NATIONAL UNIVERSITY OF SINGAPORE
MATHEMATICS SOCIETY

PAST YEAR PAPER SOLUTIONS
Written by: Lin Mingyan Simon
Audited by: Chua Hongshen

MA2101 Linear Algebra II
AY 2010/2011 Sem 2

Question 1

(a) Take $\mathbf{w}_1, \mathbf{w}_2 \in \varphi(V_1)$. Then there exist $\mathbf{v}_1, \mathbf{v}_2 \in V_1$ such that $\varphi(\mathbf{v}_1) = \mathbf{w}_1$ and $\varphi(\mathbf{v}_2) = \mathbf{w}_2$. As V_1 is a vector subspace of V, it follows that $\mathbf{v}_1 + k\mathbf{v}_2 \in V_1$, so one has $\mathbf{w}_1 + k\mathbf{w}_2 = \varphi(\mathbf{v}_1) + k\varphi(\mathbf{v}_2) = \varphi(\mathbf{v}_1 + k\mathbf{v}_2) \in \varphi(V_1)$. Hence $\varphi(V_1)$ is a vector subspace of W.

(b) (i) Suppose $\varphi^{-1}(S)$ is a subspace of V. Since ϕ is surjective, it follows that $S = \phi(\varphi^{-1}(S))$. Thus by part (a) we have S to be a vector subspace of W.
Conversely, suppose S is a subspace of W. Take $\mathbf{v}_1, \mathbf{v}_2 \in \varphi^{-1}(S)$. Then one has $\varphi(\mathbf{v}_1), \varphi(\mathbf{v}_2) \in S$. Since S is a subspace of W, it follows that $\varphi(\mathbf{v}_1 + k\mathbf{v}_2) = \varphi(\mathbf{v}_1) + k\varphi(\mathbf{v}_2) \in S$. So we have $\mathbf{v}_1 + k\mathbf{v}_2 \in \varphi^{-1}(S)$. Hence, $\varphi^{-1}(S)$ is a subspace of V.

(ii) The statement would not be true if φ is not surjective. Let $V = W = \mathbb{R}$, $S = \{0, 1\}$, and define $\varphi : \mathbb{R} \to \mathbb{R}$, $\varphi(r) = 0$. Clearly, φ is a non-surjective linear transformation. Moreover, we see that $\varphi^{-1}(0) = V$, so it follows that $\varphi^{-1}(S) = V$. We see that $\varphi^{-1}(S)$ is a subspace of V, but S is not a subspace of W, so we are done.

(c) It suffices to show that $\mathbf{w}_0 = \mathbf{0}$. Take $\mathbf{v}_1, \mathbf{v}_2 \in \varphi^{-1}(\mathbf{w}_0)$. Then one has $\varphi(\mathbf{v}_1) = \varphi(\mathbf{v}_2) = \mathbf{w}_0$. As $\varphi^{-1}(\mathbf{w}_0)$ is a subspace of V, it follows that $\mathbf{v}_1 - \mathbf{v}_2 \in \varphi^{-1}(\mathbf{w}_0)$, so we must have $\mathbf{w}_0 = \varphi(\mathbf{v}_1 - \mathbf{v}_2) = \varphi(\mathbf{v}_1) - \varphi(\mathbf{v}_2) = \mathbf{0}$. We are done.

Question 2

(i) For all $\mathbf{u}_1, \mathbf{u}_2 \in U$ and $k \in F$, one has $f(\mathbf{u}_1 + k\mathbf{u}_2) = (\mathbf{u}_1 + k\mathbf{u}_2) + W = (\mathbf{u}_1 + W) + (k\mathbf{u}_2 + W) = f(\mathbf{u}_1) + f(k\mathbf{u}_2)$. So f is a linear transformation.

(ii) Take any $\mathbf{v} \in V$. Since $V = U + W$ we must have $\mathbf{v} = \mathbf{u} + \mathbf{w}$ for some $\mathbf{u} \in U$, $\mathbf{w} \in W$. This implies that $f(\mathbf{u}) = \mathbf{u} + W = (\mathbf{v} - \mathbf{w}) + W = \mathbf{v} + W$. So f is surjective.

