Tensor theta norms

Holger Rauhut[1] Željka Stojanac[1,2]

[1]RWTH Aachen University

[2]Hausdorff Center for Mathematics
Institut für Numerische Simulation
University of Bonn

Frames and Algebraic & Computational geometry,
30th July 2015

Overview

Compressive sensing - vector case

- $\mathbf{\Phi} \in \mathbb{R}^{m \times n}$, $\mathbf{x} \in \mathbb{R}^n$, $\mathbf{x}$ sparse, $\mathbf{b} = \mathbf{\Phi x} \in \mathbb{R}^m$, $m \ll n$
- Goal: Reconstruct $\mathbf{x}$ from $\mathbf{\Phi}$ and $\mathbf{b}$
- ℓ_0-minimization (non-convex optimization problem)

$$\min_{\mathbf{z} \in \mathbb{R}^n} \|\mathbf{z}\|_0 \text{ such that } \mathbf{\Phi z} = \mathbf{b} \qquad \text{(NP-hard)}$$

  where

$$\|\mathbf{x}\|_0 = \#\{i : |x_i| \neq 0\}.$$

- ℓ_1-minimization (convex optimization problem)

$$\min_{\mathbf{z} \in \mathbb{R}^n} \|\mathbf{z}\|_1 \text{ such that } \mathbf{\Phi z} = \mathbf{b}.$$

- Gaussian matrix: with high probability all s-sparse vectors can be reconstructed provided $m \geq Cs \ln(n/s)$.

Compressive sensing - matrix case

- $\mathbf{\Phi} : \mathbb{R}^{n_1 \times n_2} \to \mathbb{R}^m$, $\mathbf{X} \in \mathbb{R}^{n_1 \times n_2}$, $\mathbf{b} = \mathbf{\Phi}(\mathbf{X}) \in \mathbb{R}^m$, $\mathbf{X}$ is of low rank, $m \ll n_1 n_2$
- Goal: Reconstruct $\mathbf{X}$ from $\mathbf{\Phi}$ and $\mathbf{b}$
- non-convex optimization problem

$$\min_{\mathbf{Z} \in \mathbb{R}^{n_1 \times n_2}} \operatorname{rank}(\mathbf{Z}) \text{ such that } \mathbf{\Phi}(\mathbf{Z}) = \mathbf{b}. \qquad \text{(NP-hard)}$$

- nuclear norm minimization (convex optimization problem)

$$\min_{\mathbf{Z} \in \mathbb{R}^{n_1 \times n_2}} \|\mathbf{Z}\|_* \text{ such that } \mathbf{\Phi}(\mathbf{Z}) = \mathbf{b},$$

where $\|\mathbf{Z}\|_* = \sum_{i=1}^{r} \sigma_i = \|\sigma(\mathbf{Z})\|_1$, σ_i are singular values of $\mathbf{Z}$.

- Gaussian map: with high probability all r-rank matrices can be reconstructed provided $m \geq C_M r \max\{n_1, n_2\}$.

Compressive sensing - tensor case

- $\mathbf{\Phi} : \mathbb{R}^{n_1 \times n_2 \times \cdots \times n_d} \to \mathbb{R}^m$, $\mathbf{X} \in \mathbb{R}^{n_1 \times n_2 \times \cdots \times n_d}$, $\mathbf{b} = \mathbf{\Phi}(\mathbf{X}) \in \mathbb{R}^m$, $m \ll n_1 n_2 \cdots n_d$, $\mathbf{X}$ of low rank
- Goal: Reconstruct $\mathbf{X}$ from $\mathbf{\Phi}$ and $\mathbf{b}$
- non-convex optimization problem

$$\min_{\mathbf{Z} \in \mathbb{R}^{n_1 \times n_2 \times \cdots \times n_d}} \operatorname{rank}(\mathbf{Z}) \text{ such that } \mathbf{\Phi}(\mathbf{Z}) = \mathbf{b}. \qquad \text{(NP-hard)}$$

- convex optimization problem

$$\min_{\mathbf{Z} \in \mathbb{R}^{n_1 \times n_2 \times \cdots \times n_d}} \|\mathbf{Z}\| \text{ such that } \mathbf{\Phi}(\mathbf{Z}) = \mathbf{b}.$$

- What is a tensor rank?
- What norm should we use?
- The number of measurements for recovery?

- Tensor of order d: $\mathbf{X} \in \mathbb{R}^{n_1 \times n_2 \times \cdots \times n_d}$.

- Tensor of order d: $\mathbf{X} \in \mathbb{R}^{n_1 \times n_2 \times \cdots \times n_d}$.
- $\mathbf{X} \in \mathbb{R}^{n_1 \times n_2 \times \cdots \times n_d}$ is a rank one tensor:
  $\exists\, \mathbf{u}_1 \in \mathbb{R}^{n_1}, \mathbf{u}_2 \in \mathbb{R}^{n_2}, \ldots, \mathbf{u}_d \in \mathbb{R}^{n_d}$ s.t.

$$\mathbf{X} = \mathbf{u}_1 \otimes \mathbf{u}_2 \otimes \cdots \otimes \mathbf{u}_d, \text{ i.e., } \mathbf{X}(i_1, i_2, \ldots, i_d) = \mathbf{u}_1(i_1)\, \mathbf{u}_2(i_2) \cdots \mathbf{u}_d(i_d).$$

Definition (Tensor rank)

The rank of a tensor $\mathbf{X} \in \mathbb{R}^{n_1 \times n_2 \times \cdots \times n_d}$ is the smallest number of rank one tensors that sum up to $\mathbf{X}$.

- Tensor of order d: $\mathbf{X} \in \mathbb{R}^{n_1 \times n_2 \times \cdots \times n_d}$.
- $\mathbf{X} \in \mathbb{R}^{n_1 \times n_2 \times \cdots \times n_d}$ is a rank one tensor: $\exists\, \mathbf{u}_1 \in \mathbb{R}^{n_1}, \mathbf{u}_2 \in \mathbb{R}^{n_2}, \ldots, \mathbf{u}_d \in \mathbb{R}^{n_d}$ s.t.

$$\mathbf{X} = \mathbf{u}_1 \otimes \mathbf{u}_2 \otimes \cdots \otimes \mathbf{u}_d, \text{ i.e., } \mathbf{X}(i_1, i_2, \ldots, i_d) = \mathbf{u}_1(i_1)\, \mathbf{u}_2(i_2) \cdots \mathbf{u}_d(i_d).$$

Definition (Tensor rank)

The rank of a tensor $\mathbf{X} \in \mathbb{R}^{n_1 \times n_2 \times \cdots \times n_d}$ is the smallest number of rank one tensors that sum up to $\mathbf{X}$.

- The set of rank-r tensors is not closed.

- Tensor of order d: $\mathbf{X} \in \mathbb{R}^{n_1 \times n_2 \times \cdots \times n_d}$.
- $\mathbf{X} \in \mathbb{R}^{n_1 \times n_2 \times \cdots \times n_d}$ is a rank one tensor: $\exists\, \mathbf{u}_1 \in \mathbb{R}^{n_1}, \mathbf{u}_2 \in \mathbb{R}^{n_2}, \ldots, \mathbf{u}_d \in \mathbb{R}^{n_d}$ s.t.

$$\mathbf{X} = \mathbf{u}_1 \otimes \mathbf{u}_2 \otimes \cdots \otimes \mathbf{u}_d, \text{ i.e., } \mathbf{X}(i_1, i_2, \ldots, i_d) = \mathbf{u}_1(i_1)\, \mathbf{u}_2(i_2) \cdots \mathbf{u}_d(i_d).$$

Definition (Tensor rank)

The rank of a tensor $\mathbf{X} \in \mathbb{R}^{n_1 \times n_2 \times \cdots \times n_d}$ is the smallest number of rank one tensors that sum up to $\mathbf{X}$.

- The set of rank-r tensors is not closed.
- Determining the rank of a tensor is NP-hard in general.

- Tensor of order d: $\mathbf{X} \in \mathbb{R}^{n_1 \times n_2 \times \cdots \times n_d}$.
- $\mathbf{X} \in \mathbb{R}^{n_1 \times n_2 \times \cdots \times n_d}$ is a rank one tensor: $\exists\, \mathbf{u}_1 \in \mathbb{R}^{n_1}, \mathbf{u}_2 \in \mathbb{R}^{n_2}, \ldots, \mathbf{u}_d \in \mathbb{R}^{n_d}$ s.t.

$$\mathbf{X} = \mathbf{u}_1 \otimes \mathbf{u}_2 \otimes \cdots \otimes \mathbf{u}_d, \text{ i.e., } \mathbf{X}(i_1, i_2, \ldots, i_d) = \mathbf{u}_1(i_1)\, \mathbf{u}_2(i_2) \cdots \mathbf{u}_d(i_d).$$

Definition (Tensor rank)

The rank of a tensor $\mathbf{X} \in \mathbb{R}^{n_1 \times n_2 \times \cdots \times n_d}$ is the smallest number of rank one tensors that sum up to $\mathbf{X}$.

