E2 212: Matrix Theory Fall 2019 – Test 2 Solutions

1. (**Jordan Form**) The following information on a 6×6 matrix is given below.

Eigenvalue	Algebraic multiplicity	Geometric multiplicity
2	4	2
-1	2	2

$$\begin{aligned} \text{rank}(\mathbf{A} - 2\mathbf{I}) &= 4 & \text{rank}(\mathbf{A} + \mathbf{I}) &= 4 \\ \text{rank}\left((\mathbf{A} - 2\mathbf{I})^2\right) &= 3 & \text{rank}\left((\mathbf{A} + \mathbf{I})^2\right) &= 4 \\ \text{rank}\left((\mathbf{A} - 2\mathbf{I})^3\right) &= 2 & & \\ \text{rank}\left((\mathbf{A} - 2\mathbf{I})^4\right) &= 2 & & \end{aligned}$$

Find the Jordan form of the matrix. (8 points)

Solution:
For $\lambda = 2$, $n - a_\lambda = 2$. Hence,

$$\begin{aligned} k &= \min_j(\text{rank}(\mathbf{A} - 2\mathbf{I})^j = n - a_\lambda) = 3 \\ r_{k-1} &= N_k + n - a_\lambda \\ \Rightarrow \quad 3 &= N_3 + 6 - 4 \\ \Rightarrow \quad N_3 &= 1 \end{aligned}$$

Thus, there is a Jordan block $J(1)$ of size 3×3.

$$\begin{aligned} r_{k-2} &= N_{k-1} + 2N_k + n - a_\lambda \\ \Rightarrow \quad r_1 &= N_2 + 2N_3 + 6 - 4 \\ \Rightarrow \quad 4 &= N_2 + 2 + 2 \\ \Rightarrow \quad N_2 &= 0 \end{aligned}$$

Hence, there is no Jordan block of size 2×2. Finally,

$$\begin{aligned} a_\lambda &= N_1 + 2N_2 + 3N_3 \\ \Rightarrow \quad 4 &= N_1 + 3 \\ \Rightarrow \quad N_1 &= 1 \end{aligned}$$

Thus, there is one Jordan block $J(1)$ of size 1×1.

For $\lambda = -1$, the algebraic multiplicity is 2 and the geometric multiplicity is 2, thus there are two Jordan blocks in the 2×2 submatrix corresponding to this eigen value, namely, two 1×1 blocks $J(-1)$. Putting all this together, we get the Jordan form as

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

2. (a) For $\mathbf{A} \in \mathbb{C}^{n \times n}$, prove that if $\mathbf{x}^H \mathbf{A} \mathbf{x} = 0$ for all $\mathbf{x} \in \mathbb{C}^{n \times 1}$, then $\mathbf{A} = 0$. (6 points)

Solution:

Consider Schur's triangularization, $\mathbf{A} = \mathbf{U}\mathbf{T}\mathbf{U}^H$, where $\mathbf{T}$ is upper triangular. Then,

$$\begin{aligned} \mathbf{x}^H \mathbf{A} \mathbf{x} &= 0 \quad \forall \mathbf{x} \in \mathbb{C}^{N \times 1} \\ \Rightarrow \mathbf{x}^H \mathbf{U}\mathbf{T}\mathbf{U}^H \mathbf{x} &= 0 \quad \forall \mathbf{x} \in \mathbb{C}^{N \times 1} \\ \Rightarrow \qquad \mathbf{y}^H \mathbf{T} \mathbf{y} &= 0 \quad \forall \mathbf{y} = \mathbf{U}^H \mathbf{x} \in \mathbb{C}^{N \times 1}. \end{aligned}$$

Choosing $\mathbf{y} = \mathbf{e}_i$, we get

$$\begin{aligned} \mathbf{e}_i^H \mathbf{T} \mathbf{e}_i &= 0 \\ \Rightarrow \qquad t_{ii} &= 0 \quad \forall i \end{aligned} \tag{1}$$

