Chapter 4 Convex optimization problems

Last update on 2025-02-27 14:11:19+08:00

Table of contents

Outline

Convex optimization problems

- relevant concepts (for general optimization problems & for convex problems)
- properties of convex problems (local implies global & optimality condition)
- operations preserving convexity (construct new from old)
- many examples of convex problems (LP, QP, QCQP, SOCP, etc.)
- extensions (quasiconvex optimization & geometric programming)
- combination with generalized inequalities (in constraints & in objective functions)

Optimization problems

Convex optimization

Linear optimization

Quadratic optimization

Geometric programming

Generalized inequality constraints

Vector optimization

Optimization problem in standard form

$$\begin{array}{lll} \text{minimize} & f_0(x) & \\ \text{subject to} & f_i(x) \leq 0, & i = 1, \cdots, m \\ & h_i(x) = 0, & i = 1, \cdots, p \end{array}$$

$x \in \mathbb{R}^n$	optimization variable
$f_0 : \mathbb{R}^n \to \mathbb{R}$	objective function (cost function)
$f_i : \mathbb{R}^n \to \mathbb{R}$	inequality constraint functions
$h_i : \mathbb{R}^n \to \mathbb{R}$	equality constraint functions

Constraints

- **implicit constraints**

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

- $\mathcal{D}$ is called the **domain** of the problem
- **explicit constraints**

$$f_i(x) \leq 0 \text{ for } 1 \leq i \leq m \quad \text{and} \quad h_i(x) = 0 \text{ for } 1 \leq i \leq p$$

- problem is **unconstrained** if it has no explicit constraints ($m = p = 0$)

Example

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

is an unconstrained problem with implicit constraints

$$a_i^T x < b_i$$

for each $1 \leq i \leq k$.

Feasibility

- x is **feasible** if $x \in \mathcal{D}$ and x satisfies all constraints
- the set of all feasible points is called the **feasible set** of the problem
- the problem is **infeasible** if the feasible set is empty
- the **feasibility problem** is to determine whether the feasible set is nonempty

$$\begin{array}{rll} \text{find} & x & \\ \text{subject to} & f_i(x) \leq 0, & i = 1, \cdots, m \\ & h_i(x) = 0, & i = 1, \cdots, p \end{array}$$

- it can be rephrased as an optimization problem

$$\begin{array}{rll} \text{minimize} & 0 & \\ \text{subject to} & f_i(x) \leq 0, & i = 1, \cdots, m \\ & h_i(x) = 0, & i = 1, \cdots, p \end{array}$$

Optimality

- The **optimal value** is

$$p^* = \textbf{inf} \left\{ f_0(x) \;\middle|\; \begin{array}{c} f_i(x) \leq 0 \text{ for } 1 \leq i \leq m \\ h_i(x) = 0 \text{ for } 1 \leq i \leq p \end{array} \right\} \in \mathbb{R} \cup \{\pm\infty\}$$

- Extreme situations

$$\begin{aligned} p^* &= \infty && \text{if problem is infeasible} \\ p^* &= -\infty && \text{if problem is unbounded below} \end{aligned}$$

- Optimal value may not be achieved.

- x is **optimal** if it is feasible and $f_0(x) = p^*$
- x is **locally optimal** if there exists $R > 0$ such that x is optimal for

$$
\begin{array}{lll}
\text{minimize} & f_0(z) & \\
\text{subject to} & f_i(z) \leq 0, & i = 1, \cdots, m \\
& h_i(z) = 0, & i = 1, \cdots, p \\
& \|z - x\|_2 \leq R &
\end{array}
$$

Examples (when $n = 1$, $m = p = 0$)

$f_0(x) = x \log x$	$\mathbf{dom}\, f_0 = \mathbb{R}_{++}$	$p^\star = -1/e$	$x = 1/e$ is optimal
$f_0(x) = -\log x$	$\mathbf{dom}\, f_0 = \mathbb{R}_{++}$	$p^\star = -\infty$	no optimal point
$f_0(x) = 1/x$	$\mathbf{dom}\, f_0 = \mathbb{R}_{++}$	$p^\star = 0$	no optimal point
$f_0(x) = x^3 - 3x$	$\mathbf{dom}\, f_0 = \mathbb{R}$	$p^\star = -\infty$	$x = 1$ is locally optimal

Convex optimization problem in standard form

$$\begin{array}{lll} \text{minimize} & f_0(x) & \\ \text{subject to} & f_i(x) \leq 0, & i = 1, \cdots, m \\ & a_i^T x = b_i, & i = 1, \cdots, p \end{array}$$

- $f_0, f_1, \cdots, f_m$ are convex
- equality constraints are affine, often written as $Ax = b$
- important property: feasible set of a convex problem is convex
- problem is **quasiconvex** if f_0 is quasiconvex (and $f_1, \cdots, f_m$ convex)

Example

$$
\begin{array}{ll}
\text{minimize} & f_0(x) = x_1^2 + x_2^2 \\
\text{subject to} & f_1(x) = x_1/(1 + x_2^2) \leq 0 \\
& h_1(x) = (x_1 + x_2)^2 = 0
\end{array}
$$

- f_0 is convex
- not a convex problem: f_1 is not convex, h_1 is not affine
- equivalent (but not identical) to the convex problem

$$
\begin{array}{ll}
\text{minimize} & x_1^2 + x_2^2 \\
\text{subject to} & x_1 \leq 0 \\
& x_1 + x_2 = 0
\end{array}
$$

Properties of convex optimization problems

- Local optima are global;
- First order optimality criterion.