- The set of rank-r tensors is not closed.
- Determining the rank of a tensor is NP-hard in general.
- **Nuclear tensor norm**: analog of matrix nuclear/trace norm

$$\|\mathbf{X}\|_* = \inf\left\{ \sum_{k=1}^{r} |c_k| : \mathbf{X} = \sum_{k=1}^{r} c_k \mathbf{u}_1^k \otimes \mathbf{u}_2^k \otimes \cdots \otimes \mathbf{u}_d^k, \; r \in \mathbb{N}, \|\mathbf{u}_i^k\|_{\ell_2} = 1, i \in [d], k \in [r] \right\}.$$

- Tensor of order d: $\mathbf{X} \in \mathbb{R}^{n_1 \times n_2 \times \cdots \times n_d}$.
- $\mathbf{X} \in \mathbb{R}^{n_1 \times n_2 \times \cdots \times n_d}$ is a rank one tensor: $\exists\, \mathbf{u}_1 \in \mathbb{R}^{n_1}, \mathbf{u}_2 \in \mathbb{R}^{n_2}, \ldots, \mathbf{u}_d \in \mathbb{R}^{n_d}$ s.t.

$$\mathbf{X} = \mathbf{u}_1 \otimes \mathbf{u}_2 \otimes \cdots \otimes \mathbf{u}_d, \text{ i.e., } \mathbf{X}(i_1, i_2, \ldots, i_d) = \mathbf{u}_1(i_1)\,\mathbf{u}_2(i_2) \cdots \mathbf{u}_d(i_d).$$

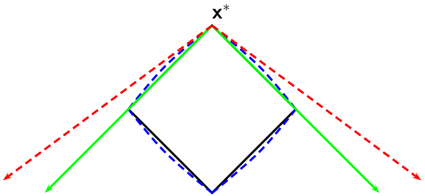
Definition (Tensor rank)

The rank of a tensor $\mathbf{X} \in \mathbb{R}^{n_1 \times n_2 \times \cdots \times n_d}$ is the smallest number of rank one tensors that sum up to $\mathbf{X}$.

- The set of rank-r tensors is not closed.
- Determining the rank of a tensor is NP-hard in general.
- **Nuclear tensor norm**: (NP-hard)

$$\|\mathbf{X}\|_* = \inf \Bigg\{ \sum_{k=1}^{r} |c_k| : \mathbf{X} = \sum_{k=1}^{r} c_k \mathbf{u}_1^k \otimes \mathbf{u}_2^k \otimes \cdots \otimes \mathbf{u}_d^k, \; r \in \mathbb{N}, \|\mathbf{u}_i^k\|_{\ell_2} = 1, i \in [d], k \in [r] \Bigg\}.$$

Relaxation



Proposition ([1])

$\hat{\mathbf{x}} = \mathbf{x}^*$ is the unique optimal solution of

$$\hat{\mathbf{x}} = \arg\min_{\mathbf{x}} \|\mathbf{x}\| \quad s.t.\ \mathbf{b} = \mathbf{\Phi x}$$

if and only if

$$\ker(\mathbf{\Phi}) \cap \mathcal{T}(\mathbf{x}^*) = \{\mathbf{0}\}.$$

- Tangent cone: $\mathcal{T}(\mathbf{x}) = \text{cone}(\mathbf{z} - \mathbf{x} : \|\mathbf{z}\| \leq \|\mathbf{x}\|)$

[1] V. Chandrasekaran, B. Recht, P. A. Parrilo, A. S. Willsky: *The Convex Geometry of Linear Inverse Problems*, 2012.

Convex relaxation via real algebraic geometry

- $\mathbb{R}[\mathbf{x}]$: set of all polynomials in variables $x_1, x_2, \ldots, x_n$
- I ideal in $\mathbb{R}[\mathbf{x}]$
- $\mathbb{R}[\mathbf{x}]_k$: set of polynomials in $\mathbb{R}[\mathbf{x}]$ of degree at most k
- real variety: $\nu_{\mathbb{R}}(I) = \{\mathbf{x} \in \mathbb{R}^n : f(\mathbf{x}) = 0, \text{ for all } f \in I\}$

Convex relaxation via real algebraic geometry

$$\mathbf{X} = \begin{pmatrix} X_{11} & X_{12} \\ X_{21} & X_{22} \end{pmatrix}$$

- $\mathbb{R}[\mathbf{x}]$: set of all polynomials in variables $X_{11}, X_{12}, X_{21}, X_{22}$
- I ideal in $\mathbb{R}[\mathbf{x}]$
- $\mathbb{R}[\mathbf{x}]_k$: set of polynomials in $\mathbb{R}[\mathbf{x}]$ of degree at most k
- real variety: $\nu_{\mathbb{R}}(I) = \{\mathbf{x} \in \mathbb{R}^n : f(\mathbf{x}) = 0, \text{ for all } f \in I\}$

Central Problem in optimization

Central Problem in optimization

- $\max_{\mathbf{x}} \langle \mathbf{c}, \mathbf{x} \rangle$ s.t. $\mathbf{x} \in \mathcal{S}$

Central Problem in optimization

- $\max_{\mathbf{x}} \langle \mathbf{c}, \mathbf{x} \rangle$ s.t. $\mathbf{x} \in \mathcal{S}$ $\quad\Leftrightarrow\quad$ $\max_{\mathbf{x}} \langle \mathbf{c}, \mathbf{x} \rangle$ s.t. $\mathbf{x} \in \overline{\text{conv}\left(\mathcal{S}\right)}$

Central Problem in optimization

- $\max_{\mathbf{x}} \langle \mathbf{c}, \mathbf{x} \rangle$ s.t. $\mathbf{x} \in \mathcal{S} \quad \Leftrightarrow \quad \max_{\mathbf{x}} \langle \mathbf{c}, \mathbf{x} \rangle$ s.t. $\mathbf{x} \in \overline{\text{conv}\,(\mathcal{S})}$
- Linear programming: $\mathcal{S} = \{\mathbf{x} : \mathbf{A}\mathbf{x} \leq \mathbf{b}\}$

Central Problem in optimization

- $\max_{\mathbf{x}} \langle \mathbf{c}, \mathbf{x} \rangle$ s.t. $\mathbf{x} \in \mathcal{S} \quad \Leftrightarrow \quad \max_{\mathbf{x}} \langle \mathbf{c}, \mathbf{x} \rangle$ s.t. $\mathbf{x} \in \overline{\text{conv}\left(\mathcal{S}\right)}$
- $\mathcal{S} = \nu_{\mathbb{R}}\left(I\right) = \{\mathbf{x} \in \mathbb{R}^n : f\left(\mathbf{x}\right) = 0, \forall f \in I\}$

Central Problem in optimization

- $\max_{\mathbf{x}} \langle \mathbf{c}, \mathbf{x} \rangle$ s.t. $\mathbf{x} \in \mathcal{S} \quad \Leftrightarrow \quad \max_{\mathbf{x}} \langle \mathbf{c}, \mathbf{x} \rangle$ s.t. $\mathbf{x} \in \overline{\mathrm{conv}\left(\mathcal{S}\right)}$
- $\mathcal{S} = \nu_{\mathbb{R}}\left(I\right) = \{\mathbf{x} \in \mathbb{R}^n : f\left(\mathbf{x}\right) = 0, \forall f \in I\}$
- $\overline{\mathrm{conv}\left(\nu_{\mathbb{R}}\left(I\right)\right)} = \left\{\mathbf{x} \in \mathbb{R}^n : \ell\left(\mathbf{x}\right) \geq 0, \forall \ell \text{ affine s.t. } \ell\mid_{\nu_{\mathbb{R}}(I)} \geq 0\right\}$

Central Problem in optimization

- $\max_{\mathbf{x}} \langle \mathbf{c}, \mathbf{x} \rangle$ s.t. $\mathbf{x} \in \mathcal{S} \quad \Leftrightarrow \quad \max_{\mathbf{x}} \langle \mathbf{c}, \mathbf{x} \rangle$ s.t. $\mathbf{x} \in \overline{\mathrm{conv}\,(\mathcal{S})}$
- $\mathcal{S} = \nu_{\mathbb{R}}(I) = \{\mathbf{x} \in \mathbb{R}^n : f(\mathbf{x}) = 0, \forall f \in I\}$
- $\overline{\mathrm{conv}\,(\nu_{\mathbb{R}}(I))} = \left\{\mathbf{x} \in \mathbb{R}^n : \ell(\mathbf{x}) \geq 0, \forall \ell \text{ affine s.t. } \ell\,|_{\nu_{\mathbb{R}}(I)} \geq 0\right\}$
- Idea: relax condition $\ell\,|_{\nu_{\mathbb{R}}(I)} \geq 0$

