Block Factor-width-two Matrices and Their Applications to Semidefinite and Sum-of-squares Optimization

Yang Zheng

Assistant Professor, ECE, UC San Diego

(Joint work with Aivar Sootla, and Antonis Papachristodoulou at University of Oxford)

UC San Diego
JACOBS SCHOOL OF ENGINEERING
Electrical and Computer Engineering

2022 INFORMS Annual Meeting

October 18, 2022

Outline

Introduction: inner/outer approximations for SDPs

A new class of block factor-width-two matrices

Applications to SDPs and SOS optimization

Conclusions

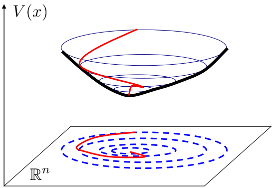
Introduction

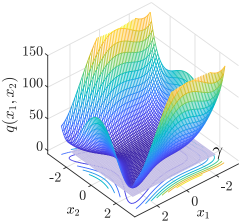
A primal standard SDP is in the form of

$$\begin{aligned} p^\star = \min_{X} \quad & \langle C, X \rangle \\ \text{subject to} \quad & \langle A_i, X \rangle = b_i, i = 1, \ldots, m, \\ & X \in \mathbb{S}^n_+. \end{aligned}$$

SDPs are more powerful than LP or SOCP

- Applications: control theory, polynomial optimization, machine learning, power grid, robotics, etc.





Introduction

A primal standard SDP is in the form of

$$\begin{aligned} p^\star = \min_{X} \quad & \langle C, X \rangle \\ \text{subject to} \quad & \langle A_i, X \rangle = b_i, i = 1, \ldots, m, \\ & X \in \mathbb{S}^n_+. \end{aligned}$$

However, SDPs are much more expensive to solve than LP or SOCP

- Standard interior-point methods scale as $\mathcal{O}(mn^3 + m^2n^2)$ per iteration
- LPs with millions of variables and constraints can be solved reliably.
- General-purpose solvers cannot efficiently handle large SDP problems ($n \approx 1000$, and m: a few thousands)
- Exploiting sparsity and structures for improving efficiency is an active research topic[1,2].

[1] Zheng, Fantuzzi, and Papachristodoulou. "Chordal and factor-width decompositions for scalable semidefinite and polynomial optimization." Annual Reviews in Control 52 (2021): 243-279.

[2] Majumdar, Hall, and Ahmadi. "Recent scalability improvements for semidefinite programming with applications in machine learning, control, and robotics." Annual Review of Control, Robotics, and Autonomous Systems 3 (2020): 331-360.

Something simpler: inner/outer approximations

Inner approximations

- Suppose we have a simpler cone $\mathcal{K} \subset \mathbb{S}^n_+$. Solving an instance of

$$\begin{aligned} \min_{X} \quad & \langle C, X \rangle \\ \text{subject to} \quad & \langle A_i, X \rangle = b_i, i = 1, \ldots, m, \\ & X \in \mathcal{K} \subset \mathbb{S}^n_+. \end{aligned}$$

gives us an upper bound on $p^\star$.

Outer approximations

- Suppose we have a simpler cone $\mathbb{S}^n_+ \subset \hat{\mathcal{K}}$. Solving an instance of

$$\begin{aligned} \min_{X} \quad & \langle C, X \rangle \\ \text{subject to} \quad & \langle A_i, X \rangle = b_i, i = 1, \ldots, m, \\ & X \in \hat{\mathcal{K}}. \end{aligned}$$

gives us a lower bound on $p^\star$.

Which cones to use?

Ahmadi and Majumdar[3] considered the cones of *diagonally dominant* and *scaled diagonally dominant* matrices

- A symmetric matrix $A \in \mathbb{S}^n$ is *diagonally dominant* if

$$a_{ii} \geq \sum_{j \neq i} |a_{ij}|, \qquad i = 1, \ldots, n.$$

- A symmetric matrix $A \in \mathbb{S}^n$ is *scaled-diagonally dominant* if there exists a diagonal matrix D with nonnegative entries, such that

$$DAD \quad \text{is diagonally dominant}$$

We denote

$$\mathcal{DD}_n = \{X \in \mathbb{S}^n \mid X \text{ is diagonally dominant}\} \subset \mathbb{S}^n_+$$
$$\mathcal{SDD}_n = \{X \in \mathbb{S}^n \mid X \text{ is scaled diagonally dominant}\} \subset \mathbb{S}^n_+$$

- Linear optimization over $\mathcal{DD}_n$ is an LP;
- Linear optimization over $\mathcal{SDD}_n$ is an SOCP;

[3] Ahmadi, and Majumdar. "DSOS and SDSOS optimization: more tractable alternatives to sum of squares and semidefinite optimization." SIAM J. Appl. Algebra Geom 3.2 (2019): 193-230.

