DSA3102 Lecture 6

Announcement: - Midterm exam on 05/10/2023 (Thu) in class.
- Topics: Everything up to and including Lecture 6 (Chapter 4)
- Time: 7pm to 8:30pm (90 mins)
- Location: LT28 (Only in-person)
- You may bring 1 cheatsheet (A4 size) written on both sides.
- You will write on paper. No iPad.

Recap: Linear Optimization/Programming (LP)

optimising a certain linear function of x

$$\min_{x \in \mathbb{R}^n} \ c^T x \ \text{(td)} \quad \text{s.t.} \quad \begin{array}{ll} Gx \le h & \text{Inequality constraints} \\ Ax = b. & \text{Equality constraints} \end{array} \Bigg] \text{ not in standard form}$$

Standard form: $\min_{x \in \mathbb{R}^n} c^T x$ s.t. $Ax = b$, $x \ge 0$.

Today: ① Quadratic Optimization Problems (QCQP)
② Second-Order Cone Programming (SOCP)
③ Geometric Programming
④ Semidefinite Programming. (SDP)

Quadratic Optimization

new term not included in LP problem

$$\min_{x \in \mathbb{R}^n} \boxed{\tfrac{1}{2}x^T P x} + q^T x + r \quad = f_0(x) \qquad \text{s.t. } Gx \le h \quad (m \text{ ineq. constraints}), \quad Ax = b \quad (p \text{ equality constraints})$$

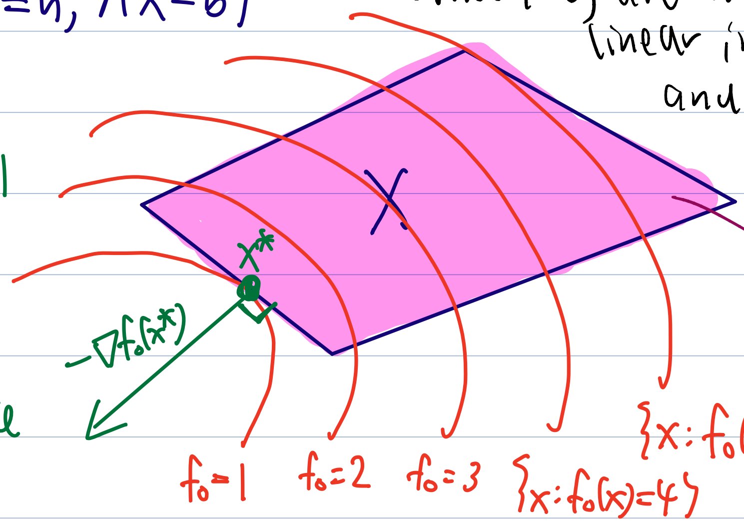
+ve semi-definite matrix

$P \in S_+^n$, $G \in \mathbb{R}^{m \times n}$, $A \in \mathbb{R}^{p \times n}$, $q \in \mathbb{R}^n$, $h \in \mathbb{R}^m$, $b \in \mathbb{R}^p$

constraint set

$X = \{x \in \mathbb{R}^n : Gx \le h, Ax = b\}$ Polyhedral constraints since they are defined by linear inequalities and equalities

By 1st order optimality condition: If $x^* \in X$ is optimal $-\nabla f_0(x^*)$ points outwards from X and is perpendicular to a face of X.



all solutions must lie inside here to be feasible

level sets are quadratics

+ve semi-definite

If $P \succeq 0$, then the above QP is convex

If $P \succ 0$, then the above QP is strictly convex. } optimum solution (minimum) is unique

+ve definite

QCQP (Quadratically constrained quadratic program)

symmetric matrix

$$\min_x \tfrac{1}{2}x^T P_0 x + q_0^T x + r_0 \quad = f_0(x)$$

s.t.

$$\tfrac{1}{2}x^T P_i x + q_i^T x + r_i \le 0 \quad \forall i = 1, \dots, m$$

Constraints are quadratic

set of all x that satisfy above inequality is a convex set

$Ax = b.$

all P_i are symmetric and +ve semi definite

$P_i \in S_+^n \Rightarrow f_i$'s are convex f^ns.

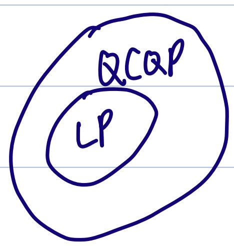
If we take P_i's $= 0$ (matrix), we obtain a linear prog.

