# MATH 217 - W24 - LINEAR ALGEBRA
# HOMEWORK 10, SOLUTIONS

**Part A (10 points)**

Solve the following problems from the book:
6.1: 20, 54
6.2: 42, 50
6.3: 14
7.1: 12, 18, 42

**Solution.**

**6.1.20** We put A into reduced row-echelon form:

$$\begin{bmatrix} 1 & k & 1 \\ 1 & k+1 & k+2 \\ 1 & k+2 & 2k+4 \end{bmatrix}$$

Det $= (k+1)(2k+4) + k(k+2) + (k+2) - (k+1) - k(2k+4) - (k+2)^2 = 1$
Thus the matrix is always invertible.

**6.1.54** The pattern

$$\begin{bmatrix} * & 1000 & * & * & * \\ * & * & * & 1000 & * \\ 1000 & * & * & * & * \\ * & * & * & * & 1000 \\ * & * & 1000 & * & * \end{bmatrix},$$

has positive sign, and thus contributes the term $1000^5 = 10^{15}$ to $\det(A)$. Every other pattern $P$ will contribute a term of the form $sgn(P)prod(P)$ where at most three of the five factors in $prod(P)$ are $= 1000$. Indeed, if four of the five factors are $= 1000$, then the fifth factor must also $= 1000$. All other factors in $prod(P)$ have absolute value $< 10$. Thus $|sgn(P)prod(P)| < 1000^3 \cdot 10 = 10^{10}$. Because there are $5! = 120$ terms in $set(A)$, this means that the sum of all these other terms has absolute value strictly less than $119 \cdot 10^{10} < 10^{13}$ . Thus $\det(A) = 10^{15} + r$, where $|r| < 10^{13}$, hence $\det(A) > 0$.

**6.2.42** We have $A^TA = R^TQ^TQR$. Because the columns $u_1, \cdots, u_m$ of $Q$ form an orthonormal basis, we have $Q^TQ = I_m$. So, $A^TA = R^TR$. Because $R$ is $m \times m$ and upper-triangular with diagonal entries $r_{ii}$, we have $\det(R) = r_{11}r_{22}\cdots r_{mm}$. Because $\det(R^T) = \det(R)$, this implies that $\det(A^TA) = r_{11}^2 r_{22}^2 \cdots r_{mm}^2$.

**6.2.50** Let $R_i$ be the $i$th row of the matrix. By replacing $R_n$ with $R_n - R_{n-1}$, then $R_{n-1}$ with $R_{n-1} - R_{n-2}$, $\cdots$, $R_2$ with $R_2 - R_1$ we get an upper triangular matrix with 1 in all the diagonal entries.

$$\begin{bmatrix} 1 & 1 & 1 & \cdots & 1 \\ 0 & 1 & 1 & \cdots & 1 \\ 0 & 0 & 1 & \cdots & 1 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \end{bmatrix}$$

These operations do not change the determinant. So, $\det(M_n) = 1$.

**6.3.14** We have $A = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 4 & 10 \\ 1 & 10 & 30 \end{bmatrix}$.

$\det(A^T A) = 6$. Thus, the 3-volume of the 3-parallelepiped is equal to $\sqrt{det(A^T A)} = \sqrt{6}$.

**7.1.12** Recall that for $\lambda$ an eigenvalue and $v$ an eigenvector of the $n \times n$ matrix $A$, we have $Av = \lambda v$. Here, let $v = [v_1, v_2]^T$. Solving $Av = 2v$ we obtain $v_1 = t$ and $v_2 = -\frac{3}{2}t$ with $t \neq 0$, and solving $Av = 4v$ we obtain $v_1 = 0$ and $v_2 = t$ with $t \neq 0$. We know that the diagonalization of $A$ is the diagonal matrix with the eigenvalues on the diagonal. In general, we obtain the diagonal matrix $B$ from $S^{-1}AS = B$ where $S$ is the matrix with the corresponding eigenvectors in the columns. Thus, $S^{-1}AS$ is

$$\begin{bmatrix} 2 & 0 \\ 0 & 4 \end{bmatrix}$$

**7.1.18** Any nonzero vector in the plane $V$ is unchanged, hence is an eigenvector with the eigenvalue 1. Since any nonzero vector in $V^\perp$ is flipped about the origin, it is an eigenvector with eigenvalue $-1$. Pick any two noncollinear vectors from $V$ and a nonzero vector from $V^\perp$ to form an eigenbasis. This transformation is diagonalizable.

