Lecture: Convex Optimization Problems

http://bicmr.pku.edu.cn/~wenzw/opt-2019-fall.html

Acknowledgement: this slides is based on Prof. Lieven Vandenberghe's lecture notes

Introduction

- optimization problem in standard form
- convex optimization problems
- quasiconvex optimization
- linear optimization
- quadratic optimization
- generalized inequality constraints
- semidefinite programming
- composite program

Optimization problem in standard form

$$\begin{array}{ll} \min & f_0(x) \\ \text{s.t.} & f_i(x) \le 0, \quad i = 1, ..., m \\ & h_i(x) = 0, \quad i = 1, ..., p \end{array}$$

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

optimal value:

$$p^* = \inf\{f_0(x) | f_i(x) \le 0, i = 1, ..., m, h_i(x) = 0, i = 1, ..., p\}$$

- $p^* = \infty$ 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

$$\begin{array}{ll} \mathrm{min(over}z) & f_0(z) \\ \text{s.t.} & f_i(z) \leq 0, \quad i = 1, ..., m \\ & h_i(z) = 0, \quad i = 1, ..., p \\ & \|z - x\|_2 \leq R \end{array}$$

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

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

Implicit constraints

the standard form optimization problem has an **implicit constraint**

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

- we call $\mathcal{D}$ 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:

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

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

Feasibility problem

$$\begin{array}{ll} \text{find} & x \\ \text{s.t.} & f_i(x) \leq 0, \quad i = 1, ..., m \\ & h_i(x) = 0, \quad i = 1, ..., p \end{array}$$

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

$$\begin{array}{ll} \min & 0 \\ \text{s.t.} & f_i(x) \leq 0, \quad i = 1, ..., m \\ & h_i(x) = 0, \quad i = 1, ..., p \end{array}$$

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

Convex optimization problem

standard form convex optimization problem

$$\begin{array}{ll} \min & f_0(x) \\ \text{s.t.} & f_i(x) \leq 0, \quad i = 1, ..., m \\ & a_i^T x = b_i, \quad i = 1, ..., p \end{array}$$

- $f_0, f_1, ..., f_m$ are convex; equality constraints are affine
- problem is *quasiconvex* if f_0 is quasiconvex (and $f_1, ..., f_m$ convex)

often written as

$$\begin{array}{ll} \min & f_0(x) \\ \text{s.t.} & f_i(x) \leq 0, \quad i = 1, ..., m \\ & Ax = b \end{array}$$

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

example

$$\begin{array}{ll} \min & f_0(x) = x_1^2 + x_2^2 \\ \text{s.t.} & f_1(x) = x_1/(1+x_2^2) \leq 0 \\ & h_1(x) = (x_1+x_2)^2 = 0 \end{array}$$

- f_0 is convex; feasible set $\{(x_1, x_2) | x_1 = -x_2 \leq 0\}$ 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

$$\begin{array}{ll} \min & x_1^2 + x_2^2 \\ \text{s.t.} & x_1 \leq 0 \\ & x_1 + x_2 = 0 \end{array}$$

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 locally optimal means there is an $R > 0$ such that

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

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

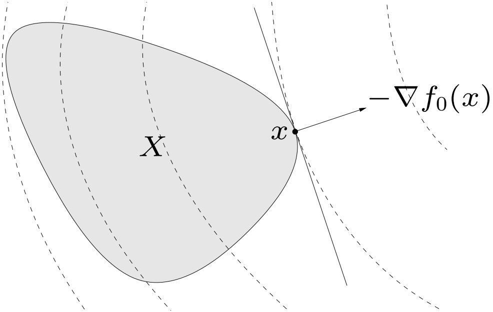
$$f_0(z) \leq \theta f_0(x) + (1 - \theta) f_0(y) < f_0(x)$$

