Convex Optimization
CMU-10725

Semidefinite Programming

Barnabás Póczos & Ryan Tibshirani

MACHINE LEARNING DEPARTMENT
Carnegie Mellon.
School of Computer Science

Semidefinite Programming

Outline

- SDP definition
- SDP basic properties
- SDP applications
- SDP solvers

Papers to Read

- **F. Alizadeh**: Interior point methods in SDP with applications to combinatorial optimization
- **Vandenberghe & Boyd**: Semidefinite Programming
- **L. Lovasz**: Semidefinite programs and combinatorial optimization
- **R. Freund**: Introduction to Semidefinite Programming

Cone of PSD Matrices

Definition [Symmetric matrices, PSD matrices]:

$$S^n = \{ X \in \mathbb{R}^{n \times n} \mid X = X^T \}$$ SYMMETRIC MATRICES

$$S^n_+ = \{ X \in S^n \mid X \succeq 0 \}$$ PSD MATRICES

Lemma [Set of PSD is a closed convex cone]:

S^n_+ IS A CLOSED CONVEX CONE IN $\mathbb{R}^{n \times n}$

DIMENSION = NUMBER OF FREE PARAMETERS: $n(n+1)/2$

Proof:

CONE & CONVEXITY:

$$\left.\begin{array}{l} \alpha, \beta \geq 0 \\ v \in \mathbb{R}^n \\ X, W \in S^n_+ \end{array}\right\} \Rightarrow v^T(\alpha X + \beta W)v = \underbrace{\alpha}_{\geq 0} \underbrace{v^T X v}_{\geq 0} + \underbrace{\beta}_{\geq 0} \underbrace{v^T W v}_{\geq 0} \geq 0$$

Basic Properties of PSD matrices

- $\bullet\ X \in S^n \Rightarrow X = QDQ^T$

  $QQ^T = I$

  D IS DIAGONAL

- $\bullet\ X \succeq 0 \Rightarrow X = QDQ^T$

  $D_{ii} \geq 0$

- $\bullet\ X \succeq 0,\ X_{ii} = 0$ FOR SOME i $\Rightarrow X_{ij} = X_{ji} = 0 \quad \forall j = 1 \dots n$

- $\bullet\ X = \begin{pmatrix} A & B \\ B^T & C \end{pmatrix} \succ 0 \iff \begin{cases} A \succ 0 \\ \underbrace{C - B^T A^{-1} B}_{\text{SCHUR COMPLEMENT}} \succ 0 \end{cases}$

Inner product of PSD matrices

Theorem:

$$A \succeq 0, \; B \succeq 0 \Rightarrow \underbrace{A \bullet B}_{TR(A^T B)} \geq 0$$

Theorem:

LET $A \succeq 0$, $B \succeq 0$ THEN $A \bullet B = 0 \iff AB = 0$

Semidefinite Programming

Definition: **[**Inner product between symmetric matrices**]**

$$X \in S^n, C \in S^n$$

$$\text{LET } C \bullet X = \sum_{i=1}^{n} \sum_{j=1}^{n} C_{ij} X_{ij}$$

LP: $\min c^T x$

$\text{s.t. } a_i^T x = b_i \quad i = 1..m$

$x \geq 0$

Definition: **[**SDP**]**

$$\min_X \; C \bullet X$$

$$\text{s.t. } A_i \bullet X = b_i$$

$$X \succeq 0$$

$X \in \mathbb{R}^{n \times n}$

$A_i \in \mathbb{R}^{n \times n} \quad i = 1 \ldots m$

$b_i \in \mathbb{R}$

$C \in \mathbb{R}^{n \times n}$

LOOKS LIKE LP BUT X MUST LIE IN THE CONE OF PSD MATRICES

SDP Duality

LP duality

(P) $\min_x \; c^T x$ s.t. $Ax = b$, $x \geq 0$

$\max_{y,s} \; y^T b$ s.t. $y^T A + s^T = c^T$, $s \geq 0$, $s \in \mathbb{R}^n$, $y \in \mathbb{R}^m$

Duality gap

$$c^T x - b^T y = y^T A x + s^T x - x^T A^T y = s^T x \geq 0$$

SDP duality

(SDP) $\min \; C \bullet X$ s.t. $A_i \bullet X = b_i \quad i = 1 \ldots m$, $X \succeq 0$

