1. (7 pts)[Apostol IV.8.12, 13, 14]
(12.) Let $A$ be an $n \times n$ matrix with characteristic polynomial $f(\lambda)$. Prove (by induction) that the coefficient of $\lambda^{n-1}$ in $f(\lambda)$ is $-\text{tr } A$.

**Proof.** Let $A$ be $(a_{ij})_{n \times n}$, we will prove the statement by induction on $n$.

When $n = 1$, $f(\lambda) = \lambda - a_{11}$ and $\text{tr}(A) = a_{11}$. The statement is true.

Now we assume that when $k = n - 1$ ($n \geq 2$) the statement is true, then we consider the case when $k = n$. Notice that

$$f(\lambda) = \det(\lambda I - A).$$

We denote $M = \lambda I - A$, then the bottom row of $M$ is $(-a_{n,1}, \cdots, -a_{n,n-1}, \lambda - a_{n,n})$.

We use the Theorem 3.9 in the textbook, we have that

$$\det M = (-1)^n a_{n,1} \det(M_{n,1}) - \cdots + a_{n,n-1} \det(M_{n,n-1}) - (a_{n,n} - \lambda) \det(M_{n,n}).$$

Since the only entries of $M$ that contain $\lambda$ are in the diagonal, so $\det(M_{n,i})$ ($1 \leq i \leq n-1$) is at most a degree $(n-2)$ polynomial on $\lambda$ as $M_{n,i}$ won’t contain $\lambda - a_{n,n}$. Then $\det(M_{n,i})$ ($1 \leq i \leq n-1$) will contribute nothing in the coefficient of $\lambda^{n-1}$.

Observe that $M_{n,n} = \lambda - (a_{ij})_{(n-1)\times(n-1)}$, so we can use our assumption on the characteristic polynomial of $(a_{ij})_{(n-1)\times(n-1)}$. Then we get that the coefficient of $\lambda^{n-2}$ in $\det M_{n,n}$ is $-\sum_{i=1}^{n-1} a_{ii}$. We also know that characteristic polynomial is monic, so $\det M_{n,n} = \lambda^{n-1} - \sum_{i=1}^{n-1} a_{ii} \lambda^{n-2} +$ some lower terms. We plug in the equation of $\det M$ above, we get that the coefficient of $\lambda^{n-1}$ in $\det M$ is

$$-\sum_{i=1}^{n-1} a_{ii} - a_{nn} = -\text{tr } A,$$

which completes the induction. \(\square\)

(13.) Let $A$ be $B$ be $n \times n$ matrices with $\det A = \det B$ and $\text{tr } A = \text{tr } B$. Prove that $A$ and $B$ have the same characteristic polynomial if $n = 2$ but that this need not be the case if $n > 2$.

**Proof.** Let $X$ be any matrix, we denote $f_X(\lambda)$ the characteristic polynomial of $X$. Then we know that $f_X$ is monic, the coefficient of $\lambda^{n-1}$ is $-\text{tr } X$ by the previous problem and the constant term in $f_X$ is $(-1)^n \det(X)$.

When $n = 2$, then $f_A = \lambda^2 - \text{tr}(A)\lambda + \det(A)$ and $f_B = \lambda^2 - \text{tr}(B)\lambda + \det(B)$.

Since $\det A = \det B$ and $\text{tr}(A) = \text{tr}(B)$, so $f_A(\lambda) = f_B(\lambda)$. In other words, $A$ and $B$ have the same characteristic polynomials.

When $n > 2$, Let $A = \text{diag}(2,0,\cdots,0)$ be an $n \times n$ matrix and $B = \text{diag}(1,1,0,\cdots,0)$ be an $n \times n$ matrix. Since $n > 2$, both matrices have determinants 0 and traces 2. But $f_A(\lambda) = (\lambda - 2)\lambda^{n-1}$ is different from $f_B(\lambda) = (\lambda - 1)^2 \lambda^{n-2}$. Hence the statement is false when $n > 2$. \(\square\)

(14.) Prove each of the following statements about the trace.

