4. Convex optimization problems

- optimization problem in standard form
- convex optimization problems
- quasiconvex optimization
- linear optimization
- quadratic optimization
- geometric programming
- generalized inequality constraints
- semidefinite programming
- vector optimization

Convex optimization problem

standard form convex optimization problem

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

- f_0, f_1, . . . , f_m are convex; equality constraints are affine
- problem is *quasiconvex* if f_0 is quasiconvex (and f_1, . . . , f_m convex)

often written as

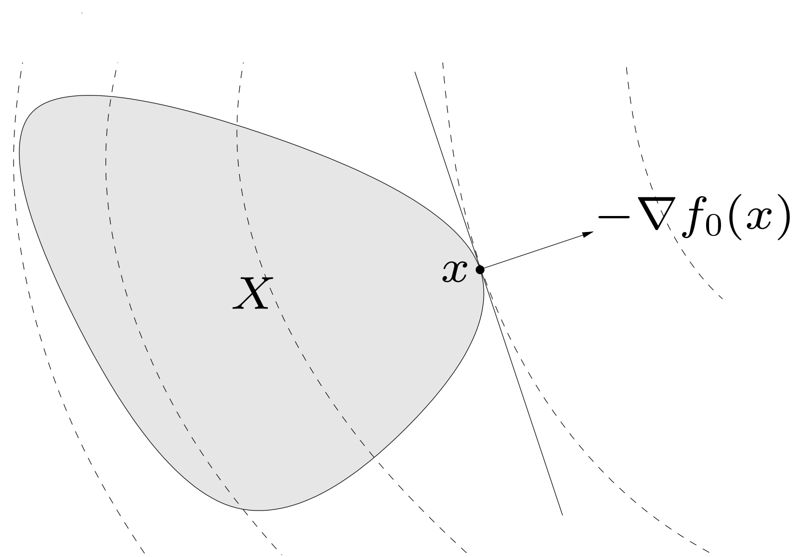
$$
\begin{array}{ll}
\text{minimize} & f_0(x) \\
\text{subject to} & f_i(x) \leq 0, \quad i = 1, \ldots, m \\
& Ax = b
\end{array}
$$

important property: feasible set of a convex optimization problem is convex

Optimality criterion for differentiable f_0

x is optimal if and only if it is feasible and

$$\nabla f_0(x)^T(y - x) \geq 0 \quad \text{for all feasible } y$$

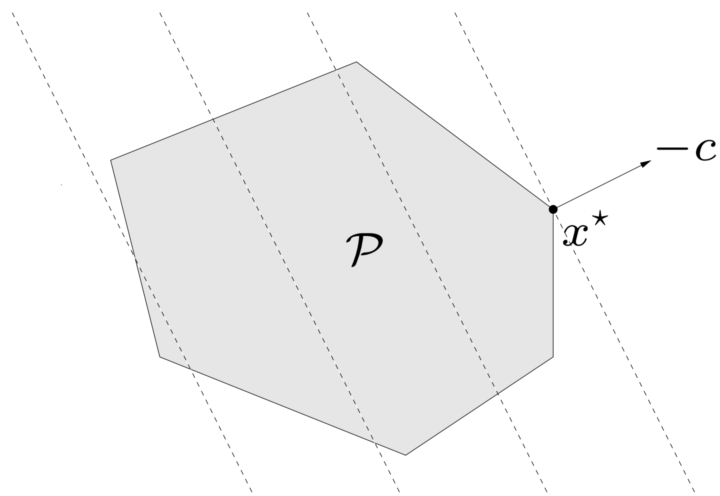


if nonzero, $\nabla f_0(x)$ defines a supporting hyperplane to feasible set X at x

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



$c^T x = \text{const}$

Linear-fractional program

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

linear-fractional program

$f_0(x) \le t \Rightarrow \frac{c^T x + d}{e^T x + f} \le t \Rightarrow c^T x + d \le t(e^T x + d)$

$f_0(x) \ge t$ $\le$

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

quasilinear

- a quasiconvex optimization problem; can be solved by bisection
- also equivalent to the LP (variables y, z)

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

Linear-fractional program

LP

$$\min \frac{c^T x + d}{e^T x + f} \quad \text{s.t. } Gx \le h, \quad Ax = b$$

$$\underset{z=1}{\Longleftrightarrow}$$

$$\min \frac{c^T x + dz}{e^T x + fz} \quad \text{s.t. } Gx \le h\cdot z, \quad Ax = b\cdot z$$

$$\Longleftrightarrow$$

$$\min\ c^T x + dz \quad \text{s.t. } Gx \le hz, \quad Ax = bz, \quad e^T x + fz = 1, \quad z \ge 0$$