(SDD) $\max_{y,S} \; y^T b$ s.t. $S = C - \sum_{i=1}^{m} y_i A_i$, $S \succeq 0$, $y \in \mathbb{R}^m$, $S \in \mathbb{R}^{n \times n}$

SDP Duality

SDD

$$(SDD) \quad \max_{y,S} \; y^T b$$

$$\text{s.t.} \quad S = C - \sum_{i=1}^{m} y_i A_i$$

$$S \succeq 0$$

$$y \in \mathbb{R}^m \quad S \in \mathbb{R}^{n \times n}$$

SYMMETRIC

Sometimes SDP is defined as

$$\min_{x \in \mathbb{R}^n} c^T x \qquad c \in \mathbb{R}^m$$

$$F_0, F_1, \ldots, F_m \in S^n \quad n \times n$$

$$\text{s.t.} \quad F_0 + x_1 F_1 + \ldots + x_m F_m \succeq 0$$

Linear matrix inequality

SDP is convex problem

Lemma [SDP is convex problem]

$$\min_{x \in \mathbb{R}^n} c^T x \qquad c \in \mathbb{R}^m$$

$$F_0, F_1, \ldots, F_m \in S^n$$

$$\text{S.T.} \quad F_0 + x_1 F_1 + \ldots + x_m F_m \succeq 0$$

Proof: COST FUNCTION $c^T x$ IS CONVEX

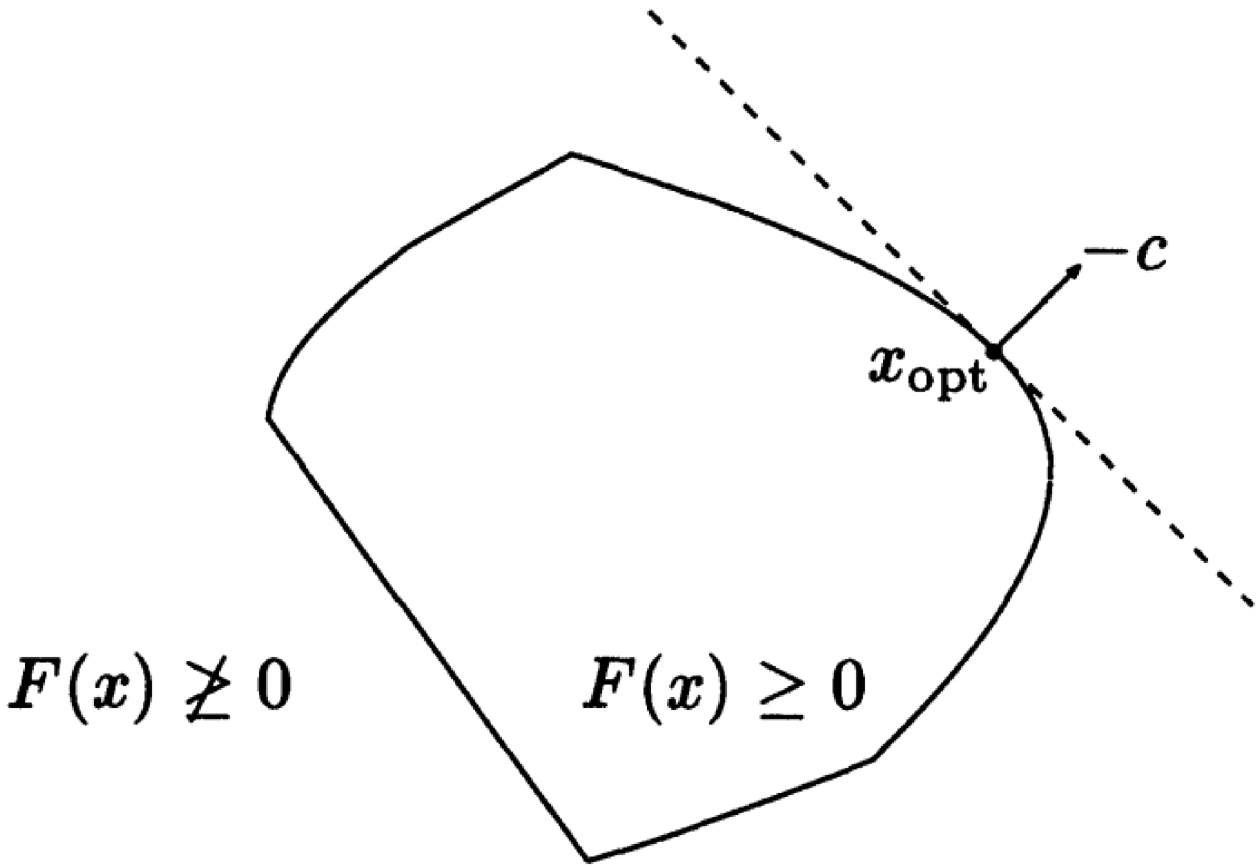
CONSTRAINTS: $F(x) = F_0 + x_1 F_1 + \ldots + x_m F_m$