By setting quadratic term to 0

$$\min_x \quad q_0^T x + r_0$$

$$\text{s.t.} \quad q_i^T x + r_i \leq 0 \quad \forall i = 1, \dots, m$$

$$Ax = b$$



Ex: Linear Regression + Least Squares

want to minimize residuals

$b = Ax + \text{noise}.$

is an example of a QP.

$$\min_x \left\{ \boxed{\|Ax - b\|_2^2} = x^T A^T A x - 2b^T A x + \|b\|^2 \right\}.$$

differentiable after you square

$x^* = A^\dagger b \qquad A^\dagger$: pseudo-inverse of A

If A has full column rank, $A^\dagger = (A^T A)^{-1} A^T$

if A does not have full column rank, do something like an SVD

Ex: Linear Program with Random Cost

go from not standard LP → standard LP: add slack variables eg. $x = x^+ + x^-$

gives rise to a quadratic program

LP: $\min_{x \in \mathbb{R}^n} c^T x$ s.t. $x \geq 0$, $Ax = b$

whole thing is a random variable

c is a random vector

covariance matrix is +ve semi-definite

$\mathbb{E}[c] = \bar{c}$ (avg), $\quad \mathrm{Cov}(c) = \mathbb{E}[(c-\bar{c})(c-\bar{c})^T] = \Sigma \succeq 0.$

$\mathbb{E}[c^T x] = x^T \mathbb{E}[c] = x^T \bar{c} = \bar{c}^T x$ linearity of expectation

$$\begin{aligned} \mathrm{Var}(c^T x) &= \mathbb{E}[(c^T x - \bar{c}^T x)^2] \\ &= \mathbb{E}[((c-\bar{c})^T x)^2] \\ &= \mathbb{E}[x^T (c-\bar{c})(c-\bar{c})^T x] \\ &= x^T \underbrace{\mathbb{E}[(c-\bar{c})(c-\bar{c})^T]}_{\text{covariance matrix of } c} x = x^T \Sigma x \end{aligned}$$

Instead of min $c^T x$, we min a linear combination of $\mathbb{E}[c^T x]$ & $\mathrm{Var}(c^T x)$

$\gamma > 0$

quadratic in x

$$\min_{x \in \mathbb{R}^n} \underbrace{\bar{c}^T x}_{\mathbb{E}[c^T x]} + \gamma \underbrace{x^T \Sigma x}_{\mathrm{Var}(c^T x)} \qquad x \geq 0, \quad Ax = b$$

γ: risk-aversion parameter

minimise the mean but we don't want too much fluctuation so also minimise some quantity related to variance

This is a quadratic program with polyhedral constraint.

modify LP

MIDTERMS start here!

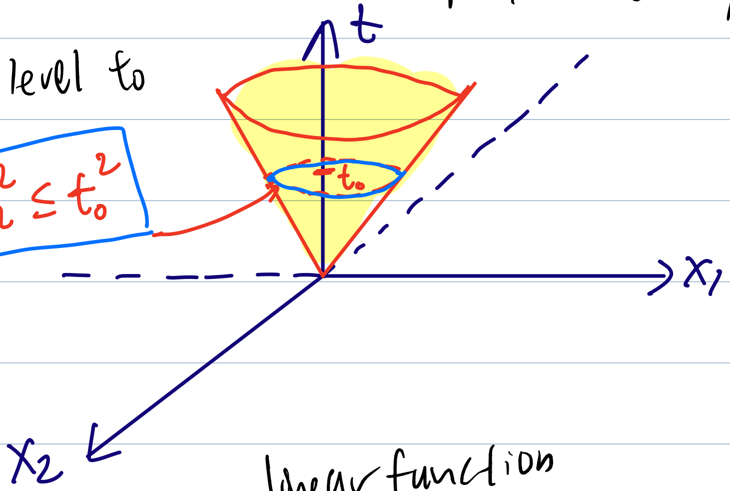
Second-Order Cone Programming (SOCP)

Second-Order Cone $K = \{(x,t) \in \mathbb{R}^{n+1} : t \geq 0, \|x\|_2 \leq t\}$

pair, x is vector, t is scalar

Slicing cone at level t_0

$\boxed{x_1^2 + x_2^2 \leq t_0^2}$



for diff levels, diff t_0, go up to obtain bigger circles

An SOCP: $\min_{x \in \mathbb{R}^n} f^T x$ s.t. $\|A_i x + b_i\| \leq c_i^T x + d_i, i \in [m]$

linear function of x

$Fx = g.$

$A_i \in \mathbb{R}^{n_i \times n}$, $b_i \in \mathbb{R}^{n_i}$, $c_i \in \mathbb{R}^n$, $d_i \in \mathbb{R}$, $F \in \mathbb{R}^{p \times n}$, $g \in \mathbb{R}^p$

We say that an inequality of the form

$$\|Ax + b\|_2 \leq c^T x + d \qquad A \in \mathbb{R}^{k \times n}, b \in \mathbb{R}^k$$

2 norm euclidean distance?

