10. Special convex optimization problems

- linear programs
- piecewise-linear minimization
- quadratic optimization
- geometric programming
- semidefinite programs
- quasiconvex optimization

Linear program

a *linear program* (LP) is an optimization problem of the form

$$\begin{array}{rl} \text{minimize (or maximize)} & \displaystyle\sum_{j=1}^{n} c_j x_j \\ \text{subject to} & \displaystyle\sum_{j=1}^{n} a_{ij} x_j \leq b_i, \quad i = 1, \ldots, m \\ & \displaystyle\sum_{j=1}^{n} g_{ij} x_j = h_i, \quad i = 1, \ldots, p \end{array}$$

- n optimization variables $x_1, \ldots, x_n$
- coefficients $c_j, a_{ij}, g_{ij}, h_i, b_i$ are given
- convex problem with linear objective and linear/affine constraints

LP in compact form

$$\begin{array}{rl} \text{minimize (or maximize)} & c^T x \\ \text{subject to} & Ax \leq b \\ & Gx = h \end{array}$$

- A is an $m \times n$ matrix with entries a_{ij}
- G is an $p \times n$ matrix with entries g_{ij}
- $b = (b_1, \ldots, b_m)$
- $h = (h_1, \ldots, h_p)$
- $c = (c_1, \ldots, c_n)$

Example: diet problem

- create meal with at least 12 units of protein, 9 units of iron, 15 units of thiamine

food	protein	iron	thiamine	cost (cents/g)
A	2 unit/g	1 unit/g	1 unit/g	30
B	1 unit/g	1 unit/g	3 unit/g	40

- how many grams of each food should be used to minimize the cost of the meal?

the problem can formulated as

$$\begin{array}{ll} \text{minimize} & 30x_1 + 40x_2 \\ \text{subject to} & 2x_1 + x_2 \geq 12 \\ & x_1 + x_2 \geq 9 \\ & x_1 + 3x_2 \geq 15 \\ & x_1, x_2 \geq 0 \end{array}$$

where x_1 and x_2 are the number of grams of food A and B used in the meal

Example: alloy mixture

- we are given four alloys that have the metal properties listed in the below table

property	alloy 1	alloy 2	alloy 3	alloy 4
% of iron	70	25	40	20
% of nickel	10	15	50	50
% of cobalt	20	60	10	30
cost ($/kg)	22	18	25	24

- goal is to create new alloy mixture with 40% iron, 35% nickel, 25% cobalt

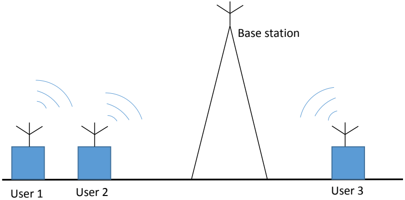
- what proportions of the alloys should be blended together while minimizing cost?

- let x_i be the proportion of alloy i that is used to produce the new alloy

- the problem can be formulated as

$$
\begin{array}{ll}
\text{minimize} & 22x_1 + 18x_2 + 25x_3 + 24x_4 \\
\text{subject to} & 0.7x_1 + 0.25x_2 + 0.4x_3 + 0.2x_4 = 0.4 \\
& 0.1x_1 + 0.15x_2 + 0.5x_3 + 0.5x_4 = 0.35 \\
& 0.2x_1 + 0.6x_2 + 0.1x_3 + 0.3x_4 = 0.25 \\
& x_1 + x_2 + x_3 + x_4 = 1 \\
& x_1, x_2, x_3, x_4 \geq 0
\end{array}
$$

Example: wireless communication



- n “mobile” users
- user i transmits signal to base station with power p_i and attenuation factor of β_i
  - signal power received at the base station from user i is $\beta_i p_i$
- total power received from all other users is considered interference
  - the interference for user i is $\sum_{j \neq i} \beta_j p_j$
- for reliable communication with user i, signal-to-interference ratio must exceed γ_i
- goal is to minimize total power transmitted by all users subject to having reliable communications for all users

Problem formulation

$$
\begin{array}{ll}
\text{minimize} & \sum_{i=1}^n p_i \\
\text{subject to} & \dfrac{\beta_i p_i}{\sum_{j \neq i} \beta_j p_j} \geq \gamma_i, \quad i = 1, \ldots, n \\
& p_i \geq 0, \quad i = 1, \ldots, n
\end{array}
$$

