Exploiting Algebraic Symmetry in Semidefinite Programs: A Tutorial

Etienne de Klerk

Tilburg University, The Netherlands

EURO Summer School 2010, Klagenfurt, August, 2010

Standard form SDP

Primal problem

$$\min_{X \succeq 0} \operatorname{trace}(A_0 X) \;\text{ subject to }\; \operatorname{trace}(A_k X) = b_k \;\; (k = 1, \ldots, m),$$

where the data matrices $A_i \in \mathbb{S}^{n\times n}$ $(i = 0, \ldots, m)$ are linearly independent.

- $\mathbb{S}^{n\times n}$: symmetric $n \times n$ matrices;
- $X \succeq 0$: X symmetric positive semi-definite.

Sometimes we will add the additional constraint $X \geq 0$ (componentwise nonnegative).

Matrix algebras

Definition

A set $\mathcal{A} \subseteq \mathbb{C}^{n\times n}$ (resp. $\mathbb{R}^{n\times n}$) is called a *matrix *-algebra* over $\mathbb{C}$ (resp. $\mathbb{R}$) if, for all $X, Y \in \mathcal{A}$:

- $\alpha X + \beta Y \in \mathcal{A} \quad \forall \alpha, \beta \in \mathbb{C}$ (resp. $\mathbb{R}$);
- $X^* \in \mathcal{A}$;
- $XY \in \mathcal{A}$.

Assumption

There is a 'low dimensional' matrix *-algebra $\mathcal{A}_{SDP} \supseteq \{A_0, \ldots, A_m\}$.

Example

The circulant matrices form a commutative matrix *-algebra.

Form of a circulant matrix C

$$C = \begin{bmatrix} c_0 & c_1 & c_2 & \cdots & & c_{n-1} \\ c_{n-1} & c_0 & c_1 & & \cdots & \\ c_{n-2} & c_{n-1} & c_0 & c_1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & \\ & & & & & c_1 \\ c_1 & & \cdots & & c_{n-1} & c_0 \end{bmatrix}.$$

Each row is a cyclic shift of the row above it, i.e: $C_{ij} = c_{i-j \text{ mod } n}$ $(i, j = 0, \ldots, n-1)$.

Further reading:

R.M. Gray. Toeplitz and Circulant Matrices: A review. *Foundations and Trends in Communications and Information Theory*, **2**(3):155–239, 2006. Available online.

Link with SDP

Recall the dual SDP problem:

Dual problem

$$\max_{y \in \mathbb{R}^m,\ S \succeq 0} b^T y \ \text{ subject to } \sum_{i=1}^{m} y_i A_i + S = A_0.$$

Clearly $S \in \mathcal{A}_{SDP}$.

What about the primal problem:

$$\min_{X \succeq 0} \operatorname{trace}(A_0 X) \ \text{ subject to } \operatorname{trace}(A_k X) = b_k \ \ (k = 1, \ldots, m)?$$

We can show (next slides) that there exists an optimal $X \in \mathcal{A}_{SDP}$ if the SDP problem and its dual meet the Slater condition.

The central path

Central path

For any $\mu > 0$, the following system has a unique solution:

$$\begin{aligned} \operatorname{trace}(A_k X) &= b_k \ \ (k = 1, \ldots, m) \\ \sum_{i=1}^{m} y_i A_i + S &= A_0 \\ XS &= \mu I. \end{aligned}$$

- The solution, denoted by $(X(\mu), y(\mu), S(\mu))$, defines an analytic curve parameterized by μ.
- This curve is called the (primal-dual) central path.

- Setting $\mu = 0$ gives the optimality conditions.
- IPM's 'follow' the central path approximately.

The central path (ctd.)

Lemma

If $A \in \mathcal{A}_{SDP}$ and $\det(A) \neq 0$, then $A^{-1} \in \mathcal{A}_{SDP}$.

- One has $X(\mu) = \mu S(\mu)^{-1} \in \mathcal{A}_{SDP}$, since $S(\mu) \in \mathcal{A}_{SDP}$
- The limit $X^* = \lim_{\mu \downarrow 0} X(\mu)$ exists and gives a minimizer. Thus $X^* \in \mathcal{A}_{SDP}$.

cf.

Y. Kanno, M. Ohsaki, K. Murota and N. Katoh, Group symmetry in interior-point methods for semidefinite programming, *Optimization and Engineering*, 2(3): 293–320, 2001.

Consequence

We may restrict the primal problem to:

$$\min_{X \succeq 0} \left\{ \mathrm{trace}(A_0 X) \ : \ \mathrm{trace}(A_k X) = b_k \quad (k = 1, \ldots, m), \ X \in \mathcal{A}_{SDP} \right\}.$$

Canonical decomposition of a matrix *-algebra $\mathcal{A}$

Theorem (Wedderburn (1907))

*Assume $\mathcal{A}$ is a matrix *-algebra over $\mathbb{C}$ that contains I. Then there is a unitary Q ($Q^*Q = I$) and some integer s such that*

$$Q^*\mathcal{A}Q = \begin{pmatrix} \mathcal{A}_1 & 0 & \cdots & 0 \\ 0 & \mathcal{A}_2 & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \cdots & 0 & \mathcal{A}_s \end{pmatrix},$$

where each $\mathcal{A}_i \sim \mathbb{C}^{n_i \times n_i}$ for some integers n_i, and takes the form

$$\mathcal{A}_i = \left\{ \begin{pmatrix} A & 0 & \cdots & 0 \\ 0 & A & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \cdots & 0 & A \end{pmatrix} \middle| \; A \in \mathbb{C}^{n_i \times n_i} \right\} \quad (i = 1, \ldots, s).$$

Joseph Wedderburn (1882 – 1942)

"He was apparently a very shy man and much preferred looking at the blackboard to looking at the students. He had the galley proofs from his book 'Lectures on Matrices' pasted to cardboard for durability, and his 'lecturing' consisted of reading this out loud while simultaneously copying it onto the blackboard."

