Solutions to Homework Set #6

1. Eigenvalues. Suppose that $A$ has $\lambda_1, \lambda_2, \ldots, \lambda_n$ as its eigenvalues.

(a) Show that $\det(A) = \lambda_1 \cdot \lambda_2 \cdots \lambda_n$.

(b) Show that the eigenvalues of $A'$ are $\lambda_1, \lambda_2, \ldots, \lambda_n$, that is, $A$ and $A'$ have the same set of eigenvalues.

(c) Show that the eigenvalues of $A^k$ are $\lambda_1^k, \lambda_2^k, \ldots, \lambda_n^k$ for $k = 1, 2, \ldots$.

(d) Show that $A$ is invertible if and only if it does not have a zero eigenvalue.

(e) Suppose that $A$ is invertible. Show that the eigenvalues of $A^{-1}$ are $\lambda_1^{-1}, \lambda_2^{-1}, \ldots, \lambda_n^{-1}$.

(f) Show that $A$ and $T^{-1}AT$ have the same set of eigenvalues, that is, eigenvalues are invariant under a similarity transformation $A \mapsto T^{-1}AT$.

Solution:

(a) Consider the characteristic polynomial of $A$, namely, $\chi_A(\lambda) := \det(\lambda I - A)$. Clearly, the highest power of $\lambda$ in $\chi_A(\lambda)$, i.e., the $n^{th}$ power, occurs only in the term $\prod_{i=1}^{n}(\lambda - A_{ii})$. Therefore, the coefficient of $\lambda^n$ equals 1. The constant term is given by $\chi_A(0) = \det(-A) = (-1)^n \det(A)$. Therefore, we have

\[
\lambda_1 \cdot \lambda_2 \cdots \lambda_n = \text{product of all roots of } \{\chi_A(\lambda) = 0\}
\]

\[
= (-1)^n \frac{\text{constant term}}{\text{coefficient of } \lambda^n}
\]

\[
= \det(A).
\]

(b) We have

\[
\chi_A'(\lambda) = \det(\lambda I - A') = \det((\lambda I - A)') = \det(\lambda I - A) = \chi_A(\lambda),
\]

which shows that $A'$ and $A$ have identical characteristic polynomials and hence, identical eigenvalues.

(c) Consider the Jordan normal form of $A$, i.e., $A = JTJ^{-1}$, where $J$ is upper-triangular and has the eigenvalues $\lambda_1, \ldots, \lambda_n$ as its diagonal entries. Then, $A^k = TJ^kT^{-1}$ and the diagonal entries of the upper-triangular matrix $J^k$ are $\lambda_1^k, \ldots, \lambda_n^k$ in the same order. Then, the eigenvalues of $J^k$ (and hence, of $A^k$, see part (f)) are given by $\lambda_1^k, \ldots, \lambda_n^k$.

Alternative proof: For any $\lambda \in \mathbb{C}$, let $\mu_1, \ldots, \mu_k$ be the $k^{th}$ roots of $\lambda$. Then we have

\[
\prod_{j=1}^{k}(A - \mu_jI) = (-1)^k c_k I + \sum_{l=0}^{k-1} (-1)^l c_l A^{k-l},
\]

where $c_0 = 1$, and for $1 \leq l \leq k$, $c_l$ is the sum of all possible products of the $\mu_j$s, taken $l$ at a time. For example, $c_k$ is simply $\prod_{j=1}^{k} \mu_j$. 