Local and global optima

Proposition

Any locally optimal point of a convex optimization problem is globally optimal.

Proof

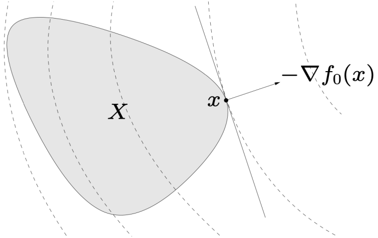
- suppose x is locally optimal, but there exists feasible y with $f_0(y) < f_0(x)$
- there exists $R > 0$ such that $f_0(z) \geq f_0(x)$ for all feasible z with $\|z - x\|_2 < R$
- consider $z = \theta y + (1 - \theta)x$ with $\theta = R/(2\|y - x\|_2)$, then $\|z - x\|_2 = R/2$
- $\|y - x\|_2 > R$ implies $0 < \theta < 1/2$, hence z is feasible by convexity of domain
- by convexity of objective $f_0(z) \leq \theta f_0(y) + (1 - \theta)f_0(x) < f_0(x)$, contradiction

Optimality criterion for differentiable objective

Optimality criterion

Suppose the problem is convex and f_0 is differentiable, then

$$x \text{ is optimal} \quad \Longleftrightarrow \quad \begin{array}{c} x \text{ is feasible and} \\ \nabla f_0(x)^T(y-x) \geq 0 \text{ for all feasible } y \end{array}$$



Geometric interpretation

Either $\nabla f_0(x) = 0$ or $\nabla f_0(x)$ defines a supporting hyperplane to the feasible set at x.

Proof

$(\Longleftarrow)$ For any feasible y, since $y \in \mathbf{dom}\, f_0$, by the convexity of f_0

$$f_0(y) \geq f_0(x) + \nabla f_0(x)^T (y - x).$$

The assumption $\nabla f_0(x)^T (y - x) \geq 0$ implies $f_0(y) \geq f_0(x)$. Hence x is optimal.

$(\Longrightarrow)$ Assume on the contrary that $\nabla f_0(x)^T (y - x) < 0$ for some feasible y, then $z(t) = ty + (1 - t)x$ is feasible for $t \in [0, 1]$ since the feasible set is convex. Then

$$\frac{\mathrm{d}}{\mathrm{d}t} f_0(z(t))\bigg|_{t=0} = \nabla f_0(x)^T (y - x) < 0,$$

hence $f_0(z(t)) < f_0(x)$ for $0 < t \ll 1$, which contradicts the optimality of x. □

unconstrained problem

$$\text{minimize} \quad f_0(x)$$

$$x \text{ is optimal} \quad \Longleftrightarrow \quad x \in \mathbf{dom}\, f_0, \quad \nabla f_0(x) = 0$$

Sample proof (for unconstrained problems)

- By optimality condition

$$x \text{ is optimal} \quad \Longleftrightarrow \quad x \in \mathbf{dom}\, f_0,\ \nabla f_0(x)^T (y - x) \geq 0 \text{ for each } y \in \mathbf{dom}\, f_0$$

- $\nabla f_0(x) = 0$ is clearly sufficient for the above statement.
- Since f_0 is differentiable, $\mathbf{dom}\, f_0$ is open, hence

$$y = x - \varepsilon \nabla f_0(x) \in \mathbf{dom}\, f_0$$

for $0 < \varepsilon \ll 1$. For such y we have

$$\nabla f_0(x)^T (y - x) = -\varepsilon \|\nabla f_0(x)\|_2^2 \leq 0.$$

- Combining above gives

$$\nabla f_0(x) = 0$$

which proves necessity.

equality constrained problem

$$\begin{array}{ll} \text{minimize} & f_0(x) \\ \text{subject to} & Ax = b \end{array}$$

$$x \text{ is optimal} \quad \Longleftrightarrow \quad \begin{array}{c} x \in \mathbf{dom}\, f_0, \quad Ax = b, \\ \nabla f_0(x) + A^T \nu = 0 \text{ for some vector } \nu \end{array}$$

minimization over nonnegative orthant

$$\begin{array}{ll} \text{minimize} & f_0(x) \\ \text{subject to} & x \succeq 0 \end{array}$$

$$x \text{ is optimal} \quad \Longleftrightarrow \quad x \in \mathbf{dom}\, f_0, \quad x \succeq 0, \quad \begin{cases} \nabla f_0(x)_i \geq 0, & \text{if } x_i = 0 \\ \nabla f_0(x)_i = 0, & \text{if } x_i > 0 \end{cases}$$

Proof for minimization over nonnegative orthant

- By optimality condition

x is optimal $\iff$ $x \in \mathbf{dom}\, f_0,\ x \succeq 0,\ \nabla f_0(x)^T(y-x) \geq 0$ for each feasible y

- We claim $\nabla f_0(x) \succeq 0$; otherwise, there exists some $v \succeq 0$ such that

$$\nabla f_0(x)^T v < 0$$

and $y = x + \varepsilon \nabla f_0(x)$ is feasible (why?), which violates the optimality condition.

