CSE203B Convex Optimization: Chapter 4: Problem Statement

CK Cheng
Dept. of Computer Science and Engineering
University of California, San Diego

Convex Optimization Formulation

1. Introduction
   I. Eliminating equality constants
   II. Slack variables
   III. Absolute values, softmax
2. Optimality Conditions
   I. Local vs. global optimum
   II. Optimality criterion for differentiable f_0
      i. Optimization without constraints
      ii. Opt. with inequality constraints
      iii. Opt. with equality constraints
   III. Quasi-convex optimization
3. Linear Optimization
4. Quadratic Optimization
5. Geometric Programming
6. Generalized Inequality Constraints

1. Introduction

Formulation: One of the most critical processes to conduct a project.

$$\min\ f_0(x)$$
$$s.t.\ f_i(x) \le 0\ \ i = 1, \dots, m$$
$$h_i(x) = 0\ i = 1, \dots, p\ (Ax = b \text{ Affine set})$$

$$x \in R^n$$
$$D_{f_0}\ f_0: R^n \to R$$
$$D_{f_i}\ f_i: R^n \to R$$
$$D_{h_i}\ h_i: R^n \to R$$
$$f_0, f_i, \dots, f_m\ are\ convex$$

$D = \cap_{i=0,m} D_f \cap_{i=0,p} D_{h_i}$ Domain of functions, but not the feasible set.

Feasible Set: The set which satisfies the constraints (is convex for convex problems).

1.1 Introduction: Eliminating Equality Constraints

$$\min\ f_0(x)$$
$$s.t.\ f_i(x) \leq 0\ \ i = 1, \dots, m$$
$$Ax = b$$

a. Convert $\{x|Ax = b\}\ to\ \{Fz + x_0 \big| z \in R^k\}$
b. We have a equivalent problem

$$\min\ f_0(Fz + x_0)$$
$$s.t.\ f_i(Fz + x_0) \leq 0$$

Remark: Matrix F contains columns of null space basis

1.2 Introduction: Slack Variables

$$\min\ f_0(x)$$
$$s.t.\ f_i(x) \le 0, i = 1, \dots, m$$
$$Ax = b$$

Add slack variables to convert to an equivalent problem

a. Convert the objective function with variable t

$$\min\ t$$
$$s.t. f_0(x) - t \le 0$$
$$f_i(x) \le 0, i = 1, \dots, m$$
$$A^T x = b$$

b. Convert the inequality with variables s_i

$$\min\ f_0(x)$$
$$s.t. f_i(x) + s_i = 0$$
$$A^T x = b$$
$$s_i \in R_+, i = 1, \dots, m$$

1.3 Introduction: Absolute values and Softmax

a. Absolute values

$$|f_i(x)| \leq b$$
$$\Rightarrow f_i(x) \leq b \; and$$
$$-f_i(x) \leq b$$

b. Maximum values

$$\max\{f_1, f_2, \dots, f_m\}$$

Softmax: $\frac{1}{\alpha}\log\left(e^{\alpha f_1} + e^{\alpha f_1} + \cdots + e^{\alpha f_m}\right)$

$Example$: $\max\{1, 5, 10, 2, 3\} \Rightarrow$ Softmax

$$\frac{1}{\alpha}\log(e^{\alpha} + e^{5\alpha} + e^{10\alpha} + e^{2\alpha} + e^{3\alpha}) \approx 10$$

2.1 Optimality Conditions: Local vs. Global Optima

Definition: Local Optima

Given a convex optimization problem and a point $\bar{x} \in R^n$

If there exists a $r > 0$

$s.t.$ $f_0(z) \geq f_0(\bar{x})$ for all $z \in$ Feasible Set, and $\|z - \bar{x}\|_2 \leq r$

Then $\bar{x}$ is a local optimum.

2.2 Optimality Conditions

Theorem: Given a convex opt. problem

If $\bar{x}$ is a local optimum, then $\bar{x}$ is a global optimum

Proof: By contradiction

Suppose that $\exists y \in Feasible\ Set$

$s.t.\quad f_0(\bar{x}) > f_0(y)$

We have $f_0(\bar{x}) > (1-\theta)f_0(\bar{x}) + \theta f_0(\bar{y})$ $\quad$ $(by\ assumption)$

$> f_0((1-\theta)\bar{x} + \theta\bar{y})$ $(f_0\ is\ convex)$

And $(1-\theta)\bar{x} + \theta\bar{y}$ is feasible (Feasible set is convex)

The inequality contradicts to the assumption of local optima.

