Iterative Valid Polynomial Inequality Generation in Polynomial Optimization

Miguel F. Anjos
Mathématiques et génie industriel
ÉCOLE POLYTECHNIQUE MONTRÉAL
GERAD

Joint work with B. Ghaddar (Waterloo) and J. Vera (Tilburg)

Workshop on Optimization, Fields Institute – 27 September 2011

Polynomial Optimization

Polynomial optimization problems (POPs) consist of optimizing a multivariate polynomial objective subject to multivariate polynomial constraints:

Polynomial Optimization Problem (POP)

$$
\begin{aligned}
z = \quad & \sup \quad f(x) \\
& \text{s.t.} \quad g_i(x) \geq 0 \qquad i = 1, \ldots, m.
\end{aligned}
$$

Numerous classes of problems can be modelled as POPs, including:

- Linear Problems
- Mixed-Binary Problems

$$x_i \in \{0, 1\} \quad \Leftrightarrow \quad x_i(1 - x_i) = 0$$

- Quadratic Problems (Convex / Non-convex)

Thus, solving POPs is in general NP-hard.

Relaxations of POPs

Polynomial Optimization Problem (POP)

$$\begin{aligned} z = \quad & \sup \quad f(x) \\ & \text{s.t.} \quad g_i(x) \geq 0 \qquad i = 1, \ldots, m. \end{aligned}$$

- Many tractable relaxations of POPs have been proposed using linear, second-order cone, and semidefinite techniques.

Relaxations of POPs

Polynomial Optimization Problem (POP)

$$\begin{array}{rll} z = & \sup & f(x) \\ & \text{s.t.} & g_i(x) \geq 0 \qquad i = 1, \ldots, m. \end{array}$$

- Many tractable relaxations of POPs have been proposed using linear, second-order cone, and semidefinite techniques.
- In particular, sum-of-squares (SOS) decompositions which lead to semidefinite programming (SDP) relaxations
  - are theoretically very strong:
    - Sequences of relaxations converging to the optimal value in the limit
    - Exact (exponential-sized) relaxations for pure binary POPs

Relaxations of POPs

Polynomial Optimization Problem (POP)

$$
\begin{aligned}
z = \quad & \sup \quad f(x) \\
& \text{s.t.} \quad g_i(x) \geq 0 \qquad i = 1, \ldots, m.
\end{aligned}
$$

- Many tractable relaxations of POPs have been proposed using linear, second-order cone, and semidefinite techniques.
- In particular, sum-of-squares (SOS) decompositions which lead to semidefinite programming (SDP) relaxations
  - are theoretically very strong:
    - Sequences of relaxations converging to the optimal value in the limit
    - Exact (exponential-sized) relaxations for pure binary POPs
  - but quickly become too expensive for practical computation.

Research objective:

Improve the SDP relaxations

• without incurring an exponential growth in their size

• by iteratively generating valid polynomial inequalities.

General POP Perspective

Given a general POP problem:

$$\text{(POP)} \quad z = \sup \; f(x) \quad \text{s.t.} \; g_i(x) \geq 0 \qquad i = 1, \ldots, m.$$

If λ is the optimal value of POP, then POP is equivalent to

$$\begin{array}{ll} \inf & \lambda \\ \text{s.t.} & \lambda - f(x) \geq 0 \quad \forall x \in S := \{x : g_i(x) \geq 0, \; i = 1, \ldots, m\} \end{array}$$

which we rewrite as

$$\begin{array}{ll} \inf & \lambda \\ \text{s.t.} & \lambda - f(x) \in \mathcal{P}_d(S) \end{array}$$

where

$$\mathcal{P}_d(S) = \{p(x) \in \mathbf{R}_d[x] : p(s) \geq 0 \text{ for all } s \in S\}$$

is the cone of polynomials of degree at most d that are non-negative over S.

Understanding $\mathcal{P}_d(S)$

The set

$$\mathcal{P}_d(S) = \{p(x) \in \mathbb{R}_d[x] : p(x) \geq 0 \text{ for all } x \in S\}$$

is in general a very complex object.

- It is always a convex cone
- In most cases the decision problem for $\mathcal{P}_d(S)$ is NP-hard:

Decision problem for $\mathcal{P}_d(S)$

Given $p(x)$, decide if $p(x) \in \mathcal{P}_d(S)$
(i.e. if $p(x) \geq 0$ for all $x \in S$)

- Idea: use algebraic geometry results to approximate (or represent) $\mathcal{P}_d(S)$ in tractable ways, i.e., using only linear, second-order, and semidefinite cones.

A General Recipe for Relaxations of POP

We relax $\lambda - f(x) \in \mathcal{P}_d(S)$ to

$$\lambda - f(x) \in \mathcal{K} \text{ for a suitable } \mathcal{K} \subseteq \mathcal{P}_d(S).$$

Then

$$\begin{array}{ll} \inf & \lambda \\ \text{s.t.} & \lambda - f(x) \in \mathcal{K} \end{array}$$

provides an upper bound for the original problem.

• The choice of $\mathcal{K}$ is a key factor in obtaining good bounds on the problem.

• We are restricted by the need for the optimization over $\mathcal{K}$ to be tractable.

