Jacobian SDP Relaxation for Polynomial Optimization

Jiawang Nie

Sino-Japan Optimization Meeting

Department of Mathematics, UCSD

September 28, 2011

Multivariate Polynomial Optimization

Given polynomials $f(x), h_i(x), g_j(x)$, solve problem

$$\begin{aligned} \min_{x\in\mathbb{R}^n} \quad & f(x) \\ s.t. \quad & h_1(x) = \cdots = h_r(x) = 0, \\ & g_1(x) \geq 0, \cdots, g_m(x) \geq 0. \end{aligned}$$

There are standard numerical methods for solving the problem globally based on semidefinite programming (SDP) and sum of squares (SOS) approximations (Lasserre, Parrilo, Sturmfels, ...)

Goal of This talk: Jacobian SDP Relaxation, the first method that can compute the global minimum exactly by SDP.

Outline of the Talk

- Some Backgrounds

- Jacobian SDP Relaxation

- Certifying Exactness

- Numerical Examples

SOS polynomials

A poly $p(x)$ is sum of squares (SOS) if $p(x) = \sum q_i^2(x)$.

$$\begin{aligned} \text{Example:} \quad & 3 \cdot \left(x_1^4 + x_2^4 + x_3^4 + x_4^4 - 4x_1x_2x_3x_4\right) \\ = & (x_1^2 - x_2^2 - x_4^2 + x_3^2)^2 + (x_1^2 + x_2^2 - x_4^2 - x_3^2)^2 + \\ & (x_1^2 - x_2^2 - x_3^2 + x_4^2)^2 + 2(x_1x_4 - x_2x_3)^2 + \\ & 2(x_1x_2 - x_3x_4)^2 + 2(x_1x_3 - x_2x_4)^2 \end{aligned}$$

SOS implies nonnegativity, but not conversely.

Theorem (Hilbert, 1888) Every nonnegative poly is SOS iff

$$(\# \, var, degree) = (1, 2d), (*, 2), \text{ or } (2, 4).$$

Hilbert'1 17th Problem: Is every nonnegative poly is a sum of squares of rational functions? (Yes, by Artin).

Testing SOS Membership

A polynomial $p(x)$ is SOS if and only if

$$\exists\, X : \qquad p(x) = [x]_d^T X [x]_d, \quad X = X^T \succeq 0.$$

The X is called a Gram matrix.

$$\begin{aligned} & 2x_1^4 + 2x_1^3 x_2 - x_1^2 x_2^2 + 5x_2^4 \\ = & \begin{bmatrix} x_1^2 \\ x_2^2 \\ x_1 x_2 \end{bmatrix}^T \begin{bmatrix} 2 & 0 & 1 \\ 0 & 5 & 0 \\ 1 & 0 & -1 \end{bmatrix} \begin{bmatrix} x_1^2 \\ x_2^2 \\ x_1 x_2 \end{bmatrix} \\ = & \begin{bmatrix} x_1^2 \\ x_2^2 \\ x_1 x_2 \end{bmatrix}^T \begin{bmatrix} 2 & -\alpha & 1 \\ -\alpha & 5 & 0 \\ 1 & 0 & -1 + 2\alpha \end{bmatrix} \begin{bmatrix} x_1^2 \\ x_2^2 \\ x_1 x_2 \end{bmatrix} \end{aligned}$$

When $\alpha = 3$, the Gram matrix is positive semidefinite.

SOS Program and SDP

- A typical SOS program is:

$$\max_{w \in \mathbb{R}^m} \quad c^T w \quad \text{s.t.} \quad f_0 + w_1 f_1 + \cdots + w_m f_m \text{ is SOS}$$

  where $f_0, f_1, \ldots, f_m$ are given polynomials.

