Announcement

- Postponed HW 6! HW5 due today midnight.
- See Canvas announcements re midterm (Fri Feb 16, in-class)
- Solving B&V book problems is good practice.
- TA review session this Fri (recorded, but with live Zoom office hours afternoon; please see Natalia's earlier announcement on Canvas.

6. Convex optimization problems: GP, SDP, and multi-objective optimization

- geometric programming (GP)
- generalized inequality constraints
- semidefinite programming
- vector (multi-objective) optimization

Geometric programming

application:
- mechanical design/ structures
- circuit design (amplifiers, PLLs)
- chemical reactions
- power allocation in wireless networks

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)

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

with f_i posynomial, h_i monomial

Geometric program in convex form

change variables to $\underline{y_i = \log x_i}$, and take logarithm of cost, constraints

- monomial $\underline{f(x) = c x_1^{a_1} \cdots x_n^{a_n}}$ transforms to $\underline{\log c} + \sum_{i=1}^{n} a_i \overbrace{\log x_i}^{y_i}$

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

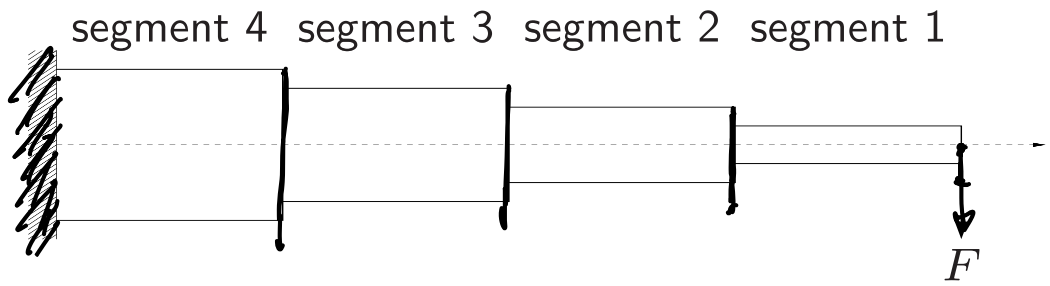
- posynomial $f(x) = \sum_{k=1}^{K} c_k x_1^{a_{1k}} x_2^{a_{2k}} \cdots x_n^{a_{nk}}$ transforms to

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

- geometric program transforms to convex problem

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

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 $w_1 h_1 + \cdots + w_N h_N$ is posynomial

$\alpha h_i w_i^{-1} \leq 1$

$\beta h_i^{-1} w_i \leq 1$

- aspect ratio h_i/w_i and inverse aspect ratio w_i/h_i are monomials

$w_i^{-1} h_i^{-2} \leq 1 \ \forall i$

- maximum stress in segment i is given by $6iF/(w_i h_i^2)$, a monomial
- the vertical deflection y_i and slope v_i of central axis at the right end of segment i are defined recursively as

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

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

$$\begin{array}{ll}\text{minimize}_{w,h} & w_1h_1 + \cdots + w_Nh_N \\ \text{subject to} & w_{\max}^{-1}w_i \leq 1, \quad w_{\min}w_i^{-1} \leq 1, \quad i = 1, \ldots, N \\ & h_{\max}^{-1}h_i \leq 1, \quad h_{\min}h_i^{-1} \leq 1, \quad i = 1, \ldots, N \\ & S_{\max}^{-1}w_i^{-1}h_i \leq 1, \quad S_{\min}w_ih_i^{-1} \leq 1, \quad i = 1, \ldots, N \\ & 6iF\sigma_{\max}^{-1}w_i^{-1}h_i^{-2} \leq 1, \quad i = 1, \ldots, N \\ & y_{\max}^{-1}y_1 \leq 1\end{array}$$

note

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

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

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

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

Minimizing spectral radius of nonnegative matrix (skip)

Perron-Frobenius eigenvalue $\lambda_{\rm pf}(A)$

- 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: $A^k \sim \lambda_{\rm pf}^k$ as $k \to \infty$
- alternative characterization: $\lambda_{\rm pf}(A) = \inf\{\lambda \mid Av \preceq \lambda v \text{ for some } v \succ 0\}$

minimizing spectral radius of matrix of posynomials

- minimize $\lambda_{\rm pf}(A(x))$, 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