Approximation and Conservativeness

$$\begin{aligned} \min_{X} \quad & \langle C, X \rangle \\ \text{subject to} \quad & \langle A_i, X \rangle = b_i, i = 1, \ldots, m, \\ & X \in \mathcal{SDD}_n \quad (\text{or } \mathcal{DD}_n). \end{aligned}$$

- **First**: the resulting upper bound might be very conservative.

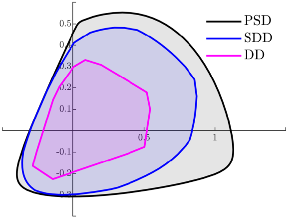


Figure: Boundary of (x, y) such that $I_6 + xA + xB$ is PSD, SDD, or DD.

- **Second**: using $\mathcal{SDD}_n$ requires a number of $\mathcal{O}(n^2)$ small SOCP constraints, which might be an issue for large n.

Outline

Introduction: inner/outer approximations for SDPs

A new class of block factor-width-two matrices

Applications to SDPs and SOS optimization

Conclusions

Factor-width-two matrices

Lemma

$\mathcal{SDD}_n$ is equivalent to the cone of factor-width-two matrices (Boman, Erik G., et al. 2005)

$$\mathcal{SDD}_n = \left\{ \sum_{i \in I} x_i x_i^\mathsf{T} \mid x_i \in \mathbb{R}^n, \textit{supp}(x_i) \leq 2 \right\}.$$

- Denote the cone of $n \times n$ *factor-width-two* matrices as $\mathcal{FW}_2^n$:

$$\mathcal{SDD}_n = \mathcal{FW}_2^n.$$

- **Another interpretation**: $P \in \mathcal{FW}_2^n$ if and only if there exists $X_{ij} \in \mathbb{S}_+^2$ s.t.

$$P = \sum_{1 \leq i < j \leq n} E_{ij}^\mathsf{T} X_{ij} E_{ij},$$

where $E_{ij} = \begin{bmatrix} E_i \\ E_j \end{bmatrix} \in \mathbb{R}^{2 \times n}, i \neq j$, and E_i is a row basis vector with 1 at the i-th entry.

$$E_i = \begin{bmatrix} 0 & \dots & 1 & \dots & 0 \end{bmatrix} \in \mathbb{R}^{1 \times n}.$$

Block factor-width-two matrices

A linear optimization problem over the FW_2^n cone can be written as an SDP over the cone product

$$\underbrace{\mathbb{S}_+^2 \times \ldots \times \mathbb{S}_+^2}_{\binom{n}{2}}$$

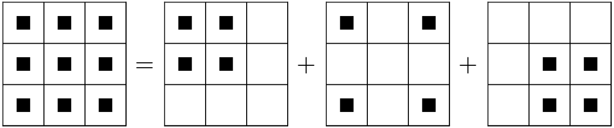


Figure: Illustration of (block) factor-width-two decomposition

- For $A \in \mathcal{SDD}_n$ or $\mathcal{FW}_2^n$, each black square represents a scalar $a_{ij} \in \mathbb{R}$.
- **Key idea – block extension**: how about each black square represents a submatrix?

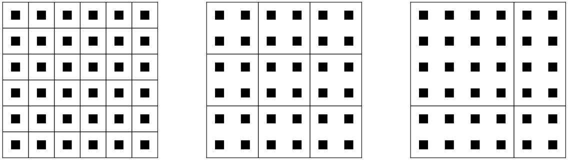
$$A_{ij} \in \mathbb{R}^{k_i \times k_j}.$$

Block-partitioned matrices

Given a matrix $A \in \mathbb{R}^{n \times n}$, we say a set of integers $\alpha = \{k_1, k_2, \ldots, k_p\}$ with $k_i \in \mathbb{N}\ (i = 1, \ldots, p)$ is a partition of A if $\sum_{i=1}^{p} k_i = n$, and A is partitioned as

$$\begin{bmatrix} A_{11} & A_{12} & \ldots A_{1p} \\ A_{21} & A_{22} & \ldots A_{2p} \\ \vdots & \vdots & \ddots \vdots \\ A_{p1} & A_{p2} & \ldots A_{pp} \end{bmatrix},$$

with $A_{ij} \in \mathbb{R}^{k_i \times k_j}, \forall i, j = 1, \ldots, p$.



