EE353: Network Optimization
Chapter 2: Mathematical Tools

Yiyin Wang
yiyinwang@sjtu.edu.cn
Department of Automation
Shanghai Jiao Tong University

Oct. 11, 2014

Table of contents

Acknowledgements

Slides borrow heavily from lectures by Stephen Boyd (Stanford), Lieven Vandenberghe (UCLA), and Mung Chiang (Princeton)

Mathematical optimization

A (mathematical) optimization problem

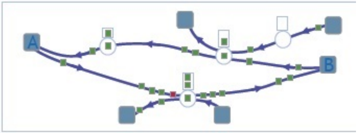
$$\begin{array}{ll} \text{minimize} & f_0(x) \\ \text{subject to} & f_i(x) \leq 0, i = 1, \ldots, m \end{array}$$

- $x = [x_1, \ldots, x_n]^T$: 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

Optimal solution x^* has smallest value of f_0 among all vectors that satisfy the constraints

Example: routing in data networks

Find best routes for forwarding traffic from source to destination



If we view data traffic as a continuous flow, optimal routing solves

$$
\begin{array}{ll}
\text{minimize} & \displaystyle\sum_{l=1}^{L} f_l(t_l) \\
\text{subject to} & \displaystyle\sum_{l\in\mathcal{I}(n)} \alpha_{pl} - \sum_{l\in\mathcal{O}(n)} \alpha_{pl} = e_{pn} \quad \text{flow conservation at nodes} \\
& \displaystyle t_l = \sum_{p} \alpha_{pl} s_p \quad \text{total traffic across links} \\
& t_l \leq c_l \quad \text{capacity constraints} \\
& \alpha_{pl} \geq 0 \quad \text{traffic is non}-\text{negative}
\end{array}
$$

Convex optimization problem if loss functions $f_l(t_l)$ are convex

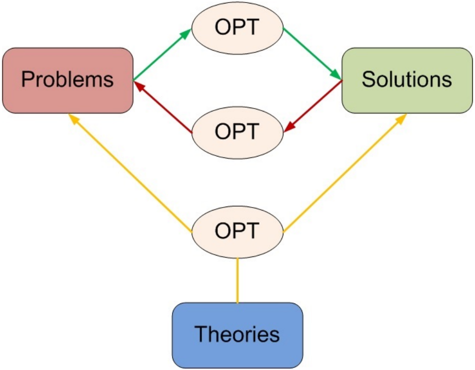
Applications

Theory and algorithms of optimization are extremely powerful:

- Information science areas: communications systems, signal/image/vedio processing, systems control, graphics, data analysis, theoretical computer science ...
- Other engineering disciplines: aerospace, mechanical, transportation, computer architecture, analog circuit design ...
- Physics, chemistry, biology ...
- Economics, finance, management ...
- Analysis, probability, statistics, differential equations ...

A powerful and widely-applicable perspective

- Traditional view: optimization is a tool
- Modern view: optimization is a principle package



Solving optimization problems

General optimization problem

- Very difficult to solve
- Methods involve some compromise, e.g., very long computation time, or not always finding solution

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

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

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

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

- Objective and constraint functions are convex:

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

  if $\alpha + \beta = 1, \alpha \geq 0, \beta \geq 0$
- Includes least-squares problems and linear programs as special cases

Convex optimization problem

Solving convex optimization problems

- No analytical solution
- Reliable and efficient algorithms
- Computation time (roughly) proportional to $\max\{n^3, n^2 m, F\}$, 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

Why does convexity matter?

"The great watershed in optimization isn't between linearity and nonlinearity, but convexity and nonconvexity", R. Rockafellar, SIAM Review 1993.

Recognize and Utilize Convexity!

