

Exploiting symmetry in optimization

Alexander Schrijver

CWI and University of Amsterdam

Based on:

A. Schrijver, New code upper bounds from the Terwilliger algebra and semidefinite programming, *IEEE Transactions on Information Theory* 51 (2005) 2859–2866

E. de Klerk, D.V. Pasechnik, A. Schrijver, Reduction of symmetric semidefinite programs using the regular $*$ -representation, *Mathematical Programming*, to appear.

DIAMANT areas:

Discrete Mathematics (graphs, codes)
Algorithmic Mathematics (optimization)
Algebra (invariant theory)

Basic fact:

Let \mathcal{G} be a finite group acting on

$$\max\{c^T x \mid x \in \mathcal{F}\},$$

where $c \in \mathbb{R}^n$ and \mathcal{F} is a compact convex subset of \mathbb{R}^n .

That is, \mathcal{G} is a group of $n \times n$ matrices such that $Mc = c$ and $M\mathcal{F} = \mathcal{F}$ for all $M \in \mathcal{G}$.

Then the maximum is attained by a vector in the *invariant space*

$$\{x \mid Mx = x \text{ for all } M \in \mathcal{G}\}.$$

Proof:

If x attains the maximum, then so does $\frac{1}{|\mathcal{G}|} \sum_{M \in \mathcal{G}} Mx$.

In particular, we consider optimization problems of type

$$\max\{\text{tr}(CX) \mid X \in \mathcal{F}\},$$

where C is some $N \times N$ matrix and \mathcal{F} is a compact convex collection of $N \times N$ matrices.

Let \mathcal{G} be a group of $N \times N$ permutation matrices such that

$$MCM^T = C \text{ and } M\mathcal{F}M^T = \mathcal{F} \text{ for each } M \in \mathcal{G}.$$

Then the maximum is attained by a matrix in

$$C_{\mathcal{G}} := \{X \in \mathbb{R}^{N \times N} \mid MXM^T = X \text{ for each } M \in \mathcal{G}\}$$

(the centralizer algebra of \mathcal{G}).

Often, $\dim(C_{\mathcal{G}})$ is much smaller than N^2 , and we might reduce the dimension of the search space. How?

First two examples.

Coding

$A(n, d) :=$ the maximum number of words in $\{0, 1\}^n$ at mutual Hamming distances $\geq d$.

$$A(n, d) \leq \max\{\text{tr}(JX) \mid X \in \mathbb{R}_+^{\{0,1\}^n \times \{0,1\}^n},$$

$$X \text{ PSD, } \text{tr}(X) = 1,$$

$$X_{u,v} = 0 \text{ if } 0 < d(u, v) < d \text{ for all } u, v \in \{0, 1\}^n,$$

$$X_{u,v} = X_{u,u+v} \text{ for all } u, v \in \{0, 1\}^n,$$

$$\tilde{X} \geq 0, \tilde{X} \text{ PSD}\}.$$

Here J is the all-one matrix, and

$$\tilde{X}_{u,v} := X_{u+v,u+v} - X_{u,v}.$$

Proof of the bound:

If $C \subseteq \{0, 1\}^n$ has minimum distance $\geq d$, define

$$X := |C|^{-2} \sum_{u \in C} \chi^{u+C} (\chi^{u+C})^T.$$

Then X is a feasible solution and $\text{tr}(JX) = |C|$.

Group acting on the problem:

For each $\pi \in S_n$, let M_π be the $\{0, 1\}^n \times \{0, 1\}^n$ permutation matrix corresponding to the permutation

$$(u_1, \dots, u_n) \mapsto (u_{\pi(1)}, \dots, u_{\pi(n)})$$

for $(u_1, \dots, u_n) \in \{0, 1\}^n$.

Let $\mathcal{G} := \{M_\pi \mid \pi \in S_n\}$.

Crossing number

Let $\text{cr}(K_{m,n})$ be the *crossing number* of $K_{m,n}$.

‘Zarankiewicz conjecture’:

$$\text{cr}(K_{m,n}) = \lfloor \frac{1}{4}(m-1)^2 \rfloor \lfloor \frac{1}{4}(n-1)^2 \rfloor.$$

Let Z_n be the collection of cyclic permutations in S_n .

Theorem (de Klerk et al.):

$$\text{cr}(K_{m,n}) \geq \frac{1}{2}m^2\alpha_n - \frac{1}{2}m \lfloor \frac{1}{4}(n-1)^2 \rfloor, \text{ where}$$

$$\alpha_n := \min\{\text{tr}(CX) \mid X \in \mathbb{R}_+^{Z_n \times Z_n} \text{ PSD}, \text{tr}(JX) = 1\}.$$

Here C is the matrix in $\mathbb{R}^{Z_n \times Z_n}$ with, for $\sigma, \tau \in Z_n$:

$C_{\sigma,\tau} :=$ the minimum number of crossings of $K_{2,n}$ such that the edges leaving the two n -degree vertices in clockwise order go to $\sigma(1), \dots, \sigma(n)$ and to $\tau(1), \dots, \tau(n)$ respectively.

(Indicating the 2-degree vertices by $1, \dots, n$.)

Proof of the bound:

Let $K_{m,n}$ be embedded with a minimum number of crossing.

For each $\sigma \in Z_n$, let d_σ be the number of n -degree vertices such that the edges leaving it go in clockwise order to $\sigma(1), \dots, \sigma(n)$.

