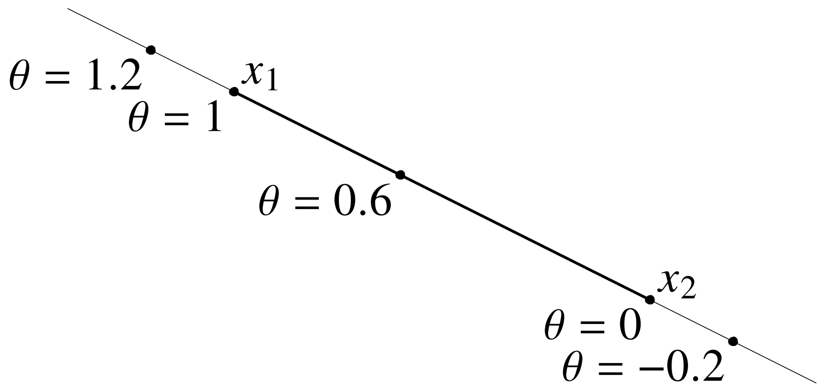
2. Convex sets

- 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

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



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

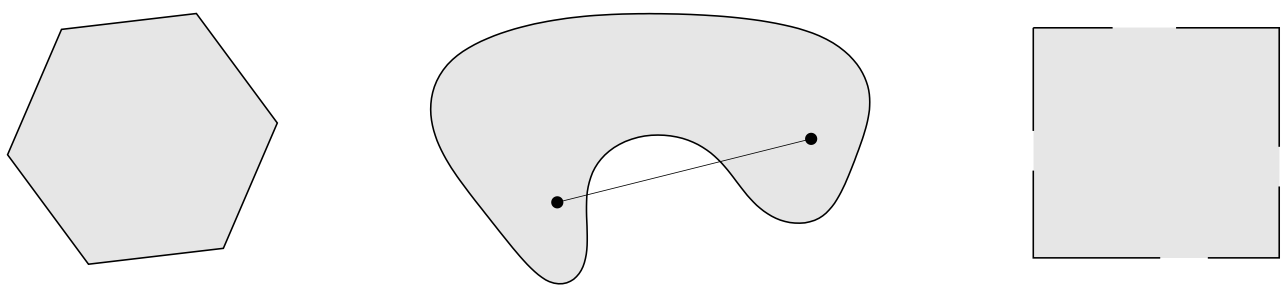
Line segment between points x_1, x_2: all points

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

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

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

Examples (one convex, two nonconvex sets)



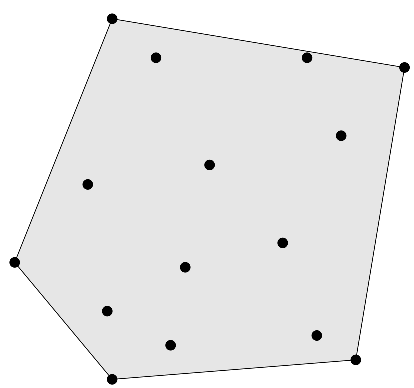
Convex combination and convex hull

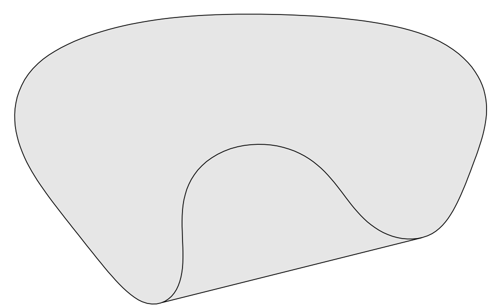
Convex combination of $x_1, \ldots, x_k$: any point x of the form