**7.1.42** $A$ has to be of the form $\begin{bmatrix} a & b & 0 \\ 0 & e & 0 \\ 0 & h & i \end{bmatrix}$, so a basis is

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

(dimension 5, linearly independent, all satisfy the condition)

## Part B (25 points)

**Problem 1.** If $V$ and $W$ are vector spaces, a function $F : V \times W \to \mathbb{R}$ is said to be *bilinear* if all of the following hold:

- for all $\vec{x}, \vec{y} \in V$ and $\vec{z} \in W$, $F(\vec{x} + \vec{y}, \vec{z}) = F(\vec{x}, \vec{z}) + F(\vec{y}, \vec{z})$;
- for all $\vec{x} \in V$ and $\vec{y}, \vec{z} \in W$, $F(\vec{x}, \vec{y} + \vec{z}) = F(\vec{x}, \vec{y}) + F(\vec{x}, \vec{z})$;
- for all $\vec{x} \in V$ and $\vec{y} \in W$ and for all $a \in \mathbb{R}$, $F(a\vec{x}, \vec{y}) = aF(\vec{x}, \vec{y})$ and $F(\vec{x}, a\vec{y}) = aF(\vec{x}, \vec{y})$.

Furthermore, if $F : V \times V \to \mathbb{R}$ is a bilinear function, we say that $F$ is *alternating* if $F(\vec{v}, \vec{v}) = 0$ for all $\vec{v} \in V$. Throughout this problem, let $F : \mathbb{R}^2 \times \mathbb{R}^2 \longrightarrow \mathbb{R}$ be a bilinear function.

(a) Prove that $F$ is alternating if and only if $F(\vec{u},\vec{v}) = -F(\vec{v},\vec{u})$ for all $\vec{u},\vec{v} \in \mathbb{R}^2$.
(b) Prove that if $F$ is alternating and $F(\vec{e}_1,\vec{e}_2) = 1$, then

$$F(\vec{u},\vec{v}) = \det[\vec{u}\ \vec{v}] \qquad \text{for all } \vec{u},\vec{v} \in \mathbb{R}^2.$$

**Solution. a)** We must show both implications.
($\Leftarrow$) Let $\vec{u}=\vec{v}$, $F(\vec{v},\vec{v}) = -F(\vec{v},\vec{v})$ implies that $F(\vec{v},\vec{v}) = 0$.
($\Rightarrow$) If $F(\vec{v},\vec{v}) = 0$ and $F$ is bilinear, then by letting $\vec{v} = \vec{v}+\vec{u}$ we have:

$$\begin{aligned} 0 = F(\vec{v}+\vec{u},\vec{v}+\vec{u}) =& F(\vec{v}+\vec{u},\vec{v}) + F(\vec{v}+\vec{u},\vec{u}) \\ =& F(\vec{v},\vec{v}) + F(\vec{u},\vec{v}) + F(\vec{v},\vec{u}) + F(\vec{u},\vec{u}) \\ =& 0 + F(\vec{u},\vec{v}) + F(\vec{v},\vec{u}) + 0 \\ =& F(\vec{u},\vec{v}) + F(\vec{v},\vec{u}) \end{aligned}$$

Therefore, $F(\vec{u},\vec{v}) = -F(\vec{v},\vec{u})$, for any $\vec{u},\vec{v} \in \mathbb{R}^2$ as desired.
**b)** By part a) we know $F$ is alternating if and only if $F(\vec{v},\vec{v}) = 0$, for all $\vec{v} \in \mathbb{R}^2$. So setting $\vec{u} = a\vec{e}_1 + c\vec{e}_2$ and $\vec{v} = b\vec{e}_1 + d\vec{e}_2$ we observe:

$$\begin{aligned} F(\begin{bmatrix} a \\ c \end{bmatrix}, \begin{bmatrix} b \\ d \end{bmatrix}) =& F(\begin{bmatrix} a \\ 0 \end{bmatrix} + \begin{bmatrix} 0 \\ c \end{bmatrix}, \begin{bmatrix} b \\ d \end{bmatrix}) \\ =& F(\begin{bmatrix} a \\ 0 \end{bmatrix}, \begin{bmatrix} b \\ d \end{bmatrix}) + F(\begin{bmatrix} 0 \\ c \end{bmatrix}, \begin{bmatrix} b \\ d \end{bmatrix}) \\ =& a \cdot F(\begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} b \\ d \end{bmatrix}) + c \cdot F(\begin{bmatrix} 0 \\ 1 \end{bmatrix}, \begin{bmatrix} b \\ d \end{bmatrix}) \\ =& a \cdot [F(\begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} b \\ 0 \end{bmatrix} + \begin{bmatrix} 0 \\ d \end{bmatrix})] + c \cdot [F(\begin{bmatrix} 0 \\ 1 \end{bmatrix}, \begin{bmatrix} b \\ 0 \end{bmatrix} + \begin{bmatrix} 0 \\ d \end{bmatrix})] \\ =& \cdots \\ =& ad \cdot F(\begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \end{bmatrix}) + (-cd) \cdot F(\begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \end{bmatrix}) \end{aligned}$$

$= ad - cd,$ as desired.

**Problem 2.** Let $M = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \in \mathbb{R}^{2\times 2}$, and consider the map $T : \mathbb{R}^{2\times 2} \to \mathbb{R}^{2\times 2}$ defined by $T(A) = AM$.

(a) Prove that $T$ is a linear transformation.
(b) Find the $\mathcal{E}$-matrix $[T]_{\mathcal{E}}$ of $T$, where $\mathcal{E}$ is the ordered basis

$$\mathcal{E} = (E_{11}, E_{12}, E_{21}, E_{22}) = \left( \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \right)$$

of $\mathbb{R}^{2\times 2}$. Your answer should be in terms of the entries of $M$.
(c) Compute $\det[T]_{\mathcal{E}}$.
(d) Compute $\det[T]_{\mathcal{B}}$, where $\mathcal{B}$ is the ordered basis of $\mathbb{R}^{2\times 2}$ given by

$$\mathcal{B} = \left( \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ \pi & 1 \end{bmatrix}, \begin{bmatrix} 0 & 7 \\ 8 & 9 \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \right).$$

(e) Either prove that $T$ is always diagonalizable no matter what $M$ is, or provide an explicit example of a matrix $M$ for which $T$ is *not* diagonalizable and briefly explain why your example works.

**Solution.**
**a)** $T(A+B) = (A+B)M = AM+BM = T(A)+T(B)$, and $T(kA) = kAM = k(AM) = kT(A)$ so $T$ respects both addition and scalar multiplication, so is linear.
**b)** Now let $M = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$. To find the matrix of $T$ in the given basis (which will be $4 \times 4$), we compute it column by column, seeing where the basis elements go and computing their coordinates:

$$\begin{bmatrix} a & c & 0 & 0 \\ b & d & 0 & 0 \\ 0 & 0 & a & c \\ 0 & 0 & b & d \end{bmatrix}$$

**c) & d)** The determinant of this matrix is $(ad-bc)^2$. This is therefore the determinant of $T$ as well as the matrix of $T$ in any basis, including the basis given in (d).

**e)** Example let $M$ be a rotation of any angle $\theta \neq n\pi$ for any $n \in \mathbb{N}$. Then $M$ is not diagonalizable over $\mathbb{R}$, hence $T$ is not diagonalizable over $\mathbb{R}$.

**Problem 3.** Let $\vec{u}, \vec{v}, \vec{w}$ be vectors in $\mathbb{R}^4$. Define the linear transformation $T : \mathbb{R}^4 \to \mathbb{R}$ by the rule $T(\vec{x}) = \det([\vec{x}\ \vec{u}\ \vec{v}\ \vec{w}])$ for all $\vec{x} \in \mathbb{R}^4$. (You do not have to prove that $T$ is linear.)

