An Introduction on SemiDefinite Program
from the viewpoint of computation

Hayato Waki
Institute of Mathematics for Industry, Kyushu University

2015-10-08

Combinatorial Optimization at Work, Berlin, 2015
Contents and Purpose of this lecture

Subject  SemiDefinite Program

Contents

Part I  Formulations & Strong duality on SDP

Part II  Algorithm on SDP – Primal-Dual Interior-Point Methods

Part III  Comments of Computation on SDP


Purpose

- Better understanding for the next lecture (MOSEK on SDP) by Dr. Dahl
- Know the difficulty in solving SDP in Part III

Message  SDP is convex, but also nonlinear
Properties and applications of SDP

**Properties**: SDP is an extension of LP
- Duality Theorem
- Solvable by primal-dual interior-point methods with up to a given tolerance

**Applications**:
- Combinatorial problems, e.g., Max-Cut by Goemans and Williams
- Control theory, e.g., $H_\infty$ control problem
- Lift-and-projection approach for nonconvex quadratic problem
- Lasserre’s hierarchy for polynomial optimization problems and complexity theory
- Embedding problems, e.g., sensor networks and molecular conformation
- Statistics and machine learning, etc...
From LP To SDP

**LP**  Primal and Dual

\[
\begin{align*}
\min_x & \quad c^T x \\
\text{s.t.} & \quad a_j^T x = b_j \quad (\forall j) \\
x & \in \mathbb{R}_+^n
\end{align*}
\]

\[
\begin{align*}
\max_{(y,s)} & \quad b^T y \\
\text{s.t.} & \quad s = c - \sum_{j=1}^m y_j a_j \\
s & \in \mathbb{R}_+^n
\end{align*}
\]

- Minimize/Maximize linear function over the intersection the affine set and \( \mathbb{R}_+^n \)
- \( \mathbb{R}_+^n \) is closed convex cone in \( \mathbb{R}^n \)

**Extension to SDP**

- Extension to the space of symmetric matrices \( \mathbb{S}^n \)

\[c \in \mathbb{R}^n \rightarrow C \in \mathbb{S}^n, \quad a_j \in \mathbb{R}^n \rightarrow A_j \in \mathbb{S}^n\]

- Minimize/Maximize linear function over the intersection the affine set and the set of positive semidefinite matrices
**LP** Primal and Dual

\[
\begin{align*}
\min_x & \quad c^T x \\
\text{s.t.} & \quad a_j^T x = b_j \ (\forall j) \\
& \quad x \in \mathbb{R}_+^n
\end{align*}
\]

\[
\begin{align*}
\max_{(y,s)} & \quad b^T y \\
\text{s.t.} & \quad s = c - \sum_{j=1}^{m} y_j a_j \\
& \quad s \in \mathbb{R}_+^n
\end{align*}
\]

**SDP** Primal and Dual

\[
\begin{align*}
\min_X & \quad C \bullet X \\
\text{s.t.} & \quad A_j \bullet X = b_j \ (\forall j) \\
& \quad X \in S_+^n
\end{align*}
\]

\[
\begin{align*}
\max_{(y,S)} & \quad b^T y \\
\text{s.t.} & \quad S = C - \sum_{j=1}^{m} y_j A_j \\
& \quad S \in S_+^n
\end{align*}
\]

- $S^n$ is the set of $n \times n$ symmetry matrices,
- $S_+^n$ is the set of $n \times n$ symmetry positive semidefinite matrices, and
- $A \bullet X := \sum_{k=1}^{n} \sum_{\ell=1}^{n} A_{k\ell} X_{k\ell}$. 
1. Definition of positive semidefinite matrices

\( \mathbf{X} \in \mathbb{S}^n \) is positive semidefinite if for all \( \mathbf{z} \in \mathbb{R}^n \), \( \mathbf{z}^T \mathbf{X} \mathbf{z} \geq 0 \). Equivalently, all eigenvalues are nonnegative.

