2. Convex sets

- 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

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

**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 $x_1, x_2$: all points

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

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

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

Examples (one convex, two nonconvex sets)
Convex combination and convex hull

**Convex combination** of \( x_1, \ldots, x_k \): 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**: \( \text{conv} \ S \) is set of all convex combinations of points in \( S \)
Convex cone

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

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

**Convex cone**: set that contains all conic combinations of points in the set.
Hyperplanes and halfspaces

**Hyperplane:** set of the form \( \{ x \mid a^T x = b \} \) where \( a \neq 0 \)

\[ a^T x = b \]

\( x_0 \) is a particular element, e.g.,

\[ x_0 = \frac{b}{a^T a} \]

\[ a^T x = b \text{ if and only if } a \perp (x - x_0) \]

**Halfspace:** set of the form \( \{ x \mid a^T x \leq b \} \) where \( a \neq 0 \)

\[ a^T x \geq b \]

\[ a^T x \leq b \]

hyperplanes are affine and convex; halfspaces are convex
Euclidean balls and ellipsoids

(Euclidean) ball with center $x_c$ and radius $r$:

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

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

**Ellipsoid:** set of the form

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

with $P$ symmetric positive definite

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

\[ \mathcal{E} = \{ x \mid (x - x_c)^T P^{-1} (x - x_c) \leq 1 \} \]

**Eigendecomposition:** \( P = Q \Lambda Q^T = \sum_{i=1}^{n} \lambda_i q_i q_i^T \)

- \( Q \) is orthogonal \((Q^T = Q^{-1})\) with columns \( q_i \)
- \( \Lambda \) is diagonal with diagonal elements \( \lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n > 0 \)

**Change of variables:** \( y = Q^T (x - x_c), \quad x = x_c + Qy \)

- after the change of variables the ellipsoid is described by

\[ y^T \Lambda^{-1} y = y_1^2/\lambda_1 + \cdots + y_n^2/\lambda_n \leq 1 \]

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 $\text{tr}(P)$ as a measure of the size of

$$\mathcal{E} = \{x \mid (x - x_c)^T P^{-1} (x - x_c) \leq 1\}$$
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
- $\| \cdot \|_{\text{symb}}$ is a particular norm
Frequently used norms

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

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

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

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

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

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

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

\[
\sigma_{\max}(X) = (\lambda_{\max}(X^T X))^{1/2} \text{ is largest singular value of } X
\]
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 \textit{finitely many} linear inequalities and equalities

\[ Ax \leq b, \quad Cx = d \]

\( \leq \) denotes componentwise inequality between vectors

polyhedron is intersection of finite number of halfspaces and hyperplanes
Positive semidefinite cone

Notation

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

$$X \in S_n^+ \iff z^T X z \geq 0 \text{ for all } z$$

$S_n^+$ is a convex cone

- $S_{++}^n = \{X \in S^n \mid X > 0\}$: positive definite $n \times n$ matrices

Example

$$\begin{bmatrix} x & y \\ y & z \end{bmatrix} \in S_2^+$$
Operations that preserve convexity

methods for establishing convexity of a set $C$

1. apply definition

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

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

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 \( m = 2 \):

\[ f(t) = 0 \]

\[ 0 \leq t \leq \pi/3 \]

\[ 0 \leq x_1 \leq 1 \]

\[ -1 \leq x_2 \leq 1 \]

\[ 0 \leq x_1 \leq 2 \]

\[ -2 \leq x_2 \leq -1 \]
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

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

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

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

- scaling, translation, projection

- image and inverse image of norm ball under affine transformation

\[ \{Ax + b \mid \|x\| \leq 1\}, \quad \{x \mid \|Ax + b\| \leq 1\} \]

- hyperbolic cone

\[ \{x \mid x^T Px \leq (c^T x)^2, \quad c^T x \geq 0\} \quad \text{with } P \in S_+^n \]

- solution set of linear matrix inequality

\[ \{x \mid x_1A_1 + \cdots + x_mA_m \leq B\} \quad \text{with } A_i, B \in S^p \]
Perspective and linear-fractional function

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

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

Images and inverse images of convex sets under perspective are convex.

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

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

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

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

\[ f(x) = \frac{1}{x_1 + x_2 + 1} x, \quad \text{dom } f = \{(x_1, x_2) \mid x_1 + x_2 + 1 > 0\} \]
Proper cone

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

- $K$ 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 $K = S_+^n$

- 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

**Generalized inequality** defined by a proper cone $K$:

\[ x \preceq_K y \iff y - x \in K, \quad x \preceq_K y \iff y - x \in \text{int } K \]

**Examples**

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

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

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

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

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

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

\[ x \preceq_K y, \quad u \preceq_K v \implies x + u \preceq_K y + v \]
Minimum and minimal elements

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

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

$$y \in S \implies x \leq_K y$$

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

$$y \in S, \quad y \leq_K x \implies y = x$$

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

$x_1$ is the minimum element of $S_1$

$x_2$ is a minimal element of $S_2$
Inner products

in this course we will use the following standard inner products

• for vectors \(x, y \in \mathbb{R}^n\):

\[
\langle x, y \rangle = x_1 y_1 + \cdots + x_n y_n = x^T y
\]

• for matrices \(X, Y \in \mathbb{R}^{m \times n}\)

\[
\langle X, Y \rangle = \sum_{i=1}^{m} \sum_{j=1}^{n} X_{ij} Y_{ij} = \text{tr}(X^T Y)
\]

• for symmetric matrices \(X, Y \in \mathbb{S}^n\)

\[
\langle X, Y \rangle = \sum_{i=1}^{n} X_{ii} Y_{ii} + 2 \sum_{i > j} X_{ij} Y_{ij} = \text{tr}(XY)
\]
Dual cones

**Dual cone** of a cone $K$:

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

note: definition depends on choice of inner product

**Examples**

<table>
<thead>
<tr>
<th></th>
<th>$K$</th>
<th>$K^*$</th>
</tr>
</thead>
<tbody>
<tr>
<td>nonnegative orthant</td>
<td>$\mathbb{R}^n_+$</td>
<td>$\mathbb{R}^n_+$</td>
</tr>
<tr>
<td>second order cone</td>
<td>${(x, t) \mid |x|_2 \leq t}$</td>
<td>${(x, t) \mid |x|_2 \leq t}$</td>
</tr>
<tr>
<td>1-norm cone</td>
<td>${(x, t) \mid |x|_1 \leq t}$</td>
<td>${(x, t) \mid |x|_\infty \leq t}$</td>
</tr>
<tr>
<td>positive semidefinite cone</td>
<td>$\mathbb{S}^n_+$</td>
<td>$\mathbb{S}^n_+$</td>
</tr>
</tbody>
</table>
Separating hyperplane theorem

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

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

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

strict separation requires additional assumptions (e.g., $C$ closed, $D$ 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:

there exists a supporting hyperplane at every boundary point of a convex set $C$