4. Convex optimization problems

- standard form (convex) optimization problem
- linear optimization
- quadratic optimization
- geometric programming
- semidefinite optimization
- quasiconvex optimization
- vector and 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
- $f_i : \mathbf{R}^n \to \mathbf{R}$, for $i = 1, \ldots, m$, are the inequality constraint functions
- $h_i : \mathbf{R}^n \to \mathbf{R}$, for $i = 1, \ldots, p$, are the equality constraint functions

Feasible and optimal points

Feasible point: x is *feasible* if $x \in \mathbf{dom}\, f_0$ and it satisfies all constraints

Optimal value

$$p^\star = \inf\{f_0(x) \mid f_i(x) \le 0,\ i = 1, \ldots, m,\ h_i(x) = 0,\ i = 1, \ldots, p\}$$

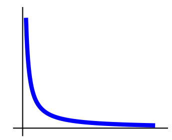
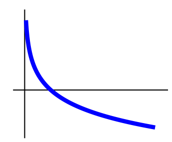
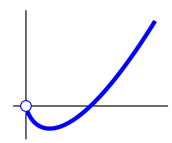
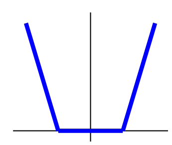
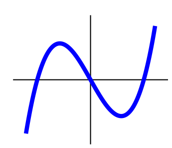
- $p^\star = \infty$ if the problem is infeasible (set of feasible x is empty)
- $p^\star = -\infty$ if the problem is unbounded below

Optimal solution

- a feasible x is *optimal* if $f_0(x) = p^\star$
- the set of optimal points will be denoted by X_{opt}
- $\hat{x}$ is *locally optimal* if there is an $R > 0$ such that $\hat{x}$ is optimal for the problem

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

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

$f_0(x)$	$\mathbf{dom}\, f_0$	graph	$p^\star$	X_{opt}
$1/x$	$\mathbf{R}_{++}$		0	empty
$-\log x$	$\mathbf{R}_{++}$		$-\infty$	empty
$x \log x$	$\mathbf{R}_{++}$		$-1/e$	$\{1/e\}$
$\max\{0, \lvert x\rvert - 1\}$	$\mathbf{R}$		0	$[-1, 1]$
$x^3 - 3x$	$\mathbf{R}$		$-\infty$	empty

Implicit constraints

the standard form optimization problem has an *implicit constraint*

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

- 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 ($m = p = 0$)
- the distinction will be important when we diccuss duality

Example

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

this 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

this formulation is not meant as a practical method for solving feasibility problems

Convex optimization problem in standard form

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

- objective and inequality constraint functions $f_0, f_1, \dots, f_m$ are convex
- equality constraints are linear, often written as $Ax = b$
- feasible set is convex: the intersection of several convex sets

$$
\mathbf{dom}\, f_0, \qquad \text{sublevel sets } \{x \mid f_i(x) \le 0\}, \qquad \text{the affine set } \{x \mid Ax = b\}
$$

- optimal set is convex: any convex combination of optimal x_1, x_2 is feasible, with

$$
\begin{aligned}
f_0(\theta x_1 + (1-\theta)x_2) &\le \theta f_0(x_1) + (1-\theta) f(x_2) \\
&= p^\star
\end{aligned}
$$

hence, $f_0(\theta x_1 + (1-\theta)x_2) = p^\star$, so the convex combination is optimal

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
- feasible set $\{(x_1, x_2) \mid x_1 = -x_2 \leq 0\}$ is convex
- not a convex problem (according to our definition): f_1 not convex, h_1 not affine
- the problem is 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

- suppose x is locally optimal: there is an $R > 0$ such that

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

- suppose x is not globally optimal: there exists a feasible y with $f_0(y) < f_0(x)$
- convex combinations of x and y are feasible
- cost function at convex combination of x and y with $0 < \theta \leq 1$ satisfies

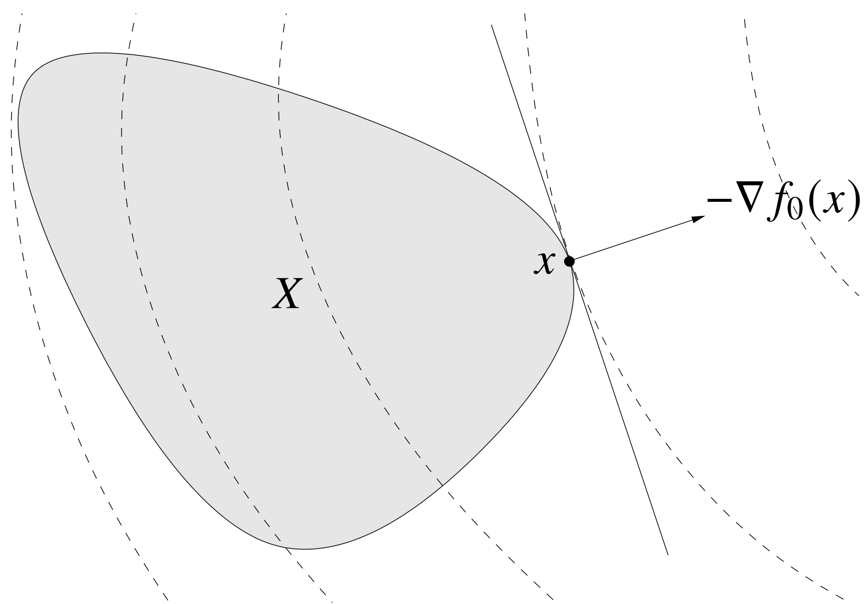
$$\begin{aligned} f_0((1-\theta)x + \theta y) &\leq (1-\theta) f_0(x) + \theta f_0(y) \\ &< (1-\theta) f_0(x) + \theta f_0(x) \\ &= f_0(x) \end{aligned}$$