- SOS program is reducible to SDP

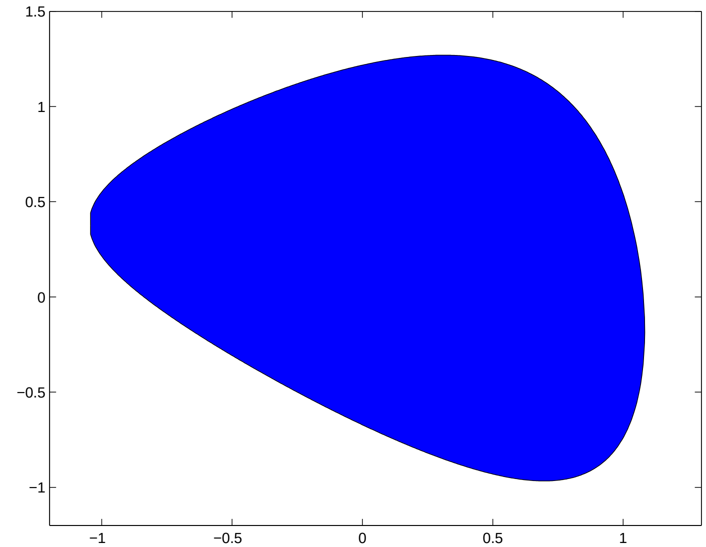
$$\begin{array}{ll} \max\limits_{w \in \mathbb{R}^3} & -w_1 + w_2 - w_3 \\ s.t. & w_1 x_1^4 + 2w_2 x_1^3 x_2 \\ & -x_1^2 x_2^2 + w_3 x_2^4 \\ & \text{is SOS} \end{array} \quad \Leftrightarrow \quad \begin{array}{cc} \max\limits_{w \in \mathbb{R}^3, \alpha \in \mathbb{R}} & -w_1 + w_2 - w_3 \\ s.t. & \begin{bmatrix} w_1 & -\alpha & w_2 \\ -\alpha & w_3 & 0 \\ w_2 & 0 & 2\alpha - 1 \end{bmatrix} \succeq 0 \end{array}$$

Semidefinite Programming (SDP)

SDP has the standard form

$$
\begin{array}{ll}
\text{min} & c_1 x_1 + \cdots + c_n x_n \\
s.t. & A_0 + \sum_{i=1}^{n} x_i A_i \succeq 0
\end{array}
$$

where $A_i \in \mathbb{R}^{M \times M}$ are symmetric.



The feasible region of SDP is convex.
Example:

$$
\begin{bmatrix}
3 - 2x_1 + x_2 & x_1 & -x_1 - x_2 \\
x_1 & 3 + x_1 - 2x_2 & x_2 \\
-x_1 - x_2 & x_2 & 1 + x_1 + x_2
\end{bmatrix} \succeq 0.
$$

Lasserre's Hierarchy of SOS Relaxations

$$
\begin{aligned}
f_{min} := \text{min} \quad & f(x) \\
s.t. \quad & h_i(x) = 0 \, (1 \leq i \leq r) \\
& g_j(x) \geq 0 \, (1 \leq j \leq m)
\end{aligned}
$$

For each integer N (relax. order), solve the SOS program

$$
\begin{aligned}
f_N := \text{max} \quad & \gamma \\
s.t. \quad & f(x) - \gamma - \sum_{i=1}^{r} \phi_i h_i - \sum_{j=1}^{m} \sigma_j g_j \text{ is SOS} \\
& \text{deg}(\phi_i h_i), \text{deg}(\sigma_j g_j) \leq 2N \text{ with } \sigma_j \text{ SOS}
\end{aligned}
$$

We get a sequence of lower bounds for $N = 1, 2, \ldots$

$$f_1 \leq f_2 \leq f_3 \leq \cdots \leq f_{min}.$$

Lasserre's Relaxation is reducible to SDP.

Convergence of Lasserre's Hierarchy

The original prob.

$$f_{min} := \min f(x)$$
$$s.t. \quad h_1(x) = \cdots = h_r(x) = 0$$
$$g_1(x), \ldots, g_m(x) \geq 0$$

N-th Lasserre's relax.

$$f_N := \max \gamma$$
$$s.t. \quad f(x) - \gamma - \sum \phi_i h_i - \sum \sigma_j g_j \text{ sos}$$
$$\deg(\phi_i h_i), \deg(\sigma_j g_j) \leq 2N$$
$$\text{each } \sigma_j(x) \text{ is SOS}$$

Theorem (Lasserre, 2001) Under archimedean condition (AC)

$$\lim_{N \to \infty} f_N \quad = \quad f_{min}.$$

AC is almost equivalent to compactness of the feasible set.

A Negative Result By Scheiderer

Scheiderer (1999) discovered a negative result: whenever the feasible set has dimension ≥ 3, there exists a "bad" polynomial f such that Lasserre's Relax. is never exact:

$$\lim_{N\to\infty} f_N = f_{min}, \qquad \text{but} \qquad f_N < f_{min} \quad \forall\, N.$$

Scheiderer proved a positive one: In the 2-D case, finite convergence holds under some general non-singularity assumptions.

Question: Can we solve polynomial optimization globally and exactly by a single SDP relaxation?

