1. Introduction

- mathematical optimization
- least-squares and linear programming
- convex optimization
- course goals and topics

Mathematical optimization

(mathematical) optimization problem

$$
\begin{array}{ll}
\text{minimize} & f_0(x) \\
\text{subject to} & f_i(x) \le b_i, \quad i = 1, \ldots, m
\end{array}
$$

- $x = (x_1, \ldots, x_n)$: optimization variables
- $f_0 : \mathbf{R}^n \to \mathbf{R}$: objective function
- $f_i : \mathbf{R}^n \to \mathbf{R}$, $i = 1, \ldots, m$: constraint functions

solution or **optimal point** $x^\star$ has smallest value of f_0 among all vectors that satisfy the constraints

Solving optimization problems

general optimization problem

- very difficult to solve
- methods involve some compromise, *e.g.*, very long computation time, or not always finding the solution (which may not matter in practice)

exceptions: certain problem classes can be solved efficiently and reliably

- least-squares problems
- linear programming problems
- convex optimization problems

- non-convex opt.

(SVD)

Least-squares

$\min_{w} \|Xw - y\|_2^2$

A x b

$$\text{minimize}_{x} \quad \|Ax - b\|_2^2$$

$x \in \mathbb{R}^n$

solving least-squares problems

$O(n^3)$ $O(n^2k)$ $O(nk)$

- analytical solution: $x^\star = (A^T A)^{-1} A^T b$
- reliable and efficient algorithms and software
- computation time proportional to $n^2 k$ ($A \in \mathbf{R}^{k \times n}$); less if structured
  ($k >> n$)
- a mature technology

$\begin{array}{cc} & n \\ k & \begin{bmatrix} A \end{bmatrix} \end{array}$

using least-squares

- least-squares problems are easy to recognize
- a few standard techniques increase flexibility ($e.g.$, including weights, adding regularization terms)

Linear programming

$$\begin{array}{ll} \text{minimize} & c^T x \\ \text{subject to} & a_i^T x \leq b_i, \quad i = 1, \ldots, m \end{array}$$

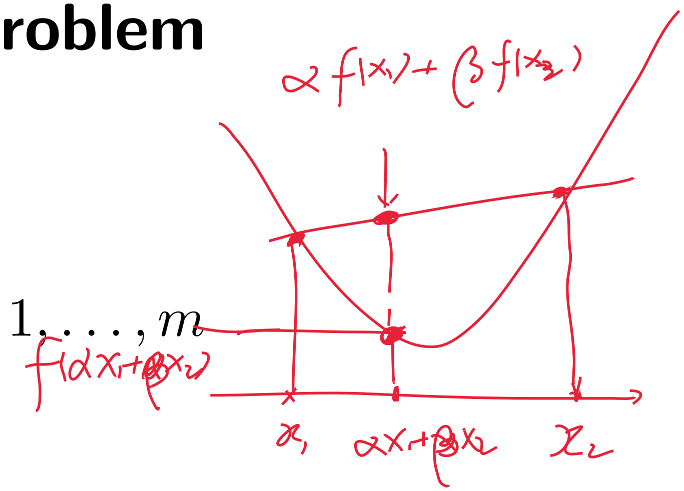
solving linear programs

- no analytical formula for solution
- reliable and efficient algorithms and software
- computation time proportional to $n^2 m$ if $m \geq n$; less with structure
- a mature technology

using linear programming

- not as easy to recognize as least-squares problems
- a few standard tricks used to convert problems into linear programs ($e.g.$, problems involving ℓ_1- or ℓ_∞-norms, piecewise-linear functions)

Convex optimization problem



$$\begin{array}{ll} \text{minimize} & f_0(x) \\ \text{subject to} & f_i(x) \leq b_i, \quad i = 1, \ldots, m \end{array}$$

- objective and constraint functions are convex:

$$f_i(\alpha x + \beta y) \leq \alpha f_i(x) + \beta f_i(y) \qquad (i = 0, 1, \ldots, m)$$

(if $\alpha + \beta = 1$, $\alpha \geq 0$, $\beta \geq 0$)

- includes least-squares problems and linear programs as special cases

solving convex optimization problems

- no analytical solution
- reliable and efficient algorithms
- computation time (roughly) proportional to $\max\{n^3, n^2m, F\}$, where F is cost of evaluating f_i's and their first and second derivatives
- almost a technology

using convex optimization

- often difficult to recognize
- many tricks for transforming problems into convex form
- surprisingly many problems can be solved via convex optimization

