Homework 5

ECE 269: Linear Algebra and Applications
Homework #5-Solution
Instructor: Behrouz Touri

1. Show that for any alternating linear form $f : (\mathbb{R}^n)^m \to \mathbb{R}$, we have the following:

   (a) for any $i \neq j$,

   $$f(v_1, \ldots, v_i, \ldots, v_j, \ldots, v_m) = -f(v_1, \ldots, v_j, \ldots, v_i, \ldots, v_m).$$

   (b) if $m = n$, and $f(e_1, \ldots, e_n) \neq 0$ (where $e_1, \ldots, e_n$ are the standard basis elements) then $f(v_1, \ldots, v_n) = 0$ only if $v_1, \ldots, v_n$ are linearly dependent (of course, this is if and only if statement, as we have shown the if part in the class).

   Solution:

   (a) Notice that

   $$\begin{aligned} 0 &= f(v_1, \ldots, v_i + v_j, \ldots, v_i + v_j, \ldots, v_m) \\ &= f(v_1, \ldots, v_i, \ldots, v_i, \ldots, v_m) + f(v_1, \ldots, v_i, \ldots, v_j, \ldots, v_m) \\ &\qquad + f(v_1, \ldots, v_j, \ldots, v_i, \ldots, v_m) + f(v_1, \ldots, v_j, \ldots, v_j, \ldots, v_m) \\ &= 0 + f(v_1, \ldots, v_i, \ldots, v_j, \ldots, v_m) + f(v_1, \ldots, v_j, \ldots, v_i, \ldots, v_m) + 0. \end{aligned}$$

   Therefore, we have

   $$f(v_1, \ldots, v_i, \ldots, v_j, \ldots, v_m) = -f(v_1, \ldots, v_j, \ldots, v_i, \ldots, v_m).$$

   (b) As mentioned in the discussion session, f should a multiple of $f(e_1, \ldots, e_n)$ times the determinant. Therefore, without loss of generality, we assume $f(e_1, \ldots, e_n) = 1$, which implies $f = \det$. Let $V = \begin{bmatrix} v_1, \ldots, v_n \end{bmatrix}$. If $v_1, \ldots, v_n$ are linearly independent, V is invertible. We also know that determinant function satisfies

   $$f(I) = f(V)f(V^{-1}) = 1.$$

   This leads to a contradiction to the fact that $f(V) \neq 0$. Now we provide a sketch of the proof for $\det(AB) = \det(A)\det(B)$. In fact, we have

   $$\begin{aligned} \det(AB) &= \det(B_{11}a_1 + \ldots + B_{n1}a_n, \ldots, B_{1n}a_1 + \ldots + B_{nn}a_n) \\ &= \sum_P \prod_{(i,j) \in P} (-1)^{\#P} B_{ij} \det(a_1, \ldots, a_n) \\ &= \det(B)\det(A), \end{aligned}$$

   where a_i, $i \in [n]$ is the i-th column of A and P is taken over all $n \times n$ permutation matrices.

2. *Practical Determinant.* In practice, one never goes over the extensive formula discussed in lecture for computing determinant, but rather the transformations involving the matrices. One of them being LU decomposition.

(a) For a $n \times n$ lower-triangular matrix Q, show that $\det(Q) = Q_{11} \cdots Q_{nn}$. (Note that since $\det(A) = \det(A')$, same result holds for upper triangular matrices)

(b) Using an LU decomposition of A (of the form $LPA = U$ or $PA = LU$), find $\det(A)$ for the following matrix.

$$A = \begin{bmatrix} 1 & 2 & 1 & 0 \\ 2 & 5 & 3 & 1 \\ 1 & 3 & 3 & 2 \\ 0 & 1 & 2 & 4 \end{bmatrix}$$

Solution:

(a) We know

$$\det(Q) = \sum_P \prod_{(i,j)\in P} (-1)^{\#P} Q_{ij},$$

where P is taken over all $n \times n$ permutation matrices. If for some P, $prod_{(i,j)\in P}(-1)^{\#P}Q_{ij}$ is nonzero, then we must have $(i \leq j)$ for any $(i,j) \in P$ since Q is LT. Also notice

$$\sum_{(i,j)\in P} i = 1 + 2 + \ldots + n = \sum_{(i,j)\in P} j.$$

Hence all inequalities $i \leq j$ become equalities. So $P = I$ and

$$\det(Q) = \prod_{i=1}^{n} Q_{ii}.$$

(b) Let u_i, $i \in [4]$ be the i-th row of the matrix U. We first set $u_1 = \begin{bmatrix} 1 & 2 & 1 & 0 \end{bmatrix}$ and $L_{11} = 1$. Next we set $u_2 = \begin{bmatrix} 0 & 1 & 1 & 1 \end{bmatrix}$ and $L_{21} = 2$, $L_{22} = 1$. Then we set $u_3 = \begin{bmatrix} 0 & 0 & 1 & 1 \end{bmatrix}$ and $L_{31} = L_{32} = L_{33} = 1$. Finally, we set $u_4 = \begin{bmatrix} 0 & 0 & 0 & 2 \end{bmatrix}$ and $L_{41} = 0$, $L_{42} = L_{43} = L_{44} = 1$. That is,

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

Therefore, $\det(A) = 1 \cdot 2 = 2$.

3. *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 $k = 1, 2, \ldots$ and if A is invertible, the result holds for all $k \in \mathbb{Z}$.

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

(e) For an invertible matrix $T \in \mathbb{R}^{n\times n}$, show that A and $T^{-1}AT$ have the same set of eigenvalues, that is, eigenvalues are invariant under a similarity transformation $A \mapsto T^{-1}AT$.

(f) Let us define the set of eigenvectors corresponding to eigenvalue λ, to be $v_\lambda(A) = \{v \in \mathbb{R}^n \mid Av = \lambda v\}$. Show that $v_\lambda(A)$ is a subspace of $\mathbb{R}^n$.

(g) The *trace* of $A \in \mathbb{R}^{n\times n}$ is defined by sum of its diagonal elements, i.e.,

$$\mathrm{trace}(A) = A_{11} + A_{22} + \cdots + A_{nn}.$$

Show that

$$\mathrm{trace}(A) = \lambda_1 + \lambda_2 + \cdots + \lambda_n.$$

Solution:

(a) Consider the characteristic polynomial of A, namely, $\chi_A(\lambda) := \det(\lambda I - A)$. Clearly, the highest power of λ 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 λ^n equals 1. The constant term is given by $\chi_A(0) = \det(-A) = (-1)^n \det(A)$. Therefore, we have

$$\begin{aligned}\lambda_1 \cdot \lambda_2 \cdots \lambda_n &= \text{product of all roots of } \{\chi_A(\lambda) = 0\}\\ &= (-1)^n \frac{\text{constant term}}{\text{coefficient of } \lambda^n}\\ &= \det(A).\end{aligned}$$

(b) We have

$$\chi_{A'}(\lambda) = \det(\lambda I - A') = \det\left((\lambda I - A)'\right) = \det(\lambda I - A) = \chi_A(\lambda),$$

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

(c) Consider the Jordan normal form of A, i.e., $A = TJT^{-1}$, 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 λ. Then we have

$$\prod_{j=1}^{k}(A - \mu_j I) = (-1)^k c_k I + \sum_{l=0}^{k-1}(-1)^l c_l A^{k-l}, \tag{1}$$

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

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, (1) becomes

$$\lambda I - A^k = (-1)^{k-1} \prod_{j=1}^{k}(\mu_j I - A). \tag{2}$$

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

$$\begin{aligned}
\chi_{A^k}(\lambda) &= \det(\lambda I - A^k) \\
&\overset{(2)}{=} (-1)^{n(k-1)} \prod_{j=1}^{k} \det(\mu_j I - A) \\
&= (-1)^{n(k-1)} \prod_{j=1}^{k} \prod_{i=1}^{n} (\mu_j - \lambda_i) \\
&= (-1)^{n(k-1)} (-1)^{nk} \prod_{i=1}^{n} \prod_{j=1}^{k} (\lambda_i - \mu_j) \\
&= (-1)^{n} \prod_{i=1}^{n} (\lambda_i^k - \lambda) \\
&= \prod_{i=1}^{n} (\lambda - \lambda_i^k),
\end{aligned}$$

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) We have

$$\begin{aligned}
\chi_{T^{-1}AT}(\lambda) &= \det(\lambda I - T^{-1}AT) \\
&= \det\left(T^{-1}(\lambda I - A)T\right) \\
&= \det(T^{-1}) \det(\lambda I - A) \det(T) \\
&= \chi_A(\lambda),
\end{aligned}$$

which shows that $T^{-1}AT$ has the same eigenvalues as A.

