Separating Doubly Nonnegative and Completely Positive Matrices

Kurt M. Anstreicher Dept. of Management Sciences University of Iowa (joint work with Hongbo Dong)

INFORMS National Meeting, Austin, November 2010

Let S_n denote the set of $n \times n$ real symmetric matrices, S_n^+ denote the cone of $n \times n$ real symmetric positive semidefinite matrices and \mathcal{N}_n denote the cone of symmetric nonnegative $n \times n$ matrices.

Let S_n denote the set of $n \times n$ real symmetric matrices, S_n^+ denote the cone of $n \times n$ real symmetric positive semidefinite matrices and \mathcal{N}_n denote the cone of symmetric nonnegative $n \times n$ matrices.

• The cone of $n \times n$ doubly nonnegative (DNN) matrices is then $\mathcal{D}_n = \mathcal{S}_n^+ \cap \mathcal{N}_n.$

Let S_n denote the set of $n \times n$ real symmetric matrices, S_n^+ denote the cone of $n \times n$ real symmetric positive semidefinite matrices and \mathcal{N}_n denote the cone of symmetric nonnegative $n \times n$ matrices.

- The cone of $n \times n$ doubly nonnegative (DNN) matrices is then $\mathcal{D}_n = \mathcal{S}_n^+ \cap \mathcal{N}_n.$
- The cone of $n \times n$ completely positive (CP) matrices is $C_n = \{X \mid X = AA^T \text{ for some } n \times k \text{ nonnegative matrix } A\}.$

Let S_n denote the set of $n \times n$ real symmetric matrices, S_n^+ denote the cone of $n \times n$ real symmetric positive semidefinite matrices and \mathcal{N}_n denote the cone of symmetric nonnegative $n \times n$ matrices.

- The cone of $n \times n$ doubly nonnegative (DNN) matrices is then $\mathcal{D}_n = \mathcal{S}_n^+ \cap \mathcal{N}_n.$
- The cone of $n \times n$ completely positive (CP) matrices is $C_n = \{X \mid X = AA^T \text{ for some } n \times k \text{ nonnegative matrix } A\}.$
- Dual of \mathcal{C}_n is the cone of $n \times n$ copositive matrices, $\mathcal{C}_n^* = \{ X \in \mathcal{S}_n \mid y^T X y \ge 0 \ \forall y \in \Re_n^+ \}.$

Let S_n denote the set of $n \times n$ real symmetric matrices, S_n^+ denote the cone of $n \times n$ real symmetric positive semidefinite matrices and \mathcal{N}_n denote the cone of symmetric nonnegative $n \times n$ matrices.

- The cone of $n \times n$ doubly nonnegative (DNN) matrices is then $\mathcal{D}_n = \mathcal{S}_n^+ \cap \mathcal{N}_n.$
- The cone of $n \times n$ completely positive (CP) matrices is $C_n = \{X \mid X = AA^T \text{ for some } n \times k \text{ nonnegative matrix } A\}.$
- Dual of C_n is the cone of $n \times n$ copositive matrices, $C_n^* = \{X \in S_n \mid y^T X y \ge 0 \ \forall y \in \Re_n^+\}.$

Clear that

$$\mathcal{C}_n \subseteq \mathcal{D}_n, \qquad \mathcal{D}_n^* = \mathcal{S}_n^+ + \mathcal{N}_n \subseteq \mathcal{C}_n^*,$$

and these inclusions are in fact strict for n > 4.

Why Bother?

Why Bother?

• [Bur09] shows that broad class of NP-hard problems can be posed as linear optimization problems over C_n .

Why Bother?

- [Bur09] shows that broad class of NP-hard problems can be posed as linear optimization problems over C_n .
- \mathcal{D}_n is a tractable relaxation of \mathcal{C}_n . Expect that solution of relaxed problem will be $X \in \mathcal{D}_n \setminus \mathcal{C}_n$.

Why Bother?

- [Bur09] shows that broad class of NP-hard problems can be posed as linear optimization problems over C_n .
- \mathcal{D}_n is a tractable relaxation of \mathcal{C}_n . Expect that solution of relaxed problem will be $X \in \mathcal{D}_n \setminus \mathcal{C}_n$.
- Note that least n where problem occurs is n = 5.

