An Extension of Sums of Squares Relaxations to Polynomial Optimization Problems over Symmetric Cones

Masakazu Kojima† and Masakazu Muramatsu⋆

† Tokyo Institute of Technology

⋆ University of Electro-Communications

Polynomial Optimization Problem (POP)

$$\begin{cases} \text{min} & a(\boldsymbol{x}) \\ \text{s. t.} & b(\boldsymbol{x}) \geq 0 \end{cases}$$

- $a : \mathbb{R}^n \to \mathbb{R}$: Polynomial
- $b : \mathbb{R}^n \to \mathbb{R}^m$: Vector of Polynomials

$$b(\boldsymbol{x}) \geq 0 \Leftrightarrow b_j(\boldsymbol{x}) \geq 0 \ (j = 1, \ldots, m)$$

POP over Various Cones

Polynomial SDP

$$\min\ x - 3y^3 \quad \text{s. t.} \begin{pmatrix} x^2 + y & 2xy \\ 2xy & x - y^3 + 1 \end{pmatrix} \in \mathcal{S}_+^{2\times 2}$$

$\mathcal{S}_+^{2\times 2}$: 2×2 real symmetric PSD matrices.

POP over Various Cones

Polynomial SDP

$$\min\ x - 3y^3 \quad \text{s. t.} \begin{pmatrix} x^2 + y & 2xy \\ 2xy & x - y^3 + 1 \end{pmatrix} \in \mathcal{S}_+^{2\times 2}$$

Polynomial SOCP

$$\min\ x + 2x^3 - 3y^4 \quad \text{s. t.} \begin{pmatrix} x + y^2 \\ 3x - x^2 - y^4 \\ -x + y + y^3 \end{pmatrix} \in SOC(3)$$

$SOC(k)$: the k dimensional second-order cone.

Symmetric Cones

- **A Generalization of Positive Orthant**

Symmetric Cones

- **A Generalization of Positive Orthant**
- **A Unified Approach**

Symmetric Cones

- **A Generalization of Positive Orthant**
- **A Unified Approach**
  - Positive Orthant $\Rightarrow$ LP

Symmetric Cones

- **A Generalization of Positive Orthant**
- **A Unified Approach**
  - Positive Orthant $\Rightarrow$ LP
  - $\mathcal{S}^{n\times n}$: $n \times n$ symmetric positive semidefinite matrices $\Rightarrow$ SDP

Symmetric Cones

- **A Generalization of Positive Orthant**
- **A Unified Approach**
  - Positive Orthant $\Rightarrow$ LP
  - $\mathcal{S}^{n\times n}$: $n \times n$ symmetric positive semidefinite matrices $\Rightarrow$ SDP
  - Second-order cones

$$\left\{ \boldsymbol{x} \in \mathbb{R}^{1+n} \mid x_0 \geq \sqrt{\sum_{j=1}^{n} x_j^2} \right\} \quad \Rightarrow \text{SOCP}$$

POP over Symmetric Cones

$$\langle POP \rangle \begin{cases} \text{min} & a(\boldsymbol{x}) \\ \text{s. t.} & b(\boldsymbol{x}) \in \mathcal{E}_+ \end{cases}$$

- $a : \mathbb{R}^n \to \mathbb{R}$, Polynomial
- $b : \mathbb{R}^n \to \mathcal{E}$, $\mathcal{E}$-valued Polynomial
- $\mathcal{E}$: Euclidean Jordan Algebra
- $\mathcal{E}_+$: Symmetric Cone associated with $\mathcal{E}$.

POP over Symmetric Cones

$$\langle POP \rangle \begin{cases} \text{min} & a(\boldsymbol{x}) \\ \text{s. t.} & b(\boldsymbol{x}) \in \mathcal{E}_+ \end{cases}$$

- $a : \mathbb{R}^n \to \mathbb{R}$, Polynomial
- $b : \mathbb{R}^n \to \mathcal{E}$, $\mathcal{E}$-valued Polynomial
- $\mathcal{E}$: Euclidean Jordan Algebra
- $\mathcal{E}_+$: Symmetric Cone associated with $\mathcal{E}$.

$\langle POP \rangle$ includes the usual POP, polynomial SDP, and polynomial SOCP.

