Sum of Squares and SemiDefinite Programmming Relaxations of Polynomial Optimization Problems

The 2006 IEICE Society Conference
Kanazawa, September 21, 2006

Masakazu Kojima
Tokyo Institute of Technology

This presentation PowerPoint file is available at http://www.is.titech.ac.jp/~kojima/talk.html

An introduction to the recent development of SOS and SDP relaxations for computing global optimal solutions of POPs

Exploiting sparsity in SOS and SDP relaxations to solve large scale POPs

Outline

1. POPs (Polynomial Optimization Problems)
2. Nonnegative polynomials and SOS (Sum of Squares) polynomials
3. SOS relaxation of unconstrained POPs
4. Conversion of SOS relaxation into an SDP (Semidefinite Program)
5. Exploiting structured sparsity
6. SOS relaxation of constrained POPs --- very briefly
7. Numerical results
8. Polynomial SDPs
9. Concluding remarks

Outline

1. POPs (Polynomial Optimization Problems)
2. Nonnegative polynomials and SOS (Sum of Squares) polynomials
3. SOS relaxation of unconstrained POPs
4. Conversion of SOS relaxation into an SDP (Semidefinite Program)
5. Exploiting structured sparsity
6. SOS relaxation of constrained POPs --- very briefly
7. Numerical results
8. Polynomial SDPs
9. Concluding remarks

POP: min $f_0(x)$ sub.to $f_j(x) \geq 0$ $(j = 1, \ldots, m)$.

$\mathbb{R}^n$: the n-dim Euclidean space.

$x = (x_1, \ldots, x_n) \in \mathbb{R}^n$: a vector variable.

$f_j(x)$: a multivariate polynomial in $x \in \mathbb{R}^n$ $(j = 0, 1, \ldots, m)$.

Example: $n = 3$

$$
\begin{array}{rl}
\text{min} & f_0(x) \equiv x_1^3 - 2x_1x_2^2 + x_1^2x_2x_3 - 4x_3^2 \\
\text{sub.to} & f_1(x) \equiv -x_1^2 + 5x_2x_3 + 1 \geq 0, \\
& f_2(x) \equiv x_1^2 - 3x_1x_2x_3 + 2x_3 + 2 \geq 0, \\
& f_3(x) \equiv -x_1^2 - x_2^2 - x_3^2 + 1 \geq 0,
\end{array}
$$

$x_1(x_1 - 1) = 0$ (0-1 integer),

$x_2 \geq 0,\ x_3 \geq 0,\ x_2x_3 = 0$ (complementarity).

POP: min $f_0(x)$ sub.to $f_j(x) \geq 0$ $(j = 1, \ldots, m)$.

[1] J.B.Lasserre, "Global optimization with polynomials and the problems of moments", ***SIAM J. on Optim.*** (2001).

[2] P.A.Parrilo, "Semidefinite programming relaxations for semialgebraic problems", ***Math. Prog.*** (2003).

[3] D.Henrion and J.B.Lasserre, GloptiPoly.

[4] S.Prajna, A.Parachristodoulou and P.A.Parrilo, SOSTOOLS.

- [1,3]⇒ "a sequence of SDP relaxations" — primal approach.
- [2,4]⇒ "a sequence of SOS relaxations" — dual approach.

(b) Lower bounds for the optimal value.

(c) Convergence to global optimal solutions in theory.

(a) Each relaxed problem can be solved as an SDP; its size gets larger rapidly along "the sequence" as we require a higher accuracy.

(d) Expensive to solve large scale POPs in practice.
⇒ Exploiting Sparsity.

POP: min $f_0(x)$ sub.to $f_j(x) \geq 0$ $(j = 1, \ldots, m)$.

[1] J.B.Lasserre, "Global optimization with polynomials and the problems of moments", ***SIAM J. on Optim.*** (2001).

[2] P.A.Parrilo, "Semidefinite programming relaxations for semialgebraic problems", ***Math. Prog.*** (2003).

[3] D.Henrion and J.B.Lasserre, GloptiPoly.

[4] S.Prajna, A.Parachristodoulou and P.A.Parrilo, SOSTOOLS.

Exploiting sparsity to solve larger scale problem in practice

[5] M. Kojima, S. Kim and H. Waki, "Sparsity in SOS Polynomials", *Math. Prog.* (2005).

[6] H. Waki, S. Kim, M. Kojima and M. Muramatsu, "SOS and SDP Relaxations for POPs with Structured Sparsity", *SIAM J. on Optim* (2006).

[7] H. Waki, S. Kim, M. Kojima and M. Muramatsu, SparsePOP (2005).

POP: min $f_0(x)$ sub.to $f_j(x) \geq 0$ $(j = 1, \ldots, m)$.