Example

The circulant matrices:

$$\begin{bmatrix} c_0 & c_1 & c_2 & \cdots & & c_{n-1} \\ c_{n-1} & c_0 & c_1 & & \cdots & \\ & c_{n-1} & c_0 & c_1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & \\ & & & & & c_1 \\ c_1 & & \cdots & & c_{n-1} & c_0 \end{bmatrix}$$

are diagonalized by the unitary (discrete Fourier transform) matrix:

$$Q_{ij} := \frac{1}{\sqrt{n}} e^{-2\pi\sqrt{-1}ij/n} \quad (i, j = 0, \ldots, n-1).$$

SDP reformulation

Assume we have a basis $B_1, \ldots, B_d$ of $\mathcal{A}_{SDP}$. Set $X = \sum_{i=1}^d x_i B_i$ to get

$$\min_{X \succeq 0} \{\mathrm{trace}(A_0 X) \;:\; \mathrm{trace}(A_k X) = b_k \quad (k = 1, \ldots, m), X \in \mathcal{A}_{SDP}\}$$

$$= \min_{x \in \mathbb{R}^d} \left\{ \sum_{i=1}^d x_i \mathrm{trace}(A_0 B_i) \;:\; \sum_{i=1}^d x_i \mathrm{trace}(A_k B_i) = b_k, \right.$$

$$\left. (k = 1, \ldots, m),\; \sum_{i=1}^d x_i B_i \succeq 0 \right\}.$$

- Replace the LMI by $\sum_{i=1}^d x_i Q^* B_i Q \succeq 0$ to get block-diagonal structure.
- Delete any identical copies of blocks in the block structure.

Proof sketch of Wedderburn's theorem

Let $\mathcal{A}$ be a matrix *-algebra over $\mathbb{C}$.

Definition

The commutant of $\mathcal{A}$ is defined as

$$\mathcal{A}' := \left\{ X \in \mathbb{C}^{n\times n} \ : \ XA = AX \ \ \forall\, A \in \mathcal{A} \right\}.$$

The commutant is a matrix *-algebra over $\mathbb{C}$.

Definition

The center of $\mathcal{A}$ is the commutative matrix *-algebra $\mathcal{A} \cap \mathcal{A}'$.

Proof sketch of Wedderburn's theorem (ctd.)

Lemma

A commutative *matrix *-algebra over* $\mathbb{C}$ *of dimension* s *containing* I *has a basis* $E_1, \ldots, E_s$ *such that:*

- $E_i^* = E_i$ *and* $E_i^2 = E_i$ *(idempotent);*
- E_i *is not a sum of other idempotents in the algebra (minimality);*
- $\sum_{i=1}^s E_i = I$.

Lemma

Let $\mathcal{A}$ *a matrix* $*$*-algebra over* $\mathbb{C}$ *containing* I*, and let* $E_1, \ldots, E_s$ *be the above basis of* $\mathcal{A} \cap \mathcal{A}'$ *(the center of* $\mathcal{A}$*). Then*

$$\mathcal{A} = \bigoplus_{i=1}^{s} \mathcal{A} E_i.$$

Proof sketch of Wedderburn's theorem (ctd.)

The lemma may be used to block diagonalize $\mathcal{A}$. Note

$$\mathcal{A}E_i = \mathcal{A}E_i^2 = E_i\mathcal{A}E_i.$$

Since $E_i = E_i^*$ (Hermitian) there is a unitary Q ($Q^*Q = I$) such that

$$E_i = Q\Lambda_i Q^* \quad (i = 1, \ldots, s)$$

and Λ_i a diagonal $0-1$ matrix for each i and $\sum_i \Lambda_i = I$.

Thus

$$\begin{aligned} Q^*\mathcal{A}Q &= \bigoplus_{i=1}^{s} Q^*E_i\mathcal{A}E_iQ \\ &= \bigoplus_{i=1}^{s} \Lambda_i\,(Q^*\mathcal{A}Q)\,\Lambda_i. \end{aligned}$$

Thus $Q^*\mathcal{A}Q$ is block diagonal. Denote block i by $\mathcal{A}_i$.

Proof sketch of Wedderburn's theorem (ctd.)

We have

$$Q^*\mathcal{A}Q = \begin{pmatrix} \mathcal{A}_1 & 0 & \cdots & 0 \\ 0 & \mathcal{A}_2 & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \cdots & 0 & \mathcal{A}_s \end{pmatrix} \text{ and } \mathcal{A}_i \cap \mathcal{A}_i' = \mathbb{C}I_{\mathrm{trace}(E_i)} \ (i = 1, \ldots, s).$$

Theorem

A matrix $*$-algebra $\mathcal{B} \subset \mathbb{C}^{n\times n}$ with center $\mathcal{B} \cap \mathcal{B}' = \mathbb{C}I$ takes the form

$$U^*\mathcal{B}U = \left\{ \begin{pmatrix} A & 0 & \cdots & 0 \\ 0 & A & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \cdots & 0 & A \end{pmatrix} \,\middle|\, A \in \mathbb{C}^{t\times t} \right\}$$

for some unitary U and integer t. The number of blocks n/t equals the dimension of a maximal commutative sub-algebra of $\mathcal{B}$.