2.2 Optimality Criterion for Differentiable $f_0(x)$

Theorem: If $\nabla f_0(x)^T(y-x) \geq 0$, for a given x ∈Feasible Set and for all $y \in$ Feasible Set, then x is optimal.
(i. e. $K = \{y - x | y \in feasible\ set\}, \nabla f_0(x) \in K^*$)

Proof: From the first order condition of convex function, we have $f_0(y) \geq f_0(x) + \nabla f_0(x)^T(y-x)$.
Given the condition that $\nabla f_0^T(x)(y-x) \geq 0$, $\forall y$ in feasible set.
We have $f_0(y) \geq f_0(x)$, $\forall y$ in feasible set, which implies that x is optimal.

Remark: $\nabla f_0^T(x)(y-x) = 0$ is a supporting hyperplane to feasible set at x.

2.2.1 Optimality Criterion without Constraints

Theorem: For problem min $f_0(x), x \in R^n$, where f_0 is convex, the optimal condition is $\nabla f_0(x) = 0$.

Proof: ($\nabla f_0(x) = 0 \Rightarrow$ Optimality)
Since $f_0(y) \geq f_0(x) + \nabla f_0(x)^T(y - x), \forall x, y \in R^n$ (first order condition of convex function)
We have $f_0(y) \geq f_0(x)$.
Therefore, x is an optimal solution.
($\nabla f_0(x) = 0 \Leftarrow$ Optimality) By contradiction

2.2.2 Opt. with Inequality Constraints

Problem: Min $f_0(x)$

$$s.t.\ Ax \leq b, A \in R^{m \times n}$$

Suppose that $A\bar{x} = b$ (one particular case).

Let $x = \bar{x} + u.$

We can write $\begin{cases} \min f_0(\bar{x} + u) \\ \quad Au \leq 0 \end{cases}$

Opt. condition: $\nabla f_0(x)^T u \geq 0,\ \forall\{u | Au \leq 0\} \equiv K$

In other words,

$$\nabla f_0(\bar{x}) \in K^*\ of\ K = \{u | Au \leq 0\}\ and\ K^* = \{-A^T v | v \geq 0\}$$

i.e. $\nabla f_0(\bar{x}) = -A^T v, \exists v \in R_+^m$

$$\nabla f_0(\bar{x}) + A^T v = 0,\ v \geq 0.$$

2.2.3 Opt. with Equality Constraints

$$\begin{cases} \min f_0(x) \\ s.t.\ Ax = b \end{cases}$$

Let $x = \bar{x} + u$ and $A\bar{x} = b$,

we have $\begin{cases} \min f_0(\bar{x} + u) \\ Au = 0 \end{cases}$ $, K = \{u | Au = 0\}$

$$\nabla f_0(\bar{x}) \in K^*,\ \ K^* = \{A^T v | v \in R^p\}$$

$$\nabla f_0(\bar{x}) + A^T v = 0$$

Let $K_1 = \{u | Au \geq 0\}$

$K_2 = \{u | -Au \geq 0\}$

$K_1 \cap K_2 = \{u | Au \geq 0, -Au \geq 0\}$

We have

$$(K_1 \cap K_2)^* = \{A^T v_1 + (-A)^T v_2 | v_1, v_2 \geq 0\}$$
$$= \{A^T v | v \in R^p\}$$

2.2.3 Opt. with Equality Constraints: Example

$$\min_x f(x) = x_1^2 + x_2^2$$

$$s.t.\ [2\ 1]\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = 3$$

We can derive $x^* = (x_1^*, x_2^*) = (\frac{6}{5}, \frac{3}{5})$

$$\nabla f(x^*) = \begin{bmatrix} 2x_1^* \\ 2x_2^* \end{bmatrix} = \begin{bmatrix} \frac{12}{5} \\ \frac{6}{5} \end{bmatrix},\ \nabla f(x^*) + A^T v = \begin{bmatrix} \frac{12}{5} \\ \frac{6}{5} \end{bmatrix} + \begin{bmatrix} 2 \\ 1 \end{bmatrix} \times \left(-\frac{6}{5}\right) = 0$$

New Problem:

$$\begin{matrix} \nabla f(x) + A^T v = 0 \\ Ax = b \end{matrix} \Rightarrow \begin{matrix} \begin{bmatrix} 2x_1 \\ 2x_2 \end{bmatrix} + \begin{bmatrix} 2 \\ 1 \end{bmatrix} v = 0 \\ [2\ 1]\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = 3 \end{matrix}$$

2.3 Quasiconvex Functions

$f: R^n \to R$ is called quasiconvex (unimodal)
sublevel set $S_t = \{x | x \in dom\ f,\ f(x) \leq t\}$
if its domain and all sublevel sets $S_t, \forall t \in R$ are convex,
$f: R^n \to R$ is called quasiconcave if $-f$ is quasiconvex.