[1] J.B.Lasserre, "Global optimization with polynomials and the problems of moments", ***SIAM J. on Optim.*** (2001).

[2] P.A.Parrilo, "Semidefinite programming relaxations for semialgebraic problems", ***Math. Prog.*** (2003).

[3] D.Henrion and J.B.Lasserre, GloptiPoly.

[4] S.Prajna, A.Parachristodoulou and P.A.Parrilo, SOSTOOLS.

Extension to polynomial SDP and SOCP

[8] M. Kojima, "SOS relaxations of polynomial SDPs" (2003).

[9] C.W.Hol and C.W.Schere,"SOS relaxations of polynomial SDPs" (2004).

[10] D. Henrion and J. B. Lasserre, "Convergent relaxations of polynomial matrix inequalities and static output feedback", *IEEE Transactions on Automatic Conrol* (2006).

[11] M. Kojima and M. Muramatsu, "An Extension of SOS Relaxations to POPs over Symmetric Cones ", to applear in *Math. Prog.*

A sparse numerical example with poly. SDP and SOCP constraints

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

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

(polynomial matrix inequality constraints)

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

(polynomial second-order inequality constraints)

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

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

Outline

1. POPs (Polynomial Optimization Problems)
2. Nonnegative polynomials and SOS (Sum of Squares) polynomials
3. SOS relaxation of unconstrained POPs
4. Conversion of SOS relaxation into an SDP (Semidefinite Program)
5. Exploiting structured sparsity
6. SOS relaxation of constrained POPs --- very briefly
7. Numerical results
8. Polynomial SDPs
9. Concluding remarks

Nonnegative polynomials and SOS polynomials

$f(x)$: a nonnegative polynomial $\Leftrightarrow$ $f(x) \geq 0$ $(\forall x \in \mathbb{R}^n)$.
$\mathcal{N}$: the set of nonnegative polynomials in $x \in \mathbb{R}^n$.

$f(x)$: an SOS (Sum of Squares) polynomial

$$\Updownarrow$$

$$\exists \text{ polynomials } g_1(x), \ldots, g_k(x);\ f(x) = \sum_{i=1}^{k} g_i(x)^2.$$

SOS_* : the set of SOS. Obviously, $\mathrm{SOS}_* \subset \mathcal{N}$.
$\mathrm{SOS}_{2r} = \{f \in \mathrm{SOS}_* : \deg f \leq 2r\}$: SOSs with degree ar most $2r$.

$n = 2$. $f(x_1, x_2) = (x_1^2 - 2x_2 + 1)^2 + (3x_1x_2 + x_2 - 4)^2 \in \mathrm{SOS}_4$.

- In theory, SOS_* (SOS) $\subset \mathcal{N}$. $\mathrm{SOS}_* \neq \mathcal{N}$ in general.
- In practice, $f(x) \in \mathcal{N} \backslash \mathrm{SOS}_*$ is rare.
- So we replace $\mathcal{N}$ by $\mathrm{SOS}_* \Longrightarrow$ SOS Relaxations.

Outline

1. POPs (Polynomial Optimization Problems)
2. Nonnegative polynomials and SOS (Sum of Squares) polynomials
3. SOS relaxation of unconstrained POPs
4. Conversion of SOS relaxation into an SDP (Semidefinite Program)
5. Exploiting structured sparsity
6. SOS relaxation of constrained POPs --- very briefly
7. Numerical results
8. Polynomial SDPs
9. Concluding remarks

$\mathcal{P}$: $\displaystyle\min_{x \in \mathbb{R}^n} f(x)$, where f is a polynomial with deg $f = 2r$

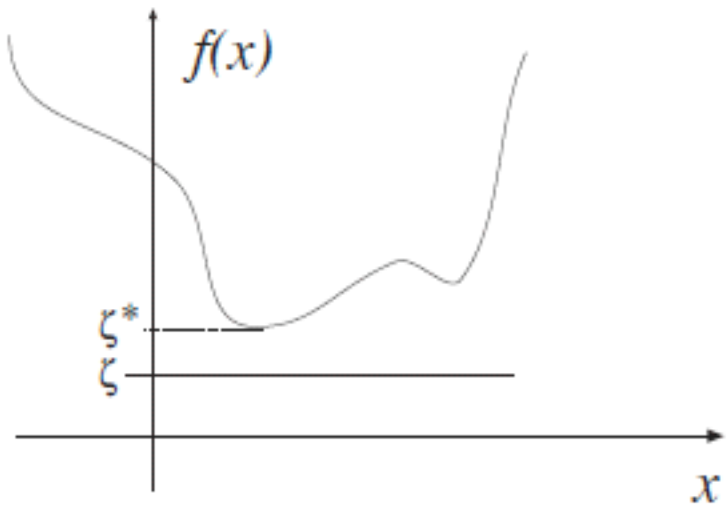
$\Updownarrow$

$\mathcal{P}$': max ζ s.t $f(x) - \zeta \geq 0$ $(\forall x \in \mathbb{R}^n)$

$\Updownarrow$

$f(x) - \zeta \in \mathcal{N}$ (the nonnegative polynomials)

Here x is a parameter (index) describing inequality constraints.