For $X \in S_n$ let G(X) denote the undirected graph on vertices $\{1, \ldots, n\}$ with edges $\{\{i \neq j\} | X_{ij} \neq 0\}.$

For $X \in S_n$ let G(X) denote the undirected graph on vertices $\{1, \ldots, n\}$ with edges $\{\{i \neq j\} | X_{ij} \neq 0\}.$

Definition 1. Let G be an undirected graph on n vertices. Then G is called a CP graph if any matrix $X \in \mathcal{D}_n$ with G(X) = G also has $X \in \mathcal{C}_n$. For $X \in S_n$ let G(X) denote the undirected graph on vertices $\{1, \ldots, n\}$ with edges $\{\{i \neq j\} | X_{ij} \neq 0\}.$

Definition 1. Let G be an undirected graph on n vertices. Then G is called a CP graph if any matrix $X \in \mathcal{D}_n$ with G(X) = G also has $X \in \mathcal{C}_n$.

The main result on CP graphs is the following:

Proposition 1. [KB93] An undirected graph on n vertices is a CP graph if and only if it contains no odd cycle of length 5 or greater. In [BAD09] it is shown that:

In [BAD09] it is shown that:

• Extreme rays of \mathcal{D}_5 are either rank-one matrices in \mathcal{C}_5 , or rankthree "extremely bad" matrices where G(X) is a 5-cycle (every vertex in G(X) has degree two). In [BAD09] it is shown that:

- Extreme rays of \mathcal{D}_5 are either rank-one matrices in \mathcal{C}_5 , or rankthree "extremely bad" matrices where G(X) is a 5-cycle (every vertex in G(X) has degree two).
- Any such extremely bad matrix can be separated from C_5 by a transformation of the Horn matrix

$$H := \begin{pmatrix} 1 & -1 & 1 & 1 & -1 \\ -1 & 1 & -1 & 1 & 1 \\ 1 & -1 & 1 & -1 & 1 \\ 1 & 1 & -1 & 1 & -1 \\ -1 & 1 & 1 & -1 & 1 \end{pmatrix} \in \mathcal{C}_5^* \setminus \mathcal{D}_5^*.$$

• Separation procedure based on transformed Horn matrix applies to $X \in \mathcal{D}_5 \setminus \mathcal{C}_5$ where X has rank three and G(X) has at least one vertex of degree 2.

- Separation procedure based on transformed Horn matrix applies to $X \in \mathcal{D}_5 \setminus \mathcal{C}_5$ where X has rank three and G(X) has at least one vertex of degree 2.
- More general separation procedure applies to any $X \in \mathcal{D}_5 \setminus \mathcal{C}_5$ that is not componentwise strictly positive.

- Separation procedure based on transformed Horn matrix applies to $X \in \mathcal{D}_5 \setminus \mathcal{C}_5$ where X has rank three and G(X) has at least one vertex of degree 2.
- More general separation procedure applies to any $X \in \mathcal{D}_5 \setminus \mathcal{C}_5$ that is not componentwise strictly positive.

An even more general separation procedure that applies to any $X \in \mathcal{D}_5 \setminus \mathcal{C}_5$ is described in [BD10]. In this talk we will describe the procedure from [DA10] for $X \in \mathcal{D}_5 \setminus \mathcal{C}_5$, $X \neq 0$, and its generalization to larger matrices having block structure.

A separation procedure for the 5×5 case

Assume that $X \in \mathcal{D}_5$, $X \not\geq 0$. After a symmetric permutation and diagonal scaling, X may be assumed to have the form

$$X = \begin{pmatrix} X_{11} & \alpha_1 & \alpha_2 \\ \alpha_1^T & 1 & 0 \\ \alpha_2^T & 0 & 1 \end{pmatrix},$$
(1)

where $X_{11} \in \mathcal{D}_3$.

A separation procedure for the 5×5 case

Assume that $X \in \mathcal{D}_5$, $X \not\geq 0$. After a symmetric permutation and diagonal scaling, X may be assumed to have the form

$$X = \begin{pmatrix} X_{11} & \alpha_1 & \alpha_2 \\ \alpha_1^T & 1 & 0 \\ \alpha_2^T & 0 & 1 \end{pmatrix},$$
(1)

where $X_{11} \in \mathcal{D}_3$.

