Operations Research & Financial Engineering

PRINCETON UNIVERSITY

Iterative LP and SOCP-based approximations to sum of squares programs

Georgina Hall
Princeton University

Joint work with:

Amir Ali Ahmadi (Princeton University)

Sanjeeb Dash (IBM)

PRINCETON UNIVERSITY

Sum of squares programs

- Problems of the type:

Decision variables are the **coefficients** of the polynomial p

$$\min_{p} C(p)$$
$$s.t. \; A(p) = b$$
$$\boldsymbol{p(x)} \textbf{ sos}$$

Linear objective and **affine constraints** in the coefficients of p (e.g., sum of coefs =1)

Sum of squares condition

Many applications for problems of this form

PRINCETON UNIVERSITY

Semidefinite programming formulation

Sum of squares program:

$$\min_{p} C(p)$$
$$s.t.\, A(p) = b$$
$$\boldsymbol{p\ sos}$$

Equivalent semidefinite programming formulation:

$$\min_{p,Q} C(p)$$
$$s.t.\, A(p) = b$$
$$p = z(x)^T Q z(x)$$
$$Q \succcurlyeq 0$$

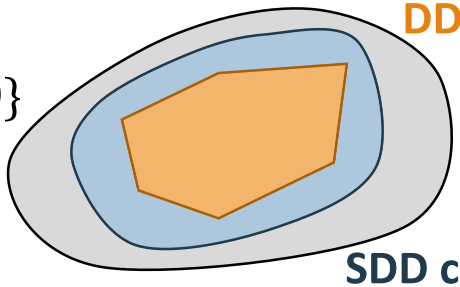
But: Size of $Q = \binom{n+d}{d} \times \binom{n+d}{d}$

PRINCETON UNIVERSITY

Alternatives to sum of squares: dsos and sdsos

Sum of squares **(sos)**	$p(x) = z(x)^T Q z(x), Q \succcurlyeq 0$	**SDP**

PSD cone $:= \{\boldsymbol{Q} \mid \boldsymbol{Q} \succcurlyeq \boldsymbol{0}\}$



DD cone $:= \{\boldsymbol{Q} \mid \boldsymbol{Q}_{ii} \geq \sum_j |\boldsymbol{Q}_{ij}|, \forall \boldsymbol{i}\}$

SDD cone $:= \{\boldsymbol{Q} \mid \exists$ **diagonal** $\boldsymbol{D}$ **with** $\boldsymbol{D}_{ii} > \boldsymbol{0}$ **s.t.** $\boldsymbol{DQD}$ $\boldsymbol{dd}\}$

Diagonally dominant sum of squares **(dsos)**	$p(x) = z(x)^T Q z(x), Q\ \textit{diagonally dominant}\ (\mathrm{dd})$	**LP**
Scaled diagonally dominant sum of squares **(sdsos)**	$p(x) = z(x)^T Q z(x), Q\ \textit{scaled diagonally dominant}\ (\mathrm{sdd})$	**SOCP**

[Ahmadi, Majumdar]

