Semidefinite programming.

$$\left[\begin{array}{l} \min\limits_{X \in Sym^{n \times n}} Tr(CX) \\ \text{such that } Tr(A_i X) = b_i \\ X \succeq 0 \end{array}\right]$$

1. Reductions

A_i and C can be assumed symmetric, since

$$Tr(AX) = Tr((AX)^T) = Tr(XA^T) = Tr(A^T X)$$

so $\quad Tr\left(\frac{A+A^T}{2} X\right) = Tr(AX)$

2. Is it a convex optimization problem?

a) Feasible set is convex : • PSD matrices form a convex set

• Intersection with the other convex set (affine constraints)

$\boxed{\text{spectrahedron}}$

(can have infinitely many vertices)

b) Does not satisfy the definition of convex opt. problem as is.
(see discussion below too)

• $X \succeq 0 \leftrightarrow y^T X y \geq 0$ is a convex constraint for $\underline{\text{every}}$ $y \in \mathbb{R}^n$

infinitely many convex conditions

• finitely many polynomial constraints (equivalent to the original problem) — from Sylvester criterion.

3. Existence of optimal value $\not\Rightarrow$ existence of optimal solution

This is an SDP problem

Example: ⊛ $\left[\begin{array}{l} \min\ y \\ \frac{1}{x} \le y \\ x, y \geq 0 \end{array}\right]$

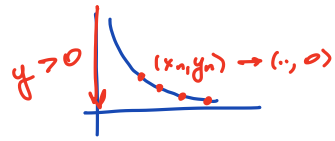
• $\left\{\begin{array}{l} xy \geq 1 \\ x, y \geq 0 \end{array}\right\} \Leftrightarrow \begin{pmatrix} x & 1 \\ 1 & y \end{pmatrix} \succeq 0$ (Sylvester)

• $Tr\left(\begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} x & 1 \\ 1 & y \end{pmatrix}\right) = Tr\begin{pmatrix} 0 & 0 \\ 1 & y \end{pmatrix} = y$

• $Tr\left(\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} x & z \\ z & y \end{pmatrix}\right) = 2z = 2$

$y_n \to 0$ as $x_n \to 0$; 0 - optimal value that is not attained at a feasible point

Convexity of the constraint set is not enough



there is no 'closeness' in $\mathrm{Tr}(CX) = y$ $(0, \infty)$

Compare with LP: $\begin{bmatrix} \min c^T x \\ Ax \le b \\ x \geq 0 \end{bmatrix}$ if $c^T x_n \to p_*$ $\exists x: \; p_* = c^T x$. The set $\{c^T x \mid x \text{ is feasible}\}$ is closed

4. Duality

Let p_* be an optimal value

$$L(X, \lambda, \mu) = \mathrm{Tr}(CX) + \sum \lambda_i (b_i - \mathrm{Tr}(A_i X)) - \mathrm{Tr}(X \cdot \mu) \quad - \text{ Lagrangian}$$

$$D(\lambda, \mu) = \min_X L(X, \lambda, \mu) \quad - \text{ dual function}$$

Lemma $\forall \lambda, \mu \succeq 0$ $D(\lambda, \mu) \le p^*$

Note: we prove it for the case when optimal solution x^* exists

(Pf)
$$L(X^*, \lambda, \mu) = \mathrm{Tr}(CX^*) - \mathrm{Tr}(X^* \mu) = p^* - \mathrm{Tr}(X^* \mu) \le p$$

[since $B, A \succeq 0 \Rightarrow \mathrm{Tr}(AB) \geq 0$ (HW 1)]

So, $\min L \le p^*$ as well ∎

Dual problem should provide the best lower bound for p_*.

Idea: maximize $D(\lambda, \mu)$

$$\begin{bmatrix} \max D(\lambda, \mu) \\ \text{s.t. } \mu \succeq 0 \end{bmatrix}$$

$$D(\lambda, \mu) = \min L(X, \lambda, \mu) = \begin{cases} \lambda^T b & \text{if } C - \sum_i \lambda_i A_i - \mu = 0 \\ -\infty & \text{otherwise} \end{cases}$$

Solving for $C - \sum \lambda_i A_i - \mu = 0$ for some $\mu \succeq 0$ we have

$$\boxed{C \succeq \sum_i \lambda_i A_i}$$

Dual problem:

$$\left[\begin{array}{l} \max\limits_{\lambda \in \mathbb{R}^m} \ b^T\lambda \\ \\ \text{s.t. } \sum_i \lambda_i A_i \preceq C \end{array}\right]$$

decision variables

$A_0 + \lambda_1 A_1 + \dots + \lambda_n A_n \succeq 0$

Linear Matrix Inequality (LMI)

Duality theorems:

Weak duality: Let X is feasible to SDP-primal, λ - feasible for SDP-dual. Then $Tr(CX) \geq b^T\lambda$

