Today, we will cover a few important facts about symmetric matrices and look at the problem of minimizing the maximum eigenvalue of a matrix as an SDP problem in dual form.

First, we recall the primal and dual forms of SDP:

\[
\begin{align*}
\text{min} & \quad C \cdot X \\
\text{s.t.} & \quad AX = b \quad (\text{i.e., } A_i \cdot X = b_i, \forall i = 1, \ldots, m) \quad (P) \\
& \quad X \succeq 0,
\end{align*}
\]

\[
\begin{align*}
\text{max} & \quad b^T y \\
\text{s.t.} & \quad A^* y + S = C \\
& \quad S \succeq 0.
\end{align*}
\]

Everything you ever want to know about symmetric matrices

Fact 1 If \( P, Q \in \mathbb{R}^{m \times n} \), then

\[
P \cdot Q := \text{trace}(P^T Q) = \text{trace}(Q P^T) = \text{trace}(Q^T P) = \text{trace}(P Q^T) = \sum_i \sum_j p_{ij} q_{ij},
\]
even though \( P^T Q \) and \( Q P^T \) have different sizes (\( n \times n \) and \( m \times m \) respectively).

Fact 2 \( A \) and \( A^* \) are adjoint mappings:

\[
(A X)^T y = (A^* y) \cdot X.
\]

Fact 3 If \( P \) is a nonsingular \( n \times n \) real matrix, then \( U \) is psd (respectively, pd) if and only if \( PUP^T \) is psd (resp., pd).

Fact 4 Suppose \( Q \in \mathbb{R}^{n \times n} \) is an orthogonal matrix; then

\[
(QUQ^T) \cdot (QVQ^T) = U \cdot V.
\]

More generally, if \( P \in \mathbb{R}^{n \times n} \) is nonsingular, then

\[
(P^{-T} U P^{-1}) \cdot (P V P^T) = U \cdot V.
\]

(Generalization of \( (Q^T u)^T (Q^T v) = (P^{-1} u)^T (P v) = u^T v. \))
Proof:

\[
(P^{-T}UP^{-1}) \cdot (PVPT) = \text{trace} \left( (P^{-T}UP^{-1}PVPT) \right), \quad \text{by definition}
\]
\[
= \text{trace} \left( (P^{-T}UPVT) \right)
\]
\[
= \text{trace} \left( (UPVTPT) \right), \quad \text{by Fact 1}
\]
\[
= \text{trace} (U) = U \cdot V.
\]

\[\square\]

Note. Facts 3 and 4 show that \((P)\) is equivalent to:

\[
\min \left(P^{-T}CP^{-1}\right) \cdot \hat{X}
\]

s.t. \((P^{-T}A_iP^{-1}) \cdot \hat{X} = b_i, \forall i = 1, \ldots, m\)

\[
\hat{X} \succeq 0.
\]

This problem arises from the change of variables \(\hat{X} = PXPT\). Thus the primal variable \(X\) transforms in a different way from the data \(C\) and the \(A_i's\) and the dual slack matrix \(S\) transforms in the same way as the data.

**Fact 5** If \(Y \in \mathbb{M}^n\), then there are an orthogonal \(Q \in \mathbb{R}^{n \times n}\) and a diagonal \(\Lambda \in \mathbb{R}^{n \times n}\) such that \(U = QAQ^T\).

**Notation.** For a diagonal matrix \(\Lambda \in \mathbb{R}^{n \times n}\) whose diagonal entries are \(\lambda_1, \ldots, \lambda_n\), we write:

\[\Lambda = \text{Diag} (\lambda),\]

where \(\lambda\) is the vector \((\lambda_1; \ldots; \lambda_n)\). We also use the notation

\[\text{diag} (U) := (u_{11}; \ldots; u_{nn}),\]

for any matrix \(U \in \mathbb{R}^{n \times n}\). This is the adjoint mapping of \(\text{Diag}\).

