Convex Optimization

Stephen Boyd   Lieven Vandenberghe

Revised slides by Stephen Boyd, Lieven Vandenberghe, and Parth Nobel

4. Convex optimization problems

Outline

Optimization problems

Some standard convex problems

Transforming problems

Disciplined convex programming

Geometric programming

Quasiconvex optimization

Multicriterion optimization

Optimization problem in standard form

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

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

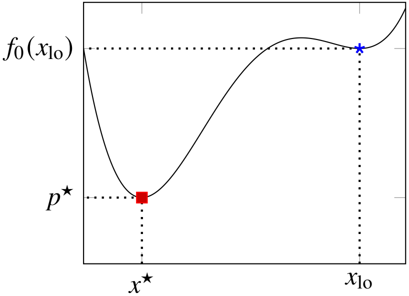
Feasible and optimal points

- $x \in \mathbf{R}^n$ is **feasible** if $x \in \mathbf{dom}\, f_0$ and it satisfies the constraints
- **optimal value** is $p^\star = \inf\{f_0(x) \mid f_i(x) \leq 0,\ i = 1, \ldots, m,\ h_i(x) = 0,\ i = 1, \ldots, p\}$
- $p^\star = \infty$ if problem is infeasible
- $p^\star = -\infty$ if problem is **unbounded below**
- a feasible x is **optimal** if $f_0(x) = p^\star$
- X_{opt} is the set of optimal points

Locally optimal points

x is **locally optimal** if there is an $R > 0$ such that x is optimal for

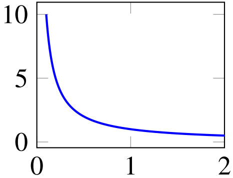
$$\begin{array}{ll} \text{minimize (over } z\text{)} & f_0(z) \\ \text{subject to} & f_i(z) \leq 0, \quad i = 1, \ldots, m, \quad h_i(z) = 0, \quad i = 1, \ldots, p \\ & \|z - x\|_2 \leq R \end{array}$$



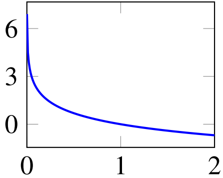
Examples

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

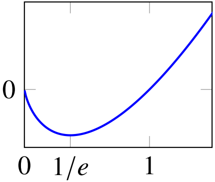
- $f_0(x) = 1/x$, $\mathbf{dom}\, f_0 = \mathbf{R}_{++}$: $p^\star = 0$, no optimal point
- $f_0(x) = -\log x$, $\mathbf{dom}\, f_0 = \mathbf{R}_{++}$: $p^\star = -\infty$
- $f_0(x) = x \log x$, $\mathbf{dom}\, f_0 = \mathbf{R}_{++}$: $p^\star = -1/e$, $x = 1/e$ is optimal
- $f_0(x) = x^3 - 3x$: $p^\star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



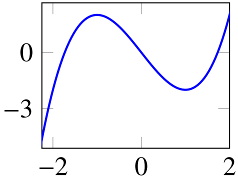
$f_0(x) = 1/x$



$f_0(x) = -\log x$



$f_0(x) = x \log x$



$f_0(x) = x^3 - 3x$

Implicit and explicit constraints

standard form optimization problem has **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:

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

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

Feasibility problem

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

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

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

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

Standard form convex optimization problem

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

- objective and inequality constraints $f_0, f_1, \ldots, f_m$ are convex
- equality constraints are affine, often written as $Ax = b$
- feasible and optimal sets of a convex optimization problem are convex

- problem is **quasiconvex** if f_0 is quasiconvex, $f_1, \ldots, f_m$ are convex, $h_1, \ldots, h_p$ are affine

Example

- standard form problem

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

- f_0 is convex; feasible set $\{(x_1, x_2) \mid x_1 = -x_2 \leq 0\}$ is convex
- not a convex problem (by our definition) since f_1 is not convex, h_1 is not affine
- equivalent (but not identical) to the convex problem

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

Local and global optima

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

proof:

- suppose x is locally optimal, but there exists a feasible y with $f_0(y) < f_0(x)$
- x locally optimal means there is an $R > 0$ such that

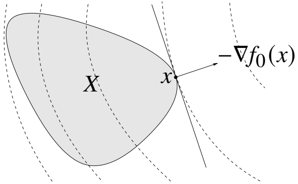
$$z \text{ feasible}, \quad \|z - x\|_2 \leq R \quad \Longrightarrow \quad f_0(z) \geq f_0(x)$$

- 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 $f_0(z) \leq \theta f_0(y) + (1 - \theta)f_0(x) < f_0(x)$, which contradicts our assumption that x is locally optimal

Optimality criterion for differentiable f_0

- x is optimal for a convex problem 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

Examples

- **unconstrained problem**: x minimizes $f_0(x)$ if and only if $\nabla f_0(x) = 0$
- **equality constrained problem**: x minimizes $f_0(x)$ subject to $Ax = b$ if and only if there exists a ν such that

$$Ax = b, \qquad \nabla f_0(x) + A^T \nu = 0$$

- **minimization over nonnegative orthant**: x minimizes $f_0(x)$ over $\mathbf{R}^n_+$ if and only if

$$x \succeq 0, \qquad \left\{ \begin{array}{ll} \nabla f_0(x)_i \geq 0 & x_i = 0 \\ \nabla f_0(x)_i = 0 & x_i > 0 \end{array} \right.$$

Outline

Optimization problems

Some standard convex problems

Transforming problems

Disciplined convex programming

Geometric programming

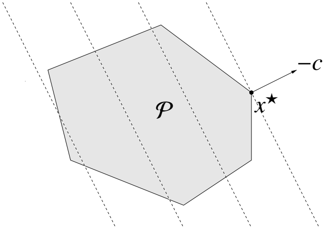
Quasiconvex optimization

Multicriterion optimization

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



Example: Diet problem

- choose nonnegative quantities $x_1, \ldots, x_n$ of n foods
- one unit of food j costs c_j and contains amount A_{ij} of nutrient i
- healthy diet requires nutrient i in quantity at least b_i
- to find cheapest healthy diet, solve

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

- express in standard LP form as

$$\begin{array}{ll} \text{minimize} & c^T x \\ \text{subject to} & \left[\begin{array}{c} -A \\ -I \end{array}\right] x \leq \left[\begin{array}{c} -b \\ 0 \end{array}\right] \end{array}$$

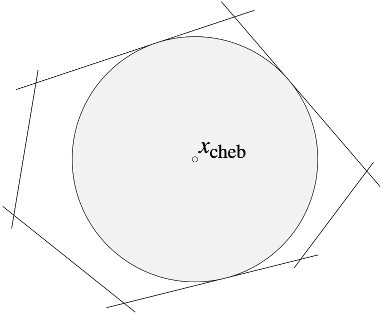
Example: Piecewise-linear minimization

- minimize convex piecewise-linear function $f_0(x) = \max_{i=1,\ldots,m}(a_i^T x + b_i)$, $x \in \mathbf{R}^n$
- equivalent to LP

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

  with variables $x \in \mathbf{R}^n$, $t \in \mathbf{R}$
- constraints describe $\mathbf{epi}\, f_0$

Example: Chebyshev center of a polyhedron



Chebyshev center of $\mathcal{P} = \{x \mid a_i^T x \leq b_i,\ i = 1, \ldots, m\}$ 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

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

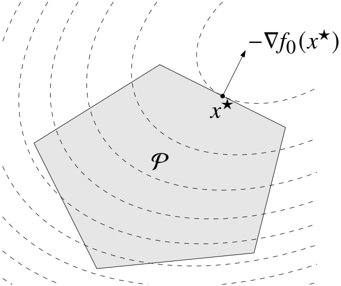
- hence, x_c, r can be determined by solving LP with variables x_c, r

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

Quadratic program (QP)