WE WANT TO SHOW:

$$\left.\begin{array}{l} F(x) \succeq 0 \\ F(y) \succeq 0 \\ \lambda \in [0,1] \end{array}\right\} \Rightarrow \underbrace{F(\lambda x + (1-\lambda) y)}_{= \lambda F(x) + (1-\lambda) F(y)} \succeq 0$$

$$\lambda F(x) + (1-\lambda) F(y) \succeq 0 \quad \square$$

Sketching SDP



Same idea as LP.

The set of feasible points is convex, but not a polytope anymore.

LP as a special SDP

Theorem: [LP as a special SDP]

ANY LP IS A SPECIAL INSTANCE OF AN SDP

Proof:

LP: $\min_x c^T x$ $\quad x \in \mathbb{R}^m \quad c \in \mathbb{R}^m$

$\text{s.t. } Ax + b \geq 0 \quad A \in \mathbb{R}^{n \times m} \quad b \in \mathbb{R}^n$

SDP: $\min_x c^T x \quad \text{s.t. } F_0 + x_1 F_1 + \dots + x_m F_m \succeq 0$

$A = [a_1 \dots a_m] \in \mathbb{R}^{n \times m}$

LET $F_0 = \begin{pmatrix} b_1 & & 0 \\ & \ddots & \\ 0 & & b_n \end{pmatrix} = \mathrm{DIAG}(b)$

$F_i = \begin{pmatrix} a_{i1} & & \\ & \ddots & \\ & & a_{in} \end{pmatrix} = \mathrm{DIAG}(a_i)$

$\Rightarrow F_0 + x_1 F_1 + \dots x_m F_m = \mathrm{DIAG}\left(\{ b_j + x_1 a_{1j} + \dots + x_m a_{mj} \}_{j=1}^{n} \right)$

$= \mathrm{DIAG}(b^T + x^T A^T) = \mathrm{DIAG}(Ax + b)$

SDP that is not LP

SDP example that is not LP:

$$\min_{x \in \mathbb{R}^n} \frac{(c^T x)^2}{d^T x}$$

$$\text{S.T.} \quad Ax + b \geq 0$$

ASSUME THAT IF $Ax + b \geq 0 \Rightarrow d^T x > 0$

This can be rewritten

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

$$\text{S.T.} \quad Ax + b \geq 0$$

$$\frac{(c^T x)^2}{d^T x} \leq t \quad \Leftrightarrow \quad t\, d^T x - (c^T x)^2 \geq 0$$

This can be rewritten

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

$$\text{S.T.} \begin{bmatrix} \mathrm{DIAG}(Ax+b) & 0 & 0 \\ 0 & t & c^T x \\ 0 & c^T x & d^T x \end{bmatrix} \succeq 0$$

LINEAR ELEMENTS IN t, x

SDP weak duality

SDP

$$\min_X C \bullet X$$

$$\text{S.T} \quad A_i \bullet X = b_i \quad i = 1 \ldots m$$

$$X \succeq 0$$

SDD

$$\max_y b^T y \qquad y \in \mathbb{R}^m$$

$$\text{S.T.} \quad \underbrace{C - \sum_{i=1}^m y_i A_i}_{S} \succeq 0$$

Theorem: [Weak duality (gap is nonnegative)**]**

- X IS FEASIBLE OF SDP
- (y, S) IS FEASIBLE OF SDD

$$\Rightarrow \underbrace{C \bullet X - \sum_{i=1}^m y_i b_i}_{S \bullet X} \geq 0$$

Proof: $C \bullet X - \sum_{i=1}^m b_i y_i = C \bullet X - \sum_{i=1}^m \underbrace{(A_i \bullet X)}_{b_i} y_i$

$$= \underbrace{\left(C - \sum_{i=1}^m y_i A_i\right)}_{\succeq 0} \bullet \underbrace{X}_{\succeq 0} = S \bullet X \geq 0$$

$$\boxed{\begin{matrix} A \succeq 0 \\ B \succeq 0 \end{matrix} \Rightarrow A \bullet B \geq 0}$$

If the SDP duality gap is zero...