$\mathcal{P}$: $\min_{x \in \mathbb{R}^n} f(x)$, where f is a polynomial with $\deg f = 2r$

$$\Updownarrow$$

$\mathcal{P}$': $\max\ \zeta$ s.t $f(x) - \zeta \geq 0\ (\forall x \in \mathbb{R}^n)$

$$\Updownarrow$$

$f(x) - \zeta \in \mathcal{N}$ (the nonnegative polynomials)

Here x is a parameter (index) describing inequality constraints.

$\Sigma \subset \mathrm{SOS}_{2r} \subset \mathrm{SOS}_* \subset \mathcal{N}$ $\Downarrow$ a subproblem of $\mathcal{P}'$ = a relaxation of $\mathcal{P}$

$\mathcal{P}$": $\max\ \zeta$ sub.to $f(x) - \zeta \in \Sigma$

SOS_* (SOS_{2r} =) the set of SOS polynomials (with degree $\leq 2r$).

- the min.val of $\mathcal{P}$ = the max.val of $\mathcal{P}'$ $\geq$ the max.val of $\mathcal{P}$".
- $\mathcal{P}$" can be solved as an SDP (Semidefinite Program) — next.
- In practice, we can exploit structured sparsity of the Hessian matrix of f to reduce the size of Σ — later.

Outline

1. POPs (Polynomial Optimization Problems)
2. Nonnegative polynomials and SOS (Sum of Squares) polynomials
3. SOS relaxation of unconstrained POPs
4. Conversion of SOS relaxation into an SDP (Semidefinite Program)
5. Exploiting structured sparsity
6. SOS relaxation of constrained POPs --- very briefly
7. Numerical results
8. Polynomial SDPs
9. Concluding remarks

Conversion of SOS relaxation into an SDP --- 1

What is an SDP (Semidefinite Program)?

- **An extension of LP (Linear Program) in $\mathbb{R}^n$ to the space $\boldsymbol{\mathcal{S}}^n$ of symmetric matrices.**

  variable a vector $x \in \mathbb{R}^n \Longrightarrow X \in \boldsymbol{\mathcal{S}}^n$.

  inequality $\mathbb{R}^n \ni x \geq 0 \Longrightarrow \boldsymbol{\mathcal{S}}^n \ni X \succeq O$ (positive semidefinite).

- **Can be solved by the interior-point method.**

- **Lots of applications.**

Conversion of SOS relaxation into an SDP --- 2

$a_p \in \mathbb{R}^n$ $(p = 0, 1, 2, \ldots, m)$, $b_p \in \mathbb{R}$ $(p = 1, 2, \ldots, m)$: data.

$x \in \mathbb{R}^n$: variable.

$a_p \cdot x = \sum_{j=1}^n [a_p]_j x_j$ (the inner product).

LP (Linear Program):

$$\begin{array}{ll} \max & a_0 \cdot x \\ \text{s.t.} & a_p \cdot x = b_p \ (p = 1, \ldots, m), \ x \geq 0. \end{array}$$

SDP (Semidefinite Program):

$$\begin{array}{ll} \max & A_0 \bullet X \\ \text{s.t.} & A_p \bullet X = b_p \ (p = 1, \ldots, m), \ X \succeq O. \end{array}$$

$A_p \in \mathcal{S}^n$ $(p = 0, 1, 2, \ldots, m)$, $b_p \in \mathbb{R}$ $(p = 1, 2, \ldots, m)$: data

$X \in \mathcal{S}^n$: variable.

$A_p \bullet X = \sum_{i=1}^n \sum_{j=1}^n [A_p]_{ij} X_{ij}$ (the inner product).

$\mathcal{S}^n$: the set of $n \times n$ real symmetric matrices.

$X \succeq O$: $X \in \mathcal{S}^n$ is positive semidefinite.