- for $0 < \theta \leq R/\|y - x\|_2$ this contradicts the assumption of local optimality of x

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

Proof (necessity)

- consider feasible $y \neq x$ and define line segment $I = \{x + t(y - x) \mid 0 \leq t \leq 1\}$
- by convexity of X, points in I are feasible
- let $g(t) = f_0(x + t(y - x))$ be the restriction of f_0 to I
- derivative at t is $g'(t) = \nabla f_0(x + t(y - x))^T(y - x)$, so

$$g'(0) = \nabla f_0(x)^T(y - x)$$

- if $g'(0) = \nabla f_0(x)^T(x - y) < 0$, the point x is not even locally optimal

Proof (sufficiency)

if y is feasible and $\nabla f_0(x)^T(y - x) \geq 0$, then, by convexity of f_0,

$$\begin{aligned} f_0(y) &\geq f_0(x) + \nabla f_0(x)^T(y - x) \\ &\geq f_0(x) \end{aligned}$$

Examples

Unconstrained problem: x is optimal if and only if

$$x \in \mathrm{dom}\, f_0, \qquad \nabla f_0(x) = 0$$

(recall our assumption that $\mathrm{dom}\, f_0$ is an open set if f_0 is differentiable)

Minimization over nonnegative orthant

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

x is optimal if and only if

$$x \in \mathrm{dom}\, f_0, \qquad x \succeq 0, \qquad \begin{cases} \nabla f_0(x)_i \geq 0 & x_i = 0 \\ \nabla f_0(x)_i = 0 & x_i > 0 \end{cases}$$

Equality constrained problem

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

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

$$x \in \mathbf{dom}\, f_0, \qquad Ax = b, \qquad \nabla f_0(x) + A^T\nu = 0$$

- first two conditions are feasibility of x
- gradient $\nabla f_0(x)$ can be decomposed as $\nabla f_0(x) + A^T\nu = w$ with $Aw = 0$
- if $w = 0$, the optimality condition on page 4.10 holds:

$$\nabla f_0(x)^T(y - x) = -\nu^T A(y - x) = 0 \quad \text{for all } y \text{ with } Ay = b$$

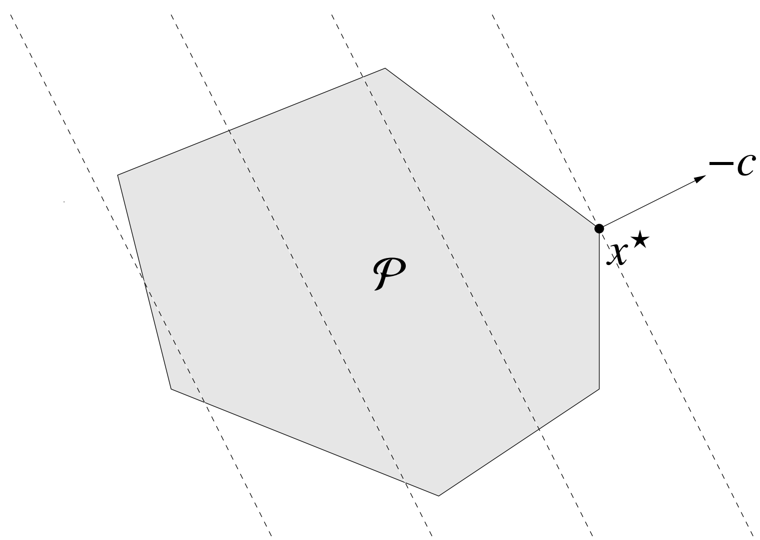
- if $w \neq 0$, condition on p. 4.10 does not hold: $y = x - tw$ is feasible for small $t > 0$,

