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

Optimization problem in standard form

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

w/o the min., this is the optimization problem

- $x \in \mathbf{R}^n$ 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

optimal value:

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

- $p^\star = \infty$ if problem is infeasible (no x satisfies the constraints) (i.e., is $\emptyset$)
- $p^\star = -\infty$ if problem is unbounded below
- ~~solutions~~ optimal points are not necessarily unique + basic solution finds just 1 sol'n. (opt. pt.)

Optimal and locally optimal points

- $X_{opt} = \emptyset$ if $p^* = \infty$ or $p^* = -\infty$.
- $X_{opt} = \emptyset$ if optimal point(s) not in dom f. (e.g., $f_0(x) = \frac{1}{x}$ on $\mathbf{R}_{++}$)

x is **feasible** if $x \in \mathbf{dom}\, f_0$ and it satisfies the constraints

a feasible x is **optimal** if $f_0(x) = p^\star$; X_{opt} is the set of optimal points

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

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

— add constraint for z to be in the closed ball centered at x w/ radius R

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

- $f_0(x) = 1/x$, $\mathbf{dom}\, f_0 = \mathbf{R}_{++}$: $p^\star = 0$, no optimal point - bc not in dom f
- $f_0(x) = -\log x$, $\mathbf{dom}\, f_0 = \mathbf{R}_{++}$: $p^\star = -\infty$
- $f_0(x) = x \log x$, $\mathbf{dom}\, f_0 = \mathbf{R}_{++}$: $p^\star = -1/e$, $x = 1/e$ is optimal
- $f_0(x) = x^3 - 3x$, $p^\star = -\infty$, local optimum at $x = 1$

differentiate and set to 0

Implicit constraints

the standard form optimization problem has an **implicit constraint**

$$x \in \mathcal{D} = \underbrace{\bigcap_{i=0}^{m} \mathbf{dom}\, f_i}_{\text{includes } f_0} \cap \bigcap_{i=1}^{p} \mathbf{dom}\, h_i,$$

• i.e., all x in domain of problem are in domain of ~~all~~ objective func. + all constraints

- we call $\mathcal{D}$ the **domain** of the problem
- the constraints $f_i(x) \leq 0$, $h_i(x) = 0$ are the explicit constraints ~~incl. $f_0(x)$~~
- a problem is **unconstrained** if it has no explicit constraints ($m = p = 0$)

example:

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

— open polyhedron

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

Feasibility problem

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

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

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

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

all x are equally good

Convex optimization problem

standard form convex optimization problem

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

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

often written as

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

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

~~Convex opt. prob~~ • Whether or not something is a convex opt. problem is an attribute of the problem, not of the feasible set.

example

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

$1 + x^2 > 0$, so $f_1(x) \equiv x_1 \le 0$, but $x_1/(1+x_2^2)$ not convex

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

by solving one you can (without much effort) construct the solution of the other

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

Local and global optima

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

proof: suppose x is locally optimal and y is optimal with $f_0(y) < f_0(x)$

x locally optimal means there is an $R > 0$ such that

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

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

$$f_0(z) \le \theta f_0(x) + (1 - \theta) f_0(y) < f_0(x)$$

which contradicts our assumption that x is locally optimal

Optimality criterion for differentiable f_0

so this ~~is a~~ determines halfspace

x is optimal if and only if it is feasible and

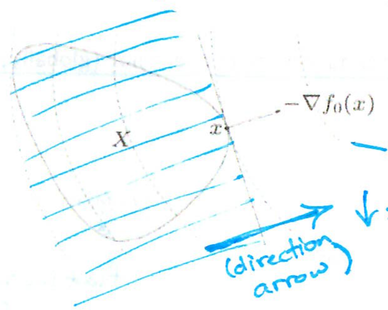
• applies only when f_0 differentiable at x

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

constant

~~or $\nabla f_0(x)^T y \geq \nabla f_0(x)^T x$~~

or $\nabla f_0(x)^T y \geq \nabla f_0(x)^T x$ $\forall$ feasible y



— level curves (rmbr. convex f → convex sublevel sets)

$\downarrow f_0$

(direction arrow)

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

from supporting hyperplane theorem

• **unconstrained problem**: x is optimal if and only if

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

• $g^T z \geq 0 \ \forall z \in \mathbb{R}^n$
→ $g = 0$
(consider $z = -g$)