is a second-order cone constraint. set of all (x, t) pairs

$\because \ (Ax + b, c^T x + d) \in K.$ such that norm of x not more than t

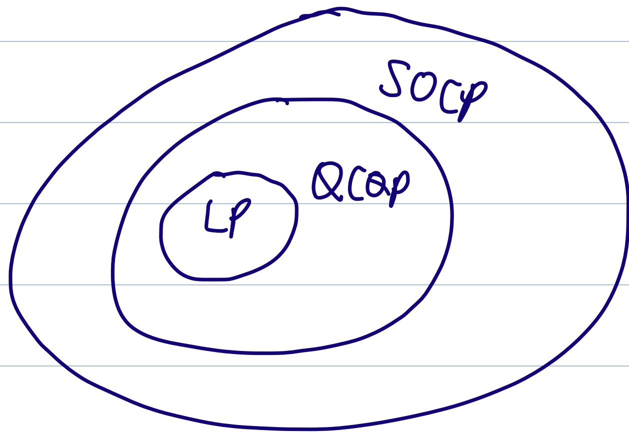
Rmks: i) If $c_i = 0$, $\|A_i x + b_i\|_2^2 \leq d_i^2$

$$\Rightarrow x^T A_i^T A_i x + 2 b_i^T A_i x + \|b_i\|_2^2 - d_i^2 \leq 0$$

$$\Rightarrow x^T P_i x + q_i^T x + r_i \leq 0.$$

a quadratic constrain

recovered a QCQP



LP

ii) If $A_i = 0$, $\quad \|b_i\| \leq c_i^T x + d_i,$

LP subset of SOCP.

$$\Rightarrow -c_i^T x + \|b_i\| - d_i \leq 0$$

$$\tilde{a}_i^T x + \tilde{b}_i \leq 0 \quad \text{affine constraint. (Linear)}$$

Example: Robust Linear Programs

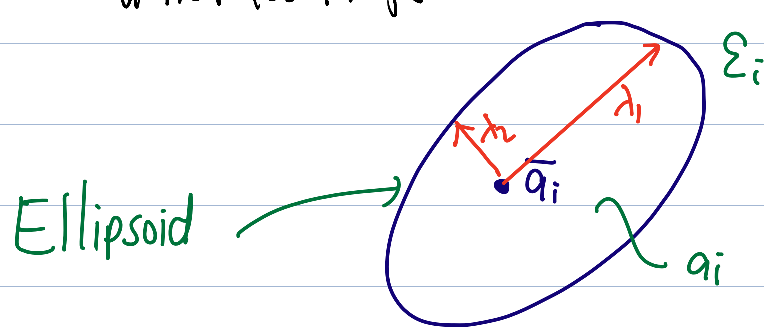
but we may never be certain about our parameters (may be random)

$$\min_x \; c^T x \quad \text{s.t.} \quad a_i^T x \leq b_i \quad \forall i \in [m]$$

$$a_i \in \mathcal{E}_i = \{ \bar{a}_i + P_i u : \|u\| \leq 1 \} \qquad P_i \in \mathbb{R}^{n \times n}.$$

u not too large

nominally $\bar{a}_i$ and can be perturbed by a little bit



$$\min_x \; c^T x \quad \text{s.t.} \quad a_i^T x \leq b_i \quad \underbrace{\forall a_i \in \mathcal{E}_i}_{\text{belonging to the set}} \quad \forall i \in [m]$$

↳ protecting against worst case outcome

maximise over all a_i
(as long as sup smaller than b_i then everyone smaller)

$$\equiv \min_x \; c^T x \quad \text{s.t.} \quad \sup_{a_i \in \Sigma_i} a_i^T x \le b_i \quad \forall i \in [m]$$

(above form)

$$\sup_{a_i \in \Sigma_i} a_i^T x = \sup_u \{ a_i^T x : a_i = \bar{a}_i + P_i u, \|u\| \le 1 \}$$

$$= \sup_u \{ (\bar{a}_i + P_i u)^T x : \|u\| \le 1 \}.$$

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

$$= \bar{a}_i^T x + \| P_i^T x \|_2.$$

b fixed, looking at all possible a that maximise inner product

$$\sup_{a : \|a\| \le 1} a^T b = \|b\| \quad \because a^* \text{ must be parallel to } b$$

$$a^* = \frac{b}{\|b\|} \quad \text{and have norm } 1.$$

$$\min_x \; c^T x \quad \text{s.t.} \quad \sup_{a_i \in \Sigma_i} a_i^T x \le b_i \quad \forall i \in [m]$$

$$\equiv \min_x \; c^T x \quad \text{s.t.} \quad \bar{a}_i^T x + \|P_i^T x\| \le b_i \quad \forall i$$

$$\underbrace{\|P_i^T x\|}_{\text{vector}} \le \underbrace{b_i - \bar{a}_i^T x}_{\text{scalar}} \quad \forall i$$

Optimising a linear function over linear functions of x belonging to 2nd order cone

$$(P_i^T x, \; b_i - \bar{a}_i^T x) \in K$$

$\Rightarrow$ This is an SOCP.