2. Convex sets

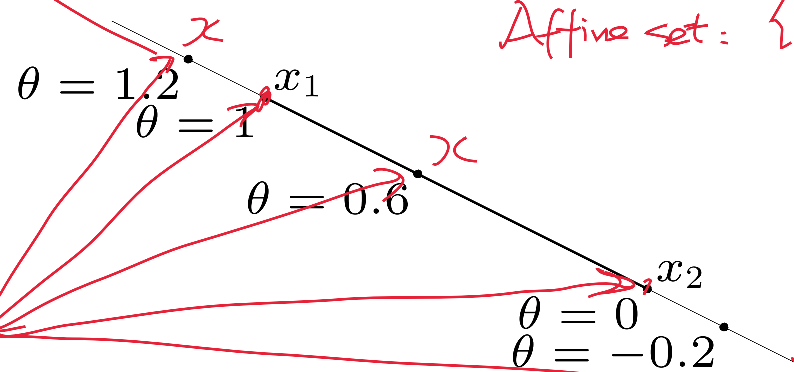
- affine and convex sets
- some important examples
- operations that preserve convexity
- generalized inequalities
- separating and supporting hyperplanes
- dual cones and generalized inequalities

Affine set

line through x_1, x_2: all points

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

Affine set: $\{x \mid x = \theta x_1 + (1-\theta)x_2,\ \theta \in R\}$



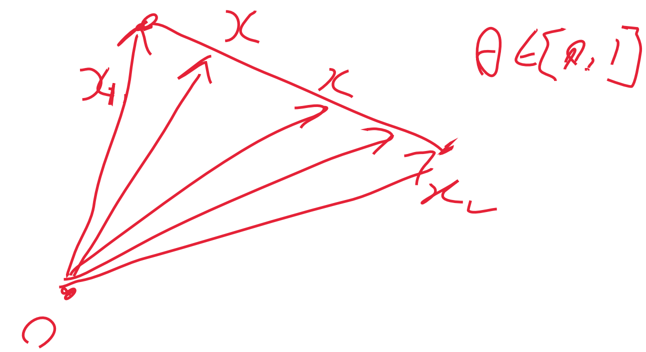
- **affine set**: contains the line through any two distinct points in the set
- **example**: solution set of linear equations $\{x \mid Ax = b\}$

  $f(x) = Ax - b$

  affine func..

  (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

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

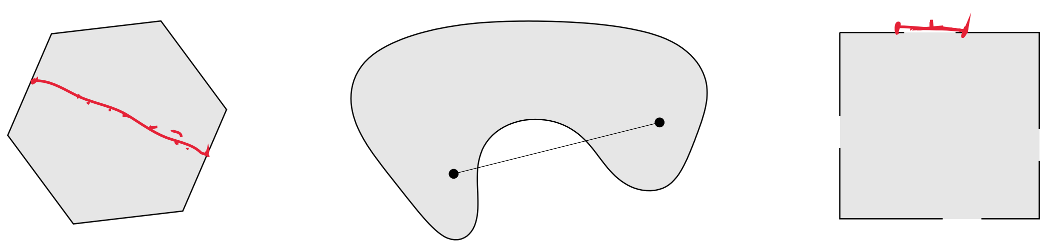
with $0 \le \theta \le 1$

$\{x \mid x = \theta x_1 + (1-\theta) x_2,\ \theta \in [0,1]\}$

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)



$S = \{x_1, x_2, x_3\}$

Convex combination and convex hull

conv S

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

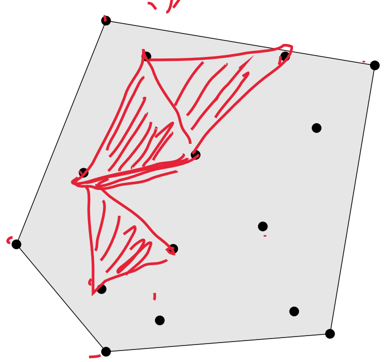
$$x = \theta_1 x_1 + \theta_2 x_2 + \cdots + \theta_k x_k = \sum_{k=1}^{k} \theta_k x_k$$