Central Problem in optimization

- $\max_{\mathbf{x}} \langle \mathbf{c}, \mathbf{x} \rangle$ s.t. $\mathbf{x} \in \mathcal{S} \quad \Leftrightarrow \quad \max_{\mathbf{x}} \langle \mathbf{c}, \mathbf{x} \rangle$ s.t. $\mathbf{x} \in \overline{\operatorname{conv}(\mathcal{S})}$
- $\mathcal{S} = \nu_{\mathbb{R}}(I) = \{\mathbf{x} \in \mathbb{R}^n : f(\mathbf{x}) = 0, \forall f \in I\}$
- $\overline{\operatorname{conv}(\nu_{\mathbb{R}}(I))} = \left\{\mathbf{x} \in \mathbb{R}^n : \ell(\mathbf{x}) \geq 0, \forall \ell \text{ affine s.t. } \ell\,|_{\nu_{\mathbb{R}}(I)} \geq 0\right\}$
- Idea: relax condition $\ell\,|_{\nu_{\mathbb{R}}(I)} \geq 0$

$$\ell(\mathbf{x}) = \sum_{i=1}^{t} h_i^2(\mathbf{x}) + g(\mathbf{x}), \quad h_i \in \mathbb{R}[\mathbf{x}], g \in I$$

- Idea: relax condition $\ell\mid_{\nu_{\mathbb{R}}(I)} \geq 0$

$$\ell(\mathbf{x}) = \sum_{i=1}^{t} h_i^2(\mathbf{x}) + g(\mathbf{x}), \quad h_i \in \mathbb{R}[\mathbf{x}], g \in I$$

- Idea: relax condition $\ell\,|_{\nu_{\mathbb{R}}(I)} \geq 0$

$$\ell\left(\mathbf{x}\right) = \sum_{i=1}^{t} h_i^2\left(\mathbf{x}\right) + g\left(\mathbf{x}\right), \quad h_i \in \mathbb{R}\left[\mathbf{x}\right], g \in I$$

1. **Lassere's method**

$$\deg\left(h_i\right) \leq k, \quad g(\mathbf{x}) = \sum_{i=1}^{m} g_j\left(\mathbf{x}\right) f_j\left(\mathbf{x}\right),$$

$$I = \langle f_1, \ldots, f_m \rangle, \; g_j f_j \in \mathbb{R}\left[\mathbf{x}\right]_{2k}$$

(ℓ is k-sos mod $\{f_1, f_2, \ldots, f_m\}$)

- Idea: relax condition $\ell \,|_{\nu_{\mathbb{R}}(I)} \geq 0$

$$\ell(\mathbf{x}) = \sum_{i=1}^{t} h_i^2(\mathbf{x}) + g(\mathbf{x}), \quad h_i \in \mathbb{R}[\mathbf{x}], g \in I$$

1. **Lassere's method**

$$\deg(h_i) \leq k, \quad g(\mathbf{x}) = \sum_{i=1}^{m} g_j(\mathbf{x}) f_j(\mathbf{x}),$$
$$I = \langle f_1, \ldots, f_m \rangle, \; g_j f_j \in \mathbb{R}[\mathbf{x}]_{2k}$$

(ℓ is k-sos mod $\{f_1, f_2, \ldots, f_m\}$)

2. **Theta bodies**

$$\deg(h_i) \leq k, \; g \in I \quad (\ell \text{ is } k\text{-sos mod I})$$

- Idea: relax condition $\ell\mid_{\nu_{\mathbb{R}}(I)} \geq 0$

$$\ell(\mathbf{x}) = \sum_{i=1}^{t} h_i^2(\mathbf{x}) + g(\mathbf{x}), \quad h_i \in \mathbb{R}[\mathbf{x}], g \in I$$

1. **Lassere's method**

$$\deg(h_i) \leq k, \quad g(\mathbf{x}) = \sum_{i=1}^{m} g_j(\mathbf{x}) f_j(\mathbf{x}),$$

$$I = \langle f_1, \ldots, f_m \rangle, \; g_j f_j \in \mathbb{R}[\mathbf{x}]_{2k}$$

(ℓ is k-sos mod $\{f_1, f_2, \ldots, f_m\}$)

2. **Theta bodies**

$$\deg(h_i) \leq k, \; g \in I \quad (\ell \text{ is } k\text{-sos mod I})$$

Theta bodies

Definition (Theta Body)

Let $I \subseteq \mathbb{R}[\mathbf{x}]$ be an ideal. For a positive integer k, the k-th theta body of an ideal I is

$$\mathrm{TH}_k(I) := \{\mathbf{x} \in \mathbb{R}^n : f(\mathbf{x}) \geq 0 \text{ for every } f \text{ affine and } k\text{-sos mod } I\}.$$

- $\mathrm{TH}_1(I) \supseteq \mathrm{TH}_2(I) \supseteq \cdots \supseteq \overline{\mathrm{conv}(\nu_{\mathbb{R}}(I))}$
- closed, convex sets

Theta bodies

Definition (Theta Body)

Let $I \subseteq \mathbb{R}[\mathbf{x}]$ be an ideal. For a positive integer k, the k-th theta body of an ideal I is

$$\mathrm{TH}_k(I) := \{\mathbf{x} \in \mathbb{R}^n : f(\mathbf{x}) \geq 0 \text{ for every } f \text{ affine and } k\text{-sos mod } I\}.$$

- $\mathrm{TH}_1(I) \supseteq \mathrm{TH}_2(I) \supseteq \cdots \supseteq \overline{\mathrm{conv}(\nu_{\mathbb{R}}(I))}$
- closed, convex sets
- Idea: define an ideal I s.t.

$$\nu_{\mathbb{R}}(I) = \{\text{all rank-one, Frobenius norm-one tensors}\}.$$

Theta bodies

Definition (Theta Body)

Let $I \subseteq \mathbb{R}[\mathbf{x}]$ be an ideal. For a positive integer k, the k-th theta body of an ideal I is

$$\mathrm{TH}_k(I) := \{\mathbf{x} \in \mathbb{R}^n : f(\mathbf{x}) \geq 0 \text{ for every } f \text{ affine and } k\text{-sos mod } I\}.$$

- $\mathrm{TH}_1(I) \supseteq \mathrm{TH}_2(I) \supseteq \cdots \supseteq \overline{\mathrm{conv}(\nu_{\mathbb{R}}(I))}$
- closed, convex sets
- Idea: define an ideal I s.t.

$$\nu_{\mathbb{R}}(I) = \{\text{all rank-one, Frobenius norm-one } (2 \times 2)\text{-matrices}\}.$$

Theta bodies

Definition (Theta Body)

Let $I \subseteq \mathbb{R}[\mathbf{x}]$ be an ideal. For a positive integer k, the k-th theta body of an ideal I is

$$\mathrm{TH}_k(I) := \{\mathbf{x} \in \mathbb{R}^n : f(\mathbf{x}) \geq 0 \text{ for every } f \text{ affine and } k\text{-sos mod } I\}.$$

- $\mathrm{TH}_1(I) \supseteq \mathrm{TH}_2(I) \supseteq \cdots \supseteq \overline{\mathrm{conv}(\nu_{\mathbb{R}}(I))}$
- closed, convex sets
- Idea: define an ideal I s.t.

$$\nu_{\mathbb{R}}(I) = \{\text{all rank-one, Frobenius norm-one } (2 \times 2)\text{-matrices}\}.$$

- $B_{\|\cdot\|_{\theta_1}} = \mathrm{TH}_1(I) \Rightarrow \|\mathbf{X}\|_{\theta_1} = \{\inf_t t \text{ s.t. } \mathbf{X} \in t\,\mathrm{TH}_1(I)\}$

Computation of $\mathrm{TH}_k(I)$

Definition (Theta Body)

Let $I \subseteq \mathbb{R}[\mathbf{x}]$ be an ideal. For a positive integer k, the k-th theta body of an ideal I is

$$\mathrm{TH}_k(I) := \{\mathbf{x} \in \mathbb{R}^n : f(\mathbf{x}) \geq 0 \text{ for every } f \text{ affine and } k\text{-sos mod } I\}.$$

- difficult to compute the theta body directly from the definition

Computation of $\mathrm{TH}_k(I)$

Definition (Theta Body)

Let $I \subseteq \mathbb{R}[\mathbf{x}]$ be an ideal. For a positive integer k, the k-th theta body of an ideal I is

$$\mathrm{TH}_k(I) := \{\mathbf{x} \in \mathbb{R}^n : f(\mathbf{x}) \geq 0 \text{ for every } f \text{ affine and } k\text{-sos mod } I\}.$$

- difficult to compute the theta body directly from the definition
- alternative approach relies on computing a Groebner basis for the ideal and results in semidefinite program for
  - computing θ_1-norm: $\|\mathbf{X}\|_{\theta_1}$, for all $\mathbf{X} \in \mathbb{R}^{n_1 \times n_2 \times n_3}$
  - low rank tensor recovery:

$$\min_{\mathbf{Z} \in \mathbb{R}^{n_1 \times n_2 \times n_3}} \|\mathbf{Z}\|_{\theta_1} \quad \text{s.t.} \quad \mathbf{\Phi}(\mathbf{Z}) = \mathbf{b}$$

