Tutorial on semidefinite programming (SDP)
Uniandes/Externado, Spring 2006

Instructor:

Javier Peña, Carnegie Mellon Univ.

Goals:

- Basic theory of semidefinite and second-order programming
- Some applications
- Acquaintance with some popular solvers
- Current challenges/trends

Course outline:

- SDP: generalization of linear programming (LP)
- Examples of SDP applications
- SDP theory: duality, complementarity
- Second-order programming (SOCP)
- SOCP/LP/SDP conic programming
- Examples of SOCP/LP/SDP applications
- Solvers: `SeDuMi`, `SDPT3`
- Polynomials and sum-of-squares (SOS), copositive matrices
- Symmetric, homogeneous, hyperbolic cones

Brief and incomplete history of SDP

- Eigenvalue optimization, LMI problems (1960s – 1970s)
- Lovász theta function (1979) in information theory
- Interior-point algorithms for SDP by Alizadeh, and by Nesterov & Nemirovski (1980s, 1990s)
- Intense development of theory, algorithms, applications (1990s)
- Extension to symmetric cones, development of general-purpose solvers (1990s)
- Currently: homogeneous, hyperbolic, polynomial programming. Solution to large problems.

Preamble: Some stuff about symmetric matrices

$\mathbf{R}^n$: n-dim Euclidean vector space
$\mathbf{R}^{n\times n}$: $n \times n$ matrices with entries in $\mathbf{R}$
$\mathbf{S}^n$: $n \times n$ symm matrices with entries in $\mathbf{R}$
$\mathbf{O}^n$: $n \times n$ orthogonal matrices

Thm (spectral decomposition). If $X \in \mathbf{S}^n$ then there exist $\lambda(X) = (\lambda_1(X), \ldots, \lambda_n(X)) \in \mathbf{R}^n$ and $Q \in \mathbf{O}^n$ such that

$$X = Q\,\mathrm{Diag}(\lambda(X))\,Q^{\mathsf{T}}.$$

Defn: Let $X \in \mathbf{S}^n$ be given.
X is *positive semidefinite* if $\lambda(X) \geq 0$ denoted as $X \succeq 0$
X is *positive definite* if $\lambda(X) > 0$, denoted as $X \succ 0$

Prop. Let $X \in \mathbf{S}^n$. Then

$$\begin{aligned} X \succeq 0 &\Leftrightarrow \lambda_{\min}(X) \geq 0 \\ &\Leftrightarrow \exists R \in \mathbf{S}^n \text{ s.t. } X = R^2 \\ &\Leftrightarrow \exists L \in \mathbf{R}^{n\times n} \text{ s.t. } X = LL^\top \\ &\Leftrightarrow \forall u \in \mathbf{R}^n \; u^\top X u \geq 0. \end{aligned}$$

Prop.

(a) If $M \in \mathbf{R}^{m\times n}$ then $\|M\|^2 = \lambda_{\max}(M^\top M) = \lambda_{\max}(MM^\top)$.

(b) If $M \in \mathbf{S}^n$ then $\|M\| = |\lambda_{\max}(M)|$.

Positive semidefinite cone:

$$\mathbf{S}^n_+ := \{X \in \mathbf{S}^n : X \succeq 0\}.$$

SDP as a generalization of linear programming (LP)

LP:

$$\begin{array}{ll} \max_y & b^\top y \\ \text{s.t.} & y_1 a_1 + \cdots + y_m a_m \leq c, \end{array}$$

where $c, a_1, \ldots, a_m \in \mathbf{R}^n$, $b \in \mathbf{R}^m$ are given and $y \in \mathbf{R}^m$.

SDP:

$$\begin{array}{ll} \max_y & b^\top y \\ \text{s.t.} & y_1 A_1 + \cdots + y_m A_m \preceq C \end{array}$$

where $C, A_1, \ldots, A_m \in \mathbf{S}^n$, $b \in \mathbf{R}^m$ are given and $y \in \mathbf{R}^m$.

Here $M \preceq N \Leftrightarrow N - M \succeq 0$.

Examples of SDP.

Example 1 (eigenvalue optimization): Suppose $A_0, A_1, \ldots A_m \in \mathbf{S}^n$ and want to solve

$$\min_y \lambda_{\max}(A_0 + y_1 A_1 + \cdots + y_m A_m).$$

Can reformulate as

$$\begin{array}{ll} \min & t \\ \text{s.t.} & A_0 + y_1 A_1 + \cdots + y_m A_m \preceq tI \end{array}$$

Example 2 (norm minimization): Suppose $A_0, A_1, \ldots A_m \in \mathbf{R}^{p\times q}$ and want to solve

$$\min_y \|A(y)\|,$$

for $A(y) := A_0 + y_1 A_1 + \cdots + y_m A_m$.