Further reading

- The proof sketch here is based on:

§2.2 in:

D. Gijswijt. *Matrix Algebras and Semidefinite Programming Techniques for Codes*. PhD thesis, Univesity of Amsterdam, 2005. Available online.

- Further reading:

Chapter X in:

J.H.M. Wedderburn. *Lectures on Matrices*. AMS publishers, 1934. Available online.

- The proof is constructive; randomized algorithms that perform the canonical decomposition using the same ideas:

W. Eberly and M. Giesbrecht, Efficient decomposition of separable algebras. *Journal of Symbolic Computation*, 37(1): 35–81, 2004. Preprint available online.

K. Murota, Y. Kanno, M. Kojima and S. Kojima, A Numerical Algorithm for Block-Diagonal Decomposition of Matrix *-Algebras, Preprint 2007 (available online).

Coherent configurations

A basis $B_1, \ldots, B_d$ of a matrix *-algebra is called a coherent configuration if:

- The B_i's are 0-1 matrices;
- For each i, $B_i^T = B_{i^*}$ for some $i^* \in \{1, \ldots, d\}$;
- $\sum_{i=1}^d B_i = J$ (the all-ones matrix).

If the B_i's also commute, and $B_1 = I$, then we speak of an association scheme.

Consequence

If $\mathcal{A}_{SDP}$ is spanned by a coherent configuration and $X = \sum_{i=1}^d x_i B_i$, then

$$X \succeq 0 \text{ and } X \geq 0 \Longleftrightarrow \sum_{i=1}^d x_i B_i \succeq 0 \text{ and } x \geq 0.$$

Example

The circulant matrices have the basis

$$\begin{bmatrix} 1 & 0 & 0 & \cdots & & 0 \\ 0 & 1 & 0 & & \cdots & \\ 0 & 0 & 1 & 0 & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & \\ & & & & & \\ 0 & & \cdots & & 0 & 1 \end{bmatrix}, \begin{bmatrix} 0 & 1 & 0 & \cdots & & 0 \\ 0 & 0 & 1 & & \cdots & \\ & 0 & 0 & 1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & \\ & & & & & 1 \\ 1 & & \cdots & & 0 & 0 \end{bmatrix}, \ldots$$

and form an association scheme.

SDP reformulation for nonnegative variables

Assume we have a coherent configuration $B_1, \ldots, B_d$ of $\mathcal{A}_{SDP}$. Then

$$\min_{X \succeq 0, X \geq 0} \{\mathrm{trace}(A_0 X) \ : \ \mathrm{trace}(A_k X) = b_k \quad (k = 1, \ldots, m), X \in \mathcal{A}_{SDP}\}$$

$$= \min_{x \geq 0} \left\{ \sum_{i=1}^{d} x_i \mathrm{trace}(A_0 B_i) \ : \ \sum_{i=1}^{d} x_i \mathrm{trace}(A_k B_i) = b_k, \right.$$

$$\left. (k = 1, \ldots, m), \ \sum_{i=1}^{d} x_i B_i \succeq 0 \right\}.$$

- We may again replace the LMI by $\sum_{i=1}^{d} x_i Q^* B_i Q \succeq 0$ to get block-diagonal structure ...
- ... and delete any identical copies of blocks in the block structure.

Lovász ϑ-function

A graph $G = (V, E)$ is given.

Lovász ϑ-function

$$\vartheta(G) := \max \operatorname{trace}(JX)$$

subject to

$$\begin{aligned} X_{ij} &= 0, \ \{i,j\} \in E \ (i \neq j) \\ \operatorname{trace}(X) &= 1 \\ X \succeq 0, & \end{aligned}$$

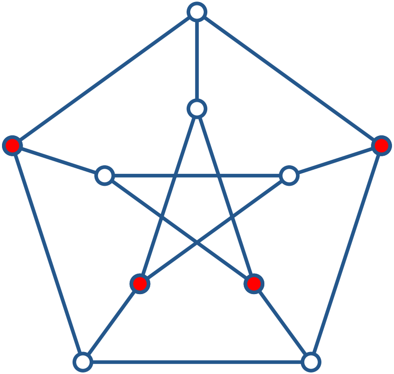
where e denotes the all-one vector.

Schrijver ϑ'-function

Add the additional constraint $X \geq 0$ to the ϑ problem.

Stable sets in graphs

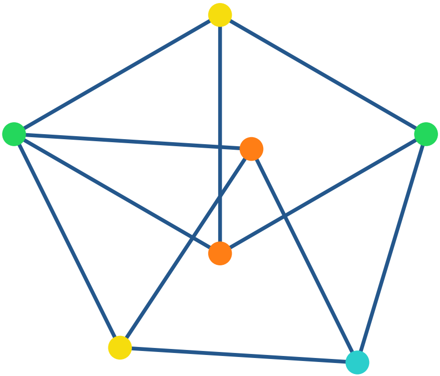
A *stable set* of $G = (V, E)$ is a subset $V' \subset V$ such that the *induced subgraph* on V' has no edges.



The stability number $\alpha(G)$ is the cardinality of the largest co-clique of G.

Vertex colourings

A (proper) vertex colouring is an assignment of colours to the vertices V of G such that endpoints of each edge are assigned different colours.



The smallest number of colours needed is called the chromatic number $\chi(G)$.

