Two remarks on previous lecture

- Fact 9 shows that pd (or psd) can be verified by checking that every principal minor is positive (or nonnegative). For pd we only need to examine leading principal minors. However, this does not work for verifying psd as $\begin{pmatrix} 0 & 0 \\ 0 & -1 \end{pmatrix}$ shows. This matrix has leading minors equal to zero but a negative trailing minor.

- Fact 10 has introduced the Schur complement. If $A \succ 0$ (\begin{pmatrix} A & B \\ B^T & C \end{pmatrix} \succeq 0$ iff $C - B^T A^{-1} B \succeq 0$. This lends itself to showing that the Cholesky factorization of a pd matrix exists.

**Fact 11 (Representing quadratics)** Given $Q \in \mathbb{S}^{n \times n}$ the quadratic form $x^T Q x = \text{trace} (x^T Q x) = \text{trace} (Q x x^T) = Q \bullet xx^T$,

while a quadratic function of $x$, is a linear function of $X := xx^T$.

**Fact 12** If $U$ and $V$ are psd, then $U \bullet V \succeq 0$ (used in showing weak duality for SDPs). In fact, we have the following theorem.

**Theorem 1** If $K = \mathbb{S}^{n \times n}_+$, then $K$ is self-dual, $K = K^* := \{ V \in \mathbb{S}^{n \times n} : U \bullet V \succeq 0, \forall U \in K \}$.

**Proof**: Part 1: showing $K \subseteq K^*$: want to show that if $U, V \succeq 0, U \bullet V \succeq 0$. Demonstration 1: By facts 3, 4, and 6 we can assume that one of $U, V$ is diagonal. But then $U \bullet V = \text{trace} (UV) = \sum_j u_{jj} v_{jj} \geq 0$ (need only look at the diagonal entries).

Demonstration 2: By Fact 1, and the existence of matrix square roots, $U \bullet V = \text{trace} (U^{1/2} V U^{1/2}) \geq 0$ since $U^{1/2} V U^{1/2}$ is psd (Fact 7) and so its trace is nonnegative.

Part 2: showing $K^* \subseteq K$ by contraposition, i.e., if $U \notin K$, $U \notin K^*$. If $U \notin K$, then by definition there exists a $z$ such that $z^T U z < 0$, i.e., $U \bullet zz^T < 0$ (by Fact 11). But $zz^T \in K$ by Corollary 1, so $U \notin K^*$.

By the same argument, if $U \succ 0, V \succeq 0, V \neq 0$, then $U \bullet V > 0$. This shows a similarity between nonnegative vectors and the cone of psd matrices.
Fact 13 If $U \succ 0$ and $\beta \geq 0$ then $\{V \in \mathbb{S}^n \in \mathbb{R}^{n \times n} : U \cdot V \leq \beta\}$ is compact.

Proof: The set is clearly closed, so need to show it is also bounded. Suppose $\mu = \lambda_{\min}(U) > 0$ (since $U \succ 0$). Then

$$U \cdot V = (U - \mu I) \cdot V + \mu I \cdot V \geq \mu I \cdot \Lambda(V) = \mu \|\lambda(V)\|_1 \geq \mu \|\lambda(V)\|_2 = \mu \|V\|_F$$

for any $V \succeq 0$. So any $V$ in the set has $\|V\|_F \leq \beta/\mu$.

Fact 14 If $U, V \succeq 0$, then $U \cdot V = 0$ iff $UV = 0$.

Proof: First note that if $UV = 0$ then clearly $U \cdot V = \text{trace}(UV) = 0$. To show the converse suppose $U, V \succeq 0$ with $U \cdot V = 0$. Then $\text{trace}(UV) = \text{trace}(U^{1/2}V^{1/2}U^{1/2}V^{1/2}U^{1/2}) = 0$. Define $A := V^{1/2}U^{1/2}$ to obtain the condition $A^T A = 0$. This implies $A = 0$ since the diagonal entries of $A^T A$ are the squares of the 2-norms of the columns of $A$. But $U^{1/2}V^{1/2} = A^{T} = 0$ easily gives $UV = 0$.

Fact 15 If $U, V \in \mathbb{S}^n$, then $U, V$ commute iff $UV$ is symmetric, and iff they can be simultaneously diagonalized, i.e., iff we can find $Q, \Lambda, M$ ($Q$ orthogonal, $\Lambda, M$ diagonal) with

$$U = QAQ^T \quad V = QMQ^T.$$ 

Proof: Omitted, since part of HW1.

