$$\{Ax + b \mid \|x\| \leq 1\}, \qquad \{x \mid \|Ax + b\| \leq 1\}$$

- hyperbolic cone

$$\{x \mid x^\top P x \leq (c^\top x)^2,\ c^\top x \geq 0\}, \qquad \text{with } P \in \mathbb{S}^n_+$$

- solution set of linear matrix inequality

$$\{x \mid x_1 A_1 + \cdots + x_m A_m \preceq B\}, \qquad \text{with } A_i, B \in \mathbb{S}^p$$

Perspective and linear-fractional function

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

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

images and inverse images of convex sets under perspective are convex

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

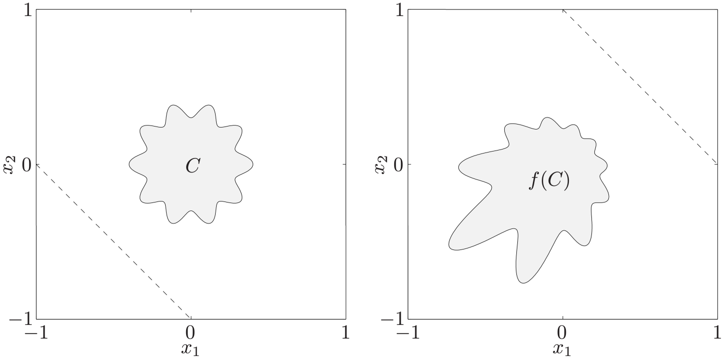
$$f(x) = \frac{Ax + b}{c^\top x + d}, \qquad \operatorname{dom} f = \{x \mid c^\top x + d > 0\}$$

image and inverse image of convex sets under linear-fractional function are convex

Example

a linear-fractional function from $\mathbb{R}^2$ to $\mathbb{R}^2$

$$f(x) = \frac{1}{x_1 + x_2 + 1} x, \qquad \operatorname{dom} f = \{(x_1, x_2) \mid x_1 + x_2 + 1 > 0\}$$



Proper cone

Proper cone: a convex cone $\mathbb{K} \in \mathbb{R}^n$ that satisifes three properties

- $\mathbb{K}$ is closed (contains its boundary)
- $\mathbb{K}$ is solid (has nonempty interior)
- $\mathbb{K}$ is pointed (contains no line)

Examples

- nonnegative orthant

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

- positive semidefinite cone $\mathbb{K} = \mathbb{S}^n_+$
- nonnegative polynomials on $[0, 1]$:

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

Generalized inequality

Generalized inequality defined by a proper cone $\mathbb{K}$:

$$x \preceq_{\mathbb{K}} y \iff y - x \in \mathbb{K}, \qquad x \prec_{\mathbb{K}} y \iff y - x \in \operatorname{int} \mathbb{K}$$

Examples

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

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

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

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

these two types are so common that we drop the subscript in $\preceq_{\mathbb{K}}$

Properties: many properties of $\preceq \mathbb{K}$ are similar to $\leq$ on $\mathbb{R}$, *e.g.*,

$$x \preceq_{\mathbb{K}} y, \quad u \preceq_{\mathbb{K}} v \implies x + u \preceq_{\mathbb{K}} y + v$$

Minimum and minimal elements

$\preceq_{\mathbb{K}}$ is not in general a *linear ordering*: we can have $x \not\preceq_{\mathbb{K}} y$ and $y \not\preceq_{\mathbb{K}} x$

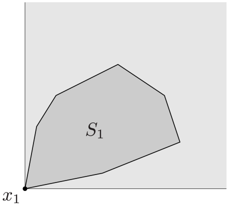
$x \in \mathbb{S}$ is **the minimum element** of $\mathbb{S}$ with respect to $\prec_{\mathbb{K}}$ if

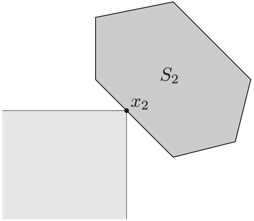
$$y \in \mathbb{S} \quad \Longrightarrow \quad x \preceq_{\mathbb{K}} y$$

$x \in \mathbb{S}$ is **the minimal element** of $\mathbb{S}$ with respect to $\prec_{\mathbb{K}}$ if

$$y \in \mathbb{S}, \quad y \preceq_{\mathbb{K}} x \quad \Longrightarrow \quad y = x$$

Example $(\mathbb{K} = \mathbb{R}^2_+)$
x_1 is the minimum element of $\mathbb{S}_1$
x_2 is the minimum element of $\mathbb{S}_2$





Inner products

in this course we will use the following standard inner products

- for vectors $x, y \in \mathbb{R}^n$:

$$\langle x, y \rangle = x_1 y_1 + \cdots + x_n y_n = x^\top y$$

- for matrices $X, Y \in \mathbb{R}^{m \times n}$:

$$\langle X, Y \rangle = \sum_{i=1}^{m} \sum_{j=1}^{n} X_{ij} Y_{ij} = \operatorname{tr}(X^\top Y)$$

- for symmetric matrices $X, Y \in \mathbb{S}^n$:

$$\langle X, Y \rangle = \sum_{i=1}^{n} X_{ii} Y_{ii} + 2 \sum_{i>j} X_{ij} Y_{ij} = \operatorname{tr}(XY)$$

Dual cones

Dual cone of a cone $\mathbb{K}$:

$$\mathbb{K}^* = \{y \mid \langle y, x \rangle \geq 0 \text{ for all } x \in \mathbb{K}\}$$

note: definition depends on choice of inner product

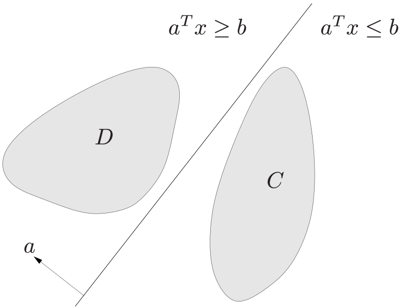
Examples

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

Separating hyperplane theorem

if C and D are nonempty disjoint convex sets, there exist $a \neq 0, b$ s.t.

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



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

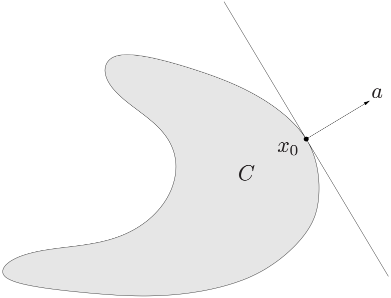
strict separation requires additional assumptions (*e.g.*, $\mathbb{C}$ closed, $\mathbb{D}$ a singleton)

Supporting hyperplane theorem

Supporting hyperplane to set $\mathbb{C}$ at boundary point x_0:

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

where $a \neq 0$ and $a^\top x \leq a^\top x_0$ for all $x \in \mathbb{C}$



Supporting hyperplane theorem:
there exists a supporting hyperplane at every boundary point of a convex set $\mathbb{C}$