NATIONAL UNIVERSITY OF SINGAPORE
MATHEMATICS SOCIETY

PAST YEAR PAPER SOLUTIONS
with credits to Lin Mingyan Simon

MA2101 Linear Algebra II
AY 2009/2010 Sem 2

Question 1

(a) It is easy to see that $\mathbf{v}_1 + \mathbf{v}_4 = (\mathbf{v}_1 + \mathbf{v}_2) + (\mathbf{v}_3 + \mathbf{v}_4) - (\mathbf{v}_2 + \mathbf{v}_3)$ so this implies that $W_1 = \text{span}\{\mathbf{v}_1 + \mathbf{v}_2, \mathbf{v}_2 + \mathbf{v}_3, \mathbf{v}_3 + \mathbf{v}_4, \mathbf{v}_1 + \mathbf{v}_4\} = \text{span}\{\mathbf{v}_1 + \mathbf{v}_2, \mathbf{v}_2 + \mathbf{v}_3, \mathbf{v}_3 + \mathbf{v}_4\}$. It remains to check that $\{\mathbf{v}_1 + \mathbf{v}_2, \mathbf{v}_2 + \mathbf{v}_3, \mathbf{v}_3 + \mathbf{v}_4\}$ is linearly independent (and hence is a basis for W_1).

Suppose that this is not the case. Then there exist $a, b, c \in \mathbb{R}$, not all zero, such that $a(\mathbf{v}_1 + \mathbf{v}_2) + b(\mathbf{v}_2 + \mathbf{v}_3) + c(\mathbf{v}_3 + \mathbf{v}_4) = \mathbf{0}_V$. Since a, b and c are not all zero, at least one of a, $a + b$, $b + c$ and c is non-zero, say a. Then it follows that at least one of $a + b$, $b + c$ and c must be non-zero as well. Then one has $\mathbf{v}_1 = -\frac{1}{a}((a + b)\mathbf{v}_2 + (b + c)\mathbf{v}_3 + c\mathbf{v}_4)$, thereby contradicting the fact that $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3, \mathbf{v}_4\}$ is linearly independent.
Likewise, we would arrive at the same contradiction if any of $a + b$, $b + c$ and c is non-zero as well. So the desired holds, and hence $\dim(W_1) = 3$.

Clearly, $\{\mathbf{v}_4\}$ is a basis for W_2. So $\dim(W_2) = 1$.

Next, let $\mathbf{u} \in W_1 \cap W_2$. Then one has $\mathbf{u} = p\mathbf{v}_4 = q(\mathbf{v}_1 + \mathbf{v}_2) + r(\mathbf{v}_2 + \mathbf{v}_3) + s(\mathbf{v}_3 + \mathbf{v}_4)$ for some $p, q, r, s \in \mathbb{R}$. This implies that $q\mathbf{v}_1 + (q + r)\mathbf{v}_2 + (r + s)\mathbf{v}_3 + (s - p)\mathbf{v}_4 = \mathbf{0}_V$. Since $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3, \mathbf{v}_4\}$ is linearly independent, we must have $q = q + r = r + s = s - p = 0$, so this implies that $p = q = r = s = 0$. Hence, one has $\mathbf{u} = \mathbf{0}_V$, so $W_1 \cap W_2 = \{\mathbf{0}_V\}$. Therefore, $\dim(W_1 \cap W_2) = 0$.

Finally, $\dim(W_1 + W_2) = \dim(W_1) + \dim(W_2) - \dim(W_1 \cap W_2) = 3 + 1 - 0 = 4$.

(b) Since $W_1, W_2 \subseteq V$, $W_1 + W_2 \subseteq V$ and $\dim(W_1 + W_2) = \dim(V) = 4$, we must have $V = W_1 + W_2$. Also, by an earlier argument in part (a), we have $W_1 \cap W_2 = \{\mathbf{0}_V\}$. So $V = W_1 \oplus W_2$ as desired.