(a) (b) (c)

Figure: Different partitions for a 6×6 matrix: (a) $\alpha = \{1, 1, 1, 1, 1, 1\}$, (b) $\beta = \{2, 2, 2\}$, (c) $\gamma = \{4, 2\}$. From right to left, we get coarser partitions, *i.e.* $\alpha \sqsubseteq \beta \sqsubseteq \gamma$.

Block factor-width-two matrices

Definition

A symmetric matrix $Z \in \mathbb{S}^n$ with partition $\alpha = \{k_1, k_2, \ldots, k_p\}$ belongs to the class of block factor-width-two matrices, denoted as $\mathcal{FW}^n_{\alpha,2}$, if and only if

$$Z = \sum_{1 \le i < j \le p} (E^\alpha_{ij})^\mathsf{T} X_{ij} E^\alpha_{ij} \qquad (1)$$

for some $X_{ij} \in \mathbb{S}^{k_i+k_j}_+$ and with $E^\alpha_{ij} \in \mathbb{R}^{(k_i+k_j)\times n}$ being an index matrix.

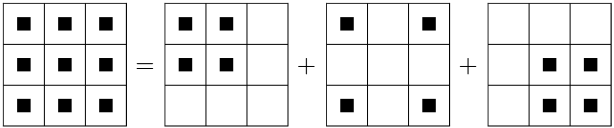


Figure: Illustration of block factor-width-two decomposition (1). The (i, j) black square represents a submatrix of dimension $k_i \times k_j, i, j = 1, 2, 3.$

A hierarchy of inner approximations

A **finer partition** (or subpartition) of $\alpha = \{k_1, k_2, \ldots, k_p\}$ is a partition that breaks some blocks of α into smaller blocks.

- Let $\alpha = \{1,1,1,1,1,1\}$, $\beta = \{2,2,2\}$ and $\gamma = \{4,2\}$. Denote $\alpha \sqsubseteq \beta \sqsubseteq \gamma$.

Theorem

Let $\alpha \sqsubseteq \beta \sqsubseteq \gamma$ be partitions of n with $\gamma = \{\gamma_1, \gamma_2\}$, and let $\mathbf{1} = \{1, \ldots, 1\}$ denote the uniform unit partition. Then,

$$\mathcal{SDD}_n = \mathcal{FW}^n_{\mathbf{1},2} \subseteq \mathcal{FW}^n_{\alpha,2} \subseteq \mathcal{FW}^n_{\beta,2} \subseteq \mathcal{FW}^n_{\gamma,2} \equiv \mathbb{S}^n_+$$

This flexibility of $\mathcal{FW}^n_{\alpha,2}$ improves the two drawbacks of $\mathcal{SDD}_n$. As the number p in a partition α decreases:

- *First – the approximation quality improves:* the largest distance between a unit-norm matrix in $\mathbb{S}^n_+$ and the cone $\mathcal{FW}^n_{\alpha,2}$ satisfies

$$\text{dist}(\mathbb{S}^n_+, \mathcal{FW}^n_{\alpha,2}) \leq \frac{p-2}{p}$$

- *Second – the number of blocks in the summation decreases* $\binom{p}{2}$ vs. $\binom{n}{2}$

Example

Consider the 5×5 matrix

$$P(x,y) = \begin{bmatrix} 1+6x+4y & 3x+y & 2x+y & x+4y & 3x+3y \\ 3x+y & 1+6y & 5x+3y & y & 2x+2y \\ 2x+y & 5x+3y & 1+2x+2y & x+2y & 5x+6y \\ x+4y & y & x+2y & 1+2x & 3x+3y \\ 3x+3y & 2x+2y & 5x+6y & 3x+3y & 1+6x+2y \end{bmatrix}$$

and partitions $\mathbf{1} = \{1,1,1,1,1\}$, $\alpha = \{2,1,1,1\}$, $\beta = \{2,1,2\}$ and $\gamma = \{2,3\}$.