Contents of this talk

- Euclidean Jordan algebra (EJA) and Symmetric Cones
- Polynomials over EJAs
- Overview of our Results
- SOS relaxation and a generalization of Putinar's lemma
- Relationship between SOS and SDP
- Concluding Remarks

Euclidean Jordan Algebra (EJA) $\mathcal{E}$

Jordan Algebra : A finite dimensional real vector space having a multiplication $\circ : \mathcal{E} \times \mathcal{E} \rightarrow \mathcal{E}$

(J1) $\boldsymbol{a} \circ \boldsymbol{b} = \boldsymbol{b} \circ \boldsymbol{a}$ for any $\boldsymbol{a}, \boldsymbol{b} \in \mathcal{E}$.

(J2) $(\boldsymbol{a}^2 \circ \boldsymbol{b}) \circ \boldsymbol{a} = \boldsymbol{a}^2 \circ (\boldsymbol{b} \circ \boldsymbol{a})$ for any $\boldsymbol{a}, \boldsymbol{b} \in \mathcal{E}$.

(In general, $(\boldsymbol{a} \circ \boldsymbol{b}) \circ \boldsymbol{c} \neq \boldsymbol{a} \circ (\boldsymbol{b} \circ \boldsymbol{c})$.)

Euclidean Jordan Algebra (EJA) $\mathcal{E}$

Jordan Algebra : A finite dimensional real vector space having a multiplication $\circ : \mathcal{E} \times \mathcal{E} \to \mathcal{E}$

(J1) $\boldsymbol{a} \circ \boldsymbol{b} = \boldsymbol{b} \circ \boldsymbol{a}$ for any $\boldsymbol{a}, \boldsymbol{b} \in \mathcal{E}$.

(J2) $(\boldsymbol{a}^2 \circ \boldsymbol{b}) \circ \boldsymbol{a} = \boldsymbol{a}^2 \circ (\boldsymbol{b} \circ \boldsymbol{a})$ for any $\boldsymbol{a}, \boldsymbol{b} \in \mathcal{E}$.

Euclidean Jordan Algebra : An associative inner product $\bullet$ is defined on $\mathcal{E}$.

$$(\boldsymbol{a} \circ \boldsymbol{b}) \bullet \boldsymbol{c} = \boldsymbol{a} \bullet (\boldsymbol{b} \circ \boldsymbol{c})$$

Euclidean Jordan Algebra (EJA) $\mathcal{E}$

Jordan Algebra : A finite dimensional real vector space having a multiplication $\circ : \mathcal{E} \times \mathcal{E} \to \mathcal{E}$

(J1) $\boldsymbol{a} \circ \boldsymbol{b} = \boldsymbol{b} \circ \boldsymbol{a}$ for any $\boldsymbol{a}, \boldsymbol{b} \in \mathcal{E}$.

(J2) $(\boldsymbol{a}^2 \circ \boldsymbol{b}) \circ \boldsymbol{a} = \boldsymbol{a}^2 \circ (\boldsymbol{b} \circ \boldsymbol{a})$ for any $\boldsymbol{a}, \boldsymbol{b} \in \mathcal{E}$.

Euclidean Jordan Algebra : An associative inner product $\bullet$ is defined on $\mathcal{E}$.

$$(\boldsymbol{a} \circ \boldsymbol{b}) \bullet \boldsymbol{c} = \boldsymbol{a} \bullet (\boldsymbol{b} \circ \boldsymbol{c})$$

Symmetric cone $\mathcal{E}_+$ associated with $\mathcal{E}$

- $\mathcal{E}_+ = \{ \boldsymbol{a}^2 \mid \boldsymbol{a} \in \mathcal{E} \}$.

Examples of EJA

$\mathbb{R}^n$:

- $(\boldsymbol{a} \circ \boldsymbol{b})_i = a_i b_i$ $(i = 1, \ldots, n)$.
- $\boldsymbol{a} \bullet \boldsymbol{b} = \boldsymbol{a}^T \boldsymbol{b}$
- $\mathbb{R}_+ = \{ \boldsymbol{a}^2 \in \mathbb{R}^n \,|\, \boldsymbol{a} \in \mathbb{R}^n \} = \{ \boldsymbol{a} \,|\, a_i \geq 0 (i = 1, \ldots, n) \}$.

Examples of EJA

$\mathbb{R}^n$:

- $(\boldsymbol{a} \circ \boldsymbol{b})_i = a_i b_i$ $(i = 1, \ldots, n)$.
- $\boldsymbol{a} \bullet \boldsymbol{b} = \boldsymbol{a}^T \boldsymbol{b}$
- $\mathbb{R}_+ = \{ \boldsymbol{a}^2 \in \mathbb{R}^n \mid \boldsymbol{a} \in \mathbb{R}^n \} = \{ \boldsymbol{a} \mid a_i \geq 0 (i = 1, \ldots, n) \}$.

