Homework 6

ECE 269: Linear Algebra and Applications
Homework #6-Solution
Instructor: Behrouz Touri

1. For any matrix $A \in \mathbb{R}^{n\times n}$ show that

$$\dim(\text{span}\{A^k \mid k \geq 1\}) = \dim(\text{span}\{I, A, A^2, \ldots, A^k, \ldots\}) \leq n.$$

Hint: Show that for any $k \geq 0$, A^k is a linear combination of $I, \ldots, A^{n-1}$.

Solution: We claim that for any $k \geq 1$, $A^k \in \text{span}\{I, A, \ldots, A^{n-1}\}$. Then we have

$$\text{span}\{A^k \mid k \geq 1\} \subset \text{span}\{I, A, \ldots, A^{n-1}\}$$

and thus $\dim(\{A^k \mid k \geq 1\}) \leq n$. In fact, for $k < n-1$, this is clearly true. The case for $k = n$ is a consequence of the Cayley–Hamilton theorem, which states that the characteristic polynomial of A has a term of degree n with coefficient 1, implying A^n can be written as a linear combination of lower powers of A. Suppose our claim is true for all $k \leq N$, where $N \geq n$. For $k = N+1$, since $A^N \in \text{span}\{I, A, \ldots, A^{n-1}$, $A^{N+1} \in \text{span}\{A, \ldots, A^n\} \subset \text{span}\{I, A, \ldots, A^{n-1}\}$ as A^n can be written as a linear combination of $I, A, \ldots, A^{n-1}$. Therefore, by induction, our claim holds for all $k \geq 1$.

2. Show that if A and B are similar, then not only their eigenvalues of the two matrices are the same, but also the algebraic and geometric multiplicity of them are the same for the two matrices.

Solution: Let c$B = PAP^{-1}$. Since $\det(\lambda I - B) = \det(P(\lambda I - A)P^{-1}) = \det(P)\det(\lambda I - A)\det(P^{-1})$, the characteristic polynomial of A and B are the same, and hence the algebraic multiplicity of them are also the same. Now, let $v_1, \ldots, v_r$ be a linearly independent set of eigenvalues of A corresponding to eigenvalue λ_0. Then we have

$$BPv_i = PAP^{-1}Pv_i = \lambda_0 v_i,\ \forall i \in [r]$$

and the equation

$$c_1Pv_1 + \ldots + c_rPv_r = 0$$

has only trivial solution since P is invertible. Therefore, $Pv_1, \ldots, Pv_r$ is a linearly independent set of eigenvalues of B. This implies that the geometric multiplicity of B is greater than or equal to A. On the other hand, since $B = P^{-1}AP$, we can apply similar arguments with P replaced by P^{-1} to show that the geometric multiplicity of A is greater than or equal to B. Therefore, the the geometric multiplicity of A and B must be the same.

3. *A computational problem.*

   (a) Find the eigenvalues of the matrix:

$$A = \begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 2 \end{bmatrix}.$$

(b) Show that A does not have an eigenvalue decomposition.

(c) Provide the Jordan Decomposition of A.

Solution:

(a) We can see that the characteristic polynomial of A is $(\lambda - 1)^2(\lambda - 2)$. The eigenvalues of matrix A are $\lambda_1 = 1$ with algebraic multiplicity 2, $\lambda_2 = 2$.

(b) The normalized eigenvectors corresponding to eigenvalue λ_1 and λ_2 are given by $v_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$ and $v_2 = \begin{bmatrix} \frac{2}{\sqrt{6}} \\ \frac{1}{\sqrt{6}} \\ \frac{1}{\sqrt{6}} \end{bmatrix}$. Now let us look at the geometric multiplicity of λ_1, given by,

$$\text{Nullity}(A - \lambda_1 I) = \text{Nullity} \begin{bmatrix} 0 & 1 & 1 \\ 0 & 0 & 1 \\ 0 & 0 & 1 \end{bmatrix} = 1.$$

Since geometric multiplicity of λ_1 is not equal to the algebraic multiplicity of λ_1, the matrix is not diagonalizable and hence it does not have an eigenvalue decomposition.