Outline of the Talk

- Some Backgrounds
- Jacobian SDP Relaxation
- Certifying Exactness
- Numerical Examples

1-Equality-1-Inequality (1E1I) Constraints

To describe Jacobian SDP relaxation, consider 1E1I case:

$$
\begin{aligned}
f_{min} := \min_{x \in \mathbb{R}^n} \quad & f(x) \\
s.t. \quad & h(x) = 0, g(x) \geq 0.
\end{aligned}
$$

The N-th Lasserre's SOS relaxation is

$$
\begin{aligned}
f_N := \max \quad & \gamma \\
s.t. \quad & f - \gamma - \phi h - \sigma g \text{ is SOS} \\
& \sigma \text{ is sos, } \deg(\phi h), \deg(\sigma g) \leq 2N.
\end{aligned}
$$

Under the archimedean condition on $\{h = 0, g \geq 0\}$(general for compact sets), we typically have only asymptotic convergence:

$$
\lim_{N \to \infty} f_N = f_{min}, \qquad f_N < f_{min} \quad \forall\, N
$$

First Order Optimality Condition

Suppose x^* is an optimizer of

$$\text{Minimize} \quad f(x) \quad s.t. \quad h(x) = 0,\ g(x) \geq 0.$$

The first order condition is

$$\nabla f(x^*) = \lambda_1 \nabla h(x^*) + \lambda_2 \nabla g(x^*), \quad \lambda_2 g(x^*) = 0.$$

There are two possibilities:

Case I: $g(x^*) > 0$.

Case II: $g(x^*) = 0$.

Case I: $g(x^*) > 0$

Suppose x^* is an optimizer of

$$\text{Minimize} \quad f(x) \quad s.t. \quad h(x) = 0,\ g(x) \geq 0.$$

If $g(x^*) > 0$, the optimality condition is reduced to

$$\nabla f(x^*) = \lambda_1 \nabla h(x^*).$$

Thus, the rank condition

$$\text{rank}\ g(x) \cdot \left[\nabla f(x) \quad \nabla h(x) \right] \leq 1$$

always holds at x^*, no matter $g(x^*) = 0$ or $g(x^*) > 0$.

Case II: $g(x^*) = 0$

Suppose x^* is an optimizer of

$$\text{Minimize} \quad f(x) \quad s.t. \quad h(x) = 0,\ g(x) \geq 0.$$

The optimality condition is

$$\nabla f(x^*) = \lambda_1 \nabla h(x^*) + \lambda_2 \nabla g(x^*), \quad \lambda_2 g(x^*) = 0.$$

Thus, the rank condition

$$\text{rank} \begin{bmatrix} \nabla f(x) & \nabla h(x) & \nabla g(x) \end{bmatrix} \leq 2$$

always holds at x^*, no matter $g(x^*) = 0$ or $g(x^*) > 0$ (if $g(x^*) > 0, \lambda_2 = 0$, the first two columns are dependant).

Characterizing Critical Points

If x^* is an optimizer of

$$\text{Minimize} \quad f(x) \quad s.t. \quad h(x) = 0,\, g(x) \geq 0,$$

then the Jacobian matrices

$$g(x) \cdot \begin{bmatrix} \nabla f(x) & \nabla g_1(x) \end{bmatrix}, \quad \begin{bmatrix} \nabla f(x) & \nabla h(x) & \nabla g(x) \end{bmatrix}$$

are always singular at x^*, no matter $g(x^*) > 0$ or $g(x^*) = 0$.

Theorem (N. 2010) Under some generic nonsingularity conditions, a point x^* is critical if and only if the matrices

$$g(x) \cdot \begin{bmatrix} \nabla f(x) & \nabla g_1(x) \end{bmatrix}, \quad \begin{bmatrix} \nabla f(x) & \nabla h(x) & \nabla g(x) \end{bmatrix}$$

are all singular at x^*.

Jacobian Type SDP relaxation:

Let $\varphi_1(x), \ldots, \varphi_K(x)$ be a minimum set (e.g., by Bruns and Vetter's method) of defining polys for the variety:

$$\left\{x : \text{rank}\left[g \cdot \nabla f \quad g \cdot \nabla g\right] \leq 1, \text{rank}\left[\nabla f \quad \nabla h \quad \nabla g\right] \leq 2\right\}.$$

Then we get an equivalent formulation

$$\begin{array}{rl} \text{Minimize} & f(x) \\ s.t. & h(x) = 0,\, g(x) \geq 0 \end{array} \iff \begin{array}{rl} \text{Minimize} & f(x) \\ s.t. & h(x) = 0,\, g(x) \geq 0 \\ & \varphi_1(x) = \cdots = \varphi_K(x) = 0 \end{array}$$

