3. Convex functions

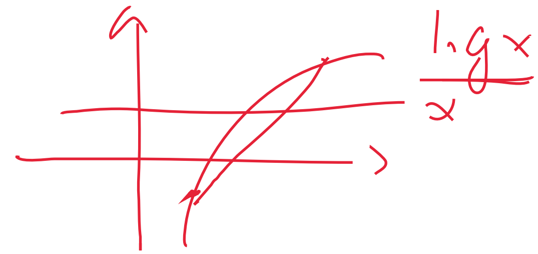
- basic properties and examples
- operations that preserve convexity
- the conjugate function
- quasiconvex functions
- log-concave and log-convex functions
- convexity with respect to generalized inequalities

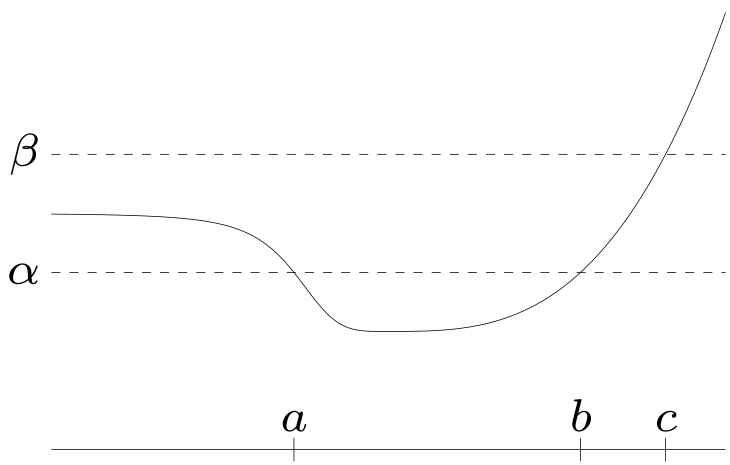
Quasiconvex functions

$f : \mathbf{R}^n \to \mathbf{R}$ is quasiconvex if $\mathbf{dom}\, f$ is convex and the sublevel sets

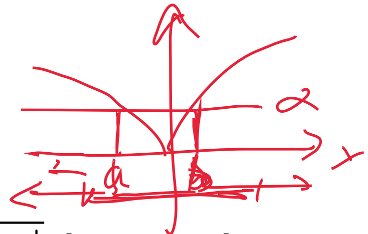
$$S_\alpha = \{x \in \mathbf{dom}\, f \mid f(x) \le \alpha\}$$

are convex for all α

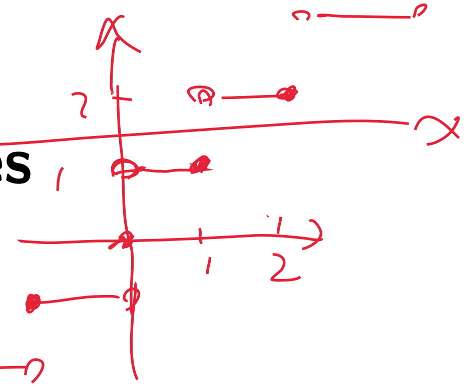




- f is quasiconcave if $-f$ is quasiconvex
- f is quasilinear if it is quasiconvex and quasiconcave



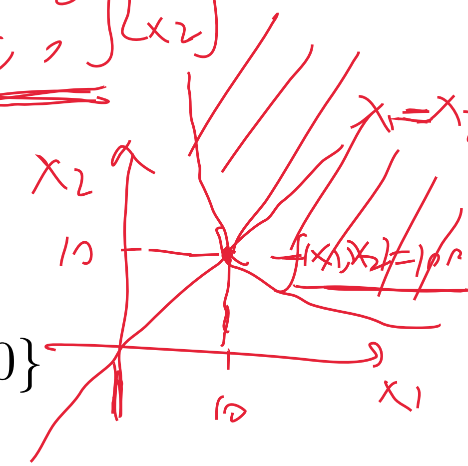
Examples



- $\sqrt{|x|}$ is quasiconvex on **R**
- $\mathrm{ceil}(x) = \inf\{z \in \mathbf{Z} \mid z \geq x\}$ is quasilinear
- $\log x$ is quasilinear on $\mathbf{R}_{++}$
- $f(x_1, x_2) = x_1 x_2$ is quasiconcave on $\mathbf{R}^2_{++}$

indefinite

$$f(x_1, x_2) = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}^T \begin{bmatrix} 0 & \frac{1}{2} \\ \frac{1}{2} & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$