(a) Prove that there exists a unique vector $\vec{z} \in \mathbb{R}^4$ such that $T(\vec{x}) = \vec{z} \cdot \vec{x}$ for all $\vec{x} \in \mathbb{R}^4$, and find the components of $\vec{z}$ in terms of the vectors $\vec{u}, \vec{v}$, and $\vec{w}$. (Hint: $\vec{x} = x_1\vec{e}_1 + x_2\vec{e}_2 + x_3\vec{e}_3 + x_4\vec{e}_4$.)
(b) Find the vector $\vec{z}$ (as in part (a)) when $\vec{u} = \vec{e}_1$, $\vec{v} = \vec{e}_2$, and $\vec{w} = \vec{e}_3$ are the first three standard basis vectors in $\mathbb{R}^4$.
(c) When is $\vec{z} = \vec{0}$? (Your answer should be in terms of $\vec{u}$, $\vec{v}$, and $\vec{w}$.)
(d) Prove that $\vec{z}$ is orthogonal to each of $\vec{u}, \vec{v}$ and $\vec{w}$, and find $\det([\vec{z}\ \vec{u}\ \vec{v}\ \vec{w}])$ in terms of $\|\vec{z}\|$.

**Solution.**
(a)

$$\begin{aligned} T(\vec{x}) &= \det\left(\begin{bmatrix} x_1\vec{e}_1 + x_2\vec{e}_2 + x_3\vec{e}_3 + x_4\vec{e}_4 & \vec{u} & \vec{v} & \vec{w} \end{bmatrix}\right). \\ &= x_1 \det\left(\begin{bmatrix} \vec{e}_1 & \vec{u} & \vec{v} & \vec{w} \end{bmatrix}\right) + x_2 \det\left(\begin{bmatrix} \vec{e}_2 & \vec{u} & \vec{v} & \vec{w} \end{bmatrix}\right) \\ &\quad + x_3 \det\left(\begin{bmatrix} \vec{e}_3 & \vec{u} & \vec{v} & \vec{w} \end{bmatrix}\right) + x_4 \det\left(\begin{bmatrix} \vec{e}_4 & \vec{u} & \vec{v} & \vec{w} \end{bmatrix}\right) \\ &= \vec{x} \cdot \vec{z} \end{aligned}$$

$$\text{where } \vec{z} = \begin{bmatrix} \det\left(\begin{bmatrix} \vec{e}_1 & \vec{u} & \vec{v} & \vec{w} \end{bmatrix}\right) \\ \det\left(\begin{bmatrix} \vec{e}_2 & \vec{u} & \vec{v} & \vec{w} \end{bmatrix}\right) \\ \det\left(\begin{bmatrix} \vec{e}_3 & \vec{u} & \vec{v} & \vec{w} \end{bmatrix}\right) \\ \det\left(\begin{bmatrix} \vec{e}_4 & \vec{u} & \vec{v} & \vec{w} \end{bmatrix}\right) \end{bmatrix}$$

(b) By (a), $\vec{z} = \begin{bmatrix} \det\left(\begin{bmatrix}\vec{e}_1 & \vec{e}_1 & \vec{e}_2 & \vec{e}_3\end{bmatrix}\right) \\ \det\left(\begin{bmatrix}\vec{e}_2 & \vec{e}_1 & \vec{e}_2 & \vec{e}_3\end{bmatrix}\right) \\ \det\left(\begin{bmatrix}\vec{e}_3 & \vec{e}_1 & \vec{e}_2 & \vec{e}_3\end{bmatrix}\right) \\ \det\left(\begin{bmatrix}\vec{e}_4 & \vec{e}_1 & \vec{e}_2 & \vec{e}_3\end{bmatrix}\right) \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ -1 \end{bmatrix} = -\vec{e}_4$

(c) If $\vec{u}, \vec{v}, \vec{w}$ are linearly independent then there exists $\vec{x}$ such that $\{\vec{x}, \vec{u}, \vec{v}, \vec{w}\}$ is a basis. Thus $\vec{z}.\vec{x} = \det\left(\begin{bmatrix}\vec{x} & \vec{u} & \vec{v} & \vec{w}\end{bmatrix}\right) \neq 0$ so $\vec{z} \neq \vec{0}$. If $\vec{u}, \vec{v}, \vec{w}$ are linearly dependent then for all $\vec{x}$ we have $\vec{z} \cdot \vec{x} = \det\left(\begin{bmatrix}\vec{x} & \vec{u} & \vec{v} & \vec{w}\end{bmatrix}\right) = 0$, so $\vec{z} = \vec{0}$.

Thereby, $\vec{z} = \vec{0}$ precisely when $\vec{u}, \vec{v}, \vec{w}$ are linearly dependent.