Remark

Eigendecomposition (Spectral decomposition); \( \exists \mathbf{Q} \in \mathbb{R}^{n \times n} \) (orthogonal) and \( \exists \lambda_i \geq 0 \) such that

\[
\mathbf{X} = \mathbf{Q} \begin{pmatrix} \lambda_1 & & \\ & \lambda_2 & \\ & & \ddots \\ & & & \lambda_n \end{pmatrix} \mathbf{Q}^T
\]

See textbooks of linear algebra for proof

\( \Rightarrow \exists \mathbf{B} \in \mathbb{R}^{n \times n} \) such that \( \mathbf{X} = \mathbf{B} \mathbf{B}^T \)

2. Zero diagonal for positive semidefinite matrices

For \( \mathbf{X} \in \mathbb{S}^n_+ \), each \( \mathbf{X}_{ii} \) is nonnegative. In addition, if \( \mathbf{X}_{ii} = 0 \) for some \( i \), then \( \mathbf{X}_{ij} = \mathbf{X}_{ji} = 0 \) for all \( j = 1, \ldots, n \).
Example of SDP

\[ \begin{align*}
C &= \begin{pmatrix} 2 \\ 1 \end{pmatrix}, A_1 = \begin{pmatrix} 10 & 4 \\ 4 & -8 \end{pmatrix}, A_2 = \begin{pmatrix} -9 \\ -9 \end{pmatrix}, A_3 = \begin{pmatrix} -9 & 2 \\ -9 & 2 \end{pmatrix}, \\
X &= \begin{pmatrix} x_{11} & x_{12} \\ x_{12} & x_{22} \end{pmatrix}, b = (42, -8, 20)^T
\end{align*} \]

Primal SDP is formulated as follows:

\[
\inf_{X} \begin{cases}
2x_{11} + x_{22} : \\
10x_{11} + 8x_{12} = 42, \\
-18x_{12} + 2x_{22} = 20,
\end{cases}
\begin{pmatrix} x_{11} & x_{12} \\ x_{12} & x_{22} \end{pmatrix} \in S^2_+
\]

(Fortunately) the primal solution is uniquely fixed:

\[ X = \begin{pmatrix} 5 & -1 \\ -1 & 1 \end{pmatrix} \] is positive definite and obj. val. = 11.
Primal SDP is formulated as follows:

\[
\inf_X \left\{ 2x_{11} + x_{22} : \begin{array}{l}
10x_{11} + 8x_{12} = 42, \\
-18x_{12} + 2x_{22} = 20,
\end{array} \right. \quad \begin{pmatrix} x_{11} & x_{12} \\ x_{12} & x_{22} \end{pmatrix} \in \mathbb{S}_+^2 \}
\]

Dual SDP is formulated as follows:

\[
\sup_{(y,S)} \left\{ 42y_1 - 8y_2 + 20y_3 : \begin{pmatrix} 2 - 10y_1 & -4y_1 + 9y_3 \\ -4y_1 + 9y_3 & 1 + 8y_2 - 2y_3 \end{pmatrix} \in \mathbb{S}_+^2 \right. \}
\]

A dual solution is \( (1/5, -37/360, 4/45) \) with the obj. val. = 11.
Application: Computation of lower bounds of nonconvex QP

\[
\theta^* := \inf_x \left\{ x^T Q x + 2c^T x : x^T Q_j x + 2c_j^T x + r_j \leq 0 \ (j = 1, \ldots, m) \right\}
\]

**SDP relaxation**: Add the following constraint and replace \( x_i x_j \rightarrow X_{ij} \):

\[
\begin{bmatrix} 1 \\ x \end{bmatrix} (1, x) \in S^{n+1}_+ \rightarrow X \in S^{n+1}_+
\]

\[
\eta^* := \inf_x \left\{ \begin{pmatrix} 0 & c^T \\ c & Q \end{pmatrix} \bullet X : \begin{pmatrix} r_j & c_j^T \\ c_j & Q_j \end{pmatrix} \bullet X \leq 0, X_{00} = 1, X \in S^{n+1}_+ \right\}
\]

**Remark**
- Handle as SDP
- \( \eta^* \leq \theta^* \)
- binary \( x \in \{0, 1\} \rightarrow x^2 - x = 0 \Rightarrow \text{MIQP with binary variables} = \text{QP} \)
Application: Lasserre’s SDP relaxation for Polynomial Optimization Problems

**POP**: \( f, g_j \) are polynomials on \( x \in \mathbb{R}^n \)

\[
\theta^* := \inf \{ f(x) : g_j(x) \geq 0 \ (j = 1, \ldots, m) \}
\]

**Lasserre’s SDP relaxation**
- Generates a sequence of SDP problems: \( \{ \mathbb{P}_r \}_{r \geq 1} \)
- Optimal value: \( \theta_r \leq \theta_{r+1} \leq \theta^* \ (\forall r) \)
- Under assumptions, \( \theta_r \rightarrow \theta^* \ (r \rightarrow \infty) \)
- \( r = 2, 3, \theta_r \approx \theta^* \) in practice
- Strongly connected to sum of square polynomials
Compared with LP