Lovász sandwich theorem

Theorem

$$\alpha(G) \le \vartheta'(G) \le \vartheta(G) \le \chi(\bar{G}),$$

where $\bar{G}$ is the complementary graph of G.

- One may approximate $\alpha(G)$ or $\chi(\bar{G})$ by $\vartheta(G)$ (or $\vartheta'(G)$).
- E.g. for the Pentagon graph (C_5) one has

$$2 \equiv \alpha(C_5) \le \vartheta'(C_5) = \sqrt{5} = \vartheta(C_5) \le \chi(\bar{C}_5) \equiv 3.$$

- We will do the symmetry reduction of the SDP to calculate $\vartheta'(C_5)$ in detail.

Example: ϑ'-of the Pentagon

Equivalent formulation for ϑ':

$$\vartheta'(G) := \max_{X \succeq 0, X \geq 0} \{\mathrm{trace}(JX) \mid \mathrm{trace}(AX) = 0, \ \mathrm{trace}(X) = 1\},$$

where A is the adjacency matrix of the graph G.

The data matrices of this SDP are J, A, and I.

For $G = C_5$ one has

$$A = \begin{bmatrix} 0 & 1 & 0 & 0 & 1 \\ 1 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 1 \\ 1 & 0 & 0 & 1 & 0 \end{bmatrix}.$$

The data matrices are 5×5 symmetric circulant matrices.

Example (ctd.)

A basis for the 5×5 symmetric circulant matrices is

$$B_1 = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}, \; B_2 = \begin{bmatrix} 0 & 1 & 0 & 0 & 1 \\ 1 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 1 \\ 1 & 0 & 0 & 1 & 0 \end{bmatrix}, \; B_3 = \begin{bmatrix} 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 & 1 \\ 1 & 0 & 0 & 0 & 1 \\ 1 & 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 \end{bmatrix}$$

Note that $A = B_2$ where A the adjacency matrix of C_5.

Example (ctd.)

For Q the discrete Fourier transform matrix we have ($B_1 = I$):

$$Q^* B_2 Q = \begin{bmatrix} 2 & 0 & 0 & 0 & 0 \\ 0 & \frac{3-\sqrt{5}}{\sqrt{5}-1} & 0 & 0 & 0 \\ 0 & 0 & -\frac{2}{\sqrt{5}-1} & 0 & 0 \\ 0 & 0 & 0 & -\frac{2}{\sqrt{5}-1} & 0 \\ 0 & 0 & 0 & 0 & \frac{3-\sqrt{5}}{\sqrt{5}-1} \end{bmatrix}$$

$$Q^* B_3 Q = \begin{bmatrix} 2 & 0 & 0 & 0 & 0 \\ 0 & -\frac{2}{\sqrt{5}-1} & 0 & 0 & 0 \\ 0 & 0 & \frac{3-\sqrt{5}}{\sqrt{5}-1} & 0 & 0 \\ 0 & 0 & 0 & \frac{3-\sqrt{5}}{\sqrt{5}-1} & 0 \\ 0 & 0 & 0 & 0 & -\frac{2}{\sqrt{5}-1} \end{bmatrix}.$$

(Block-)diagonal form with repeated blocks.

Example (ctd.)

$$\vartheta'(C_5) := \max_{X \succeq 0, X \geq 0} \{\mathrm{trace}(JX) \mid \mathrm{trace}(AX) = 0,\ \mathrm{trace}(X) = 1\}.$$

Setting $X = \sum_{i=1}^{3} x_i B_i$ with $x_i \geq 0$ $(i = 1, 2, 3)$:

$$\vartheta'(C_5) := \max_{x \geq 0} \sum_{i=1}^{3} x_i \,\mathrm{trace}(JB_i)$$

$$\text{subject to: } \sum_{i=1}^{3} x_i \,\mathrm{trace}(AB_i) = 0,\ \sum_{i=1}^{3} x_i \,\mathrm{trace}(B_i) = 1,\ \sum_{i=1}^{3} x_i Q^* B_i Q \succeq 0.$$

Note that $A = B_2$ and $\mathrm{trace}(AX) = 0$ imply $x_2 = 0$, and $\mathrm{trace}(X) = 1$ implies $x_1 = 1/5$. Thus

$$\vartheta'(C_5) := \max_{x_3 \geq 0} \{1 + 10x_3 \mid 1/5I + x_3 Q^* B_3 Q \succeq 0\}.$$

Example (ctd.)

Deleting repeated blocks, we end with an LP in one variable:

$$\vartheta'(C_5) := \max_{x_3 \geq 0} 1 + 10x_3$$

subject to

$$\frac{1}{5}\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} + x_3 \begin{bmatrix} 2 & 0 & 0 \\ 0 & -\frac{2}{\sqrt{5}-1} & 0 \\ 0 & 0 & \frac{3-\sqrt{5}}{\sqrt{5}-1} \end{bmatrix} \succeq 0.$$

Optimal solution $x_3 = \frac{\sqrt{5}-1}{10}$, $\vartheta'(C_5) = \sqrt{5}$.

Other symmetry reduction techniques

- In general we do not know the unitary Q that gives the canonical decomposition of $\mathcal{A}_{SDP}$, ...
- ... and the matrix size n may be too large to compute Q using linear algebra.
- Idea: use other faithful representations of $\mathcal{A}_{SDP}$.
- One such faithful representation is the regular $*$-representation of $\mathcal{A}_{SDP}$.