Then $X := m^{-2}dd^T$ is feasible and

$$\text{cr}(K_{m,n}) \geq \frac{1}{2}m^2 \text{tr}(CX) - \frac{1}{2}m \lfloor \frac{1}{4}(n-1)^2 \rfloor.$$

Group acting on the problem:

For each $\pi \in S_n$, let M_π be the $Z_n \times Z_n$ permutation matrix corresponding to the permutation

$$\sigma \mapsto \pi^{-1}\sigma\pi$$

for $\sigma \in Z_n$.

Let $\mathcal{G} := \{M_\pi \mid \pi \in S_n\}$.

C_G is a *matrix $*$ -algebra*, that is, it is closed under addition, scalar and matrix multiplication, and adjunction.

Two reduction methods:

- I. Block diagonalization
- II. The regular representation

Recall that a complex matrix U is called *unitary* if $U^*U = I$.

So

$$\dim(\mathcal{A}) = p_1^2 + \cdots + p_m^2$$

and

$$N = p_1 q_1 + \cdots + p_m q_m.$$

Moreover:

$A \in \mathcal{A}$ is PSD \iff each block B_i of U^*AU is PSD.

In the coding example, C_G is the *Terwilliger algebra*:

$$\mathcal{T}_n := \left\{ \sum_{k,i,j} x_{k,i,j} E_{k,i,j} \mid x_{k,i,j} \in \mathbb{R} \right\}, \text{ where}$$

$$(E_{k,i,j})_{u,v} := \begin{cases} 1 & \text{if } d(u,v) = k, |u| = i, |v| = j, \\ 0 & \text{else.} \end{cases}$$

It turns out that there are blocks $B_0, \dots, B_{\lfloor \frac{1}{2}n \rfloor}$,

such that block B_t of $U^* \left(\sum_{k,i,j} x_{k,i,j} E_{k,i,j} \right) U$ is:

$$B_t := \left(\sum_k \gamma_{k,i,j}^{(t)} x_{k,i,j} \right)_{i,j=t}^{n-t}$$

where $\gamma_{k,i,j}^{(t)} := \sum_u (-1)^{u - \frac{1}{2}(i+j-k)} \binom{u}{\frac{1}{2}(i+j-k)} \binom{n-2t}{u-t} \binom{n-t-u}{i-u} \binom{n-t-u}{j-u}$

$$A(n, d) \leq \max \left\{ \sum_k x_{0,k,k} \mid \right.$$

$$(i) \ x_{0,0,0} = 1,$$

$$(ii) \ x_{k,i,j} = x_{k,j,i} = x_{i,j,k} \geq 0 \text{ for all } k, i, j,$$

$$(iii) \ x_{k,i,j} = 0 \text{ if } \{k, i, j\} \cap \{1, \dots, d-1\} \neq \emptyset,$$

$$(iv) \text{ for each } t = 0, \dots, \lfloor \frac{1}{2}n \rfloor, \text{ the matrices}$$

$$\left(\sum_k \rho_{k,i,j}^{(t)} x_{k,i,j} \right)_{i,j=t}^{n-t}$$

and

$$\left(\sum_k \rho_{k,i,j}^{(t)} (x_{0,k,k} - x_{k,i,j}) \right)_{i,j=t}^{n-t}$$

are PSD }

$$\text{Here } \rho_{k,i,j}^{(t)} := \gamma_{k,i,j}^{(t)} \binom{n}{\frac{1}{2}(i+j-k), \frac{1}{2}(i+k-j), \frac{1}{2}(j+k-i)}^{-1}.$$

II. THE REGULAR REPRESENTATION

There exist nonzero 0, 1 matrices E_1, \dots, E_d such that

$$C_{\mathcal{G}} = \left\{ \sum_{i=1}^d x_i E_i \mid x_1, \dots, x_d \in \mathbb{R} \right\}$$

and

$$E_1 + \dots + E_d = J \text{ (the all-one matrix).}$$

(So $d = \dim(C_{\mathcal{G}})$.)

Suppose now that we can identify these matrices, and that we also can determine the *multiplication parameters* $\mu_{i,j}^k$ (for $i, j, k = 1, \dots, d$), defined by:

$$E_i E_j = \sum_{k=1}^d \mu_{i,j}^k E_k.$$

For $k = 1, \dots, d$, define the $d \times d$ matrix L_k by

$$(L_k)_{i,j} := \frac{\text{tr}(E_i E_i^T)^{1/2}}{\text{tr}(E_j E_j^T)^{1/2}} \mu_{k,j}^i$$

for $i, j = 1, \dots, d$.

$$\text{Let } \mathcal{L} := \left\{ \sum_{i=1}^d x_i L_i \mid x_1, \dots, x_d \in \mathbb{R} \right\}.$$

Then \mathcal{L} is a matrix $*$ -algebra, and $\phi : C_G \rightarrow \mathcal{L}$ defined by

$$\phi \left(\sum_i x_i E_i \right) := \sum_i x_i L_i$$

is an algebra $*$ -isomorphism.

This implies:

$$\sum_i x_i E_i \text{ PSD} \iff \sum_i x_i L_i \text{ PSD}.$$

(\mathcal{L} corresponds to the *regular representation* of C_G .)

Application to the crossing number gives:

$$\alpha_9 = 7.7352126 \dots$$

This implies, for each fixed $n \geq 9$:

$$\lim_{m \rightarrow \infty} \frac{\text{cr}(K_{m,n})}{\lfloor \frac{1}{4}(m-1)^2 \rfloor \lfloor \frac{1}{4}(n-1)^2 \rfloor} \geq 0.8303.$$