**Similar points**
- Weak and Strong duality holds
- PDIPM also works in SDP

**Different points**
- SDP may have an irrational optimal solution
  
  \[ \sup_y \left\{ y : \begin{pmatrix} 2 & y \\ y & 1 \end{pmatrix} \in S^2_+ \right\} \]

  Optimal solution \( y = \sqrt{2} \), not rational

  - Finite optimal value, but \( \not\exists \) solutions

  \[ \inf_y \left\{ y_1 : \begin{pmatrix} y_1 & 1 \\ 1 & y_2 \end{pmatrix} \in S^2_+ \right\} \]
Different points (cont’d)

∃ 2 types of infeasibility

(LP) \( \exists y; -A^T y \in \mathbb{R}_+^n, b^T y > 0 \) \iff \text{Primal LP is infeasible}

(SDP) \( \exists y; -A^T y \in \mathbb{S}_+^n, b^T y > 0 \) \implies \text{Primal SDP is infeasible}

Remark: Need to consider the following cases

- Finite optimal value, but no optimal solutions for Primal and/or Dual
- Difficult to detect the infeasibility completely
Duality on SDP

Weak duality for any $X \in \mathcal{F}_P$ and $(y, S) \in \mathcal{F}_D$,

$$C \cdot X \geq b^T y \quad \Rightarrow \quad \theta_P^* \geq \theta_D^*$$

Slater condition: $S^n_{++}$ is the set of positive definite matrices
- Primal satisfies *Slater condition* if $\exists X \in \mathcal{F}_P$ such that $X \in S^n_{++}$
- Dual *Slater condition* if $\exists (y, S) \in \mathcal{F}_D$ such that $S \in S^n_{++}$

Strong duality
- Primal satisfies Slater condition and dual is feasible. Then $\theta_P^* = \theta_D^*$ and dual has an optimal solution.
- Slater condition are required for both primal and dual for theoretical results on PDIPMs
- See survey on SDP for proof
3. Inner products on positive semidefinite matrices

For all $X, S \in \mathbb{S}_+^n$, $X \bullet S \geq 0$. Moreover, $X \bullet S = 0$ iff $XS = O_n$

Proof: \exists B \text{ s.t. } X = BB^T \text{ and } \exists D \text{ s.t. } S = DD^T. \text{ Then}

$$X \bullet S = \text{Trace}(BB^TDD^T) = \text{Trace}(D^TBB^TD) = \text{Trace}((B^TD)^T(B^TD)) \geq 0$$

Moreover, $X \bullet S = 0 \Rightarrow B^TD = O_n \Rightarrow XS = O_n$

Proof of weak duality

In fact, for $X \in \mathcal{F}_P$ and $(y, S) \in \mathcal{F}_D$,

$$C \bullet X - b^T y = \left( C - \sum_{j=1}^m y_j A_j \right) \bullet X = S \bullet X \geq 0$$

because both matrices are positive semidefinite.
Remark of 3 (cont’d)

- \( \mathbf{X} \in \mathcal{F}_P : \) optimal in primal and \( (\mathbf{y}, \mathbf{S}) \in \mathcal{F}_D : \) optimal in dual
- Then, \( \theta_P^* - \theta_D^* = \mathbf{X} \cdot \mathbf{S} = 0 \iff \mathbf{XS} = \mathbf{0}_n \)
- \( \mathbf{XS} = \mathbf{0}_n \) is used in PDIPM
SDP with multiple positive semidefinite cones

\[ \begin{align*}
\text{SDP} & \quad \inf_{X_k} \sum_{k=1}^{N} C^k \cdot X_k \\
\text{s.t.} & \quad \sum_{k=1}^{N} A^k_j \cdot X_k = b_j \quad (j = 1, \ldots, m) \\
& \quad X_k \in \mathbb{S}^{n_k}_+ \quad (k = 1, \ldots, N)
\end{align*} \]

where \( C^k, A^k_j \in \mathbb{S}^{n_k}_+ \)

Example

\[ \begin{align*}
A \cdot X \leq d, X \in \mathbb{S}^n_+ & \Rightarrow A \cdot X + s = d, X \in \mathbb{S}^n_+ \quad \text{and} \quad s \in \mathbb{S}^1_+ (= \mathbb{R}_+) 
\end{align*} \]

Dual

\[ \sup_{y, S_k} \left\{ b^T y : S_k = A^k_0 - \sum_{j=1}^{m} y_j A^k_j \in \mathbb{S}^{n_k}_+ \quad (k = 1, \ldots, N) \right\} \]
Remark