Computation of $\mathrm{TH}_k(I)$

Definition (Theta Body)

Let $I \subseteq \mathbb{R}[\mathbf{x}]$ be an ideal. For a positive integer k, the k-th theta body of an ideal I is

$$\mathrm{TH}_k(I) := \{\mathbf{x} \in \mathbb{R}^n : f(\mathbf{x}) \geq 0 \text{ for every } f \text{ affine and } k\text{-sos mod } I\}.$$

- difficult to compute the theta body directly from the definition
- alternative approach relies on computing a Groebner basis for the ideal and results in semidefinite program for
  - computing θ_1-norm: $\|\mathbf{X}\|_{\theta_1}$, for all $\mathbf{X} \in \mathbb{R}^{2\times 2}$
  - low rank tensor recovery:

$$\min_{\mathbf{Z} \in \mathbb{R}^{2\times 2}} \|\mathbf{Z}\|_{\theta_1} \quad \text{s.t.} \quad \mathbf{\Phi}(\mathbf{Z}) = \mathbf{b}$$

2×2-matrices

- Idea: define an ideal I s.t.

$$\nu_{\mathbb{R}}(I) = \{\text{all rank-one, Frobenius norm-one } (2 \times 2)\text{-matrices}\}.$$

2×2-matrices

- Idea: define an ideal I s.t.

  $$\nu_{\mathbb{R}}(I) = \{\text{all rank-one, Frobenius norm-one } (2 \times 2)\text{-matrices}\}.$$

- A matrix $\mathbf{X} \in \mathbb{R}^{2\times 2}$ is a rank-one, Frobenius norm-one matrix iff

  $$X_{21}X_{12} - X_{11}X_{22} = 0, \quad X_{11}^2 + X_{12}^2 + X_{21}^2 + X_{22}^2 = 1$$

2×2-matrices

- Idea: define an ideal I s.t.

  $$\nu_{\mathbb{R}}(I) = \{\text{all rank-one, Frobenius norm-one } (2 \times 2)\text{-matrices}\}.$$

- A matrix $\mathbf{X} \in \mathbb{R}^{2\times 2}$ is a rank-one, Frobenius norm-one matrix iff

  $$X_{21}X_{12} - X_{11}X_{22} = 0, \quad X_{11}^2 + X_{12}^2 + X_{21}^2 + X_{22}^2 = 1$$

- $I = \left\{ g \cdot (X_{21}X_{12} - X_{11}X_{22}) + h \cdot \left(\sum_{i=1}^2 \sum_{j=1}^2 X_{ij}^2 - 1 \right) : g, h \in \mathbb{R}[\mathbf{x}] \right\}$

2×2-matrices

- Idea: define an ideal I s.t.

  $$\nu_{\mathbb{R}}(I) = \{\text{all rank-one, Frobenius norm-one } (2 \times 2)\text{-matrices}\}.$$

- A matrix $\mathbf{X} \in \mathbb{R}^{2\times 2}$ is a rank-one, Frobenius norm-one matrix iff

  $$X_{21}X_{12} - X_{11}X_{22} = 0, \quad X_{11}^2 + X_{12}^2 + X_{21}^2 + X_{22}^2 = 1$$

- $I = \left\langle X_{21}X_{12} - X_{11}X_{22}, X_{11}^2 + X_{12}^2 + X_{21}^2 + X_{22}^2 - 1 \right\rangle$

1. $\mathbf{x}^{\boldsymbol{\alpha}} >_{grevlex} \mathbf{x}^{\boldsymbol{\beta}}$ if $|\boldsymbol{\alpha}| > |\boldsymbol{\beta}|$ or $|\boldsymbol{\alpha}| = |\boldsymbol{\beta}|$ and the right-most nonzero entry of $\boldsymbol{\alpha} - \boldsymbol{\beta} \in \mathbb{Z}^n_{\geq 0}$ is negative.

1. $\mathbf{x}^{\boldsymbol{\alpha}} >_{grevlex} \mathbf{x}^{\boldsymbol{\beta}}$ if $|\boldsymbol{\alpha}| > |\boldsymbol{\beta}|$ or $|\boldsymbol{\alpha}| = |\boldsymbol{\beta}|$ and the right-most nonzero entry of $\boldsymbol{\alpha} - \boldsymbol{\beta} \in \mathbb{Z}^n_{\geq 0}$ is negative.

Definition (Groebner basis)

Fix a monomial order. A basis $G = \{g_1, g_2, \ldots, g_s\}$ of a polynomial ideal $I \subset \mathbb{R}[\mathbf{x}]$ is a Groebner basis (or standard basis) if for all $f \in \mathbb{R}[\mathbf{x}]$ there exist **unique** $r \in \mathbb{R}[\mathbf{x}]$ and $g \in I$ s.t.

$$f = g + r$$

and no monomial of r is divisible by any of the leading monomials in G.

Computation of $\mathrm{TH}_k(I)$

Computation of $\mathrm{TH}_k(I)$

1. Define $\nu_{\mathbb{R}}(I)$ and a corresponding ideal I.

Computation of $\mathrm{TH}_k(I)$

1. $I = \langle X_{12}X_{21} - X_{11}X_{22}, X_{11}^2 + X_{12}^2 + X_{21}^2 + X_{22}^2 - 1 \rangle$

Computation of $\mathrm{TH}_k(I)$

1. $I = \langle X_{12}X_{21} - X_{11}X_{22}, X_{11}^2 + X_{12}^2 + X_{21}^2 + X_{22}^2 - 1 \rangle$
2. Find a Groebner basis for the ideal I (with respect to some monomial ordering).

Computation of $\mathrm{TH}_k(I)$

1. $I = \left\langle X_{12}X_{21} - X_{11}X_{22}, X_{11}^2 + X_{12}^2 + X_{21}^2 + X_{22}^2 - 1 \right\rangle$.

Computation of $\mathrm{TH}_k(I)$

1. $I = \langle X_{12}X_{21} - X_{11}X_{22}, X_{11}^2 + X_{12}^2 + X_{21}^2 + X_{22}^2 - 1 \rangle$.
2. Find a **θ-basis** $\mathcal{B} = \mathcal{B}_1 \cup \mathcal{B}_2 \cup \ldots = \{f_0 + I, f_1 + I, \ldots\}$ of $\mathbb{R}[\mathbf{x}]/I$, where
   - $\mathcal{B}_1 = \{1 + I, x_1 + I, \ldots, x_n + I\}$

Computation of $\mathrm{TH}_k(I)$

1. $I = \langle X_{12}X_{21} - X_{11}X_{22}, X_{11}^2 + X_{12}^2 + X_{21}^2 + X_{22}^2 - 1 \rangle$.
2. Find a **θ-basis** $\mathcal{B} = \mathcal{B}_1 \cup \mathcal{B}_2 \cup \ldots = \{f_0 + I, f_1 + I, \ldots\}$ of $\mathbb{R}[\mathbf{x}]/I$, where
   - $\mathcal{B}_1 = \{1 + I, X_{11} + I, X_{12} + I, X_{21} + I, X_{22} + I\}$

Computation of $\mathrm{TH}_k(I)$

1. $I = \langle X_{12}X_{21} - X_{11}X_{22}, X_{11}^2 + X_{12}^2 + X_{21}^2 + X_{22}^2 - 1 \rangle$.
2. Find a θ-**basis** $\mathcal{B} = \mathcal{B}_1 \cup \mathcal{B}_2 \cup \ldots = \{f_0 + I, f_1 + I, \ldots\}$ of $\mathbb{R}[\mathbf{x}]/I$, where
   - $\mathcal{B}_1 = \{1 + I, X_{11} + I, X_{12} + I, X_{21} + I, X_{22} + I\}$
   - if $\deg(f_i + I), \deg(f_j + I) \leq k \Rightarrow f_i f_j + I$ is in the $\mathbb{R}$-span of $\mathcal{B}_{2k}$

Computation of $\mathrm{TH}_k(I)$

1. $I = \langle X_{12}X_{21} - X_{11}X_{22}, X_{11}^2 + X_{12}^2 + X_{21}^2 + X_{22}^2 - 1 \rangle$.
2. Find a **θ-basis** $\mathcal{B} = \mathcal{B}_1 \cup \mathcal{B}_2 \cup \ldots = \{f_0 + I, f_1 + I, \ldots\}$ of $\mathbb{R}[\mathbf{x}]/I$, where
   - $\mathcal{B}_1 = \{1 + I, X_{11} + I, X_{12} + I, X_{21} + I, X_{22} + I\}$
   - if $\deg(f_i + I), \deg(f_j + I) \leq k \Rightarrow f_i f_j + I$ is in the $\mathbb{R}$-span of $\mathcal{B}_{2k}$
   - $X_{11}^2 + I = -\left(X_{12}^2 + I\right) - \left(X_{21}^2 + I\right) - \left(X_{22}^2 + I\right) + (1 + I)$