$(P \in S^n)$

P : indefinite (NP-hard)

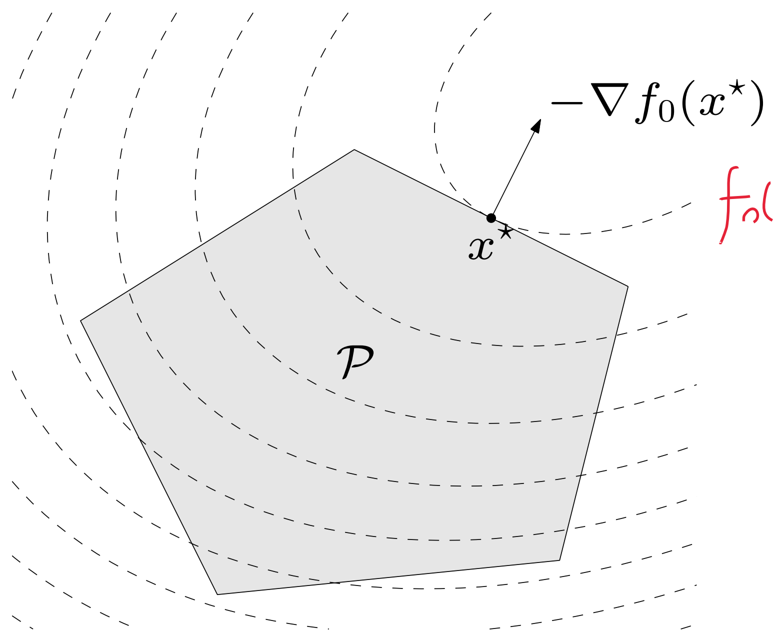
Quadratic program (QP)

(1950s)

$$
\begin{array}{ll}
\text{minimize} & \overset{f_0(x)}{(1/2)x^T P x + q^T x + r} \\
\text{subject to} & Gx \preceq h \\
& Ax = b
\end{array}
$$

$\nabla^2 f_0(x) = P \geq 0$

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



$f_0(x) = \text{const}$

$$A^+ = \begin{cases} A^{-1}, & A: \text{ square and non-singular} \\ (A^TA)^{-1}A^T, & \text{rank}(A)=n \\ A^T(AA^T)^{-1} & \text{rank}(A)=m \end{cases}$$

Examples

$A^TA \in S^n_+$

$\langle Ax-b, Ax-b\rangle = x^TA^TAx - 2x^TA^Tb + b^Tb$

least-squares

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

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

$A \in R^{m\times n}$, $\text{rank}(A) = r \le \min\{m, n\}$

$A = U\cdot\Sigma\cdot V^T$

$A^+ = V\cdot\Sigma^+\cdot U^T$

SVD

$$m\Big\{\overbrace{\begin{bmatrix} A \end{bmatrix}}^{n} = m\Big\{\overbrace{\begin{bmatrix} U \end{bmatrix}}^{m} \; m\Big\{\underbrace{\overset{\Sigma}{\begin{bmatrix} S_r & 0 \\ 0 & 0 \end{bmatrix}}}_{n} \; n\Big\{\overbrace{\begin{bmatrix} V^T \end{bmatrix}}^{n} = U\Sigma V^T$$

$Ax = b \iff U\cdot\Sigma V^T\cdot x = b$

$U^TU = UU^T = I_m$

$V^TV = V\cdot V^T = I_n.$

$\Sigma z = d$

$\Sigma^+ = n\Big\{\overbrace{\begin{bmatrix} S_r^{-1} & 0 \\ 0 & 0 \end{bmatrix}}^{m}$

$\iff \Sigma V^T x = U^T b$

$\iff V^T x = \Sigma^+ U^T b$

$z = \Sigma^+ b$

$\iff x = \underbrace{V\cdot\Sigma^+ U^T}_{A^+} b$

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

QP

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

- c is random vector with mean $\bar{c}$ and covariance Σ $\in S^n_+$
- hence, $c^T x$ is random variable with mean $\bar{c}^T x$ and variance $x^T \Sigma x$
- $\gamma > 0$ is risk aversion parameter; controls the trade-off between expected cost and variance (risk)

Quadratically constrained quadratic program (QCQP)

$P_i \in S^n_{++}$

degenerate ellipsoid

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

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

QCQP

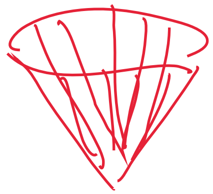
$P_i = 0$ → LP

$P_0 \in S^n_+$, $P_i = 0$, $i = 1, 2, \ldots, m$ → QP cone: $\{(x,t) \mid \|x\| \le t\}$

second-order cone: $\{(x,t) \mid \|x\|_2 \le t\}$

Second-order cone programming

(SOCP).



