Low-Rank Positive Semi-definite Programming

Wenbao Ai

mailto:wenbaoai@gmail.com

September 4, 2008

PSDP Model

$$\begin{aligned} \text{min} \quad & C \bullet X \\ \text{s.t.} \quad & A_i \bullet X = c_i,\ i = 1, \cdots, m; \\ & X \succeq 0. \end{aligned}$$

Polynomial-time problem.

Two research direction:

1. new algorithms to solve the PSDP problem (ϵ-solutions).

2. Applications of PSDP.

Low-Rank PSDP Model

$$\begin{aligned} \min \quad & C \bullet X \\ \text{s.t.} \quad & A_i \bullet X = c_i, \; i = 1, \cdots, m; \\ & X \succeq 0; \\ & \text{rank}(X) \leq p. \end{aligned}$$

Usually NP-hard problems.

Basic idea to solve the above problems:

1. Obtain a solution to the relaxed PSDP problem by PSDP solver, such as `SeDuMi`.

2. Manage to obtain a low-rank approximate solution to the original problem.

quadratic optimization with quadratic constraints: Homogeneous case

$$
\begin{array}{lll}
(QP) & \text{min} & x^T C x \\
 & \text{s.t.} & x^T A_i x \le (=) c_i, \; i = 1, \cdots, m.
\end{array}
$$

Equivalent form in low-rank PSDP

$$
\begin{array}{lll}
(QP) & \text{min} & C \bullet X \\
 & \text{s.t.} & A_i \bullet X \le (=) c_i, \; i = 1, \cdots, m; \\
 & & X \succeq 0; \\
 & & \text{rank}(X) \le 1.
\end{array}
$$

Relaxed form in PSDP

$$
\begin{array}{lll}
(SP) & \text{min} & C \bullet X \\
 & \text{s.t.} & A_i \bullet X \leq (=) c_i, \;\; i = 1, \cdots, m; \\
 & & X \succeq 0;
\end{array}
$$

Quadratic optimization with quadratic constraints: Inhomogeneous case

$$\begin{aligned} \text{min} \quad & x^T Q_0 x + b_0^T x + c_0 \\ \text{s.t.} \quad & x^T Q_i x + b_i^T x + c_i \leq (=) 0, \ i = 1, \cdots, m. \end{aligned}$$

Equivalent homogeneous form

$$\begin{aligned} \text{min} \quad & x^T Q_0 x + b_0^T x t + c_0 t^2 \\ \text{s.t.} \quad & x^T Q_i x + b_i^T x t + c_i t^2 \leq (=) 0, \ i = 1, \cdots, m. \\ & t^2 = 1. \end{aligned}$$

Question: If $t = -1$ what will happen?

Example 1. Max-cut model

$$
\begin{array}{ll}
\text{min} & x^T Q_0 x \\
\text{s.t.} & x_i^2 = 1,\ i = 1, \cdots, n.
\end{array}
$$

Equivalent form in low-rank PSDP

$$
\begin{array}{ll}
\text{min} & Q_0 \bullet X \\
\text{s.t.} & x_{ii} = 1,\ i = 1, \cdots, n; \\
 & X \succeq 0; \\
 & \text{rank}(X) \leq 1.
\end{array}
$$

An excellent software based on Goemans (1995).

Example 2. Trust-region subproblem: One-ball problem

$$\begin{aligned} \text{min} \quad & x^T Q_0 x + b_0^T x + c_0 \\ \text{s.t.} \quad & \|x\| \le 1. \end{aligned}$$

Equivalent homogeneous form

$$\begin{aligned} \text{min} \quad & x^T Q_0 x + b_0^T x t + c_0 t^2 \\ \text{s.t.} \quad & x^T x - t^2 \le 0, \\ & t^2 = 1. \end{aligned}$$

Equivalent low-rank form

$$
\begin{aligned}
\min \quad & M(q_0) \bullet X \\
\text{s.t.} \quad & M(q_1) \bullet X \le 0, \\
& x_{00} = 1; \\
& X \succeq 0; \\
& \text{rank}(X) \le 1.
\end{aligned}
$$

$$
M(q_0) := \begin{bmatrix} 0 & -b_0^T \\ -b_0 & Q_0 \end{bmatrix}, \; M(q_1) := \begin{bmatrix} c_1 & -b_1^T \\ -b_1 & Q_1 \end{bmatrix}
$$