Linear Programming with Random Constraints

$$(LP): \min_x \; c^T x \quad \text{s.t.} \quad \begin{aligned} &Ax \le b \Leftrightarrow a_i^T x \le b_i \;\; \forall i \in [m]. \\ &x \ge 0 \end{aligned}$$

Previously c was random with mean $\bar{c}$ & cov matrix Σ.

Now assume a_i's are independent Gaussian random vectors with mean $\bar{a}_i$ & covariance matrix $\Sigma_i \succeq 0$.

↳ all positive semi-definite

$$\Pr(a_i^T x \leq b_i) \geq \eta_i \qquad \eta_i = 0.99.$$

solve problem such that constraints are satisfied with probability at least 99%

$$\Pr\left(\bigcap_{i=1}^{m} \{a_i^T x \leq b_i\}\right) = \prod_{i=1}^{m} \eta_i$$

$$\Pr(a_i^T x \leq b_i) = \Pr\Bigg(\underbrace{\frac{a_i^T x - \overbrace{\bar{a}_i^T x}^{\text{mean}}}{\underbrace{\sqrt{x^T \Sigma_i x}}_{\text{sd}}}}_{N(0,1)} \leq \frac{b_i - \bar{a}_i^T x}{\sqrt{x^T \Sigma_i x}}\Bigg) \geq \eta$$

$$\Phi(t) = \int_{-\infty}^{t} \frac{1}{\sqrt{2\pi}} e^{-u^2/2} \, du \qquad = \Phi\left(\frac{b_i - \bar{a}_i^T x}{\sqrt{x^T \Sigma_i x}}\right) \geq \eta$$

$$\Rightarrow \frac{b_i - \bar{a}_i^T x}{\sqrt{x^T \Sigma_i x}} \geq \Phi^{-1}(\eta).$$

$$\min_x \; c^T x \quad \text{s.t.} \quad \frac{b_i - \bar{a}_i^T x}{\sqrt{x^T \Sigma_i x}} \geq \Phi^{-1}(\eta) \qquad \forall i \in [m]$$

$$\min_x \; c^T x \quad \text{s.t.} \quad \bar{a}_i^T x + \Phi^{-1}(\eta) \|\Sigma_i^{1/2} x\|_2 \leq b_i$$

$$\sqrt{x^T \Sigma_i x} = \sqrt{x^T \Sigma_i^{1/2} \Sigma_i^{1/2} x} = \|\Sigma_i^{1/2} x\|_2$$

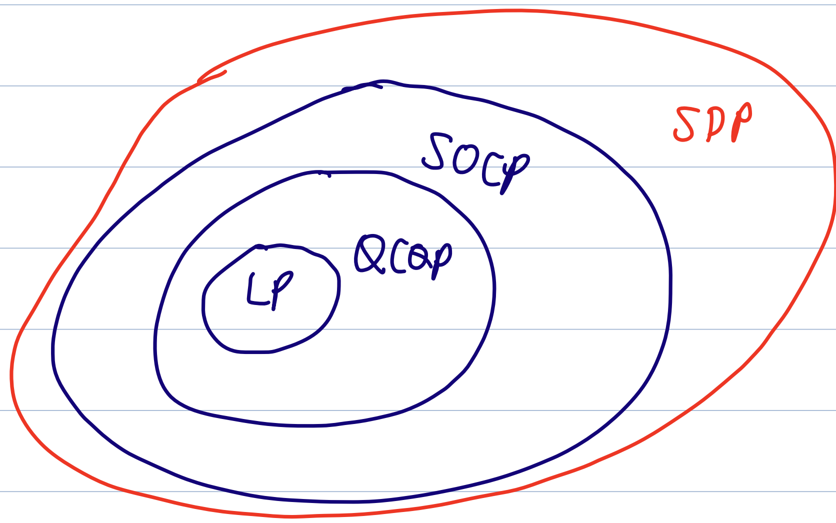
Σ_i is positive definite

vector $\qquad\qquad$ scalar

$$\min_x c^T x \quad \text{s.t.} \quad \|\Sigma_i^{1/2} x\|_2 \le \frac{1}{\Phi^{-1}(\eta)}(b_i - \bar{a}_i^T x) \;\; \forall i$$

Again the constraints represent a second-order conic constraint

$$\left(\Sigma_i^{1/2} x, \frac{1}{\Phi^{-1}(\eta)}(b_i - \bar{a}_i^T x)\right) \in K.$$



Geometric Programming (GP) – does not look convex but can be converted into something convex

Def: A function $f: \mathbb{R}^n \to \mathbb{R}$ with $\text{dom} f = \mathbb{R}^n_{++}$ defined as

$$f(x) = c x_1^{a_1} \cdots x_n^{a_n}$$

all inputs must be +ve vectors

where $c > 0$ & $a_i \in \mathbb{R}$ is a monomial.

