Solution Set 6: Diagonalization and the Cayley-Hamilton Theorem

p.218, #20. We have that $M = \begin{pmatrix} A & B \\ O & C \end{pmatrix}$. Apply elementary row operations to the top half of the matrix until A is upper triangular. This will multiply the determinant by some scalar c_1 Then apply elementary row operations to the bottom half of the matrix until C is upper triangular. This will multiply the determinant by some scalar c_2. We now have a new matrix

$$M' = \begin{pmatrix} A' & B' \\ O & C' \end{pmatrix},$$

where A' and C' are upper triangular. The determinant of this matrix is $\det A' \det C'$ because it is upper triangular, so the determinant is just the product of the diagonal entries. So the determinant of the original matrix is $\det M = \frac{\det A' \det C'}{c_1 c_2}$. But $\det A = \frac{\det A'}{c_1}$ and $\det C = \frac{\det C'}{c_2}$, as applying the row operations that turn A into A' multiplies the determinant by c_1, and likewise applying the row operations that turn C into C' multiplies the determinant by c_2. So we see that $\det M = \det A \det C$.

p.248, #3b. (i) The characteristic polynomial for A is

$$\det \begin{pmatrix} -t & -2 & -3 \\ -1 & 1-t & -1 \\ 2 & 2 & 5-t \end{pmatrix},$$

which comes out to $-t^3 + 6t^2 - 11t + 6$. The roots of this polynomial are 1, 2, and 3, so these are the eigenvalues.

(ii) $E_1 = \ker L_{A-I} = \ker L_B$ for $B = \begin{pmatrix} -1 & -2 & -3 \\ -1 & 0 & -1 \\ 2 & 2 & 4 \end{pmatrix}$. This is the space $s \begin{pmatrix} 1 \\ 1 \\ -1 \end{pmatrix}$ for $s \in \mathbb{R}$.

$E_2 = \ker L_{A-2I} = \ker L_B$ for $B = \begin{pmatrix} -2 & -2 & -3 \\ -1 & -1 & -1 \\ 2 & 2 & 3 \end{pmatrix}$. This is the space $s \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}$ for $s \in \mathbb{R}$.

$E_3 = \ker L_{A-3I} = \ker L_B$ for $B = \begin{pmatrix} -3 & -2 & -3 \\ -1 & -2 & -1 \\ 2 & 2 & 2 \end{pmatrix}$. This is the space $s \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}$ for $s \in \mathbb{R}$.

(iii) We choose an eigenvector from each space, say $\begin{pmatrix} 1 \\ 1 \\ -1 \end{pmatrix}$, $\begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}$, and $\begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}$. Since these eigenvectors each have distinct eigenvalues, they are linearly independent, and thus form a basis for $\mathbb{R}^3$.

(iv) The matrix Q is given by choosing the columns to be the eigenvectors. So if $Q = \begin{pmatrix} 1 & 1 & 1 \\ 1 & -1 & 0 \\ -1 & 0 & -1 \end{pmatrix}$, then $Q^{-1}AQ = D$, for D a diagonal matrix whose entries are the respective eigenvalues. In this case, we will have $D = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{pmatrix}$.

p.250, #20. Clearly $f(0) = a_0$, as we just plug in 0 for t. Now, $f(t)$ is defined as $\det(A - tI)$. So $f(0) = \det(A - 0I) = \det A$. So $f(0) = a_0 = \det A$. Since A is invertible if and only if $\det A \neq 0$, we see that A is invertible if and only if $a_0 \neq 0$.

p.269, #3b. Writing the matrix of T with respect to the standard basis, we obtain

$$[T]_{std} = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix}.$$

The characteristic polynomial of T then comes out to

$$\det \begin{pmatrix} -t & 0 & 1 \\ 0 & 1-t & 0 \\ 1 & 0 & -t \end{pmatrix} = -(t-1)^2(t+1).$$

Since the multiplicity of -1 is 1, we know that the dimension of E_{-1} is 1. In fact, a basis is given by $\begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}$. Since the multiplicity of 1 is 2, the dimension of E_1 could be either 1 or 2. Well, we know that the dimension of E_1 is the nullity of $T - I$, which is 3 minus the rank of $[T]_{std} - I$ or 3 minus the rank of $\begin{pmatrix} -1 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & -1 \end{pmatrix}$. The rank of this matrix is 1, so we conclude that the dimension of E_1 is 2 (a basis is given by $\begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}$ and $\begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}$). Since the dimensions of all our eigenspaces match the respective multiplicities of the eigenvalues, and the characteristic polynomial splits, T is diagonalizable. The basis β that we seek is just a basis of eigenvectors, namely $\begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}$, $\begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}$, and $\begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}$.

p. 270, #8. If $\dim E_{\lambda_1} = n - 1$, then $m_{\lambda_1} \geq n - 1$. But we know that $m_{\lambda_2} \geq 1$. Since $m_{\lambda_1} + m_{\lambda_2} \leq n$, the only possible choice is $m_{\lambda_1} = n - 1$ and $m_{\lambda_2} = 1$. This means that $\dim E_{\lambda_2} = 1$. Also, we see that, since $m_{\lambda_1} + m_{\lambda_2} = n$, the characteristic polynomial splits. So the characteristic polynomial of A must be equal to $c(t - \lambda_1)^{n-1}(t - \lambda_2)$. This splits into linear factors, and for each of the two eigenvalues, the multiplicity matches the dimension of the eigenspace. So A is diagonalizable.