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

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



Example: Least squares

- **least squares** problem: minimize $\|Ax - b\|_2^2$
- analytical solution $x^\star = A^\dagger b$ ($A^\dagger$ is pseudo-inverse)
- can add linear constraints, *e.g.*,
  - $x \geq 0$ (**nonnegative least squares**)
  - $x_1 \leq x_2 \leq \cdots \leq x_n$ (**isotonic regression**)

Example: Linear program with random cost

- LP with random cost c, with mean $\bar{c}$ and covariance Σ
- hence, LP objective $c^T x$ is random variable with mean $\bar{c}^T x$ and variance $x^T \Sigma x$
- **risk-averse** problem:

$$\begin{array}{ll} \text{minimize} & \mathbf{E}\, c^T x + \gamma\, \mathbf{var}(c^T x) \\ \text{subject to} & Gx \preceq h, \quad Ax = b \end{array}$$

- $\gamma > 0$ is **risk aversion parameter**; controls the trade-off between expected cost and variance (risk)
- express as QP

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

Quadratically constrained quadratic program (QCQP)

$$\begin{array}{ll}\text{minimize} & (1/2)x^TP_0x + q_0^Tx + r_0 \\ \text{subject to} & (1/2)x^TP_ix + q_i^Tx + r_i \leq 0, \quad i = 1, \ldots, m \\ & Ax = b\end{array}$$

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

Second-order cone programming

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

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

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

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

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

Example: Robust linear programming

suppose constraint vectors a_i are uncertain in the LP

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

two common approaches to handling uncertainty

- **deterministic worst-case**: constraints must hold for all $a_i \in \mathcal{E}_i$ (uncertainty ellipsoids)

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

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

$$\begin{array}{ll} \text{minimize} & c^T x \\ \text{subject to} & \mathbf{prob}(a_i^T x \le b_i) \ge \eta, \quad i = 1, \ldots, m \end{array}$$

Deterministic worst-case approach

- uncertainty ellipsoids are $\mathcal{E}_i = \{\bar{a}_i + P_i u \mid \|u\|_2 \leq 1\}$, ($\bar{a}_i \in \mathbf{R}^n$, $P_i \in \mathbf{R}^{n \times n}$)
- center of $\mathcal{E}_i$ is $\bar{a}_i$; semi-axes determined by singular values/vectors of P_i
- robust LP

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

- equivalent to SOCP

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

(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

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

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

where $\Phi(u) = (1/\sqrt{2\pi}) \int_{-\infty}^{u} e^{-t^2/2}\, dt$ is $\mathcal{N}(0,1)$ CDF

- $\mathbf{prob}(a_i^T x \leq b_i) \geq \eta$ can be expressed as $\bar{a}_i^T x + \Phi^{-1}(\eta) \|\Sigma_i^{1/2} x\|_2 \leq b_i$
- for $\eta \geq 1/2$, robust LP equivalent to SOCP

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

Conic form problem

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

- constraint $Fx + g \preceq_K 0$ involves a generalized inequality with respect to a proper cone K
- linear programming is a conic form problem with $K = \mathbf{R}^m_+$
- as with standard convex problem
  - feasible and optimal sets are convex
  - any local optimum is global

Semidefinite program (SDP)

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

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

- inequality constraint is called **linear matrix inequality** (LMI)
- includes problems with multiple LMI constraints: for example,

$$x_1 \hat{F}_1 + \cdots + x_n \hat{F}_n + \hat{G} \preceq 0, \qquad x_1 \tilde{F}_1 + \cdots + x_n \tilde{F}_n + \tilde{G} \preceq 0$$

  is equivalent to single LMI

$$x_1 \left[\begin{array}{cc} \hat{F}_1 & 0 \\ 0 & \tilde{F}_1 \end{array} \right] + x_2 \left[\begin{array}{cc} \hat{F}_2 & 0 \\ 0 & \tilde{F}_2 \end{array} \right] + \cdots + x_n \left[\begin{array}{cc} \hat{F}_n & 0 \\ 0 & \tilde{F}_n \end{array} \right] + \left[\begin{array}{cc} \hat{G} & 0 \\ 0 & \tilde{G} \end{array} \right] \preceq 0$$

