Sparse PSD approximation of the PSD cone

Grigoriy Blekherman[1] Santanu S. Dey[1] Marco Molinaro [2]
Kevin Shu[1] Shengding Sun[1]

[1]Georgia Institute of Technology.

[2]Pontifical Catholic University of Rio de Janeiro.

Feb 2021

1

Introduction

Semi definite programming

$$\boxed{\begin{array}{lll} \min & \langle C, X\rangle & \\ \text{s.t.} & \langle A^i, X\rangle \leq b_i \quad \forall i \in \{1, \dots, m\} & \qquad \text{(SDP)} \\ & X \in \mathcal{S}^n_+, & \end{array}}$$

where C and the A^i's are $n \times n$ matrices, $\langle M, N\rangle := \sum_{i,j} M_{ij}N_{ij}$, and

Semi definite programming

$$\boxed{\begin{array}{lll} \min & \langle C, X\rangle & \\ \text{s.t.} & \langle A^i, X\rangle \leq b_i \quad \forall i \in \{1, \dots, m\} & \qquad \text{(SDP)} \\ & X \in \mathcal{S}^n_+, & \end{array}}$$

where C and the A^i's are $n \times n$ matrices, $\langle M, N\rangle := \sum_{i,j} M_{ij}N_{ij}$, and

$$\mathcal{S}^n_+ = \{X \in \mathbb{R}^{n \times n} \,|\, X = X^T,\ u^\top X u \geq 0,\ \forall u \in \mathbb{R}^n\}.$$

Semi definite programming

$$\boxed{\begin{array}{lll} \min & \langle C, X\rangle & \\ \text{s.t.} & \langle A^i, X\rangle \leq b_i \quad \forall i \in \{1,\dots,m\} & \qquad \text{(SDP)} \\ & X \in \mathcal{S}^n_+, & \end{array}}$$

where C and the A^i's are $n \times n$ matrices, $\langle M, N\rangle := \sum_{i,j} M_{ij}N_{ij}$, and

$$\mathcal{S}^n_+ = \{X \in \mathbb{R}^{n\times n} \,|\, X = X^T,\ u^\top X u \geq 0,\ \forall u \in \mathbb{R}^n\}.$$

- Polynomial-time algorithm— but often challenging to solve in practice.

A relaxation: Sparse SDP

$$\begin{aligned} \min \quad & \langle C, X \rangle \\ \text{s.t.} \quad & \langle A^i, X \rangle \leq b_i \quad \forall i \in \{1, \ldots, m\} \\ & X \in \mathcal{S}^n_+, \end{aligned} \tag{SDP}$$

A relaxation: Sparse SDP

$$\begin{array}{ll} \min & \langle C, X\rangle \\ \text{s.t.} & \langle A^i, X\rangle \leq b_i \quad \forall i \in \{1, \ldots, m\} \\ & X \in \mathcal{S}^n_+, \end{array} \qquad \text{(SDP)}$$

$$\begin{array}{ll} \min & \langle C, X\rangle \\ \text{s.t.} & \langle A^i, X\rangle \leq b_i \; \forall i \in \{1, \ldots, m\} \\ & \text{selected } k \times k \text{ principal submatrices of } X \in \mathcal{S}^k_+. \end{array} \qquad \text{(Sparse SDP)}$$

A relaxation: Sparse SDP

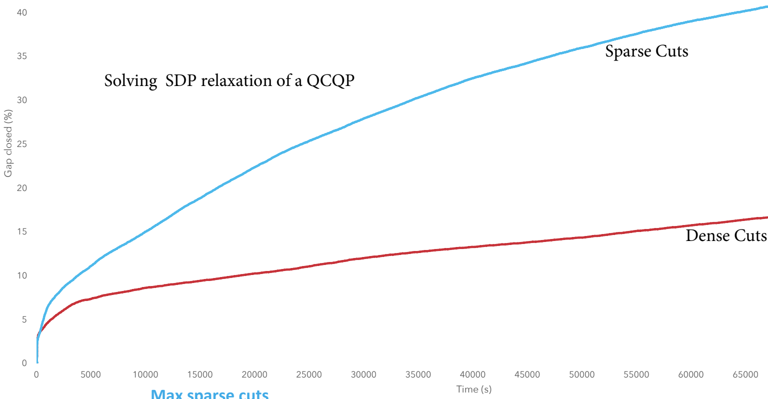
$$\begin{array}{ll} \min & \langle C, X \rangle \\ \text{s.t.} & \langle A^i, X \rangle \leq b_i \quad \forall i \in \{1, \ldots, m\} \\ & X \in \mathcal{S}^n_+, \end{array} \qquad \text{(SDP)}$$

$$\begin{array}{ll} \min & \langle C, X \rangle \\ \text{s.t.} & \langle A^i, X \rangle \leq b_i \ \forall i \in \{1, \ldots, m\} \\ & \text{selected } k \times k \text{ principal submatrices of } X \in \mathcal{S}^k_+. \end{array} \qquad \text{(Sparse SDP)}$$

Sparse cutting-plane viewpoint:

- We can enforce PSD constraints by iteratively separating linear constraints.
- Enforcing PSD-ness on smaller $k \times k$ principal submatrix leads to linear constraints that are sparser, an important property leveraged by linear programming solvers that greatly improve their efficiency.

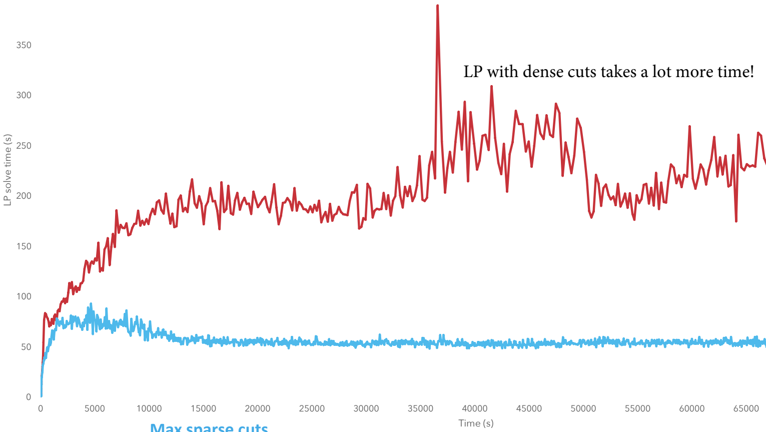
Example from [A. Kazachkov, A. Lodi, G. Munoz, SSD (2020)]



Sparsity level $k = 0.25(n+1)$

Max sparse cuts per iteration $K = 5n$

Example from [A. Kazachkov, A. Lodi, G. Munoz, SSD (2020)]



Sparsity level

$k=0.25(n+1)$

Max sparse cuts per iteration

$K=5n$

Sparse SDP

$$
\begin{array}{ll}
\min & \langle C, X\rangle \\
\text{s.t.} & \langle A^i, X\rangle \leq b_i \ \forall i \in \{1,\dots,m\} \qquad \text{(Sparse SDP)} \\
& \text{selected } k \times k \text{ principal submatrices of } X \in \mathcal{S}^k_+.
\end{array}
$$

- Sparse cutting-plane:
  - [A. Qualizza, P. Belotti, and F. Margot (2012)]
  - [R. Baltean-Lugojan, P. Bonami, R. Misener, and A. Tramontani (2018)]
  - [A. Kazachkov, A. Lodi, G. Munoz, SSD (2020)]

In many experiments, we discovered sparse SDP to give bounds quite close to that of the original SDP!