If \(Q = [q_1, \ldots, q_n]\) (that is, \(q_i\) is its \(i\)th column vector), then

\[UQ = QA = [\lambda_1 q_1, \ldots, \lambda_n q_n].\]

Hence, looking at the \(i\)th column of \(UQ\) and \(QA\),

\[Uq_i = \lambda_i q_i.\]

So, the \(q_i's\) and \(\lambda_i's\) are the eigenvectors and eigenvalues of \(U\). We call \(QAQ^T\) the **eigenvalue decomposition** of \(U\).

We will usually assume that \(\lambda_1 \geq \ldots \geq \lambda_n\), and then

\[\lambda =: \lambda(U), \quad \Lambda =: \Lambda(U).\]

**Fact 6** The following are norms on \(\mathbb{M}^n\):

..
• The 2-norm/operator norm,

\[ ||U||_2 := \max\{||Uz||_2 \mid ||z|| = 1\} \]
\[ = ||\Lambda(U)||_2 \]
\[ = \max\{||\lambda_i(U)||\} \]
\[ = ||\Lambda(U)||_\infty; \]

• the Frobenius norm,

\[ ||U||_F := (U \cdot U)^{1/2} \]
\[ = (\Lambda(U) \cdot \Lambda(U))^{1/2} \]
\[ = ||\lambda(U)||_2 \]
\[ = ||\text{vec}(U)||_2 = ||\text{svec}(U)||_2, \]

• the nuclear norm or trace norm,

\[ ||U||_* := ||\lambda(U)||_1. \]

Note. The following observation motivates why \( || \cdot ||_* \) is called the trace norm:

\[ \text{trace} (U) = \sum_i u_{ii} = I \cdot U = I \cdot \Lambda(U) \]
\[ = \sum_i \lambda_i(U). \]

Fact 7 (Theorem 1) For \( U \in \mathbb{M}^n \), the following are equivalent:

(a) \( U \) is psd (resp., pd);

(b) \( z^T Y z \geq 0 \) for all \( z \in \mathbb{R}^n \) (resp., \( z^T Y z > 0 \) for all nonzero \( z \in \mathbb{R}^n \));

(c) \( \lambda(u) \geq 0 \) (resp., \( \lambda(u) > 0 \)); and

(d) \( U = P^T P \) for some \( P \in \mathbb{R}^{n \times n} \) (resp., for some nonsingular \( P \in \mathbb{R}^{n \times n} \))

Proof: (a) \( \Leftrightarrow \) (b) by definition. For (b) \( \Leftrightarrow \) (c), note that if \( U = Q\Lambda Q^T \), then \( z^T U z = z^T Q \Lambda Q^T z = \sum_i \lambda_i z_i \), where \( z = Q^T z \). Next, (d) \( \Rightarrow \) (b) because \( z^T U z = z^T P^T P z = ||z||_2^2 \geq 0 \) for all \( z \). For the reverse, let \( U = Q\Lambda Q^T \) and let \( P = Q\Lambda^{1/2} Q^T \), where \( \Lambda^{1/2} := \text{Diag} \left( \sqrt{\lambda_1}; \ldots; \sqrt{\lambda_n} \right) \).

Example 1 (Eigenvalue optimization) Suppose that \( U(y) \) depends linearly (affinely) on \( y \in \mathbb{R}^m \) and we want to choose \( y \) to minimize the maximum eigenvalue of \( U(y) \). Introduce \( \eta \in \mathbb{R} \) and note that

\[ \lambda_{\text{max}}(U) \leq \eta \]  
\[ \Leftrightarrow \lambda_{\text{max}}(U - \eta I) \leq 0 \]  
\[ \Leftrightarrow \lambda_{\text{min}}(\eta I - U) \geq 0 \]  
\[ \Leftrightarrow \eta I - U \succeq 0. \]
So, the problem of finding \( \min(\lambda_{\max}(U(y))) \) can be formulated as the following SDP in dual form:

\[
- \max -\eta \\
\text{s.t. } -\eta I + U(y) \preceq 0.
\]

**Corollary 1** Every psd matrix \( U \) has a (unique) psd square root \( U^{1/2} \), with \( U^{1/2}U^{1/2} = U \). Every pd matrix \( U \) is nonsingular, and its inverse is pd.