$$\nabla f_0(x)^T(y - x) = -t(w - A^T\nu)^T w = -t\|w\|_2^2 < 0$$

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



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,

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

Piecewise-linear minimization

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

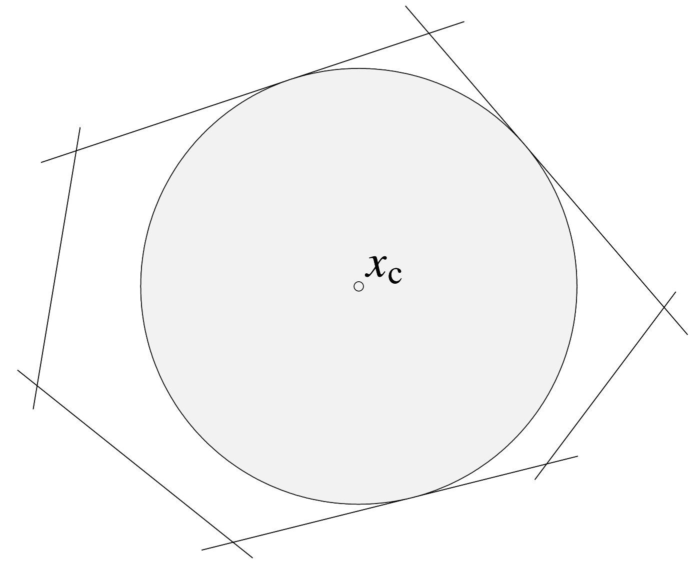
equivalent to an LP

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

Chebyshev center of a polyhedron

Chebyshev center of

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



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

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

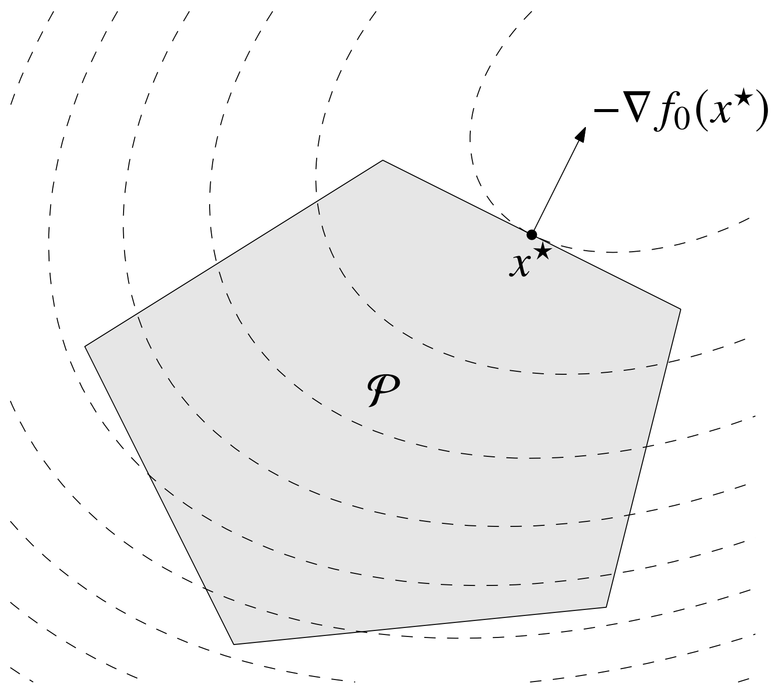
- hence, x_c, r can be determined by solving the LP

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

Quadratic program (QP)

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

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



Examples

Least squares

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

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

Linear program with random cost

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

- 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
- γ controls trade-off between expected cost and variance (risk)

Quadratically constrained quadratic program (QCQP)

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

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

Second-order cone programming

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

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

- 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

Robust linear programming

the parameters in optimization problems are often uncertain, *e.g.*, in an LP

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

there can be uncertainty in c, a_i, b_i

two common approaches to handling uncertainty (in a_i, for simplicity)

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

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

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

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

Deterministic approach via SOCP

choose an ellipsoid as $\mathcal{E}_i$:

$$\mathcal{E}_i = \{\bar{a}_i + P_i u \mid \|u\|_2 \leq 1\} \qquad (\bar{a}_i \in \mathbf{R}^n,\ P_i \in \mathbf{R}^{n\times n})$$

center is $\bar{a}_i$, semi-axes determined by singular values/vectors of P_i

SOCP formulation

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

this is equivalent to the 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 via SOCP

- assume $a_i \sim \mathcal{N}(\bar{a}_i, \Sigma_i))$: Gaussian with mean $\bar{a}_i$, covariance Σ_i
- $a_i^T x$ is Gaussian random variable with mean $\bar{a}_i^T x$, variance $x^T \Sigma_i x$
- if we denote the CDF of $\mathcal{N}(0,1)$ by $\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-t^2/2}\, dt$,

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

SOCP formulation of robust LP

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

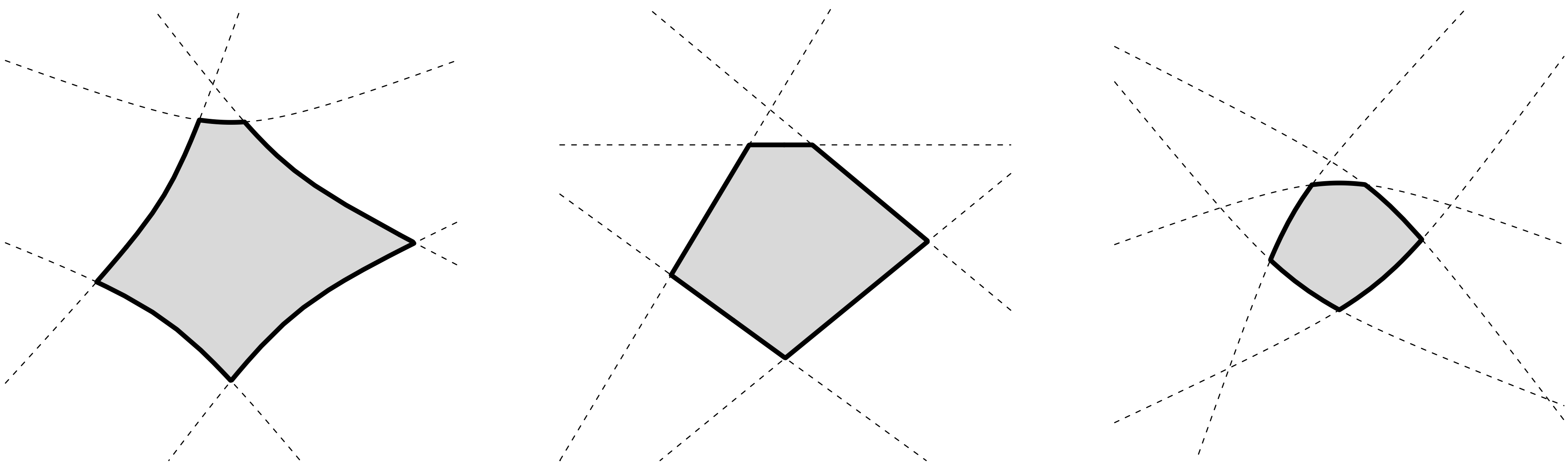
for $\eta \ge 1/2$, this is equivalent to the SOCP