1
Now, $\mu_1, \ldots, \mu_k$ are the roots of the polynomial $p(x) = x^k - \lambda = 0$, therefore $x^k - \lambda$ is identically equal to
\[
\prod_{j=1}^{n} (x - \mu_j) = \sum_{l=0}^{k} (-1)^l c_l x^{k-l}.
\]
Equating the coefficients of like powers of $x$, we therefore conclude that $c_l = 0$ for $l = 1, \ldots, k - 1$, $c_0 = 1$, and $c_k = (-1)^{k-1} \lambda$.
Using these relations, (2) becomes
\[
\lambda I - A^k = (-1)^{k-1} \prod_{j=1}^{k} (\mu_j I - A).
\]
Now, if the characteristic polynomial $\chi_A(\lambda) := \det(\lambda I - A)$ be given by
\[
\chi_A(\lambda) = \prod_{i=1}^{n} (\lambda - \lambda_i),
\]
we have
\[
\chi_{A^k}(\lambda) = \det(\lambda I - A^k)
\]
\[
= (-1)^{n(k-1)} \prod_{j=1}^{k} \det(\mu_j I - A)
\]
\[
= (-1)^{n(k-1)} \prod_{j=1}^{k} \prod_{i=1}^{n} (\mu_j - \lambda_i)
\]
\[
= (-1)^{n(k-1)} (-1)^{nk} \prod_{i=1}^{n} \prod_{j=1}^{k} (\lambda_i - \mu_j)
\]
\[
= (-1)^n \prod_{i=1}^{n} (\lambda_i^k - \lambda)
\]
\[
= \prod_{i=1}^{n} (\lambda - \lambda_i^k),
\]
which shows that the eigenvalues of $A^k$ are exactly $\lambda_1^k, \ldots, \lambda_n^k$.

(d) Let $\lambda_1, \ldots, \lambda_n$ be the eigenvalues of $A$. We have
\[
A \text{ is invertible } \iff \det(A) \neq 0 \iff \prod_{i=1}^{n} \lambda_i \neq 0 \iff \lambda_i \neq 0 \text{ for all } i.
\]
(e) If $A$ is invertible, we know that $\lambda_i \neq 0$ for all $i$. We have
\[
\chi_{A^{-1}}(\lambda) = \det(\lambda I - A^{-1}) \\
= \det((\lambda A - I)A^{-1}) \\
= \det(\lambda A - I)\det(A)^{-1} \\
= \lambda^n \det(A - \lambda^{-1}I)\det(A)^{-1} \\
= (-\lambda)^n \det(\lambda^{-1}I - A)\det(A)^{-1} \\
= (-\lambda)^n \chi_A(\lambda^{-1})\det(A)^{-1} \\
= (-1)^n \prod_{i=1}^{n} \left( \frac{\lambda}{\lambda_i} \right) \chi_A(\lambda^{-1}) \\
= (-1)^n \prod_{i=1}^{n} \frac{\lambda}{\lambda_i}(\lambda^{-1} - \lambda_i) \\
= \prod_{i=1}^{n} (\lambda - \lambda_i^{-1}) ,
\]
which shows that $\lambda_1^{-1}, \ldots, \lambda_n^{-1}$ are the eigenvalues of $A^{-1}$.

(f) We have
\[
\chi_{T^{-1}AT}(\lambda) = \det(\lambda I - T^{-1}AT) = \det(T^{-1}(\lambda I - A)T) = \det(T)^{-1} \det(\lambda I - A)\det(T) = \chi_A(\lambda),
\]
which shows that $T^{-1}AT$ has the same eigenvalues as $A$.

2. **Trace.** We define the **trace** of $A \in \mathbb{R}^{n \times n}$ as
\[
\text{tr}(A) = A_{11} + A_{22} + \cdots + A_{nn},
\]
that is, the sum of its diagonal entries.

(a) Suppose that $A$ has $\lambda_1, \lambda_2, \ldots, \lambda_n$ as its eigenvalues. Show that
\[
\text{tr}(A) = \lambda_1 + \lambda_2 + \cdots + \lambda_n.
\]

(b) Show that
\[
\text{tr}(A^k) = \lambda_1^k + \lambda_2^k + \cdots + \lambda_n^k
\]
for $k = 1, 2, \ldots$.