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

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



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

- **unconstrained problem**: x is optimal if and only if

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

- **equality constrained problem**

$$\min \quad f_0(x) \qquad \text{s.t.} \quad Ax = b$$

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

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

- **minimization over nonnegative orthant**

$$\min \quad f_0(x) \qquad \text{s.t.} \quad x \succeq 0$$

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 problems

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**

$$\begin{array}{ll} \min & f_0(x) \\ \text{s.t.} & f_i(x) \leq 0, \quad i = 1, \ldots, m \\ & Ax = b \end{array}$$

  is equivalent to

$$\begin{array}{ll} \min (\text{over} z) & f_0(Fz + x_0) \\ \text{s.t.} & f_i(Fz + x_0) \leq 0, \quad i = 1, \ldots, m \end{array}$$

  where F and x_0 are such that

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

Equivalent convex problems

- **introducing equality constraints**

$$\begin{array}{ll} \min & f_0(A_0x + b_0) \\ \text{s.t.} & f_i(A_ix + b_i) \le 0, \quad i = 1, \dots, m \end{array}$$

  is equivalent to

$$\begin{array}{ll} \min(\text{over } x, y_i) & f_0(y_0) \\ \text{s.t.} & f_i(y_i) \le 0, \quad i = 1, \dots, m \\ & y_i = A_ix + b_i, \quad i = 0, 1, \dots, m \end{array}$$

- **introducing slack variables for linear inequalities**

$$\begin{array}{ll} \min & f_0(x) \\ \text{s.t.} & a^T x \le b_i, \quad i = 1, \dots, m \end{array}$$

  is equivalent to

$$\begin{array}{ll} \min(\text{over } x, s) & f_0(x) \\ \text{s.t.} & a^T x + s_i = b_i, \quad i = 1, \dots, m \\ & s_i \ge 0, \quad i = 1, \dots m \end{array}$$

Equivalent convex problems

- **epigraph form**: standard form convex problem is equivalent to

$$\begin{array}{ll} \min(\text{over } x, t) & t \\ \text{s.t.} & f_0(x) - t \leq 0 \\ & f_i(x) \leq 0, \quad i = 1, \ldots, m \\ & Ax = b \end{array}$$

- **minimizing over some variables**

$$\begin{array}{ll} \min & f_0(x_1, x_2) \\ \text{s.t.} & f_i(x_1) \leq 0, \quad i = 1, \ldots, m \end{array}$$

is equivalent to

$$\begin{array}{ll} \min & \tilde{f}_0(x_1) \\ \text{s.t.} & f_i(x_1) \leq 0, \quad i = 1, \ldots, m \end{array}$$

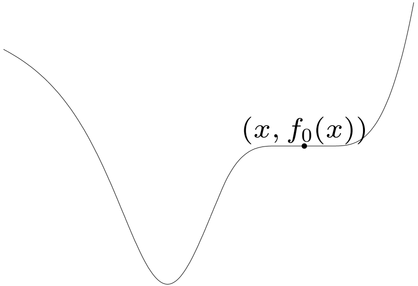
where $\tilde{f}_0(x_1) = \inf_{x_2} f_0(x_1, x_2)$

Quasiconvex optimization

$$
\begin{aligned}
\min \quad & f_0(x) \\
\text{s.t.} \quad & f_i(x) \leq 0, \quad i = 1, \ldots, m \\
& Ax = b
\end{aligned}
$$

with $f_0 : \mathbb{R}^n \to \mathbb{R}$ quasiconvex, $f_1, \ldots, f_m$ convex

can have locally optimal points that are not (globally) optimal



convex representation of sublevel sets of f_0

if f_0 is quasiconvex, there exists a family of functions ϕ_t such that:

- $\phi_t(x)$ is convex in x for fixed t
- t-sublevel set of f_0 is 0-sublevel set of ϕ_t ,$i.e.$,

$$f_0(x) \leq t \quad \Longleftrightarrow \quad \phi_t(x) \leq 0$$

example

$$f_0(x) = \frac{p(x)}{q(x)}$$

with p convex, q concave, and $p(x) \geq 0$, $q(x) > 0$ on $\mathrm{dom}\, f_0$
can take $\phi_t(x) = p(x) - tq(x)$:

- for $t \geq 0$, ϕ_t convex in x
- $p(x)/q(x) \leq t$ if and only if $\phi_t(x) \leq 0$

quasiconvex optimization via convex feasibility problems

$$\phi_t(x) \le 0, \qquad f_i(x) \le 0, \quad i = 1, \dots, m, \qquad Ax = b \tag{1}$$

- for fixed t, a convex feasibility problem in x
- if feasible, we can conclude that $t \ge p^*$; if infeasible, $t \le p^*$

Bisection method for quasiconvex optimization

given $l \le p^*$, $u \ge p^*$, tolerance $\epsilon > 0$.
repeat

1. $t := (l + u)/2$.
2. Solve the convex feasibility problem (1).
3. **if** (1) is feasible, $u := t$; **else** $l := t$.

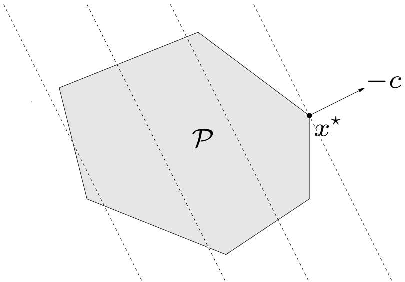
until $u - l \le \epsilon$.

requires exactly $\lceil \log_2((u - l)/\epsilon) \rceil$ iterations (where u, l are initial values)

Linear program (LP)

$$\begin{aligned} \min \quad & c^T x + d \\ \text{s.t.} \quad & Gx \le h \\ & Ax = b \end{aligned}$$

- 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 cheapest healthy diet,

$$\begin{array}{ll} \min & c^T x \\ \text{s.t.} & Ax \geq b, \quad x \geq 0 \end{array}$$

piecewise-linear minimization

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

equivalent to an LP

$$\begin{array}{ll} \min & t \\ \text{s.t.} & a_i^T x + b_i \leq t, \quad i = 1, \ldots, m \end{array}$$

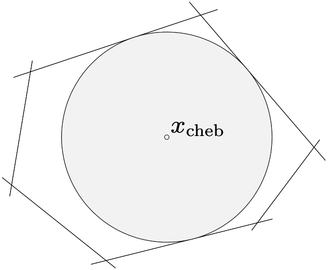
Chebyshev center of a polyhedron

Chebyshev center of

$$\mathcal{P} = \{x | a_i^T x \leq b_i, i = 1, ..., m\}$$

is center of largest inscribed ball

$$\mathcal{B} = \{x_c + u \mid \|u\|_2 \leq r\}$$



- $a_i^T x \leq b_i$ for all $x \in \mathcal{B}$ if and only if

$$\sup\{a_i^T (x_c + u) \mid \|u\|_2 \leq r\} = a_i^T x_c + r\|a_i\|_2 \leq b_i$$

- hence, x_c, r can be determined by solving the LP

$$\begin{array}{ll} \max & r \\ \text{s.t.} & a_i^T x_c + r\|a_i\|_2 \leq b_i, \quad i = 1, ..., m \end{array}$$

Linear-fractional program

$$\begin{aligned} \min \quad & f_0(x) \\ \text{s.t.} \quad & Gx \le h \\ & Ax = b \end{aligned}$$

linear-fractional program

$$f_0(x) = \frac{c^T x + d}{e^T x + f}, \qquad \operatorname{dom} f_0(x) = \{x | e^T x + f > 0\}$$