Part 1: Fundamentals of convex optimization

- 2.1 Convex sets, convex functions, and convex problems
- 2.2 Duality

2.1.1 Convex sets

- Affine and convex sets
- Some important examples
- Operations that preserve convexity

Affine set

- Line through two points x_1, x_2: all points

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

- Line segment between x_1 and x_2: all points

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

- Affine set contains the line through any two distinct points in the set
- Example: solution set of linear equations $\{x|Ax = b\}$

Convex set and convex combination

- Convex set C contains line segment between any two points in the set

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

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

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

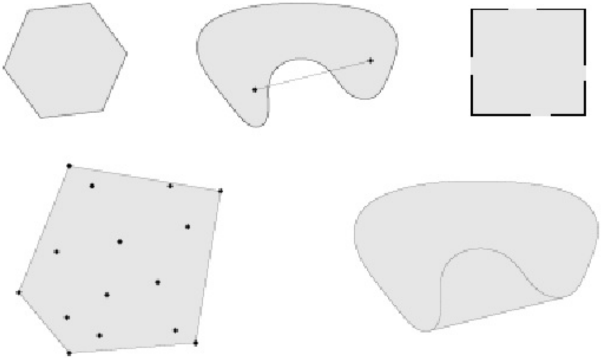
with $\theta_1 + \cdots + \theta_k = 1, \theta_i \geq 0$

Convex hull and examples

- Convex hull of a set C, denoted $\mathrm{conv}\, C$, is the set of all convex combinations of points in C

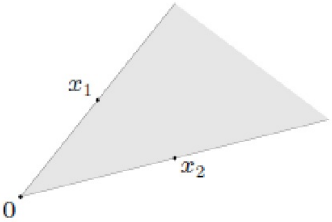
$$\mathrm{conv}\, C = \left\{ \sum_{i=1}^{k} \theta_i x_i | x_i \in C, \theta_i \geq 0, i, = 1, \ldots, k, \sum_{i=1}^{k} \theta_i = 1 \right\}$$

- Examples:



Convex cone

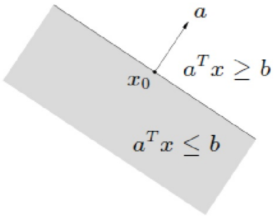
- Conic (nonnegative) combination of x_1 and x_2: any point of the form
$$x = \theta_1 x_1 + \theta_2 x_2$$
with $\theta_1 \geq 0, \theta_2 \geq 0$



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

Hyperplane and halfspaces

- Hyperplane in $\mathbf{R}^n$ is a set: $\{x \mid a^T x = b\}$ where $a \in \mathbf{R}^n, a \neq 0, b \in \mathbf{R}$. Divides $\mathbf{R}^n$ into two halfspaces: eg. $\{x \mid a^T x \leq b\}$ and $\{x \mid a^T x > b\}$

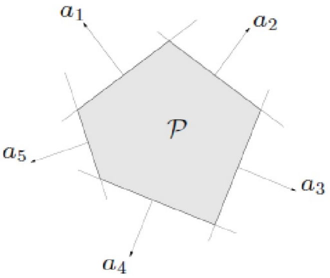


- Hyperplanes are affine and convex; halfspaces are convex

Polyhedra

- Polyhedra is the solution set of a finite number of linear equalities and inequalities (intersection of finite number of haldspaces and hyperplanes)

$$Ax \preceq b, \quad Cx = d$$
$$A \in \mathbf{R}^{m\times n},\ C \in \mathbf{R}^{p\times n}$$
$$\preceq \ \text{is componentwise inequality}$$



Euclidean ball and ellipsoid

- Euclidean ball in $\mathbf{R}^n$ with center x_c and radius r

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

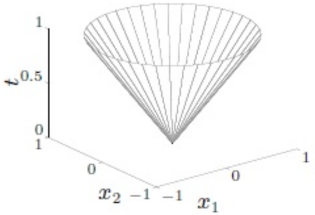
- Generalize to ellipsoids

$$\begin{aligned} \mathcal{E}(x_c, P) &= \{x \mid (x - x_c)^T P^{-1}(x - x_c) \leq 1\} \\ &= \{x_c + Au \mid \|u\|_2 \leq 1\} \end{aligned}$$

  with $P \in \mathbf{S}^n_{++}$ (P is symmetric positive definite) and A square and nonsingular
- Lengths of semi-axes of $\mathcal{E}$ are $\sqrt{\lambda_i}$ where λ_i are eigenvalues of P

