Exploiting Structured Sparsity in Large Scale Semidefinite Programming Problems

M. Kojima
Tokyo Institute of Technology

November 2010

At National Cheng-Kung University, Tainan

- Kim, Kojima, Mevissen and Yamashita, "Exploiting sparsity in linear and nonlinear inequalities via positive semidefinite matrix completion", *Mathematical Programming* to appear.

Outline

1 A simple example for 2 types of sparsities
2 Mathematics behind exploiting sparsity

SDP: min $\sum_{i=1}^{n-1} (X_{ii} + b_i(X_{i,i+1} + X_{i+1,i})) + X_{nn}$ — (1)

sub. to (Matrix inequality, diagonal+bordered)

$$\boldsymbol{M}(\boldsymbol{X}) = \begin{pmatrix} 1 - X_{11} & 0 & \dots & X_{12} \\ 0 & 1 - X_{22} & \dots & X_{23} \\ \dots & \dots & \ddots & \dots \\ X_{21} & X_{32} & \dots & 1 - X_{nn} \end{pmatrix} \succeq \boldsymbol{O} \quad \text{— (2)}$$

$$\boldsymbol{X} = \begin{pmatrix} X_{11} & X_{12} & \dots & X_{1n} \\ X_{21} & X_{22} & \dots & X_{2n} \\ \dots & \dots & \dots & \dots \\ X_{n1} & X_{n2} & \dots & X_{nn} \end{pmatrix} \succeq \boldsymbol{O} \quad \text{(positive semidefinite)}$$

- The number of variables is $n(n+1)/2$; $X_{ij} = X_{ji}$.
- domain-space sparsity — Only X_{ij} $(|i-j| \leq 1)$ are used in (1), (2) among all variables X_{ij} $(1 \leq i \leq j \leq n)$.
- range-space sparsity — (2) is diagonal + bordered.

SDP: min $\sum_{i=1}^{n-1} (X_{ii} + b_i(X_{i,i+1} + X_{i+1,i})) + X_{nn}$ — (1)

sub. to (Matrix inequality, diagonal+bordered)

$$\boldsymbol{M}(\boldsymbol{X}) = \begin{pmatrix} 1 - X_{11} & 0 & \dots & X_{12} \\ 0 & 1 - X_{22} & \dots & X_{23} \\ \dots & \dots & \ddots & \dots \\ X_{21} & X_{32} & \dots & 1 - X_{nn} \end{pmatrix} \succeq \boldsymbol{O} \quad \text{— (2)}$$

$$\boldsymbol{X} = \begin{pmatrix} X_{11} & X_{12} & \dots & X_{1n} \\ X_{21} & X_{22} & \dots & X_{2n} \\ \dots & \dots & \dots & \dots \\ X_{n1} & X_{n2} & \dots & X_{nn} \end{pmatrix} \succeq \boldsymbol{O} \quad \text{(positive semidefinite)}$$

$\Downarrow$ conversion with exploiting the domain and range sparsities
"smaller size" SDP equivalent to the original SDP

- Next, numerical results on the converted SDP

Numerical results

- SeDuMi (MATLAB, a prima-dual interior-point method)
- 2.66 GHz Dual-Core Intel Xeon with 12GB memory

	SeDuMi elapsed time (second)	
size of $\boldsymbol{X}$ $= n$	Original SDP	Converted SDP with exploiting d-space & r-space sparsities
10	0.2	0.1
100	1091.4	0.6
1000	-	6.3
10000	-	99.2

Outline

A simple example of positive semidefinite completion problems

Let

$$A(x) \quad = \quad \begin{pmatrix} 3 & 3 & x \\ 3 & 3 & 2 \\ x & 2 & 2 \end{pmatrix} \in \mathbb{S}^3.$$

Question: $\exists x$ such that $A(x)$ is positive semidefinite?

Yes, for example, $A(2) \succeq \boldsymbol{O}$. But, $A(0) \not\succeq \boldsymbol{O}$.

Let $a_{ij} \in \mathbb{R}$, and consider

$$A(x_{ij}) \quad = \quad \begin{pmatrix} a_{11} & x_{12} & x_{13} & x_{14} & x_{15} & a_{16} \\ x_{12} & a_{22} & x_{23} & x_{24} & x_{25} & a_{26} \\ x_{13} & x_{23} & a_{33} & a_{34} & x_{35} & a_{36} \\ x_{14} & x_{25} & a_{34} & a_{44} & a_{45} & a_{46} \\ x_{15} & x_{25} & x_{35} & a_{45} & a_{55} & a_{56} \\ a_{16} & a_{26} & a_{36} & a_{46} & a_{56} & a_{66} \end{pmatrix} \in \mathbb{S}^6.$$

Question: How can we check whether $\boldsymbol{A}(x_{ij})$ is positive semidefinite for some x_{ij}?