SDP

$$\min_X \; C \bullet X$$

$$\text{S.T.} \quad A_i \bullet X = b_i \quad i = 1 \ldots m$$

$$X \succeq 0$$

SDD

$$\max_y \; b^T y \qquad y \in \mathbb{R}^m$$

$$\text{S.T.} \quad \underbrace{C - \sum_{i=1}^m y_i A_i}_{S} \succeq 0$$

Theorem: [If the duality gap is zero for feasible points**]**

$$\text{IF} \quad \underbrace{C \bullet X - \sum_{i=1}^m y_i b_i}_{\text{DUALITY GAP}} = 0 \Rightarrow \begin{cases} \bullet\; X \text{ IS SDP OPTIMAL} \\ \bullet\; (y, S) \text{ IS SDD OPTIMAL} \\ \bullet\; SX = 0 \end{cases}$$

$$S \bullet X = TR(S^T X)$$

In SDP however, the duality gap might never attain zero

LP vs SDP

Theorem: [LP vs SDP**]**

IN SDP THE INF $c^T x$ MAY BE FINITE
BUT STILL MIN $c^T x$ CANNOT BE ACHIEVED

Example:

SDP FORM

$$\left.\begin{array}{l}\min_{x_1,x_2} x_1 \\ \text{s.t.} \\ \begin{pmatrix} x_1 & 1 \\ 1 & x_2 \end{pmatrix} \succeq 0 \end{array}\right\} \Leftrightarrow \left\{\begin{array}{l}\min_{x_1,x_2} x_1 \\ x_1 \cdot \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} + x_2 \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} + \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \succeq 0 \end{array}\right.$$

$\Rightarrow \begin{array}{l} x_1 \geq 0 \\ x_1 x_2 \geq 1 \end{array} \Rightarrow$ MIN $x_1 = 0$
BUT $x_1 = 0$ CAN'T BE ACHIEVED
BECAUSE $x_1 x_2 \geq 1$

SDP Duality

Theorem: **[**Strong duality**]**

ASSUME THAT X IS FEASIBLE OF SDP

(y, S) IS FEASIBLE OF SDD

$X \succ 0$ STRICTLY FEASIBLE

$S \succ 0$ STRICTLY FEASIBLE

LET $z_P^* =$ OPTIMUM FUNCTION VALUE OF SDP

$z_D^* =$ — 〃 — SDD

$$\Rightarrow \begin{cases} z_P^* = z_D^* \\ z_P^* \text{ CAN BE ATTAINED (INF = MIN)} \\ z_D^* \text{ CAN BE ATTAINED (SUP = MAX)} \end{cases}$$

A weird example

Primal

$$\min x_1$$

$$\text{s.t.} \begin{pmatrix} 0 & x_1 & 0 \\ x_1 & x_2 & 0 \\ 0 & 0 & x_1+1 \end{pmatrix} \succeq 0$$

$$x_1 \underbrace{\begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}}_{-A_1} + x_2 \underbrace{\begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}}_{-A_2} + \underbrace{\begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}}_{C} \succeq 0$$

$$\max_x \; x^T b \qquad \boxed{P}$$

$$\text{s.t. } C - \sum_{i=1}^{2} x_i A_i \succeq 0$$

$$\min x_1 = \max -x_1$$

$$b_1 = -1$$

$$b_2 = 0$$

Dual

$$\min \underbrace{\begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} \bullet Y}_{} = Y_{33} \Leftrightarrow \max -Y_{33}$$

$$\text{s.t.} \underbrace{\begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} \bullet Y}_{Y_{12}+Y_{21}+Y_{33}} = 1$$

$$Y_{12} + Y_{21} + Y_{33} = 1$$

$$\begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix} \bullet Y = 0 \Rightarrow Y_{22} = 0$$

$$\min_Y \; C \bullet Y \qquad \boxed{D}$$

$$\text{s.t. } A_i \bullet Y = b_i$$

$$i = 1, 2$$

$$Y \succeq 0$$

A weird example

Primal $\quad \text{MIN } x_1$

$$\text{SUBJECT TO} \begin{pmatrix} 0 & x_1 & 0 \\ x_1 & x_2 & 0 \\ 0 & 0 & x_1+1 \end{pmatrix} \succeq 0$$