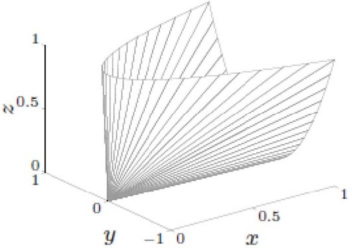
Norm ball and cone

- Norm: a function $\|\cdot\|$ that satisfies
  - $\|x\| \geq 0$; $\|x\| = 0$ if and only if $x = 0$
  - $\|tx\| = |t|\|x\|$ for $t \in \mathcal{R}$
  - $\|x + y\| \leq \|x\| + \|y\|$
- Norm ball with center x_c and radius r: $\{x | \|x - x_c\| \leq r\}$
- Norm cone: $\{(x, t) | \|x\| \leq t\}$
- Norm balls and cones are convex



Positive semidefinite cone

- $\mathbf{S}^n$ is a set of symmetric $n \times n$ matrices
- $\mathbf{S}^n_+ = \{X \in \mathbf{S}^n | X \succeq 0\}$: positive semidefinite $n \times n$ matrices
- $\mathbf{S}^n_{++} = \{X \in \mathbf{S}^n | X \succ 0\}$: positive definite $n \times n$ matrices
- Example $\begin{bmatrix} x & y \\ y & z \end{bmatrix} \in \mathbf{S}^2_+$



Convexity-Preserving Operations

Practical methods for establishing convexity of a set C

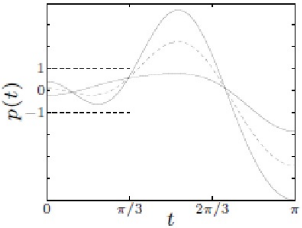
1. Apply definition

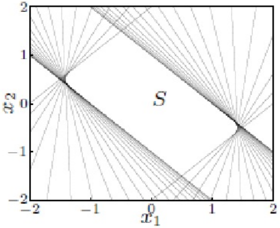
$$x_1, x_2 \in C,\ 0 \leq \theta \leq 1 \ \Rightarrow\ \theta x_1 + (1-\theta)x_2 \in C$$

2. Show that C is obtained from simple convex sets by operations that preserve convexity
   - Intersection
   - Affine functions
   - Perspective function
   - Linear-fraction 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 $p(t) = \sum_{k=1}^{m} x_k \cos kt$ for $m = 2$





Affine function

- Suppose $f : \mathbf{R}^n \to \mathbf{R}^m$ is affine

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

- The image of a convex set under f is convex

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

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

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

2.1.2 Convex functions

- Basic properties and examples
- Operations that preserve convexity
- Important examples

Definition of convex functions

$f : \mathbf{R}^n \to \mathbf{R}$ is a convex function if $\mathrm{dom}\, f$ is a convex set and for all $x, y \in \mathrm{dom}\, f$ and $0 \le \theta \le 1$, we have

$$f(\theta x + (1-\theta)y) \le \theta f(x) + (1-\theta) f(y)$$

- f is concave if $-f$ is convex
- f is strictly convex if strict inequality above for all $x \neq y$ and $0 < \theta < 1$
- Affine functions are convex and concave

Conditions of convex functions

1. First-order condition
2. Second-order conditions
3. Restriction of a convex function to a line

First-order condition

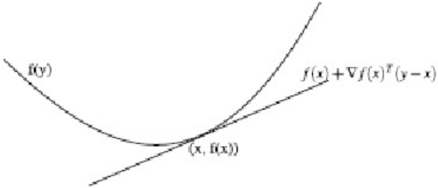
f is differentiable if $\mathrm{dom}\, f$ is open and the gradient

$$\nabla f(x) = \left(\frac{\partial f(x)}{\partial x_1}, \frac{\partial f(x)}{\partial x_2}, \ldots, \frac{\partial f(x)}{\partial x_n} \right)$$

exists at each $x \in \mathrm{dom}\, f$

1st-order condition: differentiable f with convex domain is convex iff

$$f(y) \geq f(x) + \nabla f(x)^T (y - x) \text{ for all } x, y \in \mathrm{dom}\, f$$



