An Introduction to Semidefinite Programming for Combinatorial Optimization (Problem Session)

Samuel Burer*

IPCO Summer School
May 20–21, 2019

*Department of Business Analytics, University of Iowa, Iowa City, IA, 52242-1994, USA. Email: samuel-burer@uiowa.edu.

Exercise 1. Given $X \in \mathbb{S}^p_+$, prove there exists $Y \in \mathbb{S}^p_+$ such that $Y^2 = X$.

Comments and Hints: Y is called the *square root* of X, and is usually written as $X^{1/2}$. To prove its existence, as a warm-up, first consider the case when X is diagonal. Then prove the general case by using the spectral decomposition VDV^T of X, and remember that *orthogonal* means $V^TV = VV^T = I$.

Solution. Try $Y := VD^{1/2}V^T$, where $D^{1/2}$ is the component-wise square root of D. We have

$$Y^2 = (VD^{1/2}V^T)(VD^{1/2}V^T) = VD^{1/2} \cdot I \cdot D^{1/2}V^T = VDV^T = X.$$

So this Y works.

Exercise 2. Prove $X, S \in \mathbb{S}^p_+$ implies $X \bullet S \geq 0$.

Comments and Hints: This will establish that $\mathbb{S}^p_+$ is self-dual. To prove it, use the previous exercise along with the identities

$$\begin{aligned}\text{trace}(M^T N) &= M \bullet N = N^T \bullet M^T = \text{trace}(NM^T),\\ \|M\|_F &= \sqrt{M \bullet M}\end{aligned}$$

to show that $X \bullet S = \|X^{1/2} S^{1/2}\|_F^2$.

Solution.

$$\begin{aligned} X \bullet S &= \text{trace}(X^T S)\\ &= \text{trace}(XS)\\ &= \text{trace}(X^{1/2} X^{1/2} S^{1/2} S^{1/2})\\ &= \text{trace}(S^{1/2} X^{1/2} X^{1/2} S^{1/2})\\ &= X^{1/2} S^{1/2} \bullet X^{1/2} S^{1/2}\\ &= \|X^{1/2} S^{1/2}\|_F^2\\ &\geq 0. \end{aligned}$$

Alternate solution. Let $X = VDV^T$ and $S = WEW^T$ be the spectral decompositions. Then

$$\begin{aligned} X \bullet S &= \text{trace}(VDV^T WEW^T)\\ &= \text{trace}(DV^T WEW^T V)\\ &= D \bullet (V^T WEW^T V)\\ &\geq 0, \end{aligned}$$

where the inequality follows because D is nonnegative diagonal and $V^T WEW^T V$ is PSD and hence also has a nonnegative diagonal.

Exercise 3. For $X, S \in \mathbb{S}^p_+$, prove $X \bullet S = 0 \Leftrightarrow XS = 0$.

Comments and Hints: Prove this using the prior exercise. Note that, when X and S are diagonal, this result reduces to the well-known vector condition $x^T s = 0 \Leftrightarrow x \circ s = 0$ for $x, s \geq 0$.

Solution. ($\Leftarrow$) $XS = 0$ implies $X \bullet S = \text{trace}(XS) = 0$. ($\Rightarrow$) The equation $0 = X \bullet S = \|X^{1/2} S^{1/2}\|_F^2$ implies $X^{1/2} S^{1/2} = 0$. Pre-multiplying by $X^{1/2}$ and post-multiplying by $S^{1/2}$ yields the result.

Exercise 4. What is the dimension of the feasible set of the primal problem

$$\begin{aligned} \inf \quad & C \bullet X \\ \text{s.t.} \quad & A_i \bullet X = b_i \ \ \forall\, i = 1, \ldots, m \\ & X \succeq 0 \end{aligned} \tag{P}$$

assuming that the data matrices $\{A_1, \ldots, A_m\}$ are linearly independent in $\mathbb{S}^n$ and that (P) is interior feasible? Assuming also that the dual problem

$$\begin{aligned} \sup \quad & b^T y \\ \text{s.t.} \quad & C - \textstyle\sum_{i=1}^m y_i A_i \succeq 0. \end{aligned} \tag{D}$$

has an interior feasible solution, determine the dimension of the feasible set of (D). Here, *dimension* refers to the number of degrees of freedom of the (relative) interior of the feasible set.