SOS Approach - Lasserre (2001), Parrilo (2000)

For each $r > 0$, define the approximation $\mathcal{K}_r \subseteq \mathcal{P}_d(S)$ as

$$\mathcal{K}_r := \left(\Psi_r + \sum_{i=1}^{m} g_i(x) \Psi_{r-\deg(g_i)} \right) \cap \mathbf{R}_d[x]$$

where Ψ_d denotes the cone of real polynomials of degree at most d that are SOSs of polynomials, and $\mathbf{R}_d[x]$ denotes the set of polynomials in the variables x of degree at most d.

The corresponding relaxation can be written as

$$\begin{array}{llll} (\mathrm{L}_r) & z_r = \inf\limits_{\lambda,\sigma_i} & \lambda & \\ & \text{s.t.} & \lambda - f(x) = \sigma_0(x) + \sum_{i=1}^{m} \sigma_i(x) g_i(x) & \\ & & \sigma_0(x) \text{ is SOS of degree} \leq r & \\ & & \sigma_i(x) \text{ is SOS of degree} \leq r - \deg(g_i(x)), i = 1, \ldots, m. & \end{array}$$

Solving the SOS Relaxation

For each r, the relaxation (L_r) can be cast as an SDP problem, since $\sigma(x)$ is a SOS of degree $2k$ if and only if

$$\sigma(x) = \begin{pmatrix} 1 \\ \vdots \\ x_i \\ \vdots \\ x_i x_j \\ \vdots \\ \prod_{|k|} x \end{pmatrix}^T M \begin{pmatrix} 1 \\ \vdots \\ x_i \\ \vdots \\ x_i x_j \\ \vdots \\ \prod_{|k|} x \end{pmatrix} \quad \text{with} \quad M \succeq 0.$$

Note that $\Psi_d = \Psi_{d-1}$ for every odd degree d.

Convergence of the SOS Approach

Under mild conditions $z_r \to z$:

Lemma

Suppose that

$$\mathcal{K}_G^d \subseteq \mathcal{K}_G^{d+1} \subseteq \cdots \subseteq \mathcal{K}_G^r \subseteq \mathcal{P}_d(S),$$

where G is a compact semialgebraic set (not necessarily convex) and there exists a real-valued polynomial $u(x)$ with $u(x) \in \sum_{i=0}^m g_i(x)\Psi$ such that $\{u(x) \geq 0\}$ is compact. Then

$$\mathcal{K}_G^r \uparrow \mathcal{P}_d(S) \text{ as } r \to \infty,$$

and therefore

$$z_r \uparrow z \text{ as } r \to \infty.$$

Size of the SOS Relaxation

Good news: (L_r) can be solved using SDP techniques, and under mild conditions, $z_r \to z$.

Bad news: For a problem with n variables and m inequality constraints, the size of the relaxation is:

- One psd matrix of dimension $\binom{n+r}{r}$;
- m psd matrices, each of dimension $\binom{n+r-\deg(g_i)}{r-\deg(g_i)}$
- $\binom{n+r}{r}$ linear constraints.

Size of the SOS Relaxation

Good news: (L_r) can be solved using SDP techniques, and under mild conditions, $z_r \to z$.

Bad news: For a problem with n variables and m inequality constraints, the size of the relaxation is:

- One psd matrix of dimension $\binom{n+r}{r}$;
- m psd matrices, each of dimension $\binom{n+r-\deg(g_i)}{r-\deg(g_i)}$
- $\binom{n+r}{r}$ linear constraints.

One way around this difficulty is to exploit any available structure (sparsity, symmetry) to solve smaller SDP problems.
Much progress has been made in this direction.

Size of the SOS Relaxation

Good news: (L_r) can be solved using SDP techniques, and under mild conditions, $z_r \to z$.

Bad news: For a problem with n variables and m inequality constraints, the size of the relaxation is:

- One psd matrix of dimension $\binom{n+r}{r}$;
- m psd matrices, each of dimension $\binom{n+r-\deg(g_i)}{r-\deg(g_i)}$
- $\binom{n+r}{r}$ linear constraints.

One way around this difficulty is to exploit any available structure (sparsity, symmetry) to solve smaller SDP problems.
Much progress has been made in this direction.

Our objective

Avoid the blow-up by keeping r constant (and small).

A Small Example

$$
\begin{aligned}
\inf_{x,y} \quad & (x-1)^2 + (y-1)^2 \\
\text{s.t.} \quad & x^2 - 4xy - 1 \geq 0 \\
& yx - 3 \geq 0 \\
& y^2 - 4 \geq 0 \\
12^2 - (x-2)^2 - 4(y-1)^2 & \geq 0
\end{aligned}
$$

A Small Example

$$
\begin{aligned}
\inf_{x,y} \quad & (x-1)^2 + (y-1)^2 \\
\text{s.t.} \quad & x^2 - 4xy - 1 \geq 0 \\
& yx - 3 \geq 0 \\
& y^2 - 4 \geq 0 \\
& 12^2 - (x-2)^2 - 4(y-1)^2 \geq 0
\end{aligned}
$$