First-order condition

- $f(y) \geq \tilde{f}_x(y)$ where $\tilde{f}_x(y)$ is the first-order Taylor expansion of $f(y)$ at x ($\tilde{f}_x(y) = f(x) + \bigtriangledown f(x)^T(y - x)$)
- Local information (first-order Taylor approximation) about a convex function provides global information (global underestimator)
- if $\bigtriangledown f(x) = 0$, then $f(y) \geq f(x), \forall y$, thus x is a global minimizer of f

Second-order conditions

f is twice differentiable if $\mathrm{dom}\, f$ is open and the Hessian $\nabla^2 f(x) \in \mathbf{S}^n$,

$$\nabla^2 f(x)_{ij} = \frac{\partial^2 f(x)}{\partial x_i \partial x_j}, i, j = 1, \ldots, n,$$

exists at each $x \in \mathrm{dom}\, f$
2nd-order conditions: for twice differentiable f with convex domain

- f is convex if and only if

$$\nabla^2 f(x) \succeq 0 \text{ for all } x \in \mathrm{dom}\, f$$

- if $\nabla^2 f(x) \succ 0$ for all $x \in \mathrm{dom}\, f$, then f is strictly convex

Restriction of a convex function to a line

$f : \mathbf{R}^n \to \mathbf{R}$ is a convex function if the function $g : \mathbf{R} \to \mathbf{R}$

$$g(t) = f(x + tv), \quad \mathrm{dom}\, g = \{t \mid x + tv \in \mathrm{dom}\, f\}$$

is convex (in t) for any $x \in \mathrm{dom}\, f$, $v \in \mathbf{R}^n$

- Can check convexity of f by checking convexity of functions of one variable
- Example: $f : \mathbf{S}^n \to \mathbf{R}$ with $f(X) = \log \det X, \, , \mathrm{dom}\, f = \mathbf{S}^n_{++}$

$$\begin{aligned} g(t) &= \log\det(X + tV) = \log\det X + \log\det(I + tX^{-1/2}VX^{-1/2}) \\ &= \log\det X + \sum_{i=1}^{n} \log(1 + t\lambda_i) \end{aligned}$$

where λ_i are the eigenvalues of $X^{-1/2}VX^{-1/2}$
g is concave in t (for any choice of $X \succ 0$, V); hence f is concave

Examples of convex or concave functions

Convex:

- e^{ax} on $\mathbf{R}$, for any $a \in \mathbf{R}$
- x^α on $\mathbf{R}_{++}$, for $\alpha \geq 1$ or $\alpha \leq 0$
- $|x|^p$ on $\mathbf{R}$ for $p \geq 1$
- $x \log x$ on $\mathbf{R}_{++}$
- $f(x) = \max\{x_1, \ldots, x_n\}$ on $\mathbf{R}^n$
- Every norm on $\mathbf{R}^n$ is convex

Concave:

- x^α on $\mathbf{R}_{++}$, for $0 \leq \alpha \leq 1$
- $\log x$ on $\mathbf{R}_{++}$

Examples of convex or concave functions

- Quadratic function: $f(x) = (1/2)x^T Px + q^T x + r$ (with $P \in \mathbf{S}^n$)
- Least-squares objective: $f(x) = \|Ax - b\|_2^2$
- Quadratic-over-linear: $f(x, y) = x^2/y$

Operations that preserve convexity

Practical methods for establishing convexity of a function

1. Verify definition (often simplified by restricting to a line)
2. For twice differentiable functions, show $\nabla^2 f(x) \succeq 0$
3. Show that f is obtained from simple convex functions by operations that preserve convexity
   - Nonnegative weighted sum
   - Composition with affine function
   - Pointwise maximum and supremum
   - Composition
   - Minimization
   - Perspective

