

1. Introduction

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

Mathematical optimization

(mathematical) optimization problem

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

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

solution or **optimal point** x^* 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

Least-squares

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

solving least-squares problems

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

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

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

solving linear programs

- no analytical formula for solution
- reliable and efficient algorithms and software
- computation time proportional to n^2m 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

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

- objective and constraint functions are convex:

$$f_i(\alpha x + \beta y) \leq \alpha f_i(x) + \beta f_i(y)$$

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

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



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 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: contains line segment between any two points in the set

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

examples (one convex, two nonconvex sets)



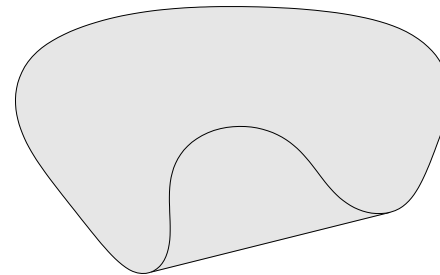
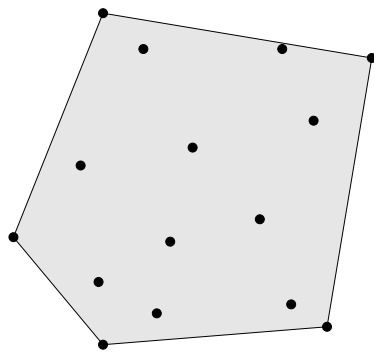
Convex combination and convex hull

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

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

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

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

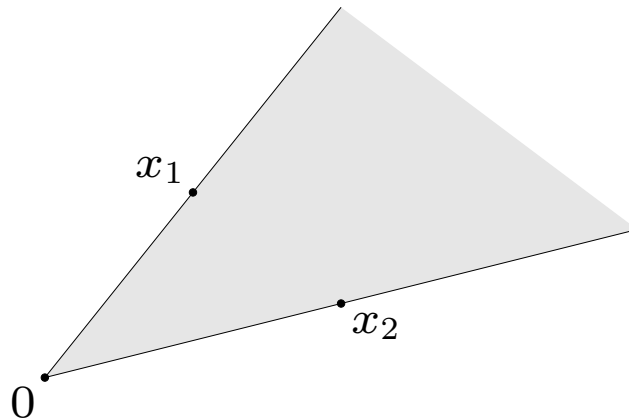


Convex cone

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

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

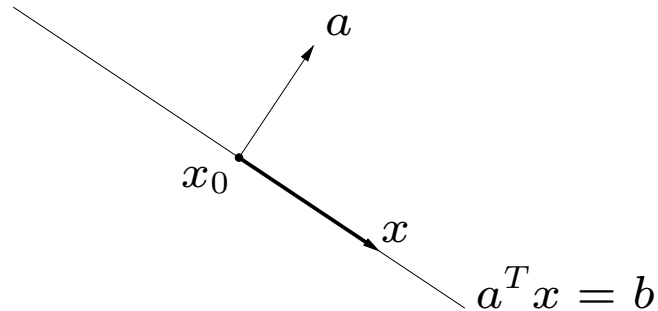
with $\theta_1 \geq 0$, $\theta_2 \geq 0$



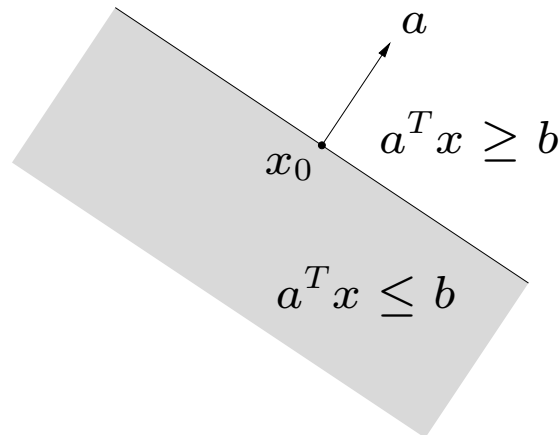
convex cone: set that contains all conic combinations of points in the set

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

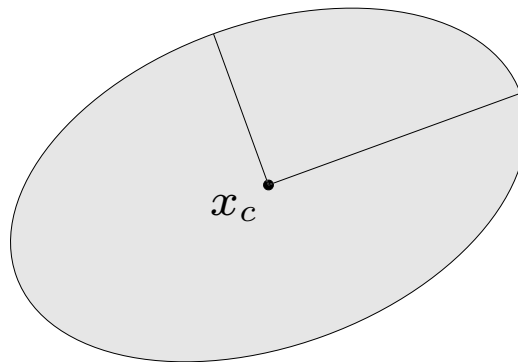
(Euclidean) ball with center x_c and radius r :