Computation of $\mathrm{TH}_k(I)$

1. $I = \langle X_{12}X_{21} - X_{11}X_{22}, X_{11}^2 + X_{12}^2 + X_{21}^2 + X_{22}^2 - 1 \rangle$.
2. Find a **θ-basis** $\mathcal{B} = \mathcal{B}_1 \cup \mathcal{B}_2 \cup \ldots = \{f_0 + I, f_1 + I, \ldots\}$ of $\mathbb{R}[\mathbf{x}]/I$, where
   - $\mathcal{B}_1 = \{1 + I, X_{11} + I, X_{12} + I, X_{21} + I, X_{22} + I\}$
   - $$\begin{aligned}\mathcal{B}_2 = \mathcal{B}_1 \cup \{&X_{12}^2 + I, X_{21}^2 + I, X_{22}^2 + I, X_{11}X_{12} + I,\\ &X_{11}X_{21} + I, X_{11}X_{22} + I, X_{12}X_{22} + I, X_{21}X_{22} + I\}\end{aligned}$$

Computation of $\mathrm{TH}_k\left(I\right)$

1. Compute a combinatorial moment matrix $M_{\mathcal{B}_k}\left(\mathbf{X}, \mathbf{y}\right)$.

Computation of $\mathrm{TH}_k(I)$

1. Compute a combinatorial moment matrix $M_{\mathcal{B}_k}(\mathbf{X}, \mathbf{y})$.
   - $[\mathbf{x}]_{\mathcal{B}_k}^T = \{$all elements of $\mathcal{B}_k$ in order$\}$

Computation of $\mathrm{TH}_k(I)$

1. Compute a combinatorial moment matrix $M_{\mathcal{B}_1}(\mathbf{X}, \mathbf{y})$.
   - $[\mathbf{x}]_{\mathcal{B}_1}^T = [1 + I, X_{11} + I, X_{12} + I, X_{21} + I, X_{22} + I]$

Computation of $\mathrm{TH}_k(I)$

1. Compute a combinatorial moment matrix $M_{\mathcal{B}_1}(\mathbf{X}, \mathbf{y})$.
   - $[\mathbf{x}]_{\mathcal{B}_1}^T = [1 + I, X_{11} + I, X_{12} + I, X_{21} + I, X_{22} + I]$
   - $\mathbf{X}_{\mathcal{B}_k} = [\mathbf{x}_{\mathcal{B}_k}][\mathbf{x}_{\mathcal{B}_k}]^T \Rightarrow [\mathbf{X}_{\mathcal{B}_k}]_{i,j} = f_i f_j + I = \sum_{f_l + I \in \mathcal{B}_{2k}} \lambda_{i,j}^l (f_l + I)$

Computation of $\mathrm{TH}_k(I)$

1. Compute a combinatorial moment matrix $M_{\mathcal{B}_1}(\mathbf{X}, \mathbf{y})$.
   - $[\mathbf{x}]_{\mathcal{B}_1}^T = [1 + I, X_{11} + I, X_{12} + I, X_{21} + I, X_{22} + I]$
   - $\mathbf{X}_{\mathcal{B}_k} = [\mathbf{x}_{\mathcal{B}_k}][\mathbf{x}_{\mathcal{B}_k}]^T \Rightarrow [\mathbf{X}_{\mathcal{B}_k}]_{i,j} = f_i f_j + I = \sum_{f_l + I \in \mathcal{B}_{2k}} \lambda_{i,j}^l (f_l + I)$

$$\mathbf{X}_{\mathcal{B}_1} = \begin{bmatrix} 1+I & X_{11}+I & X_{12}+I & X_{21}+I & X_{22}+I \\ & X_{11}^2+I & X_{11}X_{12}+I & X_{11}X_{21}+I & X_{11}X_{22}+I \\ & & X_{12}^2+I & X_{12}X_{21}+I & X_{12}X_{22}+I \\ & & & X_{21}^2+I & X_{21}X_{22}+I \\ & & & & X_{22}^2+I \end{bmatrix}$$

Computation of $\mathrm{TH}_k(I)$

1. Compute a combinatorial moment matrix $M_{\mathcal{B}_1}(\mathbf{X}, \mathbf{y})$.
   - $[\mathbf{x}]_{\mathcal{B}_1}^T = [1+I, X_{11}+I, X_{12}+I, X_{21}+I, X_{22}+I]$
   - $\mathbf{X}_{\mathcal{B}_k} = [\mathbf{x}_{\mathcal{B}_k}][\mathbf{x}_{\mathcal{B}_k}]^T \Rightarrow [\mathbf{X}_{\mathcal{B}_k}]_{i,j} = f_i f_j + I = \sum_{f_l + I \in \mathcal{B}_{2k}} \lambda_{i,j}^l (f_l + I)$

$$\mathbf{X}_{\mathcal{B}_1} = \begin{bmatrix} 1 & X_{11} & X_{12} & X_{21} & X_{22} \\ & X_{11}^2 & X_{11}X_{12} & X_{11}X_{21} & X_{11}X_{22} \\ & & X_{12}^2 & X_{12}X_{21} & X_{12}X_{22} \\ & & & X_{21}^2 & X_{21}X_{22} \\ & & & & X_{22}^2 \end{bmatrix}$$

Computation of $\mathrm{TH}_k(I)$

1. Compute a combinatorial moment matrix $M_{\mathcal{B}_1}(\mathbf{X}, \mathbf{y})$.
   - $[\mathbf{x}]_{\mathcal{B}_1}^T = [1+I, X_{11}+I, X_{12}+I, X_{21}+I, X_{22}+I]$
   - $\mathbf{X}_{\mathcal{B}_k} = [\mathbf{x}_{\mathcal{B}_k}][\mathbf{x}_{\mathcal{B}_k}]^T \Rightarrow [\mathbf{X}_{\mathcal{B}_k}]_{i,j} = f_i f_j + I = \sum_{f_l+I\in\mathcal{B}_{2k}} \lambda_{i,j}^l (f_l + I)$

$$\mathbf{X}_{\mathcal{B}_1} = \begin{bmatrix} 1 & X_{11} & X_{12} & X_{21} & X_{22} \\ & \boxed{X_{11}^2} & X_{11}X_{12} & X_{11}X_{21} & X_{11}X_{22} \\ & & X_{12}^2 & \boxed{X_{12}X_{21}} & X_{12}X_{22} \\ & & & X_{21}^2 & X_{21}X_{22} \\ & & & & X_{22}^2 \end{bmatrix}$$

Computation of $\mathrm{TH}_k(I)$

1. Compute a combinatorial moment matrix $M_{\mathcal{B}_1}(\mathbf{X}, \mathbf{y})$.
   - $[\mathbf{x}]_{\mathcal{B}_1}^T = [1 + I, X_{11} + I, X_{12} + I, X_{21} + I, X_{22} + I]$
   - $\mathbf{X}_{\mathcal{B}_k} = [\mathbf{x}_{\mathcal{B}_k}][\mathbf{x}_{\mathcal{B}_k}]^T \Rightarrow [\mathbf{X}_{\mathcal{B}_k}]_{i,j} = f_i f_j + I = \sum_{f_l + I \in \mathcal{B}_{2k}} \lambda_{i,j}^l (f_l + I)$

$$X_{11}^2 + I = -X_{12}^2 - X_{21}^2 - X_{22}^2 + 1 + I$$

$$\mathbf{X}_{\mathcal{B}_1} = \begin{bmatrix} 1 & X_{11} & X_{12} & X_{21} & X_{22} \\ & \boxed{X_{11}^2} & X_{11}X_{12} & X_{11}X_{21} & X_{11}X_{22} \\ & & X_{12}^2 & \boxed{X_{12}X_{21}} & X_{12}X_{22} \\ & & & X_{21}^2 & X_{21}X_{22} \\ & & & & X_{22}^2 \end{bmatrix}$$