Operations that preserve convexity

- Nonnegative weighted sum: $f = \sum_{i=1}^{n} w_i f_i$ convex if f_i are all convex and $w_i \geq 0$
- Composition with affine function: $f(Ax + b)$ is convex if f is convex
- Pointwise maximum: if $f_1, \ldots, f_m$ are convex, then $f(x) = \max\{f_1(x), \ldots, f_m(x)\}$ is convex
- Pointwise supremum: if $f(x, y)$ is convex in x for each $y \in \mathcal{A}$, then

  $$g(x) = \sup_{y \in \mathcal{A}} f(x, y)$$

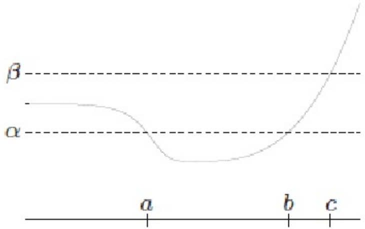
  is convex

Quasiconvex function

$f : \mathbf{R}^n \to \mathbf{R}$ is quasiconvex if $\mathrm{dom}\, f$ is convex and the sublevel sets

$$S_a = \{x \in \mathrm{dom}\, f \mid f(x) \le \alpha\}$$

are convex for all α



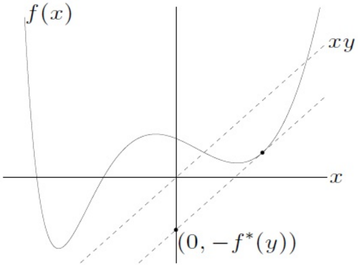
- f is quasiconcave if $-f$ is quasiconvex
- f is quasilinear if it is quasiconvex and quasiconcave

Conjugate function

Given $f : \mathbf{R}^n \to \mathbf{R}$, conjugate function $f^* : \mathbf{R}^n \to \mathbf{R}$ defined as

$$f^*(y) = \sup_{x \in \mathrm{dom}\, f}(y^T x - f(x))$$

with domain consisting of $y \in \mathbf{R}^n$ for which the supremum if finite



$f^*(y)$ is the maximum gap between the linear function yx and $f(x)$. If f is differentiable, this occurs at a point x where $f'(x) = y$

Conjugate function

- $f^*(y)$ always convex: it is the pointwise supremum of a family of affine functions of y
- Fenchel's inequality: $f(x) + f^*(y) \geq x^T y$ for all x, y (by definition)
- $f^{**} = f$ is f is convex and closed
- Useful for Lagrange duality theory

Summary

- Definitions of convex sets and convex functions
- Some examples of convex sets and convex functions
- Convexity-preserving operations
- Convexity is the watershed between "easy" and "hard" optimization problems.

2.1.3 Convex optimization problem

- Optimization problem in standard form
- Convex optimization problems
- Linear optimization
- Quadratic optimization
- Semidefinite programming

Optimization problem in standard form

$$\begin{array}{ll} \text{minimize} & f_0(x) \\ \text{subject to} & f_i(x) \leq 0, i = 1, \ldots, m \\ & h_i(x) = 0, i = 1, \ldots, p \end{array}$$

- $x \in \mathbf{R}^n$ is the optimization variable
- $f_0 : \mathbf{R}^n \to \mathbf{R}$ is the objective or cost function
- $f_i : \mathbf{R}^n \to \mathbf{R}, i = 1, \ldots, m$, are the inequality constraint functions
- $h_i : \mathbf{R}^n \to \mathbf{R}, i = 1, \ldots, m$, are the equality constraint functions

Optimal value:

$$p^* = \inf\{f_0(x) \mid f_i(x) \leq 0, i = 1, \ldots, m, h_i(x) = 0, i = 1, \ldots, p\}$$

- $p^* = \infty$ if problem is infeasible (no x satisfies the constraints)
- $p^* = -\infty$ if problem is unbounded below

Optimal and locally optimal points

x is feasible if $x \in \mathrm{dom}\, f_0$ and it satisfies the constraints
A feasible x is optimal if $f_0(x) = p^*$; X_{opt} is the set of optimal points
x is locally optimal if there is an $R > 0$ such that x is optimal for

$$\begin{array}{ll} \text{minimize (over } z) & f_0(z) \\ \text{subject to} & f_i(z) \leq 0, i = 1, \ldots, m, h_i(z) = 0, i = 1, \ldots, p \\ & \|z - x\|_2 \leq R \end{array}$$