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

Example

$$\mathbf{prob}(a_i^T x \leq b_i) \geq \eta, \quad i = 1, \ldots, 5$$

feasible set for three values of η



$\eta = 10\%$	$\eta = 50\%$	$\eta = 90\%$
$\Phi^{-1}(\eta) < 0$	$\Phi^{-1}(\eta) = 0$	$\Phi^{-1}(\eta) > 0$

Geometric programming

Monomial function

$$f(x) = c x_1^{a_1} x_2^{a_2} \cdots x_n^{a_n}, \qquad \mathbf{dom}\, f = \mathbf{R}^n_{++}$$

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

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

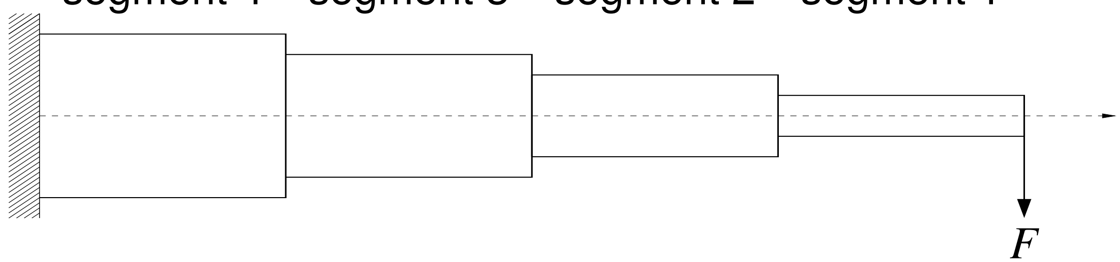
- posynomial $f(x) = \sum_{k=1}^{K} c_k x_1^{a_{1k}} x_2^{a_{2k}} \cdots x_n^{a_{nk}}$ transforms to

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

- geometric program transforms to convex problem

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

Design of cantilever beam



- N segments with unit lengths, rectangular cross-sections of size $w_i \times h_i$
- given vertical force F applied at the right end

Design problem

minimize	total weight
subject to	upper & lower bounds on w_i, h_i
	upper bound & lower bounds on aspect ratios h_i/w_i
	upper bound on stress in each segment
	upper bound on vertical deflection at the end of the beam

variables: w_i, h_i for $i = 1, \ldots, N$

Objective and constraint functions

- total weight $w_1h_1 + \cdots + w_Nh_N$ is posynomial
- aspect ratio h_i/w_i and inverse aspect ratio w_i/h_i are monomials
- maximum stress in segment i is given by $6iF/(w_ih_i^2)$, a monomial
- vertical deflection y_i and slope v_i of central axis at the right end of segment i:

$$\begin{aligned} v_i &= 12(i-1/2)\frac{F}{Ew_ih_i^3} + v_{i+1} \\ y_i &= 6(i-1/3)\frac{F}{Ew_ih_i^3} + v_{i+1} + y_{i+1} \end{aligned}$$

for $i = N, N-1, \ldots, 1$, with $v_{N+1} = y_{N+1} = 0$ (E is Young's modulus)

v_i and y_i are posynomial functions of w, h

Formulation as a GP

$$\begin{array}{ll}
\text{minimize} & w_1 h_1 + \cdots + w_N h_N \\
\text{subject to} & w_{\max}^{-1} w_i \leq 1, \quad w_{\min} w_i^{-1} \leq 1, \quad i = 1, \ldots, N \\
& h_{\max}^{-1} h_i \leq 1, \quad h_{\min} h_i^{-1} \leq 1, \quad i = 1, \ldots, N \\
& S_{\max}^{-1} w_i^{-1} h_i \leq 1, \quad S_{\min} w_i h_i^{-1} \leq 1, \quad i = 1, \ldots, N \\
& 6iF\sigma_{\max}^{-1} w_i^{-1} h_i^{-2} \leq 1, \quad i = 1, \ldots, N \\
& y_{\max}^{-1} y_1 \leq 1
\end{array}$$

note

- we write $w_{\min} \leq w_i \leq w_{\max}$ and $h_{\min} \leq h_i \leq h_{\max}$

$$w_{\min}/w_i \leq 1, \qquad w_i/w_{\max} \leq 1, \qquad h_{\min}/h_i \leq 1, \qquad h_i/h_{\max} \leq 1$$

- we write $S_{\min} \leq h_i/w_i \leq S_{\max}$ as

$$S_{\min} w_i/h_i \leq 1, \qquad h_i/(w_i S_{\max}) \leq 1$$

Minimizing spectral radius of nonnegative matrix

Perron–Frobenius eigenvalue $\lambda_{\mathrm{pf}}(A)$

- exists for (elementwise) positive $A \in \mathbf{R}^{n \times n}$
- a real, positive eigenvalue of A, equal to spectral radius $\max_i |\lambda_i(A)|$
- determines asymptotic growth (decay) rate of A^k: $A^k \sim \lambda_{\mathrm{pf}}^k$ as $k \to \infty$
- alternative characterization: $\lambda_{\mathrm{pf}}(A) = \inf\{\lambda \mid Av \preceq \lambda v \text{ for some } v \succ 0\}$

Minimizing spectral radius of matrix of posynomials

- minimize $\lambda_{\mathrm{pf}}(A(x))$, where the elements $A(x)_{ij}$ are posynomials of x
- equivalent geometric program:

$$\begin{array}{ll} \text{minimize} & \lambda \\ \text{subject to} & \sum_{j=1}^{n} A(x)_{ij} v_j / (\lambda v_i) \leq 1, \quad i = 1, \ldots, n \end{array}$$

variables λ, v, x

Conic linear optimization

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

- K is a proper convex cone in $\mathbf{R}^m$
- F is an $m \times n$ matrix, g is a m-vector
- constraint means $-(Fx + g) \in K$
- linear programming is special case with $K = \mathbf{R}^m_+$
- same properties as standard convex problem (local optimum is global, etc.)