because

power can be any number (like root) hence we want to prevent getting complex numbers

Eg: $f(x) = \pi x^{2.7}$, $f(x_1, x_2) = 3 x_1^{-1} x_2^{7.2}$

monomial as coefficient is +ve and power is real number

Def: A posynomial is a function $f: \mathbb{R}^n \to \mathbb{R}$ with $\text{dom } f = \mathbb{R}^n_{++}$ of the form.

$$f(x) = \sum_{k=1}^{K} c_k x_1^{a_{1k}} \cdots x_n^{a_{nk}}, \quad c_k > 0, \quad a_{ik} \in \mathbb{R}.$$

sum of K monomials

entire class cannot be convex

$$\boxed{f(x_1, x_2)} = 3x_1^{-1} x_2^{7.2} + 7x_1^{1/2} x_2^{-1/2}$$

concave?

Posynomials are closed under addition, multiplication & nonnegative scaling.

Geometric Programming (GP)

bound is not zero ⇒ after transformation RHS will be 0

$$\min_x \; f_0(x) \quad \text{s.t.} \quad f_i(x) \leq 1 \quad \forall i \in [m], \qquad h_i(x) = 1 \quad \forall i \in [p].$$

f_i: posynomials h_i: monomials. $\mathcal{D} = \mathbb{R}^n_{++}$, $x > 0$.

only +ve coordinates allowed

$$\max_{x,y,z} \; \frac{x}{y} \quad \text{s.t.} \quad 2 \leq x \leq 3, \quad x^2 + 3y/z \leq \sqrt{y}, \quad x/y = z^2$$

GP

write in standard GP

check both x, y non-zero (true since all +ve)

$$\equiv \min_{x,y,z} \; \frac{y}{x} \quad \frac{x}{3} \leq 1 \qquad 2x^{-1} \leq 1$$

$$x^2 y^{-1/2} + 3y^{1/2} z^{-1} \leq 1$$

$$x y^{-1} z^{-2} = 1$$

GPs are, a priori, not convex.

Consider the change-of-variables $y_i = \log x_i$, $x_i = e^{y_i}$

If f is the monomial $f(x) = c\, x_1^{a_1} \cdots x_n^{a_n}$

$$f(x) = c\,(e^{y_1})^{a_1} \cdots (e^{y_n})^{a_n}$$

$$= e^{a_1 y_1 + \cdots + a_n y_n + b}$$ exponent becomes linear, $b = \log c$

If f is the posynomial $f(x) = \sum_{k=1}^{K} c_k\, x_1^{a_{1k}} \cdots x_n^{a_{nk}}$

with $y_i = \log x_i$, $x_i = e^{y_i}$

$$f(x) = \sum_{k=1}^{K} e^{a_k^T x + b_k}$$ $a_k = (a_{1k}, \ldots, a_{nk})^T$, $b_k = \log c_k$.

convex form

So the originally non-convex GP can be rewritten as

$$\min_x \sum_{k=1}^{K_0} e^{a_{0k}^T x + b_{0k}}$$ the 0 represent the f_{00}, the coefficient for the objective

$$\text{s.t.} \quad \sum_{k=1}^{K_i} e^{a_{ik}^T x + b_{ik}} \le 1 \quad \forall i \in [m]$$

$$e^{g_i^T x + h_i} = 1 \quad \forall i \in [p].$$

$\underline{a_{ik}} \in \mathbb{R}^n$ $i = 0, 1, \ldots, m$ $g_i \in \mathbb{R}^n$

Contain the exponents of the posynomial inequality constraints.