Example: Matrix norm minimization

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

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

equivalent SDP

$$\begin{array}{ll} \text{minimize} & t \\ \text{subject to} & \left[\begin{array}{cc} tI & A(x) \\ A(x)^T & tI \end{array}\right] \succeq 0 \end{array}$$

- variables $x \in \mathbf{R}^n$, $t \in \mathbf{R}$
- constraint follows from

$$\begin{aligned} \|A\|_2 \leq t \quad & \Longleftrightarrow \quad A^T A \preceq t^2 I, \quad t \geq 0 \\ & \Longleftrightarrow \quad \left[\begin{array}{cc} tI & A \\ A^T & tI \end{array}\right] \succeq 0 \end{aligned}$$

Outline

Optimization problems

Some standard convex problems

Transforming problems

Disciplined convex programming

Geometric programming

Quasiconvex optimization

Multicriterion optimization

Change of variables

- $\phi : \mathbf{R}^n \to \mathbf{R}^n$ is one-to-one with $\phi(\mathbf{dom}\,\phi) \supseteq \mathcal{D}$
- consider (possibly non-convex) problem

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

- change variables to z with $x = \phi(z)$
- can solve equivalent problem

$$\begin{array}{ll} \text{minimize} & \tilde{f}_0(z) \\ \text{subject to} & \tilde{f}_i(z) \leq 0, \quad i = 1, \ldots, m \\ & \tilde{h}_i(z) = 0, \quad i = 1, \ldots, p \end{array}$$

  where $\tilde{f}_i(z) = f_i(\phi(z))$ and $\tilde{h}_i(z) = h_i(\phi(z))$
- recover original optimal point as $x^\star = \phi(z^\star)$

Example

- **non-convex** problem

$$\begin{array}{ll} \text{minimize} & x_1/x_2 + x_3/x_1 \\ \text{subject to} & x_2/x_3 + x_1 \leq 1 \end{array}$$

  with implicit constraint $x \succ 0$

- change variables using $x = \phi(z) = \exp z$ to get

$$\begin{array}{ll} \text{minimize} & \exp(z_1 - z_2) + \exp(z_3 - z_1) \\ \text{subject to} & \exp(z_2 - z_3) + \exp(z_1) \leq 1 \end{array}$$

  which is **convex**

Transformation of objective and constraint functions

suppose

- ϕ_0 is monotone increasing
- $\psi_i(u) \leq 0$ if and only if $u \leq 0$, $i = 1, \dots, m$
- $\varphi_i(u) = 0$ if and only if $u = 0$, $i = 1, \dots, p$

standard form optimization problem is equivalent to

$$\begin{array}{lll} \text{minimize} & \phi_0(f_0(x)) & \\ \text{subject to} & \psi_i(f_i(x)) \leq 0, & i = 1, \dots, m \\ & \varphi_i(h_i(x)) = 0, & i = 1, \dots, p \end{array}$$

example: minimizing $\|Ax - b\|$ is equivalent to minimizing $\|Ax - b\|^2$

Converting maximization to minimization

- suppose ϕ_0 is monotone decreasing
- the maximization problem

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

  is equivalent to the minimization problem

$$\begin{array}{lll} \text{minimize} & \phi_0(f_0(x)) & \\ \text{subject to} & f_i(x) \leq 0, & i = 1, \ldots, m \\ & h_i(x) = 0, & i = 1, \ldots, p \end{array}$$

- **examples**:
  - $\phi_0(u) = -u$ transforms maximizing a concave function to minimizing a convex function
  - $\phi_0(u) = 1/u$ transforms maximizing a concave positive function to minimizing a convex function

