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

M. Kojima, Tokyo Institute of Technology

M. Muramatsu, The University of Electro-Communications

Semidefinite Programming and Its Application

Institute for Mathematical Sciences

National University of Singapore

January 9-13, 2006

- This talk is based on

[A] Kojima & Muramatsu, "An extension of SOS relaxations to POPs over symmetric cones", Tokyo Inst. Tech., B-406, April 2004.

[B] Kojima & Muramatsu, "A note on sparse SOS relaxations for POPs over symmetric cones", Tokyo Inst. Tech., B-421, January 2006.

⇑

⇑

- Extensions to poly. SDPs ⊂ POPs over symmetric cones

[F] Henrion & Lasserre (2004), [G] Hol & Scherer (2004), [H] Kojima (2003)

⇑

- SDP and SOS relaxations of POPs

[C] Lasserre, "Global optimization with polynomials and the problems of moments", *SIAM J. Optim.*, 11 (2001) 796–817.

[D] Parrilo, "Semidefinite programming relaxations for semialgebraic problems", *Math. Program.*, 96 (2003) 293-320.

- Convergence proof of sparse SDP and SOS relaxations of POPs

[E] Lasserre, "Convergent Semidefinite Relaxation in Polynomial Optimization with Sparsity", LAAS-CVRS (2005).

Contents

Contents

$$\textbf{POP: } \min\ a(x)\ \text{sub.to}\ x \in F \equiv \{b(x) \in \mathbb{E}_+\}$$

$a \in \mathbb{R}[x]$ (the set of real-valued polynomials in $x = (x_1, \ldots, x_n) \in \mathbb{R}^n$),

$b \in \mathbb{E}[x]$ (the set of $\mathbb{E}$-valued polynomials in $x = (x_1, \ldots, x_n) \in \mathbb{R}^n$),

$\mathbb{E}$: a finite dimensional real vector space,

$\mathbb{E}_+$: a symmetric cone embedded in $\mathbb{E}$.

Example 1: A polynomial second-order programming problem

$$\begin{aligned} \mathbb{E} &= \mathbb{R}^{1+m}, \\ \mathbb{E}_+ &= \mathbb{Q}(m) \text{ (the second-order cone in } \mathbb{R}^{1+m}) \\ &= \{(y_0, y_1) \in \mathbb{R}^{1+m} \ : y_0 \geq \|y_1\|\} \end{aligned}$$

Let $n = 2$, $x = (x_1, x_2)$, $\mathbb{E} = \mathbb{R}^{1+2}$, $\mathbb{E}_+ = \mathbb{Q}(2)$.

$$\begin{aligned} &\text{POP: } \min\ -x_1^3 + 2x_1x_2^2 \\ &\qquad \text{sub.to} \end{aligned}$$

$$(x_1^2 - x_2, 2x_1^2x_2 - x_2, x_1 + x_2) \in \mathbb{Q}(2)$$
$$\left(\text{or } x_1^2 - x_2 \geq \|(2x_1^2x_2 - x_2, x_1 + x_2)\|\right)$$

$$\textbf{POP: min } a(x) \textbf{ sub.to } x \in F \equiv \{b(x) \in \mathbb{E}_+\}$$

$a \in \mathbb{R}[x]$ (the set of real-valued polynomials in $x = (x_1, \ldots, x_n) \in \mathbb{R}^n$),

$b \in \mathbb{E}[x]$ (the set of $\mathbb{E}$-valued polynomials in $x = (x_1, \ldots, x_n) \in \mathbb{R}^n$),

$\mathbb{E}$: a finite dimensional real vector space,

$\mathbb{E}_+$: a symmetric cone embedded in $\mathbb{E}$.

