Problem 1

Theorem 1. *Let A denote the $k \times k$ matrix*

$$\begin{bmatrix} 0 & 0 & \cdots & 0 & -a_0 \\ 1 & 0 & \cdots & 0 & -a_1 \\ 0 & 1 & \cdots & 0 & -a_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & -a_{k-1} \end{bmatrix},$$

where the a_i's are arbitrary scalars. Then the characteristic polynomial $f(t)$ of A is

$$f(t) = (-1)^k(a_0 + a_1 t + \cdots + a_{k-1}t^{k-1} + t^k).$$

Solution

Proof. We will proceed with induction on k. Starting with $k = 2$, we have

$$A = \begin{bmatrix} 0 & -a_0 \\ 1 & -a_1 \end{bmatrix}.$$

Thus,

$$\begin{aligned} |A - tI| &= \begin{vmatrix} -t & -a_0 \\ 1 & -a_1 - t \end{vmatrix} \\ &= -t(-a_1 - t) + a_0 \\ &= a_0 + a_1 t + t^2 \\ &= (-1)^2(a_0 + a_1 t + t^2), \end{aligned}$$

and we have proven the base case. Now suppose that for some $k >= 2$ the theorem holds. Now let A be a $k+1 \times k+1$ matrix of the desired form. Define B as

$$B = A - tI,$$

so that $f(t) = |B|$. Performing cofactor expansion along the first row, we have

$$|B| = -t|\tilde{B}_{1,1}| + (-1)^k(-a_0)|\tilde{B}_{1,k+1}|.$$

Noticing that $|\tilde{B}_{1,1}|$ is simply a characteristic polynomial for a $k \times k$ matrix of the desired form, our inductive hypothesis tells us that

$$|\tilde{B}_{1,1}| = (-1)^k(a_1 + a_2 t + \cdots + a_k t^{k-1} + t^k).$$

Furthermore, since $\tilde{B}_{1,k+1}$ is an upper triangular matrix, it's determinant is simply the product of the elements along its main diagonal. But all of the elements along it's main diagonal are 1, we can conclude that

$$|\tilde{B}_{1,k+1}| = 1.$$

Thus,

$$\begin{aligned} |B| &= -t(-1)^k(a_1 + a_2 t + \cdots + a_k t^{k-1} + t^k) - a_0(-1)^k \\ &= t(-1)^{k+1}(a_1 + a_2 t + \cdots + a_k t^{k-1} + t^k) + a_0(-1)^{k+1} \\ &= (-1)^{k+1}(a_0 + a_1 t + a_2 t^2 + \cdots + a_k t^k + t^{k+1}), \end{aligned}$$

and our proof is complete. □

Problem 2

Theorem 2. *Let T be a linear operator on a finite-dimensional vector space V. Then T is diagonalizable if and only if V is the direct sum of one-dimensional T-invariant subspaces.*

Solution

Proof. Suppose first that T is diagonlizable. Then, there exists a basis β of V consisting of eigenvectors of T

$$\beta = \{v_1, \cdots, v_n\}.$$

For each $1 \leq i \leq n$, define

$$\beta_i = \text{Span}\{v_i\}.$$

Clearly, by definition of an eigenvector, we have that each β_i is T-invariant. Furthermore, by linear independence of our v_i, we have

$$\beta_j \cap \sum_{i \neq j} \beta_i = \{0\}.$$

Finally, let $v \in V$. Since β is a basis of V, there exist scalars $\{a_1, \cdots, a_n\}$ such that

$$v = \sum_{i=1}^{n} a_i v_i.$$

Since $a_i v_i \in \beta_i$ for each i, it follows that

$$v \in \bigoplus_{i=1}^{n} \beta_i.$$

For the other direction, suppose V is the direct sum of one-dimensional T-invariant subspaces $\{\beta_1, \cdots, \beta_n\}$. Let $v_i \in \beta_i$, with $v_i \neq 0$, for each i. Then, since the direct sum of these subspaces equals V, it follows that $\{v_1, \cdots, v_n\}$ forms a basis of V. Furthermore, since β_i is T-invariant and one dimensional,

$$\begin{aligned} T(v_i) \in \beta_i &\implies T(v_i) \in \text{Span}\{v_i\} \\ &\implies (\exists \lambda_i \in \mathbb{F})(T(v_i) = \lambda_i v_i) \\ &\implies v_i \text{ is an eigenvector of } T. \end{aligned}$$