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
- introducing equality constraints
- introducing slack variables for linear inequalities
- epigraph form
- minimizing over some variables

eliminating equality constraints

$$\begin{array}{lll} \text{minimize} & f_0(x) & \\ \text{subject to} & f_i(x) \leq 0, & i = 1, \cdots, m \\ & Ax = b & \end{array}$$

is equivalent to

$$\begin{array}{lll} \text{minimize} & f_0(Fz + x_0) & \text{(over } z) \\ \text{subject to} & f_i(Fz + x_0) \leq 0, & i = 1, \cdots, m \end{array}$$

where F and x_0 are such that

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

In many case, however, it is better to retain the equality constraints, since eliminating them can make the problem harder to understand and analyze, or ruin the efficiency of an algorithm that solves it.

introducing equality constraints

$$\begin{array}{ll} \text{minimize} & f_0(A_0x+b_0) \\ \text{subject to} & f_i(A_ix+b_i) \leq 0, \qquad i=1,\cdots,m \end{array}$$

is equivalent to

$$\begin{array}{lll} \text{minimize} & f_0(y_0) & \text{(over } x, y_i) \\ \text{subject to} & f_i(y_i) \leq 0, & i=1,\cdots,m \\ & y_i = A_ix+b_i, & i=0,1,\cdots,m \end{array}$$

introducing slack variables for linear inequalities

$$\begin{array}{lll} \text{minimize} & f_0(x) & \\ \text{subject to} & a_i^T x \leq b_i, & i = 1, \cdots, m \end{array}$$

is equivalent to

$$\begin{array}{lll} \text{minimize} & f_0(x) & \text{(over } x, s) \\ \text{subject to} & a_i^T x + s_i = b_i, & i = 1, \cdots, m \\ & s_i \geq 0, & i = 1, \cdots, m \end{array}$$

epigraph form

$$\begin{array}{lll} \text{minimize} & f_0(x) & \\ \text{subject to} & f_i(x) \leq 0, & i = 1, \cdots, m \\ & Ax = b & \end{array}$$

is equivalent to

$$\begin{array}{lll} \text{minimize} & t & (\text{over } x, t) \\ \text{subject to} & f_0(x) - t \leq 0 & \\ & f_i(x) \leq 0, & i = 1, \cdots, m \\ & Ax = b & \end{array}$$

partial minimization

$$\begin{array}{ll} \text{minimize} & f_0(x_1, x_2) \\ \text{subject to} & f_i(x_1) \leq 0, \qquad i = 1, \cdots, m \end{array}$$

is equivalent to

$$\begin{array}{ll} \text{minimize} & \tilde{f}_0(x_1) \\ \text{subject to} & f_i(x_1) \leq 0, \qquad i = 1, \cdots, m \end{array}$$

where

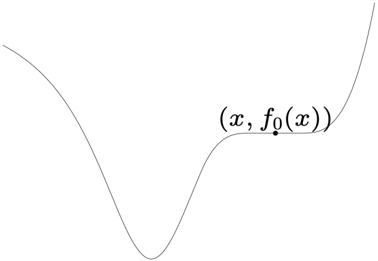
$$\tilde{f}_0(x_1) = \inf_{x_2} f_0(x_1, x_2)$$

Quasiconvex optimization

$$
\begin{array}{lll}
\text{minimize} & f_0(x) & \\
\text{subject to} & f_i(x) \le 0, & i = 1, \cdots, m \\
 & Ax = b &
\end{array}
$$

with $f_0 \colon \mathbb{R}^n \to \mathbb{R}$ quasiconvex, $f_1, \cdots, f_m$ convex.

Remark Locally optimal points may not be globally optimal



x is optimal if

$$x \in X, \ \nabla f_0(x)^T(y - x) > 0 \text{ for all } y \in X \backslash \{x\}.$$

- The condition is only sufficient for optimality
- The condition requires the gradient of f_0 to be nonzero

Convex representation of sublevel sets of f_0

For quasiconvex f_0 there exists a family of functions ϕ_t such that

- $\phi_t(x)$ is convex in x for each fixed t
- t-sublevel set of f_0 is 0-sublevel set of ϕ_t, i.e. $f_0(x) \le t \Longleftrightarrow \phi_t(x) \le 0$
- $\phi_t(x)$ is nonincreasing in t for each fixed x, namely $\phi_s(x) \le \phi_t(x)$ if $s \ge t$

In practice there are usually natural meaningful choices for ϕ_t.

Example

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

with p convex, q concave, and $p(x) \geq 0$, $q(x) > 0$ on $\mathbf{dom}\, f_0$.