Conversion of SOS relaxation into an SDP --- 3

Representation of

$$\mathrm{SOS}_{2r} \equiv \left\{ \sum_{j=1}^{k} g_j(x)^2 : \exists k \geq 1,\ g_j(x) : \text{ degree at most } r \right\} \subset \mathrm{SOS}_*.$$

$\forall$ r-degree poly. $g(x)$ $\exists a \in \mathbb{R}^{d(r)}$; $g(x) = a^T u_r(x)$, where

$$u_r(x) = (1, x_1, x_2, \ldots, x_n, x_1^2, x_1x_2, x_1x_3, \ldots, x_n^2, \ldots, x_1^r, \ldots, x_n^r)^T,$$

(a column vector of a basis of r-degree polynomial),

$$d(r) = \binom{n+r}{r} : \text{ the dimension of } u_r(x).$$

Example: $n = 2$ and $r = 2$

$$\begin{aligned} g(x_1, x_2) &= 1 - 2x_1 - 4x_1^2 + 5x_1x_2 - 6x_2^2 \\ &= (1,\ -2,\ 0,\ -4,\ 5, -6)(1,\ x_1,\ x_2,\ x_1^2,\ x_1x_2,\ x_2^2)^T \\ &= a^T u_2(x), \end{aligned}$$

where

$$\begin{aligned} a^T &= (1,\ -2,\ 0,\ -4,\ 5, -6), \\ u_2(x)^T &= (1,\ x_1,\ x_2,\ x_1^2,\ x_1x_2,\ x_2^2)^T. \end{aligned}$$

Conversion of SOS relaxation into an SDP --- 3

Representation of

$$\mathrm{SOS}_{2r} \equiv \left\{ \sum_{j=1}^{k} g_j(x)^2 : \exists k \geq 1,\ g_j(x) : \text{ degree at most } r \right\} \subset \mathrm{SOS}_*.$$

$\forall$ r-degree poly. $g(x)$ $\exists a \in \mathbb{R}^{d(r)}$; $g(x) = a^T u_r(x)$, where

$$u_r(x) = (1, x_1, x_2, \ldots, x_n, x_1^2, x_1x_2, x_1x_3, \ldots, x_n^2, \ldots, x_1^r, \ldots, x_n^r)^T,$$

(a column vector of a basis of r-degree polynomial),

$$d(r) = \binom{n+r}{r} : \text{ the dimension of } u_r(x).$$

$$\Downarrow$$

$$\begin{aligned}
\mathrm{SOS}_{2r} &= \left\{ \sum_{j=1}^{k} \left(a_j^T u_r(x) \right)^2 : k \geq 1,\ a_j \in \mathbb{R}^{d(r)} \right\} \\
&= \left\{ u_r(x)^T \left(\sum_{j=1}^{k} a_j a_j^T \right) u_r(x) : k \geq 1,\ a_j \in \mathbb{R}^{d(r)} \right\} \\
&= \left\{ u_r(x)^T V u_r(x) : V \text{ is a positive semidefinite matrix} \right\}.
\end{aligned}$$

Conversion of SOS relaxation into an SDP --- 4

Example. $n = 2$, SOS of at most deg.2 polynomials in $x = (x_1, x_2)$.

$$\mathrm{SOS}_4 \equiv \left\{ \sum_{i=1}^{k} g_i(x)^2 \;:\; k \geq 1,\; g_i(x) \text{ is at most deg.2 polynomial} \right\}$$

$$= \left\{ \begin{pmatrix} 1 \\ x_1 \\ x_2 \\ x_1^2 \\ x_1 x_2 \\ x_2^2 \end{pmatrix}^T V \begin{pmatrix} 1 \\ x_1 \\ x_2 \\ x_1^2 \\ x_1 x_2 \\ x_2^2 \end{pmatrix} \;:\; V \text{ is a } 6 \times 6 \text{ psd matrix} \right\}$$

Conversion of SOS relaxation into an SDP --- 5

Example : $f(x) = -x_1 + 2x_2 + 3x_1^2 - 5x_1^2x_2^2 + 7x_2^4$

$\max \zeta$ sub.to $f(x) - \zeta \in \mathrm{SOS}_4$ (SOS of at most deg. 2 polynomials)

$\Updownarrow$

max ζ

Sum of Squares

$$\text{s.t.}\quad f(x) - \zeta = \begin{pmatrix} 1 \\ x_1 \\ x_2 \\ x_1^2 \\ x_1x_2 \\ x_2^2 \end{pmatrix}^T \begin{pmatrix} V_{11} & V_{12} & V_{13} & V_{14} & V_{15} & V_{16} \\ V_{12} & V_{22} & V_{23} & V_{24} & V_{25} & V_{26} \\ V_{13} & V_{23} & V_{33} & V_{34} & V_{35} & V_{36} \\ V_{14} & V_{24} & V_{34} & V_{44} & V_{45} & V_{46} \\ V_{15} & V_{25} & V_{35} & V_{45} & V_{55} & V_{56} \\ V_{16} & V_{26} & V_{36} & V_{46} & V_{56} & V_{66} \end{pmatrix} \begin{pmatrix} 1 \\ x_1 \\ x_2 \\ x_1^2 \\ x_1x_2 \\ x_2^2 \end{pmatrix}$$