Sparse SDP

$$\begin{array}{lll} \min & \langle C, X\rangle & \\ \text{s.t.} & \langle A^i, X\rangle \leq b_i \ \forall i \in \{1, \ldots, m\} & \text{(Sparse SDP)} \\ & \text{selected } k \times k \text{ principal submatrices of } X \in \mathcal{S}_+^k. & \end{array}$$

- Sparse cutting-plane:
  - [A. Qualizza, P. Belotti, and F. Margot (2012)]
  - [R. Baltean-Lugojan, P. Bonami, R. Misener, and A. Tramontani (2018)]
  - [A. Kazachkov, A. Lodi, G. Munoz, SSD (2020)]

  In many experiments, we discovered sparse SDP to give bounds quite close to that of the original SDP!
- Power system optimization:
  - [S. Sojoudi and J. Lavaei (2014)]
  - [B. Kocuk, SSD, and X. A. Sun (2016)]

Our question

$$\begin{array}{llll} z^{\text{SDP}} := & \min & \langle C, X\rangle & \\ & \text{s.t.} & \langle A^i, X\rangle \leq b_i \quad \forall i \in \{1,\dots,m\} & \text{(SDP)} \\ & & X \in \mathcal{S}^n_+, & \end{array}$$

$$\begin{array}{llll} z^{\text{Sparse-SDP}} := & \min & \langle C, X\rangle & \\ & \text{s.t.} & \langle A^i, X\rangle \leq b_i \; \forall i \in \{1,\dots,m\} & \text{(Sparse SDP)} \\ & & \text{selected } k \times k \text{ principal submatrices of } X \in \mathcal{S}^k_+. & \end{array}$$

Our question

$$z^{\text{SDP}} := \begin{array}{ll} \min & \langle C, X\rangle \\ \text{s.t.} & \langle A^i, X\rangle \leq b_i \quad \forall i \in \{1,\dots,m\} \\ & X \in \mathcal{S}^n_+, \end{array} \qquad \text{(SDP)}$$

$$z^{\text{Sparse-SDP}} := \begin{array}{ll} \min & \langle C, X\rangle \\ \text{s.t.} & \langle A^i, X\rangle \leq b_i \;\forall i \in \{1,\dots,m\} \\ & \text{selected } k \times k \text{ principal submatrices of } X \in \mathcal{S}^k_+. \end{array} \qquad \text{(Sparse SDP)}$$

Relationship between z^{SDP} and $z^{\text{Sparse-SDP}}$?

Our question

$$z^{\text{SDP}} := \begin{array}{ll} \min & \langle C, X\rangle \\ \text{s.t.} & \langle A^i, X\rangle \le b_i \quad \forall i \in \{1,\dots,m\} \\ & X \in \mathcal{S}^n_+, \end{array} \qquad \text{(SDP)}$$

$$z^{\text{Sparse-SDP}} := \begin{array}{ll} \min & \langle C, X\rangle \\ \text{s.t.} & \langle A^i, X\rangle \le b_i \; \forall i \in \{1,\dots,m\} \\ & \text{selected } k \times k \text{ principal submatrices of } X \in \mathcal{S}^k_+. \end{array} \qquad \text{(Sparse SDP)}$$

Relationship between z^{SDP} and $z^{\text{Sparse-SDP}}$?

- Seems like a difficult question to analyze.

Easier question

$$\begin{array}{lll} \cancel{z^{\text{SDP}} :=} & \cancel{\min} & \cancel{\langle C, X\rangle} \\ & \cancel{\text{s.t.}} & \cancel{\langle A^i, X\rangle \leq b_i \quad \forall i \in \{1, \dots, m\}} \\ & & X \in \mathcal{S}^n_+, \end{array} \qquad \text{(SDP)}$$

$$\begin{array}{lll} \cancel{z^{\text{Sparse-SDP}} :=} & \cancel{\min} & \cancel{\langle C, X\rangle} \\ & \cancel{\text{s.t.}} & \cancel{\langle A^i, X\rangle \leq b_i \; \forall i \in \{1, \dots, m\} \qquad \text{(Sparse SDP)}} \\ & & \cancel{\text{selected}}\ k \times k \text{ principal submatrices of } X \in \mathcal{S}^k_+. \end{array}$$

Easier question

$$\cancel{z^{\mathrm{SDP}} :=} \quad \cancel{\min} \quad \cancel{\langle C, X\rangle}$$
$$\cancel{\text{s.t.}} \quad \cancel{\langle A^i, X\rangle \leq b_i \quad \forall i \in \{1, \dots, m\}} \qquad \text{(SDP)}$$
$$X \in \mathcal{S}^n_+,$$

$$\cancel{z^{\text{Sparse-SDP}} :=} \quad \cancel{\min} \quad \cancel{\langle C, X\rangle}$$
$$\cancel{\text{s.t.}} \quad \cancel{\langle A^i, X\rangle \leq b_i \; \forall i \in \{1, \dots, m\} \qquad \text{(Sparse SDP)}}$$
~~selected~~ $k \times k$ principal submatrices of $X \in \mathcal{S}^k_+$.

How much bigger is cone with all $k \times k$ submatrices PSD from $\mathcal{S}^n_+$?

Easier question

$$\begin{array}{lll} \cancel{z^{\text{SDP}} :=} & \cancel{\min} & \cancel{\langle C, X\rangle} \\ \cancel{\text{s.t.}} & \cancel{\langle A^i, X\rangle \leq b_i \quad \forall i \in \{1,\dots,m\}} & \text{(SDP)} \\ & X \in \mathcal{S}^n_+, & \end{array}$$

$$\begin{array}{lll} \cancel{z^{\text{Sparse-SDP}} :=} & \cancel{\min} & \cancel{\langle C, X\rangle} \\ \cancel{\text{s.t.}} & \cancel{\langle A^i, X\rangle \leq b_i \; \forall i \in \{1,\dots,m\} \quad \text{(Sparse SDP)}} & \\ & \cancel{\text{selected}}\ k \times k \text{ principal submatrices of } X \in \mathcal{S}^k_+. & \end{array}$$

$\boxed{\text{How much bigger is cone with all } k \times k \text{ submatrices PSD from } \mathcal{S}^n_+\text{?}}$

Dual cone is also of interest:

- [E. G. Boman, D. Chen, O. Parekh, and S. Toledo (2005)]
- [Permenter, Parrilo (2017)]
- [J. Gouveia, A. Kovačec, and M. Saee (2019)]
- [A. A. Ahmadi and A. Majumdar (2019)])

Setting-up details of precise question

[k-PSD closure]

Given positive integers n and k where $2 \leq k \leq n$, the k-PSD closure ($\mathcal{S}^{n,k}$) is the set of all $n \times n$ symmetric real matrices where all $k \times k$ principal submatrices are PSD.

Setting-up details of precise question

[k-PSD closure]

Given positive integers n and k where $2 \leq k \leq n$, the k-PSD closure ($\mathcal{S}^{n,k}$) is the set of all $n \times n$ symmetric real matrices where all $k \times k$ principal submatrices are PSD.

- $\boxed{\text{How far is matrices in } \mathcal{S}^{n,k} \text{ from } \mathcal{S}^n_+\text{?}}$

Setting-up details of precise question

[k-PSD closure]

Given positive integers n and k where $2 \leq k \leq n$, the k-PSD closure ($\mathcal{S}^{n,k}$) is the set of all $n \times n$ symmetric real matrices where all $k \times k$ principal submatrices are PSD.

- $\boxed{\text{How far is matrices in } \mathcal{S}^{n,k} \text{ from } \mathcal{S}^n_+\text{?}}$
- To measure this, we would like to consider the matrix in the k-PSD closure that is farthest from the PSD cone. We require to make two decisions here:

Setting-up details of precise question

[k-PSD closure]

Given positive integers n and k where $2 \leq k \leq n$, the k-PSD closure ($\mathcal{S}^{n,k}$) is the set of all $n \times n$ symmetric real matrices where all $k \times k$ principal submatrices are PSD.