$x \in K \qquad x - y \in K$

$x \succeq_K 0 \qquad x \succeq_K y$

Generalized inequality constraints

convex problem with generalized inequality constraints

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

$f_1(x) \preceq_{K_1} 0$
$\vdots$
$f_m(x) \preceq_{K_m} 0$

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

conic form problem: special case with affine objective and constraints

conic-linear

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

$\begin{cases} K = \mathbb{R}^m_+ \\ K = S^m_+ \end{cases}$

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

Semidefinite program (SDP)

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

negative semidefinite

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

$$\begin{bmatrix} A & 0 \\ 0 & B \end{bmatrix} \succeq 0 \iff A \succeq 0 \;\&\; B \succeq 0$$

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

$$A(x) = A_0 + \underbrace{x_1} A_1 + \underbrace{x_2} A_2 + \cdots + \underbrace{x_n} A_n \qquad x \in \mathbb{R}^n$$

$$A(x) : \mathbb{R}^n \to S^n$$

$$A(x) \preceq_{S^n_+} 0$$

$$\begin{bmatrix} 1-x_1 & x_2 & 0 \\ x_2 & 5 & -x_3+1 \\ 0 & -x_3+1 & 0 \end{bmatrix} \preceq 0 \iff \overbrace{\begin{bmatrix} 1 & 0 & 0 \\ 0 & 5 & 1 \\ 0 & 1 & 0 \end{bmatrix}}^{A_0} + \overbrace{\begin{bmatrix} -1 & 0 & \\ & & \\ & & \end{bmatrix}}^{A_1} x_1 + \overbrace{\begin{bmatrix} & & \\ & & \\ & & \end{bmatrix}}^{A_2} x_2 + \begin{bmatrix} & & \\ & & \\ & & \end{bmatrix} x_3$$

$$x = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$$

LP and SOCP as SDP

LP and equivalent SDP

LP: $\begin{array}{ll} \text{minimize} & c^T x \\ \text{subject to} & Ax \preceq_{\mathbb{R}^n_+} b \end{array}$ SDP: $\begin{array}{ll} \text{minimize} & c^T x \\ \text{subject to} & \mathbf{diag}(Ax - b) \preceq_{S^n_+} 0 \end{array}$

$$\begin{bmatrix} a_1^T x - b_1 & & 0 \\ & \ddots & \\ 0 & & a_m^T x - b_m \end{bmatrix} \preceq 0$$

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

SOCP and equivalent SDP

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

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

$\rightarrow \|A_i x + b_i\|_2^2 \le (c_i^T x + d_i)^2$

assume: $c_i^T x + d_i > 0, \forall i$

$$\rightarrow \frac{(A_i x + b_i)^T (A_i x + b_i)}{(c_i^T x + d_i)^2} \le \frac{(c_i^T x + d_i)^2}{(c_i^T x + d_i)^2}$$

Schur complements lemma:

$$\begin{bmatrix} A & B \\ B^T & C \end{bmatrix} \succeq 0, \ A \succ 0, \ C \succeq 0$$

$$\Leftrightarrow \ \underbrace{C - B^T A^{-1} B}\succeq 0$$

$$-(A_i x + b_i)^T \left(\frac{1}{(c_i^T x + d_i)^2} \right) (A_i x + b_i) + 1 \succeq 0$$

$$\Leftrightarrow$$

$$\begin{array}{cc} & \begin{array}{cc} & \overset{n}{\longleftrightarrow} \end{array} \\ \begin{array}{c} 1 \updownarrow \\ n \updownarrow \end{array} & \begin{bmatrix} (c_i^T x + d_i)^2 & (A_i x + b_i)^T \\ (A_i x + b_i) & I_{n \times n} \end{bmatrix} \end{array} \succeq 0 \rightarrow \text{not an LMI}$$

$$-(A_i x + b_i)^T \frac{1}{c_i^T x + d_i} (A_i x + b_i) + (c_i^T x + d_i) \succeq 0$$

$$\begin{array}{cc} \begin{array}{c} 1 \updownarrow \\ n \updownarrow \end{array} & \begin{bmatrix} (c_i^T x + d_i) & (A_i x + b_i)^T \\ (A_i x + b_i) & (c_i^T x + d_i) I_{n \times n} \end{bmatrix} \end{array} \succeq 0 \Rightarrow \text{LMI} \checkmark$$