Example 2: A general POP over a symmetric cone

$$\begin{aligned} \mathbb{E} &= \mathbb{R}^k \times \mathbb{S}^\ell \times \mathbb{R}^{1+m}, \\ \mathbb{E}_+ &= \mathbb{R}^k_+ \times \mathbb{S}^\ell_+ \times \mathbb{Q}(m) \end{aligned}$$

Let $n = 2$, $x = (x_1, x_2)$, $\mathbb{E} = \mathbb{R}^3 \times \mathbb{S}^2 \times \mathbb{R}^{1+2}$, $\mathbb{E}_+ = \mathbb{R}^3_+ \times \mathbb{S}^2{}_+ \times \mathbb{Q}(2)$.

$$\begin{aligned} \textbf{POP: min } & -x_1^3 + 2x_1x_2^2 \\ \textbf{sub.to } & \left(x_2 + 0.5,\ 1 - x_1^2 - x_2^2,\ -x_1^3 + x_2 \right) \in \mathbb{R}^3_+ \\ & \begin{pmatrix} 2 & -1 \\ -1 & 1 \end{pmatrix} x_1x_2^2 + \begin{pmatrix} -1 & 2 \\ 2 & 3 \end{pmatrix} x_1^2x_2 + \begin{pmatrix} 3 & 1 \\ 1 & 1 \end{pmatrix} \in \mathbb{S}^2_+ \\ & (x_1^2 - x_2, 2x_1^2x_2 - x_2, x_1 + x_2) \in \mathbb{Q}(2) \\ & (\text{or } x_1^2 - x_2 \geq \|(2x_1^2x_2 - x_2, x_1 + x_2)\|) \end{aligned}$$

$$\text{POP: min } a(x) \text{ sub.to } x \in F \equiv \{b(x) \in \mathbb{E}_+\}$$

$a \in \mathbb{R}[x]$ (the set of real-valued polynomials in $x = (x_1, \ldots, x_n) \in \mathbb{R}^n$),
$b \in \mathbb{E}[x]$ (the set of $\mathbb{E}$-valued polynomials in $x = (x_1, \ldots, x_n) \in \mathbb{R}^n$),
$\mathbb{E}$: a finite dimensional real vector space,
$\mathbb{E}_+$: a symmetric cone embedded in $\mathbb{E}$.

Importance of polynomial SOCP inequalities: Let

$f(x)$: a real valud polynomial with deg d_f in $x = (x_1, \ldots, x_n)$
$h(x)$: a $\mathbb{R}^m$-valud polynomial with deg d_h in $x = (x_1, \ldots, x_n)$

normal poly. inequalities		poly. SOCP inequalities
$\left.\begin{array}{r} f(x)^2 - h(x)^T h(x) \geq 0 \\ f(x) \geq 0 \end{array}\right\}$	$\Leftrightarrow f(x) \geq \lVert h(x) \rVert \Leftrightarrow$	$\begin{pmatrix} f(x) \\ h(x) \end{pmatrix} \in \mathbb{Q}(m)$
degree $2\max\{d_f, d_h\}$		degree $\max\{d_f, d_h\}$
$f(x) - h(x)^T h(x) \geq 0$	$\Leftrightarrow$	$\begin{pmatrix} 1 + f(x) \\ 1 - f(x) \\ h(x) \end{pmatrix} \in \mathbb{Q}(1+m)$
degree $\max\{d_f, 2d_h\}$		degree $\max\{d_f, d_h\}$

$\Rightarrow$ Applications to nonlinear least square problems

Contents

Contents

Definition. $\boldsymbol{K} \subset \mathbb{E}$ **is a symmetric cone if**

- $\boldsymbol{K}^* \equiv \{u \in \mathbb{E} \ : \ \langle u, v\rangle \geq 0 \ (\forall v \in \boldsymbol{K})\} = \boldsymbol{K}$ **(self-dual).**
- **For every pair of** u, v **of** $\mathrm{int}(\boldsymbol{K})$**, there is a linear transformation** $\boldsymbol{T} : \mathbb{E} \longrightarrow \mathbb{E}$ **such that** $\boldsymbol{T}(\boldsymbol{K}) = \boldsymbol{K}$ **and** $\boldsymbol{T}(u) = v$ **(homogeneous).**