Note that, since $\mathbf{T}$ is upper triangular,

$$t_{ji} = 0 \quad \forall i > j. \tag{2}$$

Now, we choose $\mathbf{y} = \mathbf{e}_i + \mathbf{e}_j$ with $i > j$ to get

$$\begin{aligned} \Rightarrow (\mathbf{e}_i + \mathbf{e}_j)^H \mathbf{T} (\mathbf{e}_i + \mathbf{e}_j) &= 0 \\ \Rightarrow \qquad t_{ii} + t_{jj} + t_{ij} + t_{ji} &= 0 \\ \Rightarrow \qquad t_{ij} &= 0. \end{aligned} \tag{3}$$

From (1), (2) and (3), $\mathbf{T} = \mathbf{0}$. Therefore, $\mathbf{A} = \mathbf{0}$.

(b) Show that $\mathbf{x}^T \mathbf{A} \mathbf{x} = 0$ for all $\mathbf{x} \in \mathbb{R}^{n \times 1}$ does not imply $\mathbf{A} = \mathbf{0}$, even if $\mathbf{A}$ is real. (4 points)

Solution:

Choose $\mathbf{A} = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$ and let $\mathbf{x} = [x_1 \ x_2]^T$. Then, $\mathbf{x}^T \mathbf{A} \mathbf{x} = x_1 x_2 - x_2 x_1 = 0$ for all $\mathbf{x} \in \mathbb{R}^{2 \times 1}$, even though $\mathbf{A} \neq \mathbf{0}$.

3. (**Spectral norm of Hermitian matrices**)

(a) Show that, for a Hermitian symmetric matrix $\mathbf{A}$, the spectral norm is equal to its spectral radius (i.e., its maximum absolute eigenvalue). (5 points)

Solution:

Since $\mathbf{A}$ is Hermitian symmetric, it is unitarily diagonalizable, and we can write $\mathbf{A} = \mathbf{U}\mathbf{D}\mathbf{U}^H$, where $\mathbf{U}$ is unitary and $\mathbf{D}$ is diagonal.

Hence, the spectral norm $|||\mathbf{A}|||_2 = |||\mathbf{D}|||_2 = \max_{\|\mathbf{x}\|_2=1} \|\mathbf{D}\mathbf{x}\|_2$.

Now, since $\mathbf{D}$ is diagonal, $\|\mathbf{D}\mathbf{x}\|_2^2 = \sum_{i=1}^n |\lambda_i|^2 |x_i|^2 \leqslant \max_j |\lambda_j|^2 \sum_{i=1}^n |x_i|^2 = \max_j |\lambda_j|^2 \|\mathbf{x}\|_2^2$, and equality is attained when $\mathbf{x} = \mathbf{e}_j$, where j is the index corresponding to the largest eigenvalue.

Therefore, we have that the spectral norm of $\mathbf{A}$ equals $\max_j |\lambda_j|$, which is nothing but its spectral radius.

(b) Suppose $\mathbf{S}, \mathbf{T} \in \mathbb{C}^{n \times n}$ are Hermitian symmetric matrices with (increasing) ordered eigenvalues $\{\lambda_k(\mathbf{S})\}_{k=1}^n, \{\lambda_k(\mathbf{T})\}_{k=1}^n$, respectively. Also, let $\{\lambda_k(\mathbf{S}-\mathbf{T})\}_{k=1}^n$ and $\{\lambda_k(\mathbf{S}+\mathbf{T})\}_{k=1}^n$ be the ordered eigenvalues of $\mathbf{S}-\mathbf{T}$ and $\mathbf{S}+\mathbf{T}$, respectively. Then, show that:

i. $\lambda_k(\mathbf{S}) - \lambda_n(\mathbf{T}) \leqslant \lambda_k(\mathbf{S}-\mathbf{T}) \leqslant \lambda_k(\mathbf{S}) - \lambda_1(\mathbf{T})$, $k = 1, \ldots, n$. (5 points)

Solution:

By Weyl's inequality,

$$\lambda_k(\mathbf{S}) + \lambda_1(\mathbf{T}) \leqslant \lambda_k(\mathbf{S}+\mathbf{T}) \leqslant \lambda_k(\mathbf{S}) + \lambda_n(\mathbf{T}). \tag{4}$$