PRINCETON UNIVERSITY

Alternatives to sum of squares: dsos and sdsos

$$\min C(p) \quad s.t.\ A(p) = b \quad \boldsymbol{p\ sos}$$

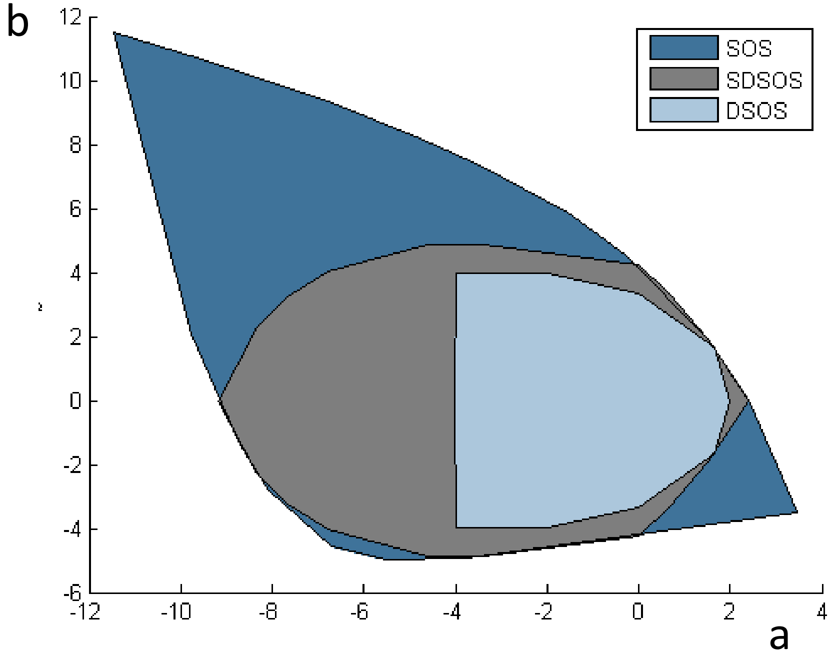
scalability →

$$\min C(p) \quad s.t.\ A(p) = b \quad \boldsymbol{p\ dsos/sdsos}$$

Example:

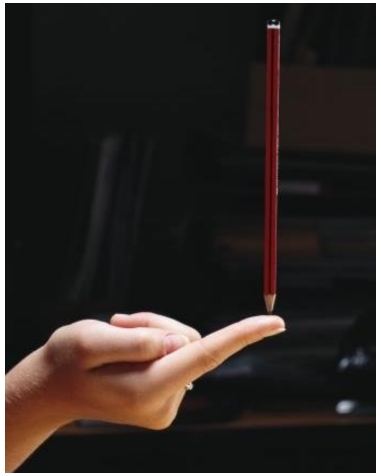
For a parametric family of polynomials:

$$p(x_1, x_2) = 2x_1^4 + 2x_2^4 + ax_1^3x_2 + (1-a)x_1^2x_2^2 + bx_1x_2^3$$

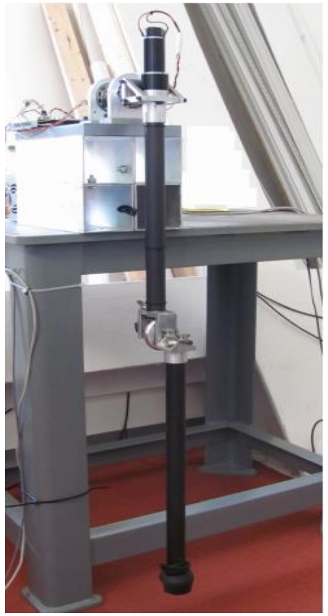


Alternatives to sum of squares: dsos and sdsos

- Example: Stabilizing the inverted N-link pendulum (2N states)

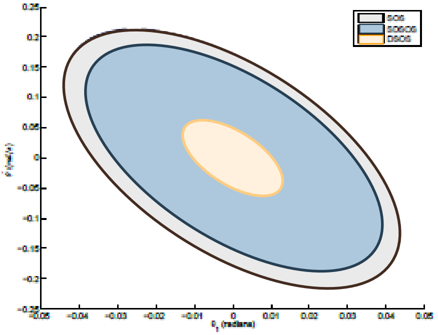


N=1

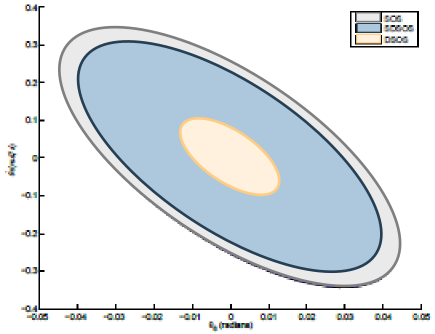


N=2

N=6



(a) θ_1-$\dot{\theta}_1$ subspace.



(b) θ_6-$\dot{\theta}_6$ subspace.