(f) Let $x, y \in v_\lambda(A)$ and $\alpha \in \mathbb{R}$. Then

$$A(x + \alpha y) = Ax + A\alpha y = \lambda x + \alpha \lambda y = \lambda(x + \alpha y).$$

Therefore, $v_\lambda(A)$ is a subspace of $\mathbb{R}^n$.

(g) 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 λ^n and λ^{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 λ^{n-1} in $\chi_A(\lambda)$ is the same as the coefficient of λ^{n-1} in $\prod_{i=1}^{n}(\lambda - A_{ii})$, which is given by $-\text{trace}(A)$. Therefore,

$$\text{trace}(A) = \lambda_1 + \cdots + \lambda_n.$$

4. *Gershgorin circles.* Let v be an eigenvector of $A \in \mathbb{C}^{n\times n}$ associated with eigenvalue λ 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
   $$\mathcal{G}_i = \{\xi \in \mathbb{C} : |A_{ii} - \xi| \le \rho_i\}, \quad i = 1, 2, \ldots, n,$$
   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.

   (c) 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.

   Solution:

   (a) Since v is an eigenvector associated with eigenvalue λ, we have $Av = \lambda v$. In particular, equating the i-th row of Av and λv, 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
   $$\begin{aligned} |(\lambda - A_{ii})||v_i| &= \left|\sum_{j\neq i} A_{ij}v_j\right| \\ &\le \sum_{j\neq i} |A_{ij}||v_j| \\ &\le \sum_{j\neq i} |A_{ij}|. \end{aligned}$$
   Thus, $|\lambda - A_{ii}| \le \rho_i$, or equivalently, $\lambda \in \mathcal{G}_i$. Similarly, the other eigenvalues lie in one of the Gergshgorin circles.

   (c) If 0 is an eigenvalue of A, $0 \in \mathcal{G}_{i^*}$ for some $i_0 \in [n]$. So, we must have
   $$|A_{i^*i^*} - 0| \le \rho_{i^*} = \sum_{j\neq i^*} A_{i^*j},$$
   which contradicts to the diagonally dominated condition. So, 0 is not an eigenvalue of A and A is nonsingular.

5. *Nilpotent matrices.* We say that a square matrix A is *nilpotent* if $A^k = 0$ for some $k \ge 1$. 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 $x, Ax, A^2x, \ldots, A^{n-1}x$ form a basis of $\mathbb{F}^n$.

(d) Continuing part (c), let

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

Show that the similarity transformation of A by T is

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

Solution:

(a) Let $\lambda_1, \lambda_2, \ldots, \lambda_n$ be the eigenvalues of A, a nilpotent matrix with index k. By Problem 3(c), the eigenvalues of A^k are $\lambda_1^k, \lambda_2^k, \ldots, \lambda_n^k$. But $A^k = 0$ and its eigenvalues $\lambda_1^k, \lambda_2^k, \cdots, \lambda_n^k$ 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 $x, Ax, A^2x, \cdots, A^{n-1}x$ are independent, consider some $a_0, a_1, \cdots, a_{n-1}$ such that $a_0x + a_1Ax + a_2A^2x + \cdots a_{n-1}A^{n-1}x = 0$. Multiplying both sides by A^{n-1}, since $A^k = 0$ for $k \geq n$, we have $a_0A^{n-1}x = 0$, which implies that $a_0 = 0$. Next, we multiply both sides by A^{n-2} and use the fact that $a_0 = 0$ to show $a_1 = 0$. Continuing this way, we can show that $a_1, a_2, \ldots, a_{n-1}$ are all zero. Thus, $x, Ax, A^2x, \cdots, A^{n-1}x$ are n independent vectors that form a basis for $\mathbb{F}^n$.

(d) Note that

$$\begin{aligned} AT &= A\begin{bmatrix} x & Ax & A^2x & \cdots & A^{n-1}x \end{bmatrix} \\ &= \begin{bmatrix} Ax & A^2x & A^3x & \cdots & A^nx \end{bmatrix} \\ &= \begin{bmatrix} Ax & A^2x & \cdots & 0 \end{bmatrix} \\ &= T\begin{bmatrix} 0 & & & \\ 1 & 0 & & \\ & \ddots & \ddots & \\ & & 1 & 0 \end{bmatrix}. \end{aligned}$$

Thus,

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