Feasible primal solutions: $x_1 = 0,\ x_2 \geq 0$

Primal optimal value: $z_P^* = \text{MIN } x_1 = 0$

Dual $\quad \text{MAX } -Y_{33}$

$$\text{S.T.} \quad Y_{12} + Y_{21} + Y_{33} = 1$$
$$Y_{22} = 0$$
$$Y \succeq 0$$

Feasible dual solutions: $\begin{pmatrix} a & 0 & b \\ 0 & 0 & 0 \\ b & 0 & 1 \end{pmatrix} \succeq 0 \quad \text{i.e. } a \geq b^2$

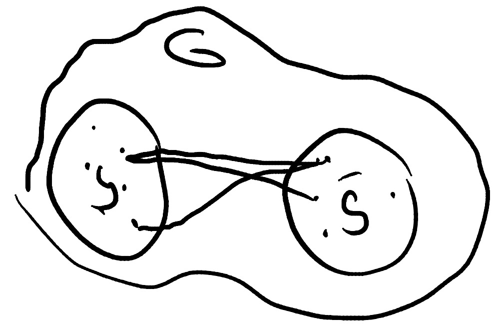
Dual optimal value: -1

The duality gap is not zero in the optimum points!

SDP Applications

SDP in Combinatorial Optimization

Definition: [MAXCUT problem]



- G IS UNDIRECTED GRAPH
- NODES $N = \{1, \dots, n\}$
- EDGE SET E
- w_{ij} WEIGHT ON EDGE $(i,j) \in E$
- GOAL OF MAXCUT:

  FIND A SUBSET $S \subset N$, $\bar{S} \doteq N \setminus S$

  SUCH THAT SUM OF WEIGHT OF EDGES FROM S TO $\bar{S}$ IS MAXIMIZED

LET $x_j = 1$ IF $j \in S$, $x_j = -1$ IF $j \in \bar{S}$

MAXCUT: $\max_{x} \sum_{i=1}^{n} \sum_{j=1}^{n} w_{ij} (1 - x_i x_j)$ s.t. $x_j \in \{-1, +1\}$

MAXCUT problem rewritten

$$\max_{x} \sum_{i=1}^{n} \sum_{j=1}^{n} w_{ij} (1 - x_i x_j)$$

NP HARD INTEGER PROGRAM

$$\text{s.t.} \quad x_j \in \{-1, 1\} \quad j = 1 \ldots n$$

LET $Y = x x^T$ $\quad Y_{ij} = x_i x_j$

MAXCUT:

$$\max_{x, Y} \sum_{i=1}^{n} \sum_{j=1}^{n} w_{ij} - \overbrace{\sum_{i=1}^{n} \sum_{j=1}^{n} w_{ij} x_i x_j}^{W \bullet Y}$$

$$\text{s.t.} \quad x_j \in \{-1, 1\} \quad j = 1 \ldots n$$

$$Y = x x^T$$

MAXCUT

$$\max_{Y, x} \sum_{i=1}^{n} \sum_{j=1}^{n} W_{ij} - W \bullet Y$$

$$\text{s.t.} \quad Y_{jj} = 1 \quad j = 1 \ldots n$$

$$Y = x x^T$$

[FROM THIS IT FOLLOWS THAT $x_j \in \{-1, 1\}$

Relaxation of Maxcut with SDP

MAXCUT: $$\max_{Y,x} \sum_{i=1}^{n}\sum_{j=1}^{n} w_{ij} - W \bullet Y$$

$$\text{s.t.} \quad Y_{jj} = 1 \quad j = 1 .. n$$

$$Y = xx^T$$

DIFFICULT BECAUSE OF THE RANK 1 CONSTRAINT

RELAX: $$\max_{Y} \sum_{i=1}^{n}\sum_{j=1}^{n} w_{ij} - W \bullet Y$$

$$\text{s.t} \quad Y_{jj} = 1$$

$$Y \succeq 0$$

SDP CAN BE SOLVED IN POLY TIME

$$\text{MAXCUT} \leq \text{RELAX}$$

Max Cut hardness

Theorem: [Negative result]