- linear-fractional function

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

is quasilinear

$$\frac{a^T x + b}{c^T x + d} \gtrless \alpha \Rightarrow \alpha^T x + b \leq (c^T x + d)\alpha$$

- distance ratio

$$f(x) = \frac{\|x - a\|_2}{\|x - b\|_2}, \leq \alpha \quad \leq 1 \quad \mathbf{dom}\, f = \{x \mid \|x - a\|_2 \leq \|x - b\|_2\}$$

is quasiconvex

$$\|x - a\|_2 \leq \alpha \|x - b\|_2$$

f: convex

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

Properties

$\le \max\{f(x), f(y)\}$

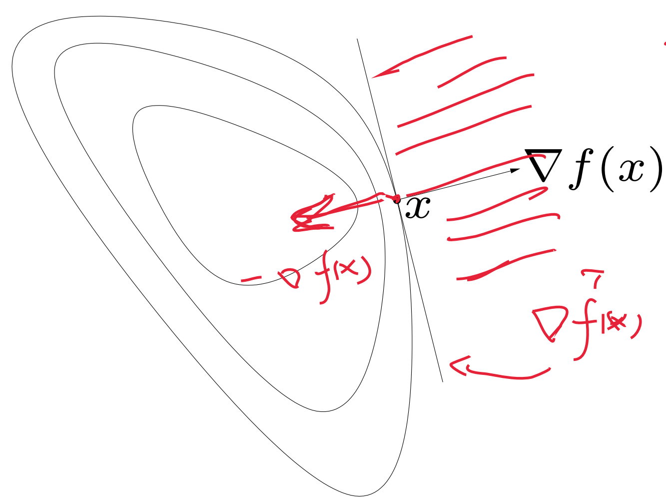
modified Jensen inequality: for quasiconvex f

$$0 \le \theta \le 1 \quad \Longrightarrow \quad f(\theta x + (1-\theta)y) \le \max\{f(x), f(y)\}$$

first-order condition: differentiable f with cvx domain is quasiconvex iff

$$f(y) \le f(x) \quad \Longrightarrow \quad \nabla f(x)^T(y-x) \le 0$$

$\{y \mid f(y) \le f(x)\}$



$\nabla f(x)^T(y-x) = 0$

sums of quasiconvex functions are not necessarily quasiconvex

$\log f(\theta x + (1-\theta) f(y)) \geq \theta \log f(x) + (1-\theta) \log f(y)$

Log-concave and log-convex functions

max log-likelihood (aka MLE)

a positive function f is log-concave if $\log f$ is concave:

$$f(\theta x + (1-\theta)y) \geq f(x)^\theta f(y)^{1-\theta} \quad \text{for } 0 \leq \theta \leq 1$$

f is log-convex if $\log f$ is convex

$f(x) = x^a$ $\qquad \log f(x) = a \log x$

- powers: x^a on $\mathbf{R}_{++}$ is log-convex for $a \leq 0$, log-concave for $a \geq 0$
- many common probability densities are log-concave, *e.g.*, normal:

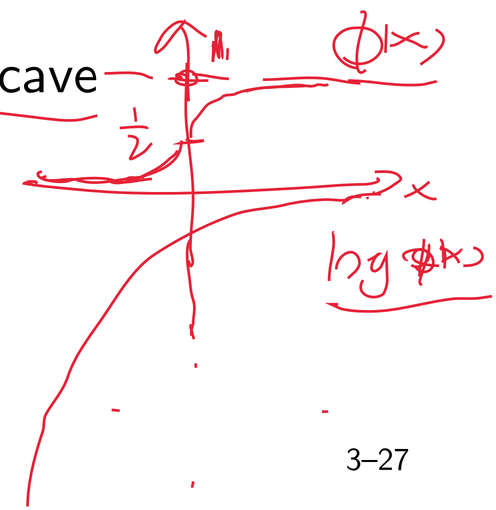
Quiz

$$f(x) = \frac{1}{\sqrt{(2\pi)^n \det \Sigma}} e^{-\frac{1}{2}(x-\bar{x})^T \Sigma^{-1}(x-\bar{x})}$$