- SDP with $\mathbb{R}^n_+$, Second order cone $\mathbb{L}_n$ and $\mathbb{S}^n_+$ can be handled as SDP and PDIPM works

$$\mathbb{L}_n := \{(x_0, x) \in \mathbb{R}^n : \|x\|_2 \leq x_0\}$$

- Free variable can be accepted

$$A \bullet X + a^T x = d, \ X \in \mathbb{S}^n_+, \ x \in \mathbb{R}^n$$

$$\Rightarrow \quad A \bullet X + a^T x_1 - a^T x_2 = d, \ X \in \mathbb{S}^n_+ \ and \ x_1, x_2 \in \mathbb{R}^n_+$$
Classification of Algorithms for SDP

Algorithms for SDP
- Ellipsoid method
- Interior-point methods
- Bundle method
- First-order methods, etc

Interior-point methods
- Path-following algorithm (= Logarithmic barrier function)
- Potential reduction algorithm
- Self-dual homogeneous embeddings

Path-following algorithm
- Primal
- Dual
- Primal-dual
Path-following method

Optimality conditions: a pair of optimal solutions \((X, y, S)\)
satisfies
\[
\begin{align*}
A_j \cdot X &= b_j, \quad X \in S_n^+, \\
S &= C - \sum_{j=1}^{m} y_j A_j, \quad S \in S_n^+,
\end{align*}
\]
\[
XS = O_n (\iff C \cdot X - b^T y = 0)
\]

Perturbed system: for \(\mu > 0\),
\[
\begin{align*}
A_j \cdot X &= b_j, \quad X \in S_{++}^n, \\
S &= C - \sum_{j=1}^{m} y_j A_j, \quad S \in S_{++}^n,
\end{align*}
\]
\[
XS = \mu I_n
\]

Remark
- for any \(\mu > 0\), \(\exists\) unique solution \((X(\mu), y(\mu), S(\mu))\)
- Central path \(\{(X(\mu), y(\mu), S(\mu)) : \mu > 0\}\) is smooth curve and go to a pair of optimal solutions of primal and dual
- Follows the central path = Path-following method
**Algorithm 1: General framework of path-following method**

**Input:** $(X^0, y^0, S^0) \in \mathcal{F}_P \times \mathcal{F}_D$ such that $X^0, S^0 \in \mathbb{S}_{++}^n$, $\epsilon > 0$, $0 < \theta < 1$ and some parameters

$X \leftarrow X^0$, $y \leftarrow y^0$ and $S \leftarrow S^0$;

while $X \cdot S > \epsilon$ do

  Compute direction $(\Delta X, \Delta y, \Delta S)$ from CPE($\mu$);
  Compute step size $\alpha_P, \alpha_D > 0$;
  $X \leftarrow X + \alpha_P \Delta X$;
  $y \leftarrow y + \alpha_D \Delta y$; $S \leftarrow S + \alpha_D \Delta S$;
  Compute $\mu \leftarrow \theta \mu$;

end

return $(X, y, S)$;

**Remark**

- Infeasible initial guess is acceptable
- # of iteration is polynomial in $n, m$ and $\log(\epsilon)$
- Computational cost = Computation of direction
Computation of direction: Find $(\Delta X, \Delta y, \Delta S)$ such that $X + \Delta X \in \mathcal{F}_P$, $(y + \Delta y, S + \Delta S) \in \mathcal{F}_D$ and

$$
\begin{align*}
A_j \cdot \Delta X &= 0, \\
\Delta S - \sum_{j=1}^{m} \Delta y_j A_j &= O_n, \\
XS + \Delta XS + X\Delta S &= \mu I_n
\end{align*}
$$

Remark

- $\Delta X$ may not be symmetry. So, change $XS = \mu I_n$ by

$$
\frac{1}{2} \left( PXSP^{-1} + P^{-T} SXP^T \right) = \mu I_n,
$$

where $P$ is nonsingular

- Possible choice of $P$

$$
\begin{align*}
P &= S^{1/2} \ (HRVW/KSH/M) \\
P &= X^{-1/2} \ (dual \ HRVW/KSH/M) \\
P &= W^{1/2}, \ W = X^{1/2}(X^{1/2} SX^{1/2})^{-1/2} X^{1/2} \ (NT) \\
P &= \ldots \text{ More than 20 types of directions by Todd}
\end{align*}
$$
## Computational cost in PDIPM

1. **Construction of linear system on $\Delta y$** for HRVW/KSH/M direction,

   $$M \Delta y = (\text{RHS}), \text{where } M = (\text{Trace}(A_i X A_j S^{-1}))_{1 \leq i,j \leq m}$$