$(\forall (x_1, x_2)^T \in \mathbb{R}^n)$, $\quad 6 \times 6\ V \succeq O$

$\Updownarrow$ Compare the coef. of $1,\ x_1,\ x_2,\ x_1^2, \ldots, x_2^4$ on both side of $=$

SDP (Semidefinite Program)

$$\max\ \zeta\ \text{s.t.}\quad -\zeta = V_{11},\ -1 = 2V_{12},\ 2 = 2V_{13},\ 3 = 2V_{14} + V_{22},$$
$$-5 = 2V_{46} + V_{55},\ 7 = V_{66},\ \text{all others}\ \ 0 = \cdots,\ V \succeq O$$

In general, each equality constraint is a linear equation in ζ and V.

Outline

1. POPs (Polynomial Optimization Problems)
2. Nonnegative polynomials and SOS (Sum of Squares) polynomials
3. SOS relaxation of unconstrained POPs
4. Conversion of SOS relaxation into an SDP (Semidefinite Program)
5. Exploiting structured sparsity
6. SOS relaxation of constrained POPs --- very briefly
7. Numerical results
8. Polynomial SDPs
9. Concluding remarks

$\mathcal{P}$: $\min_{x \in \mathbb{R}^n} f(x)$, where f is a polynomial with deg $f = 2r$

H : the sparsity pattern of the Hessian matrix of $f(x)$

$$H_{ij} = \begin{cases} \star \text{ if } i = j \text{ or } \partial^2 f(x)/\partial x_i \partial x_j \not\equiv 0, \\ 0 \text{ otherwise.} \end{cases}$$

$f(x)$: correlatively sparse $\Leftrightarrow$ $\exists$ a sparse Cholesky fact. of H.

(a) A sparse Chol. fact. is characterized as a sparse (chordal) graph $G(N, E)$; $N = \{1, \ldots, n\}$ and

$$E = \{(i, j) : H_{ij} = \star\} + \text{ "fill-in".}$$

(b) Let $C_1, C_2, \ldots, C_q \subset N$ be the maximal cliques of $G(N, E)$.

Sparse SOS relaxation

max ζ

s.t. $f(x) - \zeta \in \sum_{k=1}^{q}$ (SOS of polynomials in x_i $(i \in C_k)$)

Dense SOS relaxation

max ζ

s.t. $f(x) - \zeta \in$ (SOS of polynomials in x_i $(i \in N)$)

- Sparse relaxation is weaker but less expensive in practice.

Example: Generalized Rosenbrock function.

$$f(x) = \sum_{i=2}^{n} \left(100(x_i - x_{i-1}^2)^2 + (1 - x_i)^2\right).$$

Dense SOS relaxation
max ζ
s.t. $f(x) - \zeta \in$ (SOS of deg-2. poly. in $x_1, x_2, \ldots, x_n$)

- The size of Dense grows very rapidly, so we can't apply Dense to the case $n \geq 20$ in practice.
- The Hessian matrix is sparse (tridiagonal).
- No fill-in in the Cholesky factorization.
- $C_i = \{i-1, i\}$ $(i = 2, \ldots, n-1)$: the max. cliques.

Sparse SOS relaxation
max ζ
s.t. $f(x) - \zeta \in \sum_{i=2}^{n}$ (SOS of deg-2. poly. in $x_{i-1},\ x_i$)

- The size of Sparse grows linearly in n, and Sparse can process the case $n = 800$ in less than 10 sec.

Outline

1. POPs (Polynomial Optimization Problems)
2. Nonnegative polynomials and SOS (Sum of Squares) polynomials
3. SOS relaxation of unconstrained POPs
4. Conversion of SOS relaxation into an SDP (Semidefinite Program)
5. Exploiting structured sparsity
6. SOS relaxation of constrained POPs --- very briefly
7. Numerical results
8. Polynomial SDPs
9. Concluding remarks

POP: $\min\ f_0(x)$ sub.to $f_j(x) \geq 0\ (j = 1, \ldots, m)$.

- Rough sketch of SOS relaxation of POP

“Generalized Lagrangian Dual”

$+$

“SOS relaxation of unconstrained POPs”

$\Downarrow$

SOS relaxation of POP

- Exploiting sparsity in SOS relaxation of POP

POP: $\min\ f_0(x)$ sub.to $f_j(x) \geq 0\ (j = 1, \ldots, m)$.

Sparsity : $f_j(x)$ involves only x_i $(i \in C_j \subset N)$ $(j = 1, \ldots, n)$.