We can choose

$$\phi_t(x) = p(x) - tq(x)$$

- $\phi_t(x)$ convex in x for $t \geq 0$
- $f_0(x) \leq t \quad \Longleftrightarrow \quad \phi_t(x) \leq 0$

Quasiconvex optimization via convex feasibility problems

$$\phi_t(x) \leq 0, \qquad f_i(x) \leq 0, \quad i = 1, \cdots, m, \qquad Ax = b$$

- convex feasibility problem in x for each fixed t
- let p^* be the optimal value for the original quasiconvex problem, then

$$\begin{aligned} \text{above problem feasible} \quad &\Longrightarrow \quad p^* \leq t \\ \text{above problem infeasible} \quad &\Longrightarrow \quad p^* \geq t \end{aligned}$$

Bisection method

---

given $l \leq p^\star$, $u \geq p^\star$, tolerance $\epsilon > 0$

repeat

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

until $u - l \leq \epsilon$

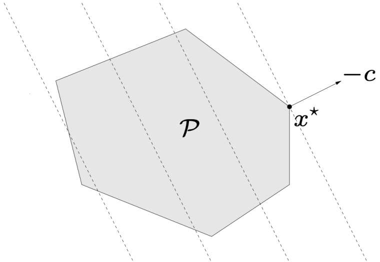
---

requires exactly $\lceil \log_2((u-l)/\epsilon) \rceil$ iterations

Linear program (LP)

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

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



Standard and inequality form linear programs

A standard form LP

$$\begin{array}{ll} \text{minimize} & c^T x \\ \text{subject to} & Ax = b \\ & x \succeq 0 \end{array}$$

An inequality form LP

$$\begin{array}{ll} \text{minimize} & c^T x \\ \text{subject to} & Ax \preceq b \end{array}$$

Converting LPs to standard form

Examples

Diet problem choose quantities $x_1, \cdots, x_n$ of n kinds of food

- 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

$$\begin{array}{ll} \text{minimize} & c^T x \\ \text{subject to} & Ax \succeq b \\ & x \succeq 0 \end{array}$$

Piecewise-linear minimization

$$\text{minimize} \quad \max \{a_i^T x + b_i \mid i = 1, \cdots, m\}$$

equivalent to the LP

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

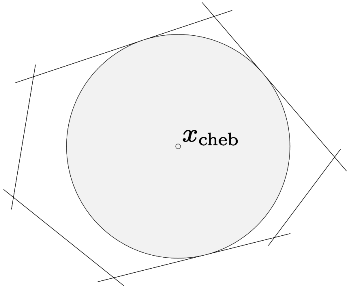
Chebyshev center of a polyhedron

Chebyshev center of

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

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

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

hence x_c and r can be determined by solving the LP

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

Linear-fractional program

$$\begin{array}{ll} \text{minimize} & f_0(x) \\ \text{subject to} & Gx \preceq h \\ & Ax = b \end{array}$$

where

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

is a quasiconvex optimization problem; can be solved by bisection method.

If the feasible set is nonempty, then the linear-fractional problem is equivalent to the LP

$$
\begin{array}{ll}
\text{minimize} & c^T y + dz \\
\text{subject to} & Gy \preceq hz \\
& Ay = bz \\
& e^T y + fz = 1 \\
& z \geq 0
\end{array}
$$

Generalized linear-fractional program

$$
\begin{array}{ll}
\text{minimize} & f_0(x) \\
\text{subject to} & Gx \preceq h \\
& Ax = b
\end{array}
$$

where

$$
f_0(x) = \mathbf{max}\left\{ \frac{c_i^T x + d_i}{e_i^T x + f_i} \,\middle|\, i = 1, \cdots, r \right\}
$$

$$
\mathbf{dom}\, f_0(x) = \left\{ x \,\middle|\, e_i^T x + f_i > 0,\ i = 1, \cdots, r \right\}
$$

is a quasiconvex optimization problem; can be solved by bisection method.

Example: von Neumann model of a growing economy

$$\begin{array}{lll} \text{maximize} & \min \left\{ x_i^+/x_i \,\middle|\, i = 1, \cdots, n \right\} & \text{(over } x, x^+\text{)} \\ \text{subject to} & x^+ \succeq 0 & \\ & Bx^+ \preceq Ax & \end{array}$$

with domain $\{(x, x^+) \mid x \succ 0\}$

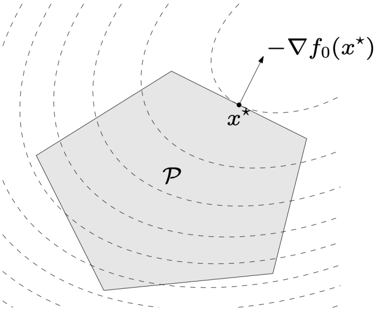
- $x, x^+ \in \mathbb{R}^n$: activity levels of n sectors, in current and next period
- $(Ax)_i$, $(Bx^+)_i$: produced resp. consumed amounts of good i
- x_i^+/x_i: growth rate of sector i

allocate activity to maximize growth rate of slowest growing sector

Quadratic program (QP)

$$\begin{array}{ll} \text{minimize} & (1/2)x^T P x + q^T x + r \\ \text{subject to} & Gx \preceq h \\ & Ax = b \end{array}$$

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



Example

Least-squares

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

- analytical solution $x^* = A^\dagger b$ (where $A^\dagger$ is pseudo-inverse)
- can add linear constraints such as $l \preceq x \preceq u$