Theorem 1. [BX04, Theorem 2.1] Let $X \in \mathcal{D}_5$ have the form (1). Then $X \in \mathcal{C}_5$ if and only if there are matrices A_{11} and A_{22} such that $X_{11} = A_{11} + A_{22}$, and

$$\begin{pmatrix} A_{ii} & \alpha_i \\ \alpha_i^T & 1 \end{pmatrix} \in \mathcal{D}_4, \quad i = 1, 2.$$

A separation procedure for the 5×5 case

Assume that $X \in \mathcal{D}_5$, $X \not\geq 0$. After a symmetric permutation and diagonal scaling, X may be assumed to have the form

$$X = \begin{pmatrix} X_{11} & \alpha_1 & \alpha_2 \\ \alpha_1^T & 1 & 0 \\ \alpha_2^T & 0 & 1 \end{pmatrix},$$
(1)

where $X_{11} \in \mathcal{D}_3$.

Theorem 1. [BX04, Theorem 2.1] Let $X \in \mathcal{D}_5$ have the form (1). Then $X \in \mathcal{C}_5$ if and only if there are matrices A_{11} and A_{22} such that $X_{11} = A_{11} + A_{22}$, and

$$\begin{pmatrix} A_{ii} & \alpha_i \\ \alpha_i^T & 1 \end{pmatrix} \in \mathcal{D}_4, \quad i = 1, 2.$$

In [BX04], Theorem 1 is utilized only as a proof mechanism, but we now show that it has algorithmic consequences as well.

$$V = \begin{pmatrix} V_{11} & \beta_1 & \beta_2 \\ \beta_1^T & \gamma_1 & 0 \\ \beta_2^T & 0 & \gamma_2 \end{pmatrix} \quad such that \quad \begin{pmatrix} V_{11} & \beta_i \\ \beta_i^T & \gamma_i \end{pmatrix} \in \mathcal{D}_4^*, \quad i = 1, 2,$$

and $V \bullet X < 0$.

$$V = \begin{pmatrix} V_{11} & \beta_1 & \beta_2 \\ \beta_1^T & \gamma_1 & 0 \\ \beta_2^T & 0 & \gamma_2 \end{pmatrix} \quad such that \quad \begin{pmatrix} V_{11} & \beta_i \\ \beta_i^T & \gamma_i \end{pmatrix} \in \mathcal{D}_4^*, \quad i = 1, 2,$$

and $V \bullet X < 0$.

• Suppose that $X \in \mathcal{D}_5 \setminus \mathcal{C}_5$, and V are as in Theorem 2. If $\tilde{X} \in \mathcal{C}_5$ is another matrix of the form (1), then Theorem 1 implies that $V \bullet \tilde{X} \ge 0$.

$$V = \begin{pmatrix} V_{11} & \beta_1 & \beta_2 \\ \beta_1^T & \gamma_1 & 0 \\ \beta_2^T & 0 & \gamma_2 \end{pmatrix} \quad such that \quad \begin{pmatrix} V_{11} & \beta_i \\ \beta_i^T & \gamma_i \end{pmatrix} \in \mathcal{D}_4^*, \quad i = 1, 2,$$

and $V \bullet X < 0$.

- Suppose that $X \in \mathcal{D}_5 \setminus \mathcal{C}_5$, and V are as in Theorem 2. If $\tilde{X} \in \mathcal{C}_5$ is another matrix of the form (1), then Theorem 1 implies that $V \bullet \tilde{X} \ge 0$.
- However cannot conclude that $V \in \mathcal{C}_5^*$ because $V \bullet \tilde{X} \ge 0$ only holds for \tilde{X} of the form (1), in particular, $\tilde{x}_{45} = 0$.

$$V = \begin{pmatrix} V_{11} & \beta_1 & \beta_2 \\ \beta_1^T & \gamma_1 & 0 \\ \beta_2^T & 0 & \gamma_2 \end{pmatrix} \quad such that \quad \begin{pmatrix} V_{11} & \beta_i \\ \beta_i^T & \gamma_i \end{pmatrix} \in \mathcal{D}_4^*, \quad i = 1, 2,$$

and $V \bullet X < 0$.