Examples (with $n = 1$, $m = p = 0$)

- $f_0(x) = 1/x$, $\mathrm{dom}\, f_0 = \mathbf{R}_{++}$: $p^* = 0$, no optimal point
- $f_0(x) = -\log x$, $\mathrm{dom}\, f_0 = \mathbf{R}_{++}$: $p^* = -\infty$
- $f_0(x) = x \log x$, $\mathrm{dom}\, f_0 = \mathbf{R}_{++}$: $p^* = -1/e, x = 1/e$ is optimal
- $f_0(x) = x^3 - 3x$, $p^* = -\infty$, local optimum at $x = 1$

Implicit constraints

The standard form optimization problem has an implicity constraint

$$x \in D = \bigcap_{i=0}^{m} \operatorname{dom} f_i \cap \bigcap_{i=1}^{p} \operatorname{dom} h_i$$

- We call D the domain of the problem
- The constraints $f_i(x) \leq 0, h_i(x) = 0$ are the explicit constraints
- A problem is unconstrained if it has no explicit constraints $(m = p = 0)$

Example:

$$\text{minimize} \quad f_0(x) = -\sum_{i=1}^{k} \log(b_i - a_i^T x)$$

is an unconstrained problem with implicit constraints $a_i^T x < b_i$

Feasibility problem

$$\begin{array}{rl} \text{find} & x \\ \text{subject to} & f_i(x) \leq 0, i = 1, \dots, m \\ & h_i(x) = 0, i = 1, \dots, p \end{array}$$

can be considered a special case of the general problem with $f_0(x) = 0$:

$$\begin{array}{rl} \text{minimize} & 0 \\ \text{subject to} & f_i(x) \leq 0, i = 1, \dots, m \\ & h_i(x) = 0, i = 1, \dots, p \end{array}$$

- $p^* = 0$ if constraints are feasible, any feasible x is optimal
- $p^* = \infty$ if constraints are infeasible

Convex optimization problem

Standard form convex optimization problem

$$\begin{array}{rl} \text{minimize} & f_0(x) \\ \text{subject to} & f_i(x) \leq 0, i = 1, \ldots, m \\ & a_i^T x = b_i, i = 1, \ldots, p \end{array}$$

- $f_0, f_1, \ldots, f_m$ are convex; equality constraints are affine
- Problem is quasiconvex if f_0 is quasiconvex (and $f_1, \ldots, f_m$ convex)

Often written as

$$\begin{array}{rl} \text{minimize} & f_0(x) \\ \text{subject to} & f_i(x) \leq 0, i = 1, \ldots, m \\ & Ax = b \end{array}$$

Important property: feasible set of a convex optimization problem is convex

Example

$$\begin{array}{ll} \text{minimize} & f_0(x) = x_1^2 + x_2^2 \\ \text{subject to} & f_1(x) = x_1/(1 + x_2^2) \leq 0 \\ & h_1(x) = (x_1 + x_2)^2 = 0 \end{array}$$

- f_0 is convex; feasible set $\{(x_1, x_2) \mid x_1 = -x_2 \leq 0\}$ is convex
- Not a convex problem (according to our definition): f_1 is not convex, h_1 is not affine
- Equivalent (but not identical) to the convex problem

$$\begin{array}{ll} \text{minimize} & x_1^2 + x_2^2 \\ \text{subject to} & x_1 \leq 0 \\ & x_1 + x_2 = 0 \end{array}$$

Local and global optima

Any locally optimal point of a convex problem is (globally) optimal
Proof:
suppose x is locally optimal and y is optimal with $f_0(y) < f_0(x)$
x is locally optimal means there is an $R > 0$ such that

$$z \text{ feasible}, \|z - x\|_2 \leq R \to f_0(z) \geq f_0(x)$$

consider $z = \theta y + (1 - \theta)x$ with $\theta = R/(2\|y - x\|_2)$

- $\|y - x\|_2 > R$, so $0 < \theta < 1/2$
- z is a convex combination of two feasible points, hence also feasible
- $\|z - x\|_2 = R/2$ and