$(\Sigma \in S^n_{++})$

- cumulative Gaussian distribution function Φ is log-concave

$$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-u^2/2}\, du$$



$\log f(x)$: concave

$\nabla \log f(x) = \frac{1}{f(x)} \nabla f(x)$

$\nabla^2 \log f(x) = \frac{1}{f(x)} \nabla^2 f(x) - \frac{1}{f(x)^2} \nabla f(x) \nabla f(x)^T \preceq 0$

$f : R^n \to R$

Properties of log-concave functions

- twice differentiable f with convex domain is log-concave if and only if

$$f(x)\nabla^2 f(x) \preceq \nabla f(x)\nabla f(x)^T$$

for all $x \in \mathbf{dom}\, f$

- product of log-concave functions is log-concave

$\log(f(x) \cdot g(x)) = \log f(x) + \log g(x)$

- sum of log-concave functions is not always log-concave
- integration: if $f : \mathbf{R}^n \times \mathbf{R}^m \to \mathbf{R}$ is log-concave, then

$$g(x) = \int f(x, y)\, dy$$

is log-concave (not easy to show)

consequences of integration property

- convolution $f * g$ of log-concave functions f, g is log-concave

$$(f * g)(x) = \int f(x-y)g(y)dy$$

- if $C \subseteq \mathbf{R}^n$ convex and y is a random variable with log-concave pdf then

$$f(x) = \mathbf{prob}(x+y \in C)$$

is log-concave

proof: write $f(x)$ as integral of product of log-concave functions

$$f(x) = \int g(x+y)p(y)\,dy, \qquad g(u) = \begin{cases} 1 & u \in C \\ 0 & u \notin C, \end{cases}$$

p is pdf of y

log g

$-\infty$ $u \in C$

example: yield function

$$Y(x) = \mathbf{prob}(x + w \in S)$$

- $x \in \mathbf{R}^n$: nominal parameter values for product
- $w \in \mathbf{R}^n$: random variations of parameters in manufactured product
- S: set of acceptable values

f(x)

log f(x)

if S is convex and w has a log-concave pdf, then

log-concave ⇒ quasi-concave

- Y is log-concave
- yield regions $\{x \mid Y(x) \geq \alpha\}$ are convex

(log: monotone)

Convexity with respect to generalized inequalities

proper cone

convex cone

1. solid
2. closed
3. pointed

R^n_+ , S^n_+

$f : \mathbf{R}^n \to \mathbf{R}^m$ is K-convex if $\mathbf{dom}\, f$ is convex and

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

for x, $y \in \mathbf{dom}\, f$, $0 \le \theta \le 1$

example $f : \mathbf{S}^m \to \mathbf{S}^m$, $f(X) = X^2$ is $\mathbf{S}^m_+$-convex $(\theta X + (1-\theta)Y)^2 \preceq_{S^n_+} \theta X^2 + (1-\theta)Y^2$

proof: for fixed $z \in \mathbf{R}^m$, $z^T X^2 z = \|Xz\|_2^2$ is convex in X, *i.e.*,

$$z^T(\theta X + (1-\theta)Y)^2 z \le \theta z^T X^2 z + (1-\theta) z^T Y^2 z$$

$z^T(\theta X^2 + (1-\theta)Y^2)z$

for $X, Y \in \mathbf{S}^m$, $0 \le \theta \le 1$

$(\theta X^2 + (1-\theta)Y^2) - (\theta X + (1-\theta)Y)^2 \in S^n_+$

therefore $(\theta X + (1-\theta)Y)^2 \preceq \theta X^2 + (1-\theta)Y^2$

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

min

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

s.t.

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

optimal value:

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

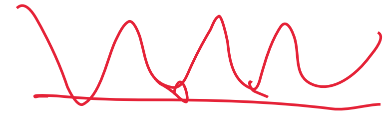
- $p^\star = \infty$ if problem is infeasible (no x satisfies the constraints)

$\inf \phi = \infty$

- $p^\star = -\infty$ if problem is unbounded below

Optimal and locally optimal points

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



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

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

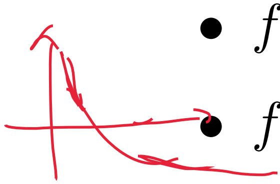
$$
\begin{array}{ll}
\text{minimize (over } z) & f_0(z) \\
\text{subject to} & f_i(z) \le 0, \quad i = 1, \ldots, m, \quad h_i(z) = 0, \quad i = 1, \ldots, p \\
& \|z - x\|_2 \le R
\end{array}
$$