L_2 relaxation

$$
\begin{aligned}
\sup_{\lambda, \sigma_i(\cdot)} \quad & \lambda \\
\text{s.t.} \quad & (x-1)^2 + (y-1)^2 - \lambda = \sigma_0(x,y) + \textstyle\sum_{i=1}^{4} \sigma_i(x,y) g_i(x,y) \\
& \\
& \sigma_0(x,y) \text{ is SOS of degree 2} \\
& \sigma_i(x,y) \text{ is SOS of degree 0}
\end{aligned}
$$

A Small Example

L_2 relaxation

$$\begin{array}{ll} \sup_{\lambda,\sigma_i(\cdot)} & \lambda \\ \text{s.t.} & (x-1)^2 + (y-1)^2 - \lambda = \sigma_0(x,y) + \sum_{i=1}^4 \sigma_i(x,y) g_i(x,y) \\ & \quad (6 \times 14 \text{ lin. system }) \\ & \sigma_0(x,y) \text{ is SOS of degree 2 } (3 \times 3 \text{ matrix}) \\ & \sigma_i(x,y) \text{ is SOS of degree 0 } (4 \text{ non-negative constants}) \end{array}$$

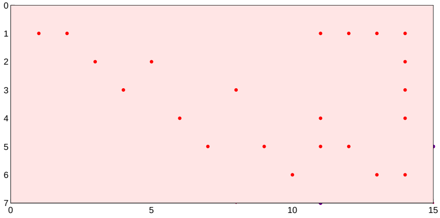


Figure: Structure of the linear system for L_2

A Small Example

L_2 relaxation (Optimal value: 9.4083)

$$\sup_{\lambda,\sigma_i(\cdot)} \quad \lambda$$

s.t. $(x-1)^2 + (y-1)^2 - \lambda = \sigma_0(x,y) + \sum_{i=1}^{4} \sigma_i(x,y) g_i(x,y)$

(6 × 14 lin. system)

$\sigma_0(x,y)$ is SOS of degree 2 (3 × 3 matrix)

$\sigma_i(x,y)$ is SOS of degree 0 (4 non-negative constants)

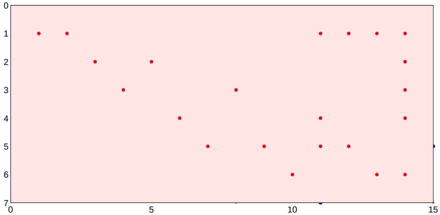


Figure: Structure of the linear system for L_2

Example (ctd)

L_4 relaxation (Optimal value: 36.0654)

$$\sup_{\lambda,\sigma_i(\cdot)} \quad \lambda$$

$$\text{s.t.} \quad (x-1)^2 + (y-1)^2 - \lambda = \sigma_0(x,y) + \sum_{i=1}^{4} \sigma_i(x,y) g_i(x,y)$$

(15×73 lin. system)

$\sigma_0(x,y)$ is SOS of degree 4 (6×6 matrix)

$\sigma_i(x,y)$ is SOS of degree 2 (3×3 SDP matrices)

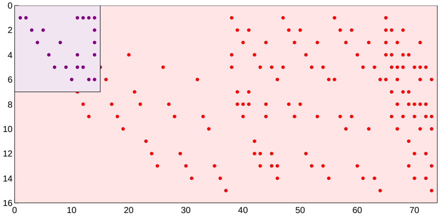


Figure: Structure of the linear system for L_4

Example (ctd)

L_6 relaxation (Optimal value: 51.7386)

$$\begin{array}{rl} \sup_{\lambda,\sigma_i(\cdot)} & \lambda \\ \text{s.t.} & (x-1)^2+(y-1)^2-\lambda = \sigma_0(x,y)+\sum_{i=1}^4 \sigma_i(x,y)g_i(x,y) \end{array}$$

(28×245 lin. system)

$\sigma_0(x,y)$ is SOS of degree 6 (10×10 matrix)

$\sigma_i(x,y)$ is SOS of degree 4 (6×6 SDP matrices)

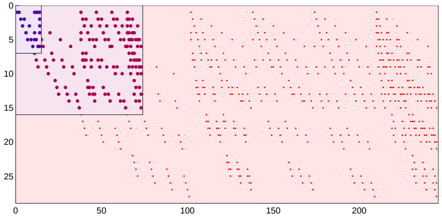


Figure: Structure of the linear system for L_6

Lasserre's Hierarchy for our Example

To solve

$$\begin{array}{ll} \inf_{x,y} & (x-1)^2+(y-1)^2 \\ \text{s.t.} & x^2-4xy-1 \geq 0 \\ & yx-3 \geq 0 \\ & y^2-4 \geq 0 \\ & 12^2-(x-2)^2-4(y-1)^2 \geq 0 \end{array}$$

r	2	4	6
# vars	14	73	245
# constraints	6	15	28
Bound	9.40	36.06	51.73

Lasserre's Hierarchy for our Example

To solve

$$\begin{array}{ll} \inf_{x,y} & (x-1)^2 + (y-1)^2 \\ \text{s.t.} & x^2 - 4xy - 1 \geq 0 \\ & yx - 3 \geq 0 \\ & y^2 - 4 \geq 0 \\ \multicolumn{2}{c}{12^2 - (x-2)^2 - 4(y-1)^2 \geq 0} \end{array}$$

r	2	4	6
# vars	14	73	245
# constraints	6	15	28
Bound	9.40	36.06	51.73

There is no need to run relaxations for $r > 6$, because an optimal solution (and optimality certificate) can be extracted from solution to L_6.