Runtime:

2N (# states)	4	6	8	10	12	14	16	18	20	22
DSOS	< 1	0.44	2.04	3.08	9.67	25.1	74.2	200.5	492.0	823.2
SDSOS	< 1	0.72	6.72	7.78	25.9	92.4	189.0	424.74	846.9	1275.6
SOS (SeDuMi)	< 1	3.97	156.9	1697.5	23676.5	∞	∞	∞	∞	∞
SOS (MOSEK)	< 1	0.84	16.2	149.1	1526.5	∞	∞	∞	∞	∞

ROA volume ratio:

2N (states)	4	6	8	10	12
ρ_{dsos}/ρ_{sos}	0.38	0.45	0.13	0.12	0.09
ρ_{sdsos}/ρ_{sos}	0.88	0.84	0.81	0.79	0.79

[Ahmadi, Majumdar, Tedrake]

Improvements on dsos and sdsos

Replacing sos polynomials by dsos/sdsos polynomials:

- +: fast bounds
- - : not always as good quality (compared to sos)

Iteratively construct a sequence of improving LP/SOCPs

Initialization: Start with the dsos/sdsos polynomials

Method 1: Cholesky change of basis

Method 2: Column Generation

Method 3: r-s/dsos hierarchy

PRINCETON UNIVERSITY

Method 1: Cholesky change of basis (1/3)

$$p(x) = x_1^4 - 6x_1^3x_2 + 2x_1^3x_3 + 6x_1^2x_3^2 + 9x_1^2x_2^2 - 6x_1^2x_2x_3 - 14x_1x_2x_3^2 + 4x_1x_3^3 + 5x_3^4 - 7x_2^2x_3^2 + 16x_2^4$$

$$p(x) = z^T(x)Qz(x)$$

$$Q = \begin{pmatrix} 1 & -3 & 0 & 1 & 0 & 2 \\ -3 & 9 & 0 & -3 & 0 & -6 \\ 0 & 0 & 16 & 0 & 0 & -4 \\ 1 & -3 & 0 & 2 & -1 & 2 \\ 0 & 0 & 0 & -1 & 1 & 0 \\ 2 & -6 & 4 & 2 & 0 & 5 \end{pmatrix}$$

psd but not dd

$$z(x) = (x_1^2, x_1x_2, x_2^2, x_1x_3, x_2x_3, x_3^2)^T$$

$$p(x) = \tilde{z}^T(x) \begin{pmatrix} \frac{1}{2} & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 4 \end{pmatrix} \tilde{z}(x)$$

dd in the “right basis”

$$\tilde{z}(x) = \begin{pmatrix} 2x_1^2 - 6x_1x_2 + 2x_1x_3 + 2x_3^2 \\ x_1x_3 - x_2x_3 \\ x_2^2 - \frac{1}{4}x_3^2 \end{pmatrix}$$

Goal: iteratively improve on basis $z(x)$.

PRINCETON UNIVERSITY

Method 1: Cholesky change of basis (2/3)

Sos problem

$\min C(p)$
$s.t.\ A(p) = b$
$\boldsymbol{p\ sos}$

Initialize

$\min C(p)$
$s.t.\ A(p) = b$
$\boldsymbol{p = z(x)^T Q z(x)},$
$\boldsymbol{Q\ dd/sdd}$

Step 1

Replace:
$U_k = chol(U_{k-1}^T Q^* U_{k-1})$

Step 2

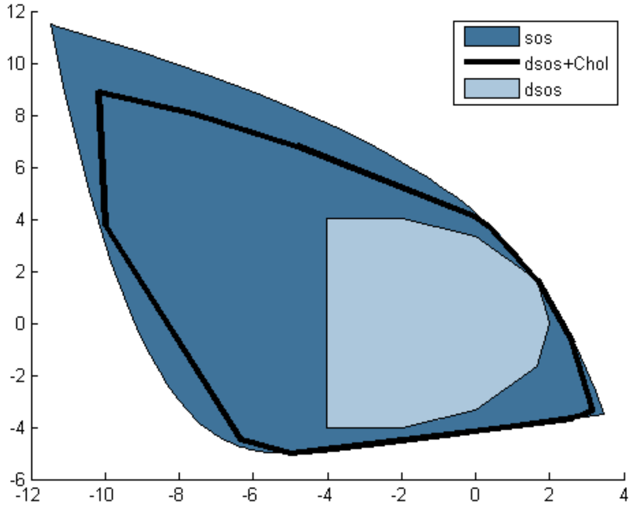
$\min C(p)$
$s.t.\ A(p) = b$
$\boldsymbol{p = z(x)^T U_k^T Q U_k z(x)},$
$\boldsymbol{Q\ dd/sdd}$