Symmetric cones are classified into the following cones

(a) the second order cone.

$$\mathbb{Q}(m) \equiv \ \left\{ u = (u_0, u_1) : u_0 \in \mathbb{R}, u_1 \in \mathbb{R}^m, u_0 \geq \|u_1\| \right\},$$

where $\|u_1\| = \sqrt{u_1^T u_1}$**.**

(b) the set $\mathbb{S}^n_+$ **of** $n \times n$ **real, symmetric positive semidefinite matrices (including the set of nonnegative numbers when** $n = 1$**).**

(c) the set of $n \times n$ **Hermitian psd matrices with complex entries.**

(d) the set of $n \times n$ **Hermitian psd matrices with quarternion entries.**

(e) the set of 3×3 **Hermitian psd matrices with octonion entries.**

(f) any cone $\boldsymbol{K}_1 \times \boldsymbol{K}_2$ **where** $\boldsymbol{K}_1$ **and** $\boldsymbol{K}_2$ **are themselves symmetric cones.**

Definition. $K \subset \mathbb{E}$ is a symmetric cone if

- $K^* \equiv \{u \in \mathbb{E} \ : \ \langle u, v \rangle \geq 0 \ (\forall v \in K)\} = K$ (self-dual).
- For every pair of u, v of $\mathrm{int}(K)$, there is a linear transformation $T : \mathbb{E} \longrightarrow \mathbb{E}$ such that $T(K) = K$ and $T(u) = v$ (homogeneous).

Theorem. A cone K is symmetric iff it is the cone of squares of some Euclidean Jordan algebra $\circ$ in $\mathbb{E}$ (Jordan algebra characterization of symmetric cones); $K = \{u \circ u : u \in \mathbb{E}\}$.

Definition. $(\mathbb{E}, \circ)$ is a Euclidean Jordan algebra if $(u, v) \in \mathbb{E} \times \mathbb{E} \longrightarrow u \circ v \in \mathbb{E}$ is a bilinear map satisfying

(i) $u \circ v = v \circ u$, (ii) $u \circ (u^2 \circ v) = u^2 \circ (u \circ v)$ where $u^2 = u \circ u$,
(iii) $\langle u \circ v, w \rangle = \langle u, v \circ w \rangle$ for $\forall u, v, w \in \mathbb{E}$.

(a) the second order cone $\mathbb{Q}(m) \equiv \left\{u = (u_0, u_1) \in \mathbb{R}^{1+m} : u_0 \geq \|u_1\|\right\}$:
$u \circ v \equiv (u_0 v_0 + u_1^T v_1, u_0 v_1 + v_0 u_1) \Rightarrow \mathbb{Q}(m) = \left\{u \circ u : u \in \mathbb{R}^{1+m}\right\}$.

(b) the set $\mathbb{S}^\ell_+$ of $\ell \times \ell$ real, symmetric positive semidefinite matrices (including the set of positive numbers as a special case when $n = 1$).

$$X \circ Y \equiv (XY + YX)/2 \Rightarrow \mathbb{S}^\ell_+ = \{X \circ X = X^2 : X \in \mathbb{S}^n\}.$$

(b)' the nonnegative orthant $\mathbb{R}^k_+ = \prod_{i=1}^k \mathbb{S}^1_+$: $u \circ v = (u_1 v_1, \ldots, u_k v_k)$
$\Rightarrow \mathbb{R}^k_+ = \{u \circ u = (u_1^2, \ldots, u_k^2) : u \in \mathbb{R}^k\}$.