$f(x)$ quasiconvex and quasiconcave → quasilinear

Ex: $\log x, x \in R_{++}$

2.3 Quasiconvex Functions

Ex: Ceiling function

$Ceil(x) = \inf\{z \in Z | z > x\}$: quasilinear

Ex: $f(x_1, x_2) = x_1 x_2 = \frac{1}{2}\begin{bmatrix} x_1 & x_2 \end{bmatrix}\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$

is quasiconcave in R_+^2, $S_t = \{x \in R_+^2 | x_1 x_2 \geq t\}$

Ex: $f(x) = \frac{a^T x + b}{c^T x + d}$ for $c^T x + d > 0$

$S_t = \{x | c^T x + d > 0, a^T x + b \leq t(c^T x + d)\}$

open halfspace   closed halfspace

→ S_t is convex (t is given here)

→ $f(x)$ is quasiconvex / quasiconcave } → quasilinear

2.3 Quasiconvex Optimization

min $f_o(x)$ ($f_o(x)$ is quasiconvex, $f_i{}'s$ are convex.)

$s.t.\ \ f_i(x) \leq 0, i = 1, \dots, m$

$Ax = b$

Remark: A locally opt. solution $(x, f_0(x))$ may not be globally opt.

Algorithm: Bisection method for quasiconvex optimization.

Given $l \leq p^* \leq u, \epsilon > 0$

Find a convex function

Repeat 1. $t = (l + u)/2$

2. Find a feasible solution x:

$s.t.\, \Phi_t(x) \leq 0$ ($f_0(x) \leq t \Leftrightarrow \Phi_\mathrm{t}(x) \leq 0$)

$f_i(x) \leq 0$

$Ax = b$

3. If solution is feasible, $u = t, else\ l = t$

Until $u - l \leq \epsilon$

Ex: $f(x) = \frac{p(x)}{q(x)} \leq t \ \rightarrow \ p(x) - tq(x) \leq 0$ (p is convex & q is concave)

3. Linear Programming: Format

General Form :

$$\min c^T x$$
$$s.t.\ \ Gx \le h,\ \ G \in R^{m*n}, A \in R^{p*n}$$
$$Ax = b$$

Standard Form :

$$\min c^T x$$
$$s.t.\ \ Ax = b$$
$$x \ge 0$$

Remark: Figure out three possible situations

1. No feasible solutions
2. Unbounded solutions
3. Bounded solutions

3. Linear Programming: Cases

$\min c^T x$

$s.t.\ \ Ax = b$

(1) No feasible solutions: $b \notin R(A)$ (b is not in the range of A)

e.g. $\begin{bmatrix} 1 & 1 \\ 1 & 2 \\ 2 & 3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 2 \\ 2 \\ 3 \end{bmatrix}$

(2) Unbounded solutions: $b \in R(A)$ but $c \notin R(A^T)$

e.g. $\min \begin{bmatrix} 1 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$

$\begin{bmatrix} 1 & 2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = 2$ (The solution → $-\infty$)

(3) Bounded solutions: b$\in R(A), c \in R(A^T)$

e.g. $\min \begin{bmatrix} 1 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$

$\begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 2 \\ 2 \end{bmatrix}$

Thus $x^* = \begin{bmatrix} 2 \\ 0 \end{bmatrix}, f(x^*) = \begin{bmatrix} 1 & 1 \end{bmatrix} \begin{bmatrix} 2 \\ 0 \end{bmatrix} = 2$

3. Linear Fractional Programming

P1: $\min f_o(x) = \frac{c^T x + d}{e^T x + f}$, $dom\ f_o = \{x | e^T x + f > 0\}$

$$s.t.\quad Gx \leq h$$
$$Ax = b$$

P1⇒P2: Let $y = \frac{x}{e^T x + f}$, $\quad z = \frac{1}{e^T x + f}$

P2: $\min c^T y + dz$

$$s.t.\quad Gy - hz \leq 0$$
$$Ay - bz = 0$$
$$e^T y + fz = 1$$
$$z \geq 0$$

4. Quadratic Opt. Problems (QP)

QP : $\min \frac{1}{2} x^T P x + q^T x + r$

$$s.t. \;\; Gx \preccurlyeq h$$

$$Ax = b$$

$$P \in S_+^n, \;\; G \in R^{m\times n}, \;\; A \in R^{p\times n}$$

QCQP : (Quadratically Constrained Quadratic Program)