Distance between polyhedra

The Euclidean distance between the polyhedra $\mathcal{P}_1 = \{x | A_1 x \preceq b_1\}$ and $\mathcal{P}_2 = \{x | A_2 x \preceq b_2\}$ in $\mathbb{R}^n$ is defined as

$$\mathbf{dist}(\mathcal{P}_1, \mathcal{P}_2) = \mathbf{inf}\{\|x_1 - x_2\|_2 \mid x_1 \in \mathcal{P}_1, x_2 \in \mathcal{P}_2\}$$

We solve the QP

$$\begin{array}{ll} \text{minimize} & \|x_1 - x_2\|_2^2 \\ \text{subject to} & A_1 x \preceq b_1, \qquad A_2 x \preceq b_2 \end{array}$$

- If the polyhedra intersect, the distance is zero
- The problem is infeasible if and only if one of the polyhedra is empty

Linear program with random cost

Consider the linear program

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

- Assume c is random vector with mean $\bar{c}$ and covariance Σ
- Then $c^T x$ is random variable with mean $\bar{c}^T x$ and variance $x^T \Sigma x$

$$
\begin{aligned}
\mathbf{E}\left(c^T x\right) &= \mathbf{E}\left(c\right)^T x = \bar{c}^T x \\
\mathbf{var}\left(c^T x\right) &= \mathbf{E}\left(c^T x - \bar{c}^T x\right)^2 = x^T \mathbf{E}\left((c-\bar{c})(c-\bar{c})^T\right) x = x^T \Sigma x
\end{aligned}
$$

We modify the above LP to the following QP

$$\begin{array}{ll} \text{minimize} & \bar{c}^T x + \gamma x^T \Sigma x \\ \text{subject to} & Gx \preceq h \\ & Ax = b \end{array}$$

- To keep both the expected cost and the cost variance (risk) under control, choose a linear combination of both as the new objective, called **risk-sensitive cost**.
- $\gamma > 0$ is the **risk-aversion parameter**, which controls the trade-off between expected cost and variance.
- Coefficient vector $(1, \gamma)$ lies in the interior of the dual cone of the nonnegative quadrant.

Quadratically constrained quadratic program (QCQP)

$$
\begin{array}{lll}
\text{minimize} & (1/2)x^T P_0 x + q_0^T x + r_0 & \\
\text{subject to} & (1/2)x^T P_i x + q_i^T x + r_i \leq 0, & i = 1, \cdots, m \\
& Ax = b &
\end{array}
$$

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

Second-order cone program (SOCP)

$$\begin{array}{ll} \text{minimize} & f^T x \\ \text{subject to} & \|A_i x + b_i\|_2 \le c_i^T x + d_i, \qquad i = 1, \cdots, m \\ & Fx = G \end{array}$$

with $A_i \in \mathbb{R}^{n_i \times n}$ and $F \in \mathbb{R}^{p \times n}$

- inequalities are called second-order cone constraints since

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

- if $n_i = 0$, reduces to LP
- if $c_i = 0$, reduces to QCQP (with linear objective)

Robust linear program

Parameters in optimization problems are often uncertain. Consider the LP

$$\begin{array}{ll} \text{minimize} & c^T x \\ \text{subject to} & a_i^T x \leq b_i, \qquad i = 1, \cdots, m \end{array}$$

- There can be uncertainty in c, a_i, b_i (in a_i for example)
- There are two common approaches to handle uncertainty
  - deterministic model
  - stochastic model

- deterministic model: constraints must hold for all $a_i \in \mathcal{E}_i$

$$\begin{array}{ll} \text{minimize} & c^T x \\ \text{subject to} & a_i^T x \leq b_i \text{ for all } a_i \in \mathcal{E}_i, \qquad i = 1, \cdots, m \end{array}$$

- stochastic model: a_i is random variable; constraints must hold with probability η

$$\begin{array}{ll} \text{minimize} & c^T x \\ \text{subject to} & \mathbf{prob}(a_i^T x \leq b_i) \geq \eta, \qquad i = 1, \cdots, m \end{array}$$

deterministic approach via SOCP

- choose ellipsoid as $\mathcal{E}_i$ with $\bar{a}_i \in \mathbb{R}^n$ and $P_i \in \mathbb{R}^{n \times n}$

$$\mathcal{E}_i = \{\bar{a}_i + P_i u \mid \|u\|_2 \leq 1\}$$

- robust LP

$$\begin{array}{ll} \text{minimize} & c^T x \\ \text{subject to} & a_i^T x \leq b_i \text{ for all } a_i \in \mathcal{E}_i, \qquad i = 1, \cdots, m \end{array}$$