   - Use of sparsity in $A_j$ is necessary for computation of $M$
   - Almost the same for other search directions

2. **Solving the linear system**

   - $M$ is dense $\Rightarrow$ takes $O(m^3)$ computation by Cholesky decomposition
   - $M$ is often sparse in SDP relax for POP $\Rightarrow$ sparse Cholesky decomposition works well

After them, $\Delta S = \sum_{j=1}^{m} \Delta y_j A_j$ and obtain $\Delta X$. 
Example: Q is nonsingular and dense. Then $P_1$ is equivalent to $P_2$:

\[
P_1 : \inf_X \left\{ C \cdot X : E_i \cdot X = 1 \ (i = 1, \ldots, n), \ X \in S^n_+ \right\},
\]

\[
P_2 : \inf_X \left\{ (Q^T C Q) \cdot X : (Q^T E_i Q) \cdot X = 1 \ (i = 1, \ldots, n), \ X \in S^n_+ \right\}
\]

where

\[
(E_i)_{pq} = \begin{cases} 
1 & \text{if } p = q = i \\
0 & \text{o.w.} 
\end{cases} \quad (p, q = 1, \ldots, n)
\]
CPU time: Solved by SeDuMi 1.3 on the MacBook Air (1.7 GHz Intel Core i7)

Figure: CPU time on $\mathbb{P}_1$ and $\mathbb{P}_2$
Software

Information from http://plato.asu.edu/ftp/sparse_sdp.html
- SeDuMi, SDPT3 (MATLAB)
- SDPA (C++, MATLAB)
- CSDP (C, MATLAB)
- DSDP (C, MATLAB)
- MOSEK

Remark
- Based on PDIPM for almost all software
- Performance depends on SDP problems

Modelling languages on SDP: they can call the above software
- YALMIP
- CVX
Slater conditions

**Strong duality**
- Require Slater conditions for **Primal** or **Dual**
- PDIPM requires Slater conditions for both **Primal** and **Dual**
- Sufficient conditions for optimal solutions
- If either **Primal** or **Dual** does not satisfy Slater conditions, ...

E.g., Lasserre’s SDP relaxation

\[ \mathbb{P} : \inf_{x} \left\{ x : x^2 - 1 \geq 0, x \geq 0 \right\} \]

- Generate SDP relaxation problems \( \mathbb{P}_1, \mathbb{P}_2, \ldots \),
- Slater condition fails in all SDP relaxation & all optimal values are 0
- SeDuMi and SDPA returns wrong value 1
- All SDP relaxation problems are sensitive to numerical errors in the computation of floating points
E.g., Graph Equipartition

- $G(V, E)$: a weighted undirected graph ⇒ Partition the vertex set $V$ into $L$ and $R$
- the minimum total weight of the cut subject to $|L| = |R|$
- QOP formulation

$$\inf_{x \in \mathbb{R}^n} \left\{ \frac{1}{2} \sum w_{ij}(1 - x_i x_j) : \sum_{i=1}^{n} x_i = 0, x_i^2 = 1 \ (i = 1, \ldots, n) \right\}$$
E.g., Graph Equipartition (cont’d)

- SDP relaxation problem: constant matrices $W$, $E$ and $E_i$

$$\inf_{\mathbf{X} \in \mathbb{S}_+^n} \{ W \bullet \mathbf{X} \mid E \bullet \mathbf{X} = 0, E_i \bullet \mathbf{X} = 1 \}$$

- Since $E \in \mathbb{S}_+^n$, $\not\exists \mathbf{X} \in \mathbb{S}_+^{n \times n}$ s.t. $E \bullet \mathbf{X} = 0$ $\Rightarrow$ Slater cond. fails

- Inaccurate value and/or many iterations

Table: SeDuMi 1.3 with $\epsilon=1.0e-8$

<table>
<thead>
<tr>
<th>SDPLIB</th>
<th>iter</th>
<th>cpusec</th>
<th>duality gap</th>
</tr>
</thead>
<tbody>
<tr>
<td>gpp124-1</td>
<td>30</td>
<td>2.40</td>
<td>-4.63e-05</td>
</tr>
<tr>
<td>gpp250-1</td>
<td>29</td>
<td>10.19</td>
<td>-1.60e-04</td>
</tr>
<tr>
<td>gpp500-1</td>
<td>34</td>
<td>61.58</td>
<td>-1.90e-04</td>
</tr>
<tr>
<td>gpp124-4</td>
<td>40</td>
<td>3.02</td>
<td>-2.14e-08</td>
</tr>
<tr>
<td>gpp500-2</td>
<td>40</td>
<td>76.88</td>
<td>-8.26e-06</td>
</tr>
</tbody>
</table>
E.g., Graph Equipartition (cont’d)

\[
\inf_{X \in S^n_+} \{ W \cdot X \mid E \cdot X = 0, E_i \cdot X = 1 \}
\]