Generalized Lagrangian function

$$L(x, \lambda_1, \ldots, \lambda_m) = f_0(x) - \sum_{j=1}^{m} \lambda_j(x) f_j(x)$$

for $\forall x \in \mathbb{R}^n,\ \forall \lambda_j \in \mathrm{SOS}_*$

If $\mathbb{R} \ni \lambda_j \geq 0$ then L is the standard Lagrangian function.

Generalized Lagrangian Dual

$$\max_{\lambda_1 \in \mathrm{SOS}_*, \ldots, \lambda_m \in \mathrm{SOS}_*} \ \min_{x \in \mathbb{R}^n} L(x, \lambda_1, \ldots, \lambda_m)$$

$\Updownarrow$

Generalized Lagrangian Dual

$$\max\ \zeta \text{ s.t. } L(x, \lambda_1, \ldots, \lambda_m) - \zeta \geq 0\ (\forall x \in \mathbb{R}^n),$$
$$\lambda_1 \in \mathrm{SOS}_*, \ldots, \lambda_m \in \mathrm{SOS}_*$$

POP: $\min\ f_0(x)$ sub.to $f_j(x) \geq 0\ (j = 1, \ldots, m)$.

Sparsity : $f_j(x)$ involves only x_i $(i \in C_j \subset N)$ $(j = 1, \ldots, n)$.

Generalized Lagrangian function

$$L(x, \lambda_1, \ldots, \lambda_m) = f_0(x) - \sum_{j=1}^{m} \lambda_j(x) f_j(x)$$

$$\text{for } \forall x \in \mathbb{R}^n,\ \forall \lambda_j \in \mathrm{SOS}_*$$

If $\mathbb{R} \ni \lambda_j \geq 0$ then L is the standard Lagrangian function.

Generalized Lagrangian Dual

$$\max\ \zeta \text{ s.t. } L(x, \lambda_1, \ldots, \lambda_m) - \zeta \geq 0\ (\forall x \in \mathbb{R}^n),$$
$$\lambda_1 \in \mathrm{SOS}_*, \ldots, \lambda_m \in \mathrm{SOS}_*$$

$\Downarrow$ sparse SOS relaxation

$$\max\ \zeta \text{ s.t. } L(x, \lambda_1, \ldots, \lambda_m) - \zeta \in \Sigma$$
$$\lambda_1 \in \Sigma_1, \ldots, \lambda_m \in \Sigma_m.$$

- Here $\Sigma_j \subset \mathrm{SOS}_*$ $(j = 1, \ldots, m)$: a set of SOS poly. in x_i $(i \in C_j)$. $\Rightarrow L(x, \lambda_1, \ldots, \lambda_m) - \zeta$: correlatively sparse.
- SOS relaxation of unconstrained POPs to choose $\Sigma \subset \mathrm{SOS}_*$.

Example

$$\begin{array}{ll}\min & f_0(x) = -x_1 - x_2 - x_3 - x_4 - x_5 \\ \text{s.t} & f_1(x_1, x_2) = -x_1^4 - 2x_2^2 + 1 \geq 0, \;\; f_2(x_2, x_3) = -3x_2^4 - 4x_3^2 + 1 \geq 0, \\ & f_3(x_3, x_4) = -x_3^4 - 3x_4^2 - 1 \geq 0, \;\; f_4(x_4, x_5) = -2x_4^4 - x_5^2 - 1 \geq 0.\end{array}$$

Generalized Lagrangian function

$$\begin{aligned} & L(x, \lambda_1, \ldots, \lambda_m) \\ & = f_0(x) - \lambda_1(x_1, x_2) f_1(x_1, x_2) - \lambda_2(x_2, x_3) f_2(x_3, x_4) \\ & \quad -\lambda_3(x_3, x_4) f_3(x_3, x_4) - \lambda_4(x_4, x_5) f_4(x_4, x_5). \end{aligned}$$

Here $\lambda_j \in \text{SOS}_*$.

Then the sparsity pattern of the Hessian matrix of $L(x, \lambda_1, \ldots, \lambda_m)$ becomes

$$H = \begin{pmatrix} \star & \star & 0 & 0 & 0 \\ \star & \star & \star & 0 & 0 \\ 0 & \star & \star & \star & 0 \\ 0 & 0 & \star & \star & \star \\ 0 & 0 & 0 & \star & \star \end{pmatrix}.$$

Thus $L(x, \lambda_1, \ldots, \lambda_m) - \zeta$: correlatively sparse.

Outline

1. POPs (Polynomial Optimization Problems)
2. Nonnegative polynomials and SOS (Sum of Squares) polynomials
3. SOS relaxation of unconstrained POPs
4. Conversion of SOS relaxation into an SDP (Semidefinite Program)
5. Exploiting structured sparsity
6. SOS relaxation of constrained POPs --- very briefly
7. Numerical results
8. Polynomial SDPs
9. Concluding remarks

Numerical results

Software

- SparsePOP (Waki-Kim-Kojima-Muramatsu, 2005)
  — MATLAB program for constructing sparse and dense SDP relaxation problems.
- SeDuMi to solve SDPs.