---

*Date: February 29, 2016.*
(a) \( \text{tr} (A + B) = \text{tr} A + \text{tr} B \).
(b) \( \text{tr} (cA) = c \text{tr} A \).
(c) \( \text{tr} (AB) = \text{tr} (BA) \).
(d) \( \text{tr} A^t = \text{tr} A \).

\textbf{Proof.} Denote \( A = (a_{ij})_{n \times n} \) and \( B = (b_{ij})_{n \times n} \).

(a) By definition, \( \text{tr}(A) = \sum a_{ii} \), \( \text{tr}(B) = \sum b_{ii} \) and \( \text{tr}(A + B) = \sum (a_{ii} + b_{ii}) \), so we have
\[
\text{tr}(A + B) = \text{tr}(A) + \text{tr}(B).
\]

(b) Matrix \( cA = (ca_{ij})_{n \times n} \), then we have
\[
\text{tr}(cA) = \sum (c \cdot a_{ii}) = c \sum a_{ii} = c \text{tr}(A).
\]

(c) Denote \( AB = (c_{ij})_{n \times n} \) and \( BA = (d_{ij})_{n \times n} \). Then we have
\[
c_{ij} = \sum_{k=1}^{n} a_{ik} b_{kj},
d_{ij} = \sum_{k=1}^{n} b_{ik} a_{kj}.
\]
So we can compute the traces:
\[
\text{tr}(AB) = \sum_{i,j=1}^{n} c_{ij} = \sum_{i=1}^{n} \sum_{k=1}^{n} a_{ik} b_{ki},
\]
\[
\text{tr}(BA) = \sum_{i,j=1}^{n} d_{ij} = \sum_{i=1}^{n} \sum_{k=1}^{n} b_{ik} a_{ki}.
\]
We switch \( i \) and \( k \) in the equation of \( \text{tr}(AB) \), then get
\[
\text{tr}(AB) = \sum_{i,k=1}^{n} a_{ki} b_{ik} = \sum_{k=1}^{n} \sum_{i=1}^{n} b_{ik} a_{ki}.
\]
Next, we change the order of the summation in the above equation and get
\[
\text{tr}(AB) = \sum_{i=1}^{n} \sum_{k=1}^{n} b_{ik} a_{ki} = \text{tr}(BA).
\]

\textbf{Remark.} Actually, \( AB \) and \( BA \) have the same characteristic polynomials. When \( A \) is invertible, \( AB \) and \( BA \) are similar matrices so they have the same characteristic polynomials. When \( A \) is not invertible, the characteristic polynomials are still the same by the continuity.

(d) Notice that the \((i,j)\)-th entry of \( A^t \) is \( a_{ji} \), so
\[
\text{tr}(A^t) = \sum a_{ji} = \text{tr}(A).
\]

\( \square \)

2. (6 pts)[Apostol I.13.11] In the linear space of all real polynomials, define
\[
(f, g) = \int_{0}^{\infty} e^{-t} f(t) g(t) \, dt.
\]
(a) Prove that this improper integral converges absolutely for all polynomials \( f \) and \( g \).
(b) If \( x_n(t) = t^n \) for \( n = 0, 1, 2, \ldots \), prove that \( \langle x_n, x_m \rangle = (m + n)! \).

(c) Compute \( \langle f, g \rangle \) when \( f(t) = (t + 1)^2 \) and \( g(t) = t^2 + 1 \).

(d) Find all linear polynomials \( g(t) = a + bt \) orthogonal to \( f(t) = 1 + t \).

Also, compute an orthonormal basis for the subspace consisting of polynomials of degree at most 3.