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

ellipsoid: set of the form

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

with $P \in \mathbf{S}_{++}^n$ (*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

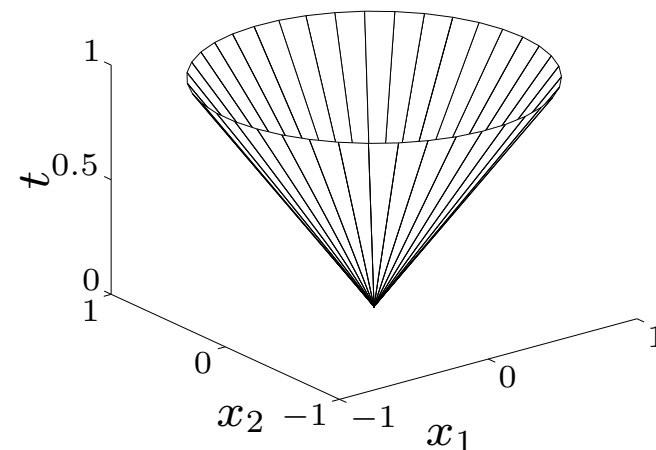
- $\|x\| \geq 0$; $\|x\| = 0$ if and only if $x = 0$
- $\|tx\| = |t| \|x\|$ for $t \in \mathbf{R}$
- $\|x + y\| \leq \|x\| + \|y\|$

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



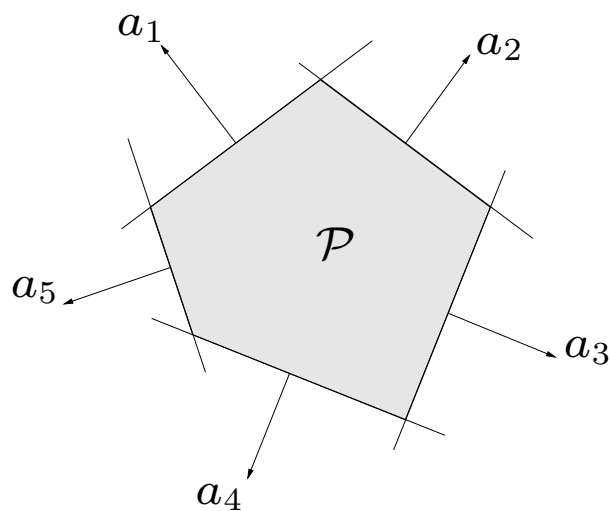
norm balls and cones are convex

Polyhedra

solution set of finitely many linear inequalities and equalities

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

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



polyhedron is intersection of finite number of halfspaces and hyperplanes

Positive semidefinite cone

notation:

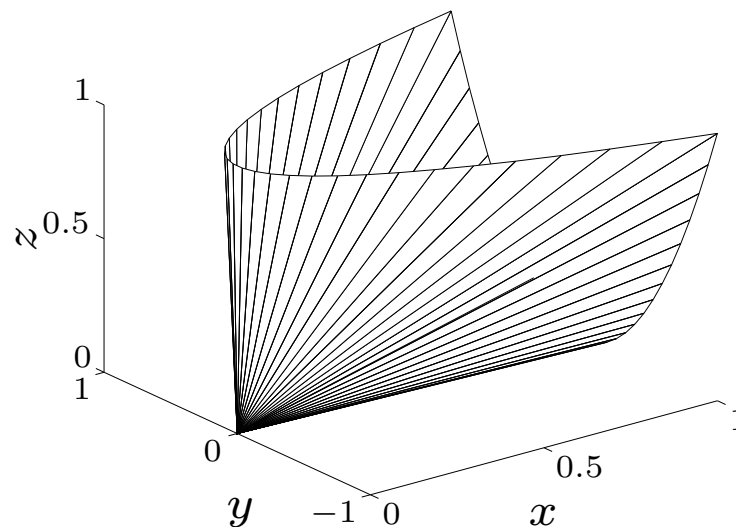
- \mathbf{S}^n is set of symmetric $n \times n$ matrices
- $\mathbf{S}_+^n = \{X \in \mathbf{S}^n \mid X \succeq 0\}$: positive semidefinite $n \times n$ matrices