Hardware

- 2.4GHz Xeon cpu with 6.0GB memory.

G.Rosenbrock function:

$$f(x) = \sum_{i=2}^{n} \left(100(x_i - x_{i-1}^2)^2 + (1 - x_i)^2\right)$$

- Two minimizers on $\mathbb{R}^n$: $x_1 = \pm 1$, $x_i = 1$ $(i \geq 2)$.
- Add $x_1 \geq 0 \Rightarrow$ a single minimizer.

		cpu in sec.	
n	ϵ_{obj}	Sparse	Dense
10	2.5e-08	0.2	10.6
15	6.5e-08	0.2	756.6
200	5.2e-07	2.2	—
400	2.5e-06	3.7	—
800	5.5e-06	6.8	—

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

An optimal control problem from Coleman et al. 1995

$$\left.\begin{array}{l} \min \ \frac{1}{M}\sum_{i=1}^{M-1}(y_i^2+x_i^2) \\ \text{s.t.} \quad y_{i+1}=y_i+\frac{1}{M}(y_i^2-x_i), \quad (i=1,\ldots,M-1), \quad y_1=1. \end{array}\right\}$$

Numerical results on sparse relaxation

M	# of variables	ϵ_{obj}	ϵ_{feas}	cpu
600	1198	3.4e-08	2.2e-10	3.4
700	1398	2.5e-08	8.1e-10	3.3
800	1598	5.9e-08	1.6e-10	3.8
900	1798	1.4e-07	6.8e-10	4.5
1000	1998	6.3e-08	2.7e-10	5.0

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

ϵ_{feas} = the maximum error in the equality constraints,

cpu : cpu time in sec. to solve an SDP relaxation problem.

alkyl.gms : a benchmark problem from globallib

$$
\begin{array}{ll}
\text{min} & -6.3x_5x_8 + 5.04x_2 + 0.35x_3 + x_4 + 3.36x_6 \\
\text{sub.to} & -0.820x_2 + x_5 - 0.820x_6 = 0, \\
& 0.98x_4 - x_7(0.01x_5x_{10} + x_4) = 0, \\
& -x_2x_9 + 10x_3 + x_6 = 0, \\
& x_5x_{12} - x_2(1.12 + 0.132x_9 - 0.0067x_9^2) = 0, \\
& x_8x_{13} - 0.01x_9(1.098 - 0.038x_9) - 0.325x_7 = 0.574, \\
& x_{10}x_{14} + 22.2x_{11} = 35.82, \\
& x_1x_{11} - 3x_8 = -1.33, \\
& \text{lbd}_i \le x_i \le \text{ubd}_i \ (i = 1, 2, \ldots, 14).
\end{array}
$$

		Sparse			Dense (Lasserre)		
problem	n	ϵ_{obj}	ϵ_{feas}	cpu	ϵ_{obj}	ϵ_{feas}	cpu
alkyl	14	5.6e-10	2.0e-08	23.0	out of memory		

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

ϵ_{feas} = the maximum error in the equality constraints,

cpu : cpu time in sec. to solve an SDP relaxation problem.

Some other benchmark problems from globallib

		Sparse			Dense (Lasserre)		
problem	n	ϵ_{obj}	ϵ_{feas}	cpu	ϵ_{obj}	ϵ_{feas}	cpu
ex3_1_1	8	6.3e-09	4.7e-04	5.5	0.7e-08	2.5e-03	597.8
st_bpaf1b	10	3.8e-08	2.8e-08	1.0	4.6e-09	7.2e-10	1.7
st_e07	10	0.0e+00	8.1e-05	0.4	0.0e+00	8.8e-06	3.0
st_jcbpaf2	10	1.1e-07	0.0e+00	2.1	1.1e-07	0.0e+00	2.0
ex2_1_3	13	5.1e-09	3.5e-09	0.5	1.6e-09	1.5e-09	7.7
ex9_1_1	13	0.0	4.5e-06	1.5	0.0	9.2e-07	7.7
ex9_2_3	16	0.0e+00	5.7e-06	2.3	0.0e+00	7.5e-06	49.7
ex2_1_8	24	1.0e-05	0.0e+00	304.6	3.4e-06	0.0e+00	1946.6
ex5_2_2_c1	9	1.0e-2	3.2e+01	1.8	1.6e-05	2.1e-01	2.6
ex5_2_2_c2	9	1.0e-02	7.2e+01	2.1	1.3e-04	2.7e-01	3.5

- ex5_2_2_c1 and ex5_2_2_c2 — Dense is better.
- Sparse attains approx. opt. solutions with the same quality as Dense except ex5_2_2_c1 and ex5_2_2_c2.
- Sparse is much faster than Dense in large dim. cases.