- Suppose that $X \in \mathcal{D}_5 \setminus \mathcal{C}_5$, and V are as in Theorem 2. If $\tilde{X} \in \mathcal{C}_5$ is another matrix of the form (1), then Theorem 1 implies that $V \bullet \tilde{X} \ge 0$.
- However cannot conclude that $V \in \mathcal{C}_5^*$ because $V \bullet \tilde{X} \ge 0$ only holds for \tilde{X} of the form (1), in particular, $\tilde{x}_{45} = 0$.
- Fortunately, by [HJR05, Theorem 1], V can easily be "completed" to obtain a copositive matrix that still separates X from C_5 .

Theorem 3. Suppose that $X \in \mathcal{D}_5 \setminus \mathcal{C}_5$ has the form (1), and V satisfies the conditions of Theorem 2. Define

$$V(s) = \begin{pmatrix} V_{11} & \beta_1 & \beta_2 \\ \beta_1^T & \gamma_1 & s \\ \beta_2^T & s & \gamma_2 \end{pmatrix}$$

Then $V(s) \bullet X < 0$ for any s, and $V(s) \in \mathcal{C}_5^*$ for $s \ge \sqrt{\gamma_1 \gamma_2}$.

Separation for larger matrices with block structure

Procedure for 5 case where $X \not\geq 0$ can be generalized to larger matrices with block structure. Assume X has the form

$$X = \begin{pmatrix} X_{11} & X_{12} & X_{13} & \dots & X_{1k} \\ X_{12}^T & X_{22} & 0 & \dots & 0 \\ X_{13}^T & 0 & \ddots & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & 0 \\ X_{1k}^T & 0 & \dots & 0 & X_{kk} \end{pmatrix},$$
(2)

where $k \ge 3$, each X_{ii} is an $n_i \times n_i$ matrix, and $\sum_{i=1}^k n_i = n$.

Lemma 1. Suppose that $X \in \mathcal{D}_n$ has the form (2), $k \geq 3$, and let

$$X^{i} = \begin{pmatrix} X_{11} & X_{1i} \\ X_{1i}^{T} & X_{ii} \end{pmatrix}, i = 2, \dots, k.$$

Then $X \in C_n$ if and only if there are matrices A_{ii} , i = 2, ..., ksuch that $\sum_{i=2}^k A_{ii} = X_{11}$, and

$$\begin{pmatrix} A_{ii} & X_{1i} \\ X_{1i}^T & X_{ii} \end{pmatrix} \in \mathcal{C}_{n_1+n_i}, \quad i = 2, \dots, k.$$

Moreover, if $G(X^i)$ is a CP graph for each i = 2, ..., k, then the above statement remains true with $C_{n_1+n_i}$ replaced by $\mathcal{D}_{n_1+n_i}$. **Theorem 4.** Suppose that $X \in \mathcal{D}_n \setminus \mathcal{C}_n$ has the form (2), where $G(X^i)$ is a CP graph, i = 2, ..., k. Then there is a matrix V, also of the form (2), such that

$$\begin{pmatrix} V_{11} & V_{1i} \\ V_{1i}^T & V_{ii} \end{pmatrix} \in \mathcal{D}_{n_1+n_i}^*, \quad i = 2, \dots, k,$$

and $V \bullet X < 0$. Moreover, if $\gamma_i = \text{diag}(\text{Diag}(V_{ii})^{.5})$, then the matrix

$$\tilde{V} = \begin{pmatrix} V_{11} & \dots & V_{1k} \\ \vdots & \ddots & \vdots \\ V_{1k}^T & \dots & V_{kk} \end{pmatrix},$$

where $V_{ij} = \gamma_i \gamma_j^T$, $2 \le i \ne j \le k$, has $V \in \mathcal{C}_n^*$ and $V \bullet X = V \bullet X < 0$.

