Convex Optimization Discussion - Week 6 (Convex Optimization Problems)

Chester Holtz

Based on slides by Prof. Stephen Boyd

Overview

- Standard form
- Classification and hierarchy of convex problems
- Graph embedding (programming assignment)

Optimization problems in standard form

$$
\begin{aligned}
&\min f_0(x) \\
&f_i(x) \leq 0, \quad i = 1, \ldots, m \\
&h_j(x) = 0, \quad j = 1, \ldots, p
\end{aligned}
$$

- $x \in \mathbb{R}^n$ is the optimization variable
- $f_0 : \mathbb{R}^n \to \mathbb{R}$ is the objective
- $f_i, h_j : \mathbb{R}^n \to \mathbb{R}$ define the *explicit* inequality and equality constraints
- Unconstrained problem with *implicit* constraints: $\min f_0(x) = -\sum_i^m \log(b_i - a_i^\top x)$

Optimization problems in standard form

$$\begin{aligned} &\min f_0(x) \\ &f_i(x) \leq 0, \quad i = 1, \dots, m \\ &h_j(x) = 0, \quad j = 1, \dots, p \end{aligned}$$

Solution:

$$p^* = \inf\{f_0(x) | f_i(x) \leq 0, h_j(x)\}$$

- $p^* = \infty \implies$ *infeasible* (no solution)
- $p^* = -\infty \implies$ *unbounded*

Optimization problems in standard form

$$\begin{aligned} &\min f_0(x) \\ &f_i(x) \leq 0, \quad i = 1, \ldots, m \\ &h_j(x) = 0, \quad j = 1, \ldots, p \end{aligned}$$

- *domain* of p: $\mathcal{D} = \bigcap_{i=0}^{m} \textbf{dom } f_i \cap \bigcap_{j=0}^{p} \textbf{dom } h_j$
- x **feasible** if $x \in \textbf{dom } f_0$ and satisfies the constraints
- x **optimal** if $f_0(x) = p^*$ (note x may not be unique)
- x **locally** optimal if feasible & $f(x) \leq f(z)$,

$$||z - x|| \leq R$$

for some $R > 0$

Convex optimization problems

Any locally optimal point of a convex problem is (globally) optimal

$$\begin{aligned} &\min f_0(x) \\ &f_i(x) \leq 0, \quad i = 1, \dots, m \\ &h_j(x) = 0, \quad j = 1, \dots, p \end{aligned}$$

- f_i, $i = 0, \dots m$ are convex, h_j are affine.
- Intersection of the constraints: *feasible set* is convex

$$\begin{aligned} &\min f_0(x) \\ &f_i(x) \leq 0, \quad i = 1, \dots, m \\ &Ax = b, A \in \mathbb{R}^{p \times n} \end{aligned}$$

Optimality criteria for differentiable f_0 (from lecture)

General idea

x is feasible and the negative gradient (descent direction) of f_0 at x is more strongly correlated with x compared to any other y:

$$\nabla f_0(x)^\top (y - x) \geq 0$$

Unconstrained problem: min $f_0(x)$

$x \in \mathbf{dom}\ f_0$ and $\nabla f_0(x) = 0$

Equality constrained problem: min $f_0(x)$ s.t. $Ax = n$

$Ax = b \quad \nabla f_0(x) + A^\top \nu = 0$ for some ν

Minimization over nonnegative orthant: min $f_0(x)$ s.t. $x \geq 0$

$$x \geq 0 \text{ and } \begin{cases} \nabla f_0(x)_i \geq 0, & x_i = 0 \\ \nabla f_0(x)_i = 0, & x_i > 0 \end{cases}$$

Example 4.5: unconstrained quadratic optimization

$f_0(x) = (1/2)x^\top Px + q^\top x + r$, $P \in S^n_{++}$ (making f_0 convex)

Necessary and sufficient conditions on x^*: $\nabla f_0(x) = Px + q = 0$

- $q \neq \mathcal{R}(P)$: f_0 is unbounded below
- $P > 0$ (f_0 strictly convex): $x^* = -P^{-1}q$ (unique)
- P is singular, but $q \in \mathcal{R}(P)$: $x^* \in -P^\dagger q + \mathcal{N}(P)$

Hierarchy and classification of convex opt. problems

Different classes of convex optimization problems.

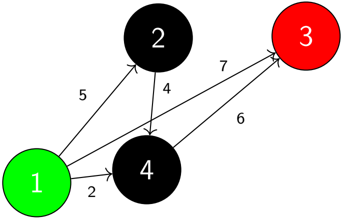
- Linear optimization
- Quadratic optimization
- Geometric programming
- Semidefinite programming

Linear program (LP)

$$\min c^\top x + d$$
$$Gx \leq h \leq 0, G \in \mathbb{R}^{m \times n}$$
$$Ax = b, A \in \mathbb{R}^{p \times n}$$

- convex problem with affine objective and constraint functions
- feasible set is a polytope