Eliminating equality constraints

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

is equivalent to

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

where F and x_0 are such that $Ax = b \iff x = Fz + x_0$ for some z

Introducing equality constraints

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

is equivalent to

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

Introducing slack variables for linear inequalities

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

is equivalent to

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

Epigraph form

standard form convex problem is equivalent to

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

Minimizing over some variables

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

is equivalent to

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

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

LP and SOCP as SDP

LP and equivalent SDP

LP: $\begin{array}{ll} \text{minimize} & c^T x \\ \text{subject to} & Ax \preceq b \end{array}$ 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 inequalities $\preceq$ in LP and SDP)

SOCP and equivalent SDP

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

SDP:
$$\begin{array}{ll} \text{minimize} & f^T x \\ \text{subject to} & \left[\begin{array}{cc} (c_i^T x + d_i)I & A_i x + b_i \\ (A_i x + b_i)^T & c_i^T x + d_i \end{array} \right] \succeq 0, \quad i = 1, \ldots, m \end{array}$$

Convex relaxation

- start with **nonconvex problem**: minimize $h(x)$ subject to $x \in C$
- find convex function $\hat{h}$ with $\hat{h}(x) \leq h(x)$ for all $x \in \mathbf{dom}\, h$ (*i.e.*, a pointwise lower bound on h)
- find set $\hat{C} \supseteq C$ (*e.g.*, $\hat{C} = \mathbf{conv}\, C$) described by linear equalities and convex inequalities

$$\hat{C} = \{x \mid f_i(x) \leq 0,\ i = 1, \ldots, m,\ f_m(x) \leq 0,\ Ax = b\}$$

- convex problem

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

  is a **convex relaxation** of the original problem
- optimal value of relaxation is lower bound on optimal value of original problem

Example: Boolean LP

- **mixed integer linear program (MILP)**:

$$\begin{array}{ll} \text{minimize} & c^T(x,z) \\ \text{subject to} & F(x,z) \preceq g, \quad A(x,z) = b, \quad z \in \{0,1\}^q \end{array}$$

  with variables $x \in \mathbf{R}^n$, $z \in \mathbf{R}^q$

- z_i are called **Boolean variables**
- this problem is in general hard to solve

- **LP relaxation**: replace $z \in \{0,1\}^q$ with $z \in [0,1]^q$
- optimal value of relaxation LP is lower bound on MILP
- can use as heuristic for approximately solving MILP, *e.g.*, **relax and round**

Outline

Optimization problems

Some standard convex problems

Transforming problems

Disciplined convex programming

Geometric programming

Quasiconvex optimization

Multicriterion optimization

Disciplined convex program

- specify objective as
  - minimize {scalar convex expression}, or
  - maximize {scalar concave expression}
- specify constraints as
  - {convex expression} <= {concave expression} or
  - {concave expression} >= {convex expression} or
  - {affine expression} == {affine expression}
- curvature of expressions are DCP certified, *i.e.*, follow composition rule
- DCP-compliant problems can be automatically transformed to standard forms, then solved

CVXPY example

math:

$$\begin{array}{ll} \text{minimize} & \|x\|_1 \\ \text{subject to} & Ax = b \\ & \|x\|_\infty \leq 1 \end{array}$$

- x is the variable
- A, b are given

CVXPY code:

```
import cvxpy as cp

A, b = ...

x = cp.Variable(n)
obj = cp.norm(x, 1)
constr = [
  A @ x == b,
  cp.norm(x, 'inf') <= 1,
]
prob = cp.Problem(cp.Minimize(obj), constr)
prob.solve()
```

How CVXPY works

- starts with your optimization problem $\mathcal{P}_1$
- finds a sequence of equivalent problems $\mathcal{P}_2, \ldots, \mathcal{P}_N$
- final problem $\mathcal{P}_N$ matches a standard form (*e.g.*, LP, QP, SOCP, or SDP)
- calls a specialized solver on $\mathcal{P}_N$
- retrieves solution of original problem by reversing the transformations

$$\begin{tikzcd}
\mathcal{P}_1 \arrow[r, Leftrightarrow] & \mathcal{P}_2 \arrow[r, Leftrightarrow] & \cdots \arrow[r, Leftrightarrow] & \mathcal{P}_{N-1} \arrow[r, Leftrightarrow] & \mathcal{P}_N \\
\text{your problem} \arrow[u] & & & & \text{standard problem} \arrow[u]
\end{tikzcd}$$