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

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

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

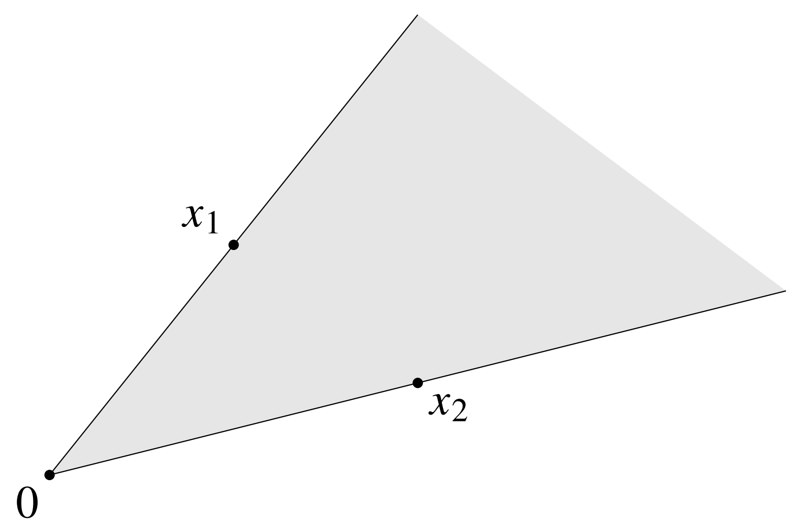




Convex cone

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

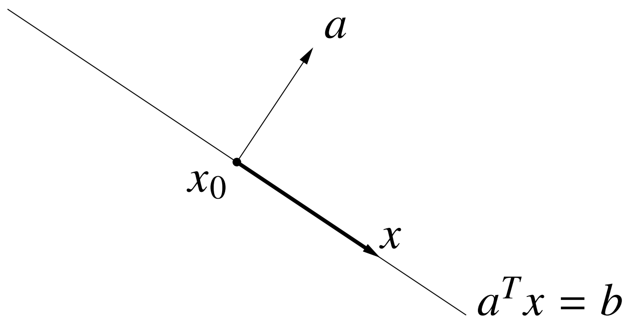
$$x = \theta_1 x_1 + \theta_2 x_2 \qquad \text{with } \theta_1 \geq 0,\ \theta_2 \geq 0$$



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

Hyperplanes and halfspaces

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

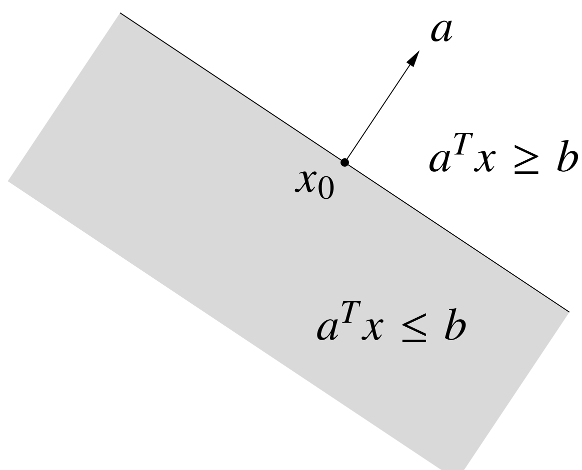


x_0 is a particular element, *e.g.*,

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

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

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



hyperplanes are affine and convex; halfspaces are convex

Euclidean balls and ellipsoids

(Euclidean) ball with center x_c and radius r:

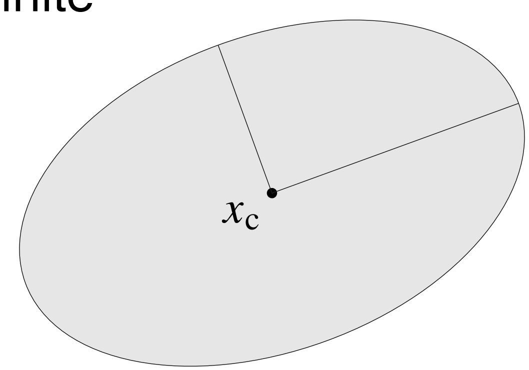
$$B(x_c, r) = \{x \mid \|x - x_c\|_2 \leq r\} = \{x_c + ru \mid \|u\|_2 \leq 1\}$$

$\|\cdot\|_2$ 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
- Λ is diagonal with diagonal elements $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n > 0$