- $\boxed{\text{How far is matrices in } \mathcal{S}^{n,k} \text{ from } \mathcal{S}^n_+\text{?}}$
- To measure this, we would like to consider the matrix in the k-PSD closure that is farthest from the PSD cone. We require to make two decisions here:
  1. The norm to measure this distance and
  2. A normalization method

Setting-up details of precise question

[k-PSD closure]

Given positive integers n and k where $2 \leq k \leq n$, the k-PSD closure ($\mathcal{S}^{n,k}$) is the set of all $n \times n$ symmetric real matrices where all $k \times k$ principal submatrices are PSD.

- $\boxed{\text{How far is matrices in } \mathcal{S}^{n,k} \text{ from } \mathcal{S}^n_+\text{?}}$
- To measure this, we would like to consider the matrix in the k-PSD closure that is farthest from the PSD cone. We require to make two decisions here:
  1. The norm to measure this distance and
  2. A normalization method

$$\overline{\text{dist}}_F(\mathcal{S}^{n,k}, \mathcal{S}^n_+) = \sup_{M \in \mathcal{S}^{n,k},\ \|M\|_F = 1} \text{dist}_F(M, \mathcal{S}^n_+)$$

$$= \sup_{M \in \mathcal{S}^{n,k},\ \|M\|_F = 1} \inf_{N \in \mathcal{S}^n_+} \|M - N\|_F.$$

Setting-up details of precise question

[k-PSD closure]

Given positive integers n and k where $2 \leq k \leq n$, the k-PSD closure ($\mathcal{S}^{n,k}$) is the set of all $n \times n$ symmetric real matrices where all $k \times k$ principal submatrices are PSD.

- $\boxed{\text{How far is matrices in } \mathcal{S}^{n,k} \text{ from } \mathcal{S}^n_+\text{?}}$
- To measure this, we would like to consider the matrix in the k-PSD closure that is farthest from the PSD cone. We require to make two decisions here:
  1. The norm to measure this distance and
  2. A normalization method

$$\overline{\text{dist}}_F(\mathcal{S}^{n,k}, \mathcal{S}^n_+) = \sup_{M \in \mathcal{S}^{n,k},\ \|M\|_F = 1} \text{dist}_F(M, \mathcal{S}^n_+)$$

$$= \sup_{M \in \mathcal{S}^{n,k},\ \|M\|_F = 1} \inf_{N \in \mathcal{S}^n_+} \|M - N\|_F.$$

Note: $\overline{\text{dist}}_F(\mathcal{S}^{n,k}, \mathcal{S}^n_+) \in [0, 1]$.

2

Main results 1

2.1
Upper bounds on $\overline{\mathrm{dist}}_F(\mathcal{S}^{n,k}, \mathcal{S}^n_+)$

Upper bound 1

Theorem (Upper Bound 1; Blekherman, D., Molinaro, Sun)

For all $2 \leq k < n$ we have

$$\overline{\text{dist}}_F(\mathcal{S}^{n,k}, \mathcal{S}^n_+) \leq \frac{n-k}{n+k-2}. \quad (1)$$

Upper bound 1

Theorem (Upper Bound 1; Blekherman, D., Molinaro, Sun)

For all $2 \leq k < n$ *we have*

$$\overline{\text{dist}}_F(\mathcal{S}^{n,k}, \mathcal{S}^n_+) \leq \frac{n-k}{n+k-2}. \quad (1)$$

- Distance between the k-PSD closure and the SDP cone is at most roughly $\approx \frac{n-k}{n}$.

Upper bound 2

- Distance between the k-PSD closure and the SDP cone is at most roughly $\approx \frac{n-k}{n}$

Upper bound 2

- Distance between the k-PSD closure and the SDP cone is at most roughly $\approx \frac{n-k}{n}$
- This appears to be weak especially when $k \approx n$

Upper bound 2

- Distance between the k-PSD closure and the SDP cone is at most roughly $\approx \frac{n-k}{n}$
- This appears to be weak especially when $k \approx n$

Theorem (Upper bound 2; Blekherman, D., Shu, Sun)

For all $2 \leq k < n$ we have

$$\overline{\text{dist}}_F(\mathcal{S}^{n,k}, \mathcal{S}^n_+) \leq \frac{(n-k)^{3/2}}{\sqrt{(n-k)^2 + (n-1)k^2}}. \quad (2)$$

Upper bound 2

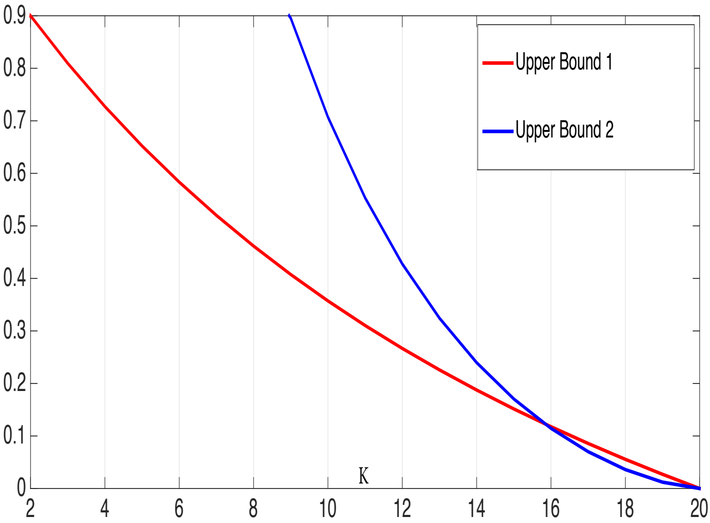
- Distance between the k-PSD closure and the SDP cone is at most roughly $\approx \frac{n-k}{n}$
- This appears to be weak especially when $k \approx n$

Theorem (Upper bound 2; Blekherman, D., Shu, Sun)
For all $2 \leq k < n$ we have

$$\overline{\text{dist}}_F(\mathcal{S}^{n,k}, \mathcal{S}^n_+) \leq \frac{(n-k)^{3/2}}{\sqrt{(n-k)^2 + (n-1)k^2}}. \quad (2)$$

- When $k \approx n$ distance between the k-PSD closure and the SDP cone is at most roughly $\approx \left(\frac{n-k}{n}\right)^{3/2}$.
- This bound dominates the previous bound when $\frac{k}{n}$ is sufficiently large.

Upper bounds on $\mathcal{S}^{20,k}$



2.2
Lower bounds on $\overline{\text{dist}}_F(\mathcal{S}^{n,k}, \mathcal{S}^n_+)$

Lower bound 1

Theorem (Lower bound 1; Blekherman, D., Molinaro, Sun)

For all $2 \leq k < n$, we have

$$\overline{\mathrm{dist}}_F(\mathcal{S}^{n,k}, \mathcal{S}^n_+) \geq \frac{n-k}{\sqrt{(k-1)^2\, n + n(n-1)}}. \tag{3}$$

Lower bound 1

Theorem (Lower bound 1; Blekherman, D., Molinaro, Sun)

For all $2 \leq k < n$, we have

$$\overline{\text{dist}}_F(\mathcal{S}^{n,k}, \mathcal{S}^n_+) \geq \frac{n-k}{\sqrt{(k-1)^2 n + n(n-1)}}. \tag{3}$$

- When k is small:

$$\frac{n-k}{\sqrt{(k-1)^2 n + n(n-1)}} \approx \frac{n-k}{n}$$

So first upper bound (Thm 1) is tight (upto constant).

