Facial reduction for symmetry reduced semidefinite programs

Hao Hu[a,b] , Renata Sotirov[a] and Henry Wolkowicz[b]

Updated: 2020/10/17

[a] Tilburg University
[b] University of Waterloo

Semidefinite programming (SDP)

- The semidefinite program in standard form is

$$\inf_{X \in \mathcal{S}^n} \{\langle C, X \rangle \mid \mathcal{A}(X) = b, \ \ X \in \mathcal{S}^n_+\}$$

- The feasible region $\mathcal{F} := \{X \mid \mathcal{A}(X) = b, \ X \in \mathcal{S}^n_+\}$ is a spectrahedron
- The SDPs are expensive to solve
- Modelling is crucial to numerical stability
- Symmetry reduction reduces the size of SDP formulations
- Facial reduction ensures an SDP instance can be solved correctly

Symmetry reduction (2 minutes crash course)

- Recall that the feasible region is

$$\mathcal{F} = \{X \mid \mathcal{A}(X) = b,\ X \in \mathcal{S}_+^n\}$$

- In symmetry reduction, we try to find an orthogonal matrix Q such that, for any feasible $X \in \mathcal{F}$, the orthogonal transformation $Q^T X Q$ is a block-diagonal matrix, i.e.,

$$Q^T X Q = \begin{bmatrix} X_1 & 0 & \cdots & 0 \\ 0 & X_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & X_t \end{bmatrix} \in \mathcal{S}_+^n,$$

where $X_i \in \mathcal{S}_+^{n_i}$ is the i-th block and $\sum_{i=1}^t n_i = n$

- As $X \in \mathcal{S}_+^n$ if and only if $X_i \in \mathcal{S}_+^{n_i}$ for $i = 1, \ldots, t$, the feasible region can be equivalently reformulated as

$$\mathcal{F} = \{X \mid \mathcal{A}(X) = b,\ X_i \in \mathcal{S}_+^{n_i} \text{ for } i = 1, \ldots, t\}$$

- For example, if $n_i = 1$ and thus $t = n$, the SDP collapses to a linear program

An example of symmetry reduction

- An SDP relaxation for the cut minimization problem *(Pong et al. '14)*

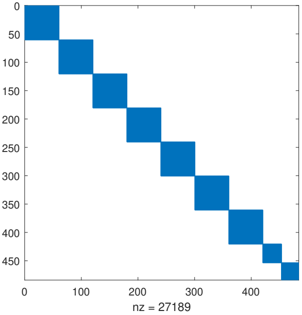
$$
\begin{array}{ll}
\min_X & \langle C, X \rangle \\
\text{s.t.} & \mathcal{A}(X) = b, X \geq 0 \\
& X \in \mathcal{S}_+^{nk},
\end{array}
$$

where n is the number of vertices and k is the number of subsets in the partition

- The problem instance **can161** has $n = 161$ vertices and $k = 3$ partitions
- The size of $X \in \mathcal{S}_+^{nk}$ is $nk = 483$ for **can161**, and very difficult to solve ☹

An example of symmetry reduction

- The feasible solutions X under certain unitary transformation, i.e., $Q^T XQ$, has the following **block-diagonal structure**



- The sizes of these 9 blocks are $60, 60, 60, 60, 60, 60, 60, 33, 30$

An example of symmetry reduction

- The symmetry reduced SDP relaxation for cut minimization problems

$$\begin{aligned} \min_X \quad & \langle C, X \rangle \\ \text{s.t.} \quad & \mathcal{A}(X) = b, X \geq 0 \\ & \cancel{X \in \mathcal{S}_+^{483}} \\ & \begin{cases} X_1 \in \mathcal{S}_+^{60} \\ \quad\quad \vdots \\ X_9 \in \mathcal{S}_+^{30} \end{cases} \end{aligned}$$

- Before and after symmetry reduction,

	the sizes of p.s.d. constraints
Original SDP	483
Symmetry reduced	60, 60, 60, 60, 60, 60, 60, 33, 30

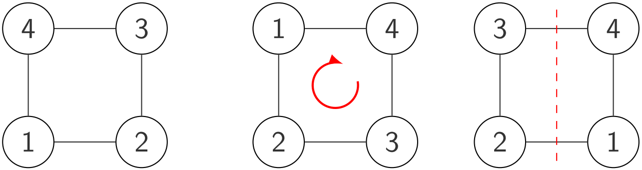
Instance **can161**

- The symmetry reduced SDP can be solved in a laptop in few minutes ☺

How do we find the orthogonal matrix Q?