Equivalent to its relaxed form (Thm by Sturm-Zhang)

$$\begin{aligned} \text{min} \quad & M(q_0) \bullet X \\ \text{s.t.} \quad & M(q_1) \bullet X \le 0, \\ & x_{00} = 1; \\ & X \succeq 0; \end{aligned}$$

How did the proof proceed? (Ye-Zhang Proof for one constraint)

Key 1: KKT condition:

$$\begin{aligned} & XZ = 0, \quad y_1 M(q_1) \bullet X = 0, \\ & Z, X \succeq 0, \\ & M(q_1) \bullet X \le 0, \\ & x_{00} = 1. \end{aligned}$$

Key 2: X's rank-one decomposition.

Lemma Let $A_1 \in \mathcal{S}^{n \times n}$ and $X \in \mathcal{S}_+^{n \times n}$ with

$$A_1 \bullet X = 0.$$

Then one can find a rank-one decomposition for X,

$$X = x_1 x_1^T + x_2 x_2^T + \cdots + x_r x_r^T,$$

such that

$$A_1 \bullet x_i x_i^T = 0 \text{ for } i = 1, ..., r.$$

What is more, constructive proof.

General Case

Theorem $v(SP) = v(QP) \iff$ $\nabla^2_{xx} L(x; y) = Q_0 + \sum_{i=1}^m y_i Q_i \succeq 0$ where y is the Lagrangian multiplier for an optimal solution of $(QP) \iff$ Strong duality.

Example 3. CDT subproblem: Two-ball problem

Equivalent low-rank form

$$\begin{aligned} \text{min} \quad & M(q_0) \bullet X \\ \text{s.t.} \quad & M(q_1) \bullet X \leq 0, \\ & M(q_2) \bullet X \leq 0, \\ & x_{00} = 1; \\ & X \succeq 0; \\ & \text{rank}(X) \leq 1. \end{aligned}$$

Also equivalent to its relaxed form?

$$
\begin{aligned}
\min \quad & M(q_0) \bullet X \\
\text{s.t.} \quad & M(q_1) \bullet X \le 0, \\
& M(q_2) \bullet X \le 0, \\
& x_{00} = 1; \\
& X \succeq 0;
\end{aligned}
$$

Can we determine if $v(SP) = v(QP)$ `only by an optimal solution of the relaxed PSDP?`

Lemma Let $A_1, A_2 \in \mathcal{S}^{n\times n}$ and $X \in \mathcal{S}_+^{n\times n}$ with

$$A_1 \bullet X = 0, \quad A_2 \bullet X = 0.$$

If $r := \operatorname{rank}(X) \geq 3$ then in polynomial-time (real-number computation) one finds a rank-one decomposition for X,

$$X = x_1 x_1^T + x_2 x_2^T + \cdots + x_r x_r^T,$$

such that

$$A_1 \bullet x_i x_i^T = 0 \text{ for } i = 1, ..., r$$

$$A_2 \bullet x_i x_i^T = 0 \text{ for } i = 1, ..., r-2.$$

Another similar result by Au-Yeung and Poon (1979)

Definition For given Lagrangian multipliers λ and μ for the quadratic program $(QP)_2$, we say that they have Property $\mathcal{I}'$ if:

(1) $\lambda > 0$ and $\mu > 0$;

(2) $H(\lambda, \mu) = Q_0 + \lambda I + \mu Q_2 \succeq 0, \mathrm{rank}\ (H(\lambda, \mu)) = n - 1$;

(3) The system of linear equations $H(\lambda, \mu)x = b_0 + \mu b_2$ has two solutions x_1 and x_2 satisfying $x_i^T x_i = 1$, $i = 1, 2$, and $q_2(x_1)q_2(x_2) < 0$.

theorem Suppose that $(QP)_2$ satisfies the Slater condition. Then, $(QP)_2$ has no strong duality if and only if there exist multipliers λ and μ such that Property $\mathcal{I}'$ holds.