(c) For the Jordan decomposition of A, note that $J = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{bmatrix}$. To find the T-matrix, we need to obtain the generalized eigenvector, v_{12}, corresponding to λ_1 by solving the following equation,

$$Av_{12} = v_1 + \lambda_1 v_{12},$$

giving us one possible $v_{12} = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$.

Using the generalized eigenvectors to build the T-matrix gives us,

$$T = \begin{bmatrix} 1 & 0 & \frac{2}{\sqrt{6}} \\ 0 & 1 & \frac{1}{\sqrt{6}} \\ 0 & 0 & \frac{1}{\sqrt{6}} \end{bmatrix}$$

We can now compute its inverse to get

$$T^{-1} = \begin{bmatrix} 1 & 0 & -2 \\ 0 & 1 & -1 \\ 0 & 0 & \sqrt{6} \end{bmatrix}.$$

4. *Properties of symmetric matrices.* Let $A = A' \in \mathbb{R}^{n \times n}$ and $B = B' \in \mathbb{R}^{n \times n}$. Prove or provide a counterexample to each of the following statements.

(a) If $A \succeq 0$, then $X'AX \succeq 0$ for every $X \in \mathbb{R}^{n \times k}$.

(b) If $A \succeq 0$ and $B \succeq 0$, then $\text{trace}(AB) \geq 0$.

(c) If $A \succeq 0$, then $A + B \succeq B$.

(d) If $A \succeq B$, then $-B \succeq -A$.

(e) If $A \succeq I$, then $I \succeq A^{-1}$.

(f) If $A \succeq B \succ 0$, then $B^{-1} \succeq A^{-1} \succ 0$.

(g) If $A \succeq B \succeq 0$, then $A^2 \succeq B^2$.

Solution:

(a) Given any $y \in \mathbb{R}^n$, $y'(X'AX)y = (Xy)'A(Xy) \geq 0$, since $A \succeq 0$. Thus, $X'AX \succeq 0$. We can in fact show that if $A \succ 0$ and X is full-rank and tall, then $X'AX \succ 0$. To see this, consider $y'X'AXy \geq 0$ with equality only if $Xy = 0$. But since the columns of X are linearly independent, $Xy = 0$ if and only if $y = 0$.

(b) Let A admit an eigendecomposition $A = Q\Lambda Q'$. Since A is PD, Λ has non-negative diagonal elements and thus has a square root. Similarly, let B admit an eigendecomposition $B = V\Sigma V'$. Then

$$\begin{aligned}
\text{trace}(AB) &= \text{trace}(Q\Lambda Q'V\Sigma V') \\
&= \text{trace}(\Lambda Q'V\Sigma V'Q) \\
&= \text{trace}(\Lambda^{1/2}Q'V\Sigma^{1/2}\Sigma^{1/2}V'Q\Lambda^{1/2}) \\
&= \text{trace}((\Sigma^{1/2}V'Q\Lambda^{1/2})'\Sigma^{1/2}V'Q\Lambda^{1/2}).
\end{aligned}$$

Besides, for any matrix $W = [w_1\ w_2\ \ldots\ w_n]$, we have

$$\begin{aligned}
\text{trace}(W'W) &= \sum_{i=1}^{n} (W'W)_{ii} \\
&= \sum_{i=1}^{n} \|w_i\|^2 \\
&\geq 0.
\end{aligned}$$

Therefore, $\text{trace}(AB) \geq 0$.

(c) Since $A = (A + B) - B \succeq 0$, we have $A + B \succeq B$.

(d) If $A \succeq B$, we have $A - B \succeq 0$, which implies that $-B - (-A) \succeq 0$, and thus that $-B \succeq -A$.

(e) Since A is symmetric, we have $A = Q\Lambda Q'$, where $QQ' = I$. Thus, $A - I = Q(\Lambda - I)Q'$, and the eigenvalues of $A - I$ are the eigenvalues of A minus 1. Thus, if $A - I$ is positive semidefinite, every eigenvalue of $A \geq 1$. Now $A^{-1} = Q\Lambda^{-1}Q'$ with eigenvalues ≤ 1. Hence, $(I - A^{-1}) = Q(I - \Lambda^{-1})Q'$ is positive semidefinite.