(iii) Suppose f is an isomorphism, and there exists some $\mathbf{u} \in U \cap W$. We have $f(\mathbf{u}) = \mathbf{u} + W = W$ (because $u \in W$). As we also have $f(0_V) = 0_V + W = W$, by the injectivity of f we necessarily have $\mathbf{u} = 0_V$. So $V = U \oplus W$.
Conversely, suppose $V = U \oplus W$, and there exist $\mathbf{u}_1, \mathbf{u}_2 \in U$ such that $f(\mathbf{u}_1) = f(\mathbf{u}_2)$. Then one has $f(\mathbf{u}_1 - \mathbf{u}_2) = f(\mathbf{u}_1) - f(\mathbf{u}_2) = 0_{V/W} = W$, so we must have $\mathbf{u}_1 - \mathbf{u}_2 \in W$. As U is a subspace of V, we must have $\mathbf{u}_1 - \mathbf{u}_2 \in U$, so this implies that $\mathbf{u}_1 - \mathbf{u}_2 \in U \cap W = \{0_V\}$. Hence, we have $\mathbf{u}_1 = \mathbf{u}_2$, so this implies that f is injective. Together with parts (a) and (b) we conclude that f is an isomorphism.

Question 3

(a) If $n = 1$, then clearly A is diagonalizable over $\mathbb{C}$. Henceforth, we assume that $n > 1$. Note that A satisfies the polynomial $f(x) = x^n - x$. Moreover, we see that the degree of the characteristic polynomial of A must be n so it follows that the characteristic polynomial $p_A(x)$ must be $p_A(x) = x^n - x = x(x^{n-1} - 1)$.
Note that the eigenvalues of A are the roots of the equation $p_A(x) = x(x^{n-1} - 1) = 0$, that is, $0, 1, \zeta, \zeta^2, \cdots, \zeta^{n-1}$, where $\zeta = e^{\frac{2\pi i}{n-1}}$. Since the values $0, 1, \zeta, \zeta^2, \cdots, \zeta^{n-1}$ are all distinct, it follows that the characteristic (and minimal) polynomial of A can be factorized into distinct linear factors in $\mathbb{C}[x]$. So A is diagonalizable and we are done.

(b) For $n = 1$, the only Jordan Canonical Form of A is A itself.
For $n > 1$, we see that by part (a) A is diagonalizable and has n distinct eigenvalues. Thus, it follows that the only Jordan Canonical Form of A, up to re-ordering of the Jordan blocks, is

$$\begin{pmatrix} 0 & 0 & 0 & 0 & \cdots & 0 \\ 0 & 1 & 0 & 0 & \cdots & 0 \\ 0 & 0 & \zeta & 0 & \cdots & 0 \\ 0 & 0 & 0 & \zeta^2 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & 0 & \cdots & \zeta^{n-1} \end{pmatrix}$$

.

Question 4

Since $B \in M_n(\mathbb{C})$, it follows that there exists a Jordan Canonical Form J of B. Hence there exists some invertible matrix $P \in M_n(\mathbb{C})$ such that $P^{-1}BP = J$. Now, let

$$J = \begin{pmatrix} J_{k_1}(\lambda_1) & 0 & \cdots & 0 \\ 0 & J_{k_2}(\lambda_2) & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & J_{k_r}(\lambda_r) \end{pmatrix},$$

where $J_{k_i}(\lambda_i)$ denotes the Jordan block associated to the eigenvalue λ_i of size k_i. Since B is invertible, it follows that $\lambda_i \neq 0$ for all $i = 1, \cdots, r$. Based on that, we observe that

