1. (12 points)

   (a) $det(A - \lambda I) = \begin{vmatrix} -\lambda & 1 \\ -1 & -\lambda \end{vmatrix} = \lambda^2 + 1 = (\lambda - i)(\lambda + i)$. The eigenvalues of A are i and $-i$.

   $$A - iI = \begin{pmatrix} -i & 1 \\ -1 & -i \end{pmatrix} \overset{RREF}{\Longrightarrow} \begin{pmatrix} 1 & i \\ 0 & 0 \end{pmatrix} \qquad A - (-iI) = \begin{pmatrix} i & 1 \\ -1 & i \end{pmatrix} \overset{RREF}{\Longrightarrow} \begin{pmatrix} 1 & -i \\ 0 & 0 \end{pmatrix}$$

   So $A \begin{bmatrix} a \\ b \end{bmatrix} = i \begin{bmatrix} a \\ b \end{bmatrix}$ if and only if $a + bi = 0$, e.g. $x = \begin{bmatrix} -i \\ 1 \end{bmatrix}$.

   Similarly, $A \begin{bmatrix} a \\ b \end{bmatrix} = -i \begin{bmatrix} a \\ b \end{bmatrix}$ if and only if $a - bi = 0$, e.g. $x = \begin{bmatrix} i \\ 1 \end{bmatrix}$.

   (b) $S = \begin{pmatrix} -i & i \\ 1 & 1 \end{pmatrix}$

   (c) By direct computation,

   $$A = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \quad A^2 = \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix} \quad A^3 = \begin{pmatrix} -i & 0 \\ 0 & i \end{pmatrix} \quad A^4 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

   By the division algorithm, we can express $k = 4s + t$ for some nonnegative integers s and t such that $0 \le t \le 3$. Thus

   $$A^k = A^{4s+t} = A^4 s A^t = (A^4)^s A^t = I^s A^t = A^t$$

   Thus,

   $$A^k = \begin{cases} I & \text{if } k = 4s \text{ for some integer } s \\ A & \text{if } k = 4s+1 \text{ for some integer } s \\ -I & \text{if } k = 4s+2 \text{ for some integer } s \\ -A & \text{if } k = 4s+3 \text{ for some integer } s \end{cases}$$

2. (24 points)

   (a) (induction) When $k = 1$, then by assumption $A = SDS^{-1}$. Suppose that a for positive integer $k > 1$, it holds that $A^{k-1} = SD^{k-1}S^{-1}$. Then

   $$\begin{aligned} A^k &= A^{k-1} \cdot A = (SD^{k-1}S^{-1}) \cdot (SDS^{-1}) = (SD^{k-1}) \underbrace{(S^{-1}S)}_{I} (DS^{-1}) \\ &= (SD^{k-1})(DS^{-1}) = S(D^{k-1}D)S^{-1} = SD^kS^{-1} \end{aligned}.$$

   By principle of mathematical induction, $A^k = SDS^{-1}$ for any positive integer k.

(b) Let x_j denote the j^{th} column of S and y_j^t denote the j^{th} column of S^{-1}, that is

$$\begin{aligned}
A^k = SD^kS^{-1} &= \left[\begin{array}{c|c|c|c} x_1 & x_2 & \cdots & x_n \end{array}\right] \left[\begin{array}{c|c|c|c} \lambda_1^k & 0 & \cdots & 0 \\ \hline 0 & \lambda_2^k & & \vdots \\ \hline \vdots & & \ddots & 0 \\ \hline 0 & \cdots & 0 & \lambda_n^k \end{array}\right] \left[\begin{array}{c} y_1^t \\ \hline y_2^t \\ \hline \vdots \\ \hline y_n^t \end{array}\right] \\
&= \left[\begin{array}{c|c|c|c} \lambda_1^k x_1 & \lambda_2^k x_2 & \cdots & \lambda_n^k x_n \end{array}\right] \left[\begin{array}{c} y_1^t \\ \hline y_2^t \\ \hline \vdots \\ \hline y_n^t \end{array}\right] \\
&= \lambda_1^k x_1 y_1^t + \lambda_2^k x_2 y_2^t + \cdots + \lambda_n^k x_n y_n^t \\
&= \textstyle\sum_{j=1}^n \lambda_j^k x_j y_j^t.
\end{aligned}$$

(c) Let k be a negative integer and let $t = -k$ (a positive integer). By (a), we know that $A^t = SD^tS^{-1}$. Since A is invertible, the eigenvalues $\lambda_1, \ldots, \lambda_n$ are nonzero and hence $\lambda_j^k = \frac{1}{\lambda^t}$ is defined. Thus D and D^t are invertible and

$$D^k = (D^t)^{-1} = \operatorname{diag}(\lambda_1^t, \ldots, \lambda_n^t)^{-1} = \operatorname{diag}\left(\frac{1}{\lambda_1^t}, \ldots, \frac{1}{\lambda_n^t}\right) = \operatorname{diag}\left(\frac{1}{\lambda_1}, \ldots, \frac{1}{\lambda_n}\right)^t = (D^{-1})^t.$$

Now,

$$A^k = (A^{-1})^t = (A^t)^{-1} = (SD^tS^{-1})^{-1} = S(D^t)^{-1}S^{-1} = SD^kS^{-1}.$$

Similar to the argument in (b), we get $A^k = SD^kS^{-1} = \sum_{j=1}^n \lambda_j^k x_n y_n^t$.