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

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

LP and SOCP as SDP

LP and equivalent SDP

LP: minimize $c^T x$
subject to $Ax \preceq b$

SDP: minimize $c^T x$
subject to $\mathbf{diag}(Ax - b) \preceq 0$

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

SOCP and equivalent SDP

SOCP: minimize $f^T x$
subject to $\|A_i x + b_i\|_2 \leq c_i^T x + d_i, \quad i = 1, \ldots, m$

SDP: minimize $f^T x$
subject to $\begin{bmatrix} (c_i^T x + d_i)I & A_i x + b_i \\ (A_i x + b_i)^T & c_i^T x + d_i \end{bmatrix} \succeq 0, \quad i = 1, \ldots, m$

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 $A_i \in \mathbf{S}^k$)

Equivalent SDP

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

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

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

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

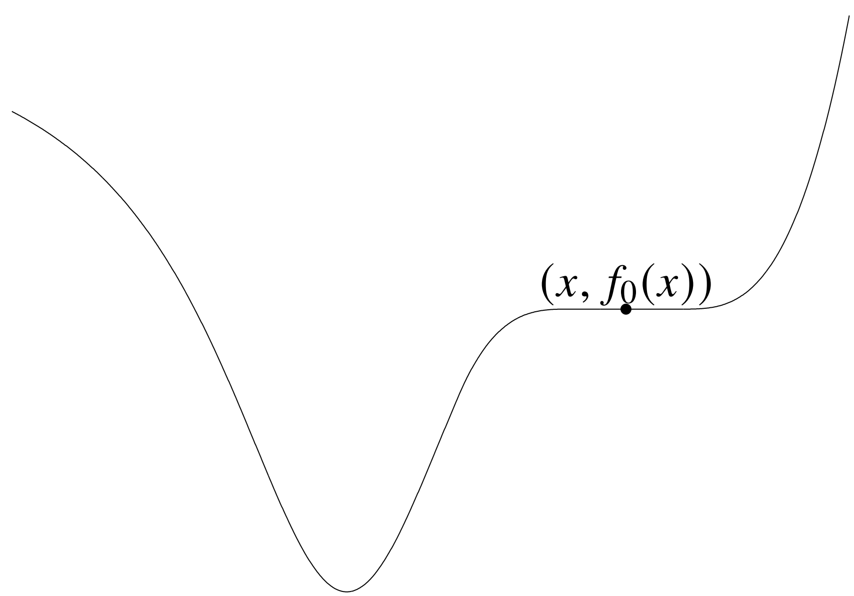
$$\begin{array}{rcl} \|A\|_2 \leq t & \Longleftrightarrow & A^T A \preceq t^2 I, \quad t \geq 0 \\ & \Longleftrightarrow & \left[\begin{array}{cc} tI & A \\ A^T & tI \end{array}\right] \succeq 0 \end{array}$$

Quasiconvex optimization

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

- f_0 is quasiconvex
- $f_1, \ldots, f_m$ are convex

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



Linear-fractional program

$$\begin{array}{ll} \text{minimize} & f_0(x) \\ \text{subject to} & Gx \preceq h \\ & Ax = b \end{array} \qquad (1)$$

where

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

- a quasiconvex optimization problem
- also equivalent to the LP (variables y, z)

$$\begin{array}{ll} \text{minimize} & c^T y + dz \\ \text{subject to} & Gy \preceq hz \\ & Ay = bz \\ & e^T y + fz = 1 \\ & z \geq 0 \end{array} \qquad (2)$$

Exercise

assume the linear-fractional program (1) is feasible

- show how to obtain the solution of (1) from the solution of the LP (2)
- what do solutions (y, z) of (2) with $z = 0$ mean for (1)?

Solution: denote the optimal values of (1) and (2) by $p^\star_{\mathrm{lfp}}$ and $p^\star_{\mathrm{lp}}$, respectively

1. for every feasible x in (1), there is a corresponding feasible (y, z) in (2):

$$y = \frac{x}{e^T x + f}, \qquad z = \frac{1}{e^T x + f}, \qquad c^T y + dz = \frac{c^T x + d}{e^T x + f}$$

2. for every feasible (y, z) in (2) with $z > 0$, there is a feasible x in (1):

$$x = \frac{y}{z}, \qquad f_0(x) = \frac{c^T x + d}{e^T x + f} = \frac{c^T y + dz}{e^T y + fz} = c^T y + d$$