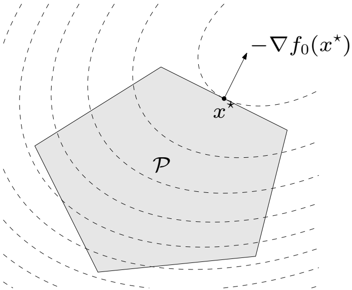
- a quasiconvex optimization problem; can be solved by bisection
- also equivalent to the LP (variables y, z)

$$\begin{aligned} \min \quad & c^T y + dz \\ \text{s.t.} \quad & Gy \le hz \\ & Ay = bz \\ & e^T y + fz = 1 \\ & z \ge 0 \end{aligned}$$

Quadratic program (QP)

$$\begin{array}{ll} \min & (1/2)x^T P x + q^T x + r \\ \text{s.t.} & Gx \leq h \\ & Ax = b \end{array}$$

- $P \in \mathbb{S}^n_+$, so objective is convex quadratic
- minimize a convex quadratic function over a polyhedron



Examples

least-squares

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

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

linear program with random cost

$$\begin{array}{ll} \min & \bar{c}^T x + \gamma x^T \Sigma x = \mathbf{E} c^T x + \gamma \mathrm{var}(c^T x) \\ \text{s.t.} & Gx \leq h, \quad Ax = b \end{array}$$

- c is random vector with mean $\bar{c}$ and covariance Σ
- hence, $c^T x$ is random variable with mean $\bar{c}^T x$ and variance $x^T \Sigma x$
- $\gamma > 0$ is risk aversion parameter; controls the trade-off between expected cost and variance (risk)

Quadratically constrained quadratic program (QCQP)

$$
\begin{array}{ll}
\min & (1/2)x^T P_0 x + q_0^T x + r_0 \\
\text{s.t.} & (1/2)x^T P_i x + q_i^T x + r_i \leq 0, \quad i = 1, \ldots, m \\
& Ax = b
\end{array}
$$

- $P_i \in \mathbb{S}_+^n$; objective and constraints are convex quadratic
- if $P_1, \ldots, P_m \in \mathbb{S}_{++}^n$, feasible region is intersection of m ellipsoids and an affine set

Generalized inequality constraints

convex problem with generalized inequality constraints

$$\begin{array}{ll} \min & f_0(x) \\ \text{s.t.} & f_i(x) \preceq_{K_i} 0, \quad i = 1, \dots, m \\ & Ax = b \end{array}$$

- $f_0 : \mathbb{R}^n \to \mathbb{R}$ convex; $f_i : \mathbb{R}^n \to \mathbb{R}^{k_i}$ K_i-convex w.r.t. proper cone K_i
- same properties as standard convex problem (convex feasible set, local optimum is global, etc.)

conic form problem: special case with affine objective and constraints

$$\begin{array}{ll} \min & c^T x \\ \text{s.t.} & Fx + g \preceq_K 0 \\ & Ax = b \end{array}$$

extends linear programming ($K = \mathbb{R}^m_+$) to nonpolyhedral cones

Second-order cone programming

$$\begin{array}{ll} \min & f^T x \\ \text{s.t.} & \|A_i x + b_i\|_2 \leq c_i^T x + d_i, \quad i = 1, \ldots, m \\ & Fx = g \end{array}$$

($A_i \in \mathbb{R}^{n_i \times n}$, $F \in \mathbb{R}^{p \times n}$)

- inequalities are called second-order cone (SOC) constraints:

$$(A_i x + b_i, c_i^T x + d_i) \in \text{ second-order cone in } \mathbb{R}^{n_i+1}$$

- for $n_i = 0$, reduces to an LP; if $c_i = 0$, reduces to a QCQP
- more general than QCQP and LP

Semidefinite program (SDP)

$$\begin{array}{ll} \min & b^T y \\ \text{s.t.} & y_1 A_1 + y_2 A_2 + \cdots + y_m A_m \preceq C \\ & By = d \end{array}$$

with $A_i, C \in \mathbb{S}^k$

- 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

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

LP and SOCP as SDP

LP and equivalent SDP

LP: $\min \quad c^T x$
s.t. $\quad Ax \leq b$

SDP: $\min \quad c^T x$
s.t. $\quad \mathrm{diag}(Ax - b) \preceq 0$

(note different interpretation of generalized inequality $\preceq$)