Lower bound 1

Theorem (Lower bound 1; Blekherman, D., Molinaro, Sun)

For all $2 \leq k < n$, we have

$$\overline{\text{dist}}_F(\mathcal{S}^{n,k}, \mathcal{S}^n_+) \geq \frac{n-k}{\sqrt{(k-1)^2\, n + n(n-1)}}. \tag{3}$$

- When k is small:

$$\frac{n-k}{\sqrt{(k-1)^2\, n + n(n-1)}} \approx \frac{n-k}{n}$$

  So first upper bound (Thm 1) is tight (upto constant).
- When k is very large: $n - k = c$ where c is very small

Lower bound 1

Theorem (Lower bound 1; Blekherman, D., Molinaro, Sun)

For all $2 \leq k < n$, we have

$$\overline{\text{dist}}_F(\mathcal{S}^{n,k}, \mathcal{S}^n_+) \geq \frac{n-k}{\sqrt{(k-1)^2\, n + n(n-1)}}. \quad (3)$$

- When k is small:

$$\frac{n-k}{\sqrt{(k-1)^2\, n + n(n-1)}} \approx \frac{n-k}{n}$$

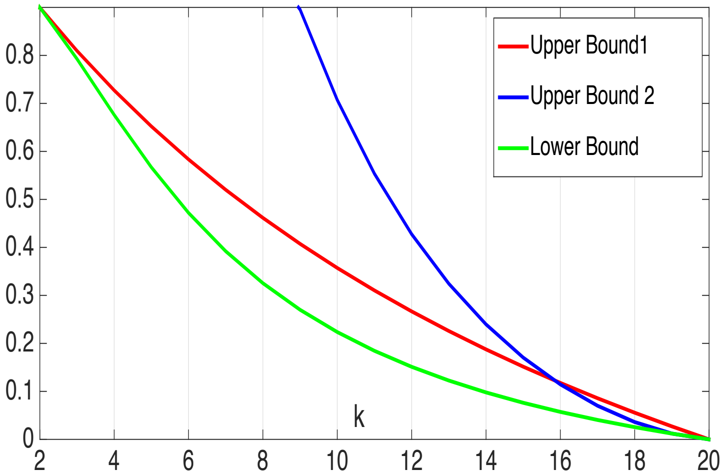
So first upper bound (Thm 1) is tight (upto constant).

- When k is very large: $n - k = c$ where c is very small

$$\frac{n-k}{\sqrt{(k-1)^2\, n + n(n-1)}} \approx \frac{c}{n^{3/2}}$$

So second upper bound (Thm 2) is tight (upto constant).

Upper and lower bounds on $\mathcal{S}^{20,k}$



Lower bound 2: What happens when $k = rn$?

- Upper bound: $\frac{n-k}{n} = 1 - r$, a constant independent of n
- Lower bound 1: $\approx (1/r - 1)\frac{1}{n^{1/2}}$.

So is upper bound weak in this regime?

Lower bound 2: What happens when $k = rn$?

- Upper bound: $\frac{n-k}{n} = 1 - r$, a constant independent of n
- Lower bound 1: $\approx (1/r - 1)\frac{1}{n^{1/2}}$.

So is upper bound weak in this regime?

Theorem (Lower bound 2; Blekherman, D., Molinaro, Sun)

Fix a constant $r < \frac{1}{93}$ and $k = rn$. Then for all $k \geq 2$,

$$\overline{\mathrm{dist}}_F(\mathcal{S}^{n,k}, \mathcal{S}^n_+) > \frac{\sqrt{r - 93r^2}}{\sqrt{162r + 3}},$$

which is independent of n.

2.3
Do we need $\binom{n}{k}$ PSD constraints?

Achieving the strength of $\mathcal{S}^{n,k}$ by a polynomial number of PSD constraints

Theorem (Blekherman, D., Molinaro, Sun)

Let $2 \leq k \leq n-1$*. Consider* $\varepsilon, \delta > 0$ *and let*

$$m = 24\left(\frac{n^2}{\varepsilon^2}\ln\frac{n}{\delta}\right).$$

Let $\mathcal{I} = (I_1, \dots, I_m)$ *be a sequence of random* k*-sets independently uniformly sampled from* $\binom{[n]}{k}$*,*

Achieving the strength of $\mathcal{S}^{n,k}$ by a polynomial number of PSD constraints

Theorem (Blekherman, D., Molinaro, Sun)

Let $2 \leq k \leq n-1$. Consider $\varepsilon, \delta > 0$ and let

$$m = 24\left(\frac{n^2}{\varepsilon^2}\ln\frac{n}{\delta}\right).$$

Let $\mathcal{I} = (I_1, \ldots, I_m)$ be a sequence of random k-sets independently uniformly sampled from $\binom{[n]}{k}$, and define $\mathcal{S}_{\mathcal{I}}$ as the set of matrices satisfying the PSD constraints for the principal submatrices indexed by the I_i's, namely

$$\mathcal{S}_{\mathcal{I}} := \{M \in \mathbb{R}^{n \times n} : M_{I_i} \succeq 0, \ \forall i \in [m]\}.$$

Achieving the strength of $\mathcal{S}^{n,k}$ by a polynomial number of PSD constraints

Theorem (Blekherman, D., Molinaro, Sun)

Let $2 \leq k \leq n-1$. Consider $\varepsilon, \delta > 0$ and let

$$m = 24\left(\frac{n^2}{\varepsilon^2}\ln\frac{n}{\delta}\right).$$

Let $\mathcal{I} = (I_1, \dots, I_m)$ be a sequence of random k-sets independently uniformly sampled from $\binom{[n]}{k}$, and define $\mathcal{S}_{\mathcal{I}}$ as the set of matrices satisfying the PSD constraints for the principal submatrices indexed by the I_i's, namely

$$\mathcal{S}_{\mathcal{I}} := \{M \in \mathbb{R}^{n\times n} : M_{I_i} \succeq 0, \ \forall i \in [m]\}.$$

Then with probability at least $1-\delta$ we have

$$\overline{\text{dist}}_F(\mathcal{S}_{\mathcal{I}}, \mathcal{S}^n_+) \leq (1+\varepsilon)\frac{n-k}{n+k-2}.$$

3
Proof sketch

3.1

Proof of:

Theorem (Upper Bound 1)

For all $2 \leq k < n$ we have

$$\overline{\mathrm{dist}}_F(\mathcal{S}^{n,k}, \mathcal{S}^n_+) \leq \frac{n-k}{n+k-2}.$$

Proof of Upper bound 1

- If

$$X = \begin{bmatrix} * & * & * & * & * \\ * & * & * & * & * \\ * & * & * & * & * \\ * & * & * & * & * \\ * & * & * & * & * \end{bmatrix} \in \mathcal{S}^{n,k}$$

then red-submatrix is $k \times k$ PSD matrix.

Proof of Upper bound 1

- If
$$X = \begin{bmatrix} * & * & * & * & * \\ * & * & * & * & * \\ * & * & * & * & * \\ * & * & * & * & * \\ * & * & * & * & * \end{bmatrix} \in \mathcal{S}^{n,k}$$
then red-submatrix is $k \times k$ PSD matrix.
- So
$$\begin{bmatrix} * & * & * & 0 & 0 \\ * & * & * & 0 & 0 \\ * & * & * & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} \in \mathcal{S}^n_+.$$

Proof of Upper bound 1

- If
$$X = \begin{bmatrix} * & * & * & * & * \\ * & * & * & * & * \\ * & * & * & * & * \\ * & * & * & * & * \\ * & * & * & * & * \end{bmatrix} \in \mathcal{S}^{n,k}$$
then red-submatrix is $k \times k$ PSD matrix.
- So
$$\begin{bmatrix} * & * & * & 0 & 0 \\ * & * & * & 0 & 0 \\ * & * & * & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} \in \mathcal{S}^n_+.$$
- Take average of all the above matrices for different principal $k \times k$ submatrices (and suitably scale with a positive number), then the resulting matrix is in S^n_+.
- The distance between this average PSD matrix and X gives bound.