3. suppose (y, z) is feasible for (2) with $z = 0$:

$$Gy \preceq 0, \qquad Ay = 0, \qquad e^T y = 1$$

let $\hat{x}$ be a feasible point for (1):

$$G\hat{x} \preceq h, \qquad A\hat{x} = b, \qquad e^T \hat{x} + f > 0$$

all points on the half-line $\{\hat{x} + \alpha y \mid \alpha \geq 0\}$ are feasible for (1),

$$G(\hat{x} + \alpha y) \preceq h, \qquad A(\hat{x} + \alpha y) = b, \qquad e^T(\hat{x} + \alpha y + f) > 0,$$

and the cost function at $\hat{x} + \alpha y$ tends to $c^T y$ as $\alpha \to \infty$:

$$f_0(\hat{x} + \alpha y) = \frac{c^T \hat{x} + d + \alpha c^T y}{e^T \hat{x} + f + \alpha} \longrightarrow c^T y$$

- 1 shows that $p^\star_{\mathrm{lfp}} \geq p^\star_{\mathrm{lp}}$ and 2, 3 show that $p^\star_{\mathrm{lp}} \geq p^\star_{\mathrm{lfp}}$; therefore $p^\star_{\mathrm{lp}} = p^\star_{\mathrm{lfp}}$
- if (y, z) is an optimal solution of (2) and $z > 0$, then $x = y/z$ is optimal for (1)
- optimal $(y, 0)$ of (2) indicates the optimal value of (1) is finite but not attained

Generalized linear-fractional program

$$f_0(x) = \max_{i=1,\dots,r} \frac{c_i^T x + d_i}{e_i^T x + f_i}, \qquad \mathrm{dom}\, f_0(x) = \{x \mid e_i^T x + f_i > 0,\ i = 1, \dots, r\}$$

- a quasiconvex optimization problem
- LP reformulation of page 4.37 does not extend to generalized problem

Example: Von Neumann model of a growing economy

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

- $x, x^+ \in \mathbf{R}^n$: activity levels of n sectors, in current and next period
- $(Ax)_i$: amount of good i produced in current period
- $(Bx^+)_i$: amount consumed in next period, cannot exceed $(Ax)_i$
- x_i^+/x_i: growth rate of sector i

allocate activity to maximize growth rate of slowest growing sector

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.*,

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

Example

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

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

can take $\phi_t(x) = p(x) - tq(x)$:

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

Quasiconvex optimization via convex feasibility problems

$$\phi_t(x) \le 0, \qquad f_i(x) \le 0, \quad i = 1, \dots, m, \qquad Ax = b \tag{3}$$

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

Bisection method

given: $l \le p^\star$, $u \ge p^\star$, tolerance $\epsilon > 0$
repeat
    1. $t := (l + u)/2$
    2. solve the convex feasibility problem (3)
    3. if (3) is feasible, $u := t$
        else $l := t$
until $u - l \le \epsilon$

requires exactly $\left\lceil \log_2 \left(\frac{u-l}{\epsilon} \right) \right\rceil$ iterations

Vector optimization

General vector optimization problem

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

vector objective $f_0 : \mathbf{R}^n \to \mathbf{R}^q$, minimized with respect to proper cone $K \in \mathbf{R}^q$

Convex vector optimization problem

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

where $f_1, \ldots, f_m$ are convex and f_0 is "K-convex", *i.e.*,

$$f_0(\theta x + (1 - \theta)y) \preceq_K \theta f_0(x) + (1 - \theta) f_0(y)$$

for all $x, y \in \mathbf{dom}\, f_0$ and $\theta \in [0, 1]$

Multicriterion optimization

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

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

- q different objectives F_i; roughly speaking we want all F_i's to be small
- feasible $x^\star$ is *optimal* if

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

  if there exists an optimal point, the objectives are noncompeting
- feasible $x^{\rm po}$ is *Pareto optimal* if

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

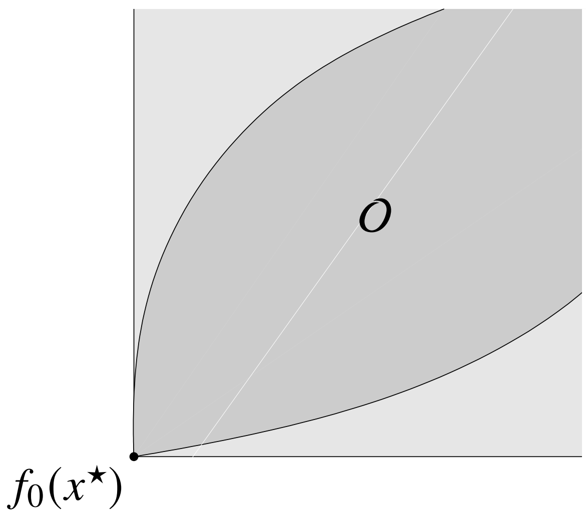
  if Pareto optimal values are not unique, there is a trade-off between objectives
- f_0 is K-convex if $F_1, \ldots, F_q$ are convex (in the usual sense)