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

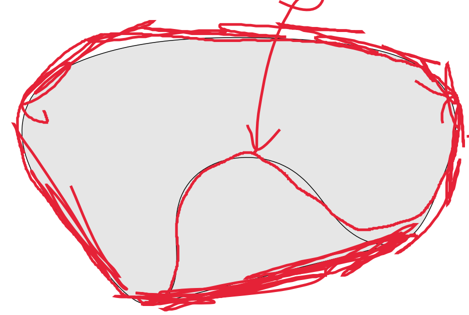
convex hull $\mathbf{conv}\, S$: set of all convex combinations of points in S

$S = \{\ldots\ldots\}$

conv S



S



conv S

Convex cone

$x_1, x_2 \in C$

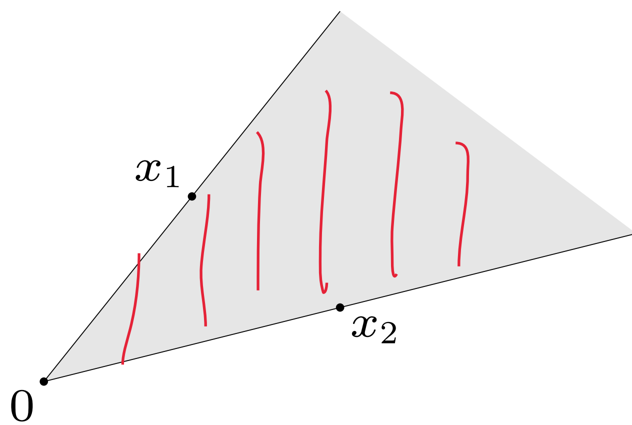
$x = \theta_1 x_1 + \theta_2 x_2 \in C$

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

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

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

affine	convex	cone
$\theta_1, \theta_2 \in \mathbb{R}$	$\theta_1 + \theta_2 = 1$ $\theta_1, \theta_2 \geq 0$	$\theta_1 \geq 0$, $\theta_2 \geq 0$



$\boxed{\text{affine} \Rightarrow \text{convex}}$

convex $\not\Rightarrow$ affine

cone $\Rightarrow$ convex

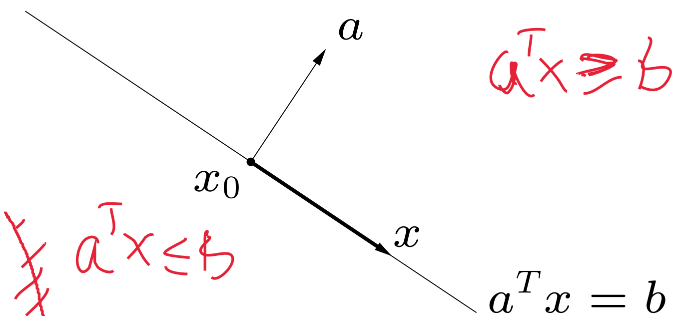
convex $\not\Rightarrow$ cone

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

$x \in \mathbb{R}^n$

Hyperplanes and halfspaces

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

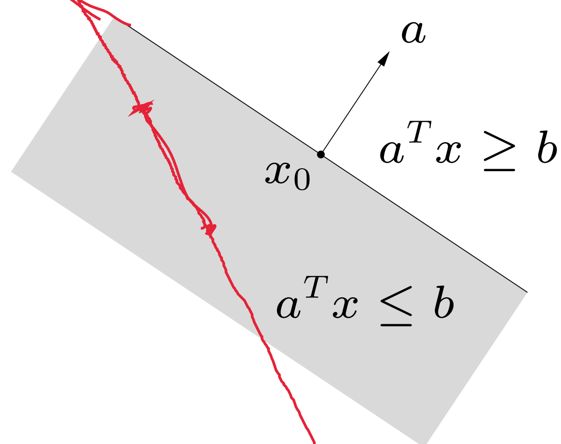


$a^T x \geq b$

$a^T x \leq b$

$\boxed{\text{convex affine}}$

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



$\boxed{\text{convex not affine}}$

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

$B(x_c, r)$:

Euclidean balls and ellipsoids

(Euclidean) ball with center x_c and radius r:

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

$(x - x_c)^T \left(\frac{1}{r^2} I\right)(x - x_c) \le 1$ $\qquad \frac{1}{r^2} \cdot I$

ellipsoid: set of the form

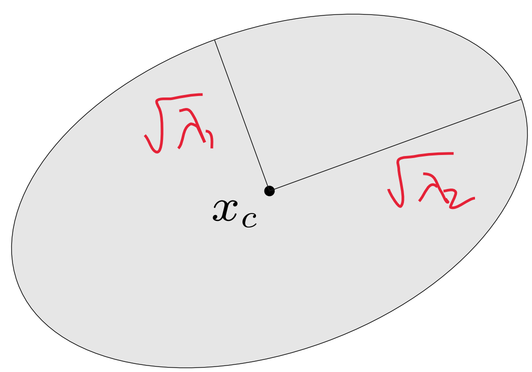
$$\{x \mid (x - x_c)^T P^{-1} (x - x_c) \le 1\} \qquad P$$

$P \in S^2_{++}$

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

① $\lambda_1, \lambda_2 > 0$

② $z^T P z > 0, \ \forall z$



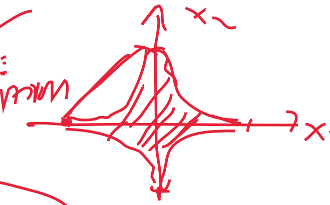
other representation: $\{x_c + Au \mid \|u\|_2 \le 1\}$ with A square and nonsingular

$f(x) = \|x\|$:

$(f: \mathbb{R}^n \to \mathbb{R})$

$\|x\|_p = \left(\sum_{i=1}^{n} |x_i|^p\right)^{\frac{1}{p}}$ $(p \geq 1)$

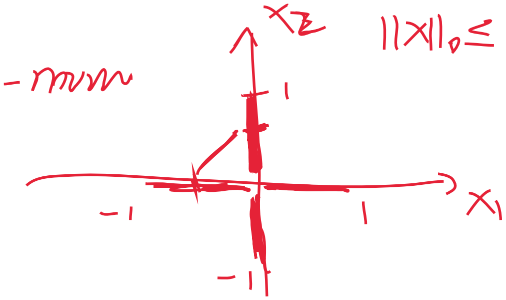
$\ell_{\frac{1}{2}}$-norm



$0 \leq p < 1$

ℓ_0-norm

$\|x\|_0 \leq 1$



Norm balls and norm cones

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

- 半正定性 $\|x\| \geq 0$; $\|x\| = 0$ if and only if $x = 0$
- 齐次性 $\|tx\| = |t| \, \|x\|$ for $t \in \mathbf{R}$
- 次可加性 $\|x + y\| \leq \|x\| + \|y\|$

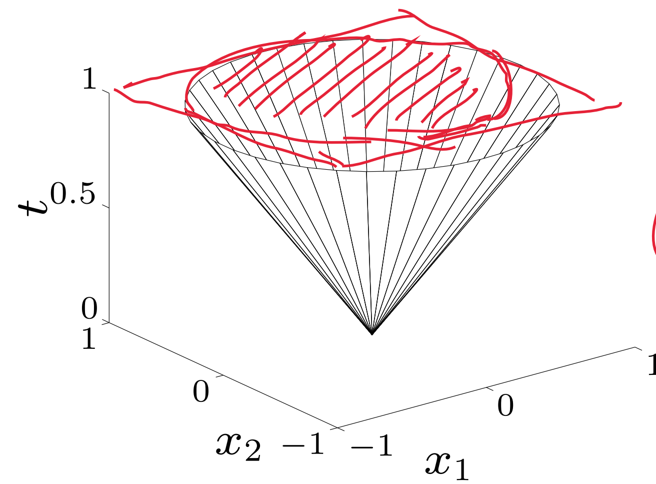
notation: $\|\cdot\|$ is general (unspecified) norm; $\|\cdot\|_{\text{symb}}$ is particular norm

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

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

ℓ_2-norm

Euclidean norm cone is called second-order cone



(ice-cream cone)

norm balls and cones are convex

$a = \begin{bmatrix} a_1 \\ a_2 \\ a_3 \end{bmatrix}$ $b = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix}$