Eigenvalue minimization

$$\begin{array}{ll} \min\limits_{x,t} & t \\ & \lambda_{max}(A(x)) \leq t \end{array}$$

$$\begin{array}{ll} \underset{x}{\text{minimize}} & \lambda_{\max}(A(x)) \end{array}$$

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

equivalent SDP

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

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

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

$A - tI \leq 0 \iff$

$\lambda_i(A - tI) \leq 0, \ \forall i \iff$

$\lambda_i(A) - t \leq 0, \ \forall i \iff$

$\lambda_{max}(A) \leq t$

Matrix norm minimization

$\sigma_{max}(A(x))$

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

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

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

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

$\lambda_{max}(A^T A) \le t^2$

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

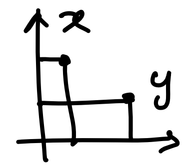
- LP — QP — QCQP — SOCP — SDP
- GP

(chap 2)

Minimum and minimal elements of a set

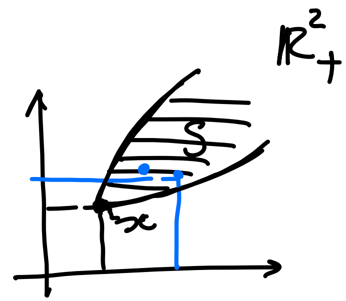
$\preceq_K$ is not in general a *linear ordering*: we can have $x \not\preceq_K y$ and $y \not\preceq_K x$

total ordering



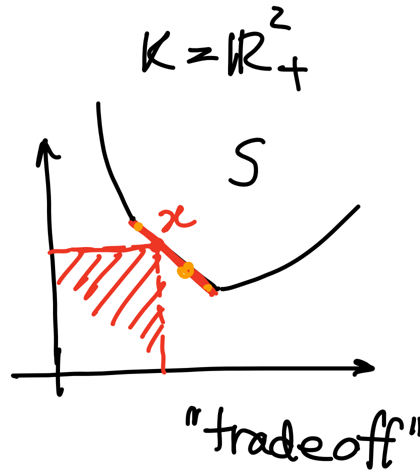
- $x \in S$ is **the minimum element** of S with respect to $\preceq_K$ if

$$y \in S \implies x \preceq_K y$$



- $x \in S$ is **a minimal element** of S with respect to $\preceq_K$ if

$$y \in S, \quad y \preceq_K x \implies y = x$$



Multiobjective (vector) optimization

• $K = \mathbb{R}^n_+$ in all our multi-obj problems in this course!

general vector optimization problem

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

vector objective $f_0 : \mathbf{R}^n \to \mathbf{R}^q$, minimized w.r.t. proper cone $K \in \mathbf{R}^q$

$$x \longrightarrow \underbrace{\begin{bmatrix} f_{01}(x) \\ \vdots \\ f_{0q}(x) \end{bmatrix}} \in \mathbb{R}^q$$