$$\begin{aligned} J_{k_i}(\lambda_i) &= \underbrace{\begin{pmatrix} \lambda_i & 1 & \cdots & 0 \\ 0 & \lambda_i & \cdots & 0 \\ \vdots & \vdots & \ddots & 1 \\ 0 & 0 & \cdots & \lambda_i \end{pmatrix}}_{k_i \text{ times}} \\ &= \underbrace{\begin{pmatrix} \lambda_i & 0 & \cdots & 0 \\ 0 & \lambda_i & \cdots & 0 \\ \vdots & \vdots & \ddots & 0 \\ 0 & 0 & \cdots & \lambda_i \end{pmatrix}}_{k_i \text{ times}} \underbrace{\begin{pmatrix} 1 & \frac{1}{\lambda_i} & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \frac{1}{\lambda_i} \\ 0 & 0 & \cdots & 1 \end{pmatrix}}_{k_i \text{ times}} \\ &= \underbrace{\begin{pmatrix} 1 & \frac{1}{\lambda_i} & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \frac{1}{\lambda_i} \\ 0 & 0 & \cdots & 1 \end{pmatrix}}_{k_i \text{ times}} \underbrace{\begin{pmatrix} \lambda_i & 0 & \cdots & 0 \\ 0 & \lambda_i & \cdots & 0 \\ \vdots & \vdots & \ddots & 0 \\ 0 & 0 & \cdots & \lambda_i \end{pmatrix}}_{k_i \text{ times}}. \end{aligned}$$

Hence, by defining

$$K_{k_i}(\lambda_i) = \underbrace{\begin{pmatrix} 1 & \frac{1}{\lambda_i} & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \frac{1}{\lambda_i} \\ 0 & 0 & \cdots & 1 \end{pmatrix}}_{k_i \text{ times}},$$

we see that $J_{k_i}(\lambda_i) = (\lambda_i I_{k_i})K_{k_i}(\lambda_i) = K_{k_i}(\lambda_i)(\lambda_i I_{k_i})$. Hence, by letting

$$C = \begin{pmatrix} \lambda_1 I_{k_1} & 0 & \cdots & 0 \\ 0 & \lambda_2 I_{k_2} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_r I_{k_r} \end{pmatrix}, \quad D = \begin{pmatrix} K_{k_1}(\lambda_1) & 0 & \cdots & 0 \\ 0 & K_{k_2}(\lambda_2) & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & K_{k_r}(\lambda_r) \end{pmatrix},$$

we see that $J = CD = DC$. Note that C is a diagonal matrix. Moreover, we see that D is upper triangular, and all of the entries on the leading diagonal of D is 1, so it follows that the only eigenvalue of D is 1.
Finally, by letting $B_s = PCP^{-1}$ and $B_u = PDP^{-1}$, we see that $B = B_sB_u = B_uB_s$, so conditions (i) and (iii) are satisfied. Moreover, we see that B_s is similar to the diagonal matrix C, so B_s is diagonalizable. Finally, as B_u is similar to D, we see that the only eigenvalue of B_u is 1. So condition (ii) is satisfied and we are done.

Question 5

(i) Let $B_1 = \{\mathbf{w}_i | i \in I\}$ and $B_2 = \{\mathbf{u}_j | j \in J\}$, where I and J are indexing sets. Take $\mathbf{v} \in W^\perp \subseteq V$. We have $\mathbf{v} = \sum\limits_{i\in I} a_i\mathbf{w}_i + \sum\limits_{j\in J} b_j\mathbf{u}_j$ for some $a_i, b_j \in \mathbb{C}$, $i \in I$, $j \in J$. Pick a $k \in I$. Then we have

$$\langle \mathbf{v}, \mathbf{w}_k\rangle = \langle \sum_{i\in I} a_i\mathbf{w}_i + \sum_{j\in J} b_j\mathbf{u}_j, \mathbf{w}_k\rangle = \sum_{i\in I} a_i\langle \mathbf{w}_i, \mathbf{w}_k\rangle + \sum_{j\in J} b_j\langle \mathbf{u}_j, \mathbf{w}_k\rangle = 0.$$

Since B is an orthonormal basis of V, it follows that $\langle \mathbf{w}_i, \mathbf{w}_k\rangle = 0$ for all $i \neq k$ and $\langle \mathbf{u}_j, \mathbf{w}_k\rangle$ for all $j \in J$. Moreover, we have $\langle \mathbf{w}_k, \mathbf{w}_k\rangle = 1$, so we must have $a_k = 0$. Hence, $a_i = 0$ for all $i \in I$ and hence $\mathbf{v} = \sum\limits_{j\in J} b_j\mathbf{u}_j \in \text{Span}(B_2)$. This implies that $W^\perp \subseteq \text{Span}(B_2)$.