Solution. Since the dimension of $\mathbb{S}^n_+$ is $\binom{n+1}{2}$, the m equations $A_i \bullet X = b_i$ remove m degrees of freedom. Hence, the primal dimension is at most $\binom{n+1}{2} - m$. The existence of an interior feasible X ensures that is exactly $\binom{n+1}{2} - m$. On the dual side, we have m degrees of freedom, and so the dimension is at most m. Again, the existence of an interior solution guarantees that it is exactly m.

Exercise 5. The following primal-dual example shows that strong duality does not hold for SDP in general:

$$\begin{aligned} 1 = \inf \quad & X_{33} \\ \text{s.t.} \quad & X_{11} = 0 \\ & X_{12} + X_{21} + 2X_{33} = 2 \\ & X \succeq 0; \end{aligned}$$

$$\begin{aligned} 0 = \sup \quad & 2\,y_2 \\ \text{s.t.} \quad & \begin{pmatrix} -y_1 & -y_2 & 0 \\ -y_2 & 0 & 0 \\ 0 & 0 & 1 - 2\,y_2 \end{pmatrix} \succeq 0. \end{aligned}$$

Argue that 1 and 0 are in fact the respective optimal values. Also, discuss the violation(s) of the strong-duality theorem.

Comments and Hints. One argument uses the following 2×2 determinant property of $X \in \mathbb{S}^p_+$: $X_{jk}^2 \le X_{jj}X_{kk}$ for all j, k. In particular, consider how $X_{jj} = 0$ affects the j-th row and column of X.

Solution. For (P), $X_{11} = 0$ causes $X_{12} = X_{21} = 0$, which in turn forces $X_{33} = 1$, which proves the optimal value. For (D), let S be the 3×3 PSD matrix. Because $S_{22} = 0$, it follows that $y_2 = 0$, which proves the optimal value. For both (P) and (D), we see that $\text{rank}(X) < 3$ and $\text{rank}(S) < 3$, and hence neither has interior. This is the "cause" of the duality gap.

Exercise 6. For a column vector x and nonnegative scalar t, prove

$$\|x\| \leq t \quad \iff \quad t^2 I - xx^T \succeq 0.$$

Comments and Hints. This is the first step in proving that any SOCP can be modeled as an SDP; the next step (which is not part of this exercise) is to apply the *Schur complement theorem*, which ensures

$$t^2 I - xx^T \succeq 0 \quad \iff \quad \begin{pmatrix} t & x^T \\ x & tI \end{pmatrix} \succeq 0.$$

To prove the exercise, apply the first definition of positive semidefiniteness: $X \in \mathbb{S}^p_+$ if and only if $v^T X v \geq 0$ for all $v \in \mathbb{R}^p$. To make the proof a little easier, you can assume without loss of generality that $\|v\| = 1$.

Solution. Let column vector v with $\|v\| = 1$ be arbitrary. Then

$$\begin{aligned}
& v^T(t^2 I - xx^T)v \geq 0 \\
\Leftrightarrow \quad & t^2 v^T v - (v^T x)^2 \geq 0 \\
\Leftrightarrow \quad & t^2 - (v^T x)^2 \geq 0 \\
\Leftrightarrow \quad & (v^T x)^2 \leq t^2.
\end{aligned}$$

By the Cauchy-Schwarz inequality, the final left-hand side $(v^T x)^2$ is maximized when $v \parallel x$. Hence, $t^2 I - xx^T \succeq 0$ if and only if

$$\left(\frac{x^T x}{\|x\|}\right)^2 \leq t^2 \quad \iff \quad \|x\|^2 \leq t^2 \quad \iff \quad \|x\| \leq t.$$

Exercise 7. For any $\mu > 0$, the point on the primal-dual central path corresponding to μ is the unique solution (X_μ, y_μ, S_μ) of the following system of equations (assuming $X, S \succ 0$):

$$\begin{aligned} A_i \bullet X &= b_i \quad \forall\, i = 1, \ldots, m \\ \sum_{i=1}^{m} y_i A_i + S &= C \\ XS &= \mu I. \end{aligned}$$

Derive a simple, closed-form expression for $X_\mu \bullet S_\mu$.