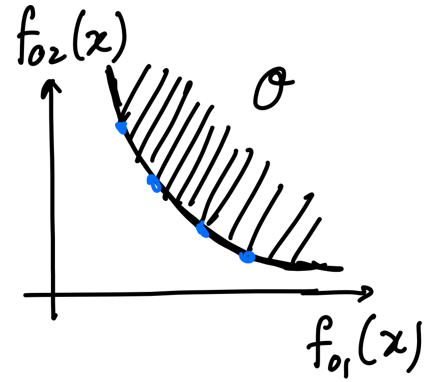
convex vector optimization problem

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

with f_0 K-convex, $f_1, \ldots, f_m$ convex

Optimal and Pareto optimal points

$$f_0(x) = \begin{bmatrix} f_{01}(x) \\ f_{02}(x) \end{bmatrix}$$

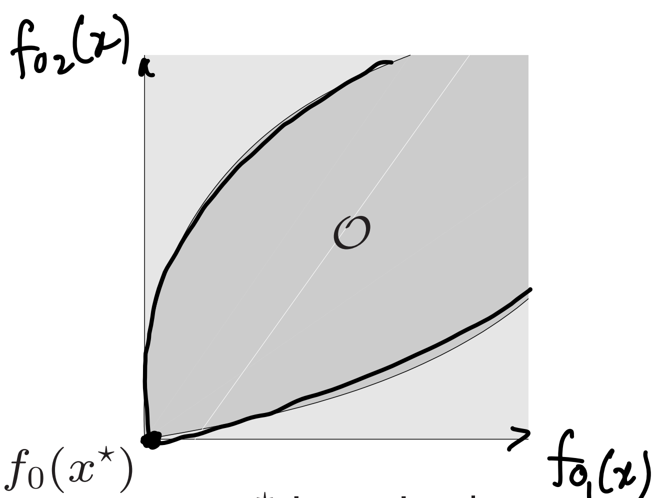


set of achievable objective values

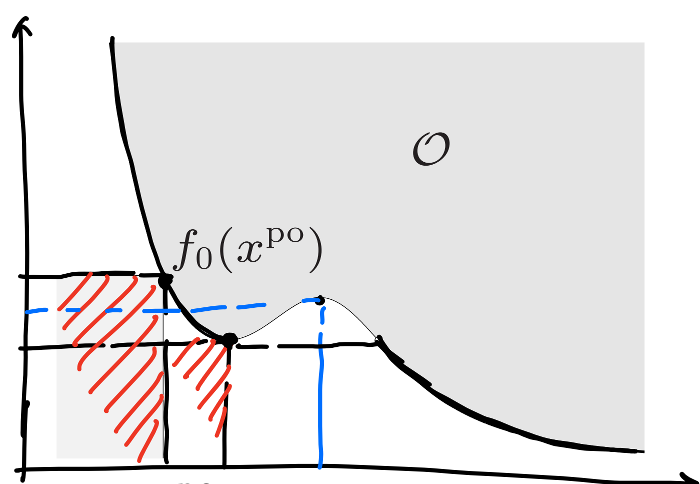
$$\mathcal{O} = \{f_0(x) \mid x \text{ feasible}\}$$

- $\mathcal{O}$ is a convex set for a cvx opt prob.

- feasible x is **optimal** if $f_0(x)$ is a 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



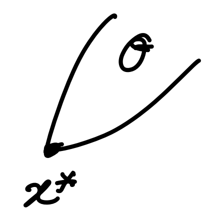
x^{po} is Pareto optimal

Multiobjective optimization

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

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

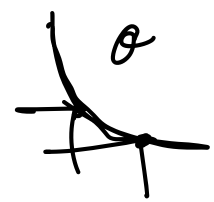
- q different objectives F_i; roughly speaking we want all F_i's to be small
- feasible $x^\star$ is optimal if