- **Transformation of SDP by** \( V \):

\[
V = \begin{pmatrix}
1 & -1 \\
1 & -1 \\
\vdots & \vdots \\
1 & 1
\end{pmatrix}
\]

- \( X \rightarrow V^{-T}XV^{-1} =: Z \) and \( E \rightarrow VEV^T \)

- Then, \( X \in S^n_+ \iff Z \in S^n_+ \) and \( E \cdot X = 0 \iff Z_{nn} = 0 \)

- Eliminate \( n \)th row and column from transformed SDP \( \Rightarrow \) Slater cond. holds
E.g., Graph Equipartition (cont’d)

Table: Numerical Results by SeDuMi 1.3 with $\epsilon=1.0e-8$.

<table>
<thead>
<tr>
<th>Problems</th>
<th>iter</th>
<th>cpusec</th>
<th>d.gap</th>
<th>Slater fails</th>
<th>iter</th>
<th>cpusec</th>
<th>d.gap</th>
<th>Slater holds</th>
</tr>
</thead>
<tbody>
<tr>
<td>gpp100</td>
<td>30</td>
<td>1.78</td>
<td>-2.46e-07</td>
<td>-4.97e-09</td>
<td>0.73</td>
<td>16</td>
<td></td>
<td></td>
</tr>
<tr>
<td>gpp124-1</td>
<td>30</td>
<td>2.34</td>
<td>-4.63e-05</td>
<td>-1.75e-08</td>
<td>1.12</td>
<td>19</td>
<td></td>
<td></td>
</tr>
<tr>
<td>gpp124-2</td>
<td>26</td>
<td>1.76</td>
<td>-1.41e-06</td>
<td>-1.11e-09</td>
<td>1.03</td>
<td>18</td>
<td></td>
<td></td>
</tr>
<tr>
<td>gpp124-3</td>
<td>30</td>
<td>2.56</td>
<td>-4.41e-07</td>
<td>-3.05e-09</td>
<td>1.01</td>
<td>17</td>
<td></td>
<td></td>
</tr>
<tr>
<td>gpp124-4</td>
<td>40</td>
<td>2.93</td>
<td>-2.14e-08</td>
<td>-9.52e-11</td>
<td>1.09</td>
<td>17</td>
<td></td>
<td></td>
</tr>
<tr>
<td>gpp250-1</td>
<td>29</td>
<td>8.81</td>
<td>-1.60e-04</td>
<td>-1.82e-08</td>
<td>4.71</td>
<td>21</td>
<td></td>
<td></td>
</tr>
<tr>
<td>gpp250-2</td>
<td>29</td>
<td>8.61</td>
<td>-1.49e-05</td>
<td>-9.74e-09</td>
<td>4.19</td>
<td>19</td>
<td></td>
<td></td>
</tr>
<tr>
<td>gpp250-3</td>
<td>34</td>
<td>9.48</td>
<td>-3.97e-07</td>
<td>-8.12e-10</td>
<td>4.08</td>
<td>18</td>
<td></td>
<td></td>
</tr>
<tr>
<td>gpp250-4</td>
<td>35</td>
<td>11.28</td>
<td>-8.80e-07</td>
<td>-7.43e-10</td>
<td>4.37</td>
<td>19</td>
<td></td>
<td></td>
</tr>
<tr>
<td>gpp500-1</td>
<td>34</td>
<td>53.45</td>
<td>-1.90e-04</td>
<td>-2.76e-08</td>
<td>31.49</td>
<td>24</td>
<td></td>
<td></td>
</tr>
<tr>
<td>gpp500-2</td>
<td>40</td>
<td>68.47</td>
<td>-8.26e-06</td>
<td>-2.20e-09</td>
<td>28.98</td>
<td>22</td>
<td></td>
<td></td>
</tr>
<tr>
<td>gpp500-3</td>
<td>28</td>
<td>54.81</td>
<td>-1.00e-05</td>
<td>-2.39e-09</td>
<td>31.35</td>
<td>21</td>
<td></td>
<td></td>
</tr>
<tr>
<td>gpp500-4</td>
<td>28</td>
<td>55.06</td>
<td>-1.02e-06</td>
<td>-8.96e-10</td>
<td>32.06</td>
<td>23</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Comments: If does not satisfy Slater conditions, ...
- PDIPM computes inaccurate values and/or spends many iter.
- But, reduce the size of SDP
A simple (?) transformation generates an SDP in which Slater cond. holds.