**Solution:**

(a) We have
\[
\lambda_1 + \cdots + \lambda_n = - (\text{coefficient of } \lambda^{n-1} \text{ in } \chi_A(\lambda)).
\]
Now, in $\chi_A(\lambda) = \det(\lambda I - A)$, the only term containing $\lambda^n$ and $\lambda^{n-1}$ is $\prod_{i=1}^{n} (\lambda - A_{ii})$. (This is immediately clear by considering the definition of a determinant in terms of permutations.) Therefore, the coefficient of $\lambda^{n-1}$ in $\chi_A(\lambda)$ is the same as the coefficient of $\lambda^{n-1}$ in $\prod_{i=1}^{n} (\lambda - A_{ii})$, which is given by $-\text{tr}(A)$. Therefore,
\[
\text{tr}(A) = \lambda_1 + \cdots + \lambda_n.
\]
(b) Using problems 2(a) and 1(c), the result is immediate.

   (a) Show that $e^{A+B} = e^A e^B$ if $A$ and $B$ commute, namely, $AB = BA$.
   (b) Show that $A$ and $e^{tA}$ commute for every $t \in \mathbb{R}$.
   (c) Suppose that $\lambda_1, \lambda_2, \ldots, \lambda_n$ are eigenvalues of $A$. Show that the eigenvalues of $e^A$ are $e^{\lambda_1}, e^{\lambda_2}, \ldots, e^{\lambda_n}$.
   (d) Show that $\det(e^A) = e^{\text{tr}(A)}$.
   (e) Show that $\frac{d}{dt} e^{tA} = Ae^{tA} = e^{tA}$. 
   (f) Show that $e^A = \lim_{k \to \infty} \left( I + \frac{A}{k} \right)^k$.

Solution:
(a) Consider
\[
e^{(A+B)} = \sum_{n=0}^{\infty} \frac{(A+B)^n}{n!} = \sum_{n=0}^{\infty} \sum_{k=0}^{n} \frac{n!}{k!(n-k)!} A^k B^{n-k}
\]
\[
= \sum_{n=0}^{\infty} \sum_{k=0}^{n} \frac{A^k B^{n-k}}{k!(n-k)!}
\]
\[
= \sum_{n,k,n \geq k} \frac{A^k B^{n-k}}{k!(n-k)!}
\]
\[
= I[I + \frac{B^1}{1!} + \frac{B^2}{2!} + \cdots] + \frac{A^1}{1!}[I + \frac{B^1}{1!} + \frac{B^2}{2!} + \cdots] + \cdots
\]
\[
= e^B[I + \frac{A}{1!} + \frac{A^2}{2!} + \cdots]
\]
\[
= e^B e^A,
\]
where (a) follows since $A$ and $B$ commute and (b) follows by grouping all terms according to the corresponding power of $A$.

(b) Note that $AA^i = A^i A = A^{(i+1)}$. We then have
\[
A e^{tA} = A[I + \frac{A}{1!} + \frac{A^2}{2!} + \cdots] = A + \frac{A^2}{1!} + \frac{A^3}{2!} + \cdots = [I + \frac{A}{1!} + \frac{A^2}{2!} + \cdots] A = e^{tA} A.
\]
Thus $A$ and $e^{tA}$ commute.
(c) Consider any eigenvector \( v_1 \) associated with \( \lambda_1 \). Note that \( A^k v_1 = A^{(k-1)} A v_1 = \lambda_1 A^{(k-1)} v_1 \). Continuing this way, we get \( A^k v_1 = \lambda_k v_1 \). We then have \( e^{A}v_1 = v_1 + A^1 v_1 + \frac{A^2 v_1}{2} + \cdots = [1 + \frac{\lambda_1}{1!} + \frac{\lambda_1^2}{2!} + \cdots] v_1 = e^{\lambda_1} v_1 \). Thus, \( e^{\lambda_1} \) is an eigenvalue for \( e^{A} \) with associated eigenvector \( v_1 \). Similarly, we can show that for all \( i = \{1, 2, \ldots, n\} \), \( e^{\lambda_i} \) is an eigenvalue of \( e^{A} \).

