Problem 1

Proof. Let n denote the size of C. We proceed with induction on n. Consider the base case $n = 1$. Then

$$\det(C - tI) = \left|-a_0 - t\right| = (-1)^n(t + a_0)$$

hence the base case holds. Assume that for a matrix of size $m \geq 1$ the given equation for the characteristic polynomial is correct. Consider the $(m+1) \times (m+1)$ matrix

$$C = \begin{pmatrix} 0 & 0 & \cdots & 0 & -a_0 \\ 1 & 0 & \cdots & 0 & -a_1 \\ 0 & 1 & \cdots & 0 & -a_2 \\ \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & \cdots & 0 & -a_{m-1} \\ 0 & 0 & \cdots & 1 & -a_m \end{pmatrix}.$$

Then finding the characteristic polynomial and expanding along the first row gives

$$\begin{aligned} \det(C - tI) &= \begin{vmatrix} -t & 0 & \cdots & 0 & -a_0 \\ 1 & -t & \cdots & 0 & -a_1 \\ 0 & 1 & \cdots & 0 & -a_2 \\ \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & \cdots & -t & -a_{m-1} \\ 0 & 0 & \cdots & 1 & -t - a_m \end{vmatrix} \\ &= (-t)\begin{vmatrix} -t & 0 & \cdots & -a_1 \\ 1 & -t & \cdots & -a_2 \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & 1 & -a_m - t \end{vmatrix} + (-1)^m(-a_0)\det \mathcal{I}_{-t} \end{aligned}$$

By the induction hypothesis,

$$\begin{aligned} &= (-t)(-1)^m(t^m + a_m t^{m-1} + \ldots + a_2 t + a_1) + (-1)^{m+1} a_0 \det \mathcal{I}_{-t} \\ &= (-1)^{m+1}(t^{m+1} + a_m t^m + \ldots + a_2 t^2 + a_1 t) + (-1)^{m+1} a_0 \det \mathcal{I}_{-t} \end{aligned}$$

where $\mathcal{I}_{-t}$ is the $m \times m$ identitiy matrix with the next upper diagonal as all $-t$. That is

$$\mathcal{I}_{-t} = \begin{pmatrix} 1 & -t & 0 & \cdots & 0 \\ 0 & 1 & -t & \cdots & 0 \\ 0 & 0 & 1 & \ddots & \vdots \\ \vdots & \vdots & & \ddots & -t \\ 0 & 0 & \cdots & 0 & 1 \end{pmatrix}.$$

Since $\mathcal{I}_{-t}$ is an upper triangular matrix, its determinant is equal to the product of its diagonal which is simply 1. Therefore

$$\begin{aligned} \det(C - tI) &= (-1)^{m+1}(t^{m+1} + a_m t^m + \ldots + a_2 t^2 + a_1 t) + (-1)^{m+1} a_0 \\ &= (-1)^{m+1}(t^{m+1} + a_m t^m + \ldots + a_1 t + a_0) \end{aligned}$$

which was to be shown. ∎

Problem 2

It is true for all matrices A.

Proof. Assume that $a_0 \neq 0$. Note that

$$p_A(0) = (-1)^n(0^n + a_{n-1} \cdot 0^{n-1} + \ldots + a_0) = (-1)^n a_0 \neq 0.$$

By the definition of the characteristic polynomial, we know $p_A(t) = \det(A - tI)$. Therefore

$$p_A(0) = \det(A - 0(I)) = \det A = (-1)^n a_0 \neq 0.$$

Since the determinant of A is non-zero, it must be invertible. ∎

Problem 3

Part A

Proof. We will show $\langle \cdot, \cdot \rangle_F$ satisfies the requirements of being an inner product. Let $A, B \in M_{n\times n}(\mathbb{R})$ and $s \in \mathbb{R}$. Some useful tools are

- The trace and transpose are linear operators.
- $(A^TA)_{ij} = \sum_{k=1}^{n} A_{ki}A_{kj} \implies \operatorname{tr}\left(A^TA\right) = \sum_{i=1}^{n}\sum_{j=1}^{n} A_{ij}^2$

We proceed with the 4 requirements of being an inner product.

1. We want to show linearity of the inner product. Take $C \in M_{n\times n}(\mathbb{R})$. Note that

$$\begin{aligned} \langle A + C, B\rangle_F &= \operatorname{tr}\left((A+C)^TB\right) \\ &= \operatorname{tr}\left((A^T + C^T)B\right) \\ &= \operatorname{tr}\left(A^TB + C^TB\right) \\ &= \operatorname{tr}\left(A^TB\right) + \operatorname{tr}\left(C^TB\right) = \langle A, B\rangle_F + \langle C, B\rangle_F \end{aligned}$$

   as a consequence of the linearity of the trace and transpose. Therefore $\langle \cdot, \cdot \rangle_F$ satisfies linearity.

2. We want to show $\langle sA, B\rangle_F = s\langle A, B\rangle_F$. This is the case since

$$\langle sA, B\rangle_F = \operatorname{tr}\left((sA)^TB\right) = \operatorname{tr}\left(sA^TB\right) = s\operatorname{tr}\left(A^TB\right) = s\langle A, B\rangle_F$$

   using the fact the trace and transpose are linear operators.

3. We want to show $\langle A, B\rangle_F = \langle B, A\rangle_F$. That is,

$$\operatorname{tr}\left(A^TB\right) = \operatorname{tr}\left(B^TA\right).$$

   Note that the trace of a matrix is the same as the trace of its transpose as the entries on the diagonal do not change. Therefore

$$\operatorname{tr}\left(A^TB\right) = \operatorname{tr}\left(\left(A^TB\right)^T\right) = \operatorname{tr}\left(B^TA\right)$$

   which was to be shown.

4. Assume that $A \neq 0$. Then there must be some entry of A that is non zero. Therefore

$$\langle A, A\rangle_F = \operatorname{tr}\left(A^T A\right) = \sum_{i=0}^{n}\sum_{j=0}^{n} A_{ij}^2 > 0$$

since some A_{ij} is non zero meaning its square must be larger than 0 and every other term is greater than or equal to 0.

$\langle\cdot,\cdot\rangle_F$ satisifies the required conditions meaning it is an inner product. ∎

Part B

Proof. Let $A \in M_{n\times n}(\mathbb{R})$ and assume that A is diagonalizable. Then $A = PDP^{-1}$ where P is unitary and D is a diagonal matrix with its entries being the eigenvalues of A. Note that

$$\begin{aligned} A^T A &= \left(PDP^{-1}\right)^T \left(PDP^{-1}\right) \\ &= \left(P^{-1}\right)^T D^T P^T PDP^{-1} \end{aligned}$$

Since P is unitary, its transpose is its inverse giving

$$\begin{aligned} &= PD^T IDP^{-1} \\ &= PD^2P^{-1}. \end{aligned}$$

Therefore $A^T A$ is also diagonalizable. Note that D^2 will be a diagonal matrix as well with entries λ_i^2 where λ_i are the original entries from D. Since D^2 is the diagonal matrix in the decomposition of $A^T A$, its entries λ_i^2 are the eigenvalues of $A^T A$. Since the trace of a matrix is equal to the sum of its eigenvalues, it follows

$$\operatorname{tr}\left(A^T A\right) = \sum_{i=1}^{n} \lambda_i^2 \implies \|A\|_F = \sqrt{\sum_{i=1}^{n} \lambda_i^2}$$ ∎