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

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



Operations that preserve convexity

practical methods for establishing convexity of a set C

1. apply definition

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

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

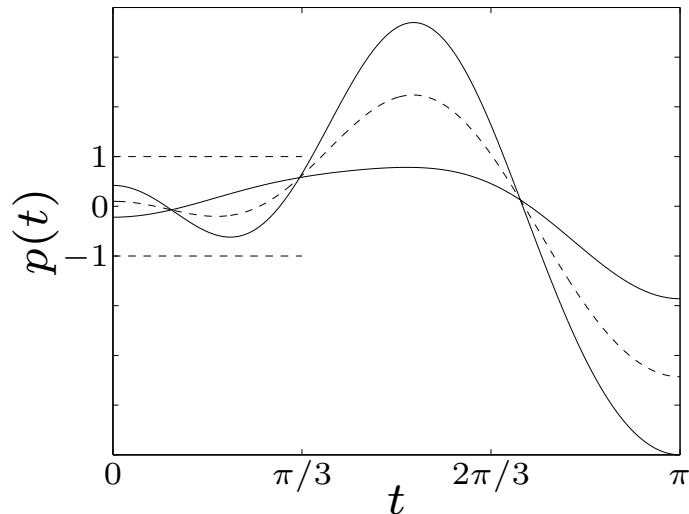
the intersection of (any number of) convex sets is convex

example:

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

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

for $m = 2$:



Affine function

suppose $f : \mathbf{R}^n \rightarrow \mathbf{R}^m$ 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

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

examples

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

Generalized inequalities

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)

examples

- nonnegative orthant $K = \mathbf{R}_+^n = \{x \in \mathbf{R}^n \mid x_i \geq 0, i = 1, \dots, n\}$
- positive semidefinite cone $K = \mathbf{S}_+^n$
- nonnegative polynomials on $[0, 1]$:

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

generalized inequality defined by a proper cone K :

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

examples

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

$$x \preceq_{\mathbf{R}_+^n} y \iff x_i \leq y_i, \quad i = 1, \dots, n$$

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

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

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

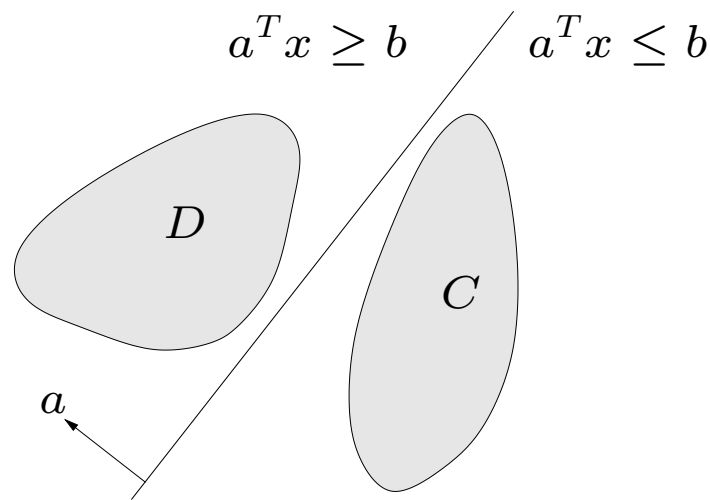
properties: many properties of \preceq_K are similar to \leq on \mathbf{R} , *e.g.*,

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

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

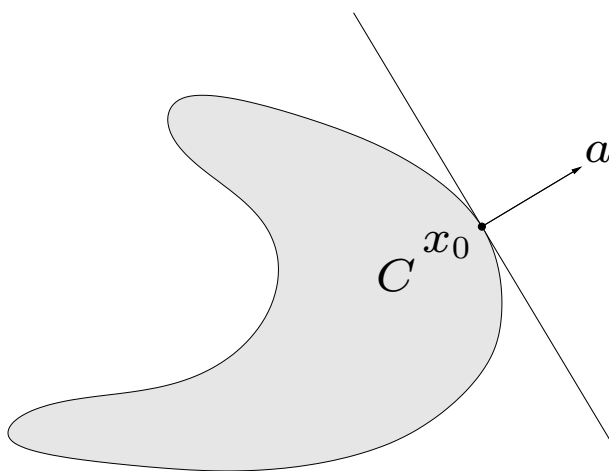
strict separation requires additional assumptions (*e.g.*, C is closed, D is a singleton)

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:

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