Strong duality ✗

Examples: ① only one of 2 optimal solutions are achieved

Dual to problem ⊛ $b = 2$, $A = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$, $C = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$

$$\left[\begin{array}{l} \max \ 2\lambda \\ \lambda A \preceq C \end{array}\right] \quad \begin{pmatrix} 0 & \lambda \\ \lambda & 0 \end{pmatrix} \preceq \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \Rightarrow \lambda = 0 \text{ optimal solution}$$

---

② SDP duality gap: (Lovasz example)

$$\min x$$
$$\begin{pmatrix} 0 & x & 0 \\ x & y & 0 \\ 0 & 0 & x+1 \end{pmatrix} \succeq 0$$

$\begin{pmatrix} -1 & 0 \end{pmatrix}$ $(x = 0)$

Let $X = \begin{pmatrix} x_{11} & x_{12} & x_{13} \\ x_{12} & x_{22} & x_{23} \\ x_{13} & x_{23} & x_{33} \end{pmatrix}$

$$\left[\begin{array}{l} \min \ -x_{33} \\ \text{s.t. } X \succeq 0 \\ \quad 2x_{12} + x_{33} = -1 \\ \quad x_{22} = 0 \end{array}\right]$$

$x_{33} = -1$

$\min -x_{33} = \underline{1}$

$$\max -x$$
$$x \underbrace{\begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}}_{A_1} + y \underbrace{\begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}}_{A_2} \succeq \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -1 \end{pmatrix}$$

Theorem: If primal and dual SDP are strictly feasible (exists positive definite feasible X) then optimal values are achieved and $Tr(CX^*) = b^T\lambda^*$.

Remark: dual form can be convenient to formulate problems:

Examples

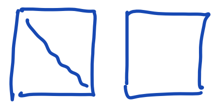
1) $\min y$: $\begin{pmatrix} x & 1 \\ 1 & y \end{pmatrix} = x\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} + \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} + \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} y \succeq 0$

2) $\begin{pmatrix} 1+x & y \\ y & 1-x \end{pmatrix} \succeq 0 \quad \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + x\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} + y\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \succeq 0 \iff$ $x \in [-1,1]$, $y \geq 0$
$x^2 + y^2 \leq 1$

6. LP and SOCP as special cases of SDP

LP

$$\left[\begin{array}{l} \min c^T x \\ a_i^T x = b_i \\ x \geq 0 \end{array}\right] \longleftrightarrow \left[\begin{array}{l} \min \operatorname{Tr}(\operatorname{diag}(c) X) \\ \operatorname{Tr}(\operatorname{diag}(a) X) = b_i \\ X \succeq 0 \end{array}\right]$$



non-diagonal entries do not participate in constraint and objective function

SOCP : we will show inclusion SOCP $\subset$ SDP

Lemma $X = \begin{pmatrix} A & B \\ B^T & C \end{pmatrix}$ is a block matrix with $A \succ 0$

Then $X \succeq 0 \iff S := C - B^T A^{-1} B \succeq 0$

(Schur complement)

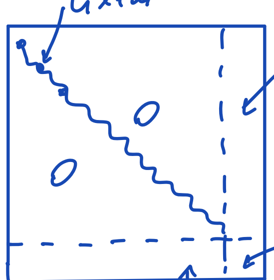
From Lemma, an SOCP problem can be rewritten as a special instance of SDP:

SOCP
$$\begin{cases} \min f^T x \\ \|A_i x + b_i\|_2 \leq c_i^T x + d_i \quad i = 1, \dots m \quad (m \text{ constraints}) \end{cases}$$

Without loss of generality $c_i^T x + d_i > 0$ (otherwise ith constraint is linear)

For each i, $(c_i^T x + d_i)^2 \geq \|A_i x + b_i\|_2^2 = \langle A_i x + b_i, A_i x + b_i \rangle$

$\iff S := (c_i^T x + d_i) - (A_i x + b_i)^T \cdot \dfrac{1}{c_i^T x + d_i} \cdot (A_i x + b_i) \geq 0$ for all $i = 1 \dots m'$

$X_i =$ 

$A = \operatorname{diag}(c_i^T x + d_i) \in \mathbb{R}^{m' \times m'}$

$\hookrightarrow A^{-1} = \operatorname{diag}((c_i^T x + d_i)^{-1})$

$\hookrightarrow B^T A^{-1} B$ is pointwise multiplication of $B^T B$ by $(c_i^T x + d_i)^{-1}$

S is Schur complement for the matrix $X = X_i$

By Lemma, $S \succeq 0 \iff X_i \succeq 0$

$(S \geq 0 \in \mathbb{R})$

So, for each i, we have an equivalent constraint $X_i \succeq 0$.