3.2

Proof of:

Theorem (Upper bound 2)

Assume $2 \leq k < n$. *Then*

$$\overline{\mathrm{dist}}_F(\mathcal{S}^{n,k}, \mathcal{S}^n_+) \leq \frac{(n-k)^{3/2}}{\sqrt{(n-k)^2 + (n-1)k^2}}.$$

Proof of upper bound 2

- Using Cauchy's Interlace Theorem for eigenvalues of symmetric matrices, we obtain that every matrix in $\mathcal{S}^{n,k}$ has at most $n - k$ negative eigenvalues.

Proof of upper bound 2

- Using Cauchy's Interlace Theorem for eigenvalues of symmetric matrices, we obtain that every matrix in $\mathcal{S}^{n,k}$ has at most $n-k$ negative eigenvalues.
- Since the PSD cone consists of symmetric matrices with non-negative eigenvalues, the distance from a unit-norm matrix $M \in \mathcal{S}^{n,k}$ to $\mathcal{S}^n_+$ is upper bounded by

  (absolute value of most negative eigenvalue of M)$\times\sqrt{n-k}$.
- So we need to upper bound absolute value of most negative eigenvalue of M for $M \in \mathcal{S}^{n,k}$ and $\|M\|_F = 1$.
- $$\sqrt{n-k} \times \left|\min\left\{\lambda_1(M) \mid \|M\|_F \leq 1, M \in \mathcal{S}^{n,k}\right\}\right|$$

Proof of upper bound 2

- Using Cauchy's Interlace Theorem for eigenvalues of symmetric matrices, we obtain that every matrix in $\mathcal{S}^{n,k}$ has at most $n-k$ negative eigenvalues.
- Since the PSD cone consists of symmetric matrices with non-negative eigenvalues, the distance from a unit-norm matrix $M \in \mathcal{S}^{n,k}$ to $\mathcal{S}^n_+$ is upper bounded by

  (absolute value of most negative eigenvalue of M)$\times\sqrt{n-k}$.
- So we need to upper bound absolute value of most negative eigenvalue of M for $M \in \mathcal{S}^{n,k}$ and $\|M\|_F = 1$.
- $$\sqrt{n-k} \times \left| \min \left\{ \lambda_1(M) \,\middle|\, \sum_{j\in[n]} (\lambda_j(M))^2 \leq 1, M \in \mathcal{S}^{n,k} \right\} \right|$$

Proof of upper bound 2

- Using Cauchy's Interlace Theorem for eigenvalues of symmetric matrices, we obtain that every matrix in $\mathcal{S}^{n,k}$ has at most $n-k$ negative eigenvalues.
- Since the PSD cone consists of symmetric matrices with non-negative eigenvalues, the distance from a unit-norm matrix $M \in \mathcal{S}^{n,k}$ to $\mathcal{S}^n_+$ is upper bounded by

  (absolute value of most negative eigenvalue of M)$\times\sqrt{n-k}$.
- So we need to upper bound absolute value of most negative eigenvalue of M for $M \in \mathcal{S}^{n,k}$ and $\|M\|_F = 1$.
- $$\sqrt{n-k} \times \left| \min \left\{ \lambda_1(M) \;\middle|\; \sum_{j\in[n]} (\lambda_j(M))^2 \leq 1, \quad \underbrace{M \in \mathcal{S}^{n,k}}_{\text{how to deal with this?}} \right\} \right|$$

Proof of upper bound 2-connection to hyperbolicity cone.

- For $S \subseteq \{1, \dots, n\}$, let $M|_S$ denote the principal submatrix of M obtained by removing rows and columns not in S.
- If $|S| = k$, and $M \in S^{n,k}$, then $M|_S$ is PSD.

Proof of upper bound 2-connection to hyperbolicity cone.

- For $S \subseteq \{1, \ldots, n\}$, let $M|_S$ denote the principal submatrix of M obtained by removing rows and columns not in S.
- If $|S| = k$, and $M \in S^{n,k}$, then $M|_S$ is PSD.
- $$M \in \mathcal{S}^{n,k} \Rightarrow c_k(M) := \sum_{S \subseteq \{1,\ldots,n\} : |S| = k} \det(M|_S) \geq 0.$$

Proof of upper bound 2-connection to hyperbolicity cone.

- For $S \subseteq \{1,\dots,n\}$, let $M|_S$ denote the principal submatrix of M obtained by removing rows and columns not in S.
- If $|S| = k$, and $M \in S^{n,k}$, then $M|_S$ is PSD.
- $$M \in \mathcal{S}^{n,k} \;\Rightarrow\; c_k(M) := \sum_{S\subseteq\{1,\dots,n\}:|S|=k} \det(M|_S) \geq 0.$$
- Let $\lambda_1(M) \leq \lambda_2(M) \leq \cdots \leq \lambda_n(M)$ are the eigenvalues of M:

$$c_k(M) = \underbrace{\sum_{1\leq i_1<i_2<\cdots<i_k\leq n} \lambda_{i_1}(M)\lambda_{i_2}(M)\dots\lambda_{i_k}(M)}_{e_k^n(\lambda_1(M),\lambda_2(M),\dots,\lambda_n(M))}$$

- $$M \in \mathcal{S}^{n,k} \;\Rightarrow\; \underbrace{e_k^n(\lambda(M))}_{\text{elementary symmetric polynomial}} \geq 0.$$

Proof of upper bound 2-connection to hyperbolicity cone.

- For $S \subseteq \{1, \dots, n\}$, let $M|_S$ denote the principal submatrix of M obtained by removing rows and columns not in S.
- If $|S| = k$, and $M \in S^{n,k}$, then $M|_S$ is PSD.
- $$M \in \mathcal{S}^{n,k} \;\Rightarrow\; c_k(M) := \sum_{S \subseteq \{1,\dots,n\} : |S| = k} \det(M|_S) \geq 0.$$
- Let $\lambda_1(M) \leq \lambda_2(M) \leq \cdots \leq \lambda_n(M)$ are the eigenvalues of M:

$$c_k(M) = \underbrace{\sum_{1 \leq i_1 < i_2 < \cdots < i_k \leq n} \lambda_{i_1}(M)\lambda_{i_2}(M)\dots\lambda_{i_k}(M)}_{e_k^n(\lambda_1(M), \lambda_2(M), \dots, \lambda_n(M))}$$

- $$M \in \mathcal{S}^{n,k} \;\Rightarrow\; \underbrace{e_k^n(\lambda(M))}_{\text{elementary symmetric polynomial}} \geq 0.$$
- We can do better...

Proof of upper bound 2- connection to hyperbolicity cone.

- Let $M \in \mathcal{S}^{n,k}$. For $t > 0$:

$$e_k^n((\lambda_1(M), \lambda_2(M), \dots, \lambda_n(M)) + t\vec{1}) = c_k(M + tI) > 0,$$

since all the $k \times k$ submatrices of $X + tI$ will be positive definite.

Proof of upper bound 2- connection to hyperbolicity cone.