Can reformulate as

$$\begin{array}{ll} \min & t \\ \text{s.t.} & \begin{bmatrix} tI & A(y) \\ A(y)^\top & tI \end{bmatrix} \succeq 0 \end{array}$$

The formulation relies on the following lemma.

Schur Complement Lemma: Suppose B, D are symmetric and $B \succ 0$. Then

$$\begin{bmatrix} B & C \\ C^\top & D \end{bmatrix} \succeq 0 \Leftrightarrow D - C^\top B^{-1} C \succeq 0.$$

LP primal and dual forms

Recall that the dual of

$$\begin{array}{rl} \max & b^\top y \\ \text{s.t.} & y_1 a_1 + \cdots + y_m a_m \le c \end{array}$$

is

$$\begin{array}{rl} \min & c^\top x \\ \text{s.t.} & a_i^\top x = b_i,\ i = 1, \ldots, m \\ & x \ge 0. \end{array}$$

The dual can be seen as the "best" upper bound on the primal problem obtained by combining the primal constraints.

SDP primal and dual forms

The dual of

$$\begin{array}{rl} \max & b^\top y \\ \text{s.t.} & y_1 A_1 + \cdots + y_m A_m \preceq C \end{array}$$

is

$$\begin{array}{rl} \min & C \bullet X \\ \text{s.t.} & A_i \bullet X = b_i,\ i = 1, \ldots, m \\ & X \succeq 0. \end{array}$$

Can do something similar for SDP.

Endow $\mathbf{S}^n$ with the following inner product: for $X, S \in \mathbf{S}^n$

$$X \bullet S := \mathrm{trace}(XS)$$

Recall: for $X \in \mathbf{R}^{n \times n}$

$$\begin{aligned} \mathrm{trace}(X) &= \sum_{i=1}^n \lambda_i(X) \\ &= \text{coefficient of } -\lambda^{n-1} \text{ in } \det(\lambda I - X) \\ &= \sum_{i=1}^n X_{ii}. \end{aligned}$$

Notice: If $X \in \mathbf{S}^n_+$ then $\mathrm{trace}(X) \ge 0$.

Frobenius norm $\|X\|_F := (X \bullet X)^{1/2}$.

Fact: For $U, V, W \in \mathbf{R}^{n \times n}$, $\mathrm{trace}(UVW) = \mathrm{trace}(VWU)$.

More examples/applications of SDP

Example 3 (positive polynomials).

$p(t) = p_1 + p_2 t + \cdots + p_{2d+1} t^{2d}$ satisfies $p(t) \ge 0 \ \forall t \in \mathbf{R}$ if and only if $\exists X \in \mathbf{S}^{d+1}_+$ such that

$$p_i = \sum_{j+k=i+1} X_{jk},\ i = 1, \ldots, 2d+1.$$

The equivalence relies on the following classical result.

Markov-Lucacs Thm. Let $p(t)$ be a $2d$-degree polynomial. Then

$$p(t) \ge 0\ \forall t \in \mathbf{R} \Leftrightarrow p(t) = q(t)^2 + r(t)^2$$

for some d-degree polynomials $q(t), r(t)$.

Equivalence SDP/SOS:

Suppose $q(t) = q_1 + q_2 t + \cdots + q_{d+1} t^d$. Can write $q(t)$ as

$$\begin{bmatrix} q_1 & \cdots & q_{d+1} \end{bmatrix} \begin{bmatrix} 1 \\ t \\ \vdots \\ t^d \end{bmatrix}.$$

Therefore

$$p(t) = q(t)^2 \Leftrightarrow \begin{bmatrix} p_1 & \cdots & p_{2d+1} \end{bmatrix} \begin{bmatrix} 1 \\ \vdots \\ t^{2d} \end{bmatrix} = \begin{bmatrix} 1 & \cdots & t^d \end{bmatrix} qq^\top \begin{bmatrix} 1 \\ \vdots \\ t^d \end{bmatrix}.$$

Hence $p(t) \geq 0 \; \forall t \in \mathbf{R}$ iff $p(t)$ is SOS iff for some $X \succeq 0$

$$\begin{bmatrix} p_1 & \cdots & p_{2d+1} \end{bmatrix} \begin{bmatrix} 1 \\ \vdots \\ t^{2d} \end{bmatrix} = \begin{bmatrix} 1 & \cdots & t^d \end{bmatrix} X \begin{bmatrix} 1 \\ \vdots \\ t^d \end{bmatrix}$$

But latter identity is the same as

$$p_i = \sum_{j+k=i+1} X_{jk}, \; i = 1, \ldots, 2d+1.$$

It follows that given a $2d$-degree polynomial $p(t)$ the problem of computing

$$p^* := \min_t p(t)$$

can be formulated as an SDP. (How?)