• **equality constrained problem**

$$\text{minimize} \quad f_0(x) \quad \text{subject to} \quad Ax = b$$

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

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

• x optimal, y feasible
→ $Ax = b$
$Ay = b$
→ $A(y-x) = 0$
→ $y - x \in \mathrm{Nul}\, A$
• $\nabla f_0(x)^T(y-x) \geq 0$
• $\nabla f_0(x)^T$ has nonneg inner prod. w/ everything in Nul A
• Nul A is a subspace
→ $\nabla f_0(x)^T(y-x) = 0$
→ $\nabla f_0(x) \perp \mathrm{Nul}\, A$
$\nabla f_0(x) \in (\mathrm{Nul}\, A)^\perp$
$\nabla f_0(x) \in \mathrm{Row}\, A$
$\nabla f_0(x) \in \mathrm{Col}\, A^T$
$\nabla f_0(x) = A^T u$ for some $u \in \mathbb{R}^m$
$\nabla f_0(x) + A^T(-u) = 0$

• **minimization over nonnegative orthant**

$$\text{minimize} \quad f_0(x) \quad \text{subject to} \quad x \succeq 0$$

x is optimal if and only if

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

• Optimality criterion for differentiable f_0 used to derive these 3 results on this page.

Equivalent convex problems

two problems are (informally) **equivalent** if the solution of one is readily obtained from the solution of the other, and vice-versa

some common transformations that preserve convexity:

• **eliminating equality constraints**

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

is equivalent to

• $A(Fz + x_0) = b$
$AFz + Ax_0 = b$
$0 + Ax_0 = b$

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

convex bc these are convex ~~fn~~ func. precomposed w/ affine.

where F and x_0 are such that

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

• col. of F ~~are in~~ span Nul A
• x_0 a solution to $Ax = b$

• **introducing equality constraints** often very useful to introduce 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**

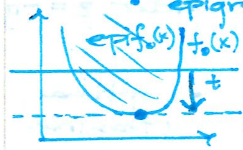
$$\begin{array}{ll} \text{minimize} & f_0(x) \\ \text{subject to} & a_i^T x \leq b_i, \quad i = 1, \ldots, m \end{array}$$

is equivalent to

change ~~switch~~ here

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

- **epigraph form**: standard form convex problem is equivalent to
  - "epigraph trick"



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

- WLOG can just minimize linear objective.

$\equiv f_0(x) \le t$

- optimal is $f_0(x) = t$ but this form or $f_0(x) - t = 0$ not convex unless $f_0(x)$ affine.

- **minimizing over some variables**

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

is equivalent to

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

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

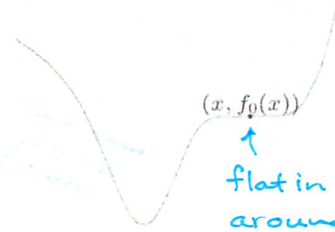
- If original objective is convex, then minimizing over any # of variables preserves convexity
  → dynamic programming

Quasiconvex optimization

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

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



flat in some neighborhood around $(x, f_0(x))$

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.*,

$$f_0(x) \le t \iff \underbrace{\phi_t(x) \le 0}$$

want to know whether the set of x that satisfies this convex fixed t

example

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

$\frac{p(x)}{q(x)} \le t$

$p(x) - \underbrace{t q(x)}_{} \le 0$

- concave → convex

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 $\phi_t(x) \le 0$

since $p(x)/q(x) \ge 0$, if $t < 0$, then t-sublevel set of $f_0(x)$ is empty, which is convex

- How to solve quasiconvex optimization problem

quasiconvex optimization via convex feasibility problems

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

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

Bisection method for quasiconvex optimization

given $l \le p^\star$, $u \ge p^\star$, 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 $u - l \le \epsilon$.

- given upper and lower bounds l and u

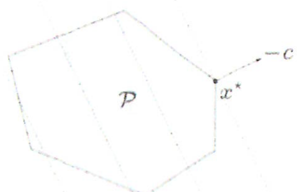
requires exactly $\lceil \log_2((u - l)/\epsilon) \rceil$ iterations (where u, l are initial values)

Linear program (LP)

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

in this form, usually called "linear" wrt optimization

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



- max # of vertices ↑ at least factorially(?) as # of variables ↑
- For LP, always a sol'n at a vertex
- level sets of affine func. are ~~affine~~ hyperplanes

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

- piecewise linear func. can be used to approx. any convex func.