$\mathcal{S}^{N \times N}$, the set of $N \times N$ real symmetric matrices :

- $X \circ Y = \frac{XY + YX}{2}$
- $X \bullet Y = \text{trace}(XY)$
- $\mathcal{S}_+^{N \times N}$: the set of $N \times N$ PSD matrices.

$\mathcal{E}$-valued Polynomials

$\mathcal{E}$-valued polynomial $f : \mathbb{R}^N \to \mathcal{E}$

$$f(\boldsymbol{x}) = \sum_{\boldsymbol{\alpha} \in \mathcal{G}} \boldsymbol{f}_{\boldsymbol{\alpha}} \boldsymbol{x}^{\boldsymbol{\alpha}}$$

- $\mathcal{G}$, a finite subset of $\mathbb{Z}^n$
- $\boldsymbol{f}_{\boldsymbol{\alpha}} \in \mathcal{E}$ for every $\boldsymbol{\alpha} \in \mathcal{G}$,
- $\boldsymbol{x}^{\boldsymbol{\alpha}} = x_1^{\alpha_1} x_2^{\alpha_2} \dots x_n^{\alpha_n}$.

$\mathcal{E}[\boldsymbol{x}]$: The set of $\mathcal{E}$-valued polynomials.

$\mathcal{E}$-valued Polynomials and EJA

- EJA $\Leftrightarrow$ $\circ$ and $\bullet$

$\mathcal{E}$-valued Polynomials and EJA

- EJA $\Leftrightarrow$ $\circ$ and $\bullet$
- For $f, g \in \mathcal{E}[\boldsymbol{x}]$,

$$(f \circ g)(\boldsymbol{x}) = f(\boldsymbol{x}) \circ g(\boldsymbol{x})$$
$$(f \bullet g)(\boldsymbol{x}) = f(\boldsymbol{x}) \bullet g(\boldsymbol{x})$$

($f \circ g \in \mathcal{E}[\boldsymbol{x}]$, $f \bullet g \in \mathbb{R}[\boldsymbol{x}]$)

$\mathcal{E}$-valued Polynomials and EJA

- EJA $\Leftrightarrow$ $\circ$ and $\bullet$
- For $f, g \in \mathcal{E}[\boldsymbol{x}]$,

$$(f \circ g)(\boldsymbol{x}) = f(\boldsymbol{x}) \circ g(\boldsymbol{x})$$
$$(f \bullet g)(\boldsymbol{x}) = f(\boldsymbol{x}) \bullet g(\boldsymbol{x})$$

- $\mathcal{E}[\boldsymbol{x}]^2$: Sums of Squares of $\mathcal{E}$-valued polynomials:

$$\mathcal{E}[\boldsymbol{x}]^2 = \left\{ \sum_{i=1}^{q} f_i \circ f_i \,\middle|\, q > 0, f_i \in \mathcal{E}[\boldsymbol{x}] \ (i = 1, \dots, q) \right\}$$

A Primal and Dual pair

$$\langle POP \rangle \begin{cases} \text{min} & a(\boldsymbol{x}) \\ \text{s. t.} & \boldsymbol{x} \in K = \{\, \boldsymbol{x} \in \mathbb{R}^n \mid b(\boldsymbol{x}) \in \mathcal{E}_+ \,\}\,. \end{cases}$$

- $a \in \mathbb{R}[\boldsymbol{x}]$: $\mathbb{R}$-valued poloynomial
- $b \in \mathcal{E}[\boldsymbol{x}]$: $\mathcal{E}$-valued poloynomial
- $\mathcal{E}_+$: the symmetric cone associated with $\mathcal{E}$.

A Primal and Dual pair

$$\langle POP \rangle \begin{cases} \text{min} & a(\boldsymbol{x}) \\ \text{s. t.} & \boldsymbol{x} \in K = \{\, \boldsymbol{x} \in \mathbb{R}^n \,|\, b(\boldsymbol{x}) \in \mathcal{E}_+ \,\}\,. \end{cases}$$

$\Updownarrow$ dual

$$\langle D \rangle \begin{cases} \text{max} & \zeta \\ \text{s. t.} & a(\boldsymbol{x}) - \zeta \geq 0 \ (\forall \boldsymbol{x} \in K). \end{cases}$$

$$\text{optval}\langle POP \rangle = \text{optval}\langle D \rangle$$

Overview of Our Results

$$
\begin{array}{rcl}
\langle POP \rangle & \stackrel{\text{dual}}{\longleftrightarrow} & \langle D \rangle \\
\text{SDP} \updownarrow & & \downarrow \text{SOS Relax.} \\
\langle SDP/POP \rangle_r & & \langle SOS/D \rangle_r \\
\text{Linear Relax.} \downarrow & & \updownarrow \text{SDP} \\
\langle LR/SDP/POP \rangle_r & \stackrel{\text{dual}}{\longleftrightarrow} & \langle SDP/SOS/D \rangle_r
\end{array}
$$

- optval$\langle SOS/D \rangle_r \to$ optval$\langle D \rangle$ as $r \to \infty$.
- optval$\langle LR/SDP/POP \rangle_r \to$ optval$\langle POP \rangle$ as $r \to \infty$.