More elementary approach:

\[
\begin{align*}
(QOP) & : \inf_{x \in \mathbb{R}^n} \left\{ \frac{1}{2} \sum w_{ij} (1 - x_i x_j) : \sum_{i=1}^{n} x_i = 0, x_i^2 = 1 \right\} \\
(QOP') & : \text{obtained by substituting } x_1 = - \sum_{i=2}^{n} x_i \text{ in (QOP)}
\end{align*}
\]

\[
\begin{align*}
(QOP) \xrightarrow{\text{equiv.}} & (QOP') \\
\downarrow \text{SDP relax.} & \quad \text{SDP relax.} \downarrow \\
(SDP) \xrightarrow{\text{equiv.}} & (SDP')
\end{align*}
\]

General case: separate \( x \) into basic and nonbasic variables & substitute basic variables \( \Rightarrow \) SDP relax

\[
\inf_{x} \left\{ x^T Q x + 2 c^T x : a_j^T x = b_j \ (j = 1, \ldots, m), x_k \in \{0, 1\} \right\}
\]
Extension

SDP

$\inf_X \{ C \bullet X : A_j \bullet X = b_j, X \in S^n_+ \}$

Slater condition fails in Primal $\iff \exists y \in \mathbb{R}^m \setminus \{0\}$ such that

$b^T y \geq 0, -\sum_j y_j A_j \in S^n_+$

Moreover, if $\exists y$ such that $b^T y > 0$, then Primal is infeasible

Proof of ($\iff$) : Suppose the contrary that Slater condition holds in Primal. $\exists \hat{X}$ such that $A_j \bullet \hat{X} = b_j$ and $\hat{X} \in S^n_{++}$.

$0 \leq b^T y = \sum_j (A_j \bullet \hat{X}) y_j = \left( \sum_j A_j y_j \right) \bullet \hat{X} < 0$ (contradiction)
Facial Reduction

**Idea**: Let $W := - \sum_j A_j y_j \in \mathbb{S}_+^n$ and $b^T y = 0$

- For any feasible solutions $X$ in Primal,
  
  $$W \cdot X = - \sum_j (A_j \cdot X) y_j = - b^T y = 0.$$  

- Primal is equivalent to
  
  $$\inf_X \left\{ C \cdot X : A_j \cdot X = b_j, X \in \mathbb{S}_+^n \cap \{W\}^\perp \right\}$$

  where $\{W\}^\perp := \{X : X \cdot W = 0\}$

- The set $\mathbb{S}_+^n \cap \{W\}^\perp$ has nice structure

\[ \mathbb{S}_+^n \cap \{W\}^\perp = \left\{ X \in \mathbb{S}^n : X = Q \begin{pmatrix} M & O \\ O & O \end{pmatrix} Q^T, M \in \mathbb{S}_+^r \right\} \]
Idea (cont’d)

\[ S_+^n \cap \{ W \}^\perp = \left\{ X \in S^n : X = Q \begin{pmatrix} M & O \\ O & O \end{pmatrix} Q^T, M \in S_r^+ \right\} \]

- Assume \( Q = I_n \). Then Primal is equivalent to

\[ \inf \left\{ \tilde{C} \bullet X : \tilde{A}_j \bullet X = b_j, X \in S_+^r \right\} \]

where \( \tilde{A}_j \) is \( r \times r \) principal matrix

- Compare this SDP with Primal \( \Rightarrow \) the size \( n \rightarrow r \)
- May not satisfy Slater cond.
- \( \Rightarrow \) Find \( y \) and \( W \) for the smaller Primal
- This procedure terminates in finitely many iterations
- This procedure is called **Facial Reduction Algorithm** and acceptable for dual
History of FRA

- Borwein-Wolkowicz in 1980 for general convex optimization
- Ramana, Ramana-Tunçel-Wolkowicz for SDP
- Pataki simplified FRA for the extension
- Apply FRA into SDP relax. for Graph Partition, Quadratic Assignment, Sensor Network by Wolkowicz group
- Apply FRA into SDP relax. for Polynomial Optimization in Waki-Muramatsu
- ...
Summary on Slater condition

- Hope that both Primal and dual satisfy Slater conditions
- Otherwise, may not have any optimal solutions, and wrong value may be obtained
- Obtain inaccurate solutions even if exists optimal solutions, but, one can reduce the size of SDP
- FRA is a general framework to remove the difficulty in Slater cond.