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

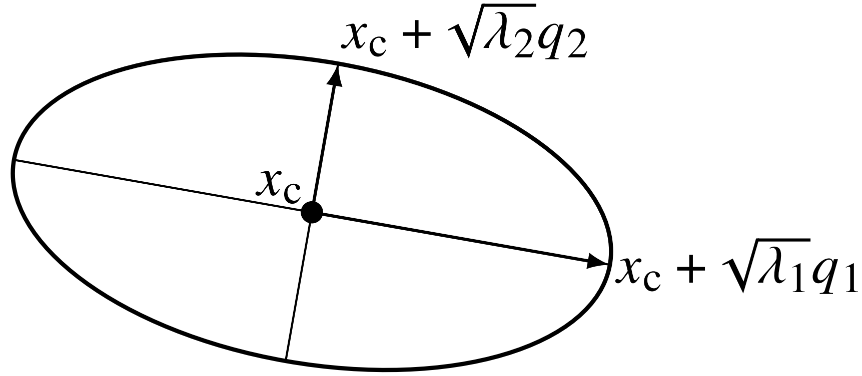
- after the change of variables the ellipsoid is described by

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

  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 $\mathbf{R}^2$



Exercise: give an interpretation of $\mathrm{tr}(P)$ 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 $t \in \mathbf{R}$
- $\|x + y\| \leq \|x\| + \|y\|$

Notation

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

Frequently used norms

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

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

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

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

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

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

Norm balls and norm cones

Norm ball with center x_c and radius r:

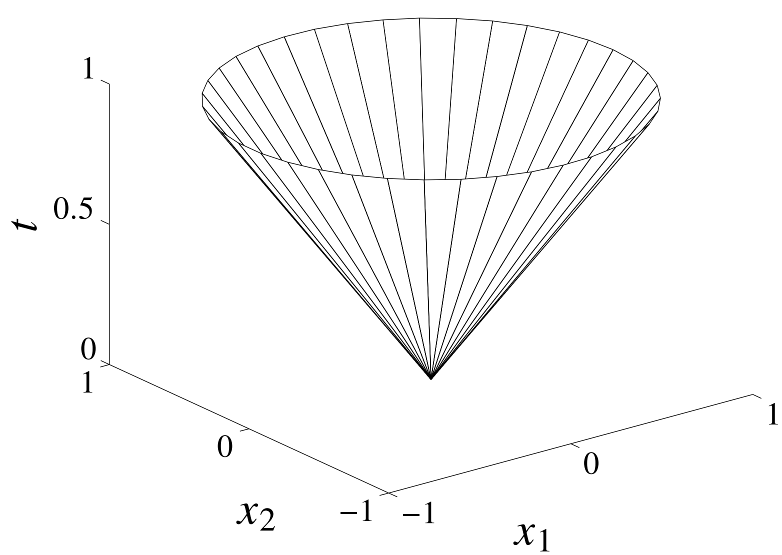
$$\{x \mid \|x - x_c\| \leq r\}$$

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)