Solution. $X_\mu \bullet S_\mu = \mathrm{trace}(X_\mu S_\mu) = \mathrm{trace}(\mu I) = n\mu$.

Exercise 8. An alternate form of the ϑ-number SDP relaxation is

$$\begin{array}{ll} \max & ee^T \bullet X \\ \text{s.t.} & \operatorname{trace}(X) = 1 \\ & X_{ij} = 0 \;\; \forall \text{ edges } (i,j) \\ & X \succeq 0. \end{array}$$

The dual of this form is

$$\begin{array}{ll} \min & \lambda \\ \text{s.t.} & \lambda I + \displaystyle\sum_{\text{edges } (i,j)} y_{ij} E_{ij} - ee^T \succeq 0. \end{array}$$

Prove that strong duality holds between this primal-dual pair by exhibiting positive definite feasible solutions in each problem.

Solution. $X = I$ is interior feasible for the primal. In the dual, take λ large enough and set all $y_{ij} = 0$ so that $\lambda I - ee^T \succ 0$.

Exercise 9. Given a MaxCut instance with adjacency matrix A, a compact way to write the SDP relaxation is

$$\max\{L \bullet X : \mathrm{diag}(X) = e, X \succeq 0\},$$

where:

- $L := \frac{1}{4}(\mathrm{Diag}(Ae) - A)$ is the *Laplacian matrix* of the graph;
- the operator $\mathrm{diag}(\cdot)$ extracts the diagonal of its matrix input;
- the operator $\mathrm{Diag}(\cdot)$ makes a diagonal matrix out of its vector input.

A compact way to write the dual is

$$\min\{e^T y : \mathrm{Diag}(y) - L = S, S \succeq 0\}.$$

Using these compact forms, prove weak duality between the primal and dual.

Comments and Hints. As with LP, sometimes it's easier to work with a specific form of your problem rather than a standard form. This particular form for the MaxCut SDP relaxation highlights the importance of the operators $\mathrm{diag}(\cdot)$ and $\mathrm{Diag}(\cdot)$, which are in fact adjoint operators, i.e., no matter the inputs X and y, it holds that $\mathrm{diag}(X)^T y = X \bullet \mathrm{Diag}(y)$.

Solution.

$$\begin{aligned} e^T y - L \bullet X &= \mathrm{diag}(X)^T y - L \bullet X \\ &= X \bullet \mathrm{Diag}(y) - L \bullet X \\ &= (\mathrm{Diag}(y) - L) \bullet X \\ &= X \bullet S \\ &\geq 0 \end{aligned}$$

Exercise 10. Referring to the previous exercise, the low-rank approach for solving the MaxCut SDP solves instead

$$\max\{L \bullet (RR^T) : \operatorname{diag}(RR^T) = e\},$$

where the number of columns p in $R \in \mathbb{R}^{n \times p}$ is approximately $\sqrt{2n}$. In terms of the n rows of R, describe the geometric interpretation of the constraint $\operatorname{diag}(RR^T) = e$.

Solution. Let e_j be the j-th coordinate vector so that the j-th row of R (as a column vector) is $R^T e_j$. The constraint $\operatorname{diag}(RR^T) = e$ breaks down into the constraints $\|R^T e_j\|^2 = e_j^T RR^T e_j = e_j e_j^T \bullet RR^T = 1$. Hence, the constraints simply say that the rows of R each have norm 1.