Definition (Chen and Yuan 2002) For given Lagrangian multipliers λ and μ for the quadratic program $(QP)_2$, we say that they have Property $\mathcal{J}$ if:

(1) $\lambda > 0$ and $\mu > 0$;

(2) $H(\lambda, \mu) = Q_0 + \lambda I + \mu Q_2 \succeq 0, \mathrm{rank}\,(H(\lambda, \mu)) = n - 1$;

(3) The following 'surrogate' problem

$$(P)_{\frac{\lambda}{\lambda+\mu}} \quad \text{minimize} \quad q_0(x)$$

$$\text{subject to} \quad \frac{\lambda}{\lambda+\mu} q_1(x) + \frac{\mu}{\lambda+\mu} q_2(x) \leq 0$$

has two solutions x_1 and x_2 satisfying

$$H(\lambda, \mu)x = b_0 + \mu b_2, x_1^T x_1 < 1, x_2^T x_2 > 1.$$

The necessarity been proved by Chen and Yuan, but the sufficiency not.

Theorem If $Q_2 \succeq 0$ then Property $\mathcal{J}$ is equivalent to Property $\mathcal{I}'$. If $Q_2 \not\succeq 0$ then Property $\mathcal{J}$ is not identical to Property $\mathcal{I}'$, the latter being a necessary and sufficient condition for $(QP)_2$ to admit a gap with its SDP relaxation.

A counterexample

$$\begin{array}{ll} \text{minimize} & q_0(x) = x_1^2 - 3x_1 \\ \text{subject to} & q_1(x) = x_1^2 + x_2^2 - 1 \le 0 \\ & q_2(x) = -x_1^2 - x_2^2 + 2x_1 \le 0 \end{array}$$

On the optimal line of the dual problem

Yuan: Is there a sufficiently good approximate solution of the primal problem on the optimal line of the dual problem?

Counterexample 6.1.

$$\begin{array}{ll} \text{minimize} & q_0(x_1,x_2) = x_1(p-x_1) \\ \text{subject to} & q_1(x_1,x_2) = x_1^2 + x_2^2 \le \frac{17}{16}p^2, \\ & q_2(x_1,x_2) = (x_1-2p)^2 + (x_2-p)^2 \le \frac{73}{16}p^2, \end{array}$$

The optimal value of the primal problem $v^* \approx -0.1544p^2$.

The optimal value of $q_0(x_1,x_2)$ On this line segment is identically 0 for any p.