$$\min_x \log\left(\sum_{k=1}^{K_0} e^{a_{0k}^T x + b_{0k}}\right)$$

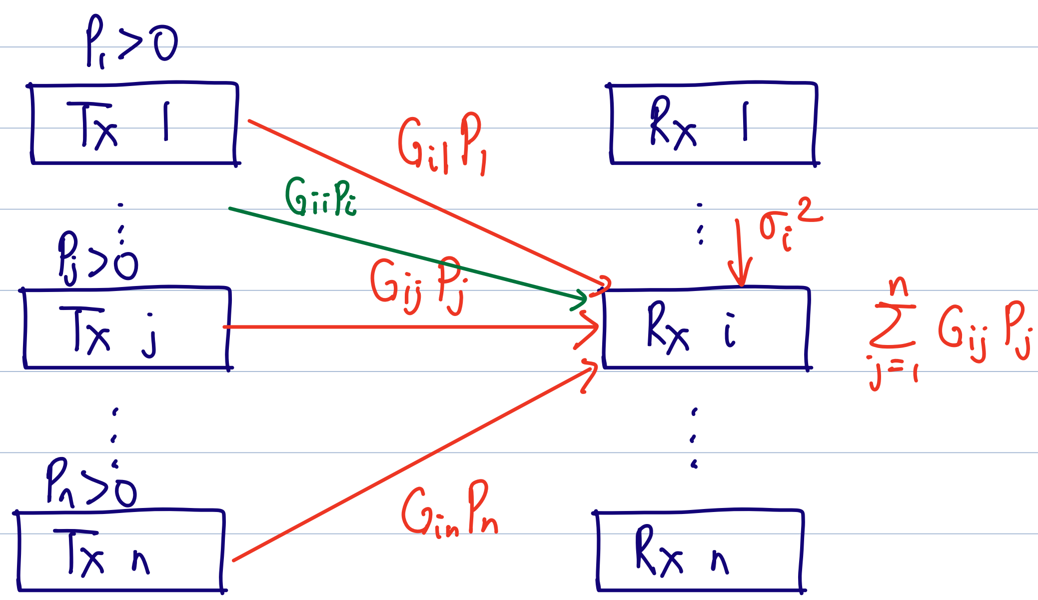
$$\log\left(\sum_{k=1}^{K_i} e^{a_{ik}^T x + b_{ik}}\right) \le 0 \quad \forall i \in [m]$$

$$g_i^T x + h_i = 0 \qquad \forall i \in [p].$$

The log-sum-exp function is convex so the objective f^n and ineq. constraint functions are convex.

$\Rightarrow$ convex optimization prob.

Example of Geometric Programming : Power Control.



n transmitters $\qquad$ n receivers

transmitting at positive powers $P_i > 0$

Power received from Tx j at Rx i is $G_{ij} P_j$.

Signal Power at Rx i is $\sum_{j=1}^{n} G_{ij} P_j$

Noise Power at Rx i is σ_i^2

intended signal power (what we are interested in)

$$SINR_i = S_i = \frac{G_{ii} P_i}{\sum_{j \neq i} G_{ij} P_j + \sigma_i^2}$$

noise

signal to interference and noise ratio

interference

Design $P_1, \ldots, P_n$ s.t. $SINR_i \geq S_{min} > 0$

(P_i cannot be too big) $0 < P_{min} \leq P_i \leq P_{max} < \infty$

$$\min_{P_1, \ldots, P_n > 0} \sum_{i=1}^{n} P_i \quad \text{s.t.} \quad P_{min} \leq P_i \leq P_{max}$$

$$\frac{G_{ii} P_i}{\sum_{j \neq i} G_{ij} P_j + \sigma_i^2} \geq S_{min}.$$

$$\equiv \quad \min_{P_1, \ldots, P_n > 0} \sum_{i=1}^{n} P_i \quad \text{s.t.} \quad P_{min} \leq P_i \leq P_{max}$$

(*) $$\underbrace{S_{min} \frac{\sum_{j \neq i} G_{ij} P_j + \sigma_i^2}{G_{ii} P_i}}_{\text{is a posynomial in } (P_1, \ldots, P_n)} \leq 1 \quad \forall i$$

variables

This is a GP in $P_1, \ldots, P_n > 0$.

$\because$ the constraints involve $P_i^{a_i}$ where $a_i \in \{-1, +1\}$.

$\downarrow$

can convert to convex programme

Use $y_i = \log p_i$ to convert (*) to a log-sum-exp convex opt. problem

---

Generalized Inequalities

K: proper cone defines a partial order

positive semi-definite cone ... positive SD with respect to K

$K = S_+^n$ $\quad A \succeq_K B \iff A - B \succeq_K 0 \iff \lambda_i (A-B) \geq 0.$

$K = \mathbb{R}_+^n$ $\quad x \succeq_K y \iff x - y \succeq_K 0 \iff x_i - y_i \geq 0 \;\; \forall i.$

non negative vector

Def: $f: \mathbb{R}^n \to \mathbb{R}^m$ is K-convex if $\forall x, y \in \operatorname{dom} f$ &

generalisation of convexity

$\theta \in [0,1]$,

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

Ex: $K = S_+^n$, set of all symmetric matrices $\quad S_+^n$ – convex

K – convex

$f: \mathbb{R}^n \to S^n$ is a matrix-valued function & is matrix-convex if

$$f(\theta x + (1-\theta) y) \preceq_{S_+^n} \theta f(x) + (1-\theta) f(y).$$

matrix-convex

To check this, check that $z^T f(X) z$ is convex in X for all z.

Ex: $f(X) = XX^T \qquad f: \mathbb{R}^{n \times m} \to S^n$

output a symmetric matrix

Claim: f is matrix-convex.