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

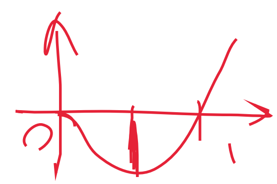
$f_0(x) = p^*$

- $f_0(x) = 1/x$, $\textbf{dom}\, f_0 = \mathbf{R}_{++}$: $p^\star = 0$, no optimal point

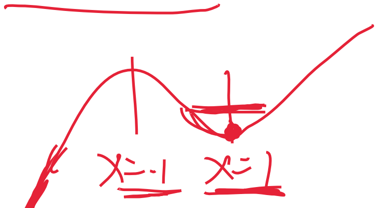
$p^* = \inf\{f_0(x) \mid \text{feasible}\}$



- $f_0(x) = -\log x$, $\textbf{dom}\, f_0 = \mathbf{R}_{++}$: $p^\star = -\infty$
- $f_0(x) = x \log x$, $\textbf{dom}\, f_0 = \mathbf{R}_{++}$: $p^\star = -1/e$, $x = 1/e$ is optimal



- $f_0(x) = x^3 - 3x$, $p^\star = -\infty$, local optimum at $x = 1$



f : convex

$f(Ax+b)$ = convex

Implicit constraints

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

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

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

$-\log x$: convex

example:

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

convex

is an unconstrained problem with implicit constraints $a_i^T x < b_i$ $\forall i \in \{1, 2, \ldots, k\}$

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

Convex optimization problem

standard form convex optimization problem

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

convex

feasible x: convex

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

$\in S^n_{++}$

example

$f_0(x) = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}^T \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = x^T I x$

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

- 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

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

Local and global optima

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

proof: suppose x is locally optimal, but there exists a feasible y with $f_0(y) < f_0(x)$

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

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

consider $z = \theta y + (1 - \theta)x$ with $\theta = R/(2\|y - x\|_2)$

$\|z - x\|_2 = \|\theta y - \theta x\|_2 = \theta \|y - x\|_2 = \frac{R}{2}$

- $\|y - x\|_2 > R$, so $0 < \theta < 1/2$
- z is a convex combination of two feasible points, hence also feasible
- $\|z - x\|_2 = R/2$ and

$, < \theta f_0(x)$

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

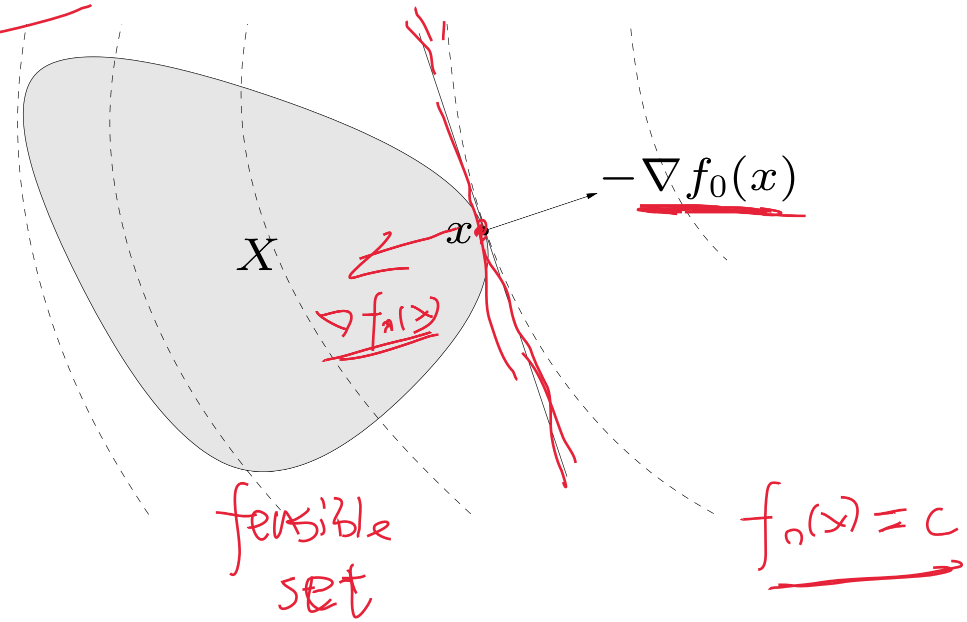
which contradicts our assumption that x is locally optimal