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

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

norm $\le$ linear

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

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

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

SOCP

QCQP

LP QP

- 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

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

$(\eta \geq \frac{1}{2})$

deterministic approach via SOCP

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

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

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

- robust LP

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

  is equivalent to the SOCP

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

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

$\sup_{\|u\|_2 \le 1} u^T (P_i^T x) = \|P_i^T x\|_2$

norm ≤ linear

stochastic approach via SOCP

$\phi\left(\frac{b_i - \bar{a}_i^T x}{\|\Sigma_i^{1/2} x\|_2}\right) \ge \eta \iff \frac{b_i - \bar{a}_i^T x}{\|\Sigma_i^{1/2} x\|_2} \ge \phi^{-1}(\eta)$

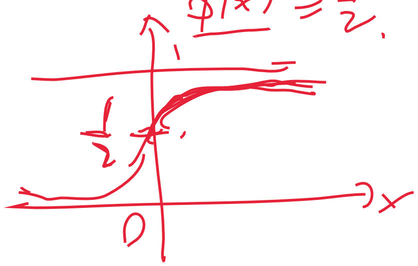
$\iff b_i - \bar{a}_i^T x \ge \phi^{-1}(\eta)\|\Sigma_i^{1/2} x\|_2$

- assume a_i is Gaussian with mean $\bar{a}_i$, covariance Σ_i ($a_i \sim \mathcal{N}(\bar{a}_i, \Sigma_i)$)

≥ 0

- $a_i^T x$ is Gaussian r.v. with mean $\bar{a}_i^T x$, variance $x^T \Sigma_i x$; hence

$\phi(x) \ge \frac{1}{2},\ x \ge 0$ $\qquad \bar{a}_i^T x - b_i \le 0$ $\qquad F(\eta) = \phi\left(\frac{\eta - \mu}{\sigma}\right)$



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

$\ge \eta$ $\quad \ge \frac{1}{2}$

$F(x) = \phi\left(\frac{x - \mu}{\sigma}\right)$

CDF

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

- robust LP

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

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

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

Geometric programming (1970s)

monomial function

positive

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

$f(x) = x_1^2 x_2^{\frac{1}{2}} x_3$

with $c > 0$; exponent a_i can be any real number

posynomial function: sum of monomials

$f(x) = x_1 + x_2 + x_1 x_2^2 + x_2 x_3^{\frac{1}{2}}$

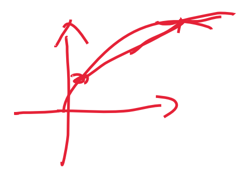
$$f(x) = \sum_{k=1}^{K} c_k x_1^{a_{1k}} x_2^{a_{2k}} \cdots x_n^{a_{nk}}, \qquad \mathbf{dom}\, f = \mathbf{R}^n_{++}$$

geometric program (GP)

does not need to be convex

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

$= \sqrt{x}$



with f_i posynomial, h_i monomial

$f(y) = c\, e^{a_1 y_1} \cdots e^{a_n y_n} = c\, e^{\sum a_i y_i} \iff \log f(y) = a^T y + \underline{b}$ $(\log c)$

$= e^{a^T y + b}, (b = \log c)$

Geometric program in convex form

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

$x_i = e^{y_i}$

- 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

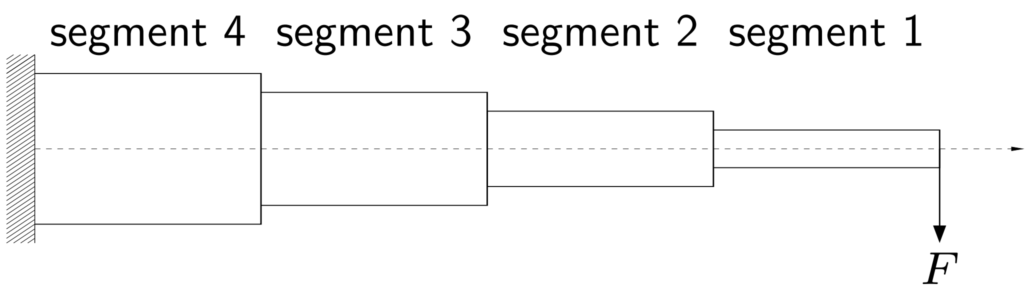
$$\log f(e^{y_1}, \ldots, e^{y_n}) = \log \left(\sum_{k=1}^K e^{a_k^T y + b_k} \right) \qquad (b_k = \log c_k)$$