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

if there exists an optimal point, the objectives are noncompeting

- feasible x^{po} is Pareto optimal if

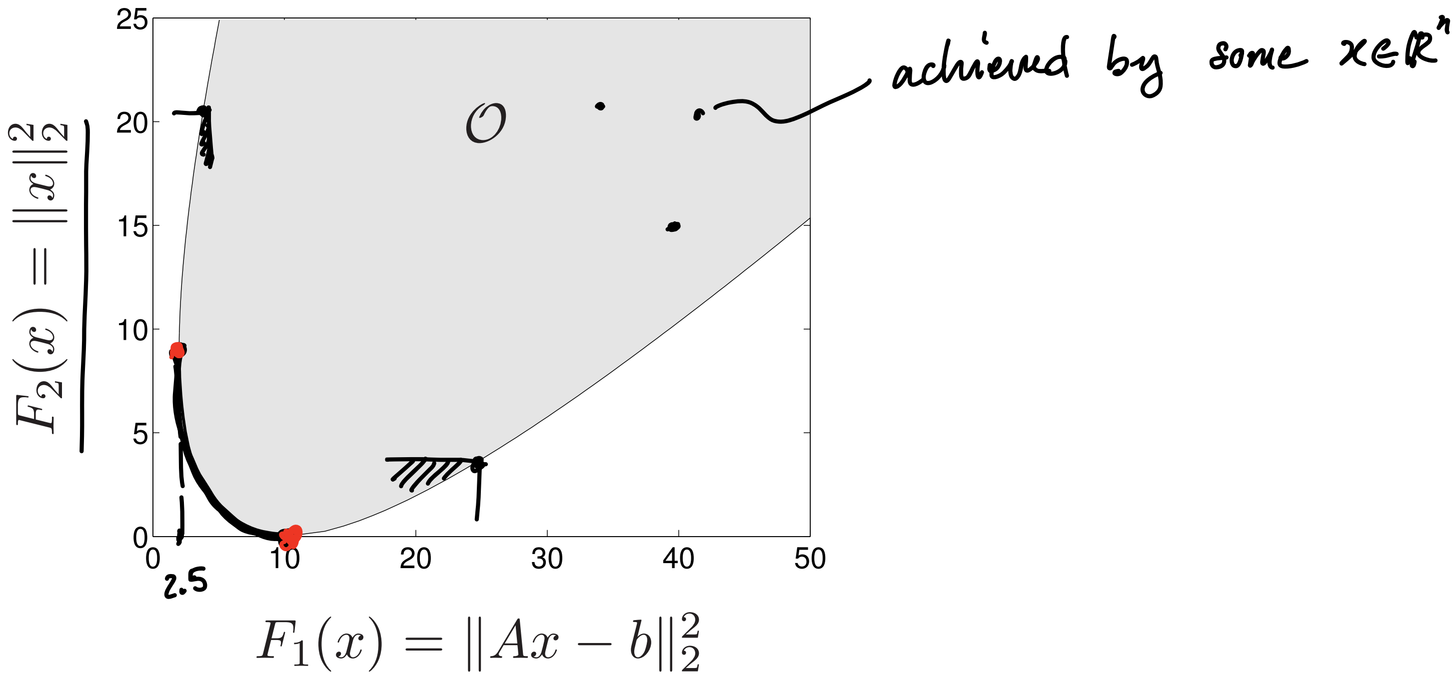


$$y \text{ feasible}, \quad f_0(y) \preceq f_0(x^{\text{po}}) \quad \Longrightarrow \quad f_0(x^{\text{po}}) = f_0(y)$$

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

Regularized least-squares

$$\underset{x \in \mathbb{R}^n}{\text{minimize (w.r.t. } \mathbf{R}^2_+)} \quad (\underbrace{\|Ax - b\|_2^2}_{F_1(x)}, \underbrace{\|x\|_2^2}_{F_2(x)})$$



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

$A_{m\times n}$, $n=10$, $b \in \mathbb{R}^m$, $m=100$ $\quad \|b\|_2^2 = 10 \quad \min_x \|Ax - b\|_2^2 \quad ? \sim 2.5$

1950's Markowitz portfolio opt.

Risk return trade-off in portfolio optimization

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

0

return & risk

- $x \in \mathbf{R}^n$ is investment portfolio; x_i is fraction invested in asset i
- $p \in \mathbf{R}^n$ is vector of relative asset price changes; modeled as a random variable with mean $\bar{p}$, covariance Σ

$p \sim D(\bar{p}, \Sigma)$

- $\bar{p}^T x = \mathbf{E}\, r$ is expected return; $x^T \Sigma x = \mathbf{var}\, r$ is return variance

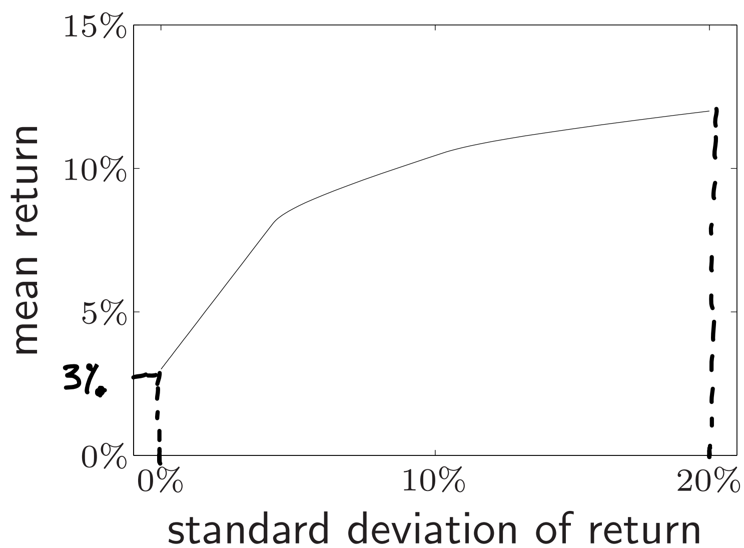
$x^T(p-\bar{p})$

return $= \sum p_i x_i = p^T x$ $\quad \mathbb{E}_p(p^T x) = \bar{p}^T x$ $\quad \mathbb{E}_p(p^T x) = \mathbb{E}_p(p^T x - \bar{p}^T x)(p^T x - \bar{p}^T x)^T$