LP formulation

$$
\begin{array}{ll}
\text{minimize} & \sum_{i=1}^n p_i \\
\text{subject to} & \beta_i p_i - \gamma_i \sum_{j \neq i} \beta_j p_j \geq 0, \quad i = 1, \ldots, n \\
& p_i \geq 0, \quad i = 1, \ldots, n
\end{array}
$$

Example: assignment problem

- we want to match N people to N tasks
- each person is assigned to one task (each task assigned to one person)
- cost of assigning person i to task j is c_{ij}
- variable $x_{ij} = 1$ if person i is assigned to task j; $x_{ij} = 0$ otherwise

Combinatorial formulation

$$
\begin{array}{ll}
\text{minimize} & \sum_{i=1}^{N} \sum_{j=1}^{N} c_{ij} x_{ij} \\
\text{subject to} & \sum_{i=1}^{N} x_{ij} = 1, \quad j = 1, \ldots, N \\
& \sum_{j=1}^{N} x_{ij} = 1, \quad i = 1, \ldots, N \\
& x_{ij} \in \{0, 1\}, \quad i, j = 1, \ldots, N
\end{array}
$$

$N!$ possible assignments (*e.g.*, $10! = 3628800$)

LP formulation

$$
\begin{array}{ll}
\text{minimize} & \sum_{i=1}^{N} \sum_{j=1}^{N} c_{ij} x_{ij} \\
\text{subject to} & \sum_{i=1}^{N} x_{ij} = 1, \quad j = 1, \ldots, N \\
& \sum_{j=1}^{N} x_{ij} = 1, \quad i = 1, \ldots, N \\
& 0 \leq x_{ij} \leq 1, \quad i, j = 1, \ldots, N
\end{array}
$$

- we have *relaxed* the constraints $x_{ij} \in \{0, 1\}$
- it can be shown that the solution $x_{ij}^{\star} \in \{0, 1\}$
- hence, we can solve this hard combinatorial problem efficiently by solving an LP

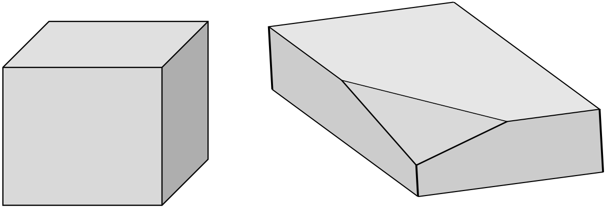
Polyhedron

a *ployhedron* is the intersection of finitely many halfspaces

$$a_1^T x \leq b_1, \ldots, a_m^T x \leq b_m$$

in matrix notation, a polyhedron can be defined as

$$\mathcal{P} = \{x \in \mathbb{R}^n \mid Ax \leq b\}$$

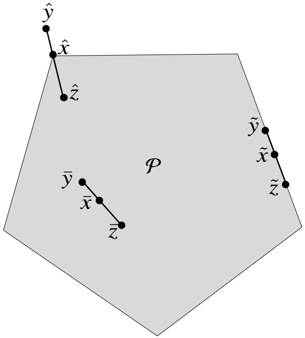


Extreme points

$x \in \mathcal{P}$ is an *extreme point* of $\mathcal{P}$ if it *cannot* be written as convex combination

$$x = \theta y + (1 - \theta)z, \quad \theta \in (0, 1)$$

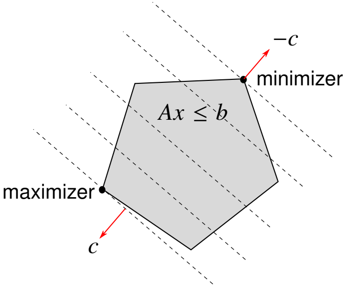
for some $y, z \in \mathcal{P}$



- $\hat{x}$ is an extreme point
- $\bar{x}$ and $\tilde{x}$ are not extreme points

Geometrical interpretation of LP

$$\begin{array}{rl} \text{minimize (or maximize)} & c^T x \\ \text{subject to} & Ax \leq b \end{array}$$



- dashed lines are level sets $c^T x = \alpha$ for different α
- feasible set is a polyhedron
- the optimal solutions occur at an extreme point

Outline

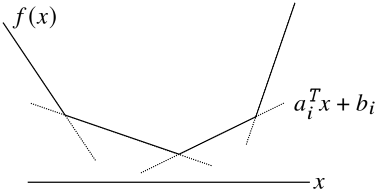
- linear programs
- **piecewise-linear minimization**
- quadratic optimization
- geometric programming
- semidefinite programs
- quasiconvex optimization

Piecewise-linear minimization

Piecewise-linear function

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

- $a_i \in \mathbb{R}^n$ and $b_i \in \mathbb{R}$
- piecewise-linear function is a pointwise maximum of affine functions