Improving the approximation without growing r

Recall

$$\text{(POP)} \quad z = \sup \; f(x) \quad \text{s.t.} \; x \in S := \{x : g_i(x) \geq 0,\ i = 1, \ldots, m\}$$

$$(\mathrm{L}_r(G)) \quad z_r(G) = \inf_{\lambda, \sigma_i} \; \lambda$$
$$\text{s.t.} \quad \lambda - f(x) = \sigma_0(x) + \sum_{i=1}^m \sigma_i(x) g_i(x)$$
$$\sigma_0(x) \text{ is SOS of degree} \leq r$$
$$\sigma_i(x) \text{ is SOS of degree} \leq r - \deg(g_i(x)),$$
$$i = 1, \ldots, m.$$

Observe that

- (L_r) is defined in terms of the functions used to describe S
- Call this set $G = \{g_i(x) : i = 1, \ldots, m\}$

Goal

Improve our description of S by growing G in such a way that the bound obtained from L_r improves, for fixed r.

Back to our Example

We start with

$$G = \{x^2 - 4xy - 1, yx - 3, y^2 - 4, 12^2 - (x-2)^2 - 4(y-1)^2\}$$

Back to our Example

We start with

$$G = \{x^2 - 4xy - 1, yx - 3, y^2 - 4, 12^2 - (x-2)^2 - 4(y-1)^2\}$$

- For all $(x, y) \in S$,

$$p_1(x, y) = 0.079x^2 + 0.072xy + 0.325x - 0.850y^2 - 0.339y - 0.213 \geq 0$$

- We say that $p_1(x, y)$ is a valid (polynomial) inequality for S.

Back to our Example

We start with

$$G = \{x^2 - 4xy - 1, yx - 3, y^2 - 4, 12^2 - (x-2)^2 - 4(y-1)^2\}$$

- For all $(x, y) \in S$,

$$p_1(x, y) = 0.079x^2 + 0.072xy + 0.325x - 0.850y^2 - 0.339y - 0.213 \geq 0$$

- We say that $p_1(x, y)$ is a valid (polynomial) inequality for S.

Let $G_1 = G \cup \{p_1(x, y)\}$

Then

$$z_2(G_1) = 22.8393 > 9.4083 = z_2(G)$$

Why Stop at p_1?

Add valid inequalities iteratively

- Start with $G_0 = G$.
- Given G_i, generate p_i valid (inequality) for S. Let $G_{i+1} = G_i \cup \{p_i\}$.

Why Stop at p_1?

Add valid inequalities iteratively

- Start with $G_0 = G$.
- Given G_i, generate p_i valid (inequality) for S. Let $G_{i+1} = G_i \cup \{p_i\}$.

$$p_1(x, y) = 0.079x^2 + 0.072xy + 0.325x - 0.850y^2 - 0.339y - 0.213$$

i	0	1
	9.4083	22.8393

Why Stop at p_1?

Add valid inequalities iteratively

- Start with $G_0 = G$.
- Given G_i, generate p_i valid (inequality) for S. Let $G_{i+1} = G_i \cup \{p_i\}$.

$$p_2(x, y) = 0.053x^2 + 0.082xy + 0.205x - 0.764y^2 - 0.533y - 0.282$$

i	0	1	2
	9.4083	22.8393	30.1062

Why Stop at p_1?

Add valid inequalities iteratively

- Start with $G_0 = G$.
- Given G_i, generate p_i valid (inequality) for S. Let $G_{i+1} = G_i \cup \{p_i\}$.

$$p_3(x, y) = 0.069x^2 + 0.002xy - 0.239x - 0.770y^2 + 0.551y - 0.200$$

i	0	1	2	3
	9.4083	22.8393	30.1062	32.2653

Why Stop at p_1?

Add valid inequalities iteratively

- Start with $G_0 = G$.
- Given G_i, generate p_i valid (inequality) for S. Let $G_{i+1} = G_i \cup \{p_i\}$.

$$p_4(x,y) = -0.019x^2 + 0.338xy + 0.097x - 0.691y^2 - 0.577y - 0.254$$

i	0	1	2	3	4
	9.4083	22.8393	30.1062	32.2653	40.1754

Why Stop at p_1?

Add valid inequalities iteratively

- Start with $G_0 = G$.
- Given G_i, generate p_i valid (inequality) for S. Let $G_{i+1} = G_i \cup \{p_i\}$.

$$p_5(x, y) = 0.070x^2 + 0.071xy - 0.158x - 0.858y^2 - 0.425y - 0.214$$

i	0	1	2	3	4
	9.4083	22.8393	30.1062	32.2653	40.1754

i	5
	43.1587

Why Stop at p_1?