- Many problems have a natural "combinatorial symmetry"

- For example, the input graph is "invariant" under certain permutations of its vertices



Input graph | rotate clock-wise | flip horizontally

- The symmetries in the problem translates into symmetries in the data matrices of its SDP relaxation

- A set $\mathcal{M} \subseteq \mathbb{C}^{n \times n}$ is a **matrix $*$-algebra** over $\mathbb{C}$ if it is closed under addition, scalar and matrix multiplication, and taking conjugate transpose

- The data matrices $C, A_1, \ldots, A_m$ are contained in a matrix $*$-algebra

How do we find the orthogonal matrix Q?

- Theorem (*de Klerk 2009)* Assume the data matrices $C, A_1, \ldots, A_m$ and the identity matrix are contained in a matrix $*$-algebra $\mathcal{M}$. Then it has an optimal solution in $\mathcal{M}$, if the SDP has an optimal solution.

- Theorem (*Wedderburn 1907*) There exists a unitary matrix Q such that the matrix $*$-algebras $\mathcal{M}$ containing the identity matrix can be decomposed as

$$Q^*\mathcal{M}Q = \begin{bmatrix} \mathcal{M}_1 & 0 & \cdots & 0 \\ 0 & \mathcal{M}_2 & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \cdots & 0 & \mathcal{M}_t \end{bmatrix}, \text{ where each } \mathcal{M}_i \subseteq \mathbb{C}^{n_i \times n_i} \text{ is } \textbf{basic}$$

- Symmetry in the problem $\Rightarrow$ Data matrices in the algebra $\Rightarrow$ Optimal solution in the algebra $\Rightarrow$ Decomposition using Wedderburn

- There are more general approaches to find the decomposition, see *Permenter and Parrilo 2016*

Facial reduction

Facial reduction

- **Slater's condition** (strict feasibility) holds for semidefinite program in standard form if there exists X such that

$$\mathcal{A}(X) = b,\ X \in \mathcal{S}^n_{++}$$

- Without strict feasibility:
  - the KKT conditions may not be necessary for the optimality
  - strong duality may not hold
  - small perturbations may render the problem infeasible
  - many solvers might run into numerical errors

- **Facial reduction** is a regularization technique that can be used for semidefinite programs that fail strict feasibility (*Borwein, Wolkowicz, '81*)

- The loss of strict feasibility is **common** in SDP relaxations for hard nonconvex problems

Facial reduction

- Given the SDP in standard form

$$\inf_X \{\langle C, X\rangle \mid \mathcal{A}(X) = b, X \in \mathcal{S}_+^n\} \tag{1}$$

Then exactly one of the following alternatives holds

1. The SDP (1) is strictly feasible:

$$\mathcal{A}(X) = b,\ X \in \mathcal{S}_{++}^n$$

2. The auxiliary system is consistent:

$$0 \neq \mathcal{A}^*(y) \in \mathcal{S}_+^n \text{ and } \langle b, y\rangle = 0$$

- We call $\mathcal{A}^*(y)$ an **exposing vector**

- The feasible region of (1) is contained in $\mathcal{A}^*(y)^\perp \cap \mathcal{S}_+^n$, thus we can reduce the problem size

- Facial reduction algorithm finds an exposing vector $\mathcal{A}^*(y)$, and repeat until a strictly feasible problem is obtained

Facial reduction for the cut minimization problem

- The SDP relaxation for the cut minimization problem *(Pong et al. '14)*

$$\begin{array}{ll} \min_X & \langle C, X \rangle \\ \text{s.t.} & \mathcal{A}(X) = b, X \geq 0 \\ & X \in \mathcal{S}_+^{nk} \end{array}$$

- After facial reduction, we obtain

$$\begin{array}{ll} \min_X & \langle C, X \rangle \\ \text{s.t.} & \mathcal{A}(X) = b, X \geq 0 \\ & \cancel{X \in \mathcal{S}_+^{nk}} \implies X = VRV^T, R \in \mathcal{S}_+^{(n-1)(k-1)} \end{array}$$

where the columns of V span $\mathcal{A}^*(y)^\perp$

	the sizes of p.s.d. constraints
Original SDP	483
Facially reduced	321
Symmetry reduced	60, 60, 60, 60, 60, 60, 60, 33, 30