Thus, we have a basis of V consisting of eigenvectors of T, which allows us to conclude that T is diagonlizable. □

Problem 3

Let A be an $n \times n$ matrix with characteristic polynomial

$$f(t) = (-1)^n t^n + a_{n-1} t^{n-1} + \cdots + a_1 t + a_0.$$

Part (a):
Prove that if A is invertible, then

$$A^{-1} = (-1/a_0)[(-1)^n A^{n-1} + a_{n-1} A^{n-2} + \cdots + a_1 I_n].$$

Part (b):
Use (a) to compute A^{-1} for

$$A = \begin{bmatrix} 1 & 2 & 1 \\ 0 & 2 & 3 \\ 0 & 0 & -1 \end{bmatrix}.$$

Solution

Part (a):
By the Cayley-Hamilton theorem, we have

$$(-1)^n A^n + a_{n-1} A^{n-1} + \cdots + a_1 A + a_0 I = 0_n.$$

Multiplying by A^{-1} on the right of both sides of this equation and rearranging, we have

$$\begin{aligned}
-1)^n A^n A^{-1} + a_{n-1}A^{n-1}A^{-1} + \cdots + a_1 AA^{-1} + a_0 IA^{-1} &= 0_n A^{-1} \\
(-1)^n A^{n-1} + a_{n-1}A^{n-2} + \cdots + a_1 I + a_0 A^{-1} &= 0_n \\
a_0 A^{-1} &= 0_n - [(-1)^n A^{n-1} + a_{n-1}A^{n-2} + \cdots + a_1 I] \\
&= -[(-1)^n A^{n-1} + a_{n-1}A^{n-2} + \cdots + a_1 I] \\
A^{-1} &= (-1/a_0)[(-1)^n A^{n-1} + a_{n-1}A^{n-2} + \cdots + a_1 I_n],
\end{aligned}$$

as desired.

Part (b):

We will start by finding the characteristic polynomial of A:

$$\begin{aligned}
f(t) &= \begin{bmatrix} 1-t & 2 & 1 \\ 0 & 2-t & 3 \\ 0 & 0 & -1-t \end{bmatrix} \\
&= (1-t)(2-t)(-1-t) \\
&= -(1-t)(2-t)(1+t) \\
&= -(t^2 - 3t + 2)(1+t) \\
&= -(t^3 - 3t^2 + 2t + t^2 - 3t + 2) \\
&= -t^3 + 2t^2 + t - 2 \\
&= (-1)^3 t^3 + 2t^2 + t + (-2).
\end{aligned}$$

From this, we see

$$a_0 = -2 \qquad a_1 = 1 \qquad a_2 = 2.$$

In order to use part (a), we must find the square of A:

$$\begin{aligned}
A^2 &= \begin{bmatrix} 1 & 2 & 1 \\ 0 & 2 & 3 \\ 0 & 0 & -1 \end{bmatrix} \begin{bmatrix} 1 & 2 & 1 \\ 0 & 2 & 3 \\ 0 & 0 & -1 \end{bmatrix} \\
&= \begin{bmatrix} 1 & 6 & 6 \\ 0 & 4 & 3 \\ 0 & 0 & 1 \end{bmatrix}.
\end{aligned}$$

Finally, plugging it all in, we see

$$\begin{aligned}
A^{-1} &= (-1/a_0)[(-1)^n A^{n-1} + a_{n-1}A^{n-2} + \cdots + a_1 I_n] \\
&= \frac{1}{2}[-A^2 + 2A + I_n] \\
&= \frac{1}{2}\left(\begin{bmatrix} -1 & -6 & -6 \\ 0 & -4 & -3 \\ 0 & 0 & -1 \end{bmatrix} + \begin{bmatrix} 2 & 4 & 2 \\ 0 & 4 & 6 \\ 0 & 0 & -2 \end{bmatrix} + \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \right) \\
&= \frac{1}{2} \begin{bmatrix} 2 & -2 & -4 \\ 0 & 1 & 3 \\ 0 & 0 & -2 \end{bmatrix}.
\end{aligned}$$