Piecewise-linear minimization

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

Equivalent LP formulation

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

- with additional variable $t \in \mathbb{R}$
- for fixed x, the optimal t is $t = f(x)$

Matrix form

$$\begin{array}{ll} \text{minimize} & \tilde{c}^T \tilde{x} \\ \text{subject to} & \tilde{A}\tilde{x} \le \tilde{b} \end{array}$$

where

$$\tilde{x} = \begin{bmatrix} x \\ t \end{bmatrix}, \quad \tilde{c} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \quad \tilde{A} = \begin{bmatrix} a_1^T & -1 \\ \vdots & \vdots \\ a_m^T & -1 \end{bmatrix}, \quad \tilde{b} = \begin{bmatrix} -b_1 \\ \vdots \\ -b_m \end{bmatrix}$$

ℓ_1-Norm approximation

$$\text{minimize} \quad \|Ax - b\|_1$$

- $A \in \mathbb{R}^{m \times n}$ and $b \in \mathbb{R}^m$
- for a vector $y \in \mathbb{R}^m$, we have

$$\|y\|_1 = \sum_{i=1}^{m} |y_i| = \sum_{i=1}^{m} \max\{y_i, -y_i\}$$

Equivalent LP formulation

$$\begin{array}{ll} \text{minimize} & \sum_{i=1}^{m} u_i \\ \text{subject to} & -u \leq Ax - b \leq u \end{array}$$

with variables $x \in \mathbb{R}^n$ and $u \in \mathbb{R}^m$

Robust curve fitting

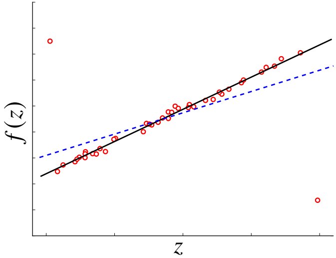
fit data points (z_i, y_i) to the straight line $x_1 + x_2 z \approx y$ using ℓ_1-norm:

$$\text{minimize} \quad \|Ax - b\|_1$$

where

$$A = \begin{bmatrix} 1 & z_1 \\ \vdots & \vdots \\ 1 & z_m \end{bmatrix}, \quad b = \begin{bmatrix} y_1 \\ \vdots \\ y_m \end{bmatrix}, \quad x = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

- red circles represent the data
- blue dotted line from minimizing $\|Ax - b\|^2$
- black line from minimizing $\|Ax - b\|_1$
- ℓ_1-norm more robust to outliers



Interview scheduling

- a company needs to schedule job interviews for n candidates $(1, 2, \ldots, n)$
- candidate i is scheduled to be the ith interview
- the starting time of candidate i must be in the interval $[\alpha_i, \beta_i]$, where $\alpha_i < \beta_i$
- goal is to find n starting times of interviews so that the minimal starting time difference between consecutive interviews is maximal

- let t_i denote the starting time of interview i
- the objective function is the minimal difference between consecutive starting times:

$$f(t) = \min\{t_2 - t_1, t_3 - t_2, \ldots, t_n - t_{n-1}\},$$

Problem formulation

$$\begin{array}{ll} \text{maximize} & \min\{t_2 - t_1, t_3 - t_2, \ldots, t_n - t_{n-1}\} \\ \text{subject to} & \alpha_i \leq t_i \leq \beta_i, \quad i = 1, 2, \ldots, n, \end{array}$$

with variable $t \in \mathbb{R}^n$

Equivalent LP

$$\begin{array}{ll} \text{maximize} & s \\ \text{subject to} & t_{i+1} - t_i \geq s, \quad i = 1, 2, \ldots, n-1 \\ & \alpha_i \leq t_i \leq \beta_i, \quad i = 1, 2, \ldots, n, \end{array}$$

with variables $\boldsymbol{t} \in \mathbb{R}^n$ and $s \in \mathbb{R}$

Outline

- linear programs
- piecewise-linear minimization
- **quadratic optimization**
- geometric programming
- semidefinite programs
- quasiconvex optimization

Quadratic optimization

Quadratic program (quadratic optimization problem)

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

- $Q \in \mathbb{S}^n_+$, so objective is convex quadratic
- $r \in \mathbb{R}^n$, $A \in \mathbb{R}^{m \times n}$, $G \in \mathbb{R}^{p \times n}$, $h \in \mathbb{R}^p$, and $b \in \mathbb{R}^m$
- minimize a convex quadratic function over a polyhedron