- Let $M \in \mathcal{S}^{n,k}$. For $t > 0$:
$$e_k^n((\lambda_1(M), \lambda_2(M), \dots, \lambda_n(M)) + t\vec{1}) = c_k(M + tI) > 0,$$
since all the $k \times k$ submatrices of $X + tI$ will be positive definite.
- Every point in the open line segment: $\{\theta\lambda(M) + (1-\theta)\vec{1} \,|\, 1 > \theta \geq 0\}$ belongs to $\underbrace{\text{connected component of } \mathbb{R}^n \setminus \{x : e_k^n(x) = 0\} \text{ containing } \vec{1}}_{H(e_k^n) \text{ hyperbolicity cone of elementary symmetric polynomial}}$.
- $$M \in \mathcal{S}^{n,k} \ \Rightarrow\ \lambda(M) \in H\left(e_k^n\right).$$

Proof of upper bound 2- connection to hyperbolicity cone.

- Let $M \in \mathcal{S}^{n,k}$. For $t > 0$:

$$e_k^n((\lambda_1(M), \lambda_2(M), \dots, \lambda_n(M)) + t\vec{1}) = c_k(M + tI) > 0,$$

  since all the $k \times k$ submatrices of $X + tI$ will be positive definite.
- Every point in the open line segment:
  $\{\theta\lambda(M) + (1-\theta)\vec{1} \,|\, 1 > \theta \geq 0\}$ belongs to $\underbrace{\text{connected component of } \mathbb{R}^n \setminus \{x : e_k^n(x) = 0\} \text{ containing } \vec{1}}_{H(e_k^n) \text{ hyperbolicity cone of elementary symmetric polynomial}}$.
-

$$M \in \mathcal{S}^{n,k} \;\Rightarrow\; \lambda(M) \in H\left(e_k^n\right).$$

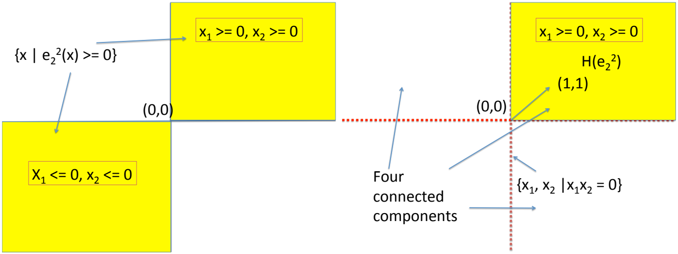
  Two nice properties:
  - $\{x \,|\, e_k^n(x) \geq 0\} \supsetneq H\left(e_k^n\right)$.
  - $H\left(e_k^n\right)$ is a convex set.

Illustration of $\{x \mid e_2^2(x) \geq 0\}$ and $H(e_2^2)$

$$e_2^2(x) = x_1 x_2$$

- $\{x \mid e_2^2(x) \geq 0\} \equiv \{x \mid x_1 x_2 \geq 0\}$.
- H_2^2 := connected component of $\mathbb{R}^2 \setminus \{x \mid x_1 x_2 = 0\}$ that contains $(1, 1)$.

Figure: $\{x \mid e_2^2(x) \geq 0\}$ Figure: $H(e_2^2)$



A quick detour to formally introduce hyperbolicity cone

- We will say that a polynomial $p \in \mathbb{R}[x_1, \dots, x_n]$ is hyperbolic with respect to a fixed vector v if
  - $p(v) > 0$, and
  - For all fixed $\hat{x} \in \mathbb{R}^n$, the univariate polynomial $p(\hat{x} - tv) \in R[t]$ has only real roots.

[1] We actually work with the closure of this set

A quick detour to formally introduce hyperbolicity cone

- We will say that a polynomial $p \in \mathbb{R}[x_1, \ldots, x_n]$ is hyperbolic with respect to a fixed vector v if
  - $p(v) > 0$, and
  - For all fixed $\hat{x} \in \mathbb{R}^n$, the univariate polynomial $p(\hat{x} - tv) \in R[t]$ has only real roots.

Example:

- $e_n^n(\hat{x} - t\vec{1}) = 0$,

[1] We actually work with the closure of this set

A quick detour to formally introduce hyperbolicity cone

- We will say that a polynomial $p \in \mathbb{R}[x_1, \ldots, x_n]$ is hyperbolic with respect to a fixed vector v if
  - $p(v) > 0$, and
  - For all fixed $\hat{x} \in \mathbb{R}^n$, the univariate polynomial $p(\hat{x} - tv) \in R[t]$ has only real roots.

Example:

- $e_n^n(\hat{x} - t\vec{1}) = 0$,
- solution for t (roots): $\hat{x}_1, \hat{x}_2, \ldots, \hat{x}_n$

[1] We actually work with the closure of this set

A quick detour to formally introduce hyperbolicity cone

- We will say that a polynomial $p \in \mathbb{R}[x_1, \ldots, x_n]$ is hyperbolic with respect to a fixed vector v if
  - $p(v) > 0$, and
  - For all fixed $\hat{x} \in \mathbb{R}^n$, the univariate polynomial $p(\hat{x} - tv) \in R[t]$ has only real roots.

Example:

- $e_n^n(\hat{x} - t\vec{1}) = 0$,
- solution for t (roots): $\hat{x}_1, \hat{x}_2, \ldots, \hat{x}_n$

- The connected set $\mathbb{R}^n \setminus \{x \mid p(x) = 0\}$ containing v is called the hyperbolicity cone of p with respect to v[1].
- The hyperbolicity cone is a convex cone!

[1] We actually work with the closure of this set

Proof of upper bound 2-contd.

- Replace:

$$\sqrt{n-k} \times \min \left\{ \lambda_1(M) \,\middle|\, \sum_{j \in [n]} (\lambda_j(M))^2 \leq 1, M \in \mathcal{S}^{n,k} \right\}$$

Proof of upper bound 2-contd.

- Replace:

$$\sqrt{n-k} \times \min \left\{ \lambda_1(M) \;\middle|\; \sum_{j \in [n]} (\lambda_j(M))^2 \leq 1, M \in \mathcal{S}^{n,k} \right\}$$

- By its relaxation:

$$\sqrt{n-k} \times \left| \min \left\{ \lambda_1(M) \;\middle|\; \sum_{j \in [n]} (\lambda_j(M))^2 \leq 1, \lambda(M) \in H(e_k^n) \right\} \right|$$

- This is a convex relaxation and can be solved in closed form. The solution is the bound we obtain.

3.3

Proof of:

Theorem (Lower bound 1)

For all $2 \leq k < n$, we have

$$\overline{\mathrm{dist}}_F(\mathcal{S}^{n,k}, \mathcal{S}^n_+) \geq \frac{n-k}{\sqrt{(k-1)^2 n + n(n-1)}}.$$

Proof of lower bound 1

- Consider the matrix:

$$G(a, b) := (a + b)I - a\mathbf{1}\mathbf{1}^\top$$

- If $u \in \mathbb{R}^n$ with $\|u\|_2 = 1$ has a support of k, then

$$u^\top Gu =$$

Proof of lower bound 1

- Consider the matrix:

$$G(a,b) := (a+b)I - a\mathbf{1}\mathbf{1}^\top$$

- If $u \in \mathbb{R}^n$ with $\|u\|_2 = 1$ has a support of k, then

$$u^\top G u = (a+b) - a\left(\sum_{i=1}^n u_i\right)^2$$

Proof of lower bound 1

- Consider the matrix:

$$G(a,b) := (a+b)I - a\mathbf{1}\mathbf{1}^\top$$

- If $u \in \mathbb{R}^n$ with $\|u\|_2 = 1$ has a support of k, then

$$u^\top Gu = (a+b) - a\left(\sum_{i=1}^{n} u_i\right)^2 \geq (a+b) - a(\|u\|_1)^2$$

