Slide 1

Bad semidefinite programs: characterizations, and complexity

Gábor Pataki

Dept. of Operations Research
UNC, Chapel Hill

Slide 2

A Pair of Linear Programs (LP)

$$
(P) \quad \begin{array}{ll} Max_x & c^T x \\ s.t. & \sum_{i=1}^m x_i a_i \le b \end{array} \qquad \begin{array}{ll} Min_y & b^T y \\ s.t. & y \ge 0 \\ & a_i^T y = c_i \; \forall i \end{array} \quad (D)
$$

where $a_i, b, y \in \mathcal{R}^n$, $c, x \in \mathcal{R}^m$.

Slide 3

A Pair of Semidefinite Programs (SDP)

$$(P) \quad \begin{array}{ll} Sup_x & c^T x \\ s.t. & \sum_{i=1}^m x_i A_i \preceq B \end{array} \qquad \begin{array}{ll} Inf_Y & B \bullet Y \\ s.t. & Y \succeq 0 \\ & A_i \bullet Y = c_i \ \forall i \end{array} \quad (D)$$

where

- A_i, B, Y are symmetric matrices, $c, x \in \mathcal{R}^m$.
- $A \preceq B$ means that $B - A$ is positive semidefinite (psd).
- $A \prec B$ means that $B - A$ is positive definite.
- A matrix is psd, if all its principal subdeterminants are nonnegative.
- $A \bullet B = \sum_{i,j} a_{ij} b_{ij}$.

Slide 4

A very simple SDP

$$\begin{array}{ll} sup_x & x_1 \\ st. & x_1 \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \preceq \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \Leftrightarrow \begin{bmatrix} 1 & -x_1 \\ -x_1 & 0 \end{bmatrix} \succeq 0. \end{array}$$

Slide 5

Why is SDP important?

Modeling power: problems from engineering to combinatorial optimization can be cast as SDPs.

Example: Minimizing over (nonconvex) quadratic constraints has an SDP relaxation.

$$\begin{array}{rlcrl} \inf_x & x^T Q_0 x & \Leftrightarrow & \inf_x & Q_0 \bullet xx^T \\ s.t. & x^T Q_i x = c_i \ \ \forall i & & s.t. & Q \bullet xx^T = c_i \ \ \forall i \end{array}$$

Slide 6

SDP relaxation:

$$\begin{array}{rl} \inf_Y & Q_0 \bullet Y \\ st. & Q_i \bullet Y = c_i \ \ \forall i \\ & Y \succeq 0 \end{array} \tag{1}$$

Applies to: max cut, max bisection, 0–1 programs, ...

Slide 7

Why is SDP important?

- A natural generalization of LP:

  $Y \succeq 0$ and Y diagonal (can be enforced by lin. constraints) $\Leftrightarrow$ $Y_{ii} \geq 0 +$ lin. constraints on Y $\rightarrow$ LP.

- Efficiently solvable (Nesterov-Nemirovskii, Nesterov-Todd, Alizadeh, ...).

Slide 8

Differences with LP:

- Optimal value may
  - be irrational;
  - require exponentially many digits to just write down;
- Perhaps the most important: **strong duality** needs certain qualifications!

Slide 9

LP duality

$$
(P) \quad \begin{array}{ll} Max_x & c^T x \\ s.t. & \sum_{i=1}^m x_i a_i \leq b \end{array} \qquad\qquad \begin{array}{ll} Min_y & b^T y \\ s.t. & y \geq 0 \\ & a_i^T y = c_i \ \forall i \end{array} \quad (D)
$$

- For any pair of feasible solutions (x, y) :

  $c^T x \leq b^T y$ (weak duality).
- Both (P) and (D) feasible $\Rightarrow$ $\exists$ (x, y) with equality (strong duality). Provides **proof of optimality** in optimization algorithms.