Regular $*$-representation

Assume that $B_1, \ldots, B_d \in \mathbb{R}^{n\times n}$ is an orthogonal basis of $\mathcal{A}_{SDP}$, seen as a matrix $*$-algebra over $\mathbb{R}$. Normalize the basis:

$$D_i := \frac{1}{\sqrt{\mathrm{trace}(B_i^T B_i)}} B_i \quad (i = 1, \ldots, d).$$

Define multiplication parameters $\gamma_{i,j}^k$ via:

$$D_i D_j = \sum_k \gamma_{i,j}^k D_k,$$

and define the $d \times d$ matrices L_k $(k = 1, \ldots, d)$ via

$$(L_k)_{ij} = \gamma_{k,j}^i.$$

Regular $*$-representation ctd.

The matrices L_k form a basis of a faithful representation of $\mathcal{A}_{SDP}$, say $\mathcal{A}_{SDP}^{reg}$, that is also a $C*$-algebra, called the *regular $*$-representation of $\mathcal{A}_{\mathcal{SDP}}$.*

Theorem

The mapping $D_i \mapsto L_i$ $(i = 1, \ldots, d)$ defines a $*$-isomorphism from $\mathcal{A}_{SDP}$ to $\mathcal{A}_{SDP}^{reg}$.

Corollary:

$\sum_{i=1}^{d} x_i D_i \succeq 0 \iff \sum_{i=1}^{d} x_i L_i \succeq 0.$

Consequence: We can work with $d \times d$ data matrices as opposed to $n \times n$.

SDP reformulation via regular $*$-representation

Setting $X = \sum_{i=1}^d x_i D_i$,

$$\min_{X \succeq 0} \{\mathrm{trace}(A_0 X) \ : \ \mathrm{trace}(A_k X) = b_k \quad (k = 1, \ldots, m)\}$$

becomes

$$\min_{x \in \mathbb{R}^d} \left\{ \sum_{i=1}^d x_i \mathrm{trace}(A_0 D_i) \ : \ \sum_{i=1}^d x_i \mathrm{trace}(A_k D_i) = b_k \ (k = 1, \ldots, m), \ \sum_{i=1}^d x_i D_i \succeq 0 \right\}.$$

... and by the corollary we can replace D_i by L_i in the LMI.

See the proof of Theorem 1 in:

E. de Klerk, D.V. Pasechnik and A. Schrijver. Reduction of symmetric semidefinite programs using the regular *-representation. *Mathematical Programming B*, 109(2-3):613-624, 2007.

Nonnegative matrix variables

As before, if $\mathcal{A}_{SDP}$ is spanned by a coherent configuration, then the additional nonnegativity constraint $X \geq 0$ becomes $x \geq 0$. Thus

$$\min_{X \succeq 0, X \geq 0} \{\mathrm{trace}(A_0 X) \ : \ \mathrm{trace}(A_k X) = b_k \quad (k = 1, \ldots, m)\}$$

reduces to

$$\min_{x \geq 0} \left\{ \sum_{i=1}^{d} x_i \mathrm{trace}(A_0 D_i) \ : \ \sum_{i=1}^{d} x_i \mathrm{trace}(A_k D_i) = b_k \ (k = 1, \ldots, m), \ \sum_{i=1}^{d} x_i L_i \succeq 0 \right\}.$$

$\vartheta'(C_5)$ example revisited

Recall the reformulation to compute $\vartheta'(C_5)$:

$$\vartheta'(C_5) := \max_{x_3 \geq 0} \{1 + 10x_3 \mid 1/5 B_1 + x_3 B_3 \succeq 0\},$$

where

$$B_1 = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}, \; B_2 = \begin{bmatrix} 0 & 1 & 0 & 0 & 1 \\ 1 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 1 \\ 1 & 0 & 0 & 1 & 0 \end{bmatrix}, \; B_3 = \begin{bmatrix} 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 & 1 \\ 1 & 0 & 0 & 0 & 1 \\ 1 & 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 \end{bmatrix}$$

was the basis for the 5×5 symmetric circulant matrices.

$\vartheta'(C_5)$ example (ctd.)

Normalize the basis B_1, B_2, B_3 to get

$$D_1 = \frac{1}{\sqrt{5}} I,\ D_2 = \frac{1}{\sqrt{10}} \begin{bmatrix} 0 & 1 & 0 & 0 & 1 \\ 1 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 1 \\ 1 & 0 & 0 & 1 & 0 \end{bmatrix},\ D_3 = \frac{1}{\sqrt{10}} \begin{bmatrix} 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 & 1 \\ 1 & 0 & 0 & 0 & 1 \\ 1 & 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 \end{bmatrix}$$

Multiplication table:

	D_1	D_2	D_3
D_1	$\frac{1}{\sqrt{5}} D_1$	$\frac{1}{\sqrt{5}} D_2$	$\frac{1}{\sqrt{5}} D_3$
D_2	$\frac{1}{\sqrt{5}} D_2$	$\frac{1}{\sqrt{5}} D_1 + \frac{1}{\sqrt{10}} D_3$	$\frac{1}{\sqrt{10}}(D_2 + D_3)$
D_3	$\frac{1}{\sqrt{5}} D_3$	$\frac{1}{\sqrt{10}}(D_2 + D_3)$	$\frac{1}{\sqrt{5}} D_1 + \frac{1}{\sqrt{10}} D_2$

$\vartheta'(C_5)$ example (ctd.)