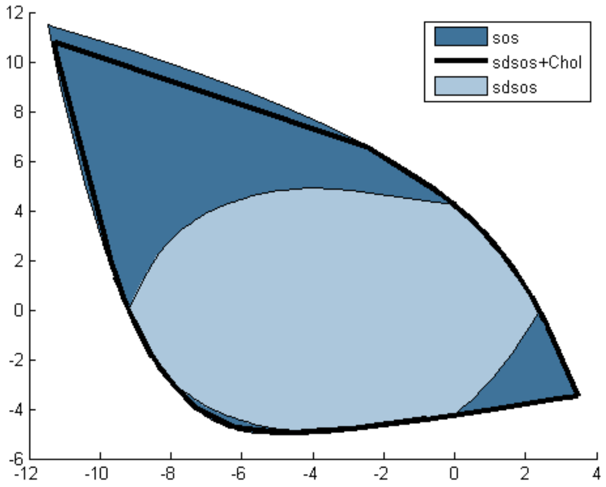
New basis

$\boldsymbol{k := k+1}$

One iteration of this method on a parametric family of polynomials:

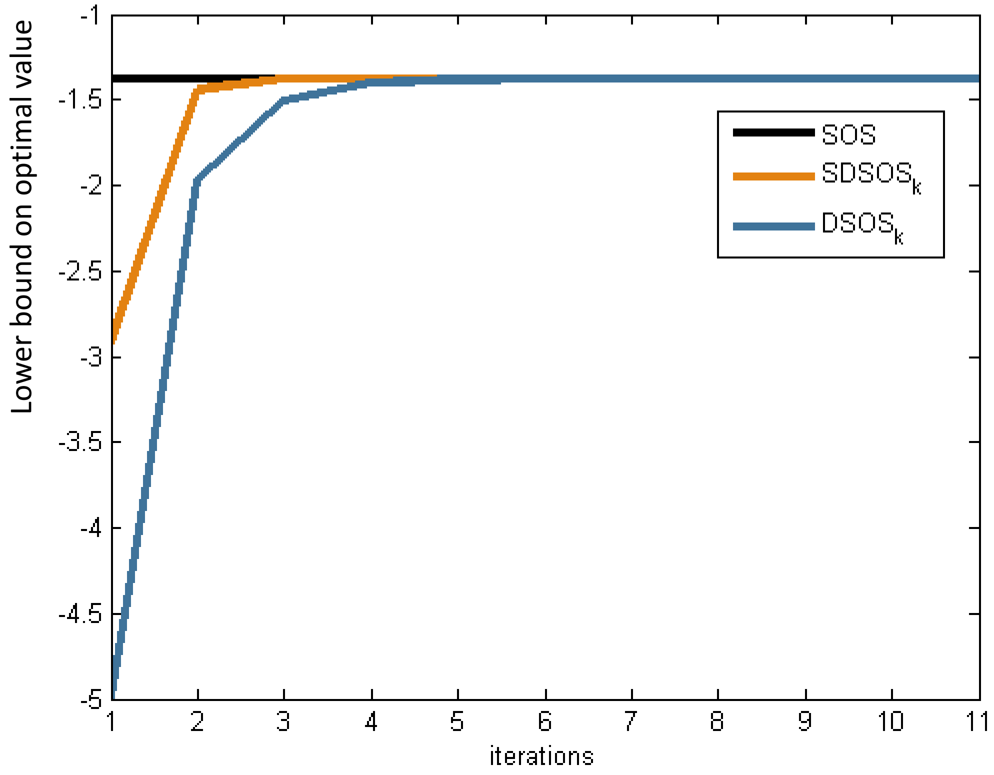
$$p(x_1, x_2) = 2x_1^4 + 2x_2^4 + a x_1^3 x_2 + (1-a) x_1^2 x_2^2 + b x_1 x_2^3$$





Method 1: Cholesky change of basis (3/3)

- Example: minimizing a degree-4 polynomial in 4 variables



PRINCETON UNIVERSITY

Method 2: Column generation (1/4)

- Focus on **LP-based version** of this method (SOCP is similar).