Slide 10

SDP duality

Weak duality is easy to prove. If x and Y are feasible, then

$$
\begin{aligned}
c^T x = \sum_{i=1}^m c_i x_i &= \sum_{i=1}^m (A_i \bullet Y) x_i \\
&= (\sum_{i=1}^m x_i A_i) \bullet Y \\
&= (B - Z) \bullet Y \quad (Z \succeq 0 \text{ is the primal slack}) \\
&\leq B \bullet Y
\end{aligned}
$$

since $Y \bullet Z \geq 0$ if $Y, Z \succeq 0$.

Slide 11

Strong duality: (Known, follows from more general results for convex programs).

If there is a feasible $\bar{x}$ for (P) such that

$$B - \sum_{i=1}^{m} \bar{x}_i A_i \succ 0 \;\; (\text{“strict feasibility''},\ \text{or “Slater} - \text{condition''}),$$

then $\inf_x = \max_Y$ for any objective c.

Fact: All known poly-time algorithms to solve SDP require a strictly feasible primal *and* dual.

Slide 12

A badly behaved semidefinite system (no unif. LP-duality)

$$x_1 \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \preceq \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \quad \Leftrightarrow \quad \begin{bmatrix} 1 & -x_1 \\ -x_1 & 0 \end{bmatrix} \succeq 0.$$

Not strictly feasible. When we seek sup $2x_1$, the dual is

$$\begin{array}{llll} \inf & \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \bullet Y & st.\ Y \succeq 0, & \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \bullet Y = 2 \\ \inf & y_{11} & st. & \begin{bmatrix} y_{11} & 1 \\ 1 & y_{22} \end{bmatrix} \succeq 0 \end{array}$$

Primal max $= 0$, dual inf $= 0$ $(y_{11} \to 0,\ y_{22} \to +\infty)$, but not attained!

Slide 13

And another:

$$st. \quad x_1 \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} + x_2 \begin{bmatrix} 0 & 0 & 1 \\ 0 & 2 & 0 \\ 1 & 0 & 0 \end{bmatrix} \quad \preceq \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

When we seek $\sup 2x_2$, the primal and dual both attain, and dual value is 1!

Slide 14

Dual is

$$\begin{array}{ll} \min & y_{22} \\ st. & \begin{bmatrix} 0 & y_{12} & 1-y_{22} \\ y_{12} & y_{22} & y_{23} \\ 1-y_{22} & y_{23} & y_{33} \end{bmatrix} \succeq 0 \end{array}$$

Any feasible solution must have $y_{22} = 1$, so its optimal value is 1, and it is attained!

Slide 15

Def: If $\sup_x = \min_Y$ for any objective c for which (P) is bounded, we say that (P), is **well behaved** (more precise terminology: "admits uniform LP-duality".)

Goal: A good characterization of badly behaved semidefinite systems! I.e. to show that to the question

$$\text{Is } \ (P) \ \ \sum_i x_i A_i \preceq B \ \ \text{well} - \text{behaved?}$$

both the "yes", and the "no" answers should have a poly-time verifiable certificate. Important: we allow operations on real numbers, regardless of how they are represented in a computer (Blum/Shub/Smale model of computing).

Slide 16

Analogy in graph theory

- "Can graph G be drawn on the plane ($\Leftrightarrow$ is it planar)?"
- Yes answer: just draw it (coordinates of vertices can have polynomial size description).
- No answer: G not planar $\Leftrightarrow$ it contains one of two minors:

Slide 17

Recall our bad system, and compare it with a good one, which is not strictly feasible either:

$$\text{(bad) } x_1 \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \preceq \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$$

$$\text{(good) } x_1 \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix} \preceq \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

After staring at the bad ones long enough, they start looking suspiciously similar.