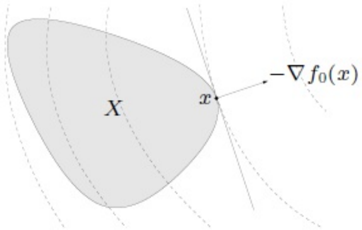
$$f_0(z) \leq \theta f_0(x) + (1 - \theta) f_0(y) < f_0(x)$$

which contradicts our assumption that x is locally optimal

Optimality criterion for differentiable f_0

x is optimal if and only if it is feasible and

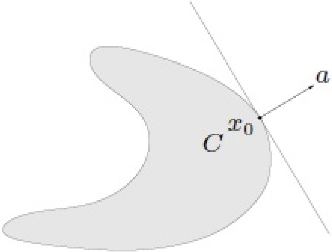
$$\nabla f_0(x)^T(y-x) \geq 0 \text{ for all feasible } y$$



If nonzero, $-\nabla f_0(x)$ defines a supporting hyperplane to feasible set X at x

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 $a^T y \leq a^T x_0$ for all $y \in C$



- Supporting hyperplane theorem: if C is convex, then there exists a supporting hyperplane at every boundary point of C

Optimality criterion for differentiable f_0

- Unconstrained problem: x is optimal if and only if

$$x \in \operatorname{dom} f_0, \quad \nabla f_0(x) = 0$$

- Equality constrained problem

$$\text{minimize } f_0(x) \quad \text{subject to } Ax = b$$

x is optimal if and only if there exists a ν such that

$$x \in \operatorname{dom} f_0, \quad Ax = b, \quad \nabla f_0(x) + A^T \nu = 0$$

- Minimization over nonnegative orthant

$$\text{minimize } f_0(x) \quad \text{subject to } x \succeq 0$$

x is optimal if and only if

$$x \in \operatorname{dom} f_0, \quad x \succeq 0, \quad \begin{cases} \nabla f_0(x)_i \geq 0 & x_i = 0 \\ \nabla f_0(x)_i = 0 & x_i > 0 \end{cases}$$

Equivalent convex problem

Two problems are (informally) equivalent if the solution of one is readily obtained from the solution of the other, and vice versa

Some common transformations that preserve convexity:

- Eliminating equality constraints

$$\begin{array}{ll} \text{minimize} & f_0(x) \\ \text{subject to} & f_i(x) \leq 0, i = 1, \ldots, m \\ & Ax = b \end{array}$$

  is equivalent to

$$\begin{array}{ll} \text{minimize} & f_0(Fz + x_0) \\ \text{subject to} & f_i(Fz + x_0) \leq 0, i = 1, \ldots, m \end{array}$$

  where F and x_0 are such that

$$Ax = b \Longleftrightarrow x = Fz + x_0 \text{ for some } z$$

Equivalent convex problem

- Introducing equality constraints

$$\begin{array}{rl} \text{minimize} & f_0(A_0x + b_0) \\ \text{subject to} & f_i(A_ix + b_i) \leq 0, i = 1, \dots, m \end{array}$$

  is equivalent to

$$\begin{array}{rl} \text{minimize (over } x, y_i) & f_0(y_0) \\ \text{subject to} & f_i(y_i) \leq 0, i = 1, \dots, m \\ & y_i = A_ix + b_i, i = 0, 1, \dots, m \ (1) \end{array}$$

- Introducing slack variables for linear inequalities

$$\begin{array}{rl} \text{minimize} & f_0(x) \\ \text{subject to} & a_i^T x \leq b_i, i = 1, \dots, m \end{array}$$

  is equivalent to

$$\begin{array}{rl} \text{minimize (over } x, s) & f_0(s) \\ \text{subject to} & a_i^T x + s_i = b_i, i = 1, \dots, m \\ & s_i \geq 0, i = 0, 1, \dots, m \qquad (2) \end{array}$$