Sos problem	Replace with	Dsos problem		Dsos problem
$\min C(p)$ $s.t.\ A(p) = b$ $\boldsymbol{p\ sos}$	→	$\min C(p)$ $s.t.\ A(p) = b$ $\boldsymbol{p\ dsos}$	⟺	$\min C(p)$ $s.t.\ A(p) = b$ $\boldsymbol{p(x) = z(x)^T Q z(x), Q\ dd}$

Two different ways of characterizing $\boldsymbol{Q\ dd}$:

① Q dd $\Leftrightarrow Q_{ii} \geq \sum_j |Q_{ij}|, \forall\ i$

② Q dd $\Leftrightarrow \exists \alpha_i \geq 0$ s.t. $Q = \sum_i \alpha_i v_i v_i^T$, where v_i **fixed vector** with at most two nonzero components $= \pm 1$

[In collaboration with Sanjeeb Dash, IBM Research]

Method 2: Column generation (2/4)

Dsos problem

$$\min C(p)$$
$$s.t.\ A(p) = b$$
$$\boldsymbol{p(x) = z(x)^T Q z(x), Q\ dd}$$

Using ②

Dsos problem

$$\min C(p)$$
$$s.t.\ A(p) = b$$
$$\boldsymbol{p(x) = \sum_i \alpha_i \left(v_i^T z(x)\right)^2, \alpha_i \geq 0}$$

Idea behind the algorithm:

Expand feasible space at each iteration by adding a new vector v and variable α

$$\min C(p)$$
$$s.t.\ A(p) = b$$
$$\boldsymbol{p(x) = \sum_i \alpha_i \left(v_i^T z(x)\right)^2 + \alpha \left(v^T z(x)\right)^2 \quad \alpha_i \geq 0, \alpha \geq 0}$$

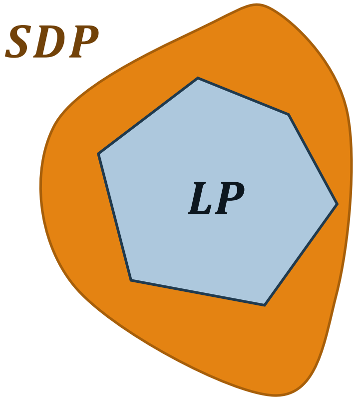
Question: How to pick v? Use the dual!

Method 2: Column generation (3/4)

PRIMAL

A general SDP

$$\max_{y \in \mathbb{R}^m} b^T y$$
$$s.t.\ \ C - \sum_{i=1}^{m} y_i A_i \succcurlyeq 0$$



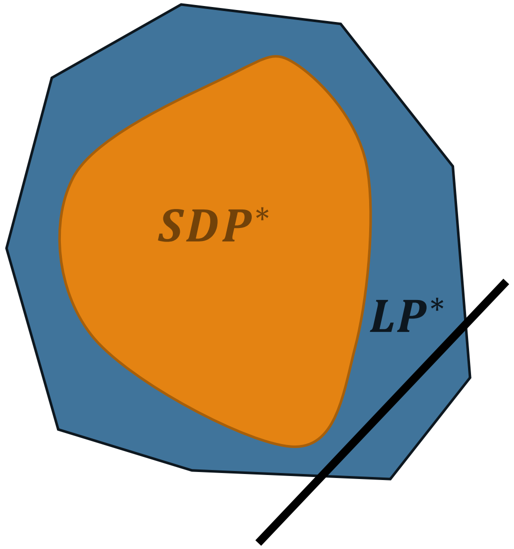
LP obtained with inner approximation of PSD by DD

$$\max_{y \in \mathbb{R}^m} b^T y$$
$$s.t.\ \ C - \sum_{i=1}^{m} y_i A_i = \sum \alpha_i v_i v_i^T$$
$$\alpha_i \geq 0$$

DUAL

Dual of SDP

$$\min_{X \in S^n} tr(CX)$$
$$s.t.\ \ tr(A_i X) = b_i$$
$$X \succcurlyeq 0$$



Dual of LP

$$\min_{X \in S^n} tr(CX)$$
$$s.t.\ \ tr(A_i X) = b_i$$
$$v_i^T X v_i \geq 0$$

Pick v s.t. $v^T X v < 0$.