$$\min \frac{1}{2} x^T P_o x + q_o^T x + r_o$$

$$s.t. \;\; \frac{1}{2} x^T P_i x + q_i^T x + r_i \leq 0, i = 1, \ldots, m$$

$$Ax = b$$

$$P_i \in S_+^n, i = 0,1, \ldots, m$$

4. Quadratic Opt. Problems (SOCP)

SOCP : (Second-Order Cone Program)

$$\min f^T x$$

$$s.t. \quad \|A_i x + b_i\|_2 \le c_i^T x + d_i, i = 1, \ldots, m$$

$$\text{F}x = \text{g}$$

$SOCP$: $(Ax + b, c^T x + d)$ lies in the second order cone

$$\{(y, t) \mid \|y\|_2 \le t, y \in R^k\}$$

QCQP viewed as SOCP

QCQP constraint: $x^T A^T A x + b^T x + c \le 0$

can be expressed as a SOCP constraint:

$$\left\| \begin{matrix} \frac{1 + b^T x + c}{2} \\ Ax \end{matrix} \right\|_2 \le (1 - b^T x - c)/2$$

4. Quadratic Opt. Problems (SOCP)

SOCP : (Second-Order Cone Program)

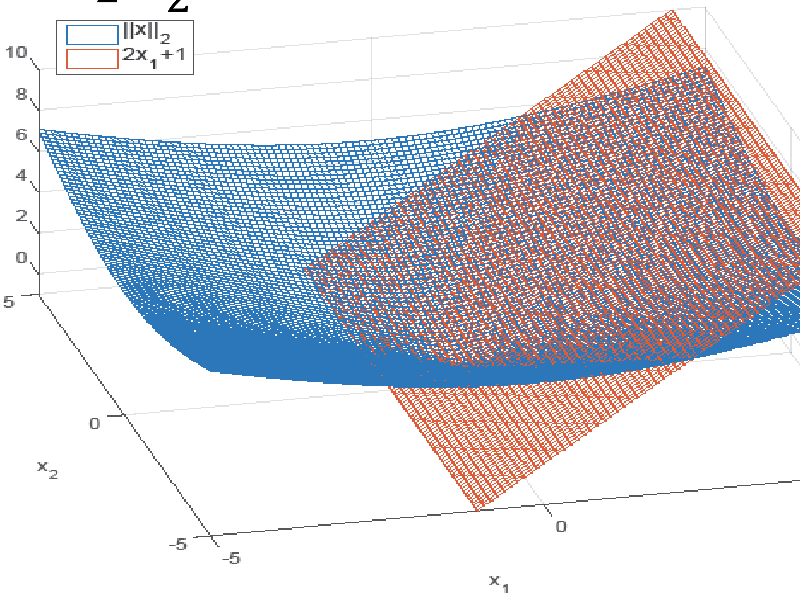
$$\min f^T x$$

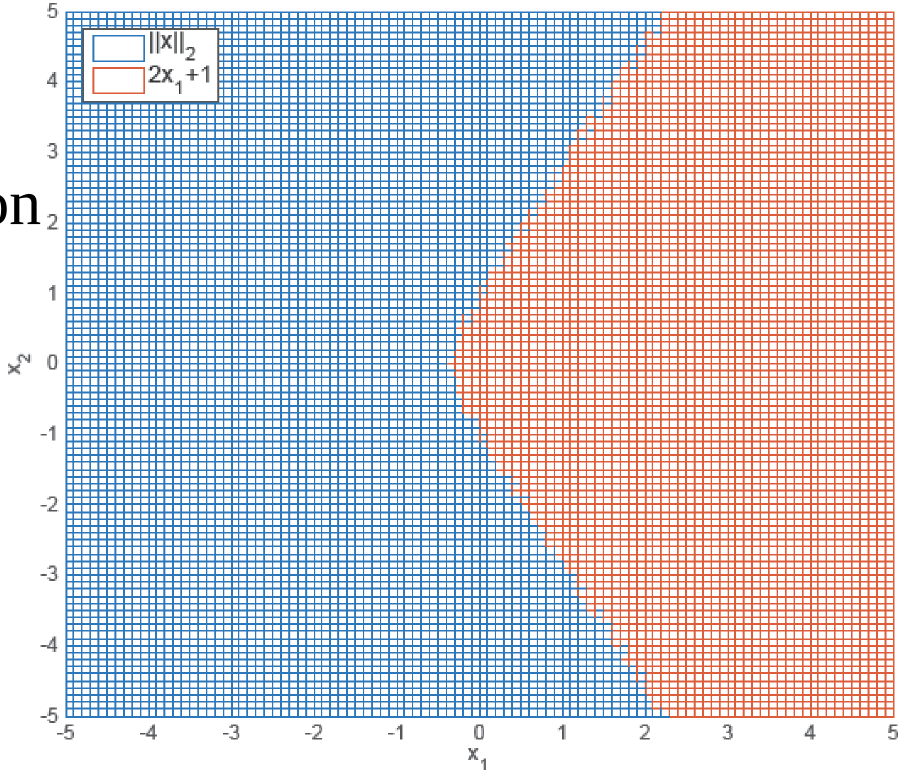
$$s.t.\quad \|A_i x + b_i\|_2 \le c_i^T x + d_i, i = 1, \ldots, m$$

$$\mathrm{F}x = \mathrm{g}$$

Example: SOCP constraint:

$\left\| \begin{matrix} x_1 \\ x_2 \end{matrix} \right\|_2 \le 2x_1 + 1$, feasible region





4. Quadratic Opt. Problems (SOCP)

SOCP : (Second-Order Cone Program)

$$\min f^T x$$
$$s.t. \quad ||A_i x + b_i||_2 \le c_i^T x + d_i, i = 1, \dots, m$$
$$Fx = g$$

$SOCP$: $(Ax + b, c^T x + d)$ lies in the second order cone

$$\{(y,t) | ||y||_2 \le t, y \in R^k\}$$

SOCP viewed as a Semidefinite Program Problem

SOCP constraint: $||Ax + b||_2 \le c^T x + d$

can be expressed as a Semidefinite Program constraint:

$$\begin{bmatrix} (c^T x + d)I & Ax + b \\ (Ax + b)^T & c^T x + d \end{bmatrix} \succcurlyeq 0$$

5. Geometric Programming

$$f(x) = \sum_{k=1}^{K} c_k x_1^{a_{1k}} x_2^{a_{2k}} \dots x_n^{a_{nk}}, \qquad c_k > 0, a_{ik} \in R, x \in R_{++}^{n}$$

Each term is called monomial

$f(x)$ is called posynomial

Geometric Program:

min $f_o(x)$

s.t.

$f_i(x) \leq 1, i = 1, \dots, m$

$h_i(x) = 1, i = 1, \dots, p$

$x > 0$

f_is are posynomials

h_is are monomials

5. Geometric programing in convex form

monomial $f(x) = cx_1^{a_1} \dots x_n^{a_n}, \quad x \in R_{++}^n$

$$\log f(e^{y_1}, \dots, e^{y_n}) = a^T y + b, \quad b = \log c$$

polynomial $f(x) = \sum_{k=1}^{K} c_k x_1^{a_{1k}} \dots x_n^{a_{nk}}$

$$\log f(e^{y_1} \dots e^{y_n}) = \log \sum_{k=1}^{K} e^{a_k^T y + b_k}, \quad b_k = \log c_k$$

Geometric program transform

$$\min \log(\sum_{k=1}^{K_0} e^{a_{ok}^T y + b_{ok}})$$

$$subject\ to \quad \log \sum_{k=1}^{K_i} e^{a_{ik}^T y + b_{ik}} \leq 0, \ i = 1, \dots, m$$

$$Gy + d = 0$$

6. Generalized Inequality Constraints

$\min f_o(x)$

$s.t.\ \ f_i(x) \preccurlyeq_{K_i} 0$

$Ax = b$

$(x \preccurlyeq_K y \ \rightarrow y - x \in K)$

Semidefinite Programming (SDP)

$\min c^T x$

$s.t.\ \ x_1 F_1 + \cdots + x_n F_n + G \preccurlyeq 0$

$Ax = b$

$G, F_1, \ldots, F_n \in S^k, A \in R^{p \times n}$

Standard Form SDP

$\min tr(CX)$

$s.t.\ \ tr(A_i X) = b_i,\ \ i = 1, \ldots, p$

$X \succcurlyeq 0$

$C, A_1, \ldots, A_p \in S^n, X \in S^n$

Summary

(1). $LP \subset QP \subset QCQP \subset SOCP \subset SDP$

(2). Software Tools (Examples)

CVX: Matlab software for disciplined convex (Boyd)

CPLEX: IP, LP, QP, SOCP (IBM)

Gurobi: LP, QP, MILP, MIQP, MIQCP (Gu, Rothberg, Bixby)

(3). Check if the problem is convex

Summary

(1). Format of the formulation

a. Follow the format of the solver (software package)

b. Find equivalent formulation for simpler approaches (coding, complexity, accuracy)

(2). Feasibility of the solution

Check if the feasible set is not empty.

(3). Boundness of the solution

Check if the solution is bounded

(reasonable, not $-\infty$)

(4). Optimality of the solution

Check the supporting hyperplane of object function