Optimal and Pareto optimal points

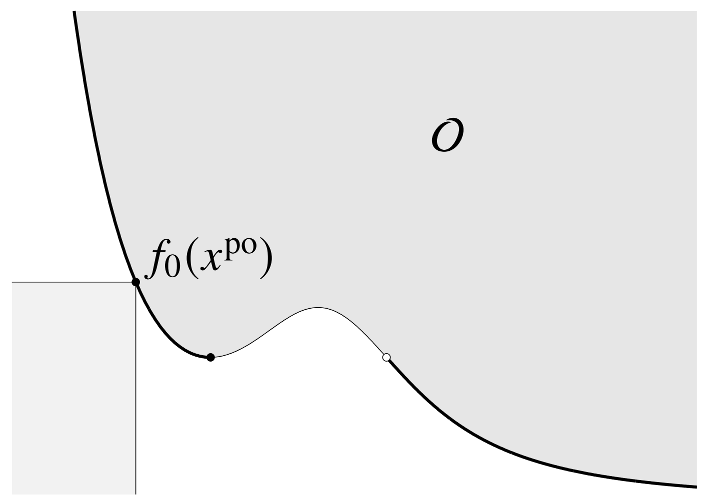
set of achievable objective values

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

- 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 $\mathcal{O}$



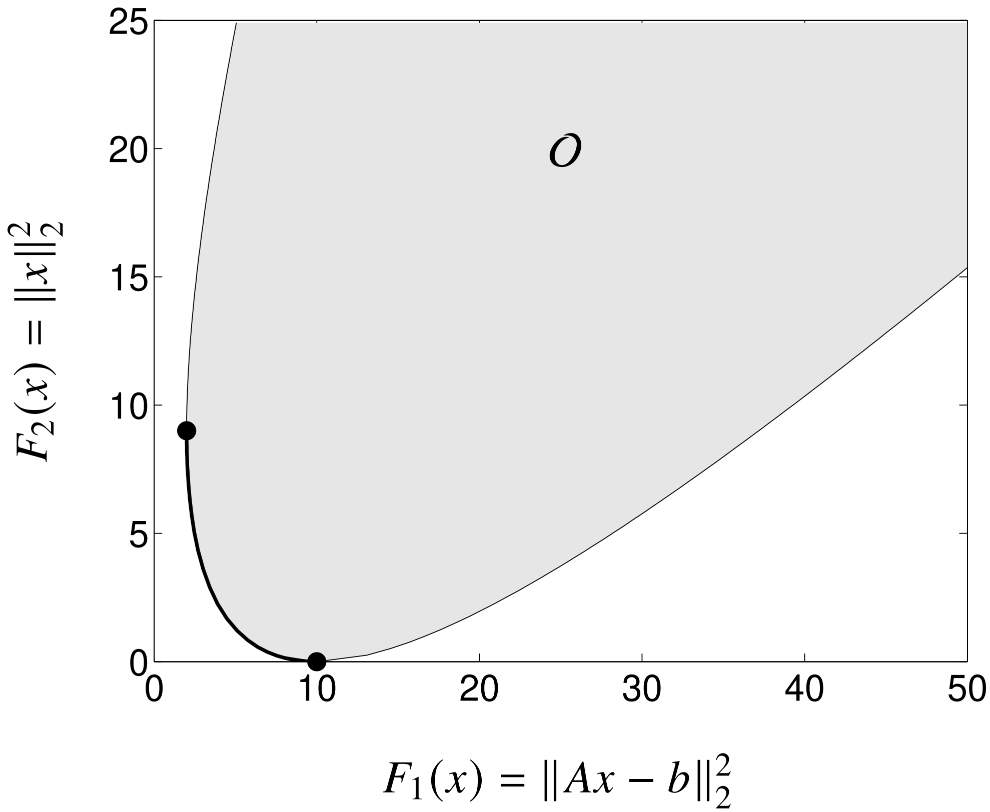
$x^\star$ is optimal



x^{po} is Pareto optimal

Regularized least-squares

$$\text{minimize (w.r.t. } \mathbf{R}_+^2) \quad (\|Ax - b\|_2^2, \|x\|_2^2)$$



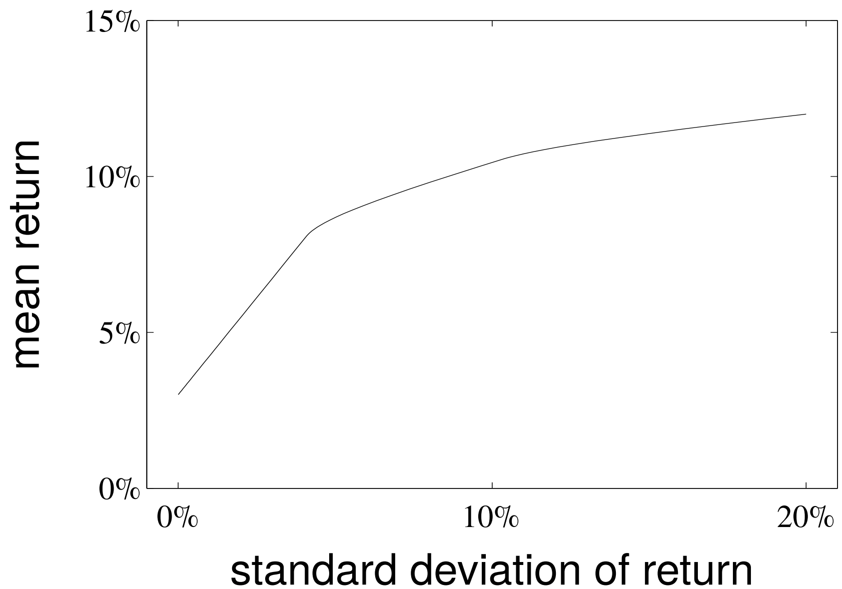
example for $A \in \mathbf{R}^{100 \times 10}$; heavy line is formed by Pareto optimal points