The N-th Jacobian SDP relaxation is

$$\begin{array}{llll} \max & \gamma & s.t. & f - \gamma - \phi h - \sigma g - \sum_j \phi_j \varphi_j \text{ is SOS} \\ & \sigma \text{ is SOS}, \deg(\phi h), \deg(\sigma g), \deg(\phi_j \varphi_j) \leq 2N. & & \end{array}$$

Exactness of Jacobian SDP relaxation:

For 1E1I polynomial optimization

$$f_{min} := \min \quad f(x) \quad s.t. \quad h(x) = 0,\, g(x) \geq 0,$$

we get a sequence of lower bounds $f_N \leq f_{min}$ from

$$\begin{aligned} f_N := \max \quad & \gamma \quad s.t. \quad f - \gamma - \phi h - \sigma g - \sum_j \phi_j \varphi_j \text{ is SOS} \\ & \sigma \text{ is SOS}, \deg(\phi h), \deg(\sigma g), \deg(\phi_j \varphi_j) \leq 2N. \end{aligned}$$

Theorem (N., 2010) Assume the feasible set is nonsingular and $f(x)$ has a global minimizer, then for all N big enough

$$f_N \quad = \quad f_{min}.$$

Jacobian SDP relaxation for General Case

For general multi-constrained polynomial optimization:

$$\begin{array}{ll} \min & f(x) \\ s.t. & h_i(x) = 0 (1 \leq i \leq r), g_j(x) \geq 0 (1 \leq j \leq m) \end{array}$$

From the fist order optimality condition, we can get redundant polynomial equations by using Jacobians, like 1E1I case.

• We have same finite convergence result, under some generic nonsingularity assumptions on h_i, g_j.

• The sizes of the Jacobian SDP relaxation grow exponentially in m (# of inequality constraints).

• Some efficient variations exist for special cases.

Outline of the Talk

- Some Backgrounds
- Jacobian SDP Relaxation
- Certifying Exactness
- Numerical Examples

How to Certify Exactness?

Jacobian SDP relaxation returns a sequence of lower bounds

$$f_1 \leq \cdots \leq f_N = f_{N+1} = \cdots = f_{min}$$

for the polynomial optimization (1E1I case)

$$f_{min} := \min \quad f(x) \quad s.t. \quad h(x) = 0,\, g(x) \geq 0.$$

The equality $f_N = f_{min}$ is certified if a feasible x^* satisfies

$$f(x^*) \quad = \quad f_N.$$

Can we always find such an x^* by an algorithm?

This is almost always true!

Duality in Jacobian SDP Relaxation

For the 1E1I polynomial optimization

$$\min \quad f(x) \quad s.t. \quad h(x) = 0,\ g(x) \geq 0,$$

the N-th Jacobian relax (SOS version) is:

$$\begin{array}{ll} \max \quad \gamma \quad s.t. & f - \gamma - \phi h - \sigma g - \sum_j \phi_j \varphi_j \text{ is SOS} \\ & \sigma \text{ is SOS}, \deg(\phi h), \deg(\sigma g), \deg(\phi_j \varphi_j) \leq 2N. \end{array}$$

Its dual problem (moment version) is:

$$\begin{array}{ll} \min \quad \sum f_\alpha y_\alpha \quad s.t. & L_h^{(N)}(y) = 0,\ L_{\varphi_j}^{(N)}(y) = 0 (1 \leq j \leq r), \\ & L_g^{(N)}(y) \succeq 0, M_N(y) \succeq 0,\ y_0 = 1. \end{array}$$

where $L_p^{(N)}(y)$ denotes the N-th localizing matrix of a poly p and a moment vector y, and $M_N(y)$ is a moment matrix.

Rank One Case

Suppose y^* is a minimizer of the dual optimization problem:

$$f_N := \min \sum f_\alpha y_\alpha \quad s.t. \quad \begin{array}{l} L_h^{(N)}(y) = 0,\, L_{\varphi_j}^{(N)}(y) = 0, (1 \le j \le r) \\ L_g^{(N)}(y) \succeq 0, M_N(y) \succeq 0,\, y_0 = 1. \end{array}$$

If $\operatorname{rank} M_N(y^*) = 1$, then

$$x^* = (y^*_{e_1}, y^*_{e_2}, \ldots, y^*_{e_n}), \qquad y^* = (1, x_1^*, \ldots, (x_1^*)^2, \ldots)$$

$$L_h^{(N)}(y^*) = 0 \quad \Rightarrow \quad h(x^*) = 0,$$

$$L_g^{(N)}(y^*) \succeq 0 \quad \Rightarrow \quad g(x^*) \ge 0.$$

$$f_{min} \quad \ge \quad f_N = \sum f_\alpha y^*_\alpha \quad = \quad f(x^*).$$

So x^* is feasible and a globally minimizer.