log-sum-exp : convex

- 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 $i = 1, \ldots, N$

non-convex

objective and constraint functions

$$\begin{bmatrix} w \\ h \end{bmatrix}^T \begin{bmatrix} ? & & \\ & \ddots & \\ & & ? \end{bmatrix} \begin{bmatrix} w \\ h \end{bmatrix}$$

- total weight $w_1 h_1 + \cdots + w_N h_N$ 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 $6iF/(w_i h_i^2)$, 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

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

note

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

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

Generalized inequality constraints

convex problem with generalized inequality constraints

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

K_i: different proper cones

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

- $f_0 : \mathbf{R}^n \to \mathbf{R}$ 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

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

if $K = R^m_+$: LP

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

Semidefinite program (SDP)

(90s)

(60s)

(LMI)

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

$K = S^n_+$

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

SDP

SOCP

QCQP

LP QP

- 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

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

LP and SOCP as SDP

SDP embedding.

$\mathrm{diag}(a) = \begin{bmatrix} a_1 & & \\ & \ddots & \\ & & a_n \end{bmatrix}$

LP and equivalent SDP

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

$(K = \mathbb{R}^m_+)$ $(K = S^n_+)$

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

SOCP and equivalent SDP

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

$\Updownarrow$

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

Quiz: how to represent QP as SDP?

LP and SOCP as SDP

$$M = \begin{bmatrix} A & B \\ C & D \end{bmatrix} \qquad M : \text{PSD} \iff \begin{array}{l} M/A : \text{PSD} \\ A : \text{PSD} \end{array}$$

Schur complement:

$$M/A = D - C^T A^{-1} B$$

$$\iff \begin{cases} \underline{M/A = c_i^T x + d_i - (A_i x + b_i)^T \dfrac{I}{c_i^T x + d_i} (A_i x + b_i) \geq 0} \\ A = c_i^T x + d_i \geq 0 \end{cases}$$

$$\Updownarrow$$

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

$$\iff \|A_i x + b_i\|_2 \leq c_i^T x + d_i$$

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

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

$(A - tI) \leq 0$

Spectral norm $\left(A\right) \in \mathbb{R}^{m\times n}$ $\quad \|A\|_2 = \sqrt{\lambda_{\max}(A^TA)} = \sigma_{\max}(A)$

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

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

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

$$tI \succ 0$$

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

$$M/A \Longleftrightarrow M/P = tI - A^T \cdot (tI)^{-1} A \succeq 0$$

$$tI \succ 0$$

Vector optimization

general vector optimization problem

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

$(q=1)$

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

convex vector optimization problem

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

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

X : minimum ; $\forall y \in S, \Rightarrow y \succeq_K x$

x : minimal : $y \in S,\ y \succeq_K x \Rightarrow x = y$

Optimal and Pareto optimal points

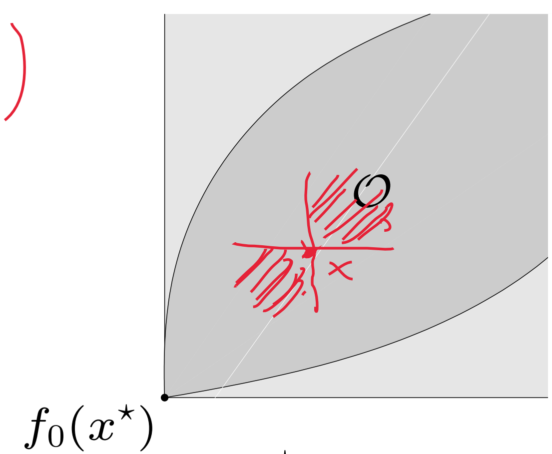
set of achievable objective values

($x \in \mathbf{dom}\, f_0$, $x \in$ constraints)