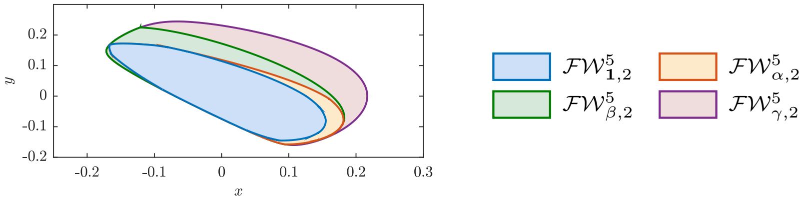


Figure: Regions of the (x,y) plane for which $P(x,y)$ belongs to the block factor-width-two cones $\mathcal{SDD}_5 \equiv \mathcal{FW}^5_{\mathbf{1},2} \subseteq \mathcal{FW}^5_{\alpha,2} \subseteq \mathcal{FW}^5_{\beta,2} \subseteq \mathcal{FW}^5_{\gamma,2} \equiv \mathbb{S}^5_+$

- The inclusions of the plotted regions reflect the inclusions of the cones and the order relation $\mathbf{1} \sqsubseteq \alpha \sqsubseteq \beta \sqsubseteq \gamma$.

Outline

Introduction: inner/outer approximations for SDPs

A new class of block factor-width-two matrices

Applications to SDPs and SOS optimization

Conclusions

Applications to SDPs

- The cones $\mathcal{FW}^n_{\alpha,2}$ and its dual $(\mathcal{FW}^n_{\alpha,2})^*$ approximate the positive semidefinite cone $\mathbb{S}^n_+$ from the inside and from the outside
- The approximation improves as the partition α is coarsened.
- This leads to convergent sequences of upper and lower bounds on the optimal value p^*

Inner approximations

$$\begin{aligned} U_\alpha := \min_X \quad & \langle C, X\rangle \\ \text{subject to} \quad & \langle A_i, X\rangle = b_i,\ i = 1,\dots,m, \\ & X \in \mathcal{FW}^n_{\alpha,2} \end{aligned}$$

Outer approximations

$$\begin{aligned} L_\alpha := \min_X \quad & \langle C, X\rangle \\ \text{subject to} \quad & \langle A_i, X\rangle = b_i,\ i = 1,\dots,m, \\ & X \in (\mathcal{FW}^n_{\alpha,2})^* \end{aligned}$$

Corollary

Let $\alpha_1 \sqsubseteq \alpha_2 \sqsubseteq \ldots \sqsubseteq \alpha_k = \{\alpha_{k1}, \alpha_{k2}\}$ *be a sequence of partitions of* n. *Then,*

$$L_{\alpha_1} \le \cdots \le L_{\alpha_k} = p^* = U_{\alpha_k} \le \cdots \le U_{\alpha_1}.$$

Applications to SOS optimization

- **Sum-of-squares (SOS) polynomials**: $p(x)$ can be represented as

$$p(x) = \sum_{i=1}^{m} f_i^2(x),$$

- **SDP characterization (Parrilo, Lasserre *etc.*):** $p(x)$ is SOS if and only if

$$p(x) = v_d(x)^T Q v_d(x), \qquad Q \succeq 0$$

where $v_d(x)$ is the standard monomial basis.

- **SOS:** $p(x) = v_d(x)^T Q v_d(x) : Q$ is PSD $\longrightarrow$ SDP
- **SDSOS:** $p(x) = v_d(x)^T Q v_d(x) : Q$ is sdd $\longrightarrow$ SOCP
- **DSOS:** $p(x) = v_d(x)^T Q v_d(x) : Q$ is dd $\longrightarrow$ LP
- **α-SDSOS:** $p(x) = v_d(x)^T Q v_d(x) : Q$ is $\mathcal{FW}_{\alpha,2}^N$ $\longrightarrow$ SDP with small blocks

Numerical examples

Consider a scalar polynomial optimization problem:

$$\begin{aligned} \min_{\gamma} \quad & \gamma \\ \text{subject to} \quad & p(x) + \gamma \geq 0, \;\; \forall x \in \mathbb{R}^n, \end{aligned}$$

Table: Computational results using SOS and α-SDSOS relaxations.

n	Full SDP	Number of blocks p in partition α: 4	10	20	50
Computational time (seconds)					
15	27.3	23.3	15.6	10.1	5.36
20	489	252	98.1	66.8	28.1
25	∞	1970	783	571	132
30	∞	∞	5680	3710	840
Objective values γ					
15	-0.92	-0.75	80.1	459	2240
20	-0.87	-0.87	-0.11	251	1910
25	∞	-1.07	-0.21	231	1360
30	∞	∞	-0.37	177	1770

Numerical results

Consider a matrix SOS program

$$\min_{\gamma} \quad \gamma$$
$$\text{subject to} \quad P(x) + \gamma I \succeq 0, \quad \forall x \in \mathbb{R}^3,$$

where $P(x)$ is an $r \times r$ polynomial matrix with each element being a quartic polynomial in three variables.