Quadratically constrained quadratic problem (QCQP)

$$
\begin{array}{ll}
\text{minimize} & (1/2)x^T Q_0 x + r_0^T x + s_0 \\
\text{subject to} & (1/2)x^T Q_i x + r_i^T x \leq 0, \quad i = 1, \ldots, p \\
& Ax = b
\end{array}
$$

- $Q_i \in \mathbb{S}^n_+$ ($i = 0, 1 \ldots, m$) are positive semidefinite
- feasible set is intersection of n ellipsoids and an affine set

Examples

Least squares

$$\text{minimize} \quad \|Ax - b\|^2 = x^T A^T A x - 2b^T A x + b^T b$$

Constrained least squares

$$\begin{array}{ll} \text{minimize} & \|Ax - b\|^2 \\ \text{subject to} & Gx = h \\ & l_i \le x_i \le u_i, \quad i = 1, \ldots, n \end{array}$$

this problem has no simple analytical solution

Example: power distribution (aggregator model)

- in electricity markets, an aggregator
  - buys wholesale p units of power (Megawatt) from power distribution utilities
  - and resells this power to a group of n business or industrial customers
- the ith customer, $i = 1, \ldots, n$, would ideally wants p_i Megawatts
- the customer i does not want to receive more or less power than needed
- the customer dissatisfaction can be modeled as

$$f_i(x_i) = c_i(x_i - p_i)^2, \quad i = 1, \ldots, n$$

  x_i is power given to customer i; c_i is a given customer parameter

- the aggregator problem is finding the power allocations x_i, $i = 1, \ldots, n$, such that
  - the average customer dissatisfaction is minimized,
  - the whole power p is sold,
  - and that the dissatisfaction level is no greater than a contract level, say d

- the aggregator problem is

$$
\begin{array}{ll}
\text{minimize} & \displaystyle\frac{1}{n}\sum_{i=1}^{n} c_i(x_i - p_i)^2 \\
\text{subject to} & \displaystyle\sum_{i=1}^{n} x_i = p, \\
& c_i(x_i - p_i)^2 \leq d, \quad i = 1, \ldots, n \\
& x_i \geq 0, \quad i = 1, \ldots, n
\end{array}
$$

this is a QCQP

Example: portfolio optimization

we want to invest on n stocks to achieve a good return while minimizing risks of losses

- let $x_i \geq 0$ be the proportion of investment on stock i
- let r_i be the return for stock i; we assume that the expected returns are known,

$$\mu_j = \mathbb{E}(r_j), \quad j = 1, 2, \ldots, n,$$

and that the covariances of all the pairs of variables are also known,

$$\sigma_{ij}^2 = \mathbb{E}\left[(r_i - \mu_i)(r_j - \mu_j)\right], \quad i, j = 1, 2, \ldots, n$$

(typically, the mean and variance are estimated from historical data)

- a high variance indicates high risk; a low variance indicates low risk
- positive covariance $\sigma_{ij}^2 > 0$ means stocks i and j prices move in the same direction
- a negative $\sigma_{ij}^2 < 0$ means they one change in opposite direction

- the overall return is the random variable

$$R = \sum_{j=1}^{n} x_j r_j$$

whose expectation and variance are given by

$$\mathbb{E}(R) = \mu^T x, \quad \mathsf{Var}(R) = x^T \Sigma x$$

- $\mu = (\mu_1, \mu_2, \ldots, \mu_n)$
- Σ is the covariance matrix whose elements are $\Sigma_{ij} = \sigma_{ij}$
- the covariance matrix is always positive semidefinite

Portfolio problem QP formulation:

$$\begin{array}{ll} \text{minimize} & x^T \Sigma x \\ \text{subject to} & \mu^T x \geq \alpha \\ & \mathbf{1}^T x = 1 \\ & x \geq 0 \end{array}$$

where α is the minimal return value

Portfolio problem QCQP formulation:

$$\begin{array}{ll} \text{maximize} & \mu^T x \\ \text{subject to} & x^T \Sigma x \leq \beta \\ & \mathbf{1}^T x = 1 \\ & x \geq 0 \end{array}$$

where β is the upper bound on the risk

Outline

- linear programs
- piecewise-linear minimization
- quadratic optimization
- **geometric programming**
- semidefinite programs
- quasiconvex optimization

Monomials and posynomials

Monomial function

$$f(x) = cx_1^{a_1} x_2^{a_2} \dots x_n^{a_n}, \quad \mathrm{dom}\, f = \mathbb{R}^n_{++}$$