- IT IS NP HARD TO FIND A CUT WITH MORE THAN $\frac{16}{17}$ MAXCUT EDGES

≈ 0.94

Theorem: [Positive result, Goemans & Williamson]

IN POLYNOMIAL TIME ONE CAN FIND A CUT WITH AT LEAST 0.878 EDGES

QCQP

(Convex) Quadratically Constrained Quadratic Programming

MIN $f_0(x)$

SUBJECT TO $f_i(x) \leq 0 \quad i = 1 \dots L$

$f_0(x) = (Ax + b)^T(Ax + b) - c^T x - d \qquad x \in \mathbb{R}^n$

$f_i(x) = (A_i x + b_i)^T(A_i x + b_i) - c_i^T x - d_i \leq 0$

Observation 1

$$(A_i x + b_i)^T(A_i x + b_i) - c_i^T x - d_i \leq 0$$

$$\Leftrightarrow \begin{bmatrix} I & A_i x + b_i \\ (A_i x + b_i)^T & c_i^T x + d_i \end{bmatrix} \succeq 0$$

Observation 2

$$\left.\begin{array}{l} \min\limits_{x} \ f_0(x) \\ \text{s.t.} \ f_i(x) \leq 0 \quad i = 1 \dots L \end{array}\right\} \Leftrightarrow \left\{\begin{array}{ll} \min\limits_{t,x} \ t & \\ \text{s.t.} & f_0(x) \leq t \\ & f_i(x) \leq 0 \quad i = 1 \dots L \end{array}\right.$$

QCQP

(Convex) QCQP as SDP

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

$$\text{s.t.} \quad \begin{bmatrix} I & A_0 x + b_0 \\ (A_0 x + b_0)^T & c_0^T x + d_0 + t \end{bmatrix} \succeq 0$$

$$\begin{bmatrix} I & A_i x + b_i \\ (A_i x + b_i)^T & c_i^T x + d_i \end{bmatrix} \succeq 0 \quad i = 1 \ldots L$$

$x \in \mathbb{R}^k$

$t \in \mathbb{R}$

$A_i \in \mathbb{R}^{n_i \times k}$

$b_i \in \mathbb{R}^{n_i}$

$c_i \in \mathbb{R}^k$

Nonconvex QCQP

Nonconvex QCQP is NP hard

Theorem:

Any integer program is a (nonconvex) QCQP

Proof:

$$x_i \in \{0,1\} \iff x_i(x_i-1)=0 \iff \begin{array}{l} x_i(x_i-1) \geq 1 \\ x_i(x_i-1) \leq 1 \end{array}$$

SDP for Eigenvalue Optimization

Min Max eigenvalue:

$A(x) = A_0 + x_1 A_1 + \dots + x_2 A_2$, WHERE $A_i = A_i^T \in \mathbb{R}^{p \times p}$

$$\min_{x} \underbrace{\max_{i} \lambda_i(A(x))}_{\text{MAX EIGENVALUE}}$$

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

SUBJECT TO $tI - A(x) \succeq 0$

Min Spectral norm $\|A\|_2 = \max_{x \neq 0} \frac{\|Ax\|}{\|x\|} = \sqrt{\max_i \lambda_i(A^T A)}$

$$\min_{x} \|A(x)\|_2 \Leftrightarrow \min_{x,t} \; t$$

S.T. $\begin{bmatrix} tI & A(x) \\ A(x)^T & tI \end{bmatrix} \succeq 0$

Log Chebyschev Approximation

Chebyschev approximation

$$\min_x \max_{i=1..p} |a_i^T x - b_i|$$

$$\Leftrightarrow \text{LP}: \quad \min t \quad \text{s.t.} \quad -t \le a_i^T x - b_i \le t \quad i=1..p$$

Logarithmic Chebyschev approximation

ASSUMING $b_i > 0$
DEF $\log(x) = -\infty$ IF $x < 0$

$$\min_x \max_{i=1..p} \underbrace{|\log(a_i^T x) - \log b_i|}_{\log \max\left(\frac{a_i^T x}{b_i},\ \frac{b_i}{a_i^T x}\right)}$$

$$\Leftrightarrow \min t \quad \text{s.t.} \quad 1/t \le \frac{a_i^T x}{b_i} \le t \quad i=1..p$$

$$\Leftrightarrow \min t \quad \text{s.t.} \quad \begin{bmatrix} t - a_i^T x / b_i & 0 & 0 \\ 0 & a_i^T x / b_i & 1 \\ 0 & 1 & t \end{bmatrix} \succeq 0 \quad i=1..p$$