Optimality criterion for differentiable f_0

x is optimal if and only if it is feasible and

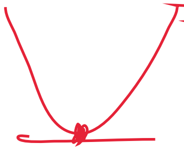
$$\nabla f_0(x)^T(y-x) \geq 0 \quad \text{for all feasible } y$$

x : optimal $\iff$ x : feasible, $\nabla_0 f(x)^T(y-x) \geq 0$, $\forall y$: feasible



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

$\langle Aa, b\rangle = \langle a, A^T b\rangle$



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

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

$\nabla f(x)^T (y-x) \geq 0$

$y, x \in R^n$

$z = y - x \in R^n$

$\nabla f(x)^T z \geq 0$

- **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

Lagrange multiplier

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

$x: Ax = b, \quad \nabla f(x)^T (y-x) \geq 0, \quad \forall\, y: Ay = b$

$v = x - y$

$Av = A(x-y) = b - b = 0$

$v \in N(A)$: null space

$\nabla f(x)^T v \geq 0$ $\qquad$ $\nabla f(x)^T v = 0$

$\nabla f(x)^T (-v) \geq 0$ $\qquad$ $\nabla f(x)^T v \leq 0$

$\nabla f(x) \perp N(A)$

$\nabla f(x) \in R(A^T)$: Range space

$\nabla f(x) = A^T w$

$R(P)$

$v \in N(A) \iff Av = 0$

$\iff \langle Av, w\rangle = 0 \quad \forall w$

$\iff \langle v, A^T w\rangle = 0 \quad \forall w$

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

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

- **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

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

$x \geq 0$, $\nabla f(\bar{x})(y - x) \geq 0$, $\forall y \geq 0$

- **minimization over nonnegative orthant**

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

$\nabla f(x)^T x \leq 0$

$\nabla f(x) \geq 0$ $\quad$ $x \geq 0$

$\Rightarrow$ $\nabla f(x)^T x = 0$

  x is optimal if and only if

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

complementary

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

$F \in N(A)$
$x_0 : Ax = b$

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

where F and x_0 are such that

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

$Ax = b \Rightarrow A(Fz + x_0) = 0 + b = b$

- **introducing equality constraints**

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

is equivalent to

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

- **introducing slack variables for linear inequalities**

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

is equivalent to

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

$\min_x \|Ax-b\|_\infty$

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

$\min_{x,t}\ t$

s.t. $\|Ax-b\|_\infty \le t$

$|a_i^T x - b_i| \le t \quad \forall i$

epigraph trick

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

- **minimizing over some variables**

$g(x_1) = \inf_{x_2} f(x_1, x_2)$

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

is equivalent to

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

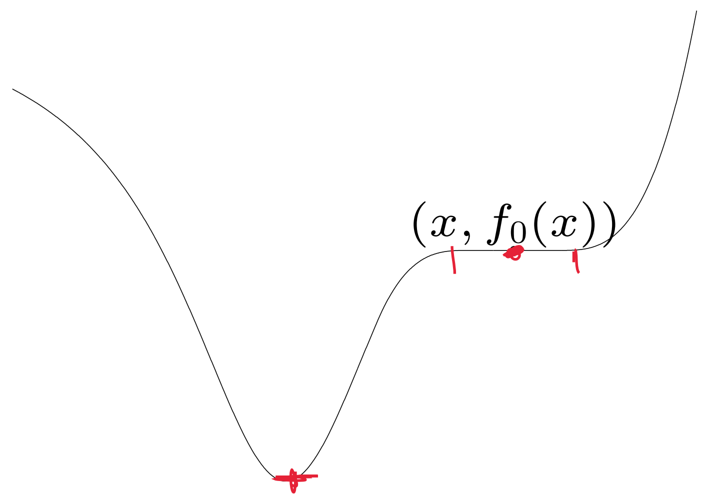
where $\boxed{\tilde{f}_0(x_1) = \inf_{x_2} f_0(x_1, x_2)}$