Add valid inequalities iteratively

- Start with $G_0 = G$.
- Given G_i, generate p_i valid (inequality) for S. Let $G_{i+1} = G_i \cup \{p_i\}$.

$$p_6(x, y) = 0.052x^2 + 0.047xy + 0.012x - 0.935y^2 - 0.130y - 0.321$$

i	0	1	2	3	4
	9.4083	22.8393	30.1062	32.2653	40.1754

i	5	6
	43.1587	49.3414

Why Stop at p_1?

Add valid inequalities iteratively

- Start with $G_0 = G$.
- Given G_i, generate p_i valid (inequality) for S. Let $G_{i+1} = G_i \cup \{p_i\}$.

$$p_7(x, y) = 0.046x^2 + 0.006xy - 0.182x - 0.707y^2 + 0.652y - 0.195$$

i	0	1	2	3	4
	9.4083	22.8393	30.1062	32.2653	40.1754

i	5	6	7
	43.1587	49.3414	51.5485

Why Stop at p_1?

Add valid inequalities iteratively

- Start with $G_0 = G$.
- Given G_i, generate p_i valid (inequality) for S. Let $G_{i+1} = G_i \cup \{p_i\}$.

$$p_8(x, y) = 0.023x^2 + 0.093xy + 0.116x - 0.566y^2 - 0.621y - 0.519$$

i	0	1	2	3	4
	9.4083	22.8393	30.1062	32.2653	40.1754

i	5	6	7	8
	43.1587	49.3414	51.5485	51.7135

Why Stop at p_1?

Add valid inequalities iteratively

- Start with $G_0 = G$.
- Given G_i, generate p_i valid (inequality) for S. Let $G_{i+1} = G_i \cup \{p_i\}$.

$$p_9(x, y) = 0.049x^2 + 0.002xy + 0.117x - 0.647y^2 - 0.592y - 0.463$$

i	0	1	2	3	4
	9.4083	22.8393	30.1062	32.2653	40.1754

i	5	6	7	8	9
	43.1587	49.3414	51.5485	51.7135	51.7382

Generating Valid Inequalities for POPs

Recall

$$\begin{aligned} z = \inf \quad & \lambda \\ \text{s.t.} \quad & \lambda - f(x) \in \mathcal{P}_d(S) \\ z_r(G) = \inf \quad & \lambda \\ \text{s.t.} \quad & \lambda - f(x) \in \mathcal{K}_r(G) \end{aligned}$$

Lemma

Let G be a description for S, and let p(x) be a valid inequality for S. Then

$$z_r(G \cup \{p(x)\}) \geq z_r(G)$$

Generating Valid Inequalities for POPs

Recall

$$\begin{aligned} z = \inf \quad & \lambda \\ \text{s.t.} \quad & \lambda - f(x) \in \mathcal{P}_d(S) \end{aligned}$$

$$\begin{aligned} z_r(G) = \inf \quad & \lambda \\ \text{s.t.} \quad & \lambda - f(x) \in \mathcal{K}_r(G) \end{aligned}$$

Lemma

Let G be a description for S, and let $p(x)$ be a valid inequality for S. Then

$$z_r(G \cup \{p(x)\}) \geq z_r(G)$$

How to generate a valid improving inequality?

Given a description G of S, find $p(x)$ valid for S such that

$$z_r(G \cup \{p(x)\}) > z_r(G)$$

Valid Inequality Generation for POPs

Goal

Given a description G of S, find $p(x) \in \mathcal{P}_d(S) \setminus \mathcal{K}_r(G)$

Issues to address:

1. Generate $p(x) \in \mathcal{P}_d(S)$.
2. Ensure that $p(x) \subsetneq \mathcal{K}_r(G)$

Valid Inequality Generation for POPs

Goal

Given a description G of S, find $p(x) \in \mathcal{P}_d(S) \setminus \mathcal{K}_r(G)$

Issues to address:

1. Generate $p(x) \in \mathcal{P}_d(S)$.
2. Ensure that $p(x) \subsetneq \mathcal{K}_r(G)$

Issue 1: Generate $p(x) \in \mathcal{P}_d(S)$

There is no tractable representation for $\mathcal{P}_d(S)$

Valid Inequality Generation for POPs

Goal

Given a description G of S, find $p(x) \in \mathcal{P}_d(S) \setminus \mathcal{K}_r(G)$

Issues to address:

1. Generate $p(x) \in \mathcal{P}_d(S)$.
2. Ensure that $p(x) \subsetneq \mathcal{K}_r(G)$

Issue 1: Generate $p(x) \in \mathcal{P}_d(S)$

There is no tractable representation for $\mathcal{P}_d(S)$

- Sol: Generate $p(x) \in \mathcal{K}_{r+2}(G) \cap \mathbf{R}_r[x] \subset \mathcal{P}_d(S)$.