- equivalent SOCP

$$\begin{array}{ll} \text{minimize} & c^T x \\ \text{subject to} & \bar{a}_i^T x + \|P_i^T x\|_2 \leq b_i, \qquad i = 1, \cdots, m \end{array}$$

which follows from

$$\sup_{\|u\|_2 \leq 1} (\bar{a}_i + P_i u)^T x = \bar{a}_i^T x + \|P_i^T x\|_2$$

stochastic approach via SOCP

- assume $a_i \sim \mathcal{N}(\bar{a}_i, \Sigma_i)$ is Gaussian, then $a_i^T x \sim \mathcal{N}(\bar{a}_i^T x, x^T \Sigma_i x)$ is also Gaussian

$$\mathbf{prob}(a_i^T x \leq b_i) = \Phi\left(\frac{b_i - \bar{a}_i^T x}{\|\Sigma_i^{1/2} x\|_2}\right)$$

  with $\Phi(x) = (1/\sqrt{2\pi}) \int_{-\infty}^{x} e^{-t^2/2}\, \mathrm{d}t$ cumulative distribution function of $\mathcal{N}(0,1)$
- robust LP

$$\begin{array}{ll} \text{minimize} & c^T x \\ \text{subject to} & \mathbf{prob}(a_i^T x \leq b_i) \geq \eta, \qquad i = 1, \cdots, m \end{array}$$

- equivalent SOCP when $\eta > 1/2$

$$\begin{array}{ll} \text{minimize} & c^T x \\ \text{subject to} & \bar{a}_i^T x + \Phi^{-1}(\eta) \|\Sigma_i^{1/2} x\|_2 \leq b_i, \qquad i = 1, \cdots, m \end{array}$$

Minimal surface p 159

Monomials and posynomials

- **monomial function**

$$f(x) = cx_1^{a_1} \cdots x_n^{a_n}, \qquad \mathbf{dom}\, f = \mathbb{R}^n_{++}$$

with $c > 0$ and $a_i \in \mathbb{R}$

- **posynomial function**

$$f(x) = \sum_{k=1}^{K} c_k x_1^{a_{1k}} \cdots x_n^{a_{nk}}, \qquad \mathbf{dom}\, f = \mathbb{R}^n_{++}$$

sum of monomials

change variables to $y_i = \log x_i$ and take logarithm

- monomial $f(x) = c x_1^{a_1} \cdots x_n^{a_n}$ transforms to

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

- posynomial $f(x) = \sum_{k=1}^{K} c_k x_1^{a_{1k}} \cdots x_n^{a_{nk}}$ transforms to

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

Geometric program (GP)

geometric program in standard form

$$
\begin{array}{ll}
\text{minimize} & f_0(x) \\
\text{subject to} & f_i(x) \leq 1, \qquad i = 1, \cdots, m \\
 & h_i(x) = 1, \qquad i = 1, \cdots, p
\end{array}
$$

with f_i posynomial, h_i monomial

geometric program in convex form

change variables to $y_i = \log x_i$ and take logarithm of objective and constraints

$$
\begin{array}{ll}
\text{minimize} & \log\left(\sum_{k=1}^{K} e^{a_{0k}^T y + b_{0k}}\right) \\
\text{subject to} & \log\left(\sum_{k=1}^{K} e^{a_{ik}^T y + b_{ik}}\right) \leq 0, \qquad i = 1, \cdots, m \\
& Gy + d = 0
\end{array}
$$

Example

Frobenius norm diagonal scaling

- Assume $M \in \mathbb{R}^{n \times n}$ defines a linear transformation. After scaling the coordinates by $D = \mathbf{diag}(d_1, \ldots, d_n) \in \mathbb{R}^{n \times n}$, the resulting matrix becomes DMD^{-1}.
- How to choose D such that DMD^{-1} is small under the Frobenius norm?

$$\|DMD^{-1}\|_F^2 = \sum_{i,j=1}^{n} (DMD^{-1})_{ij}^2 = \sum_{i,j=1}^{n} M_{ij}^2 d_i^2 / d_j^2.$$

- It is an unconstrained geometric program

$$\text{minimize} \qquad \sum_{i,j=1}^{n} M_{ij}^2 d_i^2 / d_j^2$$

with variable $d = (d_1, \ldots, d_n)$.

Convex problem with generalized inequality constraints

$$\begin{array}{ll} \text{minimize} & f_0(x) \\ \text{subject to} & f_i(x) \preceq_{K_i} 0, \qquad i = 1, \cdots, m \\ & Ax = b \end{array}$$

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

Conic form problem (cone program)

special case of above with affine objective and constraints

$$\begin{array}{ll} \text{minimize} & c^T x \\ \text{subject to} & Fx + g \preceq_K 0 \\ & Ax = b \end{array}$$

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

SOCP as equivalent conic form programming

SOCP

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

equivalent conic form programming

$$\begin{array}{ll} \text{minimize} & f^T x \\ \text{subject to} & -\left(A_i x + b_i, c_i^T x + d_i\right) \preceq_{K_i} 0, \qquad i = 1, \cdots m \end{array}$$

in which

$$K_i = \{(y, t) \in \mathbb{R}^{n_i+1} \mid \|y\|_2 \leq t\}$$

Semidefinite program (SDP)

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

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

- inequality constraint is called **linear matrix inequality** (LMI)
- includes problems with multiple LMI constrains:

$$x_1 F_1' + \cdots + x_n F_n' + G' \preceq 0 \quad \text{and} \quad x_1 F_1'' + \cdots + x_n F_n'' + G'' \preceq 0$$

  is equivalent to single LMI

$$x_1 \begin{bmatrix} F_1' & 0 \\ 0 & F_1'' \end{bmatrix} + \cdots + x_n \begin{bmatrix} F_n' & 0 \\ 0 & F_n'' \end{bmatrix} + \begin{bmatrix} G' & 0 \\ 0 & G'' \end{bmatrix} \preceq 0$$