Replacing $\mathbf{T}$ by $-\mathbf{T}$, we get

$$\begin{aligned}
\lambda_k(\mathbf{S}) + \lambda_1(-\mathbf{T}) \leqslant \lambda_k(\mathbf{S}+(-\mathbf{T})) &\leqslant \lambda_k(\mathbf{S}) + \lambda_n(-\mathbf{T}) \\
\lambda_n(\mathbf{T}) &= -\lambda_1(-\mathbf{T}) \\
\lambda_1(\mathbf{T}) &= -\lambda_n(-\mathbf{T}) \\
\lambda_k(\mathbf{S}) - \lambda_n(\mathbf{T}) \leqslant \lambda_k(\mathbf{S}-\mathbf{T})) &\leqslant \lambda_k(\mathbf{S}) - \lambda_1(\mathbf{T}),
\end{aligned} \tag{5}$$

which is the desired result.

ii. $\lambda_n(\mathbf{T}) - \lambda_1(\mathbf{T}) \geqslant \lambda_k(\mathbf{S}+\mathbf{T}) + \lambda_k(\mathbf{S}-\mathbf{T}) - \lambda_k(\mathbf{S})$, $k = 1, \ldots, n$. (5 points)

Solution:

Using upper bounds of (4) and (5),

$$\begin{aligned}
\lambda_k(\mathbf{S}+\mathbf{T}) + \lambda_k(\mathbf{S}-\mathbf{T}) &\leqslant 2\lambda_k(\mathbf{S}) + \lambda_n(\mathbf{T}) - \lambda_1(\mathbf{T}) \\
\Rightarrow \lambda_k(\mathbf{S}+\mathbf{T}) + \lambda_k(\mathbf{S}-\mathbf{T}) - 2\lambda_k(\mathbf{S}) &\leqslant \lambda_n(\mathbf{T}) - \lambda_1(\mathbf{T}).
\end{aligned}$$

iii. $|||\mathbf{S}-\mathbf{T}|||_2 \geqslant \max_i |\lambda_i(\mathbf{S}) - \lambda_i(\mathbf{T})|$. *Hint: Use part (a).* (5 points)

Solution:

Writing $\mathbf{S} = \mathbf{T} + (\mathbf{S}-\mathbf{T})$ and using Weyl's inequality, we have

$$\begin{aligned}
\lambda_k(\mathbf{T}) + \lambda_1(\mathbf{S}-\mathbf{T}) \leqslant \lambda_k(\mathbf{S}) &\leqslant \lambda_k(\mathbf{T}) + \lambda_n(\mathbf{S}-\mathbf{T}) \\
\lambda_1(\mathbf{S}-\mathbf{T}) \leqslant \lambda_k(\mathbf{S}) - \lambda_k(\mathbf{T}) &\leqslant \lambda_n(\mathbf{S}-\mathbf{T})
\end{aligned} \tag{6}$$

There are three cases to be considered:

(Case a) $\lambda_1(\mathbf{S}-\mathbf{T}), \lambda_n(\mathbf{S}-\mathbf{T})$ are both negative.
In this case, $\max_i |\lambda_i(\mathbf{S}-\mathbf{T})| = |\lambda_1(\mathbf{S}-\mathbf{T})|$. Using the lower bound of (6), we get

$$|\lambda_k(\mathbf{S}) - \lambda_k(\mathbf{T})| \leqslant |\lambda_1(\mathbf{S}-\mathbf{T})|$$

(Case b) $\lambda_1(\mathbf{S}-\mathbf{T}), \lambda_n(\mathbf{S}-\mathbf{T})$ are both positive.
In this case, $\max_i |\lambda_i(\mathbf{S}-\mathbf{T})| = |\lambda_n(\mathbf{S}-\mathbf{T})|$. Using the upper bound of (6), we get

$$|\lambda_k(\mathbf{S}) - \lambda_k(\mathbf{T})| \leqslant |\lambda_n(\mathbf{S}-\mathbf{T})|$$

(Case c) $\lambda_1(\mathbf{S}-\mathbf{T})$ is negative $\lambda_n(\mathbf{S}-\mathbf{T})$ is positive.
In this case, if $|\lambda_1(\mathbf{S}-\mathbf{T})| > |\lambda_n(\mathbf{S}-\mathbf{T})|$, (case a) follows, otherwise (case b) follows.

