Convex Optimization Problems

Goran Banjac

Large-Scale Convex Optimization
ETH Zurich

March 17, 2020

Optimization problem in standard form

$$\begin{array}{ll} \text{minimize} & f(x) \\ \text{subject to} & g_i(x) \leq 0, \quad i = 1, \ldots, m \\ & h_j(x) = 0, \quad j = 1, \ldots, p \end{array}$$

- $x \in \mathbb{R}^n$ is the *optimization variable*
- $f : \mathbb{R}^n \mapsto \mathbb{R}$ is the *objective function*
- $g_i : \mathbb{R}^n \mapsto \mathbb{R}$ are the *inequality constraint functions*
- $h_j : \mathbb{R}^n \mapsto \mathbb{R}$ are the *equality constraint functions*

optimal value:

$$p^\star = \inf \{ f(x) \mid g_i(x) \leq 0, \ i = 1, \ldots, m, \ h_j(x) = 0, \ j = 1, \ldots, p \}$$

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

Local and global minima

- x is *feasible* if it satisfies the constraints (both explicit and implicit)
- a feasible x is *optimal* if $f(x) = p^\star$
- x is *locally optimal* if there exists an $R > 0$ such that x is optimal for

$$\begin{array}{ll} \text{minimize} & f(z) \\ \text{subject to} & g_i(z) \leq 0, \quad i = 1, \ldots, m \\ & h_j(z) = 0, \quad j = 1, \ldots, p \\ & \|z - x\| \leq R \end{array}$$

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

Feasibility problem

$$
\begin{array}{ll}
\text{find} & x \\
\text{subject to} & g_i(x) \le 0, \quad i = 1, \dots, m \\
& h_j(x) = 0, \quad j = 1, \dots, p
\end{array}
$$

- can be seen as a special case of the general problem with $f(x) = 0$:

$$
\begin{array}{ll}
\text{minimize} & 0 \\
\text{subject to} & g_i(x) \le 0, \quad i = 1, \dots, m \\
& h_j(x) = 0, \quad j = 1, \dots, p
\end{array}
$$

- $p^\star = 0$ if there exists a feasible x
- $p^\star = +\infty$ if no x satisfies the constraints

Convex optimization problem

standard form convex optimization problem

$$
\begin{array}{ll}
\text{minimize} & f(x) \\
\text{subject to} & g_i(x) \leq 0, \quad i = 1, \ldots, m \\
& a_j^T x = b_j, \quad j = 1, \ldots, p
\end{array}
$$

- f and g_i are convex, equality constraints are affine
- equality constraints are often written as $Ax = b$
- feasible set and the set of minimizers are convex
- any locally optimal point of a convex problem is globally optimal

Epigraph reformulation

- standard form convex problem is equivalent to

$$\begin{array}{ll} \text{minimize} & t \\ \text{subject to} & f(x) - t \leq 0 \\ & g_i(x) \leq 0, \qquad i = 1, \ldots, m \\ & Ax = b \end{array}$$

- **example:** piecewise-affine minimization

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

equivalent to

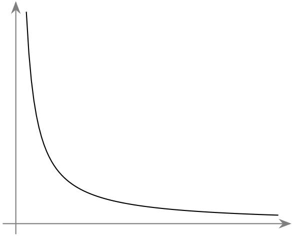
$$\begin{array}{ll} \text{minimize} & t \\ \text{subject to} & a_i^T x + b_i \leq t, \quad i = 1, \ldots, m \end{array}$$

Strict and strong convexity

- function $f : \mathbb{R}^n \mapsto \overline{\mathbb{R}}$ is *strictly convex* if for all $x, y \in \mathbb{R}^n$ and $\theta \in (0, 1)$:

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

- convexity definition with strict inequality; no flat (affine) regions
- **example:** $f(x) = 1/x$ for $x > 0$



- f is σ*-strongly convex* if $f - \frac{\sigma}{2}\|\cdot\|_2^2$ is convex, or equivalently, for all $x, y \in \mathbb{R}^n$ and $\theta \in [0, 1]$:

$$f(\theta x + (1-\theta)y) \leq \theta f(x) + (1-\theta) f(y) - \frac{\sigma}{2}\theta(1-\theta)\|x - y\|_2^2$$

Existence and uniqueness of minimizers

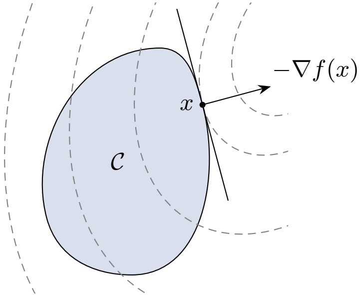
- strictly (strongly) convex functions have unique minimizers
- strictly convex functions may not have a minimizer (*e.g.*, $-\log x$)
- strongly convex functions always have a unique minimizer
- a minimizer exists if feasible set is nonempty and compact

Optimality criterion for differentiable objective

- $x \in \mathbb{R}^n$ is a minimizer if and only if it is feasible and

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

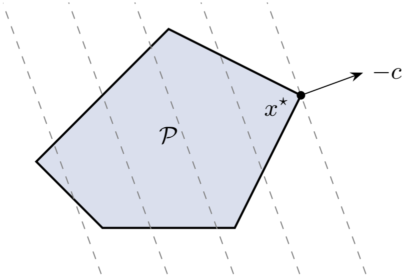
- if nonzero, $-\nabla f(x)$ defines a supporting hyperplane to feasible set $\mathcal{C}$ at x



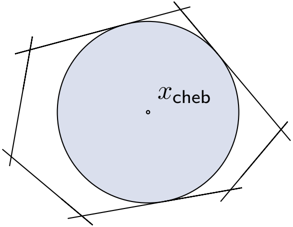
Linear program (LP)