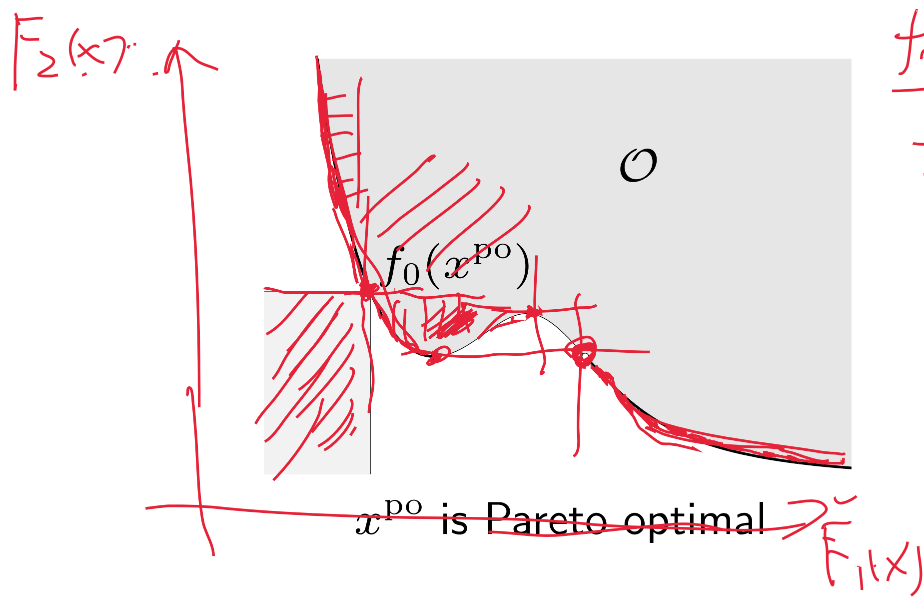
$$\mathcal{O} = \{f_0(x) \mid x \text{ feasible}\}$$

- 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 $\mathcal{O}$

$(K = \mathbb{R}^2_+)$



$x^\star$ is optimal



$x^{\rm po}$ is Pareto optimal

$x \in \mathcal{O}$

$f_0(x) \succeq f_0(x^{po})$

$f_0(x) = f_0(x^{po})$

$f_0(x) = (F_1(x), F_2(x))$

(Multi-objective)

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

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

  if there exists an optimal point, the objectives are noncompeting

- feasible x^{po} is Pareto optimal if

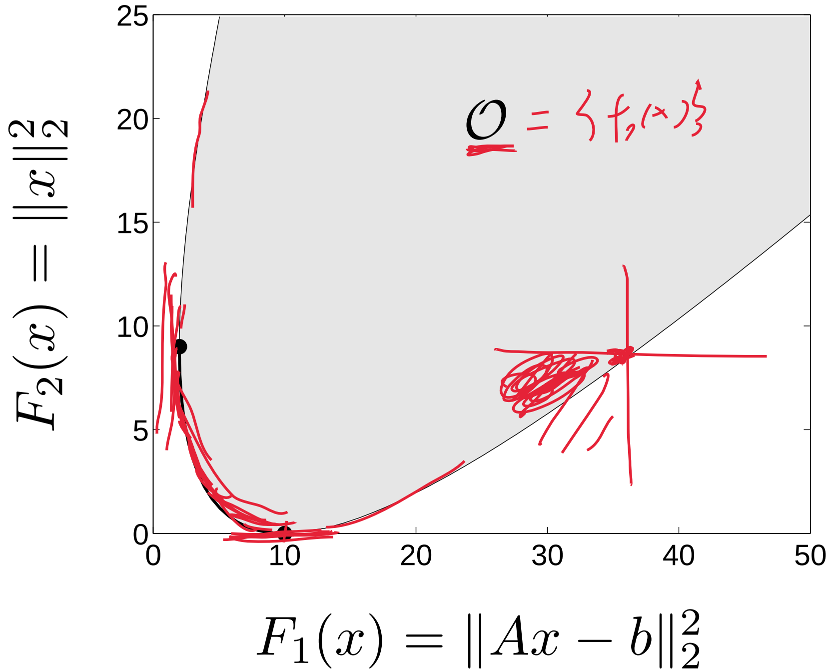
$$y \text{ feasible}, \quad f_0(y) \preceq f_0(x^{\text{po}}) \implies f_0(x^{\text{po}}) = f_0(y)$$