Contents

$\mathbb{E}[x]$: the set of $\mathbb{E}$-valued polynomials; $\varphi \in \mathbb{E}[x] \Leftrightarrow \varphi(x) = \sum_{\alpha \in \mathcal{F}} f_\alpha x^\alpha$.

$$\begin{aligned}
(\mathbb{E}, \circ) &: \text{ a Euclidean Jordan algebra} \\
\mathcal{F} &: \text{ a nonempty finite set of nonnegative integer vectors in } \mathbb{R}^n \\
f_\alpha &\in \mathbb{E} \ (\alpha \in \mathcal{F}) \\
x^\alpha &= x_1^{\alpha_1} x_2^{\alpha_2} \cdots x_n^{\alpha_n}, \text{ for example,} \\
& \quad \text{if } n = 3 \text{ and } \alpha = (2,0,4) \text{ then } x^{(2,0,4)} = x_1^2 x_2^0 x_3^4.
\end{aligned}$$

$$\begin{aligned}
\deg(\varphi) &= \max\{\textstyle\sum_{i=1}^n \alpha_i : \alpha \in \mathcal{F}\}, \\
\mathbb{E}[x]_r &= \{\varphi \in \mathbb{E}[x] : \deg(\varphi) \le r\}.
\end{aligned}$$

Specifically,

$\mathbb{R}[x]$ ($\mathbb{R}[x]_r$) : the set of $\mathbb{R}$-valued poly. (with deg. $\le r$)

Extension of $\circ$ to the $\mathbb{E}$-valued polynomials. Let

$$\varphi \in \mathbb{E}[x];\ \varphi(x) = \sum_{\alpha \in \mathcal{F}} f_\alpha x^\alpha \ \text{ and } \ \psi \in \mathbb{E}[x];\ \psi(x) = \sum_{\beta \in \mathcal{G}} g_\beta x^\beta,$$

then

$$\begin{aligned}
\varphi \circ \psi \in \mathbb{E}[x];\ (\varphi \circ \psi)(x) &= \left(\sum_{\alpha \in \mathcal{F}} f_\alpha x^\alpha\right) \circ \left(\sum_{\beta \in \mathcal{G}} g_\beta x^\beta\right) \\
&= \sum_{\alpha \in \mathcal{F}} \sum_{\beta \in \mathcal{G}} (f_\alpha \circ g_\beta)\, x^{\alpha+\beta}.
\end{aligned}$$

Contents

For $\forall$ linear subspace $\mathbb{D}[\boldsymbol{x}]$ of $\mathbb{E}[\boldsymbol{x}]$, let

$$\mathbb{D}[\boldsymbol{x}]^2 = \{\textstyle\sum_{i=1}^{q} \boldsymbol{\varphi}_i \circ \boldsymbol{\varphi}_i \ : \ \exists \boldsymbol{q}, \ \boldsymbol{\varphi}_i \in \boldsymbol{\mathcal{D}}\} \text{ (SOS poly. of } \mathbb{D}[\boldsymbol{x}]\text{).}$$

Thus we will use $\mathbb{E}[\boldsymbol{x}]^2$, $\mathbb{E}[\boldsymbol{x}]^2_r$, $\mathbb{R}[\boldsymbol{x}]^2_r$. Here

$\mathbb{E}[\boldsymbol{x}]$ ($\mathbb{E}[\boldsymbol{x}]_r$) **: the set of $\mathbb{E}[\boldsymbol{x}]$-valued poly. (with deg. $\leq \boldsymbol{r}$)**

$\mathbb{R}[\boldsymbol{x}]$ ($\mathbb{R}[\boldsymbol{x}]_r$) **: the set of $\mathbb{R}$-valued poly. (with deg. $\leq \boldsymbol{r}$)**

Contents

$$\textbf{POP: } \min\ a(x)\ \text{sub.to}\ b(x) \in \mathbb{E}_+$$

$a \in \mathbb{R}[x]$ (the set of real-valued polynomials in $x = (x_1, \ldots, x_n) \in \mathbb{R}^n$),

$b \in \mathbb{E}[x]$ (the set of $\mathbb{E}$-valued polynomials in $x = (x_1, \ldots, x_n) \in \mathbb{R}^n$),

$(\mathbb{E}, \circ)$: a Euclidean Jordan algebra,

$\mathbb{E}_+$: a symmetric cone embedded in $\mathbb{E}$,

$\omega_a = \lceil \deg(a)/2 \rceil$, $\omega_b = \lceil \deg(b)/2 \rceil$, $\omega_{\max} = \max\{\omega_a, \omega_b\}$.