$a \preceq b \Downarrow a_1 \le b_1,\ a_2 \le b_2,\ a_3 \le b_3$

Polyhedra

$\{x \mid Ax \preceq b,\ Cx = d\}$

solution set of finitely many linear inequalities and equalities

halfspace hyperplane

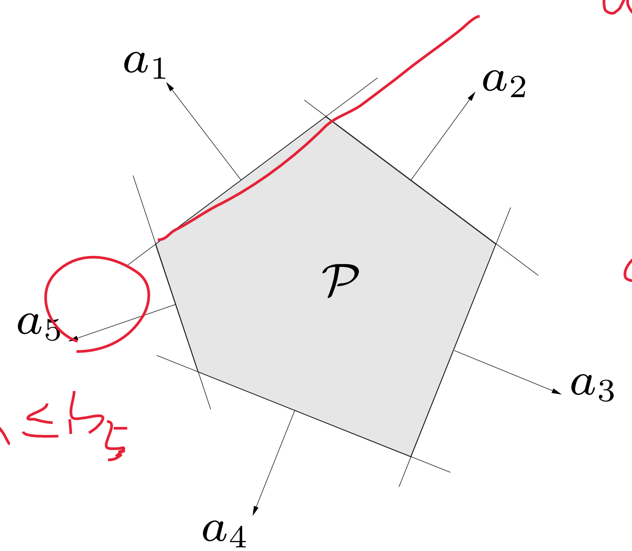
$x \in \mathbf{R}^{n\times 1}$

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

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

$a_1^T x \le b_1$

$$\begin{aligned} \min_x \quad & c^T x \\ \text{s.t.} \quad & Ax \le b \\ & Cx = d \end{aligned}$$



$c^T x = d$

$a_5^T x \le b_5$

convex

polyhedron is intersection of finite number of halfspaces and hyperplanes

$x_1, x_2 \in C$

$\theta x_1 + (1-\theta) x_2 \in C$ $(\theta \geq 0)$

(PSD)

Positive semidefinite cone

$X \succeq 0 \iff X$ is PSD

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

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

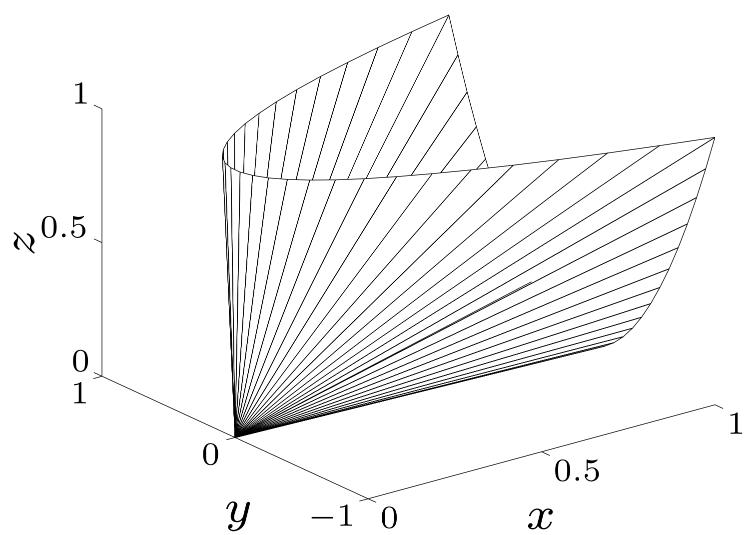
$\mathbf{S}^n_+$ is a convex cone

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

$X, Y \in S^n_+$

$\theta X + (1-\theta) Y \in S^n_+$?