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

Regularized least-squares

$f_0(x) = (F_1(x), F_2(x))$

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



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

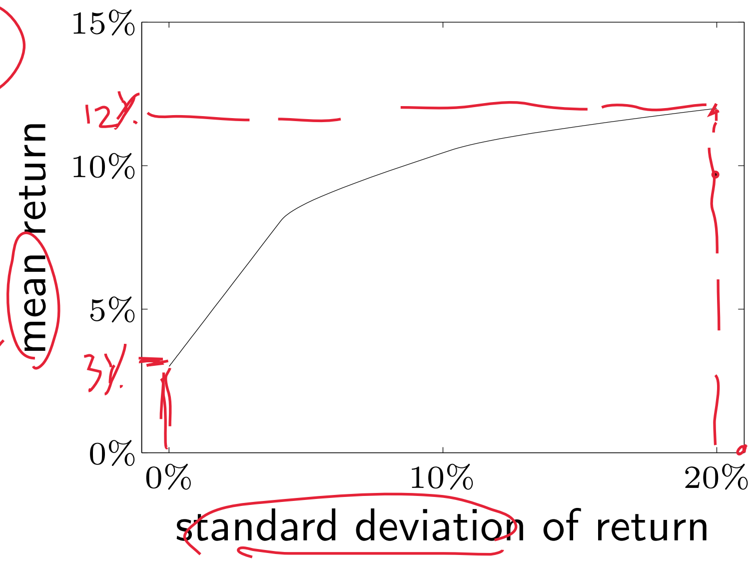
Risk return trade-off in portfolio optimization

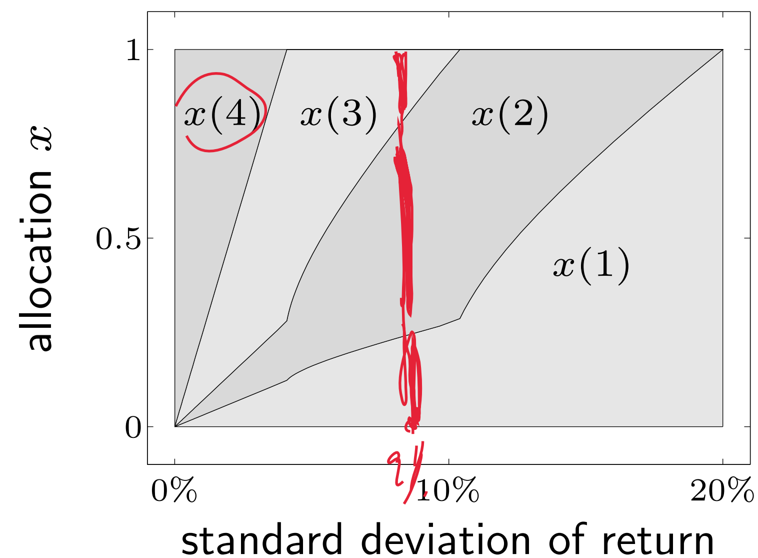
$$\begin{array}{ll} \text{minimize (w.r.t. } \mathbf{R}_+^2) \overset{f_0(x)}{=} & (-\bar{p}^T x, x^T \Sigma x) \\ \text{subject to} & \mathbf{1}^T x = 1, \quad x \succeq 0 \end{array}$$

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

example

Asset	$\bar{p}_i$	$\Sigma_{ii}^{1/2}$
1	12%	20%
2	10%	10%
3	7%	5%
4	3%	0%





Dual cone.

$K^* = \{y \mid y^T x \geq 0, \ \forall x \in K\}$

Scalarization

$\Downarrow$

x minimal (convex) $\Leftarrow$ some $\lambda \succeq_{K^*} 0$

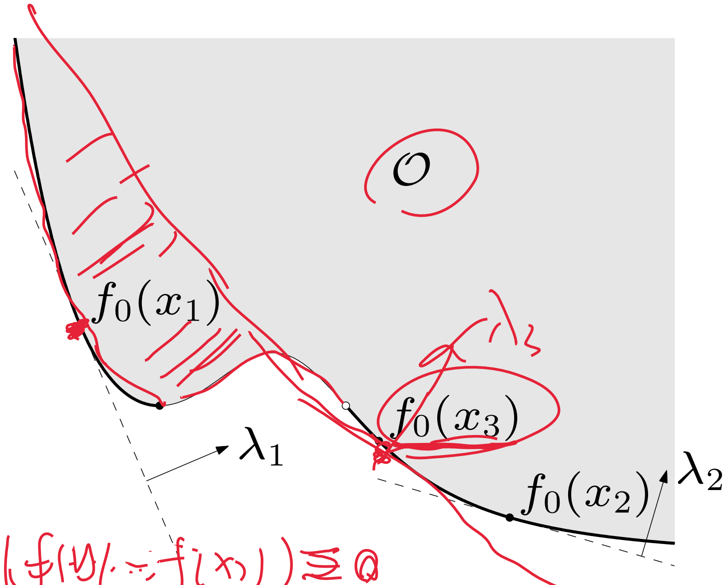
$\min_{\lambda} \lambda^T x$

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

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

$y = f_0(x)$: minimal $\Rightarrow$ x: Pareto optimal

if x is optimal for scalar problem, then it is Pareto-optimal for vector optimization problem



for convex vector optimization problems, can find (almost) all Pareto optimal points by varying $\lambda \succ_{K^*} 0$