LP as equivalent SDP

LP

$$\begin{array}{ll} \text{minimize} & c^T x \\ \text{subject to} & Ax \preceq b \end{array}$$

equivalent SDP

$$\begin{array}{ll} \text{minimize} & c^T x \\ \text{subject to} & \mathbf{diag}(Ax - b) \preceq 0 \end{array}$$

note different interpretation of generalized inequality

Eigenvalue minimization

$$\text{minimize} \quad \lambda_{\max}(A(x))$$

where $A(x) = A_0 + x_1 A_1 + \cdots + x_n A_n$ with given $A_i \in \mathbb{S}^k$

equivalent SDP with variables $(x, t) \in \mathbb{R}^{n+1}$

$$\begin{array}{ll} \text{minimize} & t \\ \text{subject to} & A(x) \preceq tI \end{array}$$

follows from

$$\lambda_{\max}(A) \leq t \quad \Longleftrightarrow \quad A \preceq tI$$

Matrix norm minimization

$$\text{minimize} \quad \|A(x)\|_2 = \left(\lambda_{\mathbf{max}}\left(A(x)^T A(x)\right)\right)^{1/2}$$

where $A(x) = A_0 + x_1 A_1 + \cdots + x_n A_n$ with given $A_i \in \mathbb{R}^{p\times q}$

equivalent SDP with variables $x \in \mathbb{R}^n$ and $t \in \mathbb{R}$

$$\begin{array}{ll} \text{minimize} & t \\ \text{subject to} & \begin{bmatrix} tI & A(x) \\ A(x)^T & tI \end{bmatrix} \succeq 0 \end{array}$$

follows from

$$\begin{aligned} \|A\|_2 \le t \quad &\Longleftrightarrow \quad A^T A \preceq t^2 I, \quad t \ge 0 \\ &\Longleftrightarrow \quad \begin{bmatrix} tI & A \\ A^T & tI \end{bmatrix} \succeq 0 \end{aligned}$$

Vector optimization

general vector optimization problem

$$\begin{array}{lll} \text{minimize (with respect to } K) & f_0(x) & \\ \text{subject to} & f_i(x) \leq 0, & i = 1, \cdots, m \\ & h_i(x) = 0, & i = 1, \cdots, p \end{array}$$

vector objective $f_0 \colon \mathbb{R}^n \to \mathbb{R}^q$ minimized with respect to proper cone $K \subseteq \mathbb{R}^q$

convex vector optimization problem

$$\begin{array}{lll} \text{minimize (with respect to } K) & f_0(x) & \\ \text{subject to} & f_i(x) \leq 0, & i = 1, \cdots, m \\ & Ax = b & \end{array}$$

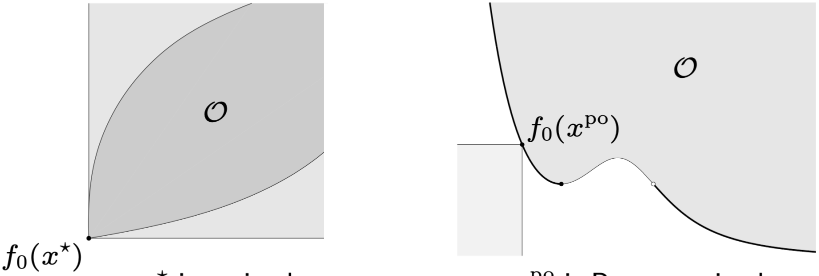
where f_0 is K-convex and $f_1, \cdots, f_m$ are convex

Optimal and Pareto optimal points

set of achievable objective values

$$\mathcal{O} = \{f_0(x) \mid x \text{ feasible}\}$$

- feasible x^* is optimal if $f_0(x^*)$ is the minimum value of $\mathcal{O}$ (optimal value)
- feasible x^{po} is Pareto optimal if $f_0(x^{\text{po}})$ is a minimal value of $\mathcal{O}$ (Pareto optimal value)



$x^\star$ is optimal $\quad\quad$ x^{po} is Pareto optimal

Example

Best linear unbiased estimator p176

Multicriterion optimization

vector optimization problem with $K = \mathbb{R}^q_+$

$$f_0(x) = (F_1(x), \cdots, F_q(x))$$

- q different objectives F_i, we want all of them to be small
- feasible x^* is optimal if

$$y \text{ feasible} \quad \Longrightarrow \quad f_0(x^*) \preceq f_0(y)$$

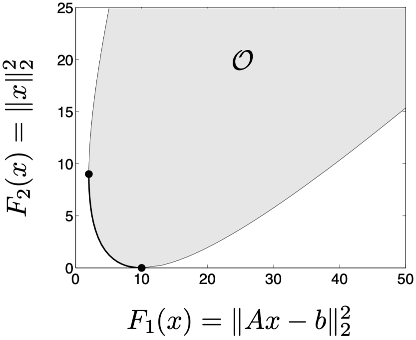
  if an optimal point exists, the objectives are noncompeting
- feasible x^{po} is Pareto optimal if

$$y \text{ feasible}, \quad f_0(y) \preceq f_0(x^{\text{po}}) \quad \Longrightarrow \quad f_0(x^{\text{po}}) = f_0(y)$$

  if multiple Pareto optimal values exist, there is a trade-off between the objectives

