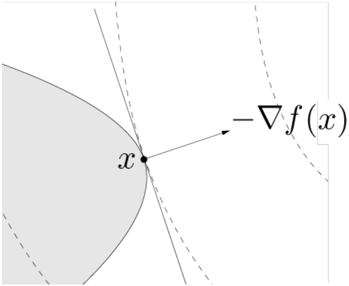
Canonical problem forms

Javier Peña (guest lecturer)
Convex Optimization 10-725/36-725

Last time: optimization basics

- Optimization terminology (e.g., criterion, constraints, feasible points, solutions)
- Properties and first-order optimality



- Equivalent transformations (e.g., partial optimization, change of variables, eliminating equality constraints)

Outline

Today:

- Linear programming
- Quadratic programming
- Semidefinite programming
- Second-order cone programming

Linear program (LP)

Optimization problem of the form

$$\begin{aligned} \min_x \quad & c^\mathsf{T} x \\ \text{subject to} \quad & Dx \le d \\ & Ax = b. \end{aligned}$$

Observe that this is a convex optimization problem.

A bit of history

- Linear programming was introduced by Dantzig in 1940s.
- Vast range of applications.
- Closely related to game theory (two-person, zero-sum games).
- Simplex method (1940s): One of the first (and still widely used) algorithms for solving linear programs.
- Interior-point methods (1980s): Theoretically fastest algorithms for solving linear programs.
- State-of-the-art solvers can easily solve problems with millions of variables and constraints.
- Polyhedra (feasible set of a linear program) have a lot of neat math properties.

Example: Diet Problem

Find cheapest combination of foods that satisfies some nutritional requirements.

$$\begin{aligned} \min_{x} \quad & c^{\mathsf{T}} x \\ \text{subject to} \quad & Dx \geq d \\ & x \geq 0. \end{aligned}$$

Here

- c_j: per-unit cost of food j
- d_i: minimum required intake of nutrient i
- D_{ij} : content of nutrient i per unit of food j
- x_j: units of food j in the diet

Example: Transportation Problem

Ship some commodity from a set of m sources to a set of n destinations at minimum cost.

$$\begin{aligned}
\min_{x} \quad & \sum_{i=1}^{m} \sum_{j=1}^{n} c_{ij} x_{ij} \\
\text{subject to} \quad & \sum_{j=1}^{n} x_{ij} \le s_i, \; i = 1, \ldots, m \\
& \sum_{i=1}^{m} x_{ij} \ge d_j, \; j = 1, \ldots, n \\
& x \ge 0.
\end{aligned}$$

Here

- s_i: supply at source i
- d_j: demand at destination j
- c_{ij} : per-unit shipping cost from i to j
- x_{ij}: units shipped from i to j

Example: ℓ_1-minimization

Heuristic to find a sparse solution to an under-determined system of equations

$$\begin{array}{ll} \min\limits_{x} & \|x\|_1 \\ \text{subject to} & Ax = b. \end{array}$$

Here $A \in \mathbb{R}^{m \times n},\ b \in \mathbb{R}^m$ with $m < n$.

Example: Dantzig selector

Tweak on previous model assuming noisy measurements:

$$b = Ax + \epsilon$$

where $\epsilon \sim N(0, \sigma^2 I)$.

Dantzig selector:

$$\begin{array}{ll} \min\limits_x & \|x\|_1 \\ \text{subject to} & \|A^\mathsf{T}(b - Ax)\|_\infty \le \lambda\sigma. \end{array}$$

Standard form

A linear program is in standard form if it is written as

$$\begin{aligned} \min_{x} \quad & c^{\mathsf{T}} x \\ \text{subject to} \quad & Ax = b \\ & x \geq 0. \end{aligned}$$

Any linear program can be rewritten in standard form.

Convex quadratic programming

Optimization problem of the form

$$\begin{aligned} \min_x \quad & c^\mathsf{T} x + \frac{1}{2} x^\mathsf{T} Q x \\ \text{subject to} \quad & Dx \le d \\ & Ax = b, \end{aligned}$$

where Q symmetric and positive semidefinite.

Example: portfolio optimization

Model to construct a financial portfolio with optimal performance/risk tradeoff:

$$\begin{aligned} \max_{x} \quad & \mu^{\mathsf{T}} x - \frac{\gamma}{2} x^{\mathsf{T}} Q x \\ \text{subject to} \quad & 1^{\mathsf{T}} x = 1 \\ & x \geq 0, \end{aligned}$$

Here

- μ : expected assets' returns
- Q : covariance matrix of assets' returns
- γ : risk aversion
- x : portfolio holdings (percentages)

Example: support vector machines

Let $y \in \{-1, 1\}^n$, and $X \in \mathbb{R}^{n \times p}$ with rows $x_1, \ldots x_n$ be given.