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



$z^T X z \geq 0$
$z^T Y z \geq 0$

$z^T(\theta X + (1-\theta) Y) z$

$= \theta z^T X z + (1-\theta) z^T Y z$

≥ 0

$\Rightarrow \theta X + (1-\theta) Y \in S^n_+$

Operations that preserve convexity

practical 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

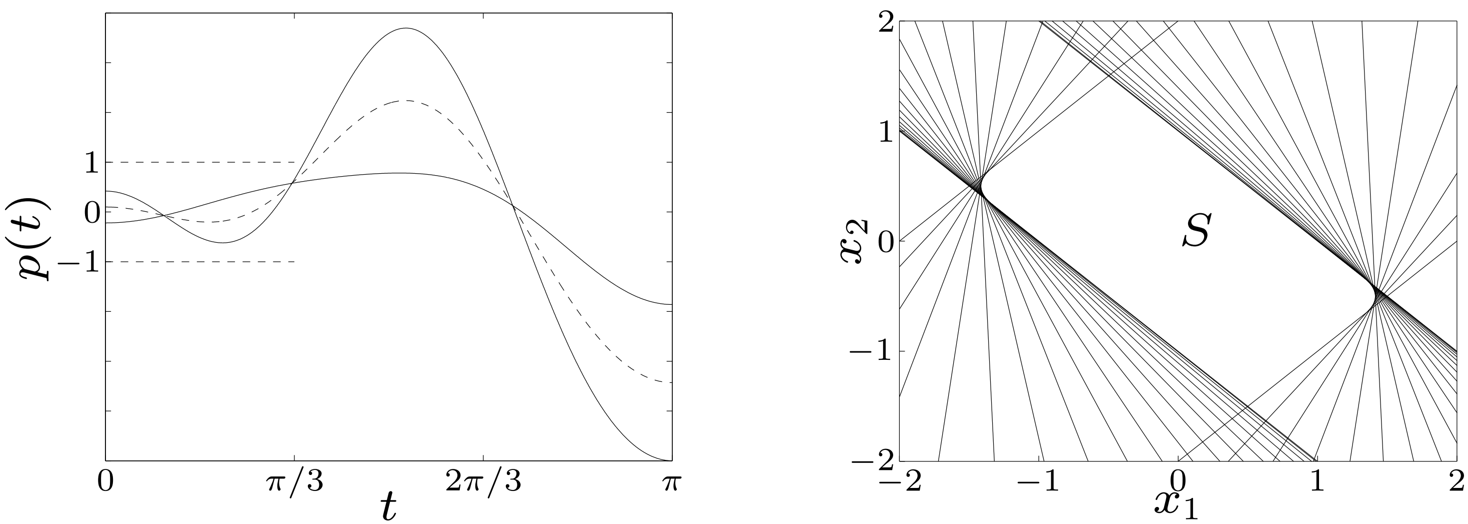
- the intersection of (any number of) convex sets is convex   Polyhedra.

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 affine ($f(x) = Ax + b$ with $A \in \mathbf{R}^{m\times n}$, $b \in \mathbf{R}^m$)

- the image of a convex set under f is convex

$$S \subseteq \mathbf{R}^n \text{ convex} \implies f(S) = \{f(x) \mid x \in S\} \text{ convex}$$

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

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

$A(x) \preceq B$

$\Rightarrow B - A(x) \succeq 0$

(PSD)

examples

$\{\alpha x \mid x \in S\}$ $\{x + \alpha \mid x \in S\}$ $\{x_1 \mid (x_1, x_2) \in C\}$

- scaling, translation, projection

$A(x)$

Convex
- solution set of linear matrix inequality $\{x \mid x_1A_1 + \cdots + x_mA_m \preceq B\}$ (with $A_i, B \in \mathbf{S}^p$)

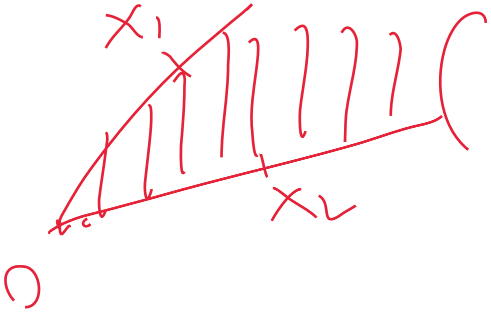
(LMI)

- hyperbolic cone $\{x \mid x^TPx \le (c^Tx)^2,\ c^Tx \ge 0\}$ (with $P \in \mathbf{S}^n_+$)

$f(x) = B - A(x)$: affine func.

$f^{-1}(S^n_+) = \{x \in \mathbb{R}^n \mid B - A(x) \succeq 0\}$

$A(x) \preceq B$



$(x = \theta_1 x_1 + \theta_2 x_2$, $\theta_1, \theta_2 \geq 0)$

C:

boundary
bd
interior
int

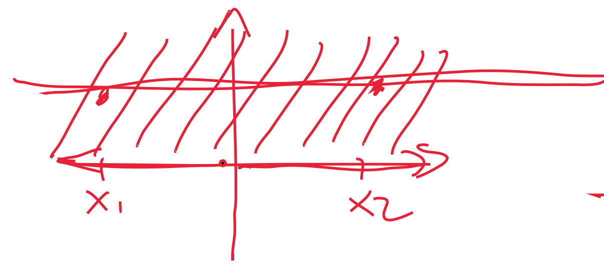
Generalized inequalities

$C = \mathrm{bd} \cup \mathrm{int}$

$\|x\|_2 < 1$

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

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



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

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 defined by a proper cone K:

$$x \preceq_K y \iff y - x \in K, \qquad x \prec_K y \iff y - x \in \textbf{int}\, K$$