Proof of lower bound 1

- Consider the matrix:

$$G(a,b) := (a+b)I - a\mathbf{1}\mathbf{1}^\top$$

- If $u \in \mathbb{R}^n$ with $\|u\|_2 = 1$ has a support of k, then

$$u^\top G u = (a+b) - a\left(\sum_{i=1}^{n} u_i\right)^2 \geq (a+b) - a(\|u\|_1)^2 \geq (a+b) - ak$$

Proof of lower bound 1

- Consider the matrix:

$$G(a,b) := (a+b)I - a\mathbf{1}\mathbf{1}^\top$$

- If $u \in \mathbb{R}^n$ with $\|u\|_2 = 1$ has a support of k, then

$$u^\top G u = (a+b) - a\left(\sum_{i=1}^{n} u_i\right)^2 \geq (a+b) - a(\|u\|_1)^2 \geq (a+b) - ak$$

- So $G(a,b) \in \mathcal{S}^{n,k}$ iff $(1-k)a + b \geq 0$.
- Use these explicit matrices to obtain lower bound from S^n_+

3.4

Proof of:

Theorem (Lower bound 2)

Fix a constant $r < \frac{1}{93}$ and $k = rn$. Then for all $k \geq 2$,

$$\overline{\text{dist}}_F(\mathcal{S}^{n,k}, \mathcal{S}^n_+) > \frac{\sqrt{r - 93r^2}}{\sqrt{162r + 3}},$$

which is independent of n.

Proof of lower bound 2

- For simplicity, assume $k = n/2$. (Actually proof does not have this value of k).

Proof of lower bound 2

- For simplicity, assume $k = n/2$. (Actually proof does not have this value of k).
- The idea is to construct a matrix M where half of its eigenvalues take the negative value $-\frac{1}{\sqrt{n}}$, with orthonormal eigenvectors $v^1, v^2, \dots, v^{n/2}$, and rest take a positive value $\frac{1}{\sqrt{n}}$, with orthonormal eigenvectors $w^1, w^2, \dots, w^{n/2}$, i.e.,

$$M = \frac{-1}{\sqrt{n}} \sum_{i=1}^{n/2} (v^i)(v^i)^\top + \frac{1}{\sqrt{n}} \sum_{i=1}^{n/2} (w^i)(w^i)^\top$$

Proof of lower bound 2

- For simplicity, assume $k = n/2$. (Actually proof does not have this value of k).
- The idea is to construct a matrix M where half of its eigenvalues take the negative value $-\frac{1}{\sqrt{n}}$, with orthonormal eigenvectors $v^1, v^2, \ldots, v^{n/2}$, and rest take a positive value $\frac{1}{\sqrt{n}}$, with orthonormal eigenvectors $w^1, w^2, \ldots, w^{n/2}$, i.e.,

$$M = \frac{-1}{\sqrt{n}} \sum_{i=1}^{n/2} (v^i)(v^i)^\top + \frac{1}{\sqrt{n}} \sum_{i=1}^{n/2} (w^i)(w^i)^\top$$

- This normalization makes $\|M\|_F = 1$.
- $\mathrm{dist}_F(M, \mathcal{S}_+^n) \geq \sqrt{\left(\frac{1}{\sqrt{n}}\right)^2 \cdot \frac{n}{2}} = \mathit{cst}$ independent of n.

Proof of lower bound 2

- For simplicity, assume $k = n/2$. (Actually proof does not have this value of k).
- The idea is to construct a matrix M where half of its eigenvalues take the negative value $-\frac{1}{\sqrt{n}}$, with orthonormal eigenvectors $v^1, v^2, \dots, v^{n/2}$, and rest take a positive value $\frac{1}{\sqrt{n}}$, with orthonormal eigenvectors $w^1, w^2, \dots, w^{n/2}$, i.e.,

$$M = \frac{-1}{\sqrt{n}} \sum_{i=1}^{n/2} (v^i)(v^i)^\top + \frac{1}{\sqrt{n}} \sum_{i=1}^{n/2} (w^i)(w^i)^\top$$

- This normalization makes $\|M\|_F = 1$.
- $\text{dist}_F(M, \mathcal{S}_+^n) \geq \sqrt{\left(\frac{1}{\sqrt{n}}\right)^2 \cdot \frac{n}{2}} = cst$ independent of n.
- So we only need to guarantee that M belongs to the k-PSD closure.

Proof of lower bound 2 –contd.

- $M = \frac{-1}{\sqrt{n}} \sum_{i=1}^{n/2} (v^i)(v^i)^\top + \frac{1}{\sqrt{n}} \sum_{i=1}^{n/2} (w^i)(w^i)^\top$
- Letting V be the matrix with rows $v^1, v^2, \ldots,$ and W the matrix with rows $w^1, w^2, \ldots$, the quadratic form $x^\top Mx$:

$$x^\top Mx = -\frac{1}{\sqrt{n}} \|Vx\|_2^2 + \frac{1}{\sqrt{n}} \|Wx\|_2^2.$$

Proof of lower bound 2 –contd.

- $M = \frac{-1}{\sqrt{n}} \sum_{i=1}^{n/2} (v^i)(v^i)^\top + \frac{1}{\sqrt{n}} \sum_{i=1}^{n/2} (w^i)(w^i)^\top$
- Letting V be the matrix with rows $v^1, v^2, \ldots,$ and W the matrix with rows $w^1, w^2, \ldots$, the quadratic form $x^\top Mx$:

$$x^\top Mx = -\frac{1}{\sqrt{n}} \|Vx\|_2^2 + \frac{1}{\sqrt{n}} \|Wx\|_2^2.$$

- $\|Vx\|_2^2 \leq \|x\|_2^2$ (because V is orthonormal)

Proof of lower bound 2 –contd.

- $M = \frac{-1}{\sqrt{n}} \sum_{i=1}^{n/2} (v^i)(v^i)^\top + \frac{1}{\sqrt{n}} \sum_{i=1}^{n/2} (w^i)(w^i)^\top$
- Letting V be the matrix with rows $v^1, v^2, \ldots,$ and W the matrix with rows $w^1, w^2, \ldots$, the quadratic form $x^\top Mx$:

$$x^\top Mx = -\frac{1}{\sqrt{n}} \|Vx\|_2^2 + \frac{1}{\sqrt{n}} \|Wx\|_2^2.$$

- $\|Vx\|_2^2 \leq \|x\|_2^2$ (because V is orthonormal)
- So *if we could construct the matrix W so that for all k-sparse vectors $x \in \mathbb{R}^n$ we had* $\boxed{\|Wx\|_2^2 \approx \|x\|_2^2}$:

$$x^\top Mx \gtrsim -\frac{1}{\sqrt{n}} \|x\|_2^2 + \frac{1}{\sqrt{n}} \|x\|_2^2 \gtrsim 0 \qquad \boxed{\text{for all } k\text{-sparse vectors } x},$$

Proof of lower bound 2 –contd.

- $M = \frac{-1}{\sqrt{n}} \sum_{i=1}^{n/2} (v^i)(v^i)^\top + \frac{1}{\sqrt{n}} \sum_{i=1}^{n/2} (w^i)(w^i)^\top$
- Letting V be the matrix with rows $v^1, v^2, \dots,$ and W the matrix with rows $w^1, w^2, \dots$, the quadratic form $x^\top Mx$:

$$x^\top Mx = -\frac{1}{\sqrt{n}} \|Vx\|_2^2 + \frac{1}{\sqrt{n}} \|Wx\|_2^2.$$

- $\|Vx\|_2^2 \leq \|x\|_2^2$ (because V is orthonormal)
- So if *we could construct the matrix W so that for all k-sparse vectors $x \in \mathbb{R}^n$ we had* $\boxed{\|Wx\|_2^2 \approx \|x\|_2^2}$:

$$x^\top Mx \gtrsim -\frac{1}{\sqrt{n}} \|x\|_2^2 + \frac{1}{\sqrt{n}} \|x\|_2^2 \gtrsim 0 \qquad \boxed{\text{for all } k\text{-sparse vectors } x},$$

- This approximate preservation of norms of sparse vectors is precisely the notion of the *Restricted Isometry Property*.