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

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



Chebyshev center of a polyhedron



- Chebyshev center of

  $$\mathcal{P} = \{x \mid a_i^T x \le b_i, i = 1, \dots, m\}$$

  is center of largest inscribed ball

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

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

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

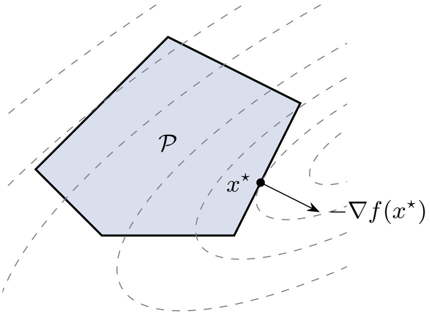
- hence, x_c and r can be determined by solving the LP

$$\begin{array}{ll} \text{maximize} & r \\ \text{subject to} & a_i^T x_c + r\|a_i\|_2 \le b_i, \quad i = 1, \dots, m \end{array}$$

Quadratic program (QP)

$$
\begin{array}{ll}
\text{minimize} & \frac{1}{2}x^T P x + q^T x \\
\text{subject to} & Gx \leq h \\
& Ax = b
\end{array}
$$

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



Examples of QPs

least-squares

$$\text{minimize} \quad \|Ax - b\|_2^2$$

- analytical solution $x^\star = (A^T A)^{-1} A^T b$
- can add linear constraints, *e.g.*, $l \leq x \leq u$

linear program with random cost

$$\begin{array}{ll} \text{minimize} & \bar{c}^T x + \gamma x^T \Sigma x = \mathbb{E}\, c^T x + \gamma \mathrm{var}(c^T x) \\ \text{subject to} & Gx \leq h \\ & Ax = b \end{array}$$

- 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)

$$
\begin{array}{ll}
\text{minimize} & \frac{1}{2}x^T P_0 x + q_0^T x \\
\text{subject to} & \frac{1}{2}x^T P_i x + q_i^T x + r_i \leq 0, \quad i = 1, \ldots, m \\
& Ax = b
\end{array}
$$

- $P_i \in \mathbb{S}^n_+$, so 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

Second-order cone program (SOCP)

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

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

- inequalities are called second-order cone constraints:

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

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

Conic programming

conic form problem

$$\begin{array}{ll} \text{minimize} & c^T x \\ \text{subject to} & Gx \preceq_{\mathcal{K}} h \\ & Ax = b \end{array}$$

- $\mathcal{K} \subseteq \mathbb{R}^m$ is a proper convex cone
- reduces to linear programs when $\mathcal{K} = \mathbb{R}^m_+$
- reduces to second-order cone programs when $\mathcal{K}$ is second-order cone

Semidefinite program (SDP)

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

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

- inequality constraint is called linear matrix inequality (LMI)
- includes problems with multiple LMI constraints, *e.g.*,

$$x_1 \bar{F}_1 + x_2 \bar{F}_2 + \ldots + x_n \bar{F}_n + \bar{G} \preceq 0, \quad x_1 \hat{F}_1 + x_2 \hat{F}_2 + \ldots + x_n \hat{F}_n + \hat{G} \preceq 0$$

  is equivalent to single LMI

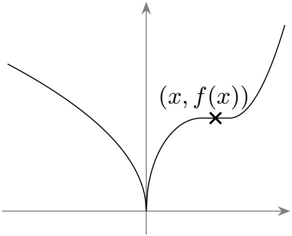
$$x_1 \begin{bmatrix} \bar{F}_1 & 0 \\ 0 & \hat{F}_1 \end{bmatrix} + x_2 \begin{bmatrix} \bar{F}_2 & 0 \\ 0 & \hat{F}_2 \end{bmatrix} + \ldots + x_n \begin{bmatrix} \bar{F}_n & 0 \\ 0 & \hat{F}_n \end{bmatrix} + \begin{bmatrix} \bar{G} & 0 \\ 0 & \hat{G} \end{bmatrix} \preceq 0$$

Quasiconvex optimization

- function $f : \mathbb{R}^n \mapsto \overline{\mathbb{R}}$ is *quasiconvex* if all its sublevel sets are convex
- quasiconvex optimization problem:

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

  where f is quasiconvex and $f_1, \ldots, f_m$ are convex
- can have local minimizers that are not global



Solving quasiconvex optimization problems

convex representation of sublevel sets of f

- if f is quasiconvex, then 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 is 0-sublevel set of ϕ_t, *i.e.*,

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

quasiconvex optimization via convex feasibility problems

$$\phi_t(x) \leq 0, \quad g_i(x) \leq 0,\ i = 1, \ldots, m, \quad Ax = b$$

- for fixed t, a convex feasibility problem in x
- if feasible, we can conclude that $t \geq p^\star$; otherwise, $t \leq p^\star$
- can use bisection and solve a sequence of problems for varying t

Composite minimization

convex composite minimization problem

$$\text{minimize} \quad f(x) + g(x)$$

(f and g are convex closed proper)

- many convex optimization problems can be represented in this form
  - constrained optimization ($g = \mathcal{I}_{\mathcal{C}}$)
  - regularized optimization ($g = \gamma \|\cdot\|$)
  - feasibility problems ($f = \mathcal{I}_{\mathcal{C}}$, $g = \mathcal{I}_{\mathcal{D}}$)
- many algorithms for large-scale optimization are designed for solving composite minimization problems

References

- these lecture notes are based to a large extent on the following material:
  - Stanford EE364a class developed by Stephen Boyd
  - Lund course on Large-Scale Convex Optimization developed by Pontus Giselsson

- the original slides can be downloaded from

  `https://web.stanford.edu/class/ee364a/lectures.html`

  `https://archive.control.lth.se/ls-convex-2015/`