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

equivalent to an LP

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

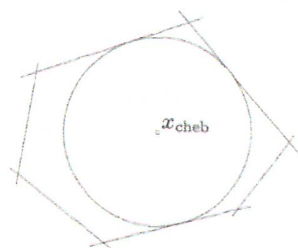
Chebyshev center of a polyhedron

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

  - general norm → ② is dual norm of ①

$$\sup\{a_i^T(x_c + u) \mid \underbrace{\|u\|_2 \le r}_{①}\} = a_i^T x_c + r\underbrace{\|a_i\|_2}_{②} \le b_i$$

sup when $a_i^T u$ in same direc → $a_i^T\left(\frac{a_i r}{\|a_i\|_2}\right) = r\|a_i\|_2$

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

$$\begin{array}{ll} \text{maximize} & r \\ \text{subject to} & a_i^T x_c + \underbrace{r\|a_i\|_2}_{\text{norms are nonlinear but this isn't a func. of } a_i} \le b_i, \quad i = 1, \ldots, m \end{array}$$

(Generalized) linear-fractional program

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

linear-fractional program

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

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

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

- "normalization" trick (see book)

generalized linear-fractional program

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

a quasiconvex optimization problem; can be solved by bisection

• can't convert to LP

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
- $(Ax)_i$, $(Bx^+)_i$: produced, resp. consumed, amounts of good i
- x_i^+/x_i: growth rate of sector i

allocate activity to maximize growth rate of slowest growing sector

Quadratic program (QP)

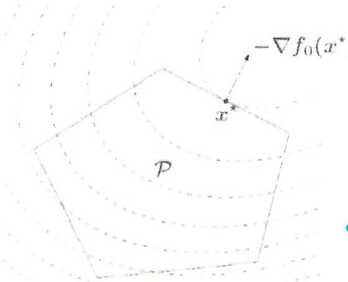
$$\begin{array}{ll} \text{minimize} & (1/2)x^T P x + q^T x + r \\ \text{subject to} & Gx \preceq h \\ & Ax = b \end{array}$$

} affine constraints

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

• extension to LP

• If P indefinite, (even w/ just one neg λ) problem becomes In general but not always NP-hard.



• level curves of a convex quadratic are ellipsoids

• May not exist a sol'n at a vertex

Examples

least-squares

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

- 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

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

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

$\gamma < 0$ is risk seeking + nonconvex

Quadratically constrained quadratic program (QCQP)

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

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

• extension to QP

• If P not full rank → degenerate ellipsoid → infinite in dim NulP directions

• a half space ~~halfspace~~ ~~hyperplane~~ is a degen. ellipsoid (shape matrix λ's are all ∞ but 1)

• rank P = degrees of curvature of degen ellipsoid.

$P \in S^3$ w rank=2 → infinite cylinder in $\mathbb{R}^3$

Second-order cone programming (SOCP)

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

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

• Note $\|A_ix - b_i\|_2 - c_i^Tx + d_i$ is not differentiable at 0, but since graph is pointed at 0, many times the sol'n is at 0.

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

$$(\underbrace{A_ix + b_i}_{\text{vec}}, \underbrace{c_i^Tx + d_i}_{\text{scalar}}) \in \text{second-order cone in } \mathbf{R}^{n_i+1}$$

bc 2nd order cone $C = \{(z,t) \in \mathbf{R}^{n+1} \mid \|z\|_2 \le t\}$

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

• ~~can~~ Almost always can solve SOCP as fast as LP

Robust linear programming

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

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

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 $a_i \in \mathcal{E}_i$ ellipsoid

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

• could be to limit worst case

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

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

• both are SOCPs

deterministic approach via SOCP

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

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

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

- robust LP

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

is equivalent to the SOCP

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

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

~~$\bar{a}_i +$~~ $\bar{a}_i^Tx + (P_iu)^Tx = \underbrace{\bar{a}_i^Tx}_{\text{constant}} + \underbrace{u^T(P_i^Tx)}_{\text{sup when } u = \frac{P_i^Tx}{\|P_i^Tx\|}}$ — from Cauchy-Schwarz: $|a^Tb| \le \sqrt{a^Ta}\sqrt{b^Tb}$ or $|\langle a, b\rangle| \le \|a\|_2\|b\|_2$

stochastic approach via SOCP

- assume a_i is Gaussian with mean $\bar{a}_i$, covariance Σ_i ($a_i \sim \mathcal{N}(\bar{a}_i, \Sigma_i)$)
- a_i^Tx is Gaussian r.v. with mean $\bar{a}_i^Tx$, variance $x^T\Sigma_ix$; hence

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

— standardized; Φ is standard 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

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

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

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

~~only works when each affine constraint~~

• works only when we impose constraint that each affine constraint is satisfied w/ prob. ≥ 0.5