Support vector machine (SVM) problem:

$$\begin{aligned} \min_{\beta, \beta_0, \xi} \quad & \frac{1}{2}\|\beta\|_2^2 + C \sum_{i=1}^{n} \xi_i \\ \text{subject to} \quad & y_i(x_i^T \beta + \beta_0) \geq 1 - \xi_i, \; i = 1, \ldots n \\ & \xi_i \geq 0, \; i = 1, \ldots n \end{aligned}$$

Standard form

A quadratic program is in standard form if it is written as

$$\begin{aligned} \min_x \quad & c^\mathsf{T} x + \frac{1}{2} x^\mathsf{T} Q x \\ \text{subject to} \quad & Ax = b \\ & x \geq 0. \end{aligned}$$

Any quadratic program can be rewritten in standard form.

Semidefinite programming

Consider linear programming again:

$$\begin{aligned} \min_x \quad & c^\mathsf{T} x \\ \text{subject to} \quad & Dx \le d \\ & Ax = b. \end{aligned}$$

Can generalize by changing "$\le$" to a different partial order.

Semidefinite program

A bit of notation:

- $\mathbb{S}^n$ = space of symmetric $n \times n$ real matrices
- Cone of positive semidefinite matrices:

$$\mathbb{S}^n_+ := \{X \in \mathbb{S}^n : u^\mathsf{T} X u \geq 0 \text{ for all } u \in \mathbb{R}^n\}.$$

- Linear algebra facts:

$$X \in \mathbb{S}^n \Rightarrow \lambda(X) \in \mathbb{R}^n$$

$$X \in \mathbb{S}^n_+ \Leftrightarrow \lambda(X) \in \mathbb{R}^n_+$$

Facts about $\mathbb{S}^n$ and $\mathbb{S}^n_+$

- Canonical inner product in $\mathbb{S}^n$: Given $X, Y \in \mathbb{S}^n$

$$\langle X, Y \rangle = X \bullet Y := \text{trace}(XY)$$

- $\mathbb{S}^n_+$ is a closed convex cone
- The interior of $\mathbb{S}^n_+$ is

$$\mathbb{S}^n_{++} := \{X \in \mathbb{S}^n : u^{\mathsf{T}} X u > 0 \text{ for all } u \in \mathbb{R}^n \setminus \{0\}\}.$$

- $X \in \mathbb{S}^n_{++} \Leftrightarrow \lambda(X) \in \mathbb{R}^n_{++}$.

Loewner ordering: Given $X, Y \in \mathbb{S}^n$

$$X \succeq Y \Leftrightarrow X - Y \in \mathbb{S}^n_+.$$

Semidefinite program (SDP)

Optimization problem of the form

$$\begin{aligned} \min_x \quad & c^\mathsf{T} x \\ \text{subject to} \quad & \sum_{j=1}^n F_j x_j \preceq F_0 \\ & Ax = b. \end{aligned}$$

Here $F_j \in \mathbb{S}^d,\ j = 0, 1, \ldots, n$ and $A \in \mathbb{R}^{m \times n}, c \in \mathbb{R}^n, b \in \mathbb{R}^m$.

A semidefinite program is a convex optimization problem.

Standard form

A semidefinite program is in standard form if it is written as

$$\begin{aligned} \min_{X} \quad & C \bullet X \\ \text{subject to} \quad & A_i \bullet X = b_i, \; i = 1, \ldots, m \\ & X \succeq 0. \end{aligned}$$

Observations:

- Any linear program can be cast as a semidefinite program.
- Any semidefinite program can be written in standard form.

A bit of history

- Eigenvalue optimization, LMI problems (1960s – 1970s)
- Lovász theta function (1979) in information theory
- Interior-point algorithms for SDP (1980s, 1990s)
- Advancements in theory, algorithms, applications (1990s)
- Extensions to symmetric cones, general-purpose solvers
- New algorithms and applications in data and imaging science (2000s–)

Example: theta function

Assume $G = (N, E)$ undirected graph, $N = \{1, \ldots, n\}$.

- $\omega(G) :=$ clique number of G
- $\chi(G) :=$ chromatic number of G

Theta function:

$$\begin{array}{rl} \vartheta(G) := \max\limits_{X} & \mathbf{1}\mathbf{1}^{\mathsf{T}} \bullet X \\ \text{subject to} & I \bullet X = 1 \\ & X_{ij} = 0, (ij) \notin E \\ & X \succeq 0. \end{array}$$

Neat property (Lovász):

$$\omega(\bar{G}) \leq \vartheta(G) \leq \chi(\bar{G}).$$

Example: nuclear norm minimization

Heuristic to find a low-rank solution to an under-determined system of matrix equations.

$$\begin{array}{ll} \min\limits_{X} & \|X\|_{\mathsf{tr}} \\ \text{subject to} & \mathcal{A}(X) = b. \end{array}$$

Here $\mathcal{A} : \mathbb{R}^{m\times n} \to \mathbb{R}^p$ linear map, $b \in \mathbb{R}^p$, and $\|\cdot\|_{\mathsf{tr}}$ is the "nuclear norm":

$$\|X\|_{\mathsf{tr}} = \|\sigma(X)\|_1$$

Nuclear norm: dual of operator norm

$$\|X\|_{\mathsf{op}} = \|\sigma(X)\|_\infty = \max\{\|Xu\|_2 : \|u\|_2 \le 1\}.$$

Conic programming

LP and SDP: special cases of conic programming.