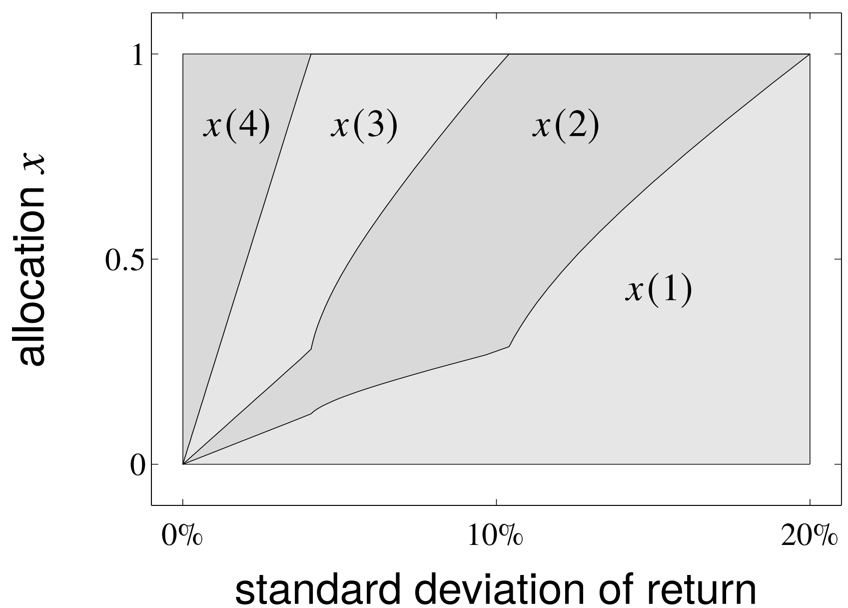
Risk–return trade-off in portfolio optimization

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

- $x \in \mathbf{R}^n$ is investment portfolio; x_i is fraction invested in asset i
- return is $r = p^T x$ where $p \in \mathbf{R}^n$ is vector of relative asset price changes
- p is modeled as a random variable with mean $\bar{p}$, covariance Σ
- $\bar{p}^T x = \mathbf{E}\, r$ is expected return; $x^T \Sigma x = \mathbf{var}\, r$ is return variance (risk)

Example





Scalarization

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

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

- solutions x of scalar problem are Pareto-optimal for vector optimization problem

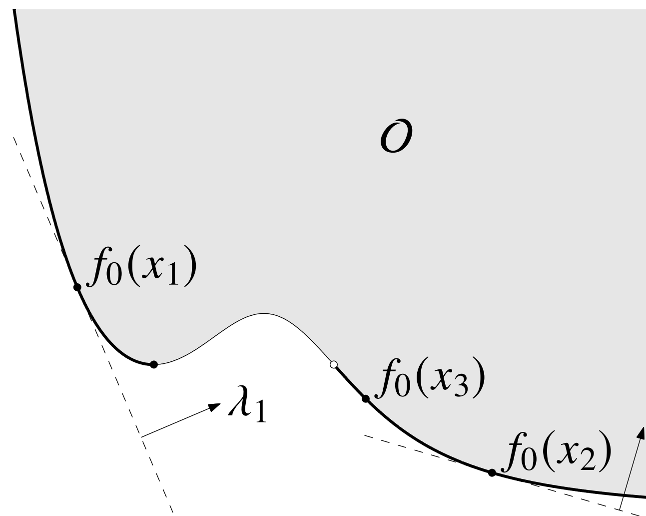
x not Pareto-optimal

$$\Downarrow$$

$\exists$ feasible $y : f_0(y) \preceq_K f_0(x),\ f_0(y) \neq f_0(x)$

$$\Downarrow$$

$\lambda^T f_0(y) < \lambda^T f_0(x)$ for $\lambda \succ_{K_*} 0$



- partial converse for convex vector optimization problems (see later in duality): can find (almost) all Pareto optimal points by varying $\lambda \succ_{K^*} 0$
- objective of scalar problem is convex if f_0 is K-convex

Scalarization for multicriterion problems

to find Pareto optimal points, minimize positive weighted sum

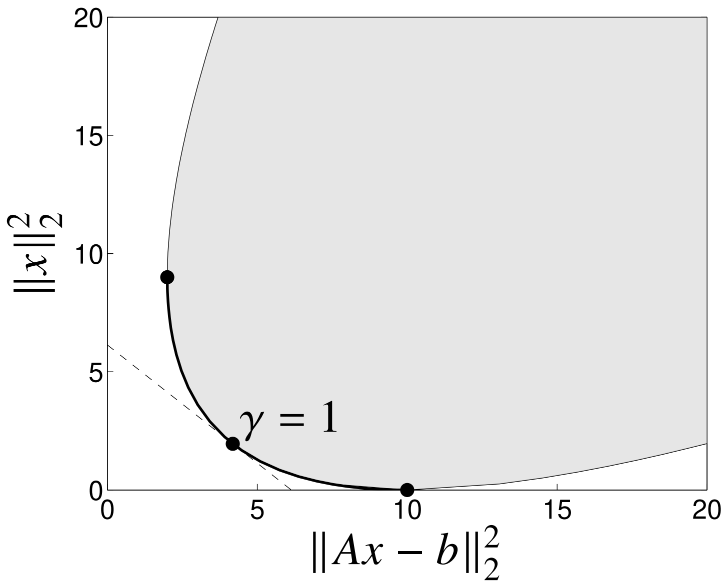
$$\lambda^T f_0(x) = \lambda_1 F_1(x) + \cdots + \lambda_q F_q(x)$$

- regularized least squares problem of page 4.46

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

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

  for fixed γ, a LS problem



- risk–return trade-off of page 4.47: with $\gamma > 0$,

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