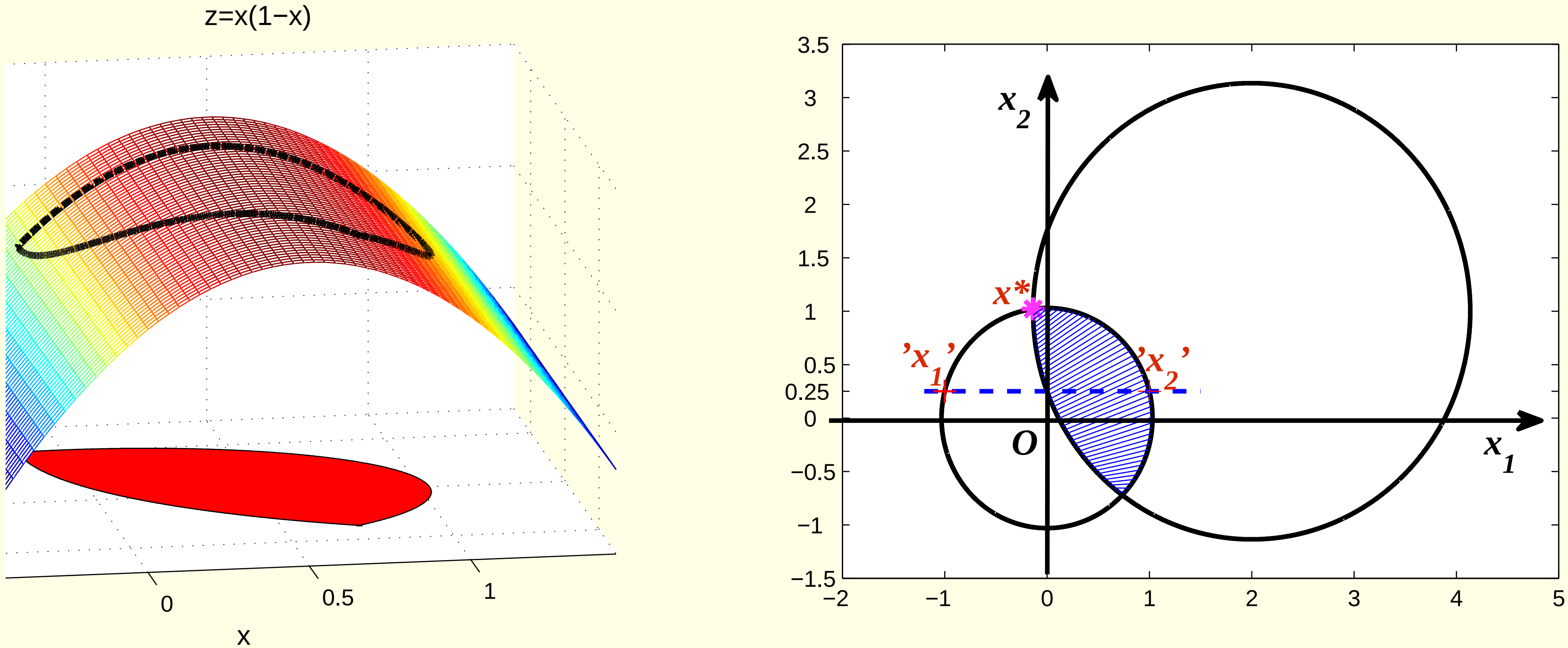

Figure 1: The graphs of $z = x_1(1 - x_1)$(on the left) and the feasible domain (on the right) at $p = 1$.

Numerical test

Table 1: Numerical results

n	value 1	value 2	gap	rank	$I(\varepsilon_2)$	n	valu
1	31.8310	31.8310	-1.9315e-008	1	V	16	-288.22
2	-61.6350	-61.6350	1.0267e-007	1	V	17	-180.26
3	-92.6195	-92.6195	1.5046e-008	1	V	18	-257.03
4	-64.3479	-64.3479	6.3392e-009	1	V	19	-307.89
5	-76.0429	-76.0429	1.2039e-007	1	V	20	-250.22
6	-148.3942	-148.3942	8.4647e-008	1	V	21	-216.68
7	-149.2147	-149.2147	1.3788e-007	1	V	22	-285.22
8	-165.2366	-165.2366	2.2856e-007	1	V	23	-305.70
9	-146.7020	-146.7020	6.5012e-010	1	V	24	-273.77
10	-193.3607	-193.3607	1.0247e-007	1	V	25	-305.12
11	-194.9409	-194.9409	4.3410e-006	1	V	26	-311.09
12	-131.2606	-131.2606	3.4186e-009	1	V	27	-269.25

Table 2: Numerical results for $n = 5$

Inst.	value 1	value 2	gap	rank	$I(\varepsilon_2)$	Ins.	va
1	-72.2487	-72.2487	-9.0962e-010	1	V	16	-195
2	-78.8733	-78.8733	3.1875e-007	1	V	17	-91
3	-129.3945	-129.3945	2.4719e-009	1	V	18	-149
4	-78.6061	-78.6061	3.1858e-007	1	V	19	-199
5	-87.7781	-87.7781	4.0048e-009	1	V	20	-96
6	-162.4757	-162.4757	3.2261e-009	1	V	21	-193
7	-181.4192	-181.4192	1.2105e-006	1	V	22	-121
8	-148.9920	-131.6450	17.3470	2	H	23	-132
9	-84.6160	-84.6160	1.2004e-007	1	V	24	-221
10	-106.1400	-106.1400	2.6063e-007	1	V	25	-69
11	-80.2952	-80.2952	8.1327e-010	1	V	26	-48
12	-93.9455	-37.5482	56.3973	2	H	27	-204