Quasiconvex optimization

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

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

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



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

$$\frac{p(x)}{q(x)} \le t \quad (t \ge 0)$$

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

$$\Rightarrow \underbrace{p(x) - t\,q(x) \le 0}_{\text{convex}}$$

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 $\phi_t(x) \le 0$

$$\min_x \; f_0(x)$$

$$\Downarrow$$

epigraph trick

$$\min_{x,t} \; t$$

$$\text{s.t.} \quad \phi_t(x) \le 0$$

quasiconvex optimization via convex feasibility problems

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

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

$$\boxed{\begin{array}{ll} \min_{x,t} & t \\ \text{s.t.} & \phi_t(x) \leq 0 \\ & f_i(x) \leq 0, \ \forall i \\ & Ax = b \end{array}}$$

$p^\star \in [l, u]$

Bisection method for quasiconvex optimization

given $l \leq p^\star$, $u \geq p^\star$, tolerance $\epsilon > 0$.
repeat
1. $t := (l + u)/2$.
2. Solve the convex feasibility problem (1).
3. **if** (1) is feasible, $u := t$; **else** $l := t$.

until $u - l \leq \epsilon$.

---

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

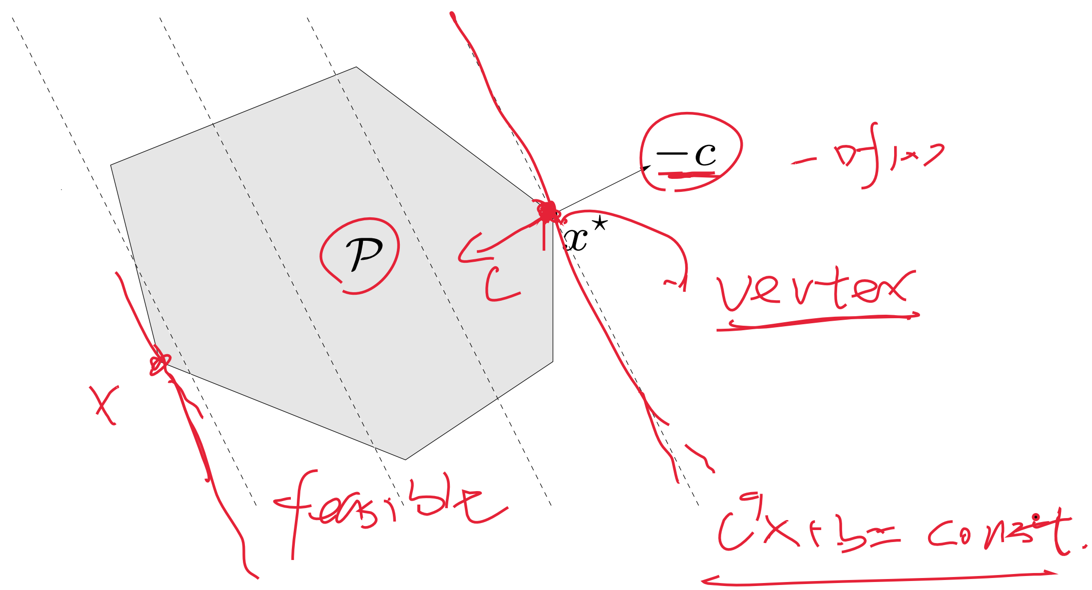
Linear program (LP)

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

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

$\nabla^T f(x)(y-x) \geq 0$

$\forall y$



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

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

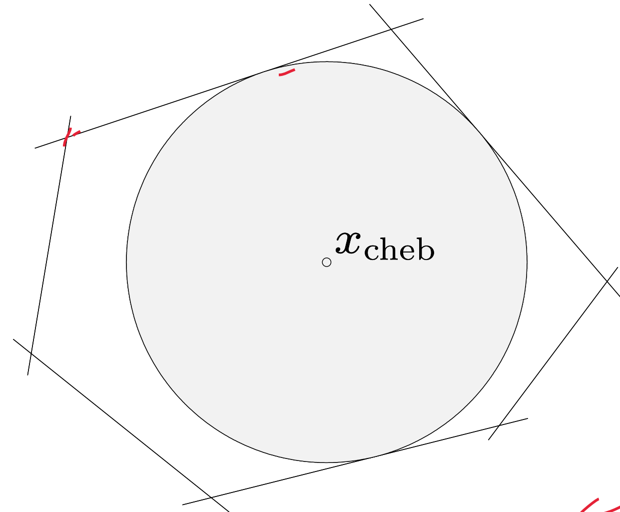
equivalent to an LP

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

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

$$\mathcal{B} = \{x_c + u \mid \|u\|_2 \le r\}$$

$r \cdot e$

$\sup_{\|u\|_2 \le r} (a_i^T u) = r\|a_i\|_2$

- $a_i^T x \le b_i$ for all $x \in \mathcal{B}$ if and only if

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

$a_i^T x_c + a_i^T u$

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

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

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)