(d)

$$\begin{aligned}
f(A) &= a_m A^m + \cdots + a_1 A + a_0 I_n \\
&= a_m(SD^mS^{-1}) + \cdots + a_1(SDS^{-1}) + a_0(SI_nS^{-1}) \\
&= S(a_m D^m)S^{-1} + \cdots + S(a_1 D)S^{-1} + S(a_0 I_n)S^{-1} \\
&= S(a_m D^m + \cdots + a_1 D + a_0 I_n)S^{-1} \\
&= S\left(\begin{pmatrix} a_m\lambda_1^m & 0 & \cdots & 0 \\ 0 & a_m\lambda_2^m & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \cdots & 0 & a_m\lambda_n^m \end{pmatrix} + \cdots + \begin{pmatrix} a_1\lambda_1 & 0 & \cdots & 0 \\ 0 & a_1\lambda_2 & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \cdots & 0 & a_1\lambda_n \end{pmatrix} + a_0 I\right) S^{-1} \\
&= S\begin{pmatrix} a_m\lambda_1^m + \cdots + a_1\lambda_1 + a_0 & 0 & \cdots & 0 \\ 0 & a_m\lambda_2^m + \cdots + a_1\lambda_2 + a_0 & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \cdots & 0 & a_m\lambda_n^m + \cdots + a_1\lambda_n + a_0 \end{pmatrix} S^{-1} \\
&= S\begin{pmatrix} f(\lambda_1) & 0 & \cdots & 0 \\ 0 & f(\lambda_2) & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \cdots & 0 & f(\lambda_n) \end{pmatrix} S^{-1} \\
&= \textstyle\sum_{j=1}^n f(\lambda_j) x_j y_j^t
\end{aligned}$$

(e)

$$\begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & 1 \end{bmatrix} I_n = S^{-1}S = \begin{bmatrix} y_1^t \\ \hline y_2^t \\ \hline \vdots \\ \hline y_n^t \end{bmatrix} \left[\begin{array}{c|c|c|c} x_1 & x_2 & \cdots & x_n \end{array}\right] = \begin{bmatrix} y_1^t x_1 & y_1^t x_2 & \cdots & y_1^t x_n \\ y_2^t x_1 & y_2^t x_2 & \cdots & y_2^t x_n \\ \vdots & & \ddots & \vdots \\ y_n^t x_1 & y_n^t x_2 & \cdots & y_n^t x_n \end{bmatrix}$$

Thus, $y_i^t x_j = \delta_{ij} = \begin{cases} 1 & \text{if } i = j, \\ 0 & \text{if } i \neq j. \end{cases}$.

(f) Note that (e) states that for any invertible matrix,

$$(i^{th} \text{ row of matrix}) \cdot (j^{th} \text{ column of inverse}) = \delta_{ij}.$$

Applying this to the invertible matrix S^{-1}, we get

$$(i^{th} \text{ row of } S^{-1}) \cdot (j^{th} \text{ column of } (S^{-1})^{-1} = S) = y_i^t x_j = \delta_{ij}.$$

3. (12 points)

(a) Since A is diagonalizable, there are n linearly independent eigenvectors corresponding to the distinct eigenvalues $\lambda_1, \ldots, \lambda_k$. Suppose there are n_j linearly independent eienvectors corresponding to λ_j. Let $S \in M_n$ with columns equal to the linearly independent eigenvectors corresponding to $\lambda_1, \ldots, \lambda_k$, etc. Then $AS = S(\lambda_1 I_{n_1} \oplus \cdots \oplus \lambda_k I_{n_k})$. It follows that $S^{-1}AS = \lambda_1 I_{n_1} \oplus \cdots \oplus \lambda_k I_{n_k}$.