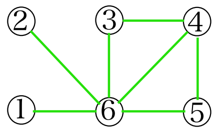
$\exists x_{ij} \Leftrightarrow$

$$A_{\{1,6\}} \equiv \begin{pmatrix} a_{11} & a_{16} \\ a_{61} & a_{66} \end{pmatrix} \succeq \boldsymbol{O}, \; A_{\{2,6\}} \succeq \boldsymbol{O}, \; A_{\{3,4,6\}} \succeq \boldsymbol{O}, \; A_{\{4,5,6\}} \succeq \boldsymbol{O}.$$

This equivalence holds when the green part has a chordal graph structure.

$$
\boldsymbol{A}(x_{ij}) \;=\; \begin{pmatrix} a_{11} & x_{12} & x_{13} & x_{14} & x_{15} & a_{16} \\ x_{12} & a_{22} & x_{23} & x_{24} & x_{25} & a_{26} \\ x_{13} & x_{23} & a_{33} & a_{34} & x_{35} & a_{36} \\ x_{14} & x_{25} & a_{34} & a_{44} & a_{45} & a_{46} \\ x_{15} & x_{25} & x_{35} & a_{45} & a_{55} & a_{56} \\ a_{16} & a_{26} & a_{36} & a_{46} & a_{56} & a_{66} \end{pmatrix} \in \mathbb{S}^6.
$$

Sparsity patter graph $G(N, E)$



- Each edge (i,j) is corresponding to an off-diagonal a_{ij}.
- If the graph $G(N, E)$ is chordal (*i.e.* any simple cycle with at least 4 edges has a chord)
- The maximal cliques are $\{1, 6\},\ \{2, 6\},\ \{3, 4, 6\},\ \{4, 5, 6\}$.
- Dual of the psd completion problem $\Rightarrow$ Next

Dual of the psd completion problem

- Duality : $\boldsymbol{X} \succeq \boldsymbol{O} \Leftrightarrow \boldsymbol{X} \bullet \boldsymbol{Y} \equiv \text{trace}\,\boldsymbol{X}^T\boldsymbol{Y} \geq 0$ for $\forall \boldsymbol{Y} \succeq \boldsymbol{O}$.
- A well-known simple case

$$
\boldsymbol{A} = \begin{pmatrix} \boldsymbol{B} & \boldsymbol{O} \\ \boldsymbol{O} & \boldsymbol{C} \end{pmatrix} \succeq \boldsymbol{O} \Leftrightarrow \boldsymbol{B} \succeq \boldsymbol{O},\ \boldsymbol{C} \succeq \boldsymbol{O}.
$$

- Extension

$$
\boldsymbol{A} = \begin{pmatrix} \boldsymbol{B}_{11} & \boldsymbol{B}_{12} & \boldsymbol{O} \\ \boldsymbol{B}_{21} & \boldsymbol{D} & \boldsymbol{C}_{12} \\ \boldsymbol{O} & \boldsymbol{C}_{21} & \boldsymbol{C}_{22} \end{pmatrix} \succeq \boldsymbol{O} \Leftrightarrow \left\{ \begin{array}{l} \begin{pmatrix} \boldsymbol{B}_{11} & \boldsymbol{B}_{12} \\ \boldsymbol{B}_{21} & \boldsymbol{B}_{22} \end{pmatrix} \succeq \boldsymbol{O}, \\ \begin{pmatrix} \boldsymbol{C}_{11} & \boldsymbol{C}_{12} \\ \boldsymbol{C}_{21} & \boldsymbol{C}_{22} \end{pmatrix} \succeq \boldsymbol{O}, \\ \boldsymbol{B}_{22} + \boldsymbol{C}_{22} = \boldsymbol{D} \end{array} \right.
$$

- Further extension to a general case $\Rightarrow$ Next

$G(N, E)$: a chordal graph with $N = \{1, \ldots, n\}$ and
the max. cliques of $C_1, \ldots, C_\ell$. $E^\bullet = E \cup \{(i, i) : i \in N\}$.
$\mathbb{S}^n(E^\bullet) = \{\boldsymbol{Y} \in \mathbb{S}^n : Y_{ij} = 0\ (i, j) \notin E^\bullet\}$.
$\mathbb{S}^C_+ = \{\boldsymbol{Y} \succeq \boldsymbol{O} : Y_{ij} = 0$ if $(i, j) \notin C \times C\}$ for $\forall C \subseteq N$.