$c > 0$ and each $a_i \in \mathbb{R}$ can be any number

Posynomial function: sum of monomials

$$f(x) = \sum_{k=1}^{K} c_k x_1^{a_{1k}} x_2^{a_{2k}} \dots x_n^{a_{nk}}, \quad \mathrm{dom}\, f = \mathbb{R}^n_{++}$$

each $c_k > 0$

Example

- wireless cellular network with n paired transmitters and receivers
- $p_1, \ldots, p_n$ are the transmit powers for these pairs
- each transmitter i is intended to communicate with its corresponding receiver i
- the signal to interference plus noise ratio (SINR) for each receiver is:

$$\gamma_i = \frac{S_i}{l_i + \sigma_i}, \quad i = 1, \ldots, n,$$

  - S_i represents the power of the desired signal received from transmitter i
  - l_i is the combined interference from all other transmitters
  - σ_i is the receiver's noise power

- the Rayleigh fading model suggests that the S_i is a linear function of $p_1, \ldots, p_n$:

$$S_i = G_{ii} p_i, \quad i = 1, \ldots, n,$$

and

$$l_i = \sum_{j \neq i} G_{ij} p_j,$$

where G_{ij} are the known path gains from transmitter j to receiver i

- therefore, the SINR expressions in terms of the powers $p_1, \ldots, p_n$ are:

$$\gamma_i(p) = \frac{G_{ii} p_i}{\sigma_i + \sum_{j \neq i} G_{ij} p_j}, \quad i = 1, \ldots, n,$$

- while the SINR functions aren't posynomials, their inverses are:

$$\gamma_i^{-1}(p) = \frac{\sigma_i}{G_{ii}} p_i^{-1} + \sum_{j \neq i} \frac{G_{ij}}{G_{ii}} p_j p_i^{-1}, \quad i = 1, \ldots, n$$

Generalized posynomials

a generalized posynomial is obtained from posynomials by various operations like

- addition
- multiplication
- pointwise maximum
- raising to a specific power

Example

$$f(x) = \max(2x_1^{2.3}x_2^7, x_1x_2x_3^{3.14}, \sqrt{x_1 + x_2^3})$$

this function qualifies as a generalized posynomial

Geometric program (GP)

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

- $f, g_1 \dots, g_m$ are posynomials
- $h_1, \dots, h_p$ are monomials
- its domain is inherently set as $\mathcal{D} = \mathbb{R}^n_{++}$ (implicit constraint $x > 0$)

Example

- consider the optimization problem:

$$\begin{array}{ll} \text{maximize} & x/y \\ \text{subject to} & 2 \le x \le 3 \\ & x^2 + 3y/z \le \sqrt{y} \\ & x/z = z^2 \end{array}$$

where $x, y, z \in \mathbb{R}$ and implicitly $x, y, z > 0$

- the problem can be recast into the standard GP form:

$$\begin{array}{ll} \text{minimize} & x^{-1}y \\ \text{subject to} & 2x^{-1} \le 1 \\ & (1/3)x \le 1 \\ & x^2y^{-1/2} + 3y^{1/2}z^{-1} \le 1 \\ & xy^{-1}z^{-2} = 1 \end{array}$$

Change of variable

- geometric programs are generally not convex optimization problems
- but, they can be recast into convex forms through suitable transformations

Change of variable: $y_i = \log x_i$ ($x_i = e^{y_i}$); take logarithm of cost, constraints

- monomial $f(x) = cx_1^{a_1} x_2^{a_2} \cdots x_n^{a_n}$ can be transformed to

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

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

$$f(y) = \sum_{k=1}^{K} e^{a_k^T y + \log c_k} \iff \log f(y) = \log(\sum_{k=1}^{K} e^{a_k^T y + b}), \quad (b_k = \log c_k)$$

with $a_k = (a_{1k}, \ldots, a_{nk})$

Geometric program in convex form

applying the logarithm to the objective/constraint functions results in

$$
\begin{array}{ll}
\text{minimize} & \bar{f}(y) = \log\left(\sum_{k=1}^{K_0} e^{a_{0k}^T y + b_{0k}}\right) \\
\text{subject to} & \bar{g}_i(y) = \log\left(\sum_{k=1}^{K_i} e^{a_{ik}^T y + b_{ik}}\right) \leq 0, \quad i = 1, \ldots, m \\
& \bar{h}_i(y) = h_i^T y + d_i = 0, \quad i = 1, \ldots, p
\end{array}
$$

- $\bar{f}$ and $\bar{g}_i$ functions are convex, and $\bar{h}_i$ functions are affine
- thus, this optimization problem is convex
- we call it *geometric program in convex form*
- the original form is called *geometric program in posynomial form*