Linear program (LP) shortest path example



$$\min_{x} \sum_{i,j \in E} w_{ij} x_{ij}$$

$$\text{s.t.} \quad \sum_{j} x_{ij} - \sum_{j} x_{ji} = \begin{cases} 1, & i = s \\ -1, & i = t \\ 0, & \text{otherwise} \end{cases}$$

$$x \geq 0$$

Quadratic program (QP)

$$\begin{aligned} &\min 1/2x^\top Px + q^\top x + r \\ &Gx \leq h \leq 0, G \in \mathbb{R}^{m \times n} \\ &Ax = b, A \in \mathbb{R}^{p \times n} \end{aligned}$$

- $P \in S_+^n$, so objective is convex quadratic
- minimize a convex quadratic function over a polytope
- LP is a subset of QP

Quadratic program (QP) examples

Least-squares

$$\min ||Ax - b||_2^2$$

- analytical solution $x^* = A^\dagger b$
- can add linear constraints, e.g. $l \leq x \leq u$

Sparsemax[1]

Euclidean projection onto the Unit Simplex (map vectors to probability distributions)

$$\begin{aligned} &\min ||x - y||_2^2 \\ &\text{s.t. } 1^\top y = 1, \quad 0 \leq y \leq 1 \end{aligned}$$

[1] Martins & Astudillo, From Softmax to Sparsemax: A Sparse Model of Attention and Multi-Label Classification

Quadratically constrained quadratic program (QCQP)

$$\min 1/2x^\top P_0 x + q_0^\top x + r_0$$
$$1/2x^\top P_i x + q_i^\top x + r_i, \quad i = 1, \ldots, m$$
$$Ax = b, A \in \mathbb{R}^{p \times n}$$

- $P_i \in S_+^n$, so objective and constraints convex quadratic
- if $P_1, \ldots, P_m \in S_{++}^n$, feasible region is an intersection of m ellipsoids and an affine set
- QP is a subset of QCQP
- Example: graph embedding (homework)

Second-order cone program (SOCP)

$$\min c^\top x$$
$$||P_i x + q_i|| \leq d_i^\top x r_i, \quad i = 1, \ldots, m$$
$$Ax = b, A \in \mathbb{R}^{p \times n}$$

- Inequalities are second-order cone (SOC) constraints:

$$(P_i x + q_i, d_i^\top + r_i) \in \text{second-order cone in } \mathbb{R}^{n_i+1}$$

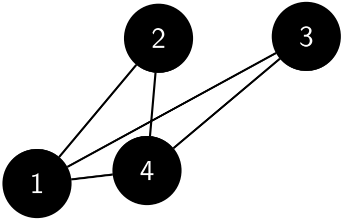
- more general than QCQP and LP (can show QCQP $\subset$ SOCP.)

Semidefinite program (SDP)

$$\min c^\top x$$
$$x_1 P_1 + x_2 P_2 + \ldots + x_n P_n + G \leq 0$$
$$Ax = b, A \in \mathbb{R}^{p \times n}$$

- Set of semidefinite matrices is a convex set (a cone)
- Linear matrix inequality (LMI) constraint
- Show LP and SOCP reduce to SDPs (via schur complement)

Programming assignment (graph embedding)



- $G = (V, E)$, $|V| = n = 4$, $|E| = 5$
- Laplacian $L = D - A$

$$A = \begin{bmatrix} 0 & 1 & 1 & 1 \\ 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & 1 \\ 1 & 1 & 1 & 0 \end{bmatrix} D = \begin{bmatrix} 3 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 \end{bmatrix} L = \begin{bmatrix} 3 & -1 & -1 & -1 \\ -1 & 2 & 0 & -1 \\ -1 & 0 & 2 & -1 \\ -1 & -1 & -1 & 3 \end{bmatrix}$$

Some properties of the Laplacian

- Symmetric: real eigenvalues, eigenspaces are mutually orthogonal
- Positive semidefinite: nonnegative eigenvalues
- Rows sum to zero: singular (at least one zero eigenvalue with unity eigenvector)

- $x^\top L x = \sum_{i,j \in E} (x_i - x_j)^2$ (show this on the hw)
- Rayleigh quotient: $\phi(x) = \frac{x^\top L x}{x^\top x}$
- Variational characterization of eigenvalues:

$$\lambda_1 = \min_x \phi(x) \quad \lambda_1 \leq \lambda_i \leq \ldots \leq \lambda_n$$

Programming assignment

Find coordinates for $v \in V$ such that:

1. Connected nodes are close together
2. Center embedding about an origin
3. We avoid trivial solutions (?)

$$\min_x x^\top L x = \min_x \sum_{i,j \in E} (x_i - x_j)^2 \qquad (1.)$$

$$1^\top x = 0, \quad x^\top x = c \qquad (2.\ \&\ 3.)$$

Is this problem convex?

Programming assignment

Two additions:

1. Convex relaxation

$$\min_x x^\top L x = \min_x \sum_{i,j \in E} (x_i - x_j)^2$$
$$x^\top x \leq c$$

2. Addition of fixed nodes
$x = [x_1 : x_2]^\top$

Programming assignment

Code walkthrough

https://colab.research.google.com/drive/1apgxNJGN1E4_W6awYbbhNxTyL0VvvMVH?usp=sharing

More examples (if time)

4.2 (logarithmic barrier), 4.3 (QP), 4.8 (LPs), 4.11 (norms), 4.12 (network flow), 4.22 (QCQP), 4.40 (SDPs)