Overview of Our Results

$$\begin{array}{ccc} \langle POP \rangle & \overset{\text{dual}}{\longleftrightarrow} & \langle D \rangle \\ \text{SDP} \updownarrow & & \downarrow \text{SOS Relax.} \\ \langle SDP/POP \rangle_r & & \langle SOS/D \rangle_r \\ \text{Linear Relax.} \downarrow & & \updownarrow \text{SDP} \\ \langle LR/SDP/POP \rangle_r & \overset{\text{dual}}{\longleftrightarrow} & \langle SDP/SOS/D \rangle_r \end{array}$$

$\langle POP \rangle$ and $\langle D \rangle$ are POP over Symmteric Cones!

An SOS relaxation of $\langle D \rangle$

$\langle D \rangle$ max ζ s. t. $a(\boldsymbol{x}) - \zeta \geq 0 \ (\forall \boldsymbol{x} \in K)$

An SOS relaxation of $\langle D \rangle$

$$\langle D \rangle \ \text{max} \ \ \zeta \ \text{s. t.} \ \ a(\boldsymbol{x}) - \zeta \geq 0 \ (\forall \boldsymbol{x} \in K)$$

$\downarrow$ SOS Relax.

$$\langle SOS/D \rangle_r \ \text{max} \ \ \zeta \ \text{s. t.} \ \ a - \zeta \in \mathbb{R}[\boldsymbol{x}]^2 + b \bullet \mathcal{E}[\boldsymbol{x}]_r^2$$

- $\mathcal{E}[\boldsymbol{x}]_r$: $\mathcal{E}$-valued Polynomials, degree $\leq r$.
- $\mathcal{E}[\boldsymbol{x}]_r^2$: Sums of Squares of $\mathcal{E}[\boldsymbol{x}]_r$.
- $\mathbb{R}[\boldsymbol{x}]^2$: SOS of $\mathbb{R}$-valued Polynomials.

An SOS relaxation of $\langle D \rangle$

$$\langle D \rangle \text{ max } \zeta \text{ s. t. } a(\boldsymbol{x}) - \zeta \geq 0 \ (\forall \boldsymbol{x} \in K)$$

$\downarrow$ SOS Relax.

$$\langle SOS/D \rangle_r \text{ max } \zeta \text{ s. t. } a - \zeta \in \mathbb{R}[\boldsymbol{x}]^2 + b \bullet \mathcal{E}[\boldsymbol{x}]_r^2$$

Theorem 1: optval$\langle SOS/D \rangle_r \nearrow$ optval$\langle D \rangle$ as $r \to \infty$.

Idea of the Proof

ζ^*: optval$\langle D \rangle$, $\epsilon > 0$: fixed.

$$a(\boldsymbol{x}) - (\zeta^* - \epsilon) > 0 \ (\forall \boldsymbol{x} \in K)$$

Idea of the Proof

ζ^*: optval$\langle D \rangle$, $\epsilon > 0$: fixed.

$$a(\boldsymbol{x}) - (\zeta^* - \epsilon) > 0 \ (\forall \boldsymbol{x} \in K)$$

$$\Downarrow$$

A Generalized Putinar's lemma

$$\exists r > 0, a - (\zeta^* - \epsilon) \in \mathbb{R}[\boldsymbol{x}]^2 + b \bullet \mathcal{E}[\boldsymbol{x}]_r^2$$

Idea of the Proof

ζ^*: optval$\langle D \rangle$, $\epsilon > 0$: fixed.

$$a(\boldsymbol{x}) - (\zeta^* - \epsilon) > 0 \ (\forall \boldsymbol{x} \in K)$$

$$\Downarrow$$

A Generalized Putinar's lemma

$$\exists r > 0, a - (\zeta^* - \epsilon) \in \mathbb{R}[\boldsymbol{x}]^2 + b \bullet \mathcal{E}[\boldsymbol{x}]^2_r$$

$$\Downarrow$$

$\langle SOS/D \rangle_r$ has a feasible solution with $\zeta = \zeta^* - \epsilon$