Instance **can161**

Facial reduction for symmetry reduced SDP

Facial reduction for symmetry reduced SDP

- The main issue:
  1. In theory, there is no problem to apply both facial reduction and symmetry reduction
  2. In practice, it is difficult as one loses the structure after facial or symmetry reduction
- The existing situation:
  1. With only facial reduction, we are only able to solve small instances
  2. With only symmetry reduction, the solution we found may be very inaccurate
- Our contribution:

A method to find the facially and symmetry reduced program

Facial reduction for symmetry reduced SDP

Theorem (H., Sotirov, Wolkowicz) Let W be an exposing vector of the minimal face of a given SDP instance. Then

1. There exists an exposing vector $W_{\mathcal{G}} \in \mathcal{M}$ of the minimal face of the input SDP instance
2. $Q^T W_{\mathcal{G}} Q$ is an exposing vector of the minimal face of the symmetry reduced SDP

- In plain words, we know how to do facial reduction for the symmetry reduced SDP

Facial reduction for the symmetry reduced program

- The SDP relaxation for the cut minimization problem *(Pong et al. '14)*

$$\min_X \langle C, X\rangle \text{ subject to } \mathcal{A}(X) = b,\ X \geq 0,\ X \in \mathcal{S}_+^{nk}$$

- The symmetry reduced program is

$$\min_X \langle C, X\rangle \text{ subject to } \mathcal{A}(X) = b,\ X \geq 0,\ X_i \in \mathcal{S}_+^{n_i} \text{ for } i = 1, \ldots, t$$

for some $n_i \ll n$ and $\sum_{i=1}^t n_i = n$

- The facially and symmetry reduced program is

$$\min_X \langle C, X\rangle \text{ subject to } \mathcal{A}(X) = b,\ X \geq 0,\ X_i = V_i R_i V_i^T,\ R_i \in \mathcal{S}_+^{r_i} \text{ for } i = 1, \ldots, t$$

for some $r_i \leq n_i \ll n$ and $\sum_{i=1}^t r_i < \sum_{i=1}^t n_i = n$

Facial reduction for symmetry reduced SDP

- For the **can161** instance in the cut minimization problem, we obtain

	the sizes of p.s.d. constraints
Original SDP	483
Facially reduced	321
Symmetry reduced	60, 60, 60, 60, 60, 60, 60, 33, 30
Facially + Symmetry	40, 40, 40, 40, 38, 40, 40, 21, 20

Instance **can161**

- Now lets check if our theory works?

Numerical results on the cut minimization problem

- We solve the SDP relaxation from Pong et al. '14 using interior point method
-

Instance		Symmetry	Facial+Symmetry
can144	bound	0.3838	0.6233
	iteration	35	18
	time	32.27s	5.8s
	solver output	fail	success
can161	bound	0.4828	0.5485
	iteration	24	20
	time	375.63s	108.05s
	solver output	fail	success

- The bounds in the symmetry reduced program is NOT CORRECT, as the solver couldn't solve it accurately due to numerical instability
- The facially and symmetry reduced program is solved CORRECTED, it also takes less iteration and time

Alternating direction method of multipliers (ADMM)

- Our technique fits surprisingly well in an ADMM approach
- Symmetry and facial reduction have a natural splitting, and it results in extremely cheap update in ADMM
- We are able to solve some huge SDP instances for the quadratic assignment problem
-

		Mittelmann&Peng		Hu,Sotirov&Wolkowicz	
Instance	Upper bound	bound	time	bound	time
Harper128	2479944	2446944	1491s	2437880	186s
Harper256	22370940	-	-	22205236	432s
Harper516	201329908	-	-	200198783	1903s

- The interior point method cannot solve this QAP with $n = 30$
- The best algorithm in the literature takes 1640 seconds to solve the same SDP relaxation for an QAP instance with $n = 64$

Summary

Input 1	Input 2	Output
SDP	matrix $*$-algebra	symmetry reduced SDP $+$ reduced problem size – numerical issues
SDP	exposing vector	facially reduced SDP $+$ numerically stable – symmetry not exploited
symmetry reduced SDP	exposing vector in the algebra	facially & symmetry reduced SDP $+$ numerically stable $+$ reduced problem size