(d)

$$\vec{z} \cdot \vec{u} = \det\left(\begin{bmatrix}\vec{u} & \vec{u} & \vec{v} & \vec{w}\end{bmatrix}\right) = 0$$
$$\vec{z} \cdot \vec{v} = \det\left(\begin{bmatrix}\vec{v} & \vec{u} & \vec{v} & \vec{w}\end{bmatrix}\right) = 0$$
$$\vec{z} \cdot \vec{w} = \det\left(\begin{bmatrix}\vec{w} & \vec{u} & \vec{v} & \vec{w}\end{bmatrix}\right) = 0$$

So $\vec{z}$ is orthogonal to each of $\vec{u}$, $\vec{v}$, and $\vec{w}$. And $\det\left(\begin{bmatrix}\vec{z} & \vec{u} & \vec{v} & \vec{w}\end{bmatrix}\right) = \vec{z} \cdot \vec{z} = \|\vec{z}\|^2$

**Problem 4.** For a polynomial $p(x)$ and an $n \times n$ matrix $A$, let $p(A)$ denote the matrix obtained by plugging in $A$ for $x$. For example, if $p(x) = x^3 + 2x^2 + 3$, then $p(A) = A^3 + 2A^2 + 3I_n$. (Note that $I_n$ behaves like the constant "1" in $\mathbb{R}^{n\times n}$.)

(a) Prove that for every $n \times n$ matrix $A$ and for every eigenvalue $\lambda$ of $A$, the real number $p(\lambda)$ is an eigenvalue of the $n \times n$ matrix $p(A)$.
(b) Let $p$ be a polynomial and let $n \in \mathbb{N}$. Prove that if $S$ is an invertible $n \times n$ matrix, then for every $A \in \mathbb{R}^{n\times n}$ we have $p(S^{-1}AS) = S^{-1}p(A)S$.
(c) Let $p$ be a polynomial and let $A$ be an $n \times n$ matrix. Prove that if $A$ is diagonalizable, then every eigenvalue of $p(A)$ is of the form $p(\lambda)$ for some eigenvalue $\lambda$ of $A$.

**Solution.** (a) As $\lambda$ is an eigenvalue of $A$, we have that $A\vec{v} = \lambda\vec{v}$ for some non-zero vector $\vec{v} \in \mathbb{R}^n$. By induction on $n$, we can prove that

$$A^n\vec{v} = A^{n-1}(A\vec{v}) = A^{n-1}(\lambda\vec{v}) = \lambda(A^{n-1}\vec{v}) = \lambda(\lambda^{n-1}\vec{v}) = \lambda^n\vec{v}.$$

We now find that, for a polynomial $p(x) = c_kx^k + c_{k-1}x^{k-1} + \cdots + c_1x + c_0$,

$$\begin{aligned} p(A)\vec{v} &= (c_kA^k + c_{k-1}A^{k-1} + \cdots + c_1A + c_0I)\vec{v} \\ &= c_k(A^k\vec{v}) + c_{k-1}(A^{k-1}\vec{v}) + \cdots + c_1(A\vec{v}) + c_0\vec{v} \\ &= (c_k\lambda^k + c_{k-1}\lambda^{k-1} + \cdots + c_1\lambda + c_0)\vec{v} \\ &= p(\lambda)\vec{v}. \end{aligned}$$

Since $\vec{v} \neq \vec{0}$, this means that $p(\lambda)$ is an eigenvalue of $p(A)$. (This proof also show that $\vec{v}$ is among the eigenvectors corresponding to this eigenvalue.)

(b) If $p(x) = c_k x^k + c_{k-1}x^{k-1} + \cdots + c_1 x + c_0$, we have

$$\begin{aligned} p(S^{-1}AS) &= c_k(S^{-1}AS)^k + c_{k-1}(S^{-1}AS)^{k-1} + \cdots + c_1 S^{-1}AS + c_0 I \\ &= c_k S^{-1}A^k S + c_{k-1}S^{-1}A^{k-1}S + \cdots + c_1 S^{-1}AS + c_0 I \\ &= S^{-1}(c_k A^k + c_{k-1}A^{k-1} + \cdots + c_1 A + c_0 I)S \\ &= S^{-1}p(A)S. \end{aligned}$$

(c) The condition that $A$ is diagonalizable means that $A = SDS^{-1}$ for some invertible $n \times n$ matrix $S$ and some diagonal $n \times n$ matrix $D$. Write

$$D = \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{bmatrix}.$$

By the Diagonalization Theorem, $\lambda_1, \lambda_2, \ldots, \lambda_n$ are the eigenvalues of $A$.