**Proof:** If \( U = QAQ^T \), then set:

\[ U^{1/2} := QA^{1/2}Q^T, \]

where \( A^{1/2} \) is as in the previous proof. If \( U \) is pd, then \( A \) has positive diagonal entries, so \( A^{-1} = \text{Diag}(\lambda^{-1}_1; \ldots; \lambda^{-1}_n) \) exists, and \( QA^{-1}Q^T \) is the inverse of \( U \), and is pd. (We won’t prove uniqueness.) \( \square \)

**Corollary 2** If \( 0 \neq u \in \mathbb{R}^n \), then \( uu^T \) is psd. (And all psd rank-one matrices are of this form—see HW1.)

**Corollary 3** \( M^n_+ \) and \( M^n_{++} \) are convex cones. \( M^n_+ \) is closed and pointed (that is, contains no one-dimensional subspaces) and its interior is \( M^n_{++} \).

**Proof:** \( M^n_+ \) is defined by the homogeneous linear (in \( U \)) inequalities

\[ z^TUz \geq 0, \forall z \in \mathbb{R}^n. \]

Hence, we see that \( M^n_+ \) is a closed convex cone.

\( M^n_{++} \) is defined by the strict homogeneous linear inequalities

\[ z^TUz > 0, \forall 0 \neq z \in \mathbb{R}^n, \]

so \( M^n_{++} \) is a convex cone.

If \( U \in (M^n_+) \cap (-M^n_+) \), then \( \lambda(U) \geq 0 \) and \( \lambda(U) \leq 0 \). So, \( U \) must be 0. This shows that \( M^n_+ \) is pointed.

If \( U \in M^n_{++} \), then \( \lambda := \lambda_{\min}(U) > 0 \). Let \( V \in M^n \) have \( ||V||_2 \leq \hat{\lambda} \). Then,

\[ z^T(U + V)z = z^T((U - \hat{\lambda}I) + \hat{\lambda}I + V)z \]
\[ \geq \hat{\lambda}z^Tz + z^TVz \]
\[ \geq \hat{\lambda} - \hat{\lambda} = 0, \]

for \( ||z||_2 = 1 \). So, \( M^n_{++} \subseteq \text{int}(M^n_+) \). Now, suppose that \( U \notin M^n_{++} \). Then, there exists a nonzero vector \( z \) such that \( z^TUz \leq 0 \). But then

\[ z^T(U - \epsilon zz^T)z \leq 0 - \epsilon(z^Tz)^2 < 0 \]

for all \( \epsilon > 0 \), which shows that \( U \notin \text{int}(M^n_+) \). Hence, \( \text{int}(M^n_+) = M^n_{++} \). \( \square \)
**Fact 8** If \( U \succeq 0 \) (resp., \( U \succ 0 \)), then \( u_{jj} \geq 0 \) (resp., \( u_{jj} > 0 \)) for all \( j = 1, \ldots, n \), and if \( u_{jj} = 0 \), then \( u_{jk} = 0 \) for all \( k \).

**Fact 9** If \( U \succeq 0 \), then \( PUP^T \) is psd for all \( P \in \mathbb{R}^{m \times n} \). If \( U \succ 0 \) and \( P \) has full row rank, then \( PUP^T \succ 0 \). Hence, if \( P \) is a permutation matrix, we see that every principal rearrangement of \( U \) is psd (resp., pd) if \( U \) is psd (resp., if \( U \) is pd). If

\[
U = \begin{bmatrix}
U_{11} & U_{12} \\
U_{21} & U_{22}
\end{bmatrix} \succeq 0, \ (\text{resp., } \succ 0),
\]

then \( U_{11} \succeq 0 \) (resp., \( \succ 0 \)), and similarly for every principal submatrix of \( U \).