(f) Since $B = Q\Lambda Q' \succ 0$, $B^{1/2} = Q\Lambda^{1/2}Q' \succ 0$. Then by part (a), $A - B \succeq 0$ implies $B^{-1/2}(A - B)B^{-1/2} = B^{-1/2}AB^{-1/2} - I \succeq 0$. Hence by part (e), $I - B^{1/2}A^{-1}B^{1/2} \succeq 0$. Finally, by part (a) once again, $B^{-1/2}(I - B^{1/2}A^{-1}B^{1/2})B^{-1/2} = B^{-1} - A^{-1} \succeq 0$.

(g) This need not be true. Let $A = \begin{bmatrix} 2 & 1 \\ 1 & 1 \end{bmatrix}$ and $B = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$. Then $A \succeq B \succeq 0$. But $A^2 - B^2 = \begin{bmatrix} 3 & 1 \\ 1 & 0 \end{bmatrix}$ is indefinite.

5. *Induced matrix norms.* We define the induced p-norm of $A \in \mathbb{C}^{m \times n}$ for $p \in [1, \infty]$ as

$$\|A\|_p = \max_{x \neq 0} \frac{\|Ax\|_p}{\|x\|_p}.$$

When $p = 2$, $\|A\|_2$ is called the *spectral norm* of the matrix. One can view such norms as the maximum attenuation of the corresponding linear mapping on the unit ball.

(a) Show that $\|A\|_p$ satisfies the axioms of matrix norms.

(b) Show that
$$\|A\|_1 = \max_j \sum_i |A_{ij}|.$$

(c) Show that
$$\|A\|_\infty = \max_i \sum_j |A_{ij}| = \|A^*\|_1.$$

Solution:

(a)
- *Absolute homogeneity:* By properties of the p norms of vectors, for all $\alpha \in \mathbb{C}$ we have $\|\alpha Ax\|_p = |\alpha|\|Ax\|_p$. Thus,
$$\|\alpha A\|_p = \max_{x\neq 0} \frac{\|\alpha Ax\|_p}{\|x\|_p} = |\alpha| \max_{x\neq 0} \frac{\|Ax\|_p}{\|x\|_p} = |\alpha|\|A\|_p.$$
- *Triangle Inequality:*
$$\|A+B\|_p = \max_{x\neq 0} \frac{\|Ax+Bx\|_p}{\|x\|_p} \le \max_{x\neq 0} \frac{\|Ax\|_p + \|Bx\|_p}{\|x\|_p} \le \max_{x\neq 0} \frac{\|Ax\|_p}{\|x\|_p} + \max_{y\neq 0} \frac{\|By\|_p}{\|y\|_p} = \|A\|_p + \|B\|_p.$$
- *Positive definiteness:* The vector p-norm $\|\cdot\|_p$ is non-negative, which immediately implies that $\|A\|_p \ge 0$. Furthermore, for $\|A\|_p = 0$, we need $\max\limits_{x\neq 0} \|Ax\|_p = 0$, which is only possible if $A = 0$. Conversely, $A = 0$ implies that $\|Ax\|_p = 0$. Thus, $\|A\|_p = 0$ iff $A = 0$.

(b) Let $x = \begin{bmatrix} x_1 & x_2 & \cdots & x_n \end{bmatrix}$, and the columns of A be $\begin{bmatrix} a_1 & a_2 & \cdots & a_n \end{bmatrix}$. Then,
$$\begin{aligned}
\|Ax\|_1 &= \|a_1x_1 + a_2x_2 + \cdots + a_nx_n\|_1 \\
&\le \|a_1x_1\|_1 + \|a_2x_2\|_1 + \cdots + \|a_nx_n\|_1 \\
&= |x_1|\|a_1\|_1 + |x_2|\|a_2\|_1 + \cdots + |x_n|\|a_n\|_1.
\end{aligned}$$

Note that an alternative way of characterizing the 1-norm is $\max\limits_{\|x\|_1=1} \|Ax\|_1$. Using the upper bound on $\|Ax\|_1$, we obtain
$$\begin{aligned}
\max_{\|x\|_1=1} \|Ax\|_1 &\le \max_{\|x\|_1=1} |x_1|\|a_1\|_1 + |x_2|\|a_2\|_1 + \cdots + |x_n|\|a_n\|_1 \\
&= \max_{\sum_{i=1}^n |x_i|=1} |x_1|\|a_1\|_1 + |x_2|\|a_2\|_1 + \cdots + |x_n|\|a_n\|_1 \\
&\le \max_{\sum_{i=1}^n |x_i|=1} \max_i \|a_i\|_1 (|x_1| + |x_2| + \cdots + |x_n|) \\
&= \max_i \|a_i\|_1 = \max_j \sum_i |A_{ij}|
\end{aligned}$$