G. Lagrangian funct.: $L(x, \varphi) = a(x) - \langle \varphi(x), b(x) \rangle$ $(\forall x \in \mathbb{R}^n, \varphi \in \mathbb{E}[x]^2)$.

$$\textbf{G.Lagrangian Dual: } \max_{\varphi \in \mathbb{E}[x]^2} \min_{x \in \mathbb{R}^n} L(x, \varphi)$$

$$\Updownarrow$$

$$\textbf{G.L. Dual: } \max \zeta \text{ sub.to } L(x, \varphi) - \zeta \geq 0\ (\forall x \in \mathbb{R}^n) \text{ and } \varphi \in \mathbb{E}[x]^2.$$

$$\text{relaxation} \Downarrow \omega \geq \omega_{\max}$$

$$\textbf{SOS relaxation: } \max \zeta \text{ sub.to } L(x, \varphi) - \zeta \in \mathbb{R}[x]^2_\omega \text{ and } \varphi \in \mathbb{E}[x]^2_{\omega - \omega_b}.$$

- An Extension of Lasserre's relaxation 2001.
- We can transform SOS relaxation to an SDP.
- We can apply an SDP' relaxation directly to POP; SDP and SDP' are dual to each other.

Contents

POP: min $a(x)$ sub.to $b(x) \in \mathbb{E}_+$, $x \in U \equiv \{x \in \mathbb{R}^n : \|x\| \le M\}$.

Let $d = \deg(b)$. Let $\epsilon > 0$. We can prove that

∃ SOS relaxation of POP; opt.val POP ≥ opt.val SOS ≥ opt.val POP$-\epsilon$

The basic idea is:

(a) Reduce POP to

P_ω: min $a_\omega(x) \equiv a(x) + \psi_\omega(x)$ sub.to $x \in U$ $(\omega = 1, 2, \ldots)$

Here $\psi_\omega \in \mathbb{R}[x]_{d+2\omega d}$ serves as a penalty function in U such that

$$x \in U \text{ and } b(x) \in \mathbb{E}_+ \Rightarrow 0 \ge \psi_\omega(x) \to 0 \text{ as } \omega \to \infty,$$
$$x \in U \text{ and } b(x) \notin \mathbb{E}_+ \Rightarrow \psi_\omega(x) \to \infty \text{ as } \omega \to \infty.$$

More specifically,

$$\psi_\omega(x) = -\langle b(x), \varphi_\omega(x)\rangle, \ \varphi_\omega(x) = (e - b(x)/\lambda_{\max})^{2\omega} \in \mathbb{E}[x]^2_\omega,$$

e denotes the identity element of $\mathbb{E}$,

$\lambda_{\max}$ denotes the max. eigenvalue of $b(x)$ over $x \in U$.

POP: min $a(x)$ sub.to $b(x) \in \mathbb{E}_+$, $x \in U \equiv \{x \in \mathbb{R}^n : \|x\| \leq M\}$.

Let $d = \deg(b)$. Let $\epsilon > 0$. We can prove that

$\exists$ **SOS relaxation** of **POP**; opt.val **POP** $\geq$ opt.val **SOS** $\geq$ opt.val **POP**$-\epsilon$

The basic idea is:

(a) Reduce **POP** to

P_ω: min $a_\omega(x) \equiv a(x) + \psi_\omega(x)$ sub.to $x \in U$ $(\omega = 1, 2, \dots)$

$\exists\omega$; opt.val **POP** $\geq$ opt.val P_ω $\geq$ opt.val **POP**$-\epsilon/2$

(b) Apply the convergence theorem by Lasserre '01 to P_ω.