Slide 18

Theorem Suppose w.l.o.g. that in the semidefinite system

$$(P) \quad \sum_i x_i A_i \preceq B$$

the maximum rank psd slack ($Z = B - \sum_i x_i A_i$) looks like

$$Z = \begin{bmatrix} I_r & 0 \\ 0 & 0 \end{bmatrix}, \text{ with } 0 \leq r \leq n.$$

Then (P) is badly behaved $\Leftrightarrow$ there is a Y of the form

$$Y = y_0 B - \sum_i y_i A_i = \begin{bmatrix} Y_{11} & Y_{12} & Y_{13} \\ Y_{12}^T & I_s & 0 \\ Y_{13}^T & 0 & 0 \end{bmatrix}, \text{ with } s \geq 0, \;\; Y_{13} \neq 0.$$

In other words, **" all bad SDP's look the same " !**

Slide 19

$$Z = \overbrace{\begin{bmatrix} I_r & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}}^{\text{max. rank slack}}, \quad Y = \overbrace{\begin{bmatrix} Y_{11} & Y_{12} & Y_{13} \\ Y_{12}^T & I_s & 0 \\ Y_{13}^T & 0 & 0 \end{bmatrix}}^{\text{certificate of the bad behaviour}}, \quad Y_{13} \neq 0$$

In the example

$$x_1 \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \preceq \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$$

these matrices are

$$Z = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \quad Y = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$$

Slide 20

Complexity of "Is (P) well behaved?"

"No" answer has poly-size certificate:

- Show Z maximum rank slack (checking that Z has maximal rank is nontrivial!).
- Show Y: checking that Y has the required shape is trivial.

How about characterizing **well behaved** semidefinite systems, and a certificate?

Slide 21

Theorem Let $Z = \begin{bmatrix} I_r & 0 \\ 0 & 0 \end{bmatrix}$, with $0 \le r \le n$ be as before.

Then (P) is well behaved $\Leftrightarrow$

(1) $\exists\, V$ s.t. $V = \begin{bmatrix} 0 & 0 \\ 0 & I_{n-r} \end{bmatrix}$, and $A_i \bullet V = B \bullet V = 0\ \forall i$.

(2) If $Y = y_0 B - \sum_i y_i A_i$, and $Y = \begin{bmatrix} Y_{11} & Y_{12} \\ Y_{12}^T & 0 \end{bmatrix}$, then $Y_{12} = 0$.

Slide 22

In the example

$$x_1 \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix} \preceq \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

we have

$$(1) \quad V = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix},$$

and (2) is trivial too.

Slide 23

Complexity of "Is (P) well behaved?", cont'd

"Yes" answer has poly-size certificate:

- Show V: checking that V has required shape and $A_i \bullet V = B \bullet V = 0 \ \forall i$ is trivial.
- Checking (2) amounts to checking the equality of two *subspaces*.

Slide 24

A more general framework: conic LPs

$$
(P(c)) \quad \begin{array}{ll} \sup & \langle c, x \rangle \\ s.t. & Ax \leq_K b \end{array} \qquad \begin{array}{ll} \inf & \langle b, y \rangle \\ s.t. & y \geq_{K^*} 0 \\ & A^* y = c \end{array} \quad (D(c))
$$

where

- K be a closed, convex cone, $(x \in K, \lambda \geq 0 \Rightarrow \lambda x \in K)$.
- $K^* = \{y \mid \langle y, s \rangle \geq 0 \ \forall s \in K\}$ the *dual* of K.
- A a linear map, A^* its adjoint (transpose).
- $x \leq_K y$ means: $y - x \in K$.

Slide 25

Now, a crash course in convex analysis ...

C is a closed, convex set, $x \in C$.

- $\operatorname{ri} C$: the *relative interior* of C:
  -

Figure 1: A set in 2D and its relative interior

  - The relative interior is the same as the usual interior, if C is full dimensional.
  - ri $\mathcal{R}^n_+ = \{ x \,|\, x_i > 0 \; i = 1, \dots, n \}$.
  - ri $\mathcal{S}^n_+ = \{ x \,|\, \text{rank } x = n \}$.

Slide 26

- dir $(x, C) := \{ y \,|\, x + \alpha y \in C \text{ for some } \alpha > 0 \}$: the *feasible directions* at x in C.
- Fact: dir (x, C) is a convex cone, but it may not be closed!