Let $j = \text{argmax}_i \|a_i\|_1$. The upper bound is achievable by taking $x = e_j$, and hence $\|A\|_1 = \max_j \sum_i |A_{ij}|$.

(c) Let the rows of A be $\tilde{a}_1', \tilde{a}_2', \cdots, \tilde{a}_m'$. We then need to find $\max_{\|x\|_\infty=1} \|Ax\|_\infty = \max_j |\tilde{a}_j' x|$ over all $\|x\|_\infty = 1$. Now for all j,

$$|\tilde{a}_j' x| = |\sum_k \tilde{a}_{jk} x_k| \leq \sum_k |\tilde{a}_{jk}||x_k| \leq \sum_k |\tilde{a}_{jk}| \leq \max_i \|\tilde{a}_i\|_1.$$

Therefore, $\|A\|_\infty \leq \max_i \|\tilde{a}_i\|_1 = \max_i \sum_j |A_{ij}|$. If $i_1 = \text{argmax}_i \sum_j |A_{ij}|$, by taking

$$x = \begin{bmatrix} \frac{\overline{A_{i_1 1}}}{|A_{i_1 1}|} & \frac{\overline{A_{i_1 2}}}{|A_{i_1 2}|} & \cdots & \frac{\overline{A_{i_1 n}}}{|A_{i_1 n}|} \end{bmatrix}^T,$$

we note that the upper bound is indeed achievable and thus $\|A\|_\infty = \max_i \sum_j |A_{ij}|$.

6. *Properties of the spectral norm.*

(a) Show that $\|A^*A\| = \|A\|^2$.

(b) Show that the spectral norm is *unitarily invariant*, namely, $\|UAV\| = \|A\|$ for any unitary matrices U and V.

(c) Show that

$$\left\| \begin{bmatrix} A & 0 \\ 0 & B \end{bmatrix} \right\| = \max(\|A\|, \|B\|).$$

Solution:

(a) Since A^*A is Hermitian, its singular values are the same as its eigenvalues. As mentioned in the discussion session, the largest eigenvalue (and hence the largest singular value) is $\sigma_1^2(A) = \|A\|^2$.

(b) Since $\|Ux\| = \|x\|$ for every unitary matrix U,

$$\|UAV\| = \max_{x\neq 0} \frac{\|UAVx\|}{\|x\|} = \max_{x\neq 0} \frac{\|AVx\|}{\|x\|} = \max_{x\neq 0} \frac{\|AVx\|}{\|Vx\|}. \quad (1)$$

Since V is a unitary transformation, $\{x|x \neq 0\} = \{x|Vx \neq 0\}$. Substituting $Vx = y$ and continuing the chain of equalities from (1) yields

$$\max_{x\neq 0} \frac{\|AVx\|}{\|Vx\|} = \max_{y\neq 0} \frac{\|Ay\|}{\|y\|} = \|A\|.$$

(c) If $U_A \Sigma_A V_A^*$ is an SVD of A, and $U_B \Sigma_B V_B^*$ is an SVD of B, then

$$\begin{bmatrix} U_A & 0 \\ 0 & U_B \end{bmatrix} \begin{bmatrix} \Sigma_A & 0 \\ 0 & \Sigma_B \end{bmatrix} \begin{bmatrix} V_A^* & 0 \\ 0 & V_B^* \end{bmatrix}$$

is an SVD of $\begin{bmatrix} A & 0 \\ 0 & B \end{bmatrix}$. This shows that the singular values of $\begin{bmatrix} A & 0 \\ 0 & B \end{bmatrix}$ are the union of the singular values of A and B (including multiplicity), which in turn implies that

$$\left\| \begin{bmatrix} A & 0 \\ 0 & B \end{bmatrix} \right\| = \max(\|A\|, \|B\|).$$