Pf: Check that $\forall z$, $z^T f(X) z$ is convex in X.

$$z^T X X^T z = (X^T z)^T (X^T z) = \|X^T z\|^2$$

This function is quadratic in X, and hence convex in X. hence shown that $f(X)$ is a K-convex function

Optimization Problem with generalized ineq. constraints.

$$\min_x f_0(x) \quad \text{s.t.} \quad Ax = b \text{ (usual equality constraint)}$$

$$-f_i(x) \in K \leftarrow f_i(x) \preceq_{K_i} 0 \quad \forall i \in [m].$$

(generalized inequality constraint defined by proper cone)

When $K = S_+^n$, the associated conic form problem is known as a semidefinite program (SDP)

semidefinite const.

$$\min_x \underbrace{c^T x}_{\text{linear function}} \quad \text{s.t.} \quad \boxed{x_1 F_1 + x_2 F_2 + \dots + x_n F_n + G} \preceq_K 0$$

in terms of the +ve semi-definite order

$$Ax = b$$

s.t p linear constraint

make sure this entire thing is negative definite

$G, F_i \in S^k$, $A \in \mathbb{R}^{p \times n}$

symmetric matrices

If G, F_i are diagonal, consider the constraint

symmetric

$$x_1 \begin{bmatrix} a & 0 \\ 0 & b \end{bmatrix} + x_2 \begin{bmatrix} c & 0 \\ 0 & d \end{bmatrix} + \begin{bmatrix} e & 0 \\ 0 & f \end{bmatrix} \preceq_K 0$$

in partial order

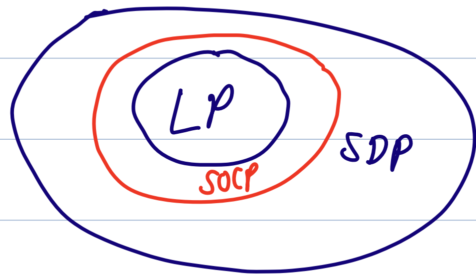
$$\begin{bmatrix} ax_1 + cx_2 + e & 0 \\ 0 & bx_1 + dx_2 + f \end{bmatrix} \preceq 0$$

diagonal matrix and -ve semi-definite, its entries all ≤ 0

linear!

$\underline{ax_1 + cx_2 + e \leq 0}$ $\quad$ $\underline{bx_1 + dx_2 + f \leq 0}$

Hence, the constraints become affine & the problem is an LP.



semi-definite constrains

Recall SOCP: $\min_x c^T x$ s.t. $\|A_i x + b_i\| \leq c_i^T x + d_i \quad \forall i$

$Fx = g$

convert into a constraint regarding +ve semi-definite matrices?

$(x,t) \in K = \{(x,t) : \|x\|_2 \leq t,\ t \geq 0\}$

$\|x\|_2 \leq t \iff \underline{x_1^2 + x_2^2 \leq t^2}$ $\rightarrow$ write in the form $(t, x_1, x_2) \succeq 0$

Arrow form:

$$\begin{array}{ccc} ① & ② & ③ \\ \end{array}$$

$$\begin{bmatrix} t & x_1 & x_2 \\ x_1 & t & 0 \\ x_2 & 0 & t \end{bmatrix} \succeq 0$$

① $t \geq 0$

② $t^2 \geq x_1^2$

convert any SOCP → SDP constrain

③

$t^3 - x_1(x_1 t) - x_2(x_2 t) \geq 0$

$\Rightarrow x_1^2 + x_2^2 \leq t^2$ //

$X = \begin{pmatrix} A & B \\ B^T & C \end{pmatrix}$ $\quad \det(A) \neq 0$

$S = C - B^T A^{-1} B$: Schur complement of A in X.

Schur complement lemma:

Consider $X = \begin{pmatrix} A & B \\ B^T & C \end{pmatrix}$ with $A > 0$ (+ve definite). Then

$$X \succeq 0 \iff S \succeq 0.$$

$X \succeq 0$: +ve semi definite

SOCP $\Rightarrow$ SDP.

$\|x\| \le t$ $\qquad \begin{pmatrix} tI & x \\ x^T & t \end{pmatrix} \succeq 0 \qquad t > 0$

the schur complement

$$\iff t - x^T (tI)^{-1} x \geq 0$$

$$t \geq t^{-1} x^T x \Rightarrow \|x\| \le t.$$

SDP in standard form

$$\min_{X \in S^n} \operatorname{tr}(CX) \quad \text{s.t.} \quad \operatorname{tr}(A_i X) = b_i \quad \forall i \in [p]. \qquad X \succeq 0$$