Outline

1. POPs (Polynomial Optimization Problems)
2. Nonnegative polynomials and SOS (Sum of Squares) polynomials
3. SOS relaxation of unconstrained POPs
4. Conversion of SOS relaxation into an SDP (Semidefinite Program)
5. Exploiting structured sparsity
6. SOS relaxation of constrained POPs --- very briefly
7. Numerical results
8. Polynomial SDPs
9. Concluding remarks

(Sparse) SOS and SDP relaxations have been extended to

PSDP (Polynomial Semidefinite Program)

$$\begin{array}{ll} \max & \sum_{i=1}^{n} c_i x_i \\ \text{sub.to} & \text{polynomial matrix inequalities.} \end{array}$$

Example:

$$\min\ 1.1x_1 + 1.2x_2 - x_1^2 - x_2^2 \text{ sub.to} \begin{pmatrix} 1 - 4x_1x_2 & x_1^3 \\ x_1^3 & 4 - x_1^2 - x_2^2 \end{pmatrix} \succeq O.$$

- Can be solved in 0.4 second with relative accuracy 3.9e-10.

[A] M.Kojima, “SOS relaxations of POPs”, 2003.

[B] C.W.Hol and C.W.Scherer, “Sum of squares relaxations for polynomial semidefinite programming”, 2004.

[C] M.Kojima and M.Muramatsu, “An extension of SOS relaxations to POPs over symmetric cones”, To appear in *Math. Prog.*

- Powerful in theory, but not practical yet.

(Sparse) SOS and SDP relaxations have been extended to

PSDP (Polynomial Semidefinite Program)

$$\begin{array}{ll} \max & \sum_{i=1}^{n} c_i x_i \\ \text{sub.to} & \text{polynomial matrix inequalities.} \end{array}$$

Example:

$$\min\ 1.1x_1 + 1.2x_2 - x_1^2 - x_2^2 \text{ sub.to } \begin{pmatrix} 1 - 4x_1x_2 & x_1^3 \\ x_1^3 & 4 - x_1^2 - x_2^2 \end{pmatrix} \succeq O.$$

- Can be solved in 0.4 second with relative accuracy 3.9e-10.

In theory:

- Convergence to a global optimal solution.
- Exploiting sparsity.

In practice:

- SDP relaxation problems become too large to solve as PSDP gets larger.
- Numerical difficulty to solve SDP relaxation problems

An example of polynomial SDPs

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

s.t. $I -$ "deg 3 poly. with $k \times k$ sym. dense matrix coefficients" $\succeq O$,
$0 \leq x_j \leq 1 \; (j = 1, \ldots, n)$.

Here I denotes the $k \times k$ identity matrix.

n	k	cpu sec.	ϵ_{obj}	ϵ_{feas}	SDP size size of A, SeDuMi	# of nonzeros
7	5	19.6	2.0-09	6.9-10	791×22,608	41,587
8	5	103.3	2.4e-09	4.0e-10	1,286×39,006	69,772
9	5	212.7	6.4e-10	1.2e-10	2,001×63,959	109,169
10	5	828.9	6.8e-10	1.8e-10	3002×100,385	171,895
7	10	23.4	2.8e-10	3.0e-10	791×27,408	75,502
7	20	38.2	3.3e-10	6.0e-09	791×46,608	210,532
7	40	123.0	2.6e-09	4.1e-08	791×123,408	749,392

A sparse numerical example with poly. SDP and SOCP constraints

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

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

(polynomial matrix inequality constraints)

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

(polynomial second-order inequality constraints)

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

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

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

Outline

1. POPs (Polynomial Optimization Problems)
2. Nonnegative polynomials and SOS (Sum of Squares) polynomials
3. SOS relaxation of unconstrained POPs
4. Conversion of SOS relaxation into an SDP (Semidefinite Program)
5. Exploiting structured sparsity
6. SOS relaxation of constrained POPs --- very briefly
7. Numerical results
8. Polynomial SDPs
9. Concluding remarks

- Lasserre's (dense) relaxation
  — Theoretical convergence but expensive in practice.
- Sparse relaxation
  = Lasserre's (dense) relaxation + sparsity
  — Theoretical convergence and very powerful in practice.

- There remain many issues to be studied further.
  - Exploiting sparsity.
  - Large-scale SDPs.
  - Numerical difficulty in solving SDP relaxations of POPs.
  - Polynomial SDPs.

This presentation material is available at

http://www.is.titech.ac.jp/~kojima/talk.html

Thank you!