Combining the three cases, for all k, we have

$$|\lambda_k(\mathbf{S}) - \lambda_k(\mathbf{T})| \leqslant \max_i |\lambda_i(\mathbf{S}-\mathbf{T})|,$$

and the right hand side is precisely the spectral norm of $\mathbf{S}-\mathbf{T}$, and thus

$$\max_k |\lambda_k(\mathbf{S}-\mathbf{T})| \leqslant \|\!|\mathbf{S}-\mathbf{T}|\!\|_2.$$

4. (**LDM^T factorization**) Show that the product of the pivots obtained during Gaussian elimination (LU factorization) of a matrix is equal to the product of its eigenvalues. (6 points)

Solution:

$$\begin{aligned}
\det(\mathbf{A}) &= \det(\mathbf{LDM}^T) \\
&= \det(\mathbf{L}) \times \det(\mathbf{D}) \times \det(\mathbf{M}^T) \\
&= 1 \times \det(\mathbf{D}) \times 1 \\
\Rightarrow \prod_{i=1}^{n} \lambda_i &= \prod_{j=1}^{n} k_j
\end{aligned}$$

where k_j are the pivot elements. The same applies to the factorization $\mathbf{A} = \mathbf{LU}$; $\det(\mathbf{L}) = 1$ since $\mathbf{L}$ is a unit lower triangular matrix, and $\det(\mathbf{U})$ is the product of the pivot elements since $\mathbf{U}$ is upper triangular with the pivot elements along its diagonal.

5. Let $\mathbf{A}, \mathbf{B} \in \mathbb{C}^{n\times n}$ be given, and suppose $\mathbf{A}$ and $\mathbf{B}$ are simultaneously similar to upper triangular matrices: that is, $\mathbf{S}^{-1}\mathbf{AS}$ and $\mathbf{S}^{-1}\mathbf{BS}$ are both upper triangular for some nonsingular $\mathbf{S}$. Show that every eigenvalue of $\mathbf{AB} - \mathbf{BA}$ must be zero. (8 points)

   Solution:
   Since $\mathbf{A}$ and $\mathbf{B}$ are simultaneously similar to upper triangular matrices, define $\mathbf{T} = \mathbf{S}^{-1}\mathbf{AS}$ and $\mathbf{R} = \mathbf{S}^{-1}\mathbf{BS}$, where $\mathbf{T}$ and $\mathbf{R}$ are upper triangular. Then, $\mathbf{STS}^{-1} = \mathbf{A}$ and $\mathbf{SRS}^{-1} = \mathbf{B}$. Thus, $\mathbf{AB} = \mathbf{STRS}^{-1}$ and $\mathbf{BA} = \mathbf{SRTS}^{-1}$.

   Define $\mathbf{C} = \mathbf{AB} - \mathbf{BA} = \mathbf{S}(\mathbf{TR} - \mathbf{RT})\mathbf{S}^{-1}$
   Thus, $\mathbf{C}$ and $(\mathbf{TR} - \mathbf{RT})$ are similar. This implies that the eigenvalues of $\mathbf{C}$ are the same as those of $\mathbf{TR} - \mathbf{RT}$. Now, $\mathbf{TR} - \mathbf{RT}$ is upper triangular, and hence its eigenvalues are its diagonal elements.

$$
\begin{aligned}
(\mathbf{TR})_{ii} &= \mathbf{T}(i,:)\mathbf{R}(:,i) \\
&= \sum_{j=1}^{n} t_{ij} r_{ji} \\
&= t_{ii} r_{ii} \quad \text{[the other elements of the sum are zero as they are upper triangular]} \\
(\mathbf{RT})_{ii} &= \mathbf{R}(i,:)\mathbf{T}(:,i) \\
&= \sum_{j=1}^{n} r_{ij} t_{ji} \\
&= r_{ii} t_{ii} \quad \text{[the other elements of the sum are zero as they are upper triangular]}
\end{aligned}
$$

$$\Rightarrow (\mathbf{TR} - \mathbf{RT})_{ii} = 0$$

   Since all the diagonal entries are equal to zero, the eigenvalues are also zero.