Conversely, take $\mathbf{u} \in \text{Span}(B_2)$. Then one has $\mathbf{u} = \sum\limits_{j\in J} c_j\mathbf{u}_j$ for some $c_j \in \mathbb{C}$, $j \in J$. Then for all $d_i \in \mathbb{C}$, $i \in I$, one has

$$\langle \mathbf{u}, \sum_{i\in I} d_i\mathbf{w}_i\rangle = \langle \sum_{j\in J} c_j\mathbf{u}_j, \sum_{i\in I} d_i\mathbf{w}_i\rangle = \sum_{i\in I}\sum_{j\in J} d_i\overline{c_j}\langle \mathbf{u}_j, \mathbf{w}_i\rangle.$$

Since B is an orthonormal basis of V, it follows that $\langle \mathbf{u}_j, \mathbf{w}_i\rangle = 0$ for all $i \in I$ and $j \in J$. Hence, we must have $\langle \mathbf{u}, \sum\limits_{i\in I} d_i\mathbf{w}_i\rangle = 0$ so this shows that $\mathbf{u} \in W^\perp$. This implies that $\text{Span}(B_2) \subseteq W^\perp$. So we have $W^\perp = \text{Span}(B_2)$ as desired.

(ii) Since $W = \text{Span}(B_1)$, $W^\perp = \text{Span}(B_2)$ and $B_1 \cup B_2$ is a basis for V, it is clear that $V = W + W^\perp$. Next, take any $\mathbf{w} \in W \cap W^\perp$. Then it follows that $\mathbf{w} = \sum\limits_{j\in J} \alpha_j\mathbf{u}_j = \sum\limits_{i\in I} \beta_i\mathbf{w}_i$ for some $\alpha_j, \beta_i \in \mathbb{C}$, $j \in J$, $i \in I$. Then it follows that

$$\langle \mathbf{w}, \mathbf{w}\rangle = \langle \sum_{j\in J} \alpha_j\mathbf{u}_j, \sum_{i\in I} \beta_i\mathbf{w}_i\rangle = \sum_{i\in I}\sum_{j\in J} \alpha_i\overline{\beta_j}\langle \mathbf{u}_j, \mathbf{w}_i\rangle.$$

Since B is an orthonormal basis of V, it follows that $\langle \mathbf{u}_j, \mathbf{w}_i\rangle = 0$ for all $i \in I$ and $j \in J$. Hence, we must have $\langle \mathbf{w}, \mathbf{w}\rangle = 0$ so this implies that $\mathbf{w} = 0_V$. Hence, we have $V = W \oplus W^\perp$ as desired.

(iii) Take $\mathbf{v} \in W^{\perp}$. Then one has $\langle \mathbf{v}, \mathbf{w} \rangle = 0$ for all $\mathbf{w} \in W$. As we have $\langle T(\mathbf{v}), \mathbf{w} \rangle = \langle \mathbf{v}, T^*(w) \rangle$, and W is T^*-invariant, it follows that we must have $T^*(\mathbf{w}) \in W$. Hence we must have $\langle T(\mathbf{v}), \mathbf{w} \rangle = 0$ for all $\mathbf{w} \in W$, so this implies that $T(\mathbf{v}) \in W^{\perp}$. Hence $W^{\perp}$ is T-invariant.

Question 6

(a) Let $U = (a_{ij})$, $B_1 = (\mathbf{v}_1, \cdots, \mathbf{v}_n)$ and $B_2 = (\mathbf{u}_1, \cdots, \mathbf{u}_n)$. Since $B_2 = B_1 U$ it follows that $\mathbf{u}_j = \sum_{k=1}^{n} a_{kj}\mathbf{v}_k$ for all $j = 1, \cdots, n$. Suppose on the contrary that U is not invertible. Then it follows that the column vectors of U form a linearly dependent set, so there exists $c_1, \cdots, c_n \in \mathbb{C}$ such that $\sum_{j=1}^{n} c_j a_{ij} = 0$ for all i. This implies that $\sum_{j=1}^{n} c_j \mathbf{u}_j = \sum_{j=1}^{n}\sum_{k=1}^{n} c_j a_{kj}\mathbf{v}_k = \sum_{k=1}^{n}\left(\sum_{j=1}^{n} c_j a_{kj}\right)\mathbf{v}_k = 0_V$, so we have $\{\mathbf{u}_1, \cdots, \mathbf{u}_n\}$ to be a linearly dependent set, a contradiction. So U is invertible as desired.