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

generalized linear-fractional program

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

a quasiconvex optimization problem; can be solved by bisection

example: Von Neumann model of a growing economy

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

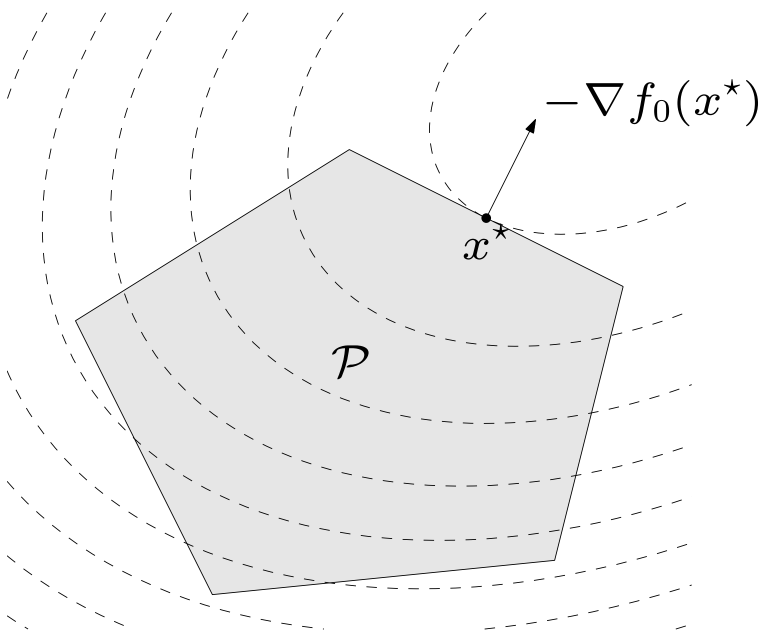
- $x, x^+ \in \mathbf{R}^n$: activity levels of n sectors, in current and next period
- $(Ax)_i$, $(Bx^+)_i$: produced, resp. consumed, amounts of good i
- x_i^+/x_i: growth rate of sector i

allocate activity to maximize growth rate of slowest growing sector

Quadratic program (QP)

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

- $P \in \mathbf{S}^n_+$, 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

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

- 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)

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 region is intersection of m ellipsoids and an affine set

Second-order cone programming

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

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

- 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 $\mathcal{E}_i$:

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

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

- robust LP

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

is equivalent to the SOCP

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

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

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

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

- robust LP

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

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

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

Geometric programming

monomial function

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

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

posynomial function: sum of monomials

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

geometric program (GP)

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

with f_i posynomial, h_i monomial

Geometric program in convex form

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

- monomial $f(x) = cx_1^{a_1} \cdots x_n^{a_n}$ transforms to

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

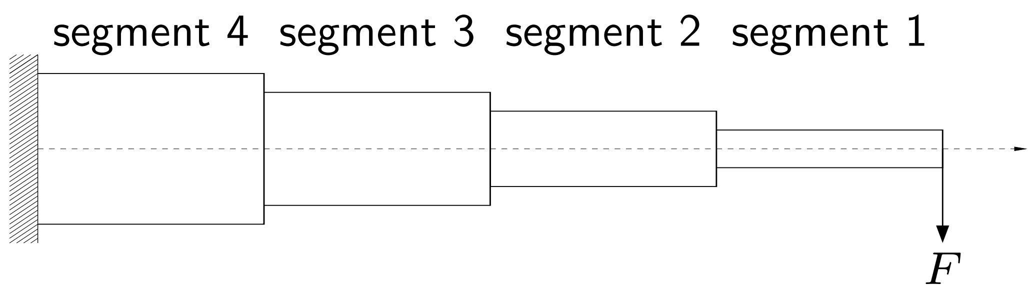
- posynomial $f(x) = \sum_{k=1}^K c_k x_1^{a_{1k}} x_2^{a_{2k}} \cdots x_n^{a_{nk}}$ transforms to

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

- geometric program transforms to convex problem

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

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
- 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

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