More generally, consider the Jordan canonical form of \( A \), i.e., \( A = T^{-1} J T \). Then, we have

\[
e^{A} = T^{-1} e^{J} T,
\]

and \( e^{J} \) consists of exponentials of the Jordan blocks arranged in the same order. Now, consider a Jordan block

\[
\tilde{J} = \begin{bmatrix}
\lambda & 1 & 0 & \cdots & 0 \\
\lambda & 1 & \cdots & 0 \\
\vdots & \ddots & \ddots & \ddots & \ddots \\
\lambda & \cdots & 1 & \lambda \\
0 & \cdots & 0 & \cdots & \lambda
\end{bmatrix}
\]

By elementary matrix multiplication, we see that for all non-negative integers \( m \), \( \tilde{J}^m \) is an upper triangular matrix with diagonals equal to \( \lambda^m \). Therefore, \( e^{J} \) is an upper triangular matrix whose diagonals are \( e^{\lambda_1}, \ldots, e^{\lambda_n} \). Therefore, since eigenvalues remain invariant under similarity transformations, and the eigenvalues of an upper-triangular matrix are simply its diagonal entries, the eigenvalues of \( e^{A} \) are \( e^{\lambda_1}, \ldots, e^{\lambda_n} \).

(d) Since determinant is the product of the eigenvalues, using the result in part (c), we have

\[
det(e^{A}) = \prod_{i=1}^{n} e^{\lambda_i} = e^{\sum_{i=1}^{n} \lambda_i} = e^{\text{tr}(A)},
\]

since \( \sum_{i=1}^{n} \lambda_i = \text{tr}(A) \).

(e) We have

\[
\frac{d}{dt} e^{tA} = \frac{d}{dt} \left( I + \sum_{k=1}^{\infty} \frac{A^k}{k!} t^k \right) = \sum_{k=1}^{\infty} \frac{A^k}{(k-1)!} t^{(k-1)} = A \left( I + \sum_{k=1}^{\infty} \frac{A^k}{k!} t^k \right) = A e^{tA} = e^{tA} A,
\]

where the final step follows from part (b).

(f) Consider the Jordan canonical form \( A = T^{-1} J T \). Then for all \( k \in \mathbb{N} \),

\[
\left( I + \frac{A}{k} \right)^k = T^{-1} \left( I + \frac{J}{k} \right)^k T.
\]

Now, consider an \( n \times n \) Jordan block \( \tilde{J} \) such that

\[
\tilde{J}_{ij} = \begin{cases} 
\lambda, & j = i \\
1, & j = i + 1 \\
0, & \text{otherwise}.
\end{cases}
\]
Note that \( kI + \tilde{J} \) is also a Jordan block with \( \lambda + k \) on the diagonal. Therefore, for \( k > n \),

\[
M^k := \left( I + \frac{J}{k} \right)^k = \frac{(kI + J)^k}{k^k}
\]

is an upper-triangular matrix whose \( (i,j)^{th} \) entry is given by (for \( j \geq i \))

\[
\left( \begin{array}{c}
k \\
\end{array} \right) \frac{(\lambda + k)^{k+i-j}}{k^k}.
\]

Taking limits, we have

\[
\lim_{k \to \infty} M^{(k)}_{ij} = \lim_{k \to \infty} \left( \begin{array}{c}
k \\
\end{array} \right) \frac{(\lambda + k)^{k+i-j}}{k^k} = \frac{1}{(j-i)!} \lim_{k \to \infty} \left( 1 + \frac{\lambda}{k} \right)^k \frac{k!}{(k + i - j)! (\lambda + k)^{j-i}}
\]

\[
= \frac{1}{(j-i)!} \lim_{k \to \infty} \left( 1 + \frac{\lambda}{k} \right)^k \prod_{m=1}^{j-i} \frac{k-1}{k+\lambda}
\]

\[
= \frac{e^\lambda}{(j-i)!}
\]

for \( j \geq i \). Therefore, \( \lim_{k \to \infty} (I + J/k)^k = \lim_{k \to \infty} M^k = e^J \), whence

\[
\lim_{k \to \infty} \left( I + \frac{A}{k} \right)^k = T^{-1} \lim_{k \to \infty} \left( I + \frac{J}{k} \right)^k T = T^{-1} e^J T = e^A.
\]