Computation of $\mathrm{TH}_k(I)$

1. Compute a combinatorial moment matrix $M_{\mathcal{B}_1}(\mathbf{X},\mathbf{y})$.
   - $[\mathbf{x}]_{\mathcal{B}_1}^T = [1+I, X_{11}+I, X_{12}+I, X_{21}+I, X_{22}+I]$
   - $\mathbf{X}_{\mathcal{B}_k} = [\mathbf{x}_{\mathcal{B}_k}][\mathbf{x}_{\mathcal{B}_k}]^T \Rightarrow [\mathbf{X}_{\mathcal{B}_k}]_{i,j} = f_i f_j + I = \sum_{f_l+I\in\mathcal{B}_{2k}} \lambda_{i,j}^l (f_l + I)$

$$X_{11}^2 + I = -X_{12}^2 - X_{21}^2 - X_{22}^2 + 1 + I$$

$$\mathbf{X}_{\mathcal{B}_1} = \begin{bmatrix} 1 & X_{11} & X_{12} & X_{21} & X_{22} \\ & \boxed{X_{11}^2} & X_{11}X_{12} & X_{11}X_{21} & X_{11}X_{22} \\ & & X_{12}^2 & \boxed{X_{12}X_{21}} & X_{12}X_{22} \\ & & & X_{21}^2 & X_{21}X_{22} \\ & & & & X_{22}^2 \end{bmatrix}$$

$$X_{12}X_{21} + I = X_{11}X_{22} + I$$

Computation of $\mathrm{TH}_k(I)$

$$\mathbf{X}_{\mathcal{B}_1} = \begin{bmatrix} 1 & X_{11} & X_{12} & X_{21} & X_{22} \\ & -X_{12}^2 - X_{21}^2 - X_{22}^2 + 1 & X_{11}X_{12} & X_{11}X_{21} & X_{11}X_{22} \\ & & X_{12}^2 & X_{11}X_{22} & X_{12}X_{22} \\ & & & X_{21}^2 & X_{21}X_{22} \\ & & & & X_{22}^2 \end{bmatrix}$$

1. $[\mathbf{X}_{\mathcal{B}_k}]_{i,j} = f_i f_j + I = \sum_{f_l + I \in \mathcal{B}_{2k}} \lambda_{i,j}^l (f_l + I)$

Computation of $\mathrm{TH}_k(I)$

$$\mathbf{X}_{\mathcal{B}_1} = \begin{bmatrix} 1 & X_{11} & X_{12} & X_{21} & X_{22} \\ & -X_{12}^2 - X_{21}^2 - X_{22}^2 + 1 & X_{11}X_{12} & X_{11}X_{21} & X_{11}X_{22} \\ & & X_{12}^2 & X_{11}X_{22} & X_{12}X_{22} \\ & & & X_{21}^2 & X_{21}X_{22} \\ & & & & X_{22}^2 \end{bmatrix}$$

1. $[\mathbf{X}_{\mathcal{B}_k}]_{i,j} = f_i f_j + I = \sum_{f_l + I \in \mathcal{B}_{2k}} \lambda_{i,j}^l (f_l + I)$
2. $[\mathbf{M}_{\mathcal{B}_k}(\mathbf{y})]_{i,j} = \sum_{f_l + I \in \mathcal{B}_{2k}} \lambda_{i,j}^l y_l$

Computation of $\mathrm{TH}_k(I)$

$$\mathbf{X}_{\mathcal{B}_1} = \begin{bmatrix} \boxed{1} & X_{11} & X_{12} & X_{21} & X_{22} \\ & -X_{12}^2 - X_{21}^2 - X_{22}^2 + \boxed{1} & X_{11}X_{12} & X_{11}X_{21} & X_{11}X_{22} \\ & & X_{12}^2 & X_{11}X_{22} & X_{12}X_{22} \\ & & & X_{21}^2 & X_{21}X_{22} \\ & & & & X_{22}^2 \end{bmatrix}$$

1. $[\mathbf{X}_{\mathcal{B}_k}]_{i,j} = f_i f_j + I = \sum_{f_l + I \in \mathcal{B}_{2k}} \lambda_{i,j}^l (f_l + I)$
2. $[\mathbf{M}_{\mathcal{B}_k}(\mathbf{y})]_{i,j} = \sum_{f_l + I \in \mathcal{B}_{2k}} \lambda_{i,j}^l y_l$

Computation of $\mathrm{TH}_k(I)$

y_0

$$\mathbf{X}_{\mathcal{B}_1} = \begin{bmatrix} \boxed{1} & X_{11} & X_{12} & X_{21} & X_{22} \\ & -X_{12}^2 - X_{21}^2 - X_{22}^2 + \boxed{1} & X_{11}X_{12} & X_{11}X_{21} & X_{11}X_{22} \\ & & X_{12}^2 & X_{11}X_{22} & X_{12}X_{22} \\ & & & X_{21}^2 & X_{21}X_{22} \\ & & & & X_{22}^2 \end{bmatrix}$$

1. $[\mathbf{X}_{\mathcal{B}_k}]_{i,j} = f_i f_j + I = \sum_{f_l + I \in \mathcal{B}_{2k}} \lambda_{i,j}^l (f_l + I)$
2. $[\mathbf{M}_{\mathcal{B}_k}(\mathbf{y})]_{i,j} = \sum_{f_l + I \in \mathcal{B}_{2k}} \lambda_{i,j}^l y_l$

Computation of $\mathrm{TH}_k\,(I)$

y_0

$$
\mathbf{X}_{\mathcal{B}_1} = \begin{bmatrix}
\boxed{1} & \boxed{X_{11}} & \boxed{X_{12}} & \boxed{X_{21}} & \boxed{X_{22}} \\
 & -X_{12}^2 - X_{21}^2 - X_{22}^2 + \boxed{1} & X_{11}X_{12} & X_{11}X_{21} & X_{11}X_{22} \\
 & & X_{12}^2 & X_{11}X_{22} & X_{12}X_{22} \\
 & & & X_{21}^2 & X_{21}X_{22} \\
 & & & & X_{22}^2
\end{bmatrix}
$$

1. $[\mathbf{X}_{\mathcal{B}_k}]_{i,j} = f_i f_j + I = \sum_{f_l + I \in \mathcal{B}_{2k}} \lambda_{i,j}^l (f_l + I)$
2. $[\mathbf{M}_{\mathcal{B}_k}(\mathbf{y})]_{i,j} = \sum_{f_l + I \in \mathcal{B}_{2k}} \lambda_{i,j}^l y_l$

Computation of $\mathrm{TH}_k(I)$

y_0 $\qquad\qquad\qquad$ y_1

$$\mathbf{X}_{\mathcal{B}_1} = \begin{bmatrix} \boxed{1} & \boxed{X_{11}} & \boxed{X_{12}} & \boxed{X_{21}} & \boxed{X_{22}} \\ & -X_{12}^2 - X_{21}^2 - X_{22}^2 + \boxed{1} & \boxed{X_{11}X_{12}} & X_{11}X_{21} & X_{11}X_{22} \\ & & X_{12}^2 & X_{11}X_{22} & X_{12}X_{22} \\ & & & X_{21}^2 & X_{21}X_{22} \\ & & & & X_{22}^2 \end{bmatrix}$$

1. $[\mathbf{X}_{\mathcal{B}_k}]_{i,j} = f_i f_j + I = \sum_{f_l + I \in \mathcal{B}_{2k}} \lambda_{i,j}^l (f_l + I)$
2. $[\mathbf{M}_{\mathcal{B}_k}(\mathbf{y})]_{i,j} = \sum_{f_l + I \in \mathcal{B}_{2k}} \lambda_{i,j}^l y_l$

Computation of $\mathrm{TH}_k\,(I)$

$$y_0 \qquad\qquad y_1 \qquad y_2$$

$$\mathbf{X}_{\mathcal{B}_1} = \left[\begin{array}{ccccc} \boxed{1} & \boxed{X_{11}} & \boxed{X_{12}} & \boxed{X_{21}} & \boxed{X_{22}} \\ & -X_{12}^2 - X_{21}^2 - X_{22}^2 + \boxed{1} & \boxed{X_{11}X_{12}} & \boxed{X_{11}X_{21}} & X_{11}X_{22} \\ & & X_{12}^2 & X_{11}X_{22} & X_{12}X_{22} \\ & & & X_{21}^2 & X_{21}X_{22} \\ & & & & X_{22}^2 \end{array}\right]$$

1. $\left[\mathbf{X}_{\mathcal{B}_k}\right]_{i,j} = f_i f_j + I = \sum_{f_l + I \in \mathcal{B}_{2k}} \lambda_{i,j}^{l} (f_l + I)$
2. $\left[\mathbf{M}_{\mathcal{B}_k}\,(\mathbf{y})\right]_{i,j} = \sum_{f_l + I \in \mathcal{B}_{2k}} \lambda_{i,j}^{l} y_l$

Computation of $\mathrm{TH}_k\,(I)$

$$\mathbf{X}_{\mathcal{B}_1} = \begin{bmatrix} 1 & X_{11} & X_{12} & X_{21} & X_{22} \\ & -X_{12}^2 - X_{21}^2 - X_{22}^2 + 1 & X_{11}X_{12} & X_{11}X_{21} & X_{11}X_{22} \\ & & X_{12}^2 & X_{11}X_{22} & X_{12}X_{22} \\ & & & X_{21}^2 & X_{21}X_{22} \\ & & & & X_{22}^2 \end{bmatrix}$$

$y_0 \quad y_1 \quad y_2 \quad y_3$

1. $[\mathbf{X}_{\mathcal{B}_k}]_{i,j} = f_i f_j + I = \sum_{f_l + I \in \mathcal{B}_{2k}} \lambda_{i,j}^l (f_l + I)$
2. $[\mathbf{M}_{\mathcal{B}_k}(\mathbf{y})]_{i,j} = \sum_{f_l + I \in \mathcal{B}_{2k}} \lambda_{i,j}^l y_l$