Outline

Optimization problems

Some standard convex problems

Transforming problems

Disciplined convex programming

Geometric programming

Quasiconvex optimization

Multicriterion optimization

Geometric programming

- **monomial function**:
$$f(x) = cx_1^{a_1} x_2^{a_2} \cdots x_n^{a_n}, \qquad \mathbf{dom}\, f = \mathbf{R}^n_{++}$$
with $c > 0$; exponent a_i can be any real number
- **posynomial function**: sum of monomials
$$f(x) = \sum_{k=1}^{K} c_k x_1^{a_{1k}} x_2^{a_{2k}} \cdots x_n^{a_{nk}}, \qquad \mathbf{dom}\, f = \mathbf{R}^n_{++}$$
- **geometric program (GP)**
$$\begin{array}{ll} \text{minimize} & f_0(x) \\ \text{subject to} & f_i(x) \leq 1, \quad i = 1, \ldots, m \\ & h_i(x) = 1, \quad i = 1, \ldots, p \end{array}$$
with f_i posynomial, h_i monomial

Geometric program in convex form

- change variables to $y_i = \log x_i$, and take logarithm of cost, constraints
- monomial $f(x) = cx_1^{a_1} \cdots x_n^{a_n}$ transforms to

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

- posynomial $f(x) = \sum_{k=1}^{K} c_k x_1^{a_{1k}} x_2^{a_{2k}} \cdots x_n^{a_{nk}}$ transforms to

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

- geometric program transforms to convex problem

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

Examples: Frobenius norm diagonal scaling

- we seek diagonal matrix $D = \mathbf{diag}(d)$, $d \succ 0$, to minimize $\|DMD^{-1}\|_F^2$
- express as

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

- a posynomial in d (with exponents 0, 2, and -2)
- in convex form, with $y = \log d$,

$$\log \|DMD^{-1}\|_F^2 = \log\left(\sum_{i,j=1}^{n} \exp\left(2(y_i - y_j + \log |M_{ij}|)\right)\right)$$

Outline

Optimization problems

Some standard convex problems

Transforming problems

Disciplined convex programming

Geometric programming

Quasiconvex optimization

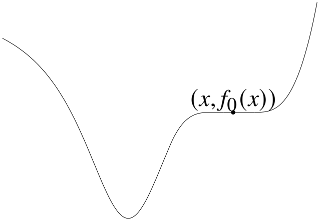
Multicriterion optimization

Quasiconvex optimization

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

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

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



Linear-fractional program

- **linear-fractional program**

$$\begin{array}{ll} \text{minimize} & (c^T x + d)/(e^T x + f) \\ \text{subject to} & Gx \preceq h, \quad Ax = b \end{array}$$

  with variable x and implicit constraint $e^T x + f > 0$

- equivalent to the LP (with variables y, z)

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

- recover $x^\star = y^\star / z^\star$

Von Neumann model of a growing economy

- $x, x^+ \in \mathbf{R}^n_{++}$: activity levels of n economic sectors, in current and next period
- $(Ax)_i$: amount of good i produced in current period
- $(Bx^+)_i$: amount of good i consumed in next period
- $Bx^+ \preceq Ax$: goods consumed next period no more than produced this period
- x_i^+/x_i: growth rate of sector i
- allocate activity to maximize growth rate of slowest growing sector

$$\begin{array}{ll} \text{maximize (over } x, x^+) & \min_{i=1,\ldots,n} x_i^+/x_i \\ \text{subject to} & x^+ \succeq 0, \quad Bx^+ \preceq Ax \end{array}$$

- a quasiconvex problem with variables x, x^+

Convex representation of sublevel sets

- 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.*, $f_0(x) \leq t \iff \phi_t(x) \leq 0$

example:

- $f_0(x) = p(x)/q(x)$, with p convex and nonnegative, q concave and positive
- 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 $\phi_t(x) \leq 0$

Bisection method for quasiconvex optimization

- for fixed t, consider convex feasiblity problem

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

  if feasible, we can conclude that $t \geq p^\star$; if infeasible, $t \leq p^\star$

- bisection method:

```
given $l \leq p^\star$, $u \geq p^\star$, 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 $u - l \leq \epsilon$.
```

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

Outline

Optimization problems

Some standard convex problems

Transforming problems

Disciplined convex programming

Geometric programming

Quasiconvex optimization

Multicriterion optimization

Multicriterion optimization

- **multicriterion** or **multi-objective** problem:

$$\begin{array}{ll} \text{minimize} & f_0(x) = (F_1(x), \ldots, F_q(x)) \\ \text{subject to} & f_i(x) \leq 0, \quad i = 1, \ldots, m, \quad Ax = b \end{array}$$

- objective is the **vector** $f_0(x) \in \mathbf{R}^q$
- q different objectives $F_1, \ldots, F_q$; roughly speaking we want all F_i's to be small
- feasible $x^\star$ is **optimal** if y feasible $\implies f_0(x^\star) \preceq f_0(y)$
- this means that $x^\star$ simultaneously minimizes each F_i; the objectives are **noncompeting**
- not surprisingly, this doesn't happen very often

Pareto optimality

- feasible x **dominates** another feasible $\tilde{x}$ if $f_0(x) \preceq f_0(\tilde{x})$ and for at least one i, $F_i(x) < F_i(\tilde{x})$
- *i.e.*, x meets $\tilde{x}$ on all objectives, and beats it on at least one

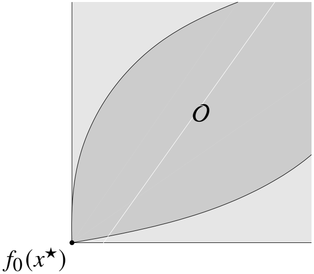
- feasible x^{po} is **Pareto optimal** if it is not dominated by any feasible point
- can be expressed as: y feasible, $f_0(y) \preceq f_0(x^{\mathrm{po}}) \implies f_0(x^{\mathrm{po}}) = f_0(y)$

- there are typically many Pareto optimal points
- for $q = 2$, set of Pareto optimal objective values is the **optimal trade-off curve**
- for $q = 3$, set of Pareto optimal objective values is the **optimal trade-off surface**

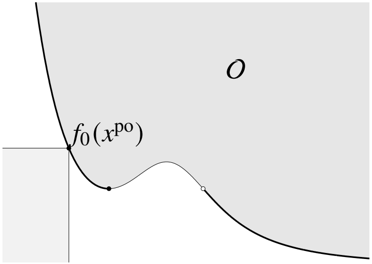
Optimal and Pareto optimal points

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



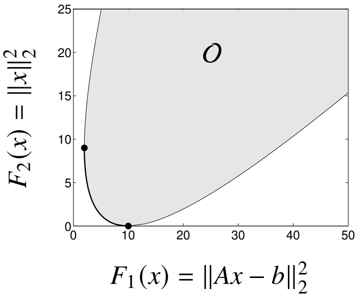
$x^\star$ is optimal



x^{po} is Pareto optimal

Regularized least-squares

- minimize $(\|Ax-b\|_2^2, \|x\|_2^2)$ (first objective is loss; second is regularization)
- example with $A \in \mathbf{R}^{100\times 10}$; heavy line shows Pareto optimal points