Method 2: Column Generation (4/4)

- Example 2: minimizing a degree-4 polynomial

	$n = 15$		$n = 20$		$n = 25$		$n = 30$		$n = 40$	
	bd	t(s)	bd	t(s)	bd	t(s)	bd	t(s)	bd	t(s)
DSOS	-10.96	0.38	-18.01	0.74	-26.45	15.51	-36.52	7.88	-62.30	10.68
$DSOS_k$	-5.57	31.19	-9.02	471.39	-20.08	600	-32.28	600	-35.14	600
SOS	-3.26	5.60	-3.58	82.22	-3.71	1068.66	NA	NA	NA	NA

PRINCETON UNIVERSITY

Method 3: r-s/dsos hierarchy (1/3)

- A polynomial p is **r-dsos** if $p(x)\left(\sum_i x_i^2\right)^r$ is **dsos**.
- A polynomial p is **r-sdsos** if $p(x)\left(\sum_i x_i^2\right)^r$ is **sdsos**.

Defines a hierarchy based on r.

Theorem

Any **even** positive definite form $\boldsymbol{p}$ **is r-dsos** for some $\boldsymbol{r}$**.**

Proof: Follows from a result by Polya.

Proof of positivity using LP.

[Ahmadi, Majumdar]

Method 3: r-s/dsos hierarchy (2/3)

- **Example:** certifying stability of a switched linear system $x_{k+1} = A_{\sigma(k)} x_k$ where $A_{\sigma(k)} \in \{A_1, \dots, A_m\}$

Recall:

Theorem 1: A switched linear system is stable if and only if $\rho(A_1 \dots, A_m) < 1$.

Theorem 2 [Parrilo, Jadbabaie]:
$\rho(A_1, \dots, A_m) < 1$
$\Leftrightarrow$
$\exists$ a pd polynomial Lyapunov function $V(x)$ such that $V(x) - V(A_i x) > 0, \forall x \neq 0$.

Method 3: r-s/dsos hierarchy (3/3)

Theorem: For nonnegative $\{A_1, \ldots, A_m\}$, $\rho(A_1, \ldots, A_m) < 1 \Leftrightarrow$ $\exists\, r \in \mathbb{N}$ and a polynomial Lyapunov function $V(x)$ such that

$$V(x.^2) \text{ r-dsos and } V(x.^2) - V(A_i x.^2) \text{ r-dsos.} \qquad (\star)$$

Proof:

$(\Leftarrow)$ $(\star) \Rightarrow V(x) \geq 0$ and $V(x) - V(A_i x) \geq 0$ for any $x \geq 0$.

Combined to $A_i \geq 0$, this implies that trajectories of $x_{k+1} = A_{\sigma(k)} x_k$ starting from $x_0 \geq 0$ go to zero.

This can be extended to any x_0 by noting that $x_0 = x_0^+ - x_0^-, x_0^+, x_0^- \geq 0$.

$(\Rightarrow)$ From Theorem 2, and using Polya's result as $V(x.^2)$ and $V(x.^2) - V(A_i x.^2)$ are even forms.

PRINCETON UNIVERSITY

Main messages

- Can construct **iterative inner approximations** of the cone of nonnegative polynomials using LPs and SOCPs.
- Presented three methods:

	Cholesky change of basis	Column Generation	r-s/dsos hierarchies
Initialization	Initialize with dsos/sdsos polynomials		
Method	Rotate existing "atoms" of the cone of dsos/sdsos polynomials	Add new atoms to the extreme rays of the cone of dsos/sdsos polynomials	Use multipliers to certify nonnegativity of more polynomials.
Size of the LP/SOCPs obtained	Does not grow (but possibly denser)	Grows slowly	Grows quickly
Objective taken into consideration	Yes	Yes	No
Can beat the SOS bound	No	No	Yes

Thank you for listening

Questions?

Want to learn more?
http://scholar.princeton.edu/ghall/