Computation of $\mathrm{TH}_k\left(I\right)$

$y_0 \quad y_1 \quad y_2 \quad y_3 \quad y_4$

$$\mathbf{X}_{\mathcal{B}_1} = \begin{bmatrix} 1 & X_{11} & X_{12} & X_{21} & X_{22} \\ & -X_{12}^2 - X_{21}^2 - X_{22}^2 + 1 & X_{11}X_{12} & X_{11}X_{21} & X_{11}X_{22} \\ & & X_{12}^2 & X_{11}X_{22} & X_{12}X_{22} \\ & & & X_{21}^2 & X_{21}X_{22} \\ & & & & X_{22}^2 \end{bmatrix}$$

1. $\left[\mathbf{X}_{\mathcal{B}_k}\right]_{i,j} = f_i f_j + I = \sum_{f_l + I \in \mathcal{B}_{2k}} \lambda_{i,j}^{l} \left(f_l + I\right)$
2. $\left[\mathbf{M}_{\mathcal{B}_k}\left(\mathbf{X}, \mathbf{y}\right)\right]_{i,j} = \sum_{f_l + I \in \mathcal{B}_{2k}} \lambda_{i,j}^{l} y_l$
3. $Q_{\mathcal{B}_k}\left(I\right) = \pi_{\mathbf{X}}\left\{\left(\mathbf{X}, \mathbf{y}\right) \in \mathbb{R}^{\mathcal{B}_{2k}} : \mathbf{M}_{\mathcal{B}_k}\left(\mathbf{X}, \mathbf{y}\right) \succeq 0, y_0 = 1\right\}$

Computation of $\mathrm{TH}_k\left(I\right)$

y_0 y_1 y_2

$$\mathbf{X}_{\mathcal{B}_1} = \begin{bmatrix} 1 & X_{11} & X_{12} & X_{21} & X_{22} \\ & -X_{12}^2 - X_{21}^2 - X_{22}^2 + 1 & X_{11}X_{12} & X_{11}X_{21} & X_{11}X_{22} \\ & & X_{12}^2 & X_{11}X_{22} & X_{12}X_{22} \\ & & & X_{21}^2 & X_{21}X_{22} \\ & & & & X_{22}^2 \end{bmatrix}$$

y_4 y_3

1. $\left[\mathbf{X}_{\mathcal{B}_k}\right]_{i,j} = f_i f_j + I = \sum_{f_l + I \in \mathcal{B}_{2k}} \lambda_{i,j}^l (f_l + I)$
2. $\left[\mathbf{M}_{\mathcal{B}_k}\left(\mathbf{X}, \mathbf{y}\right)\right]_{i,j} = \sum_{f_l + I \in \mathcal{B}_{2k}} \lambda_{i,j}^l y_l$
3. $Q_{\mathcal{B}_k}\left(I\right) = \pi_{\mathbf{X}}\left\{(\mathbf{X}, \mathbf{y}) \in \mathbb{R}^{\mathcal{B}_{2k}} : \mathbf{M}_{\mathcal{B}_k}\left(\mathbf{X}, \mathbf{y}\right) \succeq 0, y_0 = 1\right\}$

Theorem (Gouveia et al., 2010)

$\mathrm{TH}_k\left(I\right) = \overline{Q_{\mathcal{B}_k}\left(I\right)}$

Computation of $\mathrm{TH}_k(I)$

y_0 y_1 y_2

$$\mathbf{X}_{\mathcal{B}_1} = \begin{bmatrix} 1 & X_{11} & X_{12} & X_{21} & X_{22} \\ & -X_{12}^2 - X_{21}^2 - X_{22}^2 + 1 & X_{11}X_{12} & X_{11}X_{21} & X_{11}X_{22} \\ & & X_{12}^2 & X_{11}X_{22} & X_{12}X_{22} \\ & & & X_{21}^2 & X_{21}X_{22} \\ & & & & X_{22}^2 \end{bmatrix}$$

y_4 y_3

1. $[\mathbf{X}_{\mathcal{B}_k}]_{i,j} = f_i f_j + I = \sum_{f_l + I \in \mathcal{B}_{2k}} \lambda_{i,j}^l (f_l + I)$
2. $[\mathbf{M}_{\mathcal{B}_k}(\mathbf{X}, \mathbf{y})]_{i,j} = \sum_{f_l + I \in \mathcal{B}_{2k}} \lambda_{i,j}^l y_l$
3. $Q_{\mathcal{B}_k}(I) = \pi_{\mathbf{X}} \left\{ (\mathbf{X}, \mathbf{y}) \in \mathbb{R}^{\mathcal{B}_{2k}} : \mathbf{M}_{\mathcal{B}_k}(\mathbf{X}, \mathbf{y}) \succeq 0, y_0 = 1 \right\}$

Theorem (Gouveia et al., 2010)

$\mathrm{TH}_k(I) = \overline{Q_{\mathcal{B}_k}(I)}$

4. $B_{\|\cdot\|_{\theta_1}} = \mathrm{TH}_1(I) \Rightarrow \|\mathbf{X}\|_{\theta_1} = \left\{ \inf_t t \text{ s.t. } \mathbf{X} \in t\,\mathrm{TH}_1(I) \right\}$

Semidefinite program

- Given the moment matrix $\mathbf{M}_{\mathcal{B}_k}(\mathbf{X}, \mathbf{y})$ computing the θ_k-norm of a given matrix $\mathbf{X}$ is equivalent to the semidefinite program

$$\min_{\mathbf{y}} t \quad \text{subject to} \quad \mathbf{M}_{\mathcal{B}_k}(\mathbf{X}, \mathbf{y}) \succcurlyeq 0, y_0 = t.$$

- Recovery problem: the θ_k-norm minimization is equivalent to the semidefinite program

$$\min_{\mathbf{y}, \mathbf{Z}} t \quad \text{subject to } \mathbf{M}_{\mathcal{B}_k}(\mathbf{Z}, \mathbf{y}) \succeq 0, \ y_0 = t \quad \text{and} \quad \mathbf{\Phi}(\mathbf{Z}) = \mathbf{b}.$$

- A semidefinite program for computing the norm of a matrix $\mathbf{X} \in \mathbb{R}^{2\times 2}$:

$$\inf_{t,\mathbf{y}} t \quad \text{s.t.} \quad \underbrace{\begin{bmatrix} t & X_{11} & X_{12} & X_{21} & X_{22} \\ X_{11} & -y_4 - y_6 - y_8 + t & y_1 & y_2 & y_3 \\ X_{12} & y_1 & y_4 & y_3 & y_5 \\ X_{21} & y_2 & y_3 & y_6 & y_7 \\ X_{22} & y_3 & y_5 & y_7 & y_8 \end{bmatrix}}_{\text{Moment matrix}} \succeq 0$$

- $3 \times 3 \times 3$ - minimization over 216 variables (0.3705 s)
- $4 \times 4 \times 4$ - minimization over 1000 variables (7.2818 s)
- $5 \times 5 \times 5$ - minimization over 3375 variables (138.904 s)
- $6 \times 6 \times 6$ - minimization over 9261 variables

Matrix case

- $\nu_{\mathbb{R}}(I_M) = \{\mathbf{X} \in \mathbb{R}^{n_1 \times n_2} : \operatorname{rank}(\mathbf{X}) = 1, \|\mathbf{X}\|_F = 1\}$.

$$I_M = \left\{ \sum_{i<k} \sum_{j<l} g_{ijkl} \cdot (X_{il}X_{kj} - X_{ij}X_{kl}) + h \cdot \left(\sum_{i=1}^{n_1} \sum_{j=1}^{n_2} X_{ij}^2 - 1 \right) : g_{ijkl}, h \in \mathbb{R}[X_{11}, \ldots, X_{n_1 n_2}] \right\}$$

Theorem (Rauhut, S., 2015)

For the ideal I_M, $B_{\|\cdot\|_*} = B_{\|\cdot\|_{\theta_1}}$.

Basis for the ideal I

First suggestion for defining an ideal I [1]

- A rank one tensor $\mathbf{X}$: $\mathbf{X} = \mathbf{u} \otimes \mathbf{v} \otimes \mathbf{w}$
- $\nu_{\mathbb{R}}(I) = \{(\mathbf{X}, \mathbf{u}, \mathbf{v}, \mathbf{w}) : f(\mathbf{X}, \mathbf{u}, \mathbf{v}, \mathbf{w}) = 0, \text{ for all } f \in I\}$
- $\{$rank one, Frobenius norm one tensors$\}$ $= \{\mathbf{X} : (\mathbf{X}, \mathbf{u}, \mathbf{v}, \mathbf{w}) \in \nu_{\mathbb{R}}(I)\}$

$$B_1 = \Big\{ g^{ijk} = X_{ijk} - u_i v_j w_k, i \in [n_1], j \in [n_2], k \in [n_3], \\ h_1 = \sum_i u_i^2 - 1, h_2 = \sum_j v_j^2 - 1, h_3 = \sum_k w_k^2 - 1, \\ g^{ijk}, h_1, h_2, h_3 \in \mathbb{R}[\mathbf{X}, \mathbf{u}, \mathbf{v}, \mathbf{w}] \Big\}$$

[1] V. Chandrasekaran, B. Recht, P. A. Parrilo, A. S. Willsky: *The Convex Geometry of Linear Inverse Problems*, 2012.