p.309, #3b. Let $v \in \ker T$. Then $T(v) = 0$, which is also in $\ker T$. So $T(\ker T) \subseteq \ker T$, and thus $\ker T$ is T-invariant. Now, let $v \in \text{im } T$. Then $T(v) \in \text{im } T$ by definition. So $T(\text{im } T) \subseteq \text{im } T$, and thus im T is T-invariant.

p.311, #18bc. b) By the Cayley-Hamilton theorem, A satisfies its own characteristic polynomial. So $(-1)^n A^n + a_{n-1}A^{n-1} + \cdots + a_1 A + a_0(I_n) = 0$. Moving the constant term to the other side and factoring, this gives

$$-a_0(I_n) = A((-1)^n A^{n-1} + a_{n-1}A^{n-2} + \cdots + a_2A + a_1I_n).$$

Dividing both sides by $-a_0$ gives

$$I_n = A(-1/a_0)((-1)^n A^{n-1} + a_{n-1}A^{n-2} + \cdots + a_2A + a_1I_n).$$

But this just means that $(-1/a_0)((-1)^n A^{n-1} + a_{n-1}A^{n-2} + \cdots + a_2A + a_1I_n)$ is A^{-1}.

c) The characteristic polynomial of A is $(1-t)(2-t)(-1-t) = (-1)^3t^3 + 2t^2 + t - 2$. So, by the formula from part (b), $A^{-1} = (1/2)((-1)^3A^2 + 2A + I_3)$. Since $A^2 = \begin{pmatrix} 1 & 6 & 6 \\ 0 & 4 & 3 \\ 0 & 0 & 1 \end{pmatrix}$, we have that

$$A^{-1} = (1/2)\left(\begin{pmatrix} -1 & -6 & -6 \\ 0 & -4 & -3 \\ 0 & 0 & -1 \end{pmatrix} + \begin{pmatrix} 2 & 4 & 2 \\ 0 & 4 & 6 \\ 0 & 0 & -2 \end{pmatrix} + \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}\right).$$

This comes out to $\begin{pmatrix} 1 & -1 & -2 \\ 0 & 1/2 & 3/2 \\ 0 & 0 & -1 \end{pmatrix}$.

p.311, #19. We proceed by induction on k. For $k = 1$, our matrix is $-a_0$, and the characteristic polynomial is $-a_0 - t = (-1)^1(a_0 + t)$, so the formula is correct. Now, assume that the formula is correct for k. We wish to prove it for $k+1$. The characteristic polynomial is the determinant of the matrix

$$\begin{pmatrix} -t & 0 & \cdots & 0 & -a_0 \\ 1 & -t & \cdots & 0 & -a_1 \\ 0 & 1 & \cdots & 0 & -a_2 \\ \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & \cdots & -t & -a_{k-2} \\ 0 & 0 & \cdots & 1 & -a_{k-1} - t \end{pmatrix}.$$

Computing this by expansion along the first row, we get only two nonzero terms. The first term is $-t(-1)^k(a_1 + a_2t + \cdots + a_{k-1}t^{k-2} + t^{k-1})$ by the induction hypothesis. The second term is $(-1)^{k+2}(-a_0)$. Adding these two terms together, we obtain that the characteristic polynomial is equal to $(-1)^{k+1}(a_0 + a_1t + a_2t^2 + \cdots + a_{k-1}t^{k-1} + t^k)$.

Extra Credit. a) The statement is true for $n = 1$ because $F_1 = F_2 = 1$. Now, assumne the statement is true for n. We want to prove it for $n+1$. Well,

$$\begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix}^{n+1} = \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix}\begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix}^n.$$

By the induction hypothesis,

$$\begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix}^n = \begin{pmatrix} * & F_n \\ * & F_{n+1} \end{pmatrix},$$

so multiplying on the left by $\begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix}$ gives

$$\begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix}^{n+1} = \begin{pmatrix} * & F_{n+1} \\ * & F_n + F_{n+1} \end{pmatrix} = \begin{pmatrix} * & F_{n+1} \\ * & F_{n+2} \end{pmatrix}.$$

So the statement holds for $n+1$.

b) The characteristic polynomial of A is $-t(1-t)-1 = t^2-t-1$. The roots are $\frac{1\pm\sqrt{5}}{2}$, so these are the eigenvalues. Let $\varphi = \frac{1+\sqrt{5}}{2}$. Then the eigenvalues are φ and $1-\varphi$. The corresponding eigenvectors are $\binom{1}{\varphi}$ and $\binom{1}{1-\varphi}$. So letting $Q = \begin{pmatrix} 1 & 1 \\ \varphi & 1-\varphi \end{pmatrix}$ gives $Q^{-1}AQ = D$, with D equal to the matrix $\begin{pmatrix} \varphi & 0 \\ 0 & 1-\varphi \end{pmatrix}$.

c) We know that $A = QDQ^{-1}$, so $A^n = QD^nQ^{-1}$. Now, we can compute $Q^{-1} = \frac{1}{\sqrt{5}} \begin{pmatrix} \varphi - 1 & 1 \\ \varphi & -1 \end{pmatrix}$. Computing QD^nQ^{-1} yields

$$\frac{1}{\sqrt{5}} \begin{pmatrix} (\varphi-1)\varphi^n + (1-\varphi)^n\varphi & \varphi^n - (1-\varphi)^n \\ (\varphi-1)\varphi^{n+1} + (1-\varphi)^{n+1}\varphi & \varphi^{n+1} - (1-\varphi)^{n+1} \end{pmatrix}.$$

d) Since the upper-right entry is equal to F_n, we substitute back in for φ and get

$$F_n = \frac{1}{\sqrt{5}} \left(\left(\frac{1+\sqrt{5}}{2} \right)^n - \left(\frac{1-\sqrt{5}}{2} \right)^n \right).$$