Conic program

$$\begin{aligned} \min_x \quad & c^\mathsf{T} x \\ \text{subject to} \quad & d - Dx \in K \\ & Ax = b. \end{aligned}$$

Here

- $c, x \in \mathbb{R}^n$
- $D : \mathbb{R}^n \to Y$ linear, $d \in Y$ for some Euclidean space Y
- $K \subseteq Y$ is a closed convex cone.
- write $x \preceq_K y$ for $y - x \in K$

Second-order conic programming

Second-order cone (aka Lorentz cone):

$$\mathcal{Q}_n := \left\{ x = \begin{bmatrix} x_0 \\ \bar{x} \end{bmatrix} \in \mathbb{R}^n : x_0 \geq \|\bar{x}\| \right\}.$$

Second-order cone program: Optimization problem of the form

$$\begin{array}{ll} \min\limits_{x} & c^{\mathsf{T}} x \\ \text{subject to} & d - Dx \in \mathcal{Q} \\ & Ax = b \end{array}$$

where

$$\mathcal{Q} = \mathcal{Q}_{n_1} \times \cdots \times \mathcal{Q}_{n_r}.$$

Second-order cone programming (SOCP)

A second-order program is in standard form if it is written as

$$\begin{array}{rl} \min & c^\mathsf{T} x \\ \text{subject to} & Ax = b \\ & x \succeq_{\mathcal{Q}} 0, \end{array}$$

for $\mathcal{Q} = \mathcal{Q}_{n_1} \times \cdots \times \mathcal{Q}_{n_r}$.

Observations

- Case $r = 1$ can be solved in closed-form.
- Interesting case: $r \geq 2$.
- LP $\subsetneq$ SOCP $\subsetneq$ SDP.

Example: Convex QCQP

QCQP: quadratically constrained quadratic programming.

Assume $Q = LL^\mathsf{T} \in \mathbb{S}^n,\ q \in \mathbb{R}^n,\ \ell \in \mathbb{R}$. Then

$$x^\mathsf{T} Q x + q^\mathsf{T} x + \ell \le 0$$

can be recast as

$$\left\| \begin{bmatrix} L^\mathsf{T} x \\ \dfrac{1 + q^\mathsf{T} x + \ell}{2} \end{bmatrix} \right\| \le \frac{1 - q^\mathsf{T} x - \ell}{2}.$$

Therefore a QCQP problem of the form

$$\begin{array}{rl} \min & x^\mathsf{T} Q_0 x + q_0^\mathsf{T} x \\ \text{subject to} & x^\mathsf{T} Q_i x + q_i^\mathsf{T} x + \ell_i \le 0,\ i = 1, \ldots, r \end{array}$$

can be recast as an SOCP if $Q_i \succeq 0,\ i = 0, \ldots, r$.

Other conic programs

Sometimes it is useful to combine LP/SOCP/SDP:

- Given $A \in \mathbb{S}^n$ find the nearest matrix to A in $\mathbb{S}^n_+$.
- Suppose we can only change certain entries. For example, maintain zeros in

$$A = \begin{bmatrix} 1 & 0.5 & 0 & 0 \\ 0.5 & 0.2 & 0.4 & 0 \\ 0 & 0.4 & 1 & 0.6 \\ 0 & 0 & 0.6 & 1.1 \end{bmatrix}$$

Other conic programs

In general, we can consider a conic program of the form

$$\begin{array}{rl} \min & c^{\mathsf{T}} x \\ \text{subject to} & Ax = b \\ & x \in K, \end{array}$$

where $K = K_1 \times \cdots \times K_r$ and each K_i is one of

$$\mathbb{R}^n_+,\ \mathcal{Q}_n,\ \mathbb{S}^n_+,\ \mathbb{R}^n.$$

References and further reading

- D. Bertsimas and J. Tsitsiklis (1997), "Introduction to Linear Optimization," Chapters 1, 2
- A. Nemirovski and A. Ben-Tal (2001), "Lectures on Modern Convex Optimization," Chapters 1–4
- S. Boyd and L. Vandenberghe (2004), "Convex Optimization," Chapter 4