Geometric programming

monomial function

diff than math monomial

$$f(x) = cx_1^{a_1}x_2^{a_2}\cdots x_n^{a_n}, \qquad \mathbf{dom}\, f = \mathbf{R}_{++}^n$$

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)

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

also posynomial
• generally not convex

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

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

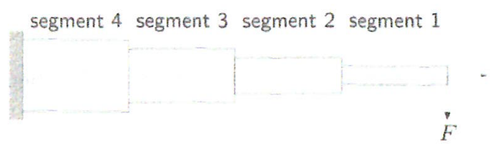
log-sum-exp (convex) precomposed w/ func. affine in y → convex in y

- geometric program transforms to convex problem

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

• can ~~put log~~ take log and min. bc log monotone

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 $i = 1, \ldots, N$

• exp of the average of logs of $x_1, x_2, \ldots, x_n$ = geometric mean of $x_1, x_2, \ldots, x_n$

objective and constraint functions

quadratic func. of (w,h) but not convex. Mat. of quad. form has 0 along diag + nonzero terms. ~~in some of~~ offdiag Matrix ~~$\in S^n$~~ w/ 0 on diag $\in S_+^n$ iff ~~all~~ row + col. w/ diag are 0.

- 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
- the vertical deflection y_i and slope v_i of central axis at the right end of segment i are defined recursively as

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

monomials

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

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

note

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

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

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

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

Minimizing spectral radius of nonnegative matrix

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:

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

variables λ, v, x

Generalized inequality constraints

convex problem with generalized inequality constraints

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

$f_i: R^n \to R^k$

vector inequality (each $-f_i(x) \in K_i$)

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

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

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

- lot of interesting problems that transform to conic problems

Semidefinite program (SDP)

(subset of cone program)

$$\begin{array}{ll} \text{minimize} & c^T x \\ \text{subject to} & x_1F_1 + x_2F_2 + \cdots + x_nF_n + G \preceq 0 \\ & Ax = b \end{array}$$

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

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

- bc if block matrix neg. semi def. then each block is neg. semidef.
  → can work w/ just 1 LMI WLOG

LP and SOCP as SDP

• SDP subsumes everything seen so far except geometric prog.

LP and equivalent SDP

SDP embedding

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

vec.

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

mat.

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

SOCP and equivalent SDP

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

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

$\begin{bmatrix} A & B \\ B^T & C \end{bmatrix}$ symm.

• SDP solver can solve these all

• Let $X/A = C - B^T A^{-1} B$ (Schur complement of A in X)

• X PD iff A is PD + X/A is PD • ~~X PSD if~~ If A is ~~PSD~~ PD, X is PSD iff X/A PSD.

X/A: $c_i^T x + d_i^2 \ge (A_i x + b_i)^T \left(\frac{I^{-1}}{c_i^T x + d_i}\right)(A x_i + b_i)$

$\rightarrow \|A_i x + b_i\|_2 \le c_i^T x + d_i{}^2$

~~can sq.~~ can mult. by $c_i^T x + d_i$ w/o switching sign bc we know it is ~~≥ 0 for $\succ 0$~~ ~~for PD X~~ for PSD X.

Eigenvalue minimization

matrix-valued func.

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

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

• constant mat. plus lin. comb. of mat. (all symm.)

equivalent SDP

minimize t
subject to $A(x) \preceq tI$

• nondifferentiable
• highly nonlinear

$\equiv A(x) - tI \preceq 0$

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

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

Matrix norm minimization

$\|A\|_2 = \|A\|_\sigma$ (spectral norm) $= \sigma_{\max} A$

largest singular value of A

• $\|A\|_1$ = max absolute col. sum
• $\|A\|_\infty$ = max abs. row sum
• $\|A\|_F$ = sqrt. of sum of squared abs. value of each elem.
• $\|A\|_2 \le \|A\|_F$

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

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

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

PSD only if
1. $t \ge 0$
2. $tI \succeq A(x)^T \left(\frac{I}{t}\right) A(x)$

$t^2 I \succeq A(x)^T A(x)$

compares λ's bc mat. inequal.