Flat Truncation (FT)

Suppose y^* is a minimizer of the dual optimization problem:

$$\begin{array}{llll} \min & \sum f_\alpha y_\alpha & s.t. & L_h^{(N)}(y) = 0,\, L_{\varphi_j}^{(N)}(y) = 0, (1 \le j \le r) \\ & & & L_g^{(N)}(y) \succeq 0, M_N(y) \succeq 0,\, y_0 = 1. \end{array}$$

We say y^* has a flat truncation (FT) or FT holds at y^* if

$$\operatorname{rank} M_t(y^*) \quad = \quad \operatorname{rank} M_{t-d}(y^*) \qquad \text{for some } t \in [d, N]$$

where $d = \max\{1, \lceil \deg(h)/2 \rceil, \lceil \deg(g)/2 \rceil\}$.

FT $\Rightarrow$ $y^*|_{2t}$ admits a finite measure (Curto-Fialkow).

FT $\Rightarrow$ global minimizers can be found (Henrion-Lasserre).

FT Holds Generally

Suppose y^* is a minimizer of the Jacobian relax.:

$$\min \sum f_\alpha y_\alpha \quad s.t. \quad \begin{array}{l} L_h^{(N)}(y) = 0,\, L_{\varphi_j}^{(N)}(y) = 0, (1 \leq j \leq r) \\ L_g^{(N)}(y) \succeq 0, M_N(y) \succeq 0,\, y_0 = 1. \end{array}$$

Theorem (N. 2011) If the optimization problem

$$\min \quad f(x) \quad s.t. \quad h(x) = 0,\, g(x) \geq 0$$

has finitely many global minimizers, then FT holds for every minimizer y^* of Jacobian relaxation for some N.

If # global minimizers $= \infty$, then FT fails.

Outline of the Talk

- Some Backgrounds
- Jacobian SDP Relaxation
- Certifying Exactness
- Numerical Examples

Example I

$$\min_{x\in\mathbb{R}^3} \quad x_1^4x_2^2+x_1^2x_2^4+x_3^6-3x_1^2x_2^2x_3^2$$
$$\text{s.t.} \quad x_1^2+x_2^2+x_3^2\le 1.$$

The objective is the Motzkin polynomial (nonnegative but not SOS). Its minimum $f_{min}=0$.

By Jacobian relaxation of order 4, we get a lower bound

$$f_{4,Jac} \quad = \quad -1.6948\times 10^{-8} \quad \approx \quad f_{min}.$$

By Lasserre's Relaxation of orders $4,5,6,7,8$, we get lower bounds respectively

$$-2\times 10^{-4}, -2.9\cdot 10^{-5}, -8.2\cdot 10^{-6}, -4.2\cdot 10^{-6}, -2.3\cdot 10^{-6}.$$

Example II

Minimizing Motzkin poly outside unit ball

$$\begin{array}{rl} \min_{x \in \mathbb{R}^3} & x_1^4 x_2^2 + x_1^2 x_2^4 + x_3^6 - 3x_1^2 x_2^2 x_3^2 \\ \text{s.t.} & x_1^2 + x_2^2 + x_3^2 \geq 1. \end{array}$$

The min. $f_{min} = 0$. By Jacobian SDP Relax. of order 4, we get a lower bound (its sign is false due to numerical issues):

$$f_{4,\,Jac} \quad = \quad 1.7633 \cdot 10^{-9} \quad \approx \quad f_{min}.$$

By Lasserre's relaxation of orders $5, 6, 7, 8$, we get lower bounds respectively

$$-4.8567 \cdot 10^5, \quad -98.4862, \quad -0.7079, \quad -0.0277.$$

Jacobian Relax. is stronger than Lasserre's Relax.

Example III

Consider the optimization

$$\min_{x\in\mathbb{R}^2} \quad x_1^2+x_2^2 \quad \text{s.t.} \quad \begin{aligned} & x_2^2-1\geq 0, \\ & x_1^2-Mx_1x_2-1\geq 0, \\ & x_1^2+Mx_1x_2-1\geq 0. \end{aligned}$$

Its minimum $f_{min}=2+\frac{1}{2}M(M+\sqrt{M^2+4})$. Let $M=5$.

By Jacobian Relaxation of order 4, we get a lower bound

$$f_{4,\,Jac} \quad = \quad 27.9629 \quad = \quad f_{min}.$$

By Lasserre's Relaxation, we get lower bounds

$$f_{N,\,Las} \quad = \quad 2 \qquad \forall\, N.$$

THIS IS THE END!

THANK YOU VERY MUCH!