In modeling to SDP...

- Need to be careful in even dual to guarantee the existence of optimal solutions in dual
- A rigorous solution for FRA is necessary
Status of infeasibility

Feasibility and infeasibility

\[
\inf_{\mathbf{X}} \left\{ \mathbf{C} \bullet \mathbf{X} : \mathbf{A}_j \bullet \mathbf{X} = \mathbf{b}_j, \mathbf{X} \in \mathbb{S}^n_+ \right\}
\]

- **Strongly feasible** if SDP satisfies Slater cond.
- **Weakly feasible** if SDP is feasible but, does not satisfies Slater cond.
- **Strongly infeasible** if \( \exists \) improving ray \( \mathbf{d} \), *i.e.*, \( \mathbf{b}^T \mathbf{d} > 0, - \sum_j d_j \mathbf{A}_j \in \mathbb{S}^n_+ \).
- **Weakly infeasible** if SDP is infeasible, but \( \nexists \) improving ray

**Remark**

- Weak infeasibility does not occur in LP
- SOCP and conic optimization also have the four status
Example: Infeasible SDPs

\[ \mathbb{P}_1 \inf_{\mathbf{X}} \left\{ \mathbf{C} \cdot \mathbf{X} : \begin{pmatrix} 1 \\ 1 \end{pmatrix} \cdot \mathbf{X} = 0, \begin{pmatrix} 1 \\ 1 \end{pmatrix} \cdot \mathbf{X} = 2, \mathbf{X} \in S^2_+ \right\}, \]

\[ \mathbb{P}_2 \inf_{\mathbf{X}} \left\{ \mathbf{C} \cdot \mathbf{X} : \begin{pmatrix} 1 \\ 1 \end{pmatrix} \cdot \mathbf{X} = 0, \begin{pmatrix} 1 \\ 1 \end{pmatrix} \cdot \mathbf{X} = 2, \mathbf{X} \in S^2_+ \right\} \]

Comments

- \( \mathbb{P}_1 \) is strongly infeasible because \( \exists \) certificate \( \mathbf{y} = (-1, 1) \)
- \( \mathbb{P}_2 \) is weakly infeasible because \( \not\exists \) certificate
Characterization of weak infeasibility

- Weakly infeasible SDP; for all $\epsilon > 0$, $\exists X \in S^n_+$

$$|A_j \cdot X - b_j| < \epsilon \ (j = 1, \ldots, m)$$

- More elementary characterization of Weak infeasibility by recent work by Liu and Pataki

**Example $P_2$**

Perturb $b_1 = 0 \rightarrow \epsilon > 0$

$$P_2 : \inf_X \left\{ C \cdot X : \begin{pmatrix} 1 \\ 1 \end{pmatrix} \cdot X = \epsilon, \begin{pmatrix} 1 \\ 1 \end{pmatrix} \cdot X = 2, X \in S^2_+ \right\}$$

Then, perturbed $P_1$ is feasible:

$$X = \begin{pmatrix} 1/\epsilon & 1 \\ 1 & \epsilon \end{pmatrix}$$
Pathological?

\[(\text{POP}) : \inf_{x,y} \{-x - y : xy \leq 1/2, x \geq 1/2, y \geq 1/2\}\]

- Optimal value is \(-1.5\)
- Apply Lasserre’s SDP hierarchy
- All SDP relaxation is weakly infeasible (in Waki 2012)
- SeDuMi and SDPA returns \(-1.5\) for higher order SDP relaxation
- Sufficient conditions of (POP) for SDP relaxation to be weakly infeasible (in Waki 2012)
Summary on infeasibility

- Weak infeasibility may occur in SDP, SOCP and conic optimization, but not in LP
- Difficult to detect this type of infeasibility by software
- But, software returns good values for weak infeasible SDP
Summary

- Introduce a part of theoretical and practical aspects in SDP
- Skip applications of SDP, e.g., SDP relaxation for combinatorial problems
- Can read papers on SDP
- Not so easy to handle SDP because it is convex but nonlinear programming
Further Reading I

M. Anjos and JB Lasserre,  
*Handbook of Semidefinite, Conic and Polynomial Optimization: Theory, Algorithms.*  

E. de Klerk,  
*Aspects of semidefinite programming : interior point algorithms and selected applications.*  

B. Gärtner and J. Matoušek  
*Approximation Algorithms and Semidefinite Programming.*  
Further Reading II