Applications

Matrix and NLP optimization problems

We have already shown how to minimize the maximum eigenvalue of a symmetric $M(y) \in \mathbb{S}^n$. What if $M(y)$ is not symmetric and not necessarily square? For example, say we want $\|M(y)\|_2$ as small as possible. Recall:

$$\|P\|_2 := \max\{\|Pz\|_2 : \|z\|_2 = 1\} = \sqrt{\max\{z^T P^T P z : \|z\|_2 = 1\}} = \sqrt{\lambda_{\max}(P^T P)}.$$

Lemma 1 If $P \in \mathbb{R}^{m \times n}$ then $\|P\|_2 < \eta$ iff

$$\begin{pmatrix} \eta I_m & P \\ P^T & \eta I_n \end{pmatrix} \succeq 0.$$

Proof: Observe that the above matrix retains linearity w.r.t. $\eta, y$. Examine two cases:

If $P = 0$, then both conditions reduce to $\eta \geq 0$.

If $P \neq 0$, 

$$\begin{pmatrix} \eta I & P \\ P^T & \eta I \end{pmatrix} \succeq 0 \text{ iff } \eta > 0 \text{ and } \eta I - P^T (\eta I)^{-1} P \succeq 0.$$

With $\eta > 0$, the last condition is equivalent to $\eta^2 I - P^T P \succeq 0$ or $\eta^2 \geq \lambda_{\max}(P^T P)$, which holds iff $\eta \geq \sqrt{\lambda_{\max}(P^T P)}$ and $\eta > 0$, which is what we needed to show.

Note that we have also shown that $\lambda_{\max}(P^T P) = \lambda_{\max}(PP^T)$, i.e., $\|P\|_2 = \|P^T\|_2$. 

2
So, \( \min_y ||M(y)||_2 \) is equivalent to

\[
\begin{align*}
\max_{(\eta,y)} & - \eta \\
\begin{pmatrix}
\eta I & M(y) \\
M(y)^T & \eta I
\end{pmatrix} & \succeq 0
\end{align*}
\]

which is the desired SDP in dual form.

**Linear programming and nonlinear programming**

Consider first an LP in dual form

\[
\begin{align*}
\max & b^T y \\
A^T y & \leq c
\end{align*}
\]

and try to rewrite it as an SDP. The constraint \( c - A^T y \geq 0 \) holds iff \( \text{Diag}(c - A^T y) \succeq 0 \), i.e.,

\[
C - \sum_i y_i A_i \preceq 0,
\]

where \( C = \text{Diag}(c) \) and \( A_i = \text{Diag}(a_i) \) and \( a_i \) is the \( i \)th column of \( A^T \). This gives the equivalent SDP in dual form.

Now consider a correspondence between primal LP and SDP. The primal SDP is given by

\[
\begin{align*}
\min_{x \in \mathbb{S}^{n \times n}} & C \cdot X \\
A_i \cdot X & = b_i, \ i = 1, \ldots, m \\
X & \succeq 0,
\end{align*}
\]

whereas the primal LP is given by

\[
\begin{align*}
\min_{x \in \mathbb{R}^n} & c^T x \\
Ax & = b \\
x & \geq 0.
\end{align*}
\]

In general, our SDPs may have block diagonal form. Suppose \( n_1, \ldots, n_k \geq 0, \sum_j n_j = n \) and define

\[
\mathcal{S} = \left\{ \begin{pmatrix} U_{11} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & U_{kk} \end{pmatrix} : U_{jj} \in \mathbb{S}^{n_j \times n_j} \right\} \subseteq \mathbb{S}^{n \times n}
\]

If \( C \) and all \( A_i \)'s are in \( \mathcal{S} \) then automatically \( S = C - \sum_i y_i A_i \in \mathcal{S} \) for any \( y \in \mathbb{R}^m \). The primal problem involves \( X \in \mathbb{S}^{n \times n} \), i.e.

\[
\begin{pmatrix}
X_{11} & X_{12} & \cdots & X_{1k} \\
\vdots & \vdots & \ddots & \vdots \\
X_{k1} & X_{k2} & \cdots & X_{kk}
\end{pmatrix}
\]
but we can restrict $X$ to $S$ without loss of generality (see HW1). Applying this to the primal-form SDP above we can assume that $X$ is diagonal, so that it reduces to the primal-form LP.