Valid Inequality Generation for POPs

Goal

Given G describing S, find $p(x) \in \mathcal{P}_d(S) \setminus \mathcal{K}_r(G)$

Issues to address:

1. Generate $p(x) \in \mathcal{P}_d(S)$.
2. Ensure that $p(x) \notin \mathcal{K}_r(G)$

Issue 2: Ensure $p(x) \notin \mathcal{K}_r(G)$

Valid Inequality Generation for POPs

Goal

Given G describing S, find $p(x) \in \mathcal{P}_d(S) \setminus \mathcal{K}_r(G)$

Issues to address:

1. Generate $p(x) \in \mathcal{P}_d(S)$.
2. Ensure that $p(x) \notin \mathcal{K}_r(G)$

Issue 2: Ensure $p(x) \notin \mathcal{K}_r(G)$

- Let Y be the dual optimal solution of $\mathrm{L}_r(G)$
- Then $Y \in \mathcal{K}_r(G)^*$
- and therefore $p \in \mathcal{K}_r(G) \Rightarrow \langle p, Y \rangle \geq 0$.

$\Rightarrow$ Look for $p(x)$ such that $\langle p, Y \rangle < 0$.

Inequality Generating Subproblem

Given G and Y

$$\begin{array}{ll} \min & \langle p, Y \rangle \\ \text{s.t.} & p(x) \in \mathcal{K}_{r+2}(G) \\ & \|p\| = 1 \end{array}$$

Inequality Generating Subproblem

Given G and Y

$$\begin{array}{ll} \min & \langle p, Y \rangle \\ \text{s.t.} & p(x) \in \mathcal{K}_{r+2}(G) \\ & \|p\| = 1 \end{array}$$

The normalization is necessary, otherwise the problem is unbounded

- For any $c > 0$, $p(x) \geq 0 \Leftrightarrow cp(x) \geq 0$.

Another Small Example

Optimal value = 0

$$\begin{aligned} \min \quad & x_1 - x_1x_3 - x_1x_4 + x_2x_4 + x_5 - x_5x_7 - x_5x_8 + x_6x_8 \\ \text{s.t.} \; & x_3 + x_4 \le 1 \\ & x_7 + x_8 \le 1 \\ & 0 \le x_i \le 1 \qquad \forall i \in \{1, \cdots, 8\}. \end{aligned}$$

Another Small Example

Optimal value = 0

$$\begin{aligned} \min \quad & x_1 - x_1x_3 - x_1x_4 + x_2x_4 + x_5 - x_5x_7 - x_5x_8 + x_6x_8 \\ \text{s.t. } & x_3 + x_4 \leq 1 \\ & x_7 + x_8 \leq 1 \\ & 0 \leq x_i \leq 1 \qquad \forall i \in \{1, \cdots, 8\}. \end{aligned}$$

r	2	4	6	8
Objective val.	unb.	-0.03550	-0.00192	-
Time (s)	1.02	2.81	726.50	> 18000

Table: Lasserre's Hierarchy

Another Small Example

Optimal value = 0

$$\begin{aligned} \min \quad & x_1 - x_1x_3 - x_1x_4 + x_2x_4 + x_5 - x_5x_7 - x_5x_8 + x_6x_8 \\ \text{s.t. } & x_3 + x_4 \le 1 \\ & x_7 + x_8 \le 1 \\ & 0 \le x_i \le 1 \qquad \forall i \in \{1, \cdots, 8\}. \end{aligned}$$

r	2	4	6	8
Objective val.	unb.	-0.03550	-0.00192	-
Time (s)	1.02	2.81	726.50	> 18000

Table: Lasserre's Hierarchy

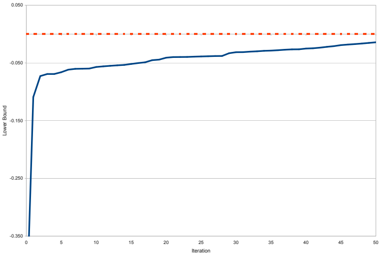
Iter.	0	1	2	3	4	5	10	50
Objective val.	unb.	-0.109	-0.073	-0.069	-0.068	-0.066	-0.057	-0.014
Time (s)								
Subproblem	-	1.5	1.8	2.1	1.9	2.0	2.5	
Master problem	0.2	0.3	0.3	0.4	0.5	0.5	0.6	
Cumulative	0.2	2.0	4.2	6.7	9.2	11.8	26.1	200.1

Table: Inequality Generation

Another Small Example

Optimal value = 0

$$\begin{aligned} \min \quad & x_1 - x_1x_3 - x_1x_4 + x_2x_4 + x_5 - x_5x_7 - x_5x_8 + x_6x_8 \\ \text{s.t.} \quad & x_3 + x_4 \leq 1 \\ & x_7 + x_8 \leq 1 \\ & 0 \leq x_i \leq 1 \qquad \forall i \in \{1, \cdots, 8\}. \end{aligned}$$



The Motzkin Polynomial

Optimal value = 0

$$\min_{x,y,z\in\mathbb{R}} \quad x^2y^2(x^2+y^2-3z^2)+z^6$$

The Motzkin Polynomial

Optimal value = 0

$$\min_{x,y,z\in\mathbb{R}} \quad x^2y^2(x^2+y^2-3z^2)+z^6$$

The Lasserre relaxations are unbounded for all values of r (since the Motzkin polynomial is not a SOS).