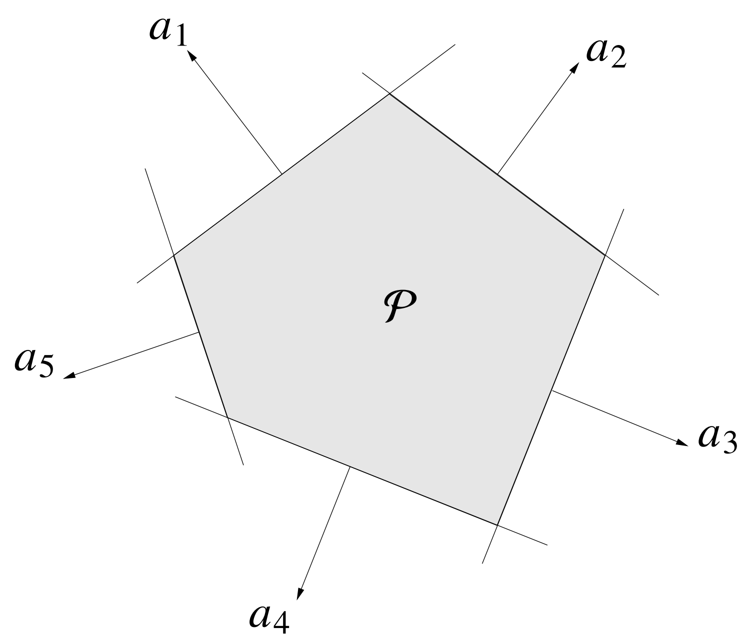


Polyhedra

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

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

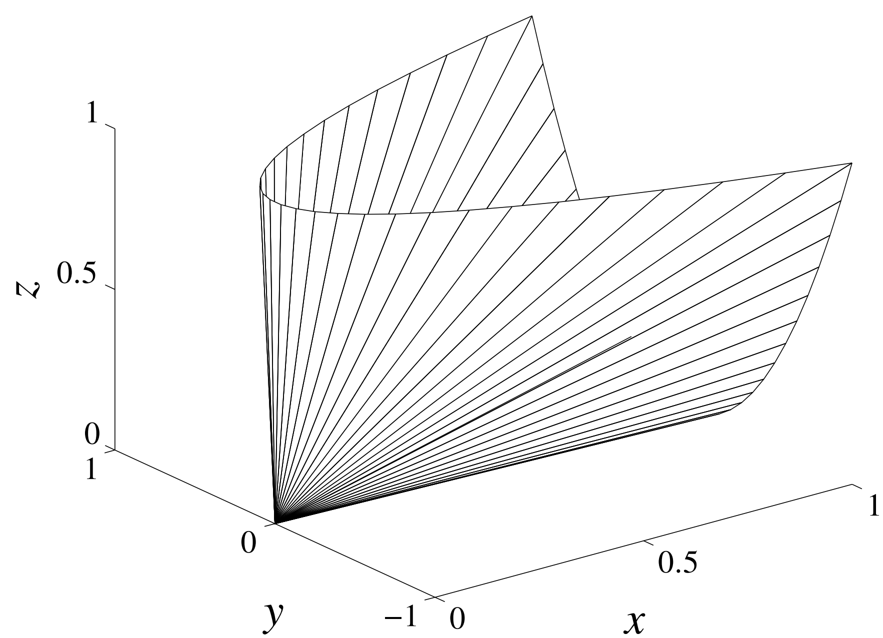
$$X \in \mathbf{S}^n_+ \quad \Longleftrightarrow \quad z^T X z \geq 0 \text{ for all } z$$

  $\mathbf{S}^n_+$ is a convex cone

- $\mathbf{S}^n_{++} = \{X \in \mathbf{S}^n \mid X \succ 0\}$: positive definite $n \times n$ matrices

Example

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



Operations that preserve convexity

methods for establishing convexity of a set C

1. apply definition

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

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

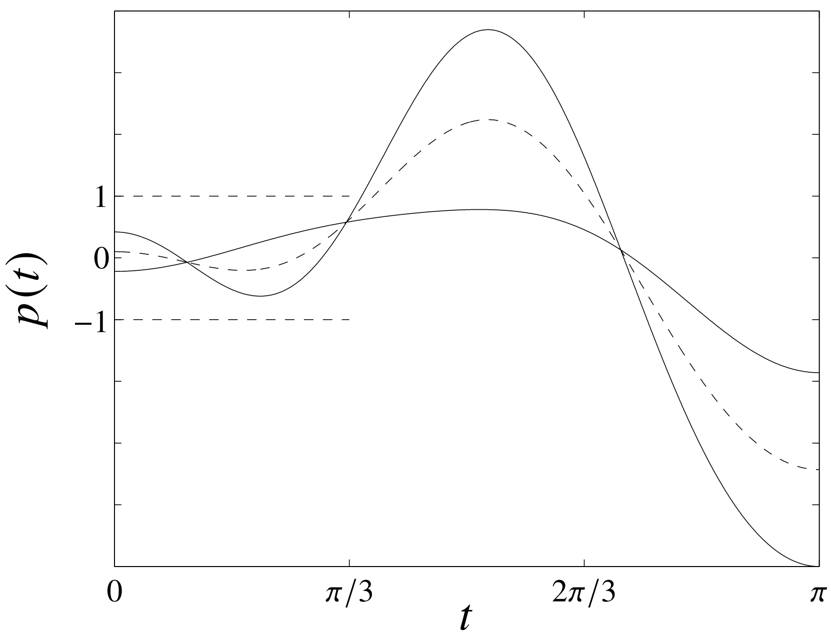
the intersection of (any number of) convex sets is convex