4. **Square root of a matrix.** Let \( A \in \mathbb{R}^{n \times n} \) be diagonalizable with eigenvalue decomposition \( T \Lambda T^{-1} \). We say that a matrix \( B \) is a square root of \( A \) if \( B^2 = A \). Show that \( B = T \Lambda^{1/2} T^{-1} \) is a square root of \( A \), where \( \Lambda^{1/2} \) is a diagonal matrix with entries \( \gamma_1, \gamma_2, \ldots, \gamma_n \) such that \( \gamma_i^2 = \lambda_i \).

**Solution:** \( B^2 = T \Lambda^{1/2} T^{-1} T \Lambda^{1/2} T^{-1} = T \Lambda^{1/2} \Lambda^{1/2} T^{-1} \). Since \( \Lambda^{1/2} \) is a diagonal matrix, \( \Lambda^{1/2} \Lambda^{1/2} \) will simply be \( \text{diag}(\gamma_1^2, \gamma_2^2, \ldots, \gamma_n^2) = \text{diag}(\lambda_1, \lambda_2, \ldots, \lambda_n) \). Hence,

\[
T \Lambda^{1/2} \Lambda^{1/2} T^{-1} = T \Lambda T^{-1} = A.
\]

Thus, \( B^2 = A \) and \( B \) is a square root of \( A \).

5. **Gershgorin circles.** Let \( v \) be an eigenvector of \( A \in \mathbb{C}^{n \times n} \) associated with eigenvalue \( \lambda \) such that \( \|v\|_\infty = |v_i| = 1 \).

(a) Show that \( (\lambda - A_{ii})v_i = \sum_{j \neq i} A_{ij}v_j \).

(b) Let the Gershgorin circles of \( A \) be defined as

\[
\mathcal{G}_i = \{ \xi \in \mathbb{C} : |\xi - A_{ii}| \leq \rho_i \}, \quad i = 1, 2, \ldots, n,
\]

where the radius of the \( i \)-th circle centered at \( A_{ii} \) is

\[
\rho_i = \sum_{j \neq i} |A_{ij}|.
\]

Show that all eigenvalues of \( A \) are contained in the union of the Gershgorin circles.
Solution:

(a) Since $v$ is an eigenvector associated with eigenvalue $\lambda$, we have $Av = \lambda v$. In particular, equating the $i$-th element of $Av$ and $\lambda v$, we have $\lambda v_i = \sum_{j=1}^{n} A_{ij} v_j = v_i A_{ii} + \sum_{j \neq i} A_{ij} v_j$, which yields $(\lambda - A_{ii}) v_i = \sum_{j \neq i} A_{ij} v_j$.

(b) From part (a), we have

$$|(\lambda - A_{ii})| v_i = \left| \sum_{j \neq i} A_{ij} v_j \right| \leq \sum_{j \neq i} |A_{ij}| v_j \leq \sum_{j \neq i} |A_{ij}|.$$ 

Thus, $|\lambda - A_{ii}| \leq \rho_i$, or equivalently, $\lambda \in G_i$. Similarly, the other eigenvalues lie in one of the Gershgorin circles.

6. Diagonally dominated matrices. We say that $A \in \mathbb{C}^{n \times n}$ is diagonally dominated if $A_{ii} > \sum_{j \neq i} |A_{ij}|, \quad i = 1, 2, \ldots, n.$

Show that a diagonally dominated matrix $A$ is nonsingular. (Hint: Use Gershgorin circles.)

Solution: Consider the Gershgorin circles $G_i$ from Problem #5. Since $0 \notin G_i$ for every $i$, no eigenvalue is 0. Hence, $A$ is nonsingular.