The Motzkin Polynomial

Optimal value = 0

$$\min_{x,y,z\in\mathbb{R}} \quad x^2y^2(x^2+y^2-3z^2)+z^6$$

The Lasserre relaxations are unbounded for all values of r (since the Motzkin polynomial is not a SOS).

Iter.	0	1	2	3	4	5	10
Objective val.	unb.	-8591.8	-5687.1	-663.8	-643.8	-640.7	-613.5
Time (s)							
Subproblem	-	0.4	0.3	0.3	0.4	0.4	0.5
Master problem	0.3	0.3	0.3	0.3	0.3	0.3	0.3
Cumulative	0.3	1.0	1.6	2.2	2.9	3.6	7.5

Table: Inequality Generation

Special Case: Binary Quadratic POPs

Consider the general binary quadratic POP:

$$\begin{array}{lll} \max & f(x) & \\ \text{s.t.} & g_i(x) \geq 0 & \forall i \in I = \{1, \ldots, m\} \\ & x \in \{-1, 1\}^n. & \end{array}$$

where $f(x)$ and $g_i(x)$ are polynomials of degree at most 2.

We write the following equivalent formulation:

$$\begin{array}{ll} \min & \lambda \\ \text{s.t.} & \lambda - f(x) \in \mathcal{P}_2(S \cap \{-1, 1\}^n) \end{array}$$

where $S = \{x : g_i(x) \geq 0\}$.

Valid Inequality Generation for Binary Quadratic POPs

We make use of the following theorem:

Theorem (Peña-Vera-Zuluaga (2006))

Let S be a compact set. For any degree d,

$$p(x) \in \mathcal{P}_d(x \in S : x_j \in \{-1, 1\})$$

$$\Leftrightarrow$$

$$p(x) = (1 + x_j)r_+(x) + (1 - x_j)r_-(x) + (1 - x_j^2)c(x),$$

where $r_+(x), r_-(x) \in \mathcal{P}_d(S)$ and $c(x) \in \mathbf{R}_{d-1}[x]$.

Valid Inequality Generation for Binary Quadratic POPs

We make use of the following theorem:

Theorem (Peña-Vera-Zuluaga (2006))

Let S be a compact set. For any degree d,

$$p(x) \in \mathcal{P}_d(x \in S : x_j \in \{-1, 1\})$$
$$\Leftrightarrow$$
$$p(x) = (1 + x_j)r_+(x) + (1 - x_j)r_-(x) + (1 - x_j^2)c(x),$$

where $r_+(x), r_-(x) \in \mathcal{P}_d(S)$ and $c(x) \in \mathbf{R}_{d-1}[x]$.

We can approximate $\mathcal{P}_2(S \cap \{-1, 1\}^n)$ by

$$\mathcal{Q}_2^j(G) = \{(1 + x_j)r_+(x) + (1 - x_j)r_-(x) + (1 - x_j^2)c(x) : r^+(x), r^-(x) \in \mathcal{K}_2(G), \quad c(x) \in \mathbf{R}_1[x]\}$$

and we have that

$$\mathcal{K}_2(G) \subset \mathcal{Q}_2^j(G) \subset \mathcal{P}_2(S \cap \{-1, 1\}^n)$$

Valid Inequality Generation for Binary Quadratic POPs

Goal

Given G describing S, find $p(x) \in \mathcal{P}_2(S \cap \{-1,1\}^n) \setminus \mathcal{K}_2(G)$

Given G and Y

$$\begin{aligned} \min \quad & \langle p, Y \rangle \\ \text{s.t.} \quad & p(x) \in \mathcal{Q}_2^j(G) \\ & \|p\| = 1 \end{aligned}$$

Valid Inequality Generation for Binary Quadratic POPs

Goal

Given G describing S, find $p(x) \in \mathcal{P}_2(S \cap \{-1, 1\}^n) \setminus \mathcal{K}_2(G)$

Given G and Y

$$\begin{array}{ll} \min & \langle p, Y \rangle \\ \text{s.t.} & p(x) \in \mathcal{Q}_2^j(G) \\ & \|p\| = 1 \end{array}$$

Note that there is exactly one subproblem per binary variable j.
Moreover,

- the size of $\mathcal{Q}_2^j(G)$ is only twice size of $\mathcal{K}_2(G)$
- while the size of $\mathcal{K}_4(G)$ is $\sim n^2$ times size of $\mathcal{K}_2(G)$

Convergence Result

Theorem

When the polynomial inequality generation scheme is applied to a binary quadratic optimization problem with linear constraints $Ax = b$, and the initial set is

$$G_0 = \left\{ n - \|x\|^2, \sum_i (A_i^T x - b_i)^2, -\sum_i (A_i^T x - b_i)^2 \right\},$$

then if all the subproblems have an optimal value 0, then the algorithm has converged to a global optimal solution.