$\exists$ **SOS relaxation** of P_ω; opt.val P_ω $\geq$ opt.val **SOS** $\geq$ opt.val $\mathrm{P}_\omega-\epsilon/2$

$$\Downarrow$$

opt.val.of **POP** $\geq$ opt.val.of **SOS** $\geq$ opt.val.of **POP** $-\epsilon$

Contents

$$\begin{array}{rl} \textbf{POP}: \min & \sum_{p=1}^{q} a_p(x_{N_p}) \\ \text{sub.to} & b_p(x_{N_p}) \in \mathbb{E}_{p+}, \\ & x_{N_p} \in U_p \equiv \{x_{N_p} : \|x_{N_p}\| \leq M_p\} \ (p = 1, \ldots, q). \end{array}$$

Here $N_p \subset N \equiv \{1, \ldots, n\}$ and $x_{N_p} = (x_i : i \in N_p)$;

if $N_p \equiv \{1, 4\} \subset N \equiv \{1, 2, 3, 4\}$ then $x_{N_p} = (x_1, x_4)$.

- Each a_p & each b_p involve only variables x_i $(i \in N_p)$ among x_i $(i \in N)$.
- Ball constraint $x_{N_p} \in U_p$ $(p = 1, \ldots, q)$.
- We can extend the sparse relaxation (Waki et al. 04) to POP.
- We can prove the convergence of the extension under Assumption using the same argument as in the dense case and Lasserre 05.

Assumption (Lasserre 05, Waki et.al 04 as a chordal graph structure). N_p $(p = 1, \ldots, q)$ are the "maximal" cliques of a chordal graph;

$$\forall p \in \{1, \ldots, q-1\} \ \exists r \geq p+1; N_p \cap \left(\cup_{k=p+1}^{q} N_k\right) \subset N_r$$

(the running intersection property of the max.cliques of a chordal graph)

[Lasserre 05] "Convergent semidefinite relaxation in polinomial optimization with sparseity", November 2004.

Proof is given in: M.Kojima and Muramatsu, "A note on sparse SOS relaxations for POPs over symmetric cones", B-421, January 2006.

A sparse numerical example

$$\min \sum_{i=1}^{n} a_i x_i$$

$$\text{s.t.} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + \begin{pmatrix} b_j & c_j \\ c_j & d_j \end{pmatrix} x_j + \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} x_j x_{j+1} + \begin{pmatrix} -2 & 1 \\ 1 & -1 \end{pmatrix} x_{j+1} \succeq O,$$

$$\left(0.3(x_k^3 + x_n) + 1\right) - \|(x_k + \beta_i, x_n)\| \geq 0 \ (j, k = 1, \ldots, n-1),$$

$$1 - x_p^2 - x_{p+1}^2 - x_n^2 \geq 0 \ (p = 1, \ldots, n-2).$$

Here $a_i, b_j, d_j \in (-1, 0)$, $c_j, \beta_j \in (0, 1)$ are random numbers.

$N_p \equiv \{p, p+1, n\} \subset N \equiv \{1, 2, \ldots, n\}$ $(p = 1, 2, \ldots, n-2)$.

n	cpu sec.	ω	ϵ_{obj}	ϵ_{feas}	SDP size size of A, SeDuMi	# of nonzeros
600	25.7	2	4.0e-12	0.0	11,974 × 113,022	235,612
800	34.8	2	3.2e-12	0.0	15,974 × 150,822	314,412
1000	44.5	2	1.6e-12	0.0	19,974 × 188,622	393,212

$$\epsilon_{\text{obj}} = \frac{|\text{the lower bound for opt. value} - \text{the approx. opt. value}|}{\max\{1, |\text{the lower bound for opt. value}|\}},$$

$$\epsilon_{\text{feas}} = -\min\{\text{the left side (min.eigen)values over all constraints}, 0\}.$$

- # of nonzero elements in A increases linearly as n increases.