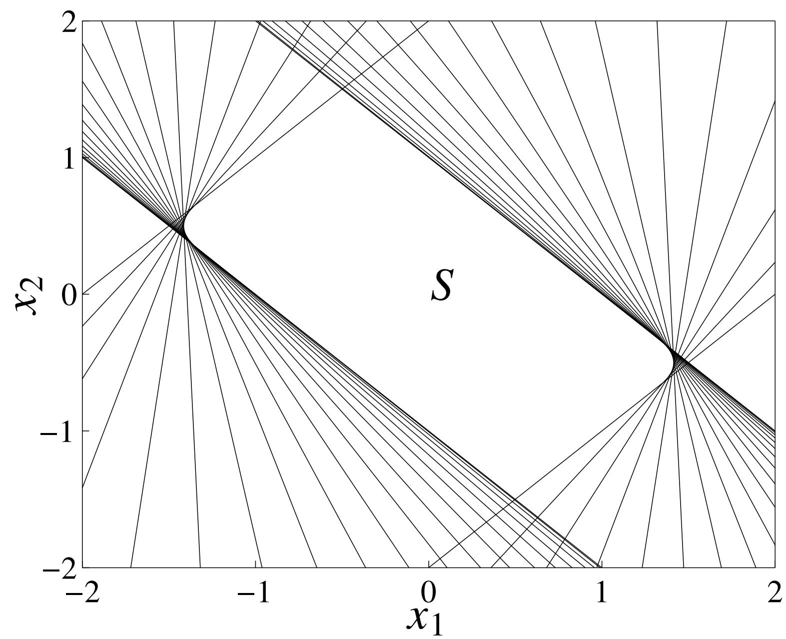
Example

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

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

for $m = 2$:





Affine function

suppose $f : \mathbf{R}^n \to \mathbf{R}^m$ is an affine function:

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

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

- the image of a convex set under f is convex

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

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

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

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

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

- solution set of linear matrix inequality

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

Perspective and linear-fractional function

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

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

images and inverse images of convex sets under perspective are convex

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

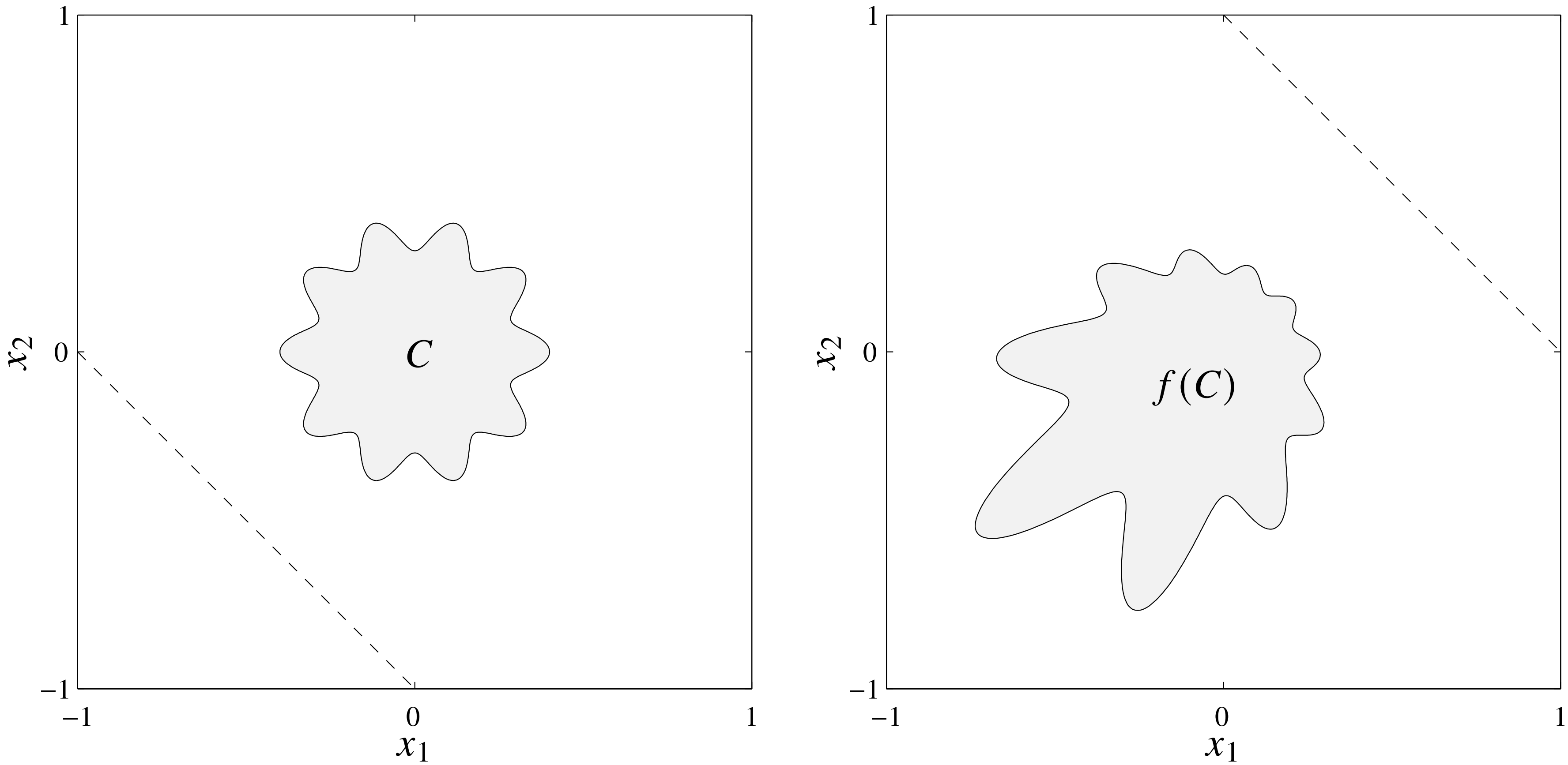
$$f(x) = \frac{Ax + b}{c^T x + d}, \qquad \mathbf{dom}\, f = \{x \mid c^T x + d > 0\}$$

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

Example

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

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



Proper cone

Proper cone: a convex cone $K \subseteq \mathbf{R}^n$ 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

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

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

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

Generalized inequality

Generalized inequality defined by a proper cone K:

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

Examples

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

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

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

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

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 $\mathbf{R}$, *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

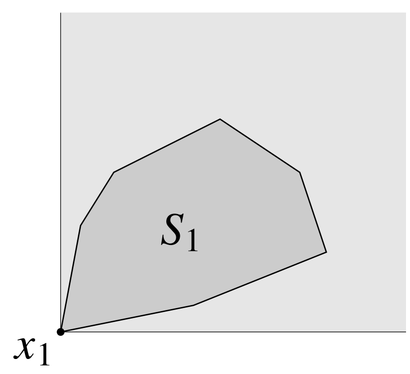
$$y \in S \quad \Longrightarrow \quad x \preceq_K y$$

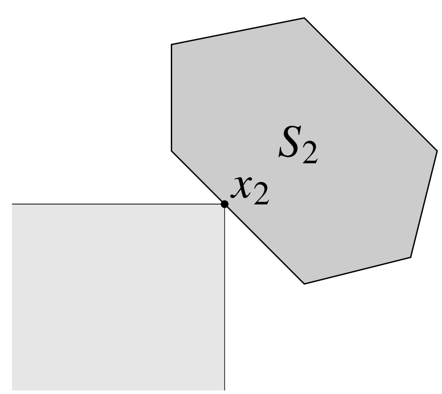
$x \in S$ is **a minimal element** of S with respect to $\preceq_K$ if