Computational Results

Quadratic Knapsack Problem

$$\max x^T P x$$
$$\text{s.t. } w^T x \le c$$
$$x \in \{-1, 1\}^n$$

		Lasserre $r = 4$		Lasserre $r = 2$		Poly. Ineq. Gen.				
n	Optimal	Obj.	Time (s)	Obj.	Time (s)	Iter. 0	Iter. 1	Iter. 5	Iter. 10	Time (s)
10	1653	1707.3	28.1	1857.7	0.8	1857.7	1821.9	1797.4	1784.8	5.8
20	8510	8639.7	17269.1	9060.3	2.9	9060.3	9015.3	8925.9	8850.3	35.4
30	18229	-	-	19035.9	4.3	19035.9	18920.2	18791.7	18727.2	196.6
40	2679	-	-	4735.9	6.8	4735.9	4590.7	4248.2	4126.7	1009.7
50	16192	-	-	21777.9	19.2	21777.9	21390.3	20162.1	19407.1	7014.3
60	58451	-	-	62324.4	126.6	62324.4	62019.1	60906.0	60585.5	17961.1
70	16982	-	-	23884.9	231.4	23884.9	23484.0	22852.8	-	15582.2
80	-	-	-	80482.7	365.4	80482.7	79738.9	-	-	11072.3

(5-hour time limit)

Computational Results

Quadratic Assignment Problem

$$\begin{aligned} \min \ & \textstyle\sum_{i\neq k, j\neq l} f_{ik} d_{jl} x_{ij} x_{kl} & \\ \text{s.t. } & \textstyle\sum_i x_{ij} = 1 & 1 \leq j \leq n \\ & \textstyle\sum_j x_{ij} = 1 & 1 \leq i \leq n \\ & x \in \{0,1\}^{n\times n}. & \end{aligned}$$

		Lasserre $r=4$		Lasserre $r=2$		Poly. Ineq. Gen.				
n	Optimal	Obj.	Time (s)	Obj.	Time (s)	Iter. 0	Iter. 1	Iter. 5	Iter. 10	Time (s)
3	46			**46.0**	0.3	**46.0**				0.3
4	52	**52.0**	1154.8	50.8	1.0	50.8	51.8	**52.0**		6.3
5	110	-	-	104.3	3.4	104.3	105.1	106.3	106.8	68.5
6	272	-	-	268.9	9.3	268.9	269.4	269.8	270.2	404.4
7	356	-	-	344.2	18.1	344.2	344.9	345.6	346.0	3331.3
8	100	-	-	77.2	73.2	77.2	77.8	78.9	-	11413.9
9	280	-	-	247.5	281.7	247.5	248.6	-	-	13171.5

(5-hour time limit)

Computational Results

Degree Three Binary POPs

$$\max \sum_{|\alpha|\leq 3} c_\alpha x^\alpha$$
$$\text{s.t. } a^T x \leq b$$
$$x \in \{-1, 1\}^n$$

		Lasserre $r = 6$		Lasserre $r = 4$		Poly. Ineq. Gen.				
n	Optimal	Obj.	Time (s)	Obj.	Time (s)	Iter. 0	Iter. 1	Iter. 5	Iter. 10	Time (s)
5	58	**58.00**	9.6	59.37	2.1	67.16	58.45	**58.00**		5.2
10	139	**139.00**	4866.0	148.97	35.9	154.59	148.85	143.41	139.12	75.3
15	1371	-	-	1524.71	1436.2	1582.04	1575.49	1519.88	1494.01	1319.9
20	1654	-	-	1707.95	18106.6	1718.53	1716.00	1708.66	1705.15	15763.9
25	-	-	-	-	-	3967.12	3960.78	-	-	14287.3

(5-hour time limit)

Plenty of Future Research!

- Theoretical issues:
  - Prove convergence for (some scheme of) the general case

Plenty of Future Research!

- Theoretical issues:
  - Prove convergence for (some scheme of) the general case
  - Specialize the methodology to other interesting problem classes, such as:
    - complementarity problems
    - quadratic constrained quadratic problems

Plenty of Future Research!

- Theoretical issues:
  - Prove convergence for (some scheme of) the general case
  - Specialize the methodology to other interesting problem classes, such as:
    - ⋆ complementarity problems
    - ⋆ quadratic constrained quadratic problems
- Algorithmic issues:
  - Choice of index for j for inequality generation in the binary case

Plenty of Future Research!

- Theoretical issues:
  - Prove convergence for (some scheme of) the general case
  - Specialize the methodology to other interesting problem classes, such as:
    - complementarity problems
    - quadratic constrained quadratic problems
- Algorithmic issues:
  - Choice of index for j for inequality generation in the binary case
  - Add multiple inequalities at each iteration

Plenty of Future Research!

- Theoretical issues:
  - Prove convergence for (some scheme of) the general case
  - Specialize the methodology to other interesting problem classes, such as:
    - complementarity problems
    - quadratic constrained quadratic problems
- Algorithmic issues:
  - Choice of index for j for inequality generation in the binary case
  - Add multiple inequalities at each iteration
  - Find ways to reduce size of SDP subproblems

Plenty of Future Research!

- Theoretical issues:
  - Prove convergence for (some scheme of) the general case
  - Specialize the methodology to other interesting problem classes, such as:
    - complementarity problems
    - quadratic constrained quadratic problems
- Algorithmic issues:
  - Choice of index for j for inequality generation in the binary case
  - Add multiple inequalities at each iteration
  - Find ways to reduce size of SDP subproblems
  - Avoid SDP altogether: second-order cone optimization?