Note that:

Note that:

• Matrix X may have numerical entries that are small but not exactly zero. Can then apply Lemma 1 to perturbed matrix \tilde{X} where entries of X below a specified tolerance are set to zero. If a cut V separating \tilde{X} from C_n is found, then $V \bullet X \approx V \bullet \tilde{X} <$ 0, and V is very likely to also separate X from C_n . Note that:

- Matrix X may have numerical entries that are small but not exactly zero. Can then apply Lemma 1 to perturbed matrix \tilde{X} where entries of X below a specified tolerance are set to zero. If a cut V separating \tilde{X} from \mathcal{C}_n is found, then $V \bullet X \approx V \bullet \tilde{X} <$ 0, and V is very likely to also separate X from \mathcal{C}_n .
- Theorem 4 may provide a cut separating a given $X \in \mathcal{D}_n \setminus \mathcal{C}_n$ even when the sufficient conditions for generating such a cut are not satisfied. In particular, a cut may be found even when the condition that X^i is a CP graph for each *i* is not satisfied.

A second case where block structure can be used to generate cuts for a matrix $X \in \mathcal{D}_n \setminus \mathcal{C}_n$ is when X has the form

$$X = \begin{pmatrix} I & X_{12} & X_{13} & \dots & X_{1k} \\ X_{12}^T & I & X_{23} & \dots & X_{2k} \\ X_{13}^T & X_{23}^T & \ddots & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & X_{(k-1)k} \\ X_{1k}^T & X_{2k}^T & \dots & X_{(k-1)k}^T & I \end{pmatrix},$$
(3)

where $k \ge 2$, each X_{ij} is an $n_i \times n_j$ matrix, and $\sum_{i=1}^k n_i = n$. The structure in (3) corresponds to a partitioning of the vertices $\{1, 2, \ldots, n\}$ into k stable sets in G(X), of size n_1, \ldots, n_k (note that $n_i = 1$ is allowed).

Applications

Example 1: Consider the box-constrained QP problem (QPB) max $x^TQx + c^Tx$ s.t. $0 \le x \le e$.

Applications

Example 1: Consider the box-constrained QP problem (QPB) max $x^TQx + c^Tx$ s.t. $0 \le x \le e$.

To describe a CP-formulation of QPB, define matrices

$$Y = \begin{pmatrix} 1 & x^T \\ x & X \end{pmatrix}, \qquad Y^+ = \begin{pmatrix} 1 & x^T & s^T \\ x & X & Z \\ s & Z^T & S \end{pmatrix}.$$

Applications

Example 1: Consider the box-constrained QP problem (QPB) max $x^TQx + c^Tx$ s.t. $0 \le x \le e$.

To describe a CP-formulation of QPB, define matrices

$$Y = \begin{pmatrix} 1 & x^T \\ x & X \end{pmatrix}, \qquad Y^+ = \begin{pmatrix} 1 & x^T & s^T \\ x & X & Z \\ s & Z^T & S \end{pmatrix}.$$

By the result of [Bur09], QPB can then be written in the form

(QPB_{CP}) max
$$Q \bullet X + c^T x$$

s.t. $x + s = e$, $\text{Diag}(X + 2Z + S) = e$,
 $Y^+ \in \mathcal{C}_{2n+1}$.

• Can be shown that DNN relaxation is equivalent to "SDP+RLT" relaxation, and for n = 2 problem is equivalent to QPB [AB10].

- Can be shown that DNN relaxation is equivalent to "SDP+RLT" relaxation, and for n = 2 problem is equivalent to QPB [AB10].
- Constraints from Boolean Quadric Polytope (BQP) are valid for off-diagonal components of X [BL09]. For n = 3, BQP is completely determined by triangle inequalities and RLT constraints. However, can still be a gap when using SDP+RLT+TRI relaxation.

- Can be shown that DNN relaxation is equivalent to "SDP+RLT" relaxation, and for n = 2 problem is equivalent to QPB [AB10].
- Constraints from Boolean Quadric Polytope (BQP) are valid for off-diagonal components of X [BL09]. For n = 3, BQP is completely determined by triangle inequalities and RLT constraints. However, can still be a gap when using SDP+RLT+TRI relaxation.
- Example from [BL09] with n = 3 has optimal value for QPB of 1.0, value for SDP+RLT+TRI relaxation of 1.093. Solution matrix Y^+ has 5×5 principal submatrix that is *not* strictly positive, and is *not* CP. Can obtain cut from Theorem 3, resolve problem, and repeat.