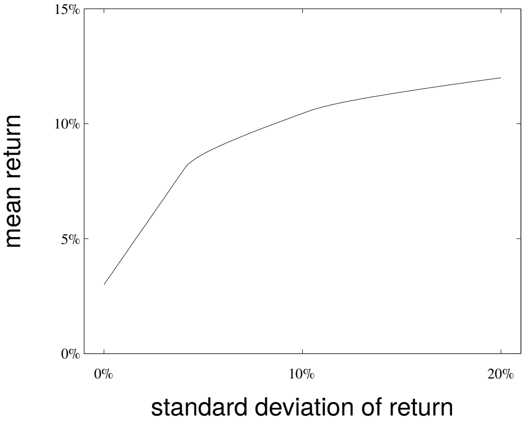


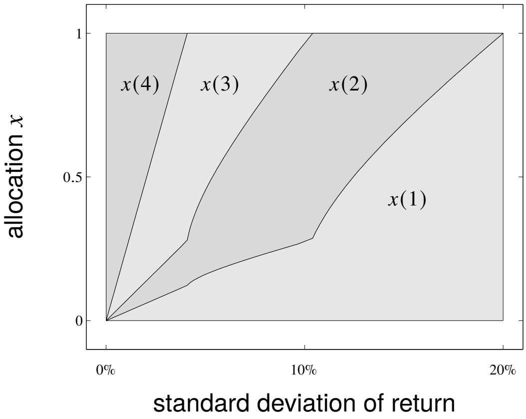
Risk return trade-off in portfolio optimization

- variable $x \in \mathbf{R}^n$ is investment portfolio, with x_i fraction invested in asset i
- $\bar{p} \in \mathbf{R}^n$ is mean, Σ is covariance of asset returns
- portfolio return has mean $\bar{p}^T x$, variance $x^T \Sigma x$

- minimize $(-\bar{p}^T x, x^T \Sigma x)$, subject to $\mathbf{1}^T x = 1$, $x \succeq 0$
- Pareto optimal portfolios trace out optimal risk-return curve

Example


15%
10%
5%
0%
0%
10%
20%
mean return
standard deviation of return


1
0.5
0
0%
10%
20%
$x(4)$
$x(3)$
$x(2)$
$x(1)$
allocation x
standard deviation of return

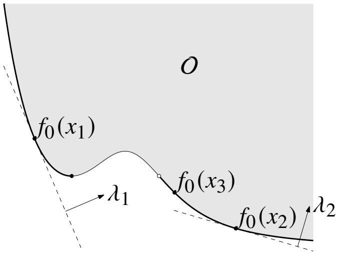
Scalarization

- **scalarization** combines the multiple objectives into one (scalar) objective
- a standard method for finding Pareto optimal points
- choose $\lambda \succ 0$ and solve scalar problem

$$\begin{array}{ll} \text{minimize} & \lambda^T f_0(x) = \lambda_1 F_1(x) + \cdots + \lambda_q F_q(x) \\ \text{subject to} & f_i(x) \leq 0, \quad i = 1, \ldots, m, \quad h_i(x) = 0, \quad i = 1, \ldots, p \end{array}$$

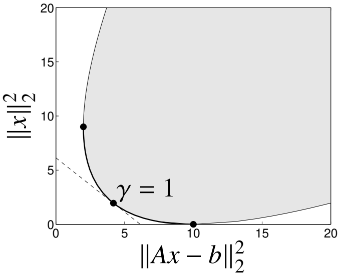
- λ_i are relative weights on the objectives
- if x is optimal for scalar problem, then it is Pareto-optimal for multicriterion problem
- for convex problems, can find (almost) all Pareto optimal points by varying $\lambda \succ 0$

Example



Example: Regularized least-squares

- regularized least-squares problem: minimize $(\|Ax - b\|_2^2, \|x\|_2^2)$
- take $\lambda = (1, \gamma)$ with $\gamma > 0$, and minimize $\|Ax - b\|_2^2 + \gamma\|x\|_2^2$



Example: Risk-return trade-off

- risk-return trade-off: minimize $(-\bar{p}^T x, x^T \Sigma x)$ subject to $\mathbf{1}^T x = 1$, $x \succeq 0$
- with $\lambda = (1, \gamma)$ we obtain scalarized problem

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

- objective is negative **risk-adjusted return**, $\bar{p}^T x - \gamma x^T \Sigma x$
- γ is called the **risk-aversion parameter**