4

Examining the hyperbolicity relaxations

How good is the hyperbolicity relaxation?

- $\{\lambda(M) \mid M \in \mathcal{S}^{n,k}\} \subseteq H(e_k^n)$

How good is the hyperbolicity relaxation?

- $\{\lambda(M) \mid M \in \mathcal{S}^{n,k}\} \subseteq H(e_k^n)$
- In fact, $H(e_n^n) = \mathbb{R}_+^n = \{\lambda(M) \mid M \in \mathcal{S}^{n,n}\}$.

The case of $k = n - 1$

Theorem ($k = n - 1$; Blekherman, D., Shu, Sun)

Let $n \geq 3$. Then:

$$H(e_{n-1}^{n}) = \{\lambda(M) \mid M \in \mathcal{S}^{n,n-1}\}.$$

The case of $k = n - 1$

Theorem ($k = n - 1$; Blekherman, D., Shu, Sun)

Let $n \geq 3$. Then:

$$H(e^n_{n-1}) = \{\lambda(M) \mid M \in \mathcal{S}^{n,n-1}\}.$$

Corollary

Let $n \geq 3$. Then: $\{\lambda(M) \mid M \in \mathcal{S}^{n,n-1}\}$ is a convex set.

How good is the hyperbolicity relaxation?

$$G(1,k-1) := \begin{bmatrix} k-1 & -1 & \dots & -1 \\ -1 & k-1 & \dots & -1 \\ \cdot & \cdot & \cdot & \cdot \\ -1 & \dots & -1 & k-1 \end{bmatrix} \in \mathcal{S}^{n,k}$$

- Every $k \times k$ principal submatrix of $G(1,k)$ is singular. Thus,

$$e_k^n(G(1,k-1)) = \sum_{S \subseteq [n]:|S|=k} \det(G(1,k-1)|_S) = 0.$$

How good is the hyperbolicity relaxation?

$$G(1, k-1) := \begin{bmatrix} k-1 & -1 & \dots & -1 \\ -1 & k-1 & \dots & -1 \\ \cdot & \cdot & \cdot & \cdot \\ -1 & \dots & -1 & k-1 \end{bmatrix} \in \mathcal{S}^{n,k}$$

- Every $k \times k$ principal submatrix of $G(1, k)$ is singular. Thus,

$$e_k^n(G(1, k-1)) = \sum_{S \subseteq [n] : |S| = k} \det(G(1, k-1)|_S) = 0.$$

- So $G(1, k-1)$ belongs to the boundary of $H(e_k^n)$.

How good is the hyperbolicity relaxation?

$$G(1, k-1) := \begin{bmatrix} k-1 & -1 & \dots & -1 \\ -1 & k-1 & \dots & -1 \\ \cdot & \cdot & \cdot & \cdot \\ -1 & \dots & -1 & k-1 \end{bmatrix} \in \mathcal{S}^{n,k}$$

- Every $k \times k$ principal submatrix of $G(1, k)$ is singular. Thus,

$$e_k^n(G(1, k-1)) = \sum_{S \subseteq [n]:|S|=k} \det(G(1, k-1)|_S) = 0.$$

- So $G(1, k-1)$ belongs to the boundary of $H(e_k^n)$.
- Let D be a diagonal matrix (where every diagonal entry is non-zero). Then $DG(1, k-1)D \in \mathcal{S}^{n,k}$ and $DG(1, k-1)D \in \text{bnd}(H(e_k^n))$.

How good is the hyperbolicity relaxation?

$$G(1,k-1) := \begin{bmatrix} k-1 & -1 & \dots & -1 \\ -1 & k-1 & \dots & -1 \\ \cdot & \cdot & \cdot & \cdot \\ -1 & \dots & -1 & k-1 \end{bmatrix} \in \mathcal{S}^{n,k}$$

- Every $k \times k$ principal submatrix of $G(1,k)$ is singular. Thus,

$$e_k^n(G(1,k-1)) = \sum_{S \subseteq [n]:|S|=k} \det(G(1,k-1)|_S) = 0.$$

- So $G(1,k-1)$ belongs to the boundary of $H(e_k^n)$.
- Let D be a diagonal matrix (where every diagonal entry is non-zero). Then $DG(1,k-1)D \in \mathcal{S}^{n,k}$ and $DG(1,k-1)D \in \mathrm{bnd}(H(e_k^n))$.
- Finally, $DG(1,k-1)D$ is a non-singular matrix.

How good is the hyperbolicity relaxation?

$$G(1, k-1) := \begin{bmatrix} k-1 & -1 & \dots & -1 \\ -1 & k-1 & \dots & -1 \\ \vdots & \vdots & \vdots & \vdots \\ -1 & \dots & -1 & k-1 \end{bmatrix} \in \mathcal{S}^{n,k}$$

- Every $k \times k$ principal submatrix of $G(1, k)$ is singular. Thus,

$$e_k^n(G(1, k-1)) = \sum_{S \subseteq [n] : |S| = k} \det(G(1, k-1)|_S) = 0.$$

- So $G(1, k-1)$ belongs to the boundary of $H(e_k^n)$.
- Let D be a diagonal matrix (where every diagonal entry is non-zero). Then $DG(1, k-1)D \in \mathcal{S}^{n,k}$ and $DG(1, k-1)D \in \text{bnd}(H(e_k^n))$.
- Finally, $DG(1, k-1)D$ is a non-singular matrix.
- So, $DG(1, k-1)D \in \text{bnd}(H(e_k^n))$, is a non-singular matrix and belongs to $\mathcal{S}^{n,k}$.

How good is the hyperbolicity relaxation — Comparing boundary

Theorem (Blekherman, D., Shu, Sun)

Let $2 < k < n-1$ or $n = 4$ and $k = 2$. Let $M \in \mathcal{S}^{n,k}$. If M is non-singular and M belongs to the boundary of $H(e_k^n)$, then there exists a diagonal matrix D such that $M = DG(1, k-1)D$.

How good is the hyperbolicity relaxation — Comparing boundary

Theorem (Blekherman, D., Shu, Sun)

Let $2 < k < n-1$ or $n = 4$ and $k = 2$. Let $M \in \mathcal{S}^{n,k}$. If M is non-singular and M belongs to the boundary of $H(e_k^n)$, then there exists a diagonal matrix D such that $M = DG(1, k-1)D$.

- There exist points on the boundary of $H(e_k^n)$ with as many as $n-k$ negative entries and no zero entries.

Corollary

Let $2 < k < n-1$ or $n = 4$ and $k = 2$. Then the set of "eigenvalue vectors" for matrices in $\mathcal{S}^{n,k}$ is strictly contained in $H(e_k^n)$.

Thank You.

- Grigoriy Blekherman, Santanu S. Dey, Marco Molinaro, Shengding Sun, “Sparse PSD approximation of the PSD cone," To appear in *Mathematical Programming*.
- Grigoriy Blekherman, Santanu S. Dey, Kevin Shu, Shengding Sun, “Hyperbolic Relaxation of k-Locally Positive Semidefinite Matrices," *https://arxiv.org/abs/2012.04031.*
- Santanu S. Dey, Aleksandr M. Kazachkov, Andrea Lodi, Gonzalo Muñoz, "Cutting Plane Generation Through Sparse Principal Component Analysis" http://www.optimization-online.org/DBHTML/2021/02/8259.html