Example 4 (Lovász's theta function): Let $G = (N, E)$ be an undirected graph, where $N = \{1, \ldots, n\}$. Put:

$\omega(G)$: clique number of G

$\chi(G)$: chromatic number of G.

Notice that $\omega(G) \leq \chi(G)$.

Define

$$\begin{array}{rcll} \vartheta(\bar{G}) & := & \max & ee^\top \bullet X \\ & & & I \bullet X = 1 \\ & & & X_{ij} = 0, \; ij \notin E \\ & & & X \succeq 0. \end{array}$$

Thm (Lovász): $\omega(G) \leq \vartheta(\bar{G}) \leq \chi(G)$.

Proof of $\omega(G) \leq \vartheta(\bar{G})$.

Assume K is a clique in G. Let $\chi_K \in \{0,1\}^n$ be the "indicator" vector of S. Consider

$$X := \frac{1}{|K|} \chi_K \chi_K^\top.$$

Observe that X is feasible for the SDP and

$$ee^\top \bullet X = e^\top X e = \frac{(e^\top \chi_K)^2}{|K|} = \frac{|K|^2}{|K|} = |K|.$$

Thus $\vartheta(G) \geq |K|$. $\square$

Example 5 (Max-cut): Let $G = (N, E)$ be an undirected graph, where $N = \{1, \ldots, n\}$. Assume have edge-weights $w = (w_{ij}) \in \mathbf{R}^E_+$. For $K \subseteq N$, $\delta(K) := \{ij \in E : i \in K, j \notin K\}$.

MAX-CUT problem:

$$\max_{K \subseteq N} \sum_{ij \in \delta(K)} w_{ij}$$

Can formulate MAX-CUT as

$$\begin{array}{ll} \max & \frac{1}{4} \sum_{i,j=1}^{n} w_{ij}(1 - x_i x_j) \\ \text{s.t.} & x_i \in \{-1, 1\},\ i \in N. \end{array}$$

Putting $X = xx^{\mathsf{T}}$, can reformulate the latter as

$$\begin{array}{ll} \max & \frac{1}{4} \sum_{i,j=1}^{n} w_{ij}(1 - X_{ij}) \\ \text{s.t.} & X_{ii} = 1,\ i \in N \\ & X \succeq 0 \\ & X \ \text{rank-one}. \end{array}$$

SDP relaxation

$$\begin{array}{ll} \max & \frac{1}{4} \sum_{i,j=1}^{n} w_{ij}(1 - X_{ij}) \\ \text{s.t.} & X_{ii} = 1,\ i \in N \\ & X \succeq 0. \end{array}$$

Thm (Goemans & Williamson): $0.87856 \le \frac{\text{MAX-CUT}}{\text{SDP}} \le 1$.

$\alpha = 0.87856 := \min_{t \in [-1,1]} \frac{2 \arccos(t)}{\pi(1-t)}$.

Proof. Let X be feasible for the SDP relaxation. Then $X = V^{\mathsf{T}} V$ for some $V = \begin{bmatrix} v_1 & \cdots & v_n \end{bmatrix} \in \mathbf{R}^{n \times n}$, where each $\|v_i\| = 1$.

Construct a cut as follows: Pick $v \in \mathbf{R}^n$, $\|v\| = 1$ uniformly at random and put

$$K := \left\{ i \in N : v^{\mathsf{T}} v_i \ge 0 \right\}.$$

It turns out that the expected value of this cut is

$$W[X] = \frac{1}{2} \sum_{i,j=1}^{n} \frac{w_{ij} \arccos X_{ij}}{\pi} \ge \frac{\alpha}{4} \sum_{i,j=1}^{n} w_{ij}(1 - X_{ij}).$$

It follows that MAX-CUT $\ge \alpha\,$SDP.