Figure 2: Feasible directions

Slide 27

Theorem Suppose that K belongs to the class of **nice** cones (technical condition) and in (P), the feasible set of $(P(c))$ the most interior slack is

$$z = b - Ax.$$

Then (P) is badly behaved $\Leftrightarrow$ there is a y such that

$$y = v_0 b - Av \in \operatorname{cl} \operatorname{dir}\,(z, K) \setminus \operatorname{dir}\,(z, K).$$

Slide 28

How general is our class of cones?

- K is nice cone, if
  - K is polyhedral,
  - K is the cone of psd matrices,
  - $K = \{\, u \in \mathcal{R}^k \,|\, u_1 \geq \| u_{2:k} \|_p \,\}$ (another important cone in optimization).
  - . . . essentially all cones interesting from an optimization viewpoint!

For all these, our main theorem gives a very simple, and poly-time checkable characterization, when the conic system is badly behaved.

Slide 29

Specialization:

- If K is polyhedral: the set dir (z, K) is closed $\Rightarrow$ we recover the good behaviour of polyhedra.
- If K is the cone of psd matrices:

$$Z = \overbrace{\begin{bmatrix} I_r & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}}^{\text{max. rank slack}}, \quad Y = \overbrace{\begin{bmatrix} Y_{11} & Y_{12} & Y_{13} \\ Y_{12}^T & I_s & 0 \\ Y_{13}^T & 0 & 0 \end{bmatrix}}^{\text{certificate of the bad behaviour}}$$

- A bit of linear algebra shows: $Y \in$ cldir $(Z, K) \setminus$ dir (Z, K).

Slide 30

The background: the classical part

Theorem: (Duffin, Jeroslow, Karlovitz):

The feasible conic system

$$(P) \quad Ax \leq_K b$$

yields uniform LP duality $\Leftrightarrow$

The set

$$\begin{pmatrix} A & b \\ 0 & 1 \end{pmatrix}^* \begin{pmatrix} K^* \\ \mathcal{R}_+ \end{pmatrix} \tag{2}$$

is closed.

Slide 31

The background: the new part

Theorem: (P 2001): Let K be a nice cone, M a linear map, $x \in \text{ri}(\mathcal{R}(M) \cap K)$ (nonneg. orthant: max # of nonzeros; semidef. cone: max. rank).

Then

$$
\begin{array}{rcl}
M^*K^* \text{ is closed} & \Leftrightarrow & \\
\mathcal{R}(M) \cap \text{cl dir } (x, K) & = & \mathcal{R}(M) \cap \text{dir } (x, K) \ (\textbf{Condition 1})
\end{array}
$$

Obviously,

K is polyhedral **or** $x \in \text{int } K \Rightarrow \text{dir } (x, K)$ is closed $\Rightarrow$ Condition 1.

Slide 32

Example The image of $\mathcal{S}^2_+$ (2×2 semidefinite cone) is not closed. Now $K = K^* = \mathcal{S}^2_+$,

$$
\begin{array}{rcl}
M \begin{bmatrix} z_1 \\ z_2 \end{bmatrix} & = & z_1 \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} + z_2 \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \\
M^* \begin{bmatrix} a & c \\ c & b \end{bmatrix} & = & \begin{bmatrix} a \\ 2c \end{bmatrix}
\end{array}
$$

(M^*K^* is not closed (e.g. $(0, 2)^T \in \text{cl } (M^*K^*) \setminus M^*K^*$).

Slide 33

Example 1 cont'd Indeed

$$x = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \in \mathrm{ri}\,(\mathcal{R}(M) \cap K))$$

$$y = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \in \mathcal{R}(M) \cap (\mathrm{cl\ dir}\ (x, K) \setminus \mathrm{dir}\ (x, K))$$

Slide 34

Conclusion, and further work

- SDP: a useful, and exciting optimization paradigm: many similarities, and challenging differences with LP.
- An area of crucial differences: "badly behaved systems", that lack strong duality for some objective.
- Our result: a very simple, exact, and poly-time *checkable* characterization of such systems: the result actually holds for many other conic optimization problems.
- Background: a new result on a basic problem in convex analysis.
- Next: Is the recognition of bad systems in P?