Example

- consider a cylindrical liquid storage tank with height, h, and diameter, d
- unlike the main body of the tank, its base is made from a distinct material
- assume the height of base remains unchanged irrespective of tank's height
- V_{tank} is the volume of the tank
- V_{supp} is the volume supplied within a designated time frame
- total costs associated with manufacturing/operating the tank over a set duration (e.g., a year) is divided into
  - filling cost
  - construction cost
- goal is to minimize cost subject to some constraints

Filling costs

$$C_{\text{fill}}(d,h) = \alpha_1 \frac{V_{\text{supp}}}{V_{\text{tank}}} = c_1 h^{-1} d^{-2}$$

- α_1 is a positive constant (in dollars), and $c_1 = \frac{4\alpha_1 V_{\text{supp}}}{\pi}$
- tied to supplying a certain volume, V_{supp}, of a liquid within the time-frame
- $V_{\text{supp}}/V_{\text{tank}}$ determines the frequency of tank refilling; hence its cost
- as the volume of the tank diminishes relative to the supply volume, filling costs rise

Construction costs:

$$C_{\text{constr}}(d,h) = c_2 d^2 + c_3 dh,$$

- $c_2 = \alpha_2 \frac{\pi}{4}$ and $c_3 = \alpha_3 \pi$ (α_2, α_3 are +ve dollar-per-square-meter constants)
- include the expenses of constructing the tank's and its base
- the base's cost is proportional to its area, $\frac{\pi d^2}{4}$
- the tank's cost correlates with its surface area, πdh

Total cost

$$\begin{aligned} C_{\text{total}}(d,h) &= C_{\text{fill}}(d,h) + C_{\text{constr}}(d,h) \\ &= c_1 h^{-1} d^{-2} + c_2 d^2 + c_3 dh \end{aligned}$$

this posynomial objective function is subject to constraints such as upper and lower limits on the diameter and height, represented as:

$$0 < d \leq d_{\max}, \quad 0 < h \leq h_{\max}$$

GP formulation

$$\begin{array}{ll} \text{minimize} & c_1 h^{-1} d^{-2} + c_2 d^2 + c_3 dh \\ \text{subject to} & 0 < d_{\max}^{-1} d \leq 1 \\ & 0 < h_{\max}^{-1} h \leq 1 \end{array}$$

with variables d, h

Example: Frobenius norm diagonal scaling

we seek diagonal matrix $D = \mathrm{diag}(d)$, $d > 0$, with

$$\text{minimize} \quad \|DMD^{-1}\|_F^2$$

- express as

$$\|DMD^{-1}\|_F^2 = \sum_{i,j=1}^{n} (DMD^{-1})_{ij}^2 = \sum_{i,j=1}^{n} M_{ij}^2 d_i^2 / d_j^2$$

- a posynomial in d (with exponents 0, 2 , and -2)

- in convex form, with $y_i = \log d_i$,

$$\log \|DMD^{-1}\|_F^2 = \log \left(\sum_{i,j=1}^{n} \exp\left(2(y_i - y_j + \log |M_{ij}|)\right) \right)$$

Outline

- linear programs
- piecewise-linear minimization
- quadratic optimization
- geometric programming
- **semidefinite programs**
- quasiconvex optimization

Semidefinite program

a *linear matrix inequality* (LMI) constrains a vector of variables $x \in \mathbb{R}^n$ as

$$F(x) = F_0 + \sum_{i=1}^{m} x_i F_i \preceq 0 \tag{10.1}$$

with symmetric coefficient matrices $F_0, \ldots, F_n$ of size $m \times m$

a **semidefinite program** (SDP) is a particular type of convex optimization problem:

$$\begin{array}{ll} \text{minimize} & c^T x \\ \text{subject to} & F(x) = F_0 + \sum_{i=1}^{n} x_i F_i \preceq 0 \end{array} \tag{10.2}$$

- $x \in \mathbb{R}^n$ is the optimization variable and $c \in \mathbb{R}^n$
- each F_i is a known $m \times m$ symmetric matrices
- if $F_0, F_1, \ldots, F_m$ are diagonal matrices the SDP becomes a linear program

General form SDP

$$\begin{array}{ll} \text{minimize} & c^T x \\ \text{subject to} & F^{(i)}(x) = x_1 F_1^{(i)} + \cdots + x_n F_n^{(i)} + F_0^{(i)} \preceq 0, \quad i = 1, \ldots, K \\ & Gx \leq h \\ & Ax = b \end{array}$$

can be equivalently represented as an SDP

$$\begin{array}{ll} \text{minimize} & c^T x \\ \text{subject to} & \mathrm{diag}(Gx - h, F^{(1)}(x), \ldots, F^{(K)}(x)) \preceq 0 \\ & Ax = b \end{array}$$