7. Nilpotent matrices. We say that a square matrix $A$ is nilpotent if $A^k = 0$ for some $k \geq 1$. We define the smallest $k$ for which $A^k = 0$ to be its (nilpotent) index. For example,

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

is nilpotent of index 3.

(a) Show that every nilpotent matrix $A \in \mathbb{F}^{n \times n}$ has no nonzero eigenvalue and thus that its characteristic function is $\chi(\lambda) = \det(\lambda I - A) = \lambda^n$.

(b) Show that the index of a nilpotent matrix $A \in \mathbb{F}^{n \times n}$ is always $\leq n$.

(c) Suppose that $A \in \mathbb{F}^{n \times n}$ is nilpotent of index $n$. Show that if $A^{n-1}x \neq 0$, then $x, Ax, A^2 x, \ldots, A^{n-1} x$ form a basis of $\mathbb{F}^n$.

(d) Continuing part (c), let

$$T = \begin{bmatrix} x & Ax & A^2 x & \cdots & A^{n-1} x \end{bmatrix} \in \mathbb{F}^{n \times n}.$$

Show that the similarity transformation of $A$ by $T$ is

$$T^{-1} AT = \begin{bmatrix} 0 & 1 & 0 & \cdots & 0 \\ 1 & 0 & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & \cdots & 0 \end{bmatrix}.$$
Solution:

(a) Let \( \lambda_1, \lambda_2, \ldots, \lambda_n \) be the eigenvalues of \( A \), a nilpotent matrix with index \( k \). By Problem 1(c), the eigenvalues of \( A^k \) are \( \lambda_1^k, \lambda_2^k, \ldots, \lambda_n^k \). But \( A^k = 0 \) and its eigenvalues \( \lambda_1^k, \lambda_2^k, \ldots, \lambda_n^k \) must be all zero, which in turn implies that \( \lambda_1, \lambda_2, \ldots, \lambda_n \) must be all zero. Therefore all \( n \) eigenvalues of \( A \) are zero, implying that \( \chi(\lambda) = \lambda^n \).

(b) By the Cayley–Hamilton theorem, \( \chi(A) = A^n = 0 \). Thus, the index of a nilpotent matrix is at most \( n \).

(c) To show that \( x, Ax, A^2x, \ldots, A^{n-1}x \) are independent, consider some \( \alpha_0, \alpha_1, \ldots, \alpha_{n-1} \) such that \( \alpha_0x + \alpha_1Ax + \alpha_2A^2x + \cdots + \alpha_{n-1}A^{n-1}x = 0 \). Multiplying both sides by \( A^{n-1} \), since \( A^k = 0 \) for \( k \geq n \), we have \( \alpha_0A^{n-1}x = 0 \), which implies that \( \alpha_0 = 0 \). Next, we multiply both sides by \( A^{n-2} \) and use the fact that \( \alpha_0 = 0 \) to show \( \alpha_1 = 0 \). Continuing this way, we can show that \( \alpha_1, \alpha_2, \ldots, \alpha_{n-1} \) are all zero. Thus, \( x, Ax, A^2x, \ldots, A^{n-1}x \) are \( n \) independent vectors that form a basis for \( \mathbb{F}^n \).

(d) Note that

\[
AT = A \begin{bmatrix} x & Ax & A^2x & \cdots & A^{n-1}x \end{bmatrix} \\
= \begin{bmatrix} Ax & A^2x & A^3x & \cdots & A^n x \end{bmatrix} \\
= \begin{bmatrix} Ax & A^2x & \cdots & 0 \end{bmatrix} \\
= T \begin{bmatrix} 0 & 1 & 0 & \cdots & \cdots & \cdots & 1 & 0 \end{bmatrix}.
\]

Thus,

\[
T^{-1}AT = \begin{bmatrix} 0 & 1 & 0 & \cdots & \cdots & \cdots & 1 & 0 \end{bmatrix}.
\]