(b) Note that

$$f(\lambda_j) = (\lambda_j - \lambda_1) \cdots (\lambda_j - \lambda_{j-1}) \cdot 0 \cdot (\lambda_j - \lambda_{j+1}) \cdots (\lambda_j - \lambda_k) = 0.$$

By (a), we can write $A = T(\lambda_1 I_{n_1} \oplus \cdots \oplus \lambda_k I_{n_k})T^{-1}$. Let x_j be the j^{th} column of T and y_j^t be the j^{th} row of T^{-1}. By (2b), we have

$$\begin{aligned} f(A) &= \underbrace{f(\lambda_1)}_{0}\underbrace{(x_1 y_1^t + \cdots + x_{n_1} y_{n_1}^t)}_{n \times n} + \underbrace{f(\lambda_2)}_{0}\underbrace{(x_{n_1+1} y_{n_1+1}^t + \cdots + x_{n_1+n_2} y_{n_1+n_2}^t)}_{n \times n} + \cdots \\ &\quad \cdots + \underbrace{f(\lambda_k)}_{0}\underbrace{(x_{n_1+\cdots+n_{k-1}+1} y_{n_1+\cdots+n_{k-1}+1}^t + \cdots + x_n y_n^t)}_{n \times n} = 0_n \end{aligned}$$

(c) By a similar argument above,

$$\begin{aligned} g(A) &= \underbrace{g(\lambda_1)}_{0}(x_1 y_1^t + \cdots + x_{n_1} y_{n_1}^t) + \underbrace{g(\lambda_2)}_{0}(x_{n_1+1} y_{n_1+1}^t + \cdots + x_{n_1+n_2} y_{n_1+n_2}^t) + \cdots \\ &\quad \cdots + \underbrace{g(\lambda_k)}_{0}(x_{n_1+\cdots+n_{k-1}+1} y_{n_1+\cdots+n_{k-1}+1}^t + \cdots + x_n y_n^t) = 0_n \end{aligned}$$

4. (12 points)

(a) Let v_j be the j^{th} column of S^{-1}. Then

$$y(s) = S^{-1}x(s) = \sum_{j=1}^{n} v_j x_j(s) \quad \Longrightarrow \quad y'(s) = \sum_{j=1}^{n} v_j x_j'(s) = S^{-1}x'(s)$$

So

$$y'(s) = S^{-1}x'(s) = S^{-1}(Ax(s)) = S^{-1}(SDS^{-1})x(s) = D(S^{-1}x(s)) = Dy(s).$$

(b) From (a), $y'(s) = Dy(s) = (\lambda_1 y_1(s), \ldots, \lambda_n y_n(s))^t$. Thus for any $j = 1, \ldots, n$, we have $y_j'(s) = \lambda_j y_j(s)$. Equivalently, $\frac{dy_j}{ds} = \lambda_j y_j$. So

$$\frac{1}{y_j} dy_j = \lambda_j ds \Longrightarrow \int \frac{1}{y_j} dy_j = \int \lambda_j ds \Longrightarrow \ln |y_j| = \lambda_j s + C_j \Longrightarrow y_j(s) = e^{s\lambda_j + C_j} = c_j e^{\lambda_j s}$$

where $y_j(0) = c_j e^0 = c_j$.

(c) By assumption $y(s) = S^{-1}x(s)$. Multiplying both sides of this equation by the matrix S, we get $x(s) = Sy(s) = S(y_1(s), \ldots, y_n(s))^t = S(c_1 e^{s\lambda_1}, \ldots, c_n e^{s\lambda_n})^t$, based on (b).

5. (Extra 6 points)
Comparing columns in the equation $AX - XB = C$, we get

$$\begin{aligned} Col_j(C) = Col_j(AX - XB) &= Col_j(AX) - Col_j(XB) \\ &= ACol_j(X) - XCol_j(B) \\ &= ACol_j(X) - \begin{bmatrix} Col_1(X) & Col_2(X) & \cdots & Col_n(X) \end{bmatrix} \begin{bmatrix} b_{1j} \\ b_{2j} \\ \vdots \\ b_{nj} \end{bmatrix} \\ &= ACol_j(X) - (b_{1j}Col_1(X) + b_{2j}Col_2(X) + \cdots + b_{nj}Col_n(X)) \end{aligned}$$

Therefore,

$$Col_j(C) = ACol_j(X) - \sum_{i=1}^{n} b_{ij} Col_i(X),$$

Remark Consequently,

$$\begin{bmatrix} A & & \\ & \ddots & \\ & & A \end{bmatrix} \begin{bmatrix} Col_1(X) \\ \vdots \\ Col_n(X) \end{bmatrix} - \begin{bmatrix} b_{11}I_m & \cdots & b_{n1}I_m \\ \vdots & \ddots & \vdots \\ b_{1n}I_m & \cdots & b_{nn}I_m \end{bmatrix} \begin{bmatrix} Col_1(X) \\ \vdots \\ Col_n(X) \end{bmatrix} = \begin{bmatrix} Col_1(C) \\ \vdots \\ Col_n(C) \end{bmatrix}.$$

In particular, if A is in upper triangular form and B is in lower triangular form, and the two matrices have no common eigenvalues, then

$$\begin{bmatrix} A & & \\ & \ddots & \\ & & A \end{bmatrix} - \begin{bmatrix} b_{11}I_m & \cdots & b_{n1}I_m \\ \vdots & \ddots & \vdots \\ b_{1n}I_m & \cdots & b_{nn}I_m \end{bmatrix}$$

is in upper triangular form with nonzero diagonal entries.