SDP

Solving SDP

SDP Algorithms

- No simplex algorithm
  (the domain is not polytope anymore)
- Ellipsoid method
- Barrier methods
- and many more…

Barrier methods for SDP

SDP: $\min \; C \bullet X$

$$\text{s.t. } A_i \bullet X = b_i \qquad i = 1 \ldots m$$

$$X \succeq 0 \qquad X \in \mathbb{R}^{n \times n}$$

We need barrier function for the inequality constraint:

$$X \succeq 0 \quad \text{i.e. } X \in S_+^n$$

$$\text{INT } S_+^n = \{ X \in S^n \mid \lambda_1(X) > 0, \ldots \lambda_n(X) > 0 \}$$

A natural barrier function:

$$-\sum_{j=1}^{n} \log(\lambda_i(X)) = -\log\left(\prod_{j=1}^{n} \lambda_i(X)\right) = -\log(\det(X))$$

Barrier SDP

BSDP

$$BSDP(\mu): \quad \min_X \overbrace{C \bullet X - \mu \log(\det(X))}^{f(X)}$$

$$\text{s.t.} \quad A_i \bullet X = b_i \quad i = 1 \ldots m$$

$$X \succ 0$$

Repeat the steps we did with primal dual LP!

Lagrange function:

$$L(X, y) = C \bullet X - \mu \log(\det(X)) + \sum_{i=1}^{m} y_i (b_i - A_i \bullet X)$$

KKT stationarity condition:

$$\frac{\partial L(X, y)}{\partial X} = 0$$

Lagrange function

$$L(X,y) = C \bullet X - \mu \log(\det(X)) + \sum_{i=1}^{m} y_i (b_i - A_i \bullet X)$$

KKT stationarity condition: $\dfrac{\partial L(X,y)}{\partial X} = 0$

Derivatives:

$$\frac{\partial\, C \bullet X}{\partial X} = C, \qquad \frac{\partial \log \det(X)}{\partial X} = (X^T)^{-1} = X^{-1}$$

$$\frac{\partial \sum_{k,l} C_{kl} X_{kl}}{\partial X_{ij}} = C_{ij} \qquad \frac{\partial\, y_i (b_i - A_i \bullet X)}{\partial X} = -y_i A_i$$

KKT stationarity condition: $C - \mu X^{-1} = \sum_{i=1}^{m} y_i A_i$

KKT conditions

$$A_i \bullet X = b_i \quad i = 1 \ldots m$$

$$X \succeq 0$$

$$C - \mu X^{-1} = \sum_{i=1}^{m} y_i A_i$$

Tricks similarly to the LP case:

$$X \succeq 0 \Rightarrow X = LL^T$$

$$\text{LET} \quad S \doteq \mu \underbrace{X^{-1}}_{(LL^T)^{-1} = L^{-T} L^{-1}} = C - \sum_{i=1}^{m} y_i A_i \succeq 0 \quad \left[\begin{array}{l} \text{LP CASE} \\ S = \mu D_x^{-1} e \end{array}\right]$$

$$\Rightarrow \frac{1}{\mu} L^T S L = I$$

Solving SDP

$$\left.\begin{aligned} &A_i \bullet X = b_i \quad i = 1 \dots m \\ &X = LL^T \\ &\sum_{i=1}^{m} y_i A_i + S = C \\ &I - \frac{1}{\mu} L^T S L = 0 \end{aligned}\right\} \text{SOLVE IT FOR } L, X, y, S$$

Algorithm:

1. This is a nonlinear system:
   Use (one step of) Newton method to find an approximate solution
2. Update L, X, y, S
3. Decrease μ, and go back to 1.

Duality gap:

$$S \bullet X = TR(\underset{\uparrow}{S}^T \underset{\uparrow}{X})_{\text{SYMMETRIC}} = \sum_{j=1}^{n} \underbrace{(SX)_{jj}}_{\mu} = n\mu \quad \left[\text{SAME AS THE LP CASE}\right]$$

Summary

- SDP definition
- SDP basic properties
- SDP applications
- SDP solvers