(b) Note that the (i,j)-th entry of U^*U is $\sum_{k=1}^{n} \overline{a_{ki}}a_{kj}$ for all $i \neq j$, and the (i,i)-th entry of U^*U is $\sum_{k=1}^{n} |a_{ki}|^2$ for $i = 1, \cdots, n$. Since B_1 and B_2 are orthonormal bases of V it follows that $\langle \mathbf{v}_i, \mathbf{v}_i \rangle = \langle \mathbf{u}_i, \mathbf{u}_i \rangle = 1$ for all $i = 1, \cdots, n$, and $\langle \mathbf{v}_j, \mathbf{v}_i \rangle = \langle \mathbf{u}_j, \mathbf{u}_i \rangle = 0$ for all $i \neq j$. Thus, one has

$$\begin{aligned}
\langle \mathbf{u}_j, \mathbf{u}_i \rangle &= \langle \sum_{k=1}^{n} a_{kj}\mathbf{v}_k, \sum_{k=1}^{n} a_{ki}\mathbf{v}_k \rangle \\
&= \sum_{k=1}^{n} a_{kj}\overline{a_{ki}}\langle \mathbf{v}_k, \mathbf{v}_k \rangle + \sum_{k\neq\ell} a_{kj}\overline{a_{\ell i}}\langle \mathbf{v}_k, \mathbf{v}_\ell \rangle \\
&= \sum_{k=1}^{n} \overline{a_{ki}}a_{kj} = 0 \quad \text{for all } i \neq j, \\
\langle \mathbf{u}_i, \mathbf{u}_i \rangle &= \langle \sum_{k=1}^{n} a_{ki}\mathbf{v}_k, \sum_{k=1}^{n} a_{ki}\mathbf{v}_k \rangle \\
&= \sum_{k=1}^{n} a_{ki}\overline{a_{ki}}\langle \mathbf{v}_k, \mathbf{v}_k \rangle + \sum_{k\neq\ell} a_{ki}\overline{a_{\ell i}}\langle \mathbf{v}_k, \mathbf{v}_\ell \rangle \\
&= \sum_{k=1}^{n} |a_{ki}|^2 = 1 \quad \text{for all } i = 1, \cdots, n.
\end{aligned}$$

This implies that the (i,j)-th entry of U^*U is 0 for all $i \neq j$, and the (i,i)-th entry of U^*U is 1 for $i = 1, \cdots, n$. So we have $U^*U = I_n$, and hence U is unitary as required.

Question 7

(a) Take $\mathbf{v} \in \mathrm{Ker}(T - \lambda I_V)^{m_1}$. Then one has $(T - \lambda I_V)^{m_1}(\mathbf{v}) = 0_V$, so one has $(T - \lambda I_V)^r(\mathbf{v}) = (T - \lambda I_V)^{r-m_1}((T - \lambda I_V)^{m_1}(\mathbf{v})) = (T - \lambda I_V)^{r-m_1}(0_V) = 0_V$ for all $r > m_1$. This shows that $\mathbf{v} \in \mathrm{Ker}(T - \lambda I_V)^r$ so one has $\mathrm{Ker}(T - \lambda I_V)^{m_1} \subseteq \mathrm{Ker}(T - \lambda I_V)^r$ for all $r \geq m_1$.

Conversely, suppose $\mathbf{u} \in \mathrm{Ker}(T - \lambda I_V)^r$ with $r > m_1$. Write $m_T(x) = (x - \lambda)^{m_1} g(x)$. Since λ is a zero of $m_T(x)$ of multiplicity m_1, it follows that $(x - \lambda)^k \nmid g(x)$ for all positive integers k. Hence, there exist polynomials $u(x), v(x)$ such that $g(x)u(x) + (x - \lambda)^{r-m_1}v(x) = 1$, so this implies that $(x - \lambda)^{m_1} = (x - \lambda)^{m_1}\left(g(x)u(x) + (x - \lambda)^{r-m_1}v(x)\right) = u(x)m_T(x) + v(x)(x - \lambda)^r$.