To get one SDP constraint, all X_i's can be joined in a block diagonal matrix

$$X := \begin{pmatrix} X_1 & & & & 0 \\ & X_2 & & & \\ & & X_3 & & \\ & & & \ddots & \\ 0 & & & & X_{m'} \end{pmatrix}$$

Exercise: $X \succeq 0 \iff$ all $X_i \succeq 0$, $i \in 1, \dots, m'$

(Proof of Lemma)

$$f(u,v) = u^T A u + 2 v^T B^T u + v^T C v = \begin{pmatrix} u \\ v \end{pmatrix}^T X \begin{pmatrix} u \\ v \end{pmatrix}$$

f is strictly convex in u.

$$\frac{\partial f}{\partial u} = 2Au + 2Bv = 0 \qquad u = -A^{-1}Bv \to \text{gives global minimum in } u.$$

$$f_v^* = \min_u f = v^T (C - B^T A^{-1} B) v = v^T S v$$

$\Rightarrow$ If $S \not\succeq 0$ $\quad v^T S v < 0$ implies $\min_u f < 0$ and $z^T X z < 0$ for $z = \begin{pmatrix} -A^{-1}Bv \\ v \end{pmatrix}$

$\Leftarrow$ $\begin{pmatrix} u \\ v \end{pmatrix}^T X \begin{pmatrix} u \\ v \end{pmatrix} = f(u,v) \geq 0$ is S is PSD. ∎

6. Spectahedron is a convex and closed set S

Convex optimization problems

(*)
$$\begin{cases} \min f(x) \\ g_i(x) \leq 0 & \text{convex} \\ h_j(x) = 0 & \text{affine} \end{cases}$$

Consider $d_S(x) = \inf_{s \in S} \|x - s\|$

This function is convex!

Indeed,

$d_S((1-\lambda)x + \lambda y) \leq \quad \forall s_1, s_2 \in S$

$\|(1-\lambda)x + \lambda y - (1-\lambda)s_1 - \lambda s_2\|$

$\leq (1-\lambda)\|x - s_1\| + \lambda \|y - s_2\|$

$= (1-\lambda) d_S(x) + \lambda d_S(y)$

for proper s_1, s_2 ∎

Now, $x \in S \iff d_S(x) \leq 0$

($\Rightarrow$) take $s = x$

($\Leftarrow$) If $d_S(x) = 0$ $\exists$ sequence $s_i \in S$: $\|x - s_i\| \to 0$

$\Rightarrow s_i \to x \Rightarrow x \in S$ (closedness)

So one can take $g_i(x) = d_S(x)$ in ⊛. But this does not make the description more constructive then $x \succeq 0$.

Recall: $x^3 \le 0$ can be written as $x \le 0$

no! convex constraint

7. Cone programs duality

$$\begin{bmatrix} \min c^T x \\ Ax = b \\ x \in K \end{bmatrix}$$

K - convex cone

$K^* = \{y \mid x^T y \ge 0 \ \forall x \in K\}$ - dual cone

Examples
- $K = \mathbb{R}^n$ $\quad K^* = \{0\}$
- $K = \mathbb{R}^n_{\ge 0}$ $\quad K^* = \mathbb{R}^n_{\ge 0}$
- $K = PSD_n$ $\quad K^* = PSD_n$
- K - Lorentz cone $\quad K^* = K$

Dual program:
$$\begin{bmatrix} \max b^T y \\ \sum y_i a_i + s = c \\ s \in K^* \end{bmatrix}$$

weak duality:

$$c^T x = \left(\sum y_i a_i + s\right)^T x = \sum y_i b_i + s^T x \ge b^T y$$

$s^T x \ge 0$ since $s \in K^*$, $x \in K$

For $\mathbb{R}^n_{\ge 0}$ cone: $P: \begin{bmatrix} \min c^T x \\ Ax = b \\ x \ge 0 \end{bmatrix}$ $\quad D: \begin{bmatrix} \max b^T y \\ \sum y_i a_i + s = c \\ s \ge 0 \end{bmatrix} \iff \begin{bmatrix} \max b^T y \\ \sum y_i a_i \le c \end{bmatrix}$

For SDP cone: $P: \begin{Bmatrix} \min C \cdot X \\ A_i X = b_i \\ X \succeq 0 \end{Bmatrix}$ $\quad D: \begin{bmatrix} \max b^T y \\ \sum A_i y_i + S = C \\ S \succeq 0 \end{bmatrix}$

$A \cdot X := \operatorname{Tr}(A X)$

Exercise: Prove that Lorentz cone is self-dual. Derive SOCP dual form

---

- All the norms are convex, but not strictly convex
- $\|x\|^2$ is strictly convex
- Not true for the squares of all norms

For example, $\|x\|_\infty^2$ is <u>not</u> strictly convex

$x = (1, 0)$
$y = (1, 1)$

- $\|\frac{1}{2}x + \frac{1}{2}y\|_\infty^2 = \|(1, \frac{1}{2})\|_\infty^2 = 1^2 = 1$
- $\frac{1}{2}\|x\|_\infty^2 + \frac{1}{2}\|y\|_\infty^2 = \frac{1}{2}\cdot 1^2 + \frac{1}{2}\cdot 1^2 = 1$

no strict convexity