$\boxed{\sigma \text{ of } A \text{ are sqrts of } \lambda\text{'s of this}}$

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

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

• $A(x)^T A(x)$ is quadratic and not linear as req. by SDP

→ idea of SDP is to make something bigger but that is linear in var. of interest

Vector optimization

general vector optimization problem

• here we generalize the objective to vectors

minimize (w.r.t. K) $f_0(x)$
subject to $f_i(x) \le 0, \quad i = 1, \dots, m$
$h_i(x) \le 0, \quad i = 1, \dots, p$

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

minimize (w.r.t. K) $f_0(x)$
subject to $f_i(x) \le 0, \quad i = 1, \dots, m$
$Ax = b$

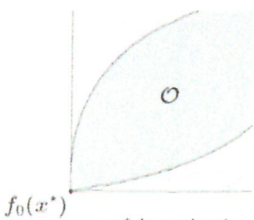
with f_0 K-convex, $f_1, \dots, f_m$ convex

Optimal and Pareto optimal points

set of achievable objective values

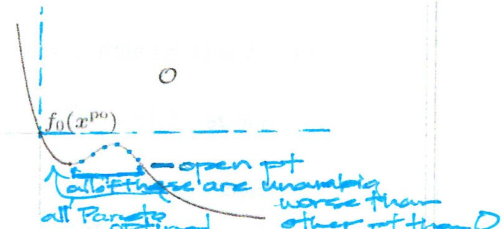
$$\mathcal{O} = \{f_0(x) \mid x \text{ feasible}\}$$

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

for every $y \in \mathcal{O}$ ~~$\neq$~~, $x^\star \preceq_K y$



x^{po} is Pareto optimal

no $y \in \mathcal{O}$ s.t. $y \preceq_K x^{\text{po}}$,

~~or if $y \preceq_K x^{\text{po}} \rightarrow y =$~~

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

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

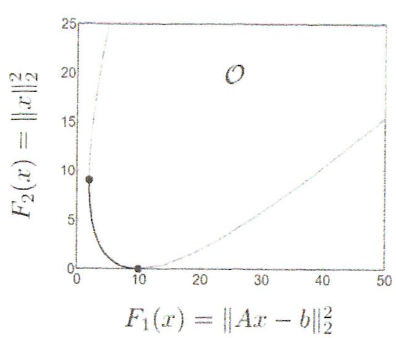
  if there exists an optimal point, the objectives are noncompeting
- feasible x^{po} is Pareto optimal if

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

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

Regularized least-squares

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



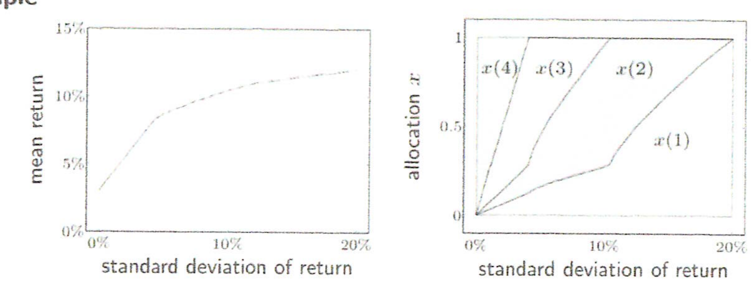
example for $A \in \mathbf{R}^{100 \times 10}$; heavy line is formed by Pareto optimal points

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
- $p \in \mathbf{R}^n$ is vector of relative asset price changes; 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

example



Scalarization

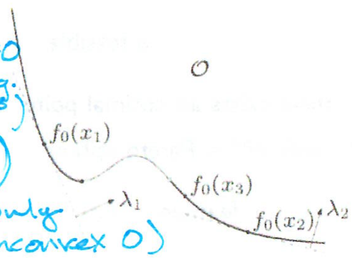
to find Pareto optimal points: choose $\lambda \succ_{K^*} 0$ and solve scalar problem

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

- changing λ changes slope of dashed line below
  - shift hyperplane until it supports $\mathcal{O}$ at some $x \in \mathcal{O}$
  - some x^{po} can't be found (e.g. x^3)

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

(issue only for nonconvex $\mathcal{O}$)



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

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

$\lambda \succeq 0$

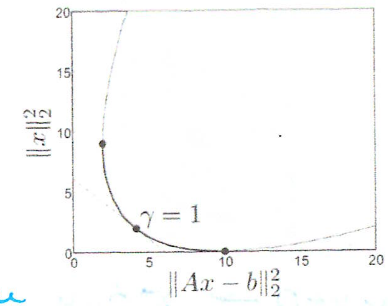
examples

- regularized least-squares problem of page 1–43

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

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

for fixed γ, a LS problem



- weights (i.e., λ) are a joystick that moves you along Pareto optimal surface.

- soft irritation func. like an objec. func.
- hard irritation func. like a constraint.

- risk-return trade-off of page 1–44

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

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