$(y - x \succeq_K 0)$.

y

examples

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

$y - x \in \mathbf{R}^n_+$

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

K

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

PSD

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

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

properties: many properties of $\preceq_K$ are similar to $\le$ on $\mathbf{R}$, *e.g.*,

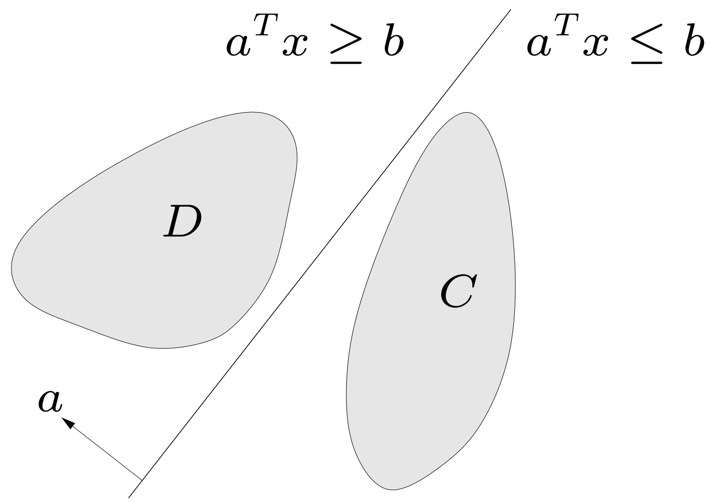
$$x \preceq_K y, \quad u \preceq_K v \implies x + u \preceq_K y + v$$

Separating hyperplane theorem

$C, D \notin \phi$ $C \cap D = \phi$

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

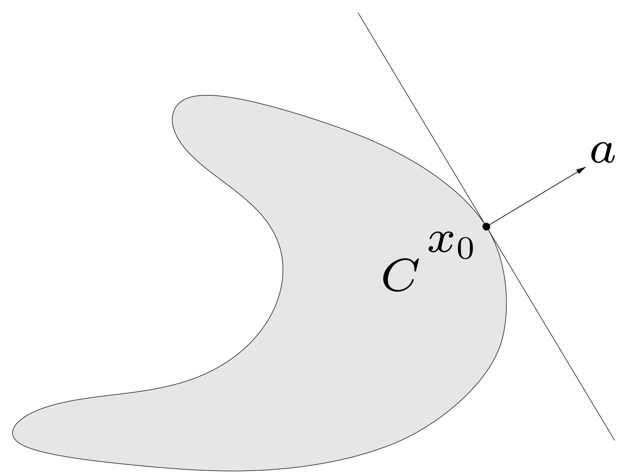
strict separation requires additional assumptions ($e.g.$, C is closed, D is a singleton)

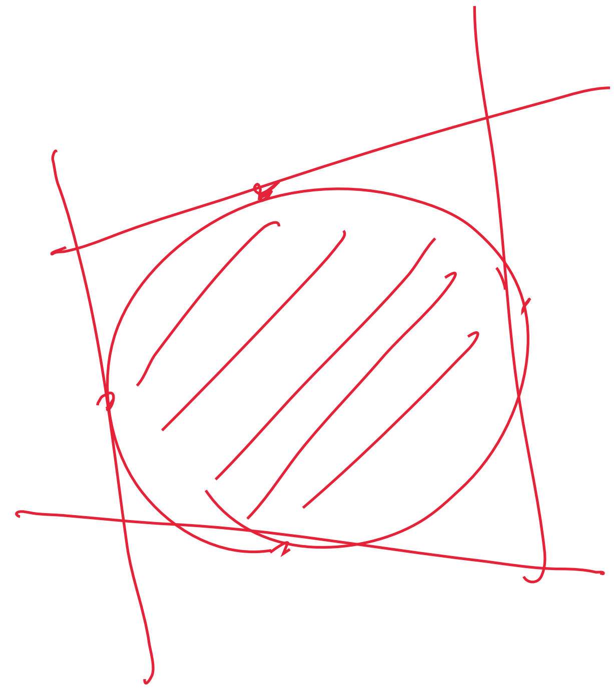
Supporting hyperplane theorem

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

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

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

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

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