Equivalent convex problem

- Epigraph form: standard form convex problem is equivalent to

$$
\begin{array}{rl}
\text{minimize (over } x, t) & t \\
\text{subject to} & f_0(x) - t \leq 0 \\
& f_i(x) \leq 0, i = 1, \ldots, m \\
& Ax = b
\end{array}
\tag{3}
$$

- Minimizing over some variables

$$
\begin{array}{rl}
\text{minimize} & f_0(x_1, x_2) \\
\text{subject to} & f_i(x_1) \leq 0, i = 1, \ldots, m
\end{array}
$$

is equivalent to

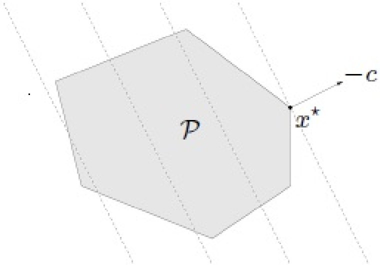
$$
\begin{array}{rl}
\text{minimize} & \tilde{f}_0(x_1) \\
\text{subject to} & f_i(x_1) \leq 0, i = 1, \ldots, m
\end{array}
$$

where $\tilde{f}_0(x_1) = \inf_{x_2} f_0(x_1, x_2)$

Linear program (LP)

$$
\begin{array}{ll}
\text{minimize} & c^T x + d \\
\text{subject to} & Gx \preceq h \\
& Ax = b
\end{array}
$$

- Convex problem with affine objective and constraint functions
- Feasible set is a polyhedron



Examples

- Diet problem: choose quantities $x_1, \ldots, x_n$ of n foods
  - One unit of food j costs c_j, contains amount a_{ij} of nutrient i
  - Healthy diet requires nutrient i in quantity at least b_i

  To find the cheapest healthy diet

$$\begin{array}{ll} \text{minimize} & \sum_{j=1}^{n} c_j x_j \\ \text{subject to} & \sum_{j=1}^{n} a_{ij} x_j \geq b_i, x_j \geq 0, j = 1, \ldots, n \end{array}$$

- Piecewise-linear minimization

$$\text{minimize} \quad \max_{i=1,\ldots,m} (a_i^T x + b_i)$$

  Equivalent to an LP

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

Semidefinite program (SDP)

$$\begin{array}{ll} \text{minimize} & c^T x \\ \text{subject to} & x_1 F_1 + x_2 F_2 + \cdots + x_n F_n + G \preceq 0 \\ & Ax = b \end{array}$$

with $F_i, G \in \mathbf{S}^k$

- Inequality constraint is called linear matrix inequality (LMI)
- Includes problems with multiple LMI constraints: for example,

$$x_1 \hat{F}_1 + \cdots + x_n \hat{F}_n + \hat{G} \preceq 0, \quad x_1 \tilde{F}_1 + \cdots + x_n \tilde{F}_n + \tilde{G} \preceq 0$$

is equivalent to single LMI

$$x_1 \begin{bmatrix} \hat{F}_1 & 0 \\ 0 & \tilde{F}_1 \end{bmatrix} + x_2 \begin{bmatrix} \hat{F}_2 & 0 \\ 0 & \tilde{F}_2 \end{bmatrix} + \cdots + x_n \begin{bmatrix} \hat{F}_n & 0 \\ 0 & \tilde{F}_n \end{bmatrix} + \begin{bmatrix} \hat{G} & 0 \\ 0 & \tilde{G} \end{bmatrix} \preceq 0$$

Examples

The second order cone programming

$$\begin{array}{ll} \text{minimize} & f^T x \\ \text{subject to} & \|A_i x + b_i\|_2 \leq c_i^T x + d_i, i = 1, \dots, m \\ & Ax = b \end{array}$$

can be expressed by SDP

$$\begin{array}{ll} \text{minimize} & f^T x \\ \text{subject to} & \begin{bmatrix} (c_i^T x + d_i) I & A_i x + b_i \\ (A_i x + b_i)^T & (c_i^T x + d_i) \end{bmatrix} \succeq 0, i = 1, \dots, m \end{array}$$