A Generalization of Putinar's Lemma

$$
\begin{aligned}
K &= \{\, \boldsymbol{x} \in \mathbb{R}^n \mid b(\boldsymbol{x}) \in \mathcal{E}_+ \,\} \\
\tilde{\mathcal{C}} &= \mathbb{R}[\boldsymbol{x}]^2 + b \bullet \mathcal{E}[\boldsymbol{x}]^2
\end{aligned}
$$

If K is compact, then

any positive polynomials on K belongs to $\tilde{\mathcal{C}}$

$$\Updownarrow$$

$\exists p \in \tilde{\mathcal{C}}$ such that $\{\, \boldsymbol{x} \mid p(\boldsymbol{x}) \in \mathcal{E}_+ \,\}$ is compact.

- Putinar's original Lemma : $\mathcal{E} = \mathbb{R}^N$.
- We use this lemma assuming the latter condition.

Overview of Our Results

$$\begin{array}{rcl}
\langle POP\rangle & \overset{\text{dual}}{\longleftrightarrow} & \langle D\rangle \\
\text{SDP} \updownarrow & & \downarrow \text{SOS Relax.} \\
\langle SDP/POP\rangle_r & & \langle SOS/D\rangle_r \\
\text{Linear Relax.} \downarrow & & \updownarrow \text{SDP} \\
\langle LR/SDP/POP\rangle_r & \overset{\text{dual}}{\longleftrightarrow} & \langle SDP/SOS/D\rangle_r
\end{array}$$

- optval$\langle SOS/D\rangle_r \to$ optval$\langle D\rangle$ as $r \to \infty$.
- optval$\langle LR/SDP/POP\rangle_r \to$ optval$\langle POP\rangle$ as $r \to \infty$.

An Equivalent SDP of SOS problem

- $\mathcal{D}$: a finite dimensional linear subspace of $\mathcal{E}[\boldsymbol{x}]$.
  (e.g. $\mathcal{D} = \mathcal{E}[\boldsymbol{x}]_r$)
- $\mathcal{A}$: a basis of $\mathcal{D}$.
- $\mathcal{D}^2$: Sums of Squares of polynomials in $\mathcal{D}$.

An Equivalent SDP of SOS problem

- $\mathcal{D}$: a finite dimensional linear subspace of $\mathcal{E}[\boldsymbol{x}]$.
  (e.g. $\mathcal{D} = \mathcal{E}[\boldsymbol{x}]_r$)
- $\mathcal{A}$: a basis of $\mathcal{D}$.
- $\mathcal{D}^2$: Sums of Squares of polynomials in $\mathcal{D}$.

Theorem 2

$$\mathcal{D}^2 = \left\{ \sum_{f \in \mathcal{A}} \sum_{g \in \mathcal{A}} V_{fg} f \circ g \,\middle|\, V \in \mathcal{S}_+^{|\mathcal{A}| \times |\mathcal{A}|} \right\}.$$

- $\mathcal{S}_+^{|\mathcal{A}| \times |\mathcal{A}|}$: the set of $|\mathcal{A}| \times |\mathcal{A}|$ PSD matrices.

Overview of Our Results

$$
\begin{array}{rcl}
\langle POP \rangle & \overset{\text{dual}}{\longleftrightarrow} & \langle D \rangle \\
\text{SDP} \updownarrow & & \downarrow \text{SOS Relax.} \\
\langle SDP/POP \rangle_r & & \langle SOS/D \rangle_r \\
\text{Linear Relax.} \downarrow & & \updownarrow \text{SDP} \\
\langle LR/SDP/POP \rangle_r & \overset{\text{dual}}{\longleftrightarrow} & \langle SDP/SOS/D \rangle_r
\end{array}
$$

- optval$\langle SOS/D \rangle_r \to$ optval$\langle D \rangle$ as $r \to \infty$.
- optval$\langle LR/SDP/POP \rangle_r \to$ optval$\langle POP \rangle$ as $r \to \infty$.

Notes

- This work is motivated by Kojima [1] which deals with polynomial SDP. Some relationships between [1] and ours are also established in the paper [2].
- In [2], a more detailed discussion is done to exploit sparsity of the polynomials to shrink the size of SDPs.

[1] M. Kojima : “Sums of Squares Relaxations of Polynomial Semidefi nite Programs”, Research Report B-397, TIT.
[2] M. Kojima and M. Muramatsu : “An Extension of Sums of Squares Relaxations to Polynomial Optimization Problems over Symmetric Cones”, Research Report B-406, TIT.