Table: Computational results using SOS, α-SDSOS and SDSOS relaxations.

r	25	30	35	40	45	50
Computational time (seconds)						
SOS	14.4	35.9	87.2	175.0	316.0	487.8
α-SDSOS	10.8	16.6	25.3	36.0	57.4	71.4
SDSOS	1.1	1.3	1.6	2.1	2.6	3.3
Objective value γ						
SOS	266.5	316.2	460.8	562.0	746.9	919.8
α-SDSOS	266.5	316.2	460.8	562.0	746.9	919.8
SDSOS	270.3	324.8	477.7	570.9	762.2	961.7

An iterative scheme

- Ahmadi and Hall[4] introduced an iterative method over $\mathcal{SDD}_n$ and $\mathcal{DD}_n$.
- This iterative scheme can be naturally extended to the class of block factor-width-two matrices.

$$\begin{aligned} \min_{x,y} \quad & -x-y \\ \text{subject to} \quad & I + xA + yB \succeq 0, \end{aligned} \tag{2}$$

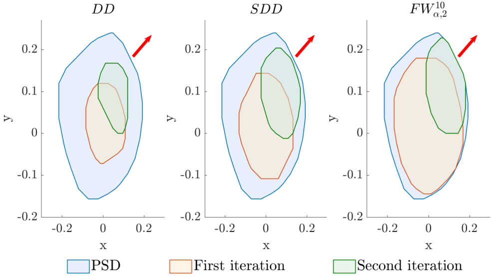


Figure: Feasible regions of inner approximations using $\mathcal{DD}_n$, $\mathcal{SDD}_n$, and $\mathcal{FW}_{\alpha,2}^n$.

[4] Ahmadi and Hall. "Sum of squares basis pursuit with linear and second order cone programming." Algebraic and geometric methods in discrete mathematics 2017: 27-53.

An iterative scheme

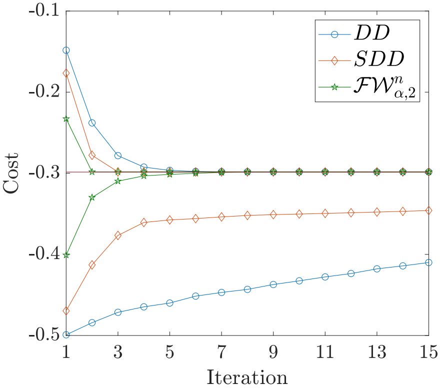


Figure: Inner/Outer approximations of the SDP (2) using $\mathcal{DD}_n$, $\mathcal{SDD}_n$, and $\mathcal{FW}^n_{2,\alpha}$ with $\alpha = \{2, \ldots, 2\}$[5].

[5] Liao, Feng-Yi, and Yang Zheng. "Iterative Inner/outer Approximations for Scalable Semidefinite Programs using Block Factor-width-two Matrices." arXiv preprint arXiv:2204.06759 (2022).

Outline

Introduction: inner/outer approximations for SDPs

A new class of block factor-width-two matrices

Applications to SDPs and SOS optimization

Conclusions

Summary

A new class of block factor-width-two matrices

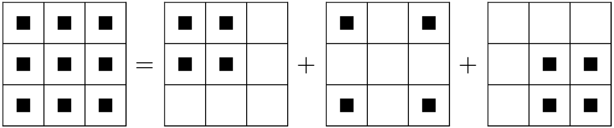


Figure: Illustration of (block) factor-width-two decomposition

- It offers a new hierarchy of inner/outer approximations of PSD matrices

$$\mathcal{SDD}_n = \mathcal{FW}^n_{\mathbf{1},2} \subseteq \mathcal{FW}^n_{\alpha,2} \subseteq \mathcal{FW}^n_{\beta,2} \subseteq \mathcal{FW}^n_{\gamma,2} \equiv \mathbb{S}^n_+$$

- Applications to inner/outer approximations of SDPs and SOS optimization

Future work

- Develop special algorithms (such as non-symmetric IPM, non-smooth optimization algorithms) for problems with factor-width-two matrices.

Thank you for your attention!

Q & A

- Yang Zheng, Aivar Sootla, and Antonis Papachristodoulou. "Block factor-width-two matrices and their applications to semidefinite and sum-of-squares optimization." IEEE Transactions on Automatic Control (2022).
- Feng-Yi Liao, and Yang Zheng. "Iterative Inner/outer Approximations for Scalable Semidefinite Programs using Block Factor-width-two Matrices." arXiv preprint arXiv:2204.06759 (2022).