A sparse numerical example

$$\min \sum_{i=1}^{n} a_i x_i$$

$$\text{s.t.} \quad \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + \begin{pmatrix} b_j & c_j \\ c_j & d_j \end{pmatrix} x_j + \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} x_j x_{j+1} + \begin{pmatrix} -2 & 1 \\ 1 & -1 \end{pmatrix} x_{j+1} \succeq O,$$

$$\left(0.3(x_k^3 + x_n) + 1\right) - \|(x_k + \beta_i, x_n)\| \geq 0 \ (j, k = 1, \ldots, n-1),$$

$$\left(\left(0.3(x_k^3 + x_n) + 1\right)^2 - (x_k + \beta_i)^2 - x_n^2 \geq 0\right) \ \text{(degree 6)}$$

$$1 - x_p^2 - x_{p+1}^2 - x_n^2 \geq 0 \ (p = 1, \ldots, n-2).$$

Here $a_i, b_j, d_j \in (-1, 0)$, $c_j, \beta_j \in (0, 1)$ are random numbers.

$$N_p \equiv \{p, p+1, n\} \subset N \equiv \{1, 2, \ldots, n\} \ (p = 1, 2, \ldots, n-2).$$

n	cpu sec.	ω	ϵ_{obj}	ϵ_{feas}	SDP size size of A, SeDuMi	# of nonzeros
600	25.7	2	4.0e-12	0.0	11,974 × 113,022	235,612
800	34.8	2	3.2e-12	0.0	15,974 × 150,822	314,412
1000	44.5	2	1.6e-12	0.0	19,974 × 188,622	393,212
600	137.7	3	5.6e-12	0.0	33,515 × 539,199	1,318,200
800	218.2	3	2.0e-12	0.0	44,715 × 719,399	1,758,600
1000	229.2	3	4.8e-12	0.0	55,915 × 899,198	2,197,596

Contents

Concluding remarks

(i) Applications to polynomial least square problems: Let $f_i \in \mathbb{R}[x]$ $(i = 1, \ldots, m)$, $d = \max_i \deg(f_i)$ and $f = (f_1, \ldots, f_m)^T$.

$$\min \ \sum_{i=1}^m f_i(x)^2 \text{ or } \min \ \|f(x)\|.$$

Three different formulations for SOS relaxations.

(a) A normal POP $\Rightarrow$ degree $= 2d$: $\min \ \sum_{i=1}^m f_i(x)^2$.

(b) A polynomial SOCP $\Rightarrow$ degree $= d$: $\min \ \|f(x)\| \Leftrightarrow$

$$\min \ t \text{ sub.to } (t, f_1(x), \ldots, f_m(x)) \in \mathbb{Q}(m).$$

(c) A polynomial SDP $\Rightarrow$ degree $= d$: $\min \ \|f(x)\|^2 \Leftrightarrow$

$$\min \ t \text{ sub.to } \begin{pmatrix} I & f(x) \\ f(x)^T & t \end{pmatrix} \succeq O.$$

- (b) and (c) are better than (a) because of the difference in degrees.
- (b) is better than (c)?
  - Given the max degree of SOS multiplier polynomials, the size of SOS relaxations of (b) is smaller than that of (c).
  - effectiveness of SOS relaxation.
- SOS and SDP relaxations of (b) and (c) have structured sparsity.

Concluding remarks — continued

(ii) **POPs over symmetric cone covers wide range of nonconvex optimization problems**

(iii) **SOS relaxations proposed for POPs over symmetric cones covers are very powerful in theory — global converence**

(iv) **Computationally very expensive — large scale SDPs**

(v) **Exploiting sparsity is necessary!**