**Proof.** (a) Since \( f(t)g(t) \) is a polynomial, we can find \( N \) such that if \( t \geq N \) we have \( |f(t)g(t)| \leq e^\frac{t}{2} \). Then

\[
\int_0^\infty |e^{-t}f(t)g(t)|dt \leq \int_0^N |e^{-t}f(t)g(t)|dt + \int_N^\infty |e^{-t}e^\frac{t}{2}|dt < \infty.
\]

(b) We argue by induction. We have \( \langle x_0, x_0 \rangle = \int_0^\infty e^{-t}dt = 1 = 0! \). Assume \( \langle x_m, x_n \rangle = (m + n)! \). Then integrating by parts we have

\[
\langle x_m, x_{n+1} \rangle = \int_0^\infty e^{-t}t^{m+n+1}dt
= \int_0^\infty \left( \frac{d}{dt} (-e^{-t}) \right) t^{m+n+1}dt
= \left[ -e^{-t}t^{n+m+1} \right]_0^\infty - \int_0^\infty -e^{-t} \left( \frac{d}{dt} t^{n+m+1} \right) dt
= (n + m + 1) \int_0^\infty e^{-t}t^{n+m}dt
= (n + m + 1) \langle x_n, x_m \rangle
= (n + m + 1)(n + m)!
= (n + m + 1)!.
\]

(c) By linearity and (b) we have

\[
\langle (t + 1)^2, t^2 + 1 \rangle = \langle t^2 + 2t + 1, t^2 + 1 \rangle
= \langle t^2, t^2 \rangle + 2 \langle t, t^2 \rangle + 2 \langle t, 1 \rangle + \langle 1, t^2 \rangle + \langle 1, 1 \rangle
= 4! + 2! + 2 \cdot 3! + 2 \cdot 1! + 2! + 0!
= 43.
\]

(d) Consider the set \( \{1 + t, 1\} \) which spans the space of linear polynomials. We apply Gram-Schmidt to get an orthogonal basis \( \{v_1, v_2\} \) for the same space, by
letting \( v_1 = 1 + t \) and
\[
v_2 = 1 - \frac{(1, 1 + t)}{(1 + t, 1 + t)}(1 + t)
\]
\[
= 1 - \frac{(1, 1) + (1, t)}{(1, 1) + 2(1, t) + \langle t, t \rangle(1 + t)}(1 + t)
\]
\[
= 1 - \frac{0! + 1!}{0! + 2! - 1! + 2!}(1 + t)
\]
\[
= 1 - \frac{2}{5}(1 + t)
\]
\[
= \frac{1}{5}(3 - 2t).
\]

Since the space of linear polynomials is two dimensional, the orthogonal complement of \( 1 + t \) is one-dimensional. Therefore every linear polynomial orthogonal to \( 1 + t \) is of the form \( c(3 - 2t) \) for some \( c \in \mathbb{R} \).

An orthonormal basis for the subspace consisting of polynomials of degree at most 3 is
\[
\left\{ 1, t - 1, \frac{1}{2} t^2 - 2t + 3, \frac{1}{6} t^3 - \frac{3}{2} t^2 + 2t - 1 \right\}.
\]

3. (5 pts) Let \( V \) be a vector space with an inner product \((,\). Let \( \{v_1, v_2, \ldots, v_n\} \) be a generating set of \( V \). Prove
(a) if \( (x, v) = 0 \) for all \( v \in V \), then \( x = 0 \).
(b) if \( (x, v_i) = 0 \) for all \( i = 1, \ldots, n \), then \( x = 0 \).
(c) if \( (x, v_i) = (y, v_i) \) for all \( i = 1, \ldots, n \), then \( x = y \).