Scalarization for multicriterion problems

to find Pareto optimal points, minimize positive weighted sum

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

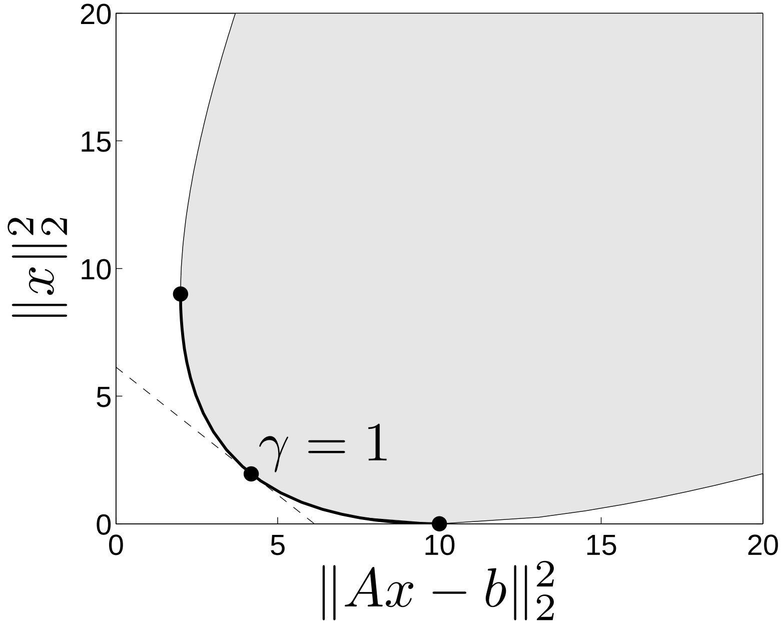
examples

- regularized least-squares problem of page 4–43

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 4–44

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

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

5. Duality

- Lagrange dual problem
- weak and strong duality
- geometric interpretation
- optimality conditions
- perturbation and sensitivity analysis
- examples
- generalized inequalities

Lagrangian

standard form problem (not necessarily convex)

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

variable $x \in \mathbf{R}^n$, domain $\mathcal{D}$, optimal value $p^\star$

Lagrangian: $L : \mathbf{R}^n \times \mathbf{R}^m \times \mathbf{R}^p \to \mathbf{R}$, with $\mathbf{dom}\, L = \mathcal{D} \times \mathbf{R}^m \times \mathbf{R}^p$,

$$
L(x, \lambda, \nu) = f_0(x) + \sum_{i=1}^{m} \lambda_i f_i(x) + \sum_{i=1}^{p} \nu_i h_i(x)
$$

- weighted sum of objective and constraint functions
- λ_i is Lagrange multiplier associated with $f_i(x) \leq 0$
- ν_i is Lagrange multiplier associated with $h_i(x) = 0$

Lagrange dual function

Lagrange dual function: $g : \mathbf{R}^m \times \mathbf{R}^p \to \mathbf{R}$,

$$\begin{aligned} g(\lambda, \nu) &= \inf_{x \in \mathcal{D}} L(x, \lambda, \nu) \\ &= \inf_{x \in \mathcal{D}} \left(f_0(x) + \sum_{i=1}^{m} \lambda_i f_i(x) + \sum_{i=1}^{p} \nu_i h_i(x) \right) \end{aligned}$$

g is concave, can be $-\infty$ for some λ, ν

lower bound property: if $\lambda \succeq 0$, then $g(\lambda, \nu) \leq p^\star$

proof: if $\tilde{x}$ is feasible and $\lambda \succeq 0$, then

$$f_0(\tilde{x}) \geq L(\tilde{x}, \lambda, \nu) \geq \inf_{x \in \mathcal{D}} L(x, \lambda, \nu) = g(\lambda, \nu)$$

minimizing over all feasible $\tilde{x}$ gives $p^\star \geq g(\lambda, \nu)$

Least-norm solution of linear equations

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

dual function

- Lagrangian is $L(x, \nu) = x^T x + \nu^T(Ax - b)$
- to minimize L over x, set gradient equal to zero:

$$\nabla_x L(x, \nu) = 2x + A^T \nu = 0 \quad \Longrightarrow \quad x = -(1/2)A^T \nu$$