Construct the matrices L_1, L_2, and L_3 via:

$$D_i D_j = \sum_k \gamma_{i,j}^k D_k,$$

and

$$(L_k)_{ij} = \gamma_{k,j}^i \quad (k = 1, \ldots, 3).$$

For example:

$$L_3 := \begin{bmatrix} \gamma_{3,1}^1 & \gamma_{3,2}^1 & \gamma_{3,3}^1 \\ \gamma_{3,1}^2 & \gamma_{3,2}^2 & \gamma_{3,3}^2 \\ \gamma_{3,1}^3 & \gamma_{3,2}^3 & \gamma_{3,3}^3 \end{bmatrix} = \begin{bmatrix} 0 & 0 & \frac{1}{\sqrt{5}} \\ 0 & \frac{1}{\sqrt{10}} & \frac{1}{\sqrt{10}} \\ \frac{1}{\sqrt{5}} & \frac{1}{\sqrt{10}} & 0 \end{bmatrix}.$$

$\vartheta'(C_5)$ example (ctd.)

We had

$$\begin{aligned}\vartheta'(C_5) &:= \max_{x_3\geq 0}\{1+10x_3 \mid 1/5B_1 + x_3B_3 \succeq 0\},\\ &= \max_{x_3\geq 0}\{1+10x_3 \mid 1/\sqrt{5}D_1 + \sqrt{10}x_3D_3 \succeq 0\}.\end{aligned}$$

Via the regular $*$-reduction we may replace the D_i's by the L_i's:

$$\vartheta'(C_5) = \max_{x_3\geq 0} 1+10x_3$$

subject to

$$\frac{1}{5}I + \sqrt{10}x_3\begin{bmatrix} 0 & 0 & \frac{1}{\sqrt{5}} \\ 0 & \frac{1}{\sqrt{10}} & \frac{1}{\sqrt{10}} \\ \frac{1}{\sqrt{5}} & \frac{1}{\sqrt{10}} & 0 \end{bmatrix} \succeq 0.$$

Matrix $*$-algebras from groups

Symmetric group

Let $\mathcal{S}_n$ denote the symmetric group on n elements, i.e. the group of all permutations of $\{1, \ldots, n\}$.

We may represent any sub-group $\mathcal{G} \subseteq \mathcal{S}_n$ as a multiplicative group of $n \times n$ permutation matrices via

$$(P_\pi)_{i,j} := \begin{cases} 1 & \text{if } \pi(i) = j \\ 0 & \text{else.} \end{cases} \qquad \pi \in \mathcal{G},\ i,j = 1, \ldots, n.$$

Commutant (centralizer ring)

The commutant of the representation is

$$\{A \in \mathbb{C}^{n\times n}\ :\ AP_\pi = P_\pi A \quad \forall \pi \in \mathcal{G}\},$$

and forms a matrix $*$-algebra over $\mathbb{C}$.

Matrix $*$-algebras from groups

Lemma

The commutant of the representation has a basis that is a coherent configuration.

One may construct this 0-1 basis of the commutant from the orbitals of $\mathcal{G}$.

Definition

The two-orbit or orbital of an index pair (i, j) is defined as

$$\{(\pi(i), \pi(j)) \; : \; \pi \in \mathcal{G}\}.$$

The orbitals partition $\{1, \ldots, n\} \times \{1, \ldots, n\}$ and this partition yields the $0 - 1$ matrices of the coherent configuration.

Matrix automorphism groups

Definition

We define the *automorphism group* $\mathrm{aut}(Z)$ of a matrix $Z \in \mathbb{S}^{n\times n}$ as all $\pi \in \mathcal{S}_n$ such that

$$Z_{ij} = Z_{\pi(i)\pi(j)} \quad \forall i,j = 1,\ldots,n.$$

Thus

$$Z = P_\pi Z P_\pi^T \quad \forall \pi \in \mathrm{aut}(Z).$$

Note that

$$ZP_\pi = P_\pi Z \quad \forall \pi \in \mathrm{aut}(Z),$$

i.e. Z belongs to the commutant of the permutation representation of $\mathrm{aut}(Z)$.

SDP symmetry assumption

SDP symmetry assumption:

The multiplicative matrix group $\mathcal{G}_{SDP} := \bigcap_{i=0}^{m} \mathrm{aut}(A_i)$ is non-trivial.

Thus we may take $\mathcal{A}_{SDP}$ as the commutant of the permutation representation of $\mathcal{G}_{SDP}$.

Then $\mathcal{A}_{SDP}$ will have a 0-1 basis (coherent configuration) given by the orbitals of $\mathcal{G}_{SDP}$.

ϑ' example once more

$$\vartheta'(C_5) := \max_{X \succeq 0, X \geq 0} \{\text{trace}(JX) \mid \text{trace}(AX) = 0,\ \text{trace}(X) = 1\},$$

where

$$A = \begin{bmatrix} 0 & 1 & 0 & 0 & 1 \\ 1 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 1 \\ 1 & 0 & 0 & 1 & 0 \end{bmatrix}.$$

The data matrices of this SDP are J, A, and I and $\mathcal{G}_{SDP} = \text{aut}(A) = D_5$, where D_5 is the dihedral group on 5 elements.

$\vartheta'(C_5)$ example revisited

The orbitals of D_5 correspond to the basis we had before:

$$B_1 = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix},\ B_2 = \begin{bmatrix} 0 & 1 & 0 & 0 & 1 \\ 1 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 1 \\ 1 & 0 & 0 & 1 & 0 \end{bmatrix},\ B_3 = \begin{bmatrix} 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 & 1 \\ 1 & 0 & 0 & 0 & 1 \\ 1 & 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 \end{bmatrix}.$$

Thus $\mathcal{A}_{SDP}$ is the set 5×5 symmetric circulant matrices.