Proof. (i) If \( x \neq 0 \), then by the axiom of the inner product, \( (x, x) > 0 \).
(ii) Let \( v \in V \). Then \( v = \sum_{i=1}^{n} a_i v_i \) for some \( a_i \in \mathbb{R} \). Then
\[
(x, v) = (x, \sum_{i=1}^{n} a_i v_i) = \sum_{i=1}^{n} (x, a_i v_i) = \sum_{i=1}^{n} a_i (x, v_i) = 0
\]
By part (i), \( x = 0 \).
(iii) The assumption implies that \( (x - y, v_i) = 0 \) for all \( i \). By part (ii), \( x - y = 0 \).
\( x = y \). □

4. (7 pts) Let \( V = \mathbb{R}^4 \) and \( U = L(v_1, v_2) \) be the subspace of \( V \) generated by the vectors \( v_1 = (1, 3, 1, 1) \) and \( v_2 = (3, 2, 2, 1) \).
(a) Find an orthogonal basis of \( U \).

Solution. Since the generating set for \( U \) is linearly independent, we can use the Gram–Schmidt process to get an orthogonal basis. We set
\[
w_1 = (1, 3, 1, 1).
\]
Then we produce an element in \( U \) orthogonal to \( w_1 \) by taking
\[
w_2 = (3, 2, 2, 1) - \text{proj}_{w_1}(3, 2, 2, 1)
\]
\[
= (3, 2, 2, 1) - (1, 3, 1, 1)
\]
\[
= (2, -1, 1, 0),
\]
where we used the calculation that 
\[
\text{proj}_{w_1}(3, 2, 2, 1) = \frac{\langle w_1, (3, 2, 2, 1) \rangle}{\langle w_1, w_1 \rangle} w_1 \\
= \frac{3 + 6 + 2 + 1}{12} w_1 \\
= w_1.
\]
Thus, \(\{w_1, w_2\}\) is an orthogonal basis for \(U\). \(\Box\)

(b) Find the orthogonal projection of \(v = (1, 1, 1)^T\) onto \(U\).

\textit{Solution.} Since \(\{w_1, w_2\}\) is an orthogonal basis for \(U\), we have 
\[
\text{proj}_U(v) = \frac{\langle v, w_1 \rangle}{\langle w_1, w_1 \rangle} w_1 + \frac{\langle v, w_2 \rangle}{\langle w_2, w_2 \rangle} w_2.
\]
We calculate 
\[
\text{proj}_U(1, 1, 1, 1) = \frac{\langle (1, 1, 1, 1), (1, 3, 1, 1) \rangle}{\langle w_1, w_1 \rangle} w_1 + \frac{\langle (1, 1, 1, 1), (2, -1, 1, 0) \rangle}{\langle w_2, w_2 \rangle} w_2 \\
= \frac{1 + 3 + 1 + 1}{12} w_1 + \frac{2 - 1 + 1}{6} w_2 \\
= \frac{1}{2} (1, 3, 1, 1) + \frac{1}{3} (2, -1, 1, 0) \\
= \frac{3}{6} (1, 3, 1, 1) + \frac{2}{6} (2, -1, 1, 0) \\
= (7/6, 7/6, 5/6, 3/6).
\]
\(\Box\)

(c) Find the distance of \(v = (1, 1, 0, 0)^T\) from \(U\).

\textit{Solution.} This is just finding \(\|v - \text{proj}_U(v)\|\). We compute 
\[
\text{proj}_U(1, 1, 0, 0) = \frac{\langle (1, 1, 0, 0), (1, 3, 1, 1) \rangle}{\langle w_1, w_1 \rangle} w_1 + \frac{\langle (1, 1, 0, 0), (2, -1, 1, 0) \rangle}{\langle w_2, w_2 \rangle} w_2 \\
= \frac{2}{6} (1, 3, 1, 1) + \frac{1}{6} (2, -1, 1, 0) \\
= (4/6, 5/6, 3/6, 2/6).
\]
Therefore, 
\[
\|\text{proj}_U(1, 1, 0, 0) - (1, 1, 0, 0)\| = \| (4/6, 5/6, 3/6, 2/6) - (1, 1, 0, 0) \|
= \|(2/6, -1/6, 3/6, 2/6)\|
= (4/36 + 1/36 + 9/36 + 4/36)^{1/2}
= (18/36)^{1/2}
= 1/\sqrt{2}.
\] \(\Box\)