SOCP and equivalent SDP

SOCP:
$$\begin{array}{ll} \min & f^T x \\ \text{s.t.} & \|A_i x + b_i\|_2 \leq c_i^T x + d_i, \quad i = 1, \ldots, m \end{array}$$

SDP:
$$\begin{array}{ll} \min & f^T x \\ \text{s.t.} & \begin{bmatrix} (c_i^T x + d_i) I & A_i x + b_i \\ (A_i x + b_i)^T & c_i^T x + d_i \end{bmatrix} \succeq 0, \quad i = 1, \ldots, m \end{array}$$

复合优化问题

复合优化问题一般可以表示为如下形式：

$$\min_{x\in\mathbb{R}^n} \quad \psi(x) = f(x) + h(x),$$

其中$f(x)$ 是光滑函数，$h(x)$ 可能是非光滑的（比如ℓ_1范数正则项，约束集合的示性函数，或他们的线性组合）。令$h(x) = \mu\|x\|_1$:

- ℓ_1 范数正则化回归分析问题：$f(x) = \|Ax - b\|_2^2$或$\|Ax - b\|_1$.
- ℓ_1 范数正则化逻辑回归问题：$f(x) = \sum_{i=1}^m \log(1 + \exp(-b_i \cdot a_i^T x))$.
- ℓ_1 范数正则化支持向量机：$f(x) = C\sum_{i=1}^m \max\{1 - b_i a_i^T x, 0\}$.
- ℓ_1范数正则化精度矩阵估计：$f(x) = -(\log\det(X) - \mathrm{tr}XS)$.
- 矩阵分离问题：$f(X) = \|X\|_*$.

低秩矩阵恢复

令Ω 是矩阵M 中所有已知元素的下标的集合

- 低秩矩阵恢复

$$\begin{aligned} \min_{X\in\mathbb{R}^{m\times n}} \quad & \operatorname{rank}(X) \\ \text{s.t.} \quad & X_{ij} = M_{ij},\ (i,j)\in\Omega. \end{aligned}$$

- 核范数松弛问题：

$$\begin{aligned} \min_{X\in\mathbb{R}^{m\times n}} \quad & \|X\|_*, \\ \text{s.t.} \quad & X_{ij} = M_{ij},\ (i,j)\in\Omega. \end{aligned}$$

- 二次罚函数形式：

$$\min_{X\in\mathbb{R}^{m\times n}} \quad \mu\|X\|_* + \frac{1}{2}\sum_{(i,j)\in\Omega}(X_{ij}-M_{ij})^2.$$

随机优化问题

- 随机优化问题可以表示成以下形式：

$$\min_{x\in\mathcal{X}} \quad \mathbb{E}_{\xi}[f(x,\xi)] + h(x),$$

其中$\mathcal{X} \subseteq \mathbb{R}^n$ 表示决策变量x 的可行域，ξ 是一个随机变量（分布一般是未知的）。对于每个固定的ξ，$f(x,\xi)$ 表示样本ξ 上的损失或者奖励。正则项$h(x)$ 用来保证解的某种性质。由于变量ξ 分布的未知性，其期望$\mathbb{E}_{\xi}[f(x,\xi)]$一般是不可计算的。为了得到目标函数值的一个比较好的估计，实际问题中往往利用ξ 的经验分布来代替其真实分布。

- 假设有N 个样本$\xi_1,\xi_2,\ldots,\xi_N$，令$f_i(x) = f(x,\xi_i)$，我们得到下面的优化问题

$$\min_{x\in\mathcal{X}} \quad f(x) := \frac{1}{N}\sum_{i=1}^{N} f_i(x) + h(x),$$

其也被称作经验风险极小化问题或者采样平均极小化问题。