Example: maximum eigenvalue minimization

$$\text{minimize} \quad \lambda_{\max}(F(x))$$

- the function $\lambda_{\max}(\cdot)$ is nonconvex
- this problem can be equivalently reformulated as:

$$\begin{array}{ll} \text{minimize} & t \\ \text{subject to} & F(x) - tI \preceq 0 \end{array}$$

  where the variables are $x \in \mathbb{R}^n$ and $t \in \mathbb{R}$

- this is a specific instance of an SDP in the augmented (vector) variable:

$$\hat{x} = \left[\begin{array}{c} t \\ x \end{array} \right], \quad \hat{c} = (1, 0, \ldots, 0)\,, \quad \hat{F}(\hat{x}) = F(x) - tI$$

Example: spectral matrix norm minimization

$$\text{minimize} \quad \|A(x)\|_2$$

- $A(x) = A_0 + x_1 A_1 + \cdots + x_n A_n \in \mathbb{R}^{p \times m}$
- this problem is equivalent to the following SDP:

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

with decision variables $x \in \mathbb{R}^n$ and $t \in \mathbb{R}$ ($t \geq 0$)

- to show this, recall that the spectral norm is

$$\|A(x)\|_2 = \sqrt{\lambda_{\max}(A^T(x)A(x))}$$

- it follows that

$$\|A(x)\|_2 \le t \iff A^T(x)A(x) \preceq t^2 I, \quad t \ge 0$$

- using the Schur complement rule, this matrix inequality is same as

$$\begin{bmatrix} t^2 I_m & A^T(x) \\ A(x) & I_p \end{bmatrix} \succeq 0 \iff \begin{bmatrix} tI_m & A^T(x) \\ A(x) & tI_p \end{bmatrix} \succeq 0$$

right inequality obtained by congruence transformation with

$$\mathrm{diag}(1/\sqrt{t} I_m, \sqrt{t} I_p)$$

for $t > 0$

Example: Frobenius norm minimization

$$\text{minimize} \quad \|A(x)\|_F^2$$

- equivalent to SDP:

$$\begin{array}{ll} \text{minimize} & \operatorname{tr}(Y) \\ \text{subject to} & \begin{bmatrix} Y & A(x) \\ A^T(x) & I_m \end{bmatrix} \succeq 0 \end{array}$$

  where the variables are $x \in \mathbb{R}^n$ and $Y \in \mathbb{R}^{p \times p}$ is positive semidefinite

- the equivalence of this formulation can be established by noting the relationship:

$$\|A(x)\|_F^2 = \operatorname{tr}(A(x)A^T(x))$$

- using the Schur complement, the matrix condition can be written as:

$$\begin{bmatrix} Y & A(x) \\ A^T(x) & I_m \end{bmatrix} \succeq 0 \iff A(x)A^T(x) \preceq Y$$

  this validation links the original objective with the SDP representation

Outline

- linear programs
- piecewise-linear minimization
- quadratic optimization
- geometric programming
- semidefinite programs
- **quasiconvex optimization**

Quasiconvex function

$f : \mathbb{R}^n \to \mathbb{R}$ is *quasiconvex* if its domain and all of its sublevel sets

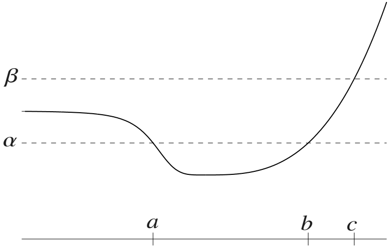
$$\mathcal{S}_\gamma = \{x \in \operatorname{dom} f \mid f(x) \leq \gamma\}$$

are convex for every real number γ

- every convex function naturally possesses convex level sets
- there exist non-convex functions that have convex level sets
- a function is *quasiconcave* if its negative $(-f)$ is quasiconvex
- a function that's both quasiconvex and quasiconcave is called *quasilinear*
  - both their domain and each level set $\{x \mid f(x) = \alpha\}$ are convex