Regularized least-squares

$$\text{minimize (with respect to } \mathbb{R}^2_+) \qquad \left(\|Ax-b\|_2^2, \|x\|_2^2\right)$$

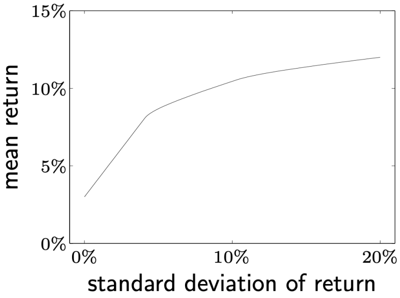


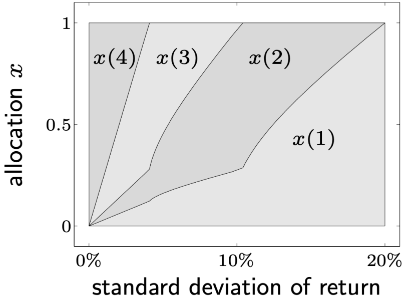
the optimal trade-off curve, shown darker, is formed by Pareto optimal points

Risk-return trade-off in portfolio optimization

$$\begin{array}{ll} \text{minimize (with respect to } \mathbb{R}^2_+) & (-\bar{p}^T x, x^T \Sigma x) \\ \text{subject to} & \mathbf{1}^T x = 1 \\ & x \succeq 0 \end{array}$$

- $x \in \mathbb{R}^n$ investment portfolio; x_i fraction invested in asset i
- $p \in \mathbb{R}^n$ (relative) asset price, random variable with mean $\bar{p}$ and covariance Σ
- $r = p^T x$ (relative) return, random variable with mean $\bar{p}^T x$ and variance $x^T \Sigma x$





Scalarization

To find Pareto optimal points, choose $\lambda \succ_{K^*} 0$ and solve scalar problem

$$\begin{array}{lll} \text{minimize} & \lambda^T f_0(x) & \\ \text{subject to} & f_i(x) \le 0, & i = 1, \cdots, m \\ & h_i(x) = 0, & i = 1, \cdots, p \end{array}$$

- if x is optimal for scalar problem, then it is Pareto optimal for vector optimization problem
- for convex vector optimization problem, can find (almost) all Pareto optimal points by varying $\lambda \succ_{K^*} 0$

Scalarization for multicriterion problems

In this more concrete situation

$$K = K^* = \mathbb{R}^q_+.$$

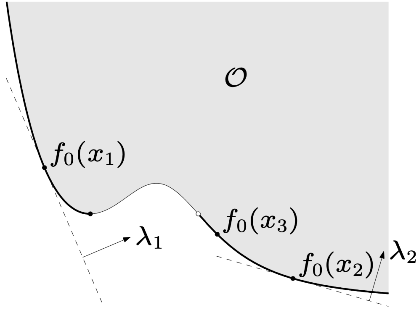
To find Pareto optimal points, write

$$\lambda = \begin{bmatrix} a_1 \\ \vdots \\ a_q \end{bmatrix} \in \mathbb{R}^q_{++} \qquad \text{and} \qquad f_0(x) = \begin{bmatrix} F_1(x) \\ \vdots \\ F_q(x) \end{bmatrix},$$

then minimize the positive weighted sum

$$\lambda^T f_0(x) = a_1 F_1(x) + \cdots + a_q F_q(x)$$

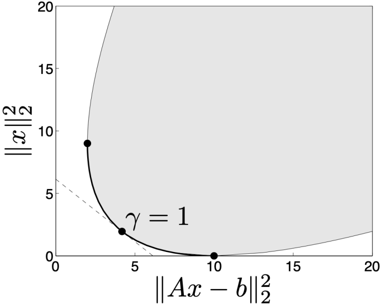
Geometric interpretation



- $\mathcal{O}$ is the set of achievable objective values
- Pareto optimal values $f_0(x_1)$ and $f_0(x_2)$ can both be obtained by scalarization: $f_0(x_1)$ minimizes $\lambda_1^T u$ and $f_0(x_2)$ minimizes $\lambda_2^T u$ over all $u \in \mathcal{O}$
- $f_0(x_3)$ is Pareto optimal, but cannot be found by scalarization

Examples

Regularized least-square problem



Take $\lambda = (1, \gamma)$ with $\gamma > 0$

$$\text{minimize} \qquad \|Ax - b\|_2^2 + \gamma \|x\|_2^2$$

least-square problem for fixed $\gamma > 0$

Risk-return trade-off problem

Take $\lambda = (1, \gamma)$ with $\gamma > 0$

$$\begin{array}{ll} \text{minimize} & -\bar{p}^T x + \gamma x^T \Sigma x \\ \text{subject to} & \mathbf{1}^T x = 1 \\ & x \succeq 0 \end{array}$$

quadratic program for each fixed $\gamma > 0$