Table 3: Numerical results for $n = 50$

ins.	value 1	value 2	gap	rank	$I(\varepsilon_2)$	Ins.	valu
1	-329.1350	-329.1350	1.3564e-008	1	V	16	-353.2
2	-418.0411	-418.0411	1.3500e-010	1	V	17	-422.6
3	-334.9108	-334.9108	8.4879e-010	1	V	18	-373.7
4	-314.3538	-314.3538	4.0116e-007	1	V	19	-356.4
5	-406.6970	-406.6970	1.8738e-008	1	V	20	-449.4
6	-376.4849	-376.4849	6.9003e-009	1	V	21	-363.3
7	-436.8686	-436.8686	1.1316e-009	1	V	22	-422.4
8	-456.1419	-456.1419	1.0745e-009	1	V	23	-376.0
9	-420.0406	-420.0406	2.3637e-009	1	V	24	-399.0
10	-443.0921	-443.0921	2.8577e-009	1	V	25	-428.4
11	-398.1299	-398.1299	1.2683e-008	1	V	26	-422.2
12	-381.3000	-381.3000	2.2239e-009	1	V	27	-422.8

A new proof been presented by me jointly with my student for the I' condition, based on the mathematical analysis.

Problem 1. Can any bound be obtained in practice due to the ϵ-approximate in PSDP?

Problem 2. Can we present other algorithms based on the I' condition?

Problem 3. Can any bound be obtained for quadratic programming if using PSDP to solve it?

Low-Rank Solutions for m Linear Matrix Inequalities or equalities

Theorem (Bohnenblust, Barvinok)

$$\mathcal{L} = \{X \in \mathcal{SF}^n \mid A_i \bullet X = 0, i = 1, ..., m\},$$

Let $\mathcal{L}$ be a subspace of $\mathcal{SF}^n$. Suppose that $\dim(\mathcal{L} \cap \mathcal{SF}^n_+) \geq 1$. If $\mathrm{codim}(\mathcal{L}) \leq d_{\mathcal{F}}(p+1) - 1$ and $1 \leq p \leq n-2$, then in polynomial time (in terms of n and $\log \frac{1}{\epsilon}$) one can find $X \succeq 0$, with $\|X\|_F = 1$, $\mathrm{rank}\ (X) \leq p$, and $\mathrm{dist}\ (X, \mathcal{L}) \leq \epsilon$. By taking any cluster point as $\epsilon \to 0$, we conclude that there is a nontrivial $X \in \mathcal{L} \cap \mathcal{SF}^n_+$, such that $\mathrm{rank}\ (X) \leq p$.

No constructive proof there exists!

If $\operatorname{codim}(\mathcal{L}) \leq d_{\mathcal{F}}(p+1)-2$, rank-reduction method is constructive.

Basic idea

Before we present our construction, let us first comment on the result and discuss the main ideas of the new proof here. As Barvinok remarked in [**?**], Theorem **??** and Theorem **??** (or as an exact counterpart, Theorem **??**) are quite different in nature. Theorem **??** gives a bound on the rank that is satisfied by any extremal solution of the LMI system, whereas Theorem **??** (or Theorem **??**) only assures the existence of one extremal solution of the LMI system with an even lower rank. This lower rank solution is in some way an analog of a degenerate vertex in the context of a polyhedron. As is well known, degeneracy occurs in a polyhedron like an accident, and it will almost surely not happen for a generally positioned polyhedron. Interestingly, Theorems **??** and **??** suggest that this rank degeneracy is inevitable for a general LMI system. To look for this particular 'degenerate matrix solution', we introduce a single parameter in the range space of the matrix solution. It is possible to confine this parameter to a finite interval, say $[0, 1]$. Moreover, it is easy to find the corresponding matrix solutions with respect to the parameters at the ends of the interval, initially being 0 and 1. One of the matrix solution is known to be positive definite with rank $p+1$, and the other has a rank no more than $p+1$ and is indefinite. Now, we can apply bisection to reduce the parameter search interval, retaining the property that one end yields a matrix in $\mathcal{L} \cap \mathcal{SF}_+^n$ with rank $p+1$, and the other end yields a matrix that is indefinite. It is also possible to combine the positive definite matrix with the indefinite matrix (after modifying its range space to fit that of the positive definite one), hence further reduce the rank of the positive definite matrix. However, this will lead to a matrix solution that does not precisely reside in $\mathcal{L}$, because of the nonlinear modification made on the indefinite matrix. The main effort is then to show that the distance to $\mathcal{L}$, nonetheless, converges to zero as the bisection reduces the size of the parameter search interval to zero.

Problem 1. How about the two-ball optimal solution when gap is too big?

Problem 2. How about the three ball problem?

Problem 3. Under what condition, lower rank solution obtained for m matrix equations?