- plug in in L to obtain g:

$$g(\nu) = L((-1/2)A^T \nu, \nu) = -\frac{1}{4}\nu^T A A^T \nu - b^T \nu$$

  a concave function of ν

lower bound property: $p^\star \geq -(1/4)\nu^T A A^T \nu - b^T \nu$ for all ν

Standard form LP

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

dual function

- Lagrangian is

$$
\begin{array}{rcl}
L(x, \lambda, \nu) & = & c^T x + \nu^T (Ax - b) - \lambda^T x \\
& = & -b^T \nu + (c + A^T \nu - \lambda)^T x
\end{array}
$$

- L is affine in x, hence

$$
g(\lambda, \nu) = \inf_x L(x, \lambda, \nu) = \begin{cases} -b^T \nu & A^T \nu - \lambda + c = 0 \\ -\infty & \text{otherwise} \end{cases}
$$

g is linear on affine domain $\{(\lambda, \nu) \mid A^T \nu - \lambda + c = 0\}$, hence concave

lower bound property: $p^\star \geq -b^T \nu$ if $A^T \nu + c \succeq 0$

Equality constrained norm minimization

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

dual function

$$g(\nu) = \inf_x (\|x\| - \nu^T Ax + b^T\nu) = \begin{cases} b^T\nu & \|A^T\nu\|_* \le 1 \\ -\infty & \text{otherwise} \end{cases}$$

where $\|v\|_* = \sup_{\|u\|\le 1} u^T v$ is dual norm of $\|\cdot\|$

proof: follows from $\inf_x(\|x\| - y^T x) = 0$ if $\|y\|_* \le 1$, $-\infty$ otherwise

- if $\|y\|_* \le 1$, then $\|x\| - y^T x \ge 0$ for all x, with equality if $x = 0$
- if $\|y\|_* > 1$, choose $x = tu$ where $\|u\| \le 1$, $u^T y = \|y\|_* > 1$:

$$\|x\| - y^T x = t(\|u\| - \|y\|_*) \to -\infty \quad \text{as } t \to \infty$$

lower bound property: $p^\star \ge b^T\nu$ if $\|A^T\nu\|_* \le 1$

Two-way partitioning

$$\begin{array}{ll} \text{minimize} & x^T W x \\ \text{subject to} & x_i^2 = 1, \quad i = 1, \ldots, n \end{array}$$

- a nonconvex problem; feasible set contains 2^n discrete points
- interpretation: partition $\{1, \ldots, n\}$ in two sets; W_{ij} is cost of assigning i, j to the same set; $-W_{ij}$ is cost of assigning to different sets

dual function

$$\begin{aligned} g(\nu) = \inf_x (x^T W x + \sum_i \nu_i (x_i^2 - 1)) &= \inf_x x^T (W + \mathbf{diag}(\nu)) x - \mathbf{1}^T \nu \\ &= \begin{cases} -\mathbf{1}^T \nu & W + \mathbf{diag}(\nu) \succeq 0 \\ -\infty & \text{otherwise} \end{cases} \end{aligned}$$

lower bound property: $p^\star \geq -\mathbf{1}^T \nu$ if $W + \mathbf{diag}(\nu) \succeq 0$

example: $\nu = -\lambda_{\min}(W)\mathbf{1}$ gives bound $p^\star \geq n\lambda_{\min}(W)$

Lagrange dual and conjugate function

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

dual function

$$\begin{aligned} g(\lambda, \nu) &= \inf_{x \in \mathbf{dom}\, f_0} \left(f_0(x) + (A^T\lambda + C^T\nu)^T x - b^T\lambda - d^T\nu \right) \\ &= -f_0^*(-A^T\lambda - C^T\nu) - b^T\lambda - d^T\nu \end{aligned}$$

- recall definition of conjugate $f^*(y) = \sup_{x \in \mathbf{dom}\, f}(y^Tx - f(x))$
- simplifies derivation of dual if conjugate of f_0 is known

example: entropy maximization

$$f_0(x) = \sum_{i=1}^n x_i \log x_i, \qquad f_0^*(y) = \sum_{i=1}^n e^{y_i - 1}$$