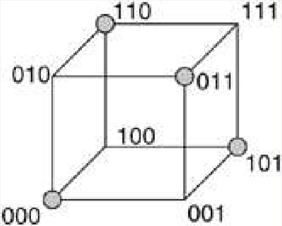
Appendix: Applications

- Error correcting binary codes;
- Crossing numbers of completely bipartite graphs;
- Quadratic assignment problems;
- Truss topology optimization.

The Hamming graph and binary codes

The Hamming graph $G(k,\delta)$ has vertices indexed by $\{0,1\}^k$ and vertices adjacent if they are at Hamming distance less than δ.

Hamming graph with $k = 3$ and $\delta = 2$.



Usual notation: $\alpha(G(k,\delta)) =: A(k,\delta)$. Thus $A(3,2) = 4$ (see picture).

$A(k,\delta)$ is the maximum size of a binary code on k letters such that any two words are at a Hamming distance of at least δ.

ϑ'-of the Hamming graph

Equivalent formulation for ϑ':

$$\vartheta'(G) := \max_{X \succeq 0, X \geq 0} \{\mathrm{trace}(JX) \mid \mathrm{trace}(A+I)X = 1\},$$

where A is the adjacency matrix of the graph G. Thus $\mathcal{G}_{SDP} = \mathrm{aut}(A)$.

- For the Hamming graph $|\mathrm{aut}(A)| = 2^k k!$, and ....
- ... the commutant of $\mathrm{aut}(A)$ is the commutative Bose-Mesner algebra of the Hamming scheme ...
- ... that has dimension $k+1$.
- Thus the SDP matrices may be reduced from the original size $n = 2^k$ to diagonal matrices of size $k+1$.

The resulting LP coincides with the LP bound of Delsarte.

A. Schrijver. A comparison of the Delsarte and Lovász bounds. *IEEE Trans. Inform. Theory*, 25:425–429, 1979.

Improvements

A. Schrijver. New code upper bounds from the Terwilliger algebra. *IEEE Transactions on Information Theory*, 51:2859–2866, 2005,

In this paper, a stronger SDP bound for $A(k, \delta)$ is obtained as follows:

- a stronger SDP relaxation is constructed via 'lift-and-project' ...
- ... such that some symmetry is retained in the resulting SDP.
- $\mathcal{A}_{SDP}$ becomes the Terwilliger algebra of the Hamming scheme, a non-commutative algebra that contains the Bose-Mesner algebra of the Hamming scheme.
- The Terwilliger algebra has dimension $\binom{k+3}{3}$ and its canonical block-diagonalization is known.

Thus, improved upper bounds were computed for $A(19, 6)$, $A(23, 6)$, $A(25, 6)$, ...

Further improvements

Using other lift-and-project schemes, slightly better SDP bounds may be obtained.

M. Laurent. Strengthened semidefinite bounds for codes. *Mathematical Programming*, 109(2-3):239–261, 2007,

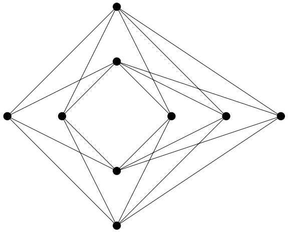
... and the approach may be extended to non-binary codes:

D. Gijswijt, A. Schrijver, H. Tanaka, New upper bounds for nonbinary codes based on the Terwilliger algebra and semidefinite programming, *Journal of Combinatorial Theory, Series A*, 113, 1719–1731, 2006.

Crossing numbers

The complete bipartite graph $K_{r,s}$ can be drawn in the plane with at most $Z(r,s)$ edges crossing, where

$$Z(r,s) = \left\lfloor \frac{r-1}{2} \right\rfloor \left\lfloor \frac{r}{2} \right\rfloor \left\lfloor \frac{s-1}{2} \right\rfloor \left\lfloor \frac{s}{2} \right\rfloor .$$



A drawing of $K_{4,5}$ with $Z(4,5) = 8$ crossings.

The smallest possible number of crossings is called the crossing number: $\mathrm{cr}(K_{r,s})$.

Crossing numbers (ctd.)

Conjecture (Zarankiewicz)

$$\mathrm{cr}(K_{r,s}) = Z(r,s).$$

(Open problem since Turán posed it in the 1940's).

- Lower bounds on $K_{r,s}$ may be obtained by solving

$$\mathrm{cr}(K_{r,s}) \geq \min_{X \succeq 0, X \geq 0} \{\mathrm{trace}(DX) \; : \; \mathrm{trace}(JX) = 1\} - \left\lfloor \frac{s}{2} \right\rfloor \left\lfloor \frac{s-1}{2} \right\rfloor \qquad (\heartsuit)$$

- ... where D has $(r-1)!$ columns indexed by the cyclic orderings on r elements, ...
- and the entries of D are 'distances' between pairs of cyclic orderings.
- Here, $\mathcal{G}_{SDP} = \mathrm{aut}(D)$, and $|\mathcal{G}_{SDP}| = 2r!$.

Crossing numbers (ctd.)

Using the regular $*$-representation of the SDP ($\heartsuit$), it was shown that

$$0.859 Z(r,s) \leq \mathrm{cr}(K_{r,s}) \leq Z(r,s)$$

if r or s is sufficiently large.