Theorem (Agler, Helton, McCulough and Rodman 1988)
Suppose $\boldsymbol{M} \in \mathbb{S}^n(E^\bullet)$. $\boldsymbol{M} \succeq \boldsymbol{O}$ iff
$\boldsymbol{M} = \boldsymbol{Y}^1 + \boldsymbol{Y}^2 + \cdots + \boldsymbol{Y}^\ell$ for $\exists \boldsymbol{Y}^k \in \mathbb{S}^{C_k}_+$ $(k = 1, \ldots, \ell)$.

① — ② — ③ $C_1 = \{1, 2\},\ C_2 = \{2, 3\}$. $\boldsymbol{M} : \mathbb{R}^m \to \mathbb{S}^3(E^\bullet)$.

$$
\boldsymbol{M}(\boldsymbol{u}) = \begin{pmatrix} M_{11}(\boldsymbol{u}) & M_{12}(\boldsymbol{u}) & 0 \\ M_{21}(\boldsymbol{u}) & M_{22}(\boldsymbol{u}) & M_{23}(\boldsymbol{u}) \\ 0 & M_{32}(\boldsymbol{u}) & M_{33}(\boldsymbol{u}) \end{pmatrix}
$$

$G(N, E)$: a chordal graph with $N = \{1, \ldots, n\}$ and
the max. cliques of $C_1, \ldots, C_\ell$. $E^\bullet = E \cup \{(i, i) : i \in N\}$.
$\mathbb{S}^n(E^\bullet) = \{\boldsymbol{Y} \in \mathbb{S}^n : Y_{ij} = 0\ (i, j) \notin E^\bullet\}$.
$\mathbb{S}^C_+ = \{\boldsymbol{Y} \succeq \boldsymbol{O} : Y_{ij} = 0$ if $(i, j) \notin C \times C\}$ for $\forall C \subseteq N$.

Theorem (Agler, Helton, McCulough and Rodman 1988)
Suppose $\boldsymbol{M} \in \mathbb{S}^n(E^\bullet)$. $\boldsymbol{M} \succeq \boldsymbol{O}$ iff
$\boldsymbol{M} = \boldsymbol{Y}^1 + \boldsymbol{Y}^2 + \cdots + \boldsymbol{Y}^\ell$ for $\exists \boldsymbol{Y}^k \in \mathbb{S}^{C_k}_+$ $(k = 1, \ldots, \ell)$.

① — ② — ③ $C_1 = \{1, 2\},\ C_2 = \{2, 3\}$. $\boldsymbol{M} : \mathbb{R}^m \to \mathbb{S}^3(E^\bullet)$.

$$
\boldsymbol{M}(\boldsymbol{u}) = \begin{pmatrix} Y^1_{11} & Y^1_{12} & 0 \\ Y^1_{12} & Y^1_{22} & 0 \\ 0 & 0 & 0 \end{pmatrix} + \begin{pmatrix} 0 & 0 & 0 \\ 0 & Y^2_{22} & Y^2_{23} \\ 0 & Y^2_{32} & Y^2_{33} \end{pmatrix}
$$

$$
\boldsymbol{M}(\boldsymbol{u}) \succeq \boldsymbol{O}
$$

$$
\Updownarrow
$$

$$
\left. \begin{array}{l} M_{11} = Y^1_{11},\ M_{12} = Y^1_{12}, \\ M_{22} = Y^1_{22} + Y^2_{22}, \\ M_{23} = Y^2_{23},\ M_{33} = Y^2_{33}, \\ \square \succeq \boldsymbol{O},\ \square \succeq \boldsymbol{O} \end{array} \right\} \Leftrightarrow \left\{ \begin{array}{l} \begin{pmatrix} M_{11}(\boldsymbol{u}) & M_{12}(\boldsymbol{u}) \\ M_{21}(\boldsymbol{u}) & Y^1_{22} \end{pmatrix} \succeq \boldsymbol{O}, \\ \begin{pmatrix} M_{22}(\boldsymbol{u}) - Y^1_{22} & M_{23}(\boldsymbol{u}) \\ M_{32}(\boldsymbol{u}) & M_{33}(\boldsymbol{u}) \end{pmatrix} \succeq \boldsymbol{O} \end{array} \right.
$$