Graphical illustration

quasiconvex function that is non-convex



- $S_\alpha = [a, b]$ is convex
- $S_\alpha = (\infty, c)$ is convex

Examples

- $f(x) = \sqrt{|x|}$ is nonconvex, but it is quasiconvex
  - when $\gamma < 0$, then $\mathcal{S}_\gamma = \emptyset$
  - for $\gamma \geq 0$, the sublevel set is given by:

$$\mathcal{S}_\gamma = \{x \mid \sqrt{|x|} \leq \gamma\} = \{x \mid |x| \leq \gamma^2\} = [-\gamma^2, \gamma^2]$$

- $\log x$ over $\mathbb{R}_{++}$ is both quasiconvex and quasiconcave, making it quasilinear

- $\mathrm{ceil}(x) = \inf\{z \in \mathbb{Z} \mid z \geq x\}$, is quasiconvex and quasiconcave

- the nonconvex $f(x_1, x_2) = x_1 x_2$ is quasiconcave on $\mathbb{R}_+^2$ but not on $\mathbb{R}^2$

- the function

$$f(x) = \frac{a^T x + b}{c^T x + d}, \quad \mathrm{dom}\, f = \{x \in \mathbb{R}^n \mid c^T x + d > 0\}, c \neq 0$$

is quasiconvex since

$$\mathcal{S}_\gamma = \{x \mid f(x) \leq \gamma\} = \{x \in \mathbb{R}^n \mid (a - \gamma c)^T x + (b - \gamma d) \leq 0\}$$

is a convex set

- given points $a, b \in \mathbb{R}^n$, the function

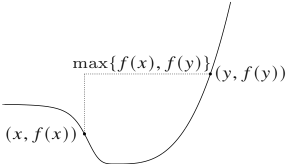
$$f(x) = \frac{\|x - a\|}{\|x - b\|}$$

is quasiconvex since its sublevel set represents the halfspace where the distance to a is less than or equal to the distance to b

Properties of quasiconvex function

- f is quasiconvex iff $\mathrm{dom}\, f$ is convex and for any $x, y \in \mathrm{dom}\, f$ with $0 \le \theta \le 1$,

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



- a differentiable f with convex domain is quasiconvex if and only if

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

- a sum of quasiconvex functions is not necessarily quasiconvex

Examples

- the cardinality $x \in \mathbb{R}^n$, denoted $\mathrm{card}(x)$, is the no. of its non-zero entries

  $\mathrm{card}(x)$ is quasiconcave on $\mathbb{R}^n_+$ but not on $\mathbb{R}^n$; this stems from the fact:

$$\mathrm{card}(x+y) \geq \min\{\mathrm{card}(x), \mathrm{card}(y)\},$$

  valid for non-negative vectors x, y

- the rank is quasiconcave on positive semidefinite matrices since

$$\mathrm{rank}(X+Y) \geq \min\{\mathrm{rank}\, X, \mathrm{rank}\, Y\}$$

  holds for positive semidefinite matrices X, Y

Quasiconvex optimization

a quasiconvex optimization problem in standard form is represented as

$$\begin{array}{ll} \text{minimize} & f(x) \\ \text{subject to} & g_i(x) \leq 0, \quad i = 1, \ldots, m \\ & Ax = b \end{array} \tag{10.3}$$

- the objective f is quasiconvex
- g_i are convex
- can have locally optimal points that are not (globally) optimal

Convex representation of sublevel sets of f

if f is quasiconvex, there exists a family of functions $\phi_t(x)$ such that:

- $\phi_t(x)$ os convex in x for fixed t
- t-sublevel set of f is 0-sublevel set of $\phi_t(x)$:

$$f(x) \leq t \Longleftrightarrow \phi_t(x) \leq 0$$

  where for every x, we have $\phi_s(x) \leq \phi_t(x)$ for any $s \geq t$

Example

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

with p convex, q concave, and $p(x) \geq 0$, $q(x) > 0$ on $\mathrm{dom}\, f$

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

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

Quasiconvex optimization via convex feasibility problems

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

- if feasible then $p^\star \leq t$; $p^\star$ is optimal solution of original quasiconvex problem
- if infeasible, then $p^\star \geq t$;

Bisection for quasiconvex problems

---

given: $l \leq p^\star, u \geq p^\star$ and a tolerance $\epsilon > 0$

repeat

1. $t := \frac{l+u}{2}$
2. solve the convex feasibility problem
3. if feasible, set $u := t$; else, set $l := t$

until $u - l \leq \epsilon$

References and further readings

- S. Boyd and L. Vandenberghe. *Convex Optimization.* Cambridge University Press, 2004. (chapters 2.2.1, 2.2.4, 4.3)

- G. C. Calafiore and L. El Ghaoui. *Optimization Models.* Cambridge University Press, 2014. (chapter 9).