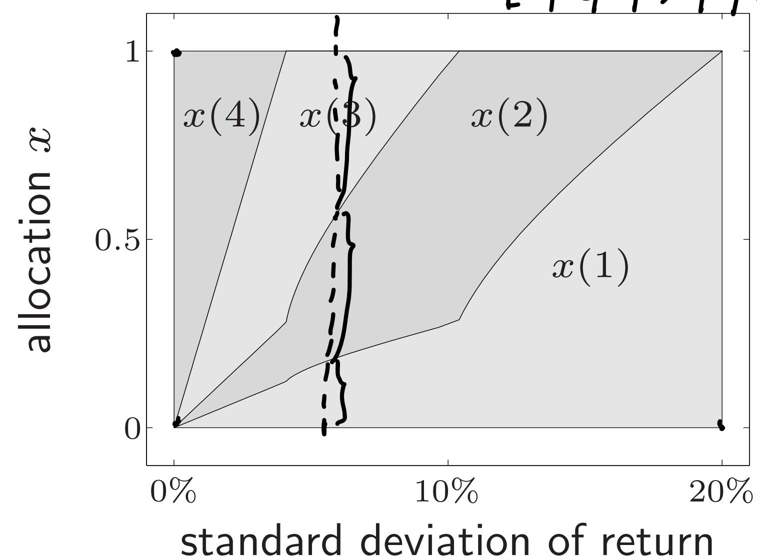
$= x^T[\mathbb{E}_p(p-\bar{p})(p-\bar{p})^T]\, x$

$= x^T \Sigma x$

example



$= \sqrt{x^T \Sigma x}$



$x \in \mathbb{R}^4$

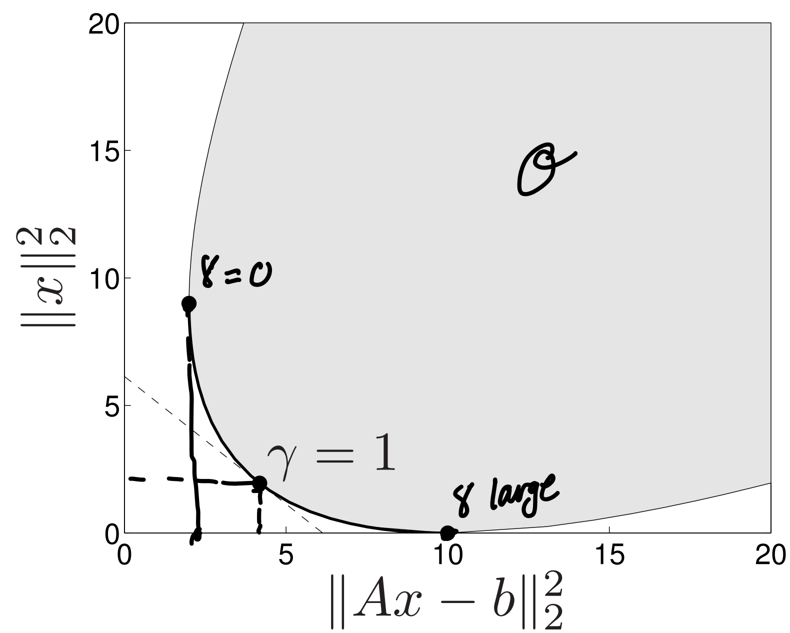
Scalarization for multicriterion problems

to find Pareto optimal points, minimize positive weighted sum

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

examples

- regularized least-squares problem of page 6–17



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

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

for fixed γ, a LS problem

$$\left\| \underbrace{\begin{bmatrix} A \\ \sqrt{\gamma} I \end{bmatrix}}_{\tilde{A}} x - \underbrace{\begin{bmatrix} b \\ 0 \end{bmatrix}}_{\tilde{b}} \right\|_2^2 \qquad x^* = \tilde{A}^\dagger \tilde{b}$$

$$= (A^T A + \gamma I)^{-1} A^T b$$

- risk-return trade-off of page 6–18

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

for fixed $\gamma > 0$, a quadratic program

in practice, discretize γ, solve one QP for each γ