The inequality MAX-CUT $\le$ SDP is immediate. $\square$

Example 6

$$\begin{array}{ll} \max & x^{\mathsf{T}} Q x \\ \text{s.t.} & x_i \in \{-1, 1\},\ i = 1, \ldots, n. \end{array}$$

SDP relaxation

$$\begin{array}{ll} \max & Q \bullet X \\ \text{s.t.} & X_{ii} = 1,\ i = 1, \ldots, n \\ & X \succeq 0. \end{array}$$

Thm (Nesterov). If $Q \succeq 0$ then

$$\frac{2}{\pi} \le \frac{\text{OPT}}{\text{SDP}} \le 1.$$

There are more schemes to get SDP-relaxations of combinatorial problems.

Duality and complementarity

Recall LP duality: For conciseness put

$$A := \begin{bmatrix} a_1 & \cdots & a_m \end{bmatrix}^\top.$$

Primal-dual pair

$$\text{(P)} \quad \begin{array}{ll} \min & c^\top x \\ \text{s.t.} & Ax = b \\ & x \geq 0, \end{array} \qquad \text{(D)} \quad \begin{array}{ll} \max & b^\top y \\ \text{s.t.} & A^\top y + s = c \\ & s \geq 0. \end{array}$$

Prop (weak duality). If x is (P)-feas, and (y, s) is (D)-feasible then

$$b^\top y \leq c^\top x.$$

Thm (strong duality). If either (P) or (D) is feasible and bounded, then both have optimal solutions. In that case x and (y, s) solve (P) and (D) respectively iff

$$b^\top y = c^\top x \;\Leftrightarrow\; x^\top s = 0.$$

Prop (complementarity). Let $x, s \in \mathbf{R}^n_+$. Then

$$x^\top s = 0 \;\Leftrightarrow\; x_i s_i = 0,\; i = 1, \ldots, n.$$

Given $x, s \in \mathbf{R}^n$ define $x \circ s \in \mathbf{R}^n$ as

$$(x \circ s)_i = x_i s_i,\; i = 1, \ldots, n.$$

Can recast (P) and (D) as

$$\begin{aligned} A^\top y + s &= c \\ Ax &= b \\ x \circ s &= 0 \\ x, s &\geq 0. \end{aligned}$$

Simplex method: maintain first three conditions; aim for the last one.

Interior-point methods: maintain first two and last one, aim for the third one.

Similar duality results for SDP?

For conciseness let $\mathcal{A} : \mathbf{S}^n \to \mathbf{R}^m$ be the mapping

$$X \mapsto \begin{bmatrix} A_1 \bullet X & \cdots & A_m \bullet X \end{bmatrix}^\top.$$

The adjoint is $\mathcal{A}^* : \mathbf{R}^m \to \mathbf{S}^n$ is defined by

$$y \mapsto y_1 A_1 + \cdots + y_m A_m.$$

Consider the SDP primal-dual pair.

$$\text{(P)} \quad \begin{array}{ll} \min & C \bullet X \\ \text{s.t.} & \mathcal{A}X = b \\ & X \succeq 0, \end{array} \qquad \text{(D)} \quad \begin{array}{ll} \max & b^\top y \\ \text{s.t.} & \mathcal{A}^* y + S = C \\ & S \succeq 0. \end{array}$$

Prop (weak duality). If X is (P)-feas, and (y, S) is (D)-feasible then

$$b^\top y \le C \bullet X.$$

Proof.

$$C \bullet X - b^\top y = (\mathcal{A}^* y + S) \bullet X - (\mathcal{A}X)^\top y = S \bullet X.$$

Because $S \in \mathbf{S}^n_+$, can put $S = LL^\top$ for some $L \in \mathbf{R}^{n \times n}$.

Hence, since $X \in \mathbf{S}^n_+$

$$S \bullet X = \text{trace}(LL^\top X) = \text{trace}(L^\top X L) \ge 0.$$

$\square$

Strong duality may not necessarily hold...

Example 7. Consider the SDP

$$\begin{array}{ll} \min & X_{12} \\ \text{s.t.} & \begin{bmatrix} 0 & X_{12} & 0 \\ X_{12} & X_{22} & 0 \\ 0 & 0 & 1 + X_{12} \end{bmatrix} \succeq 0 \end{array}$$

and its dual

$$\begin{array}{ll} \max & y_1 \\ \text{s.t.} & \begin{bmatrix} -y_2 & \frac{1+y_1}{2} & -y_3 \\ \frac{1+y_1}{2} & 0 & -y_4 \\ -y_3 & -y_4 & -y_1 \end{bmatrix} \succeq 0. \end{array}$$

Primal optimal value is 0;
dual optimal value is -1.