Figure 1: Gap to optimal value for Burer-Letchford QPB problem (n = 3)

Example 2: Let A be the adjacency matrix of a graph G on n vertices, and let α be the maximum size of a stable set in G. Known [dKP02] that **Example 2:** Let A be the adjacency matrix of a graph G on n vertices, and let α be the maximum size of a stable set in G. Known [dKP02] that

$$\alpha^{-1} = \min\left\{ (I+A) \bullet X : ee^T \bullet X = 1, X \in \mathcal{C}_n \right\}.$$
(4)

Example 2: Let A be the adjacency matrix of a graph G on n vertices, and let α be the maximum size of a stable set in G. Known [dKP02] that

$$\alpha^{-1} = \min\left\{ (I+A) \bullet X : ee^T \bullet X = 1, X \in \mathcal{C}_n \right\}.$$
(4)

Relaxing C_n to D_n results in the Lovász-Schrijver bound

$$(\vartheta')^{-1} = \min\left\{ (I+A) \bullet X : ee^T \bullet X = 1, X \in \mathcal{D}_n \right\}.$$
 (5)

Let G_{12} be the complement of the graph corresponding to the vertices of a regular icosahedron [BdK02]. Then $\alpha = 3$ and $\vartheta' \approx 3.24$.

Let G_{12} be the complement of the graph corresponding to the vertices of a regular icosahedron [BdK02]. Then $\alpha = 3$ and $\vartheta' \approx 3.24$.

• Using the cone \mathcal{K}_{12}^1 to approximate the dual of (4) provides no improvement [BdK02].

Let G_{12} be the complement of the graph corresponding to the vertices of a regular icosahedron [BdK02]. Then $\alpha = 3$ and $\vartheta' \approx 3.24$.

- Using the cone \mathcal{K}_{12}^1 to approximate the dual of (4) provides no improvement [BdK02].
- For the solution matrix $X \in \mathcal{D}_{12}$ from (5), cannot find cut based on first block structure (2). However *can* find a cut based on (3). Adding this cut and re-solving, gap to $1/\alpha = \frac{1}{3}$ is approximately 2×10^{-8} .

References

- [AB10] Kurt M. Anstreicher and Samuel Burer, Computable representations for convex hulls of lowdimensional quadratic forms, Math. Prog. B **124** (2010), 33–43.
- [BAD09] Samuel Burer, Kurt M. Anstreicher, and Mirjam Dür, *The difference between* 5 × 5 *doubly* nonnegative and completely positive matrices, Linear Algebra Appl. **431** (2009), 1539–1552.
- [BD10] Samuel Burer and Hongbo Dong, Separation and relaxation for cones of quadratic forms, Dept. of Management Sciences, University of Iowa (2010).
- [BdK02] Immanuel M. Bomze and Etienne de Klerk, Solving standard quadratic optimization problems via linear, semidefinite and copositive programming, J. Global Optim. **24** (2002), 163–185.
- [BL09] Samuel Burer and Adam N. Letchford, On nonconvex quadratic programming with box constriants, SIAM J. Optim. **20** (2009), 1073–1089.
- [Bur09] Samuel Burer, On the copositive representation of binary and continuous nonconvex quadratic programs, Math. Prog. **120** (2009), 479–495.
- [BX04] Abraham Berman and Changqing Xu, 5 × 5 Completely positive matrices, Linear Algebra Appl. **393** (2004), 55–71.
- [DA10] Hongbo Dong and Kurt Anstreicher, Separating doubly nonnegative and completely positive matrices, Dept. of Management Sciences, University of Iowa (2010).
- [dKP02] Etienne de Klerk and Dmitrii V. Pasechnik, Approximation of the stability number of a graph via copositive programming, SIAM J. Optim. **12** (2002), 875–892.
- [HJR05] Leslie Hogben, Charles R. Johnson, and Robert Reams, *The copositive completion problem*, Linear Algebra Appl. **408** (2005), 207–211.
- [KB93] Natalia Kogan and Abraham Berman, Characterization of completely positive graphs, Discrete Math. 114 (1993), 297–304.