E. de Klerk, J. Maharry, D.V. Pasechnik, B. Richter and G. Salazar. Improved bounds for the crossing numbers of $K_{m,n}$ and K_n. *SIAM J. Discr. Math.* 20:189–202, 2006.

E. de Klerk, D.V. Pasechnik and A. Schrijver. Reduction of symmetric semidefinite programs using the regular *-representation. *Mathematical Programming B*, 109(2-3):613-624, 2007.

Quadratic assignment problem (QAP)

Definition (Trace formulation (Edwards 1977))

Given are symmetric $k \times k$ matrices A (distance matrix) and B (flow matrix).

$$\min_{X \in \Pi_k} \operatorname{trace}(AXBX^T)$$

where Π_k is the set of $k \times k$ permutation matrices.

- QAP is NP-hard in the strong sense;
- Many applications, but very hard to solve in practice for $k \geq 30$.

SDP relaxation of QAP

$$\left.\begin{array}{rl} \min & \operatorname{trace}(B \otimes A)Y \\ \text{subject to} & \\ & \operatorname{trace}((I \otimes (J - I))Y + ((J - I) \otimes I)Y) = 0 \\ & \operatorname{trace}(Y) - 2e^T y = -k \\ & \begin{pmatrix} 1 & y^T \\ y & Y \end{pmatrix} \succeq 0, \quad Y \geq 0. \end{array}\right\} \qquad (\diamondsuit)$$

- J (resp. e) is the all-ones matrix (resp. vector);
- Y corresponds to $\operatorname{vec}(X)\operatorname{vec}(X)^T$ for an optimal assignment $X \in \Pi_k$.
- This relaxation is equivalent to an SDP relaxation studied in:

Q. Zhao, S.E. Karisch, F. Rendl, and H. Wolkowicz. Semidefinite Programming Relaxations for the Quadratic Assignment Problem. *Journal of Combinatorial Optimization*, **2**, 71–109, 1998.

- Y is $k^2 \times k^2$ — size reduction of the SDP essential if $k \geq 15$ ...

SDP relaxation of QAP: symmetry

For the QAP relaxation $(\diamondsuit)$, $n = k^2 + 1$ and the data matrices are:

$$\begin{pmatrix} 0 & 0^T \\ 0 & B \otimes A \end{pmatrix}, \quad \begin{pmatrix} 0 & 0^T \\ 0 & I \otimes (J - I) + (J - I) \otimes I, \end{pmatrix}, \quad \text{and} \quad \begin{pmatrix} 0 & -e^T \\ -e & I \end{pmatrix},$$

... and the group $\mathcal{G}_{SDP}$ is given by

$$\mathcal{G}_{SDP} := \left\{ \begin{pmatrix} 1 & 0^T \\ 0 & P_A \otimes P_B \end{pmatrix} \; : \; P_A \in \mathrm{aut}(A), \; P_B \in \mathrm{aut}(B) \right\},$$

where A and B are the distance and flow matrices of the QAP as before.

SDP relaxation of QAP: numerical results

Several instances in the QAPlib library have algebraic symmetry, e.g. the distance matrix is a Hamming distance matrix.

Some numerical results, after doing the SDP symmetry reduction:

instance	k	previous l.b.	SDP l.b. ($\diamondsuit$)	best known u.b.	time(s)
esc64a	64	47	98	116	13
esc128a	128	2	54	64	140

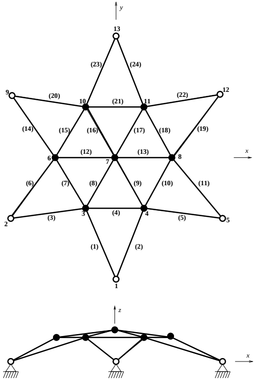
E. de Klerk and R. Sotirov. Exploiting Group Symmetry in Semidefinite Programming Relaxations of the Quadratic Assignment Problem. *Mathematical Programming*, 122(2), 2010.

The traveling salesman problem (TSP) may be formulated as QAP with a circulant distance matrix — new SDP relaxation of TSP via symmetry reduction.

E. de Klerk, D.V. Pasechnik and R. Sotirov. On semidefinite programming relaxations of the traveling salesman problem. *SIAM Journal on Optimization*, 19(4), 1559–1573, 2008.

A truss topology optimization problem

Design a truss of minimum volume such that the fundamental frequency of vibrations is higher that some prescribed critical value.



Top and front view of a dome-shaped truss

A truss topology optimization problem (ctd.)

SDP formulation introduced in:

M. Ohsaki, K. Fujisawa, N. Katoh and Y. Kanno, Semi-definite programming for topology optimization of trusses under multiple eigenvalue constraints, *Comp. Meth. Appl. Mech. Engng.*, 180: 203–217, 1999.

- The SDP has algebraic symmetry if the ground structure of nodes has isometries ...
- E.g., for the dome example the symmetry group $\mathcal{G}_{SDP}$ of the SDP is (a certain representation of) the dihedral group.

Further reading:

Y.Q. Bai, E. de Klerk, D.V. Pasechnik, R. Sotirov. Exploiting Group Symmetry in Truss Topology Optimization. *Optimization and Engineering*, **10**(3), 331–349, 2009.

And, finally ...

- Symmetry reduction in SDP is the application of representation theory to reduce the size of specially structured SDP instances.
- The most notable applications are in computer assisted proofs (bounds on crossing numbers, kissing numbers, error correcting codes, ...)
- ... but also pre-processing of some SDP's arising in optimal design (truss design, QAP, ...)
- More applications in polynomial optimization, graph coloring, ...

The End

THANK YOU!