$\operatorname{tr}(CX)$: linear f^n in X

$C, A_1, \ldots, A_p \in S^n$ symmetric

Suppose C, A_1 diagonal $\quad C = \begin{bmatrix} c_1 & & 0 \\ & \ddots & \\ 0 & & c_n \end{bmatrix}, \quad A_1 = \begin{bmatrix} a_1 & & 0 \\ & \ddots & \\ 0 & & a_n \end{bmatrix}$

$$\operatorname{tr}(CX) = \sum c_i x_i \quad \text{s.t.} \quad \operatorname{tr}(A_1 X) = \sum a_i x_i$$

trace: multiply the 2 together and sum the diagonals

recover an LP. linear in the decision variable x_i

Examples: Experimental Design

Estimate a vector from noisy linear measurements

when A has full column rank,

$y = Ax + w$ $\qquad w \sim N(0, I)$

design matrix

Least squares estimator $\hat{x} = (A^TA)^{-1}A^Ty$

Error (Residuals) $e = \hat{x} - x$

$$\hat{x} = (A^TA)^{-1}A^T(Ax + w) = \cancel{(A^TA)^{-1}}(\cancel{A^TA}x) + (A^TA)^{-1}A^Tw$$
$$= x + (A^TA)^{-1}A^Tw$$
$$e = \hat{x} - x = (A^TA)^{-1}A^Tw$$

only thing random

Error covariance $= \mathbb{E}[(x - \hat{x})(x - \hat{x})^T]$

$$= \mathbb{E}[(A^TA)^{-1}A^T w w^T A(A^TA)^{-1}]$$
$$= (A^TA)^{-1}\cancel{A^T}\, \cancel{\mathbb{E}[ww^T]}\, \cancel{A(A^TA)^{-1}} = (A^TA)^{-1}$$

calc this in terms of the design matrix

$$A = \begin{array}{c c} \overset{\longleftarrow\; d \;\longrightarrow}{} & \\ \begin{bmatrix} - & a_1 & - \\ & \vdots & \\ - & a_n & - \end{bmatrix} & \updownarrow n \end{array}$$

want to minimize error covariance

Rows of A $a_1, \dots, a_n \in \mathbb{R}^d$ can only be chosen from the set of M possible test vectors $a_i \in \{v^{(1)}, \dots, v^{(M)}\} \subset \mathbb{R}^d$.

choose matrix: only choose rows from this allowable set

can only choose between these m different choices

cannot minimise matrix,
can only minimise function of
~~the~~ matrix

We can write

$$A^TA = n \cdot \sum_{i=1}^{M} \mu_i v^{(i)} (v^{(i)})^T$$

where μ_i is the fraction of times, out of n, that vector $v^{(i)}$ is used from $\{v^{(1)}, \dots, v^{(M)}\}$

spectral norm   like error covariance

E-optimal design: $\min \lambda_{max}((A^TA)^{-1})$
$\equiv \max \lambda_{min}(A^TA)$

✓ turn out to be SDP

Rewrite as the following SDP $\lambda_{min}\left(\sum_{i=1}^{M} \mu_i v^{(i)} (v^{(i)})^T\right) \geq t.$

$$\max_{t, \mu_1, \dots, \mu_M} t \qquad \boxed{\sum_{i=1}^{M} \mu_i v^{(i)} (v^{(i)})^T \succeq tI.}$$

maximise matrix has min eigenvalue t

linear in $(t, \mu_1, \dots, \mu_M)$

s.t. $\sum_{i=1}^{M} \mu_i = 1$, $\mu_i \geq 0 \ \forall i \in [M]$.

probability distribution?

semidefinite constraint on t, μ.

E-optimal design is an SDP.

sum of diagonal elements

A-optimal design $\min \mathrm{Tr}((A^TA)^{-1}) = \sum_{j=1}^{d} [(A^TA)^{-1}]_{jj}$

$$\min_{\substack{t_1, \dots, t_d \\ \mu_1, \dots, \mu_M}} \sum_{j=1}^{d} t_j \qquad \text{s.t. } \sum_{i=1}^{M} \mu_i = 1, \ \mu_i \geq 0 \ \forall i \in [M].$$

$\Rightarrow$ is also an SDP!

semidefinite constrain

$$(*) \text{ — } \begin{bmatrix} \sum_{i=1}^{m} \mu_i v^{(i)}(v^{(i)})^T & e_j \\ e_j^T & t_j \end{bmatrix} \succeq 0 \quad \forall j = 1, \dots, d.$$

→ results in a bound on the j component

Constraint $(*)$ by the Schur-complement lemma is equivalent to $t_j \geq 0$ &

$$t_j - e_j^T \left(\sum_{i=1}^{m} \mu_i v^{(i)}(v^{(i)})^T \right)^{-1} e_j \geq 0$$

$$\underbrace{e_j^T \left(\sum_{i=1}^{m} \mu_i v^{(i)}(v^{(i)})^T \right)^{-1} e_j}_{j^{th} \text{ entry on the diagonal of } \left(\sum_{i=1}^{m} \mu_i v^{(i)}(v^{(i)})^T \right)^{-1}} \leq t_j$$

want to minimize the sum of t_j s
⇒ minimise sum of diagonal elements of inverse matrix

Midterm:
- Schur complement lemma
- convert something in SOCP to SDP