On the other hand, we have that

$$D^n = \begin{bmatrix} \lambda_1^n & 0 & \cdots & 0 \\ 0 & \lambda_2^n & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n^n \end{bmatrix},$$

and so

$$\begin{aligned} p(D) &= c_k D^k + c_{k-1}D^{k-1} + \cdots + c_1 D + c_0 I \\ &= c_k \begin{bmatrix} \lambda_1^k & 0 & \cdots & 0 \\ 0 & \lambda_2^k & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n^k \end{bmatrix} + c_{k-1} \begin{bmatrix} \lambda_1^{k-1} & 0 & \cdots & 0 \\ 0 & \lambda_2^{k-1} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n^{k-1} \end{bmatrix} + \cdots \\ &\quad \cdots + c_1 \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{bmatrix} + c_0 \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{bmatrix} \\ &= \begin{bmatrix} c_k\lambda_1^k + \cdots + c_1\lambda_1 + c_0 & 0 & \cdots & 0 \\ 0 & c_k\lambda_2^k + \cdots + c_1\lambda_2 + c_0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & c_k\lambda_n^k + \cdots + c_1\lambda_n + c_0 \end{bmatrix} \\ &= \begin{bmatrix} p(\lambda_1) & 0 & \cdots & 0 \\ 0 & p(\lambda_2) & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & p(\lambda_n) \end{bmatrix}. \end{aligned}$$

This calculation shows that $p(D)$ is a diagonal matrix with diagonal entries $p(\lambda_1), p(\lambda_2), \ldots, p(\lambda_n)$, and so by the Diagonalization Theorem the eigenvalues of $p(D)$ are precisely $p(\lambda_1), p(\lambda_2), \ldots, p(\lambda_n)$.

Since $A$ is similar to $D$, Part (b) shows that $p(A)$ and $p(D)$ are similar and therefore they have the same eigenvalues. So the eigenvalues of $p(A)$ are also precisely $p(\lambda_1), p(\lambda_2), \ldots, p(\lambda_n)$. This shows that every eigenvalue of $p(A)$ is of the form $p(\lambda)$ for some eigenvalue $\lambda$ of $A$.

**Problem 5.**

(a) Let $A \in \mathbb{R}^{2\times 2}$ be a $2 \times 2$ matrix such that $A^2 = I_2$. Prove that $A$ is diagonalizable. (Hint: try factoring $A^2 - I_2$, and consider the possible ranks of the factors.)

(b) Does the same result hold for larger matrices? That is, if $A \in \mathbb{R}^{n\times n}$ is an $n \times n$ matrix for which $A^2 = I_n$, must $A$ be diagonalizable? Either prove this or give a counterexample.

**Solution.**

a) Since $A^2 = I_2$ we have $(A - I_2)(A + I_2) = 0$. If rank of $A + I_2$ is 2, then it is invertible, we can multiply $(A - I_2)(A + I_2) = 0$ by $(A + I_2)^{-1}$ on the right and get $A - I_2 = 0$. In this case $A = I_2$, so $A$ is diagonal. Similarly, if rank of $A - I_2$ is 2, then it is invertible, and we get $A = -I_2$, so $A$ is diagonal. Assume that both ranks are less than 2. Then both $A - I_2$ and $A + I_2$ are not invertible. This implies that both 1 and $-1$ are eigenvalues of $A$. Since $A$ is $2 \times 2$ and $A$ has 2 distinct eigenvalues, $A$ is diagonalizable.

b) Let $A$ be $n \times n$. As in part (a) we see that both 1 and $-1$ are eigenvalues of $A$. For any $v \in \mathbb{R}^n$, $Av + v$ is an eigenvector with eigenvalue 1, $Av - v$ is an eigenvector with eigenvalue $-1$, and $v = \frac{1}{2}(Av + v) - \frac{1}{2}(Av - v)$. So, every vector $v \in \mathbb{R}^n$ can be written as a sum of vectors in eigenspaces $E_1$ and $E_{-1}$. By choosing a basis of $\mathbb{R}^n$ consisting of eigenvectors from $E_1$ and $E_2$ we see that $A$ is diagonalizable.