Basis for the ideal II

Our suggestion for defining an ideal I

- $\nu_{\mathbb{R}}(I) = \{\mathbf{X} : f(\mathbf{X}) = 0, \text{ for all } f \in I\}$
- $\{\text{rank one, Frobenius norm one tensors}\} = \{\mathbf{X} : \mathbf{X} \in \nu_{\mathbb{R}}(I)\}$

$$\begin{aligned} B_2 = \Big\{ & f_1^{ijk\hat{i}\hat{j}\hat{k}} = -X_{ijk}X_{\hat{i}\hat{j}\hat{k}} + X_{ij\hat{k}}X_{\hat{i}\hat{j}k}, i < \hat{i}, j \leq \hat{j}, k < \hat{k}, \\ & f_2^{ijk\hat{i}\hat{j}\hat{k}} = -X_{ijk}X_{\hat{i}\hat{j}\hat{k}} + X_{i\hat{j}k}X_{\hat{i}j\hat{k}}, i \leq \hat{i}, j < \hat{j}, k < \hat{k}, \\ & f_3^{ijk\hat{i}\hat{j}\hat{k}} = -X_{ijk}X_{\hat{i}\hat{j}\hat{k}} + X_{i\hat{j}\hat{k}}X_{\hat{i}jk}, i < \hat{i}, j < \hat{j}, k \leq \hat{k}, \\ & g = \sum_{i,j,k} X_{ijk}^2 - 1, \ f_1^{ijk\hat{i}\hat{j}\hat{k}}, f_2^{ijk\hat{i}\hat{j}\hat{k}}, f_3^{ijk\hat{i}\hat{j}\hat{k}}, g \in \mathbb{R}[\mathbf{X}] \Big\} \end{aligned}$$

	$\mathbf{X} \in \mathbb{R}^{2\times2\times2}$	$\|\mathbf{X}^{\{1\}}\|_*$	$\|\mathbf{X}^{\{2\}}\|_*$	$\|\mathbf{X}^{\{3\}}\|_*$	$\|\mathbf{X}\|_{\theta_1}$
1	$\left[\begin{array}{cc\|cc} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{array}\right]$	2	2	2	2
2	$\left[\begin{array}{cc\|cc} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{array}\right]$	2	2	$\sqrt{2}$	2
3	$\left[\begin{array}{cc\|cc} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{array}\right]$	2	$\sqrt{2}$	2	2
4	$\left[\begin{array}{cc\|cc} 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{array}\right]$	$\sqrt{2}$	2	2	2
5	$\left[\begin{array}{cc\|cc} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \end{array}\right]$	$\sqrt{2}+1$	$\sqrt{2}+1$	$\sqrt{2}+1$	3

Table: Tensors $\mathbf{X} \in \mathbb{R}^{2\times2\times2}$ are represented as $\mathbf{X} = [\mathbf{X}(:,:,1) \,|\, \mathbf{X}(:,:,2)]$.

$$\mathbf{X}(i_1, i_2, i_3) = \mathbf{X}^{\{1\}}(i_1, (i_2, i_3)) = \mathbf{X}^{\{2\}}(i_2, (i_1, i_3)) = \mathbf{X}^{\{3\}}(i_3, (i_1, i_2))$$

Recovery of order-3 tensors

- $\mathbf{X}^* = \arg\min_{\mathbf{Z}:\mathbf{\Phi}(\mathbf{Z})=\mathbf{\Phi}(\mathbf{X})} \|\mathbf{Z}\|_{\theta_1}$
- $\mathbf{\Phi} : \mathbb{R}^{n_1 \times n_2 \times n_3} \to \mathbb{R}^m$ Gaussian, $\mathbf{X} \in \mathbb{R}^{n_1 \times n_2 \times n_3}$
- Tensor $\mathbf{X}$ is recovered if $\left|(\mathbf{X} - \mathbf{X}^*)(i,j,k)\right| < 10^{-6}$, $\forall\ i,j,k$.
- Fixed dim, rank and m: 200 trials
- m_{max} maximal m s.t. recovery: 0/200
- m_{min} minimal m s.t. recovery: 200/200

	rank	$m_{\max}$	$m_{\min}$	deg. of freedom
$2 \times 2 \times 3$	1	4	12	12
$3 \times 3 \times 3$	1	7	21	27
$3 \times 4 \times 5$	1	10	31	60
$4 \times 4 \times 4$	1	12	34	64
$4 \times 5 \times 6$	1	18	42	120
$5 \times 5 \times 5$	1	18	43	125
$3 \times 4 \times 5$	2	38	47	60
$4 \times 4 \times 4$	2	31	51	64
$4 \times 5 \times 6$	2	41	85	120

	$\mathbf{X} \in \mathbb{R}^{2\times2\times2\times2}$	$\lVert\mathbf{X}^{\{1\}}\rVert_*$, $\lVert\mathbf{X}^{\{2\}}\rVert_*$, $\lVert\mathbf{X}^{\{3\}}\rVert_*$, $\lVert\mathbf{X}^{\{4\}}\rVert_*$	$\lVert\mathbf{X}\rVert_{\theta_1}$, $\lVert\mathbf{X}\rVert_{u,\theta_1}$
1	$\mathbf{X}(:,:,:,1) = \left[\begin{array}{cc\|cc} 1 & 0 & 0 & 2 \\ 0 & 0 & 0 & 0 \end{array}\right]$, $\mathbf{X}(:,:,:,2) = \left[\begin{array}{cc\|cc} 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 \end{array}\right]$	$1+\sqrt{6}$, $2+\sqrt{3}$, $\sqrt{2}+\sqrt{5}$, $\sqrt{2}+\sqrt{5}$	5 , 4.2361
2	$\mathbf{X}(:,:,:,1) = \left[\begin{array}{cc\|cc} 1 & 0 & 0 & 2 \\ 0 & 0 & 0 & 0 \end{array}\right]$, $\mathbf{X}(:,:,:,2) = \left[\begin{array}{cc\|cc} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{array}\right]$	$1+\sqrt{6}$, $\sqrt{2}+\sqrt{5}$, $2+\sqrt{3}$, $\sqrt{2}+\sqrt{5}$	5 , 4.2361
3	$\mathbf{X}(:,:,:,1) = \left[\begin{array}{cc\|cc} 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{array}\right]$, $\mathbf{X}(:,:,:,2) = \left[\begin{array}{cc\|cc} 0 & 1 & 0 & 0 \\ 2 & 0 & 0 & 0 \end{array}\right]$	$2+\sqrt{3}$, $\sqrt{2}+\sqrt{5}$, $1+\sqrt{6}$, $\sqrt{2}+\sqrt{5}$	5 , 4.2361
4	$\mathbf{X}(:,:,:,1) = \left[\begin{array}{cc\|cc} 1 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{array}\right]$, $\mathbf{X}(:,:,:,2) = \left[\begin{array}{cc\|cc} 0 & 1 & 0 & 0 \\ 2 & 0 & 0 & 0 \end{array}\right]$	$\sqrt{3}+\sqrt{5}$, $\sqrt{2}+\sqrt{6}$, $\sqrt{2}+\sqrt{6}$, $\sqrt{3}+\sqrt{5}$	6 , 4.6503
5	$\mathbf{X}(:,:,:,1) = \left[\begin{array}{cc\|cc} 1 & 0 & 0 & 2 \\ 0 & 0 & 1 & 0 \end{array}\right]$, $\mathbf{X}(:,:,:,2) = \left[\begin{array}{cc\|cc} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{array}\right]$	$\sqrt{2}+\sqrt{6}$, $\sqrt{3}+\sqrt{5}$, $\sqrt{3}+\sqrt{5}$, $\sqrt{2}+\sqrt{6}$	6 , 4.6503

References

D. A. Cox , J. Little, D. O'Shea.
Ideals, Varieties and Algorithms: An Introduction to Computational Algebraic Geometry and Commutative Algebra.
Springer, 1991.

V. Chandrasekaran, B. Recht, P. A. Parrilo, A. S. Willsky.
The Convex Geometry of Linear Inverse Problems.
Foundations of Computational Mathematics, 12(6):805–849, 2012.

J. Gouveia, P. A. Parrilo, R. R. Thomas.
Theta Bodies for Polynomial Ideals.
SIAM Journal on Optimization, 20(4):2097–2118, 2010.

H. Rauhut, Ž. Stojanac.
Tensor theta norms and low rank recovery.
arXiv:1505.05175, 2015.

Thanks for your attention!