6. Let $\lambda, a \in \mathbb{R}$, $\mathbf{y} \in \mathbb{C}^n$, and $\mathbf{A} = \begin{bmatrix} \lambda\mathbf{I}_n & \mathbf{y} \\ \mathbf{y}^H & a \end{bmatrix} \in \mathbb{C}^{(n+1)\times(n+1)}$. Use the Cauchy interlacing theorem to show that λ is an eigenvalue of $\mathbf{A}$ with multiplicity at least $n-1$. What are the other two eigenvalues? (8 points)

   Solution:
   The matrix $\lambda\mathbf{I}_n$ has n repeated eigenvalues equal to λ. Therefore, by the Cauchy interlacing theorem, the eigenvalues $(\lambda_1, \lambda_2, \ldots, \lambda_{n+1})$ of $\mathbf{A}$ satisfy the inequalities $\lambda_1 \leqslant \lambda \leqslant \lambda_2 \leqslant \lambda \leqslant \ldots \leqslant \lambda_n \leqslant \lambda \leqslant \lambda_{n+1}$

$$\Rightarrow \lambda_2 = \lambda_3 = \ldots = \lambda_n = \lambda$$

   Therefore, λ is eigenvalue of $\mathbf{A}$ with multiplicity at least $n-1$.

   To find the value of λ_1 and λ_{n+1}, we use the trace and determinant of $\mathbf{A}$

$$\begin{aligned}\text{Trace}(\mathbf{A}) &= n\lambda + a \\ \Rightarrow \lambda_1 + \lambda_{n+1} + (n-1)\lambda &= n\lambda + a \\ \Rightarrow \lambda_1 + \lambda_{n+1} &= \lambda + a \end{aligned} \tag{7}$$

Similarly, using

$$\det(\mathbf{A}) = \det\left(\begin{bmatrix} \lambda \mathbf{I}_n & \mathbf{y} \\ \mathbf{y}^H & a \end{bmatrix}\right)$$

Using Schur's complement,

$$\begin{aligned} &= \det\left(\begin{bmatrix} \lambda \mathbf{I}_n & \mathbf{0} \\ \mathbf{0} & a - \frac{1}{\lambda}\mathbf{y}^H \mathbf{I}_n \mathbf{y} \end{bmatrix}\right) \\ \Rightarrow \lambda_1 \lambda_{n+1} \lambda^{n-1} &= \lambda^n (a - \frac{1}{\lambda}\mathbf{y}^H \mathbf{y}) \\ \Rightarrow \lambda_1 \lambda_{n+1} &= a\lambda - \mathbf{y}^H \mathbf{y} \end{aligned} \tag{8}$$

From (7) and (8), we form the quadratic equation

$$\begin{aligned} \frac{a\lambda - \mathbf{y}^H \mathbf{y}}{\lambda_{n+1}} + \lambda_{n+1} &= \lambda + a \\ \Rightarrow \lambda_{n+1}^2 - (\lambda + a)\lambda_{n+1} + (a\lambda - \mathbf{y}^H \mathbf{y}) &= 0. \end{aligned}$$

Solving,

$$\begin{aligned} \lambda_{n+1} &= \frac{\lambda + a \pm \sqrt{(\lambda + a)^2 - 4a\lambda + 4\mathbf{y}^H \mathbf{y}}}{2} \\ &= \frac{\lambda + a \pm \sqrt{(\lambda - a)^2 + 4\mathbf{y}^H \mathbf{y}}}{2} \\ \Rightarrow \lambda_1 &= \frac{(\lambda + a) \mp \sqrt{(\lambda - a)^2 + 4\mathbf{y}^H \mathbf{y}}}{2}. \end{aligned}$$

It is clear that the "+" yields λ_{n+1} and "−" yields λ_1. An alternative, perhaps simpler way to obtain the same solution for λ_1 and λ_{n+1} is to use the fact that the squared Frobenius norm of the matrix $\mathbf{A}$ is the sum of the squares of its eigenvalues.