Hence, we have

$$\begin{aligned}(T-\lambda I_V)^{m_1}(\mathbf{u}) &= \left(u(T)m_T(T)+v(T)(T-\lambda I_V)^r\right)(\mathbf{u})\\ &= (u(T)m_T(T))(\mathbf{u})+(v(T)(T-\lambda I_V)^r)(\mathbf{u})\\ &= u(T)(m_T(T)(\mathbf{u}))+v(T)((T-\lambda I_V)^r(\mathbf{u}))\\ &= u(T)(0_V)+v(T)(0_V) \quad = 0_V.\end{aligned}$$

This shows that $\mathbf{u}\in \operatorname{Ker}(T-\lambda I_V)^{m_1}$ so one has $\operatorname{Ker}(T-\lambda I_V)^r \subseteq \operatorname{Ker}(T-\lambda I_V)^{m_1}$ for all $r \geq m_1$. So we have $\operatorname{Ker}(T-\lambda I_V)^{m_1} = \operatorname{Ker}(T-\lambda I_V)^r$ for all $r\geq m_1$ as desired.

(b) We shall prove the following lemmas:

Lemma 1. If A and B are commuting $n\times n$ diagonalizable complex matrices then there exists some invertible matrix P such that $P^{-1}AP$ and $P^{-1}BP$ are both diagonal.
Lemma 2. If C and D are commuting $n\times n$ nilpotent complex matrices then $C-D$ is nilpotent as well.
Lemma 3. If E is a $n\times n$ diagonalizable and nilpotent complex matrix then E is necessarily the zero matrix.

Proof of Lemma 1. Let $T_A, T_B : \mathbb{C}^n \to \mathbb{C}^n$ be linear operators on $\mathbb{C}^n$ whose representation matrices with respect to the standard ordered basis for $\mathbb{C}^n$ are A and B respectively. Let $\lambda_1,\cdots,\lambda_r$ be the eigenvalues of A (and T_A). Since A (and hence T_A) is diagonalizable it follows that $\mathbb{C}^n = \oplus_{i=1}^r E_{\lambda_i}$.

For each $i=1,\cdots,r$ and $v\in E_{\lambda_i}$, we have $T_A((T_B)(\mathbf{v})) = A(B\mathbf{v}) = B(A\mathbf{v}) = B(\lambda_i\mathbf{v}) = \lambda_i(B\mathbf{v}) = \lambda_i T_B(\mathbf{v})$, so this shows that the eigenspace E_{λ_i} is T_B-invariant. Hence, $T_B|_{E_{\lambda_i}}$ is a diagonalizable linear operator on E_{λ_i}. Hence, there exists a basis β_i for E_{λ_i} consisting of the eigenvectors of $T_B|_{E_{\lambda_i}}$.

However, the eigenvectors of $T_B|_{E_{\lambda_i}}$ are also eigenvectors for T_B. Moreover, any vector (in particular, basis vector) in E_{λ_i} is an eigenvector of T_A. Hence, the basis β_i consists of vectors that are both eigenvectors of T_A and T_B. So by concatenating the β_i's, we get an ordered basis β for $\mathbb{C}^n$ that consists of vectors that are both eigenvectors of T_A and T_B (and hence A and B). So there exists some invertible matrix P such that $P^{-1}AP$ and $P^{-1}BP$ are both diagonal. We are done.

Proof of Lemma 2. Since C and D are nilpotent it follows that there exist positive integers r, s such that $C^r = D^s = 0_n$. By the Binomial Theorem, this implies that

$$(C-D)^{r+s} = \sum_{k=0}^{r+s}\binom{r+s}{k} C^k D^{r+s-k} \quad \text{(because } C \text{ and } D \text{ commutes, hence } CD=DC\text{)}.$$

For all $k\geq r$, we have $C^k = 0_n$ thus C^kD^{r+s-k} for all $k\geq r$. Else, if $k<r$, then one has $r+s-k>s$ so we have $D^k=0_n$. Thus C^kD^{r+s-k} for all $k<r$, and therefore, we have $(C-D)^{r+s}=0_n$. So $C-D$ is nilpotent as desired.