Other pathologies may occur: inf/sup may not be attained. One of the problems may be feasible and bounded while the other is infeasible.

Example 8. Consider the SDP

$$\begin{array}{ll} \min & X_{12} \\ \text{s.t.} & \begin{bmatrix} 0 & X_{12} \\ X_{12} & X_{22} \end{bmatrix} \succeq 0 \end{array}$$

and its dual

$$\begin{array}{ll} \max & 0 \\ \text{s.t.} & \begin{bmatrix} -y_1 & 1/2 \\ 1/2 & 0 \end{bmatrix} \succeq 0. \end{array}$$

Primal optimal value is 0; dual is infeasible.

Example 9.

$$\begin{array}{ll} \min & X_{11} \\ \text{s.t.} & \begin{bmatrix} X_{11} & 1 \\ 1 & X_{22} \end{bmatrix} \succeq 0. \end{array}$$

Optimal value is 0, but it is not attained.

Crux of strong duality in LP:
Farkas Lemma: Assume $A \in \mathbf{R}^{m \times n}$, $b \in \mathbf{R}^m$. Then either

$$\exists x \text{ s.t. } Ax = b, x \ge 0$$

or

$$\exists y \text{ s.t. } A^\top y \ge 0,\ b^\top y < 0$$

but not both.

For SDP:
Asymptotic Farkas Lemma: Assume $\mathcal{A} \in L(\mathbf{S}^n, \mathbf{R}^m)$, $b \in \mathbf{R}^m$. Then either

$$b \in \overline{\{\mathcal{A}X : X \succeq 0\}}$$

or

$$\exists y \text{ s.t. } \mathcal{A}^* y \succeq 0,\ b^\top y < 0$$

but not both.

Get strong duality under additional assumptions:

(P) strongly feasible if $b \in \text{int}\{\mathcal{A}X : X \succeq 0\}$.
(D) strongly feasible if $C \in \text{int}\{\mathcal{A}^* y + S : S \succeq 0\}$.

Usual assumption: $\mathcal{A}$ surjective, i.e., $A_1, \ldots, A_m$ linearly independent.

Under such assumption (P) strongly feasible iff $\exists X \succ 0$ s.t. $\mathcal{A}X = b$.

Strong Duality Thm. Assume (P) and (D) are strongly feasible. Then both (P) and (D) have optimal solutions. In that case X and (y, S) are optimal sols to (P) and (D) respectively iff

$$b^\mathsf{T} y = C \bullet X \;\Leftrightarrow\; X \bullet S = 0.$$

Hence under the strong feasibility assumptions can recast (P) and (D) as

$$\begin{aligned} \mathcal{A}^* y + S &= C \\ \mathcal{A}X &= b \\ XS &= 0 \\ X, S &\succeq 0. \end{aligned}$$

Strong Duality Thm follows from Asymptotic Farkas Lemma and the following lemma.

Lemma. If $\mathcal{A}^* z \in \mathbf{S}^n_{++}$ for some $z \in \mathbf{R}^m$ then $\mathcal{A}\mathbf{S}^n_{+} := \{\mathcal{A}X : X \succeq 0\}$ is closed.

SDP Complementarity

Prop (complementarity). Let $X, S \in \mathbf{S}^n_+$. Then $X \bullet S = 0$ iff $XS = 0$ iff there exists $Q \in \mathbf{O}^n$, $\lambda, \omega \in \mathbf{R}^n_+$ such that

$$X = Q\text{Diag}(\lambda)Q^\mathsf{T},\; S = Q\text{Diag}(\omega)Q^\mathsf{T},$$
$$\lambda_i \omega_i = 0,\; i = 1, \ldots, n.$$

References for today's material

- A. Ben-Tal and A. Nemirovski, "Lectures on Modern Convex Optimization," MPS-SIAM Series on Optimization, 2001. Related material available from `http://www2.isye.gatech.edu/~nemirovs/`
- E. de Klerk, "Aspects of Semidefinite Programming," Kluwer Academic Publishers, 2002.
- C. Helmberg, "Semidefinite Programming for Combinatorial Optimization," available from `http://www-user.tu-chemnitz.de/~helmberg/`
- H. Wolkowicz, R. Saigal, L. Vandenberghe eds. "Handbook of Semidefinite Programming," Kluwer International Series, 2000.