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

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

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 $x, y \in \mathbf{R}^n$:

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

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

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

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

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

Dual cones

Dual cone of a cone K:

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

note: definition depends on choice of inner product

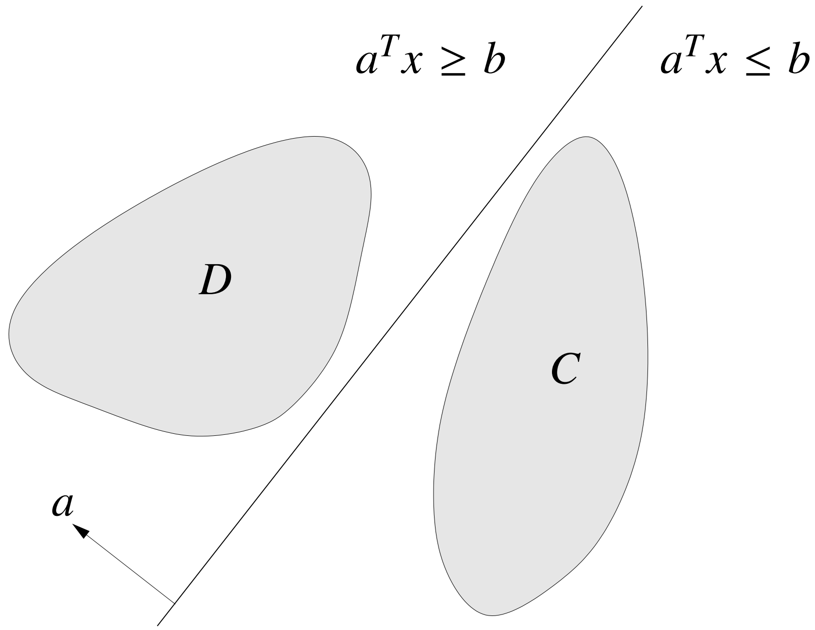
Examples

	K	K^*
nonnegative orthant	$\mathbf{R}^n_+$	$\mathbf{R}^n_+$
second order cone	$\{(x,t) \mid \|x\|_2 \leq t\}$	$\{(x,t) \mid \|x\|_2 \leq t\}$
1-norm cone	$\{(x,t) \mid \|x\|_1 \leq t\}$	$\{(x,t) \mid \|x\|_\infty \leq t\}$
positive semidefinite cone	$\mathbf{S}^n_+$	$\mathbf{S}^n_+$

Separating hyperplane theorem

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

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



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

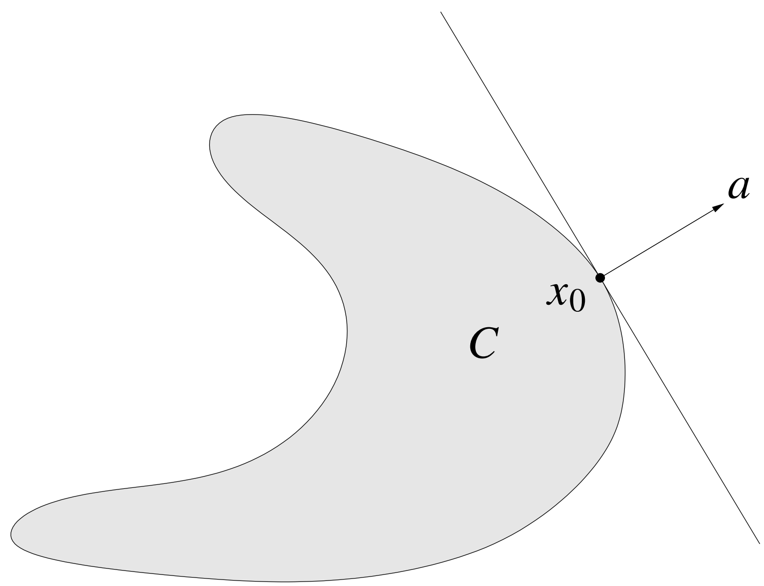
strict separation requires additional assumptions (*e.g.*, C closed, D a singleton)

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