Proof of Lemma 3. Since E is nilpotent, there exists some positive integer m such that $E^m = 0_n$. Moreover, as E is diagonalizable, there exists some invertible matrix P such that $F = P^{-1}EP$ is diagonal. This implies that $F^m = (P^{-1}EP)^m = P^{-1}E^mP = 0_n$, so we must have $F=0_n$. Hence $E = PFP^{-1} = 0_n$ and we are done.

With the above lemmas proven, we shall proceed to prove the main statement.

Write $f(x)$ as $f(x) = a_p x^p + a_{p-1}x^{p-1} \cdots + a_0$. Firstly, we have $A'_s A = A'_s(A'_s + A'_n) = (A'_s)^2 + A'_s A'_n = (A'_s)^2 + A'_n A'_s = (A'_s + A'_n)A'_s = AA'_s$, so A'_s commutes with A. Thus, one has

$$\begin{aligned}
A_s A'_s &= f(A)A'_s \\
&= (a_p A^p + a_{p-1}A^{p-1} \cdots + a_0 I_n)A'_s \\
&= a_p A^p A'_s + a_{p-1}A^{p-1}A'_s \cdots + a_0 A'_s \\
&= A'_s\left(a_p A^p\right) + A'_s\left(a_{p-1}A^{p-1}\right) \cdots + a_0 A'_s \quad \text{(because } AA'_s = A'_s A) \\
&= A'_s(a_p A^p + a_{p-1}A^{p-1} \cdots + a_0 I_n) \\
&= A'_s f(A) = A'_s A_s.
\end{aligned}$$

So this implies that A_s commutes with A'_s. By a similar argument above, we have A_n to commute with A'_n.

Now, we have $A_s + A_n = A = A'_s + A'_n$ so it follows that $A_s - A'_s = A'_n - A_n$. Since A_s commutes with A'_s, by Lemma 1 there exists some invertible matrix P such that $P^{-1}A_sP$ and $P^{-1}A'_sP$ are both diagonal. So $P^{-1}(A_s - A'_s)P = P^{-1}A_sP - P^{-1}A'_sP$ is diagonal, and hence $A_s - A'_s$ is diagonalizable. Moreover, A_n commutes with A'_n so by Lemma 2, $A'_n - A_n$ is nilpotent and hence $A_s - A'_s$ is nilpotent. By Lemma 3, we must have $A_s - A'_s$ to be the zero matrix. Hence, we have $A_s = A'_s$ and $A'_n = A_n$ as desired.

Question 8

(a) For all $X = (x_1, \cdots, x_n)^t$ and $Y = (y_1, \cdots, y_n)^t$, define the map $f : W \times W \to \mathbb{C}$ to be $f\left(\sum_{i=1}^n x_i\mathbf{w}_i, \sum_{j=1}^n y_j\mathbf{w}_j\right) = X^tD\overline{Y}$. We shall show that f defines a complex inner product on W.

Conjugate Symmetry

Note that the notion of positive-definiteness of D is well-defined if and only if D is self-adjoint. We have

$$\begin{aligned}
f\left(\sum_{i=1}^n y_i\mathbf{w}_i, \sum_{j=1}^n x_j\mathbf{w}_j\right) &= Y^tD\overline{X} \\
&= \overline{(\overline{Y}^t\overline{D}X)} \\
&= \overline{(\overline{Y}^tD^tX)} \quad \text{(because } D = (\overline{D})^t) \\
&= \overline{(X^tD\overline{Y})^t} \\
&= \overline{(X^tD\overline{Y})} = \overline{f\left(\sum_{i=1}^n x_i\mathbf{w}_i, \sum_{j=1}^n y_j\mathbf{w}_j\right)}.
\end{aligned}$$

So f is conjugate-symmetric.

Linearity in the first argument

For all $a \in \mathbb{C}$, $X = (x_1, \cdots, x_n)^t$, $X' = (x'_1, \cdots, x'_n)^t$ and $Y = (y_1, \cdots, y_n)^t$, we have

$$\begin{aligned}
f\left(a\sum_{i=1}^n x_i\mathbf{w}_i, \sum_{j=1}^n x_j\mathbf{w}_j\right) &= f\left(\sum_{i=1}^n ax_i\mathbf{w}_i, \sum_{j=1}^n x_j\mathbf{w}_j\right) \\
&= (aX)^tD\overline{Y} \\
&= a(X^tD\overline{Y}) = af\left(\sum_{i=1}^n x_i\mathbf{w}_i, \sum_{j=1}^n x_j\mathbf{w}_j\right),
\end{aligned}$$

$$\begin{aligned}
f\left(\sum_{i=1}^n x_i\mathbf{w}_i + \sum_{i=1}^n x_i'\mathbf{w}_i, \sum_{j=1}^n x_j\mathbf{w}_j\right) &= f\left(\sum_{i=1}^n (x_i + x_i')\mathbf{w}_i, \sum_{j=1}^n x_j\mathbf{w}_j\right)\\
&= (X+X')^t D\overline{Y}\\
&= X^t D\overline{Y} + (X')^t D\overline{Y}\\
&= f\left(\sum_{i=1}^n x_i\mathbf{w}_i, \sum_{j=1}^n x_j\mathbf{w}_j\right) + f\left(\sum_{i=1}^n x_i'\mathbf{w}_i, \sum_{j=1}^n x_j\mathbf{w}_j\right).
\end{aligned}$$

This shows that f is linear in the first argument.

Positive Definiteness
Since D is positive definite, it follows that $X^t D\overline{X} \geq 0$ for all X with equality if and only if $X = (0, \cdots, 0)^t$. Thus, one has $f\left(\sum\limits_{i=1}^n x_i\mathbf{w}_i, \sum\limits_{j=1}^n x_j\mathbf{w}_j\right) = X^t D\overline{X} \geq 0$ for all $X = (x_1, \cdots, x_n)^t$, with equality if and only if $X = (0, \cdots, 0)^t$.

Therefore, f defines a complex inner product on W as desired.

(b) (i) We have

$$\begin{aligned}
X^t A\overline{Y} &= (x_1, \cdots, x_n)\begin{pmatrix} a_{11} & \cdots & a_{1n}\\ \vdots & \ddots & \vdots\\ a_{n1} & \cdots & a_{nn}\end{pmatrix}\begin{pmatrix}\overline{y_1}\\ \vdots\\ \overline{y_n}\end{pmatrix}\\
&= \left(\sum_{i=1}^n a_{i1}x_i, \cdots, \sum_{i=1}^n a_{in}x_i\right)\begin{pmatrix}\overline{y_1}\\ \vdots\\ \overline{y_n}\end{pmatrix}\\
&= \sum_{j=1}^n \overline{y_j}\sum_{i=1}^n a_{ij}x_i\\
&= \sum_{i=1}^n\sum_{j=1}^n x_i\overline{y_j}a_{ij}\\
&= \sum_{i=1}^n\sum_{j=1}^n x_i\overline{y_j}\langle\mathbf{v}_i, \mathbf{v}_j\rangle = \langle\sum_{i=1}^n x_i\mathbf{v}_i, \sum_{j=1}^n y_j\mathbf{v}_j\rangle.
\end{aligned}$$

(ii) Since we have $a_{ji} = \langle\mathbf{v}_j, \mathbf{v}_i\rangle = \overline{\langle\mathbf{v}_i, \mathbf{v}_j\rangle} = \overline{a_{ij}}$ for all $i, j = 1, \cdots, n$, it follows that A is self-adjoint.

(iii) As we have $\langle\sum\limits_{i=1}^n x_i\mathbf{v}_i, \sum\limits_{j=1}^n x_j\mathbf{v}_j\rangle \geq 0$ for all $x_1, \cdots, x_n \in \mathbb{C}$ with equality if and only if $x_1 = \cdots = x_n = 0$, it follows that $X^t A\overline{X} = \langle\sum\limits_{i=1}^n x_i\mathbf{v}_i, \sum\limits_{j=1}^n x_j\mathbf{v}_j\rangle > 0$ for all non-zero X. So A is positive definite.