(c) Consider $U = \text{span}\{\mathbf{v}_1\}$. Clearly, $\dim(U) = 1$. By a similar argument in part (a), one can show that $W_1 \cap U = \{\mathbf{0}_V\}$ so one has $\dim(W_1 + U) = 4$. By a similar argument in part (b), we see that $V = W_1 \oplus U$, but $U \neq W_2$, so we are done.

Question 2

(a) Since X is a subspace of V, we must have $\mathbf{0}_V \in X$, so one has $\mathbf{0}_W = T(\mathbf{0}_V) \in T(X)$. Hence $T(X) \neq \phi$. Let $\mathbf{w}_1, \mathbf{w}_2 \in T(X)$. Then one has $\mathbf{w}_1 = T(\mathbf{v}_1)$ and $\mathbf{w}_2 = T(\mathbf{v}_2)$ for some $\mathbf{v}_1, \mathbf{v}_2 \in X$. As X is a subspace of V, it follows that $\mathbf{v}_1 + k\mathbf{v}_2 \in X$ for all $k \in \mathbb{R}$. Therefore, we have $\mathbf{w}_1 + k\mathbf{w}_2 = T(\mathbf{v}_1) + k\,T(\mathbf{v}_2) = T(\mathbf{v}_1 + k\mathbf{v}_2) \in T(X)$, and hence $T(X)$ is a subspace of W.

(b) (i) We shall show that $\text{Ker}(T) = \{(k, k, k) | k \in \mathbb{R}\}$. Let $(x, y, z) \in \mathbb{R}^3$. Then we have

$$\begin{aligned} & (x, y, z) \in \text{Ker}(T) \\ \Leftrightarrow\ & T(x, y, z) = (0, 0, 0) \\ \Leftrightarrow\ & (x - y, y - z, z - x) = (0, 0, 0) \\ \Leftrightarrow\ & x = y = z \\ \Leftrightarrow\ & (x, y, z) \in \{(k, k, k) | k \in \mathbb{R}\}. \end{aligned}$$

So the desired holds. Now, we see that $\{(1,1,1)\}$ is a basis for $\text{Ker}(T)$, so one has $\text{nullity}(T) = 1$. Then by the Rank-Nullity Theorem, one has $\text{rank}(T) = \dim(V) - \text{nullity}(T) = 3 - 1 = 2$.

(ii) Clearly, X is a subspace of $\mathbb{R}^3$. So $T(X)$ is a subspace of $\mathbb{R}^3$ as well, and hence $\dim(T(X))$ is well-defined. Now, we claim that $T(X) = \{(0,k,-k)|k \in \mathbb{R}\}$. Let $(x,y,z) \in \mathbb{R}^3$. We have

$$\begin{aligned}
& (x,y,z) \in T(X) \\
\Leftrightarrow\ & (x,y,z) = T(a,a,b) \quad \text{for some } a,b \in \mathbb{R} \\
\Leftrightarrow\ & (x,y,z) = (a-a, a-b, b-a) \quad \text{for some } a,b \in \mathbb{R} \\
\Leftrightarrow\ & x = 0, y = a-b, z = -(a-b) \quad \text{for some } a,b \in \mathbb{R} \\
\Leftrightarrow\ & x = 0, y = c, z = -c \quad \text{for some } c \in \mathbb{R} \\
\Leftrightarrow\ & (x,y,z) \in \{(0,k,-k)|k \in \mathbb{R}\}.
\end{aligned}$$

So we have $T(X) = \{(0,k,-k)|k \in \mathbb{R}\}$ as desired. Now, we see that $\{(0,1,-1)\}$ is a basis for $T(X)$, so one has $\dim(T(X)) = 1$.

Question 3

(a) The characteristic polynomial of $\mathbf{A}$, $\chi_{\mathbf{A}}(x)$, is

$$\begin{aligned}
\chi_{\mathbf{A}}(x) &= \det(x\mathbf{I}_4 - \mathbf{A}) \\
&= \det\begin{pmatrix} x-3 & -1 & 2 & -1 \\ 1 & x-1 & -2 & -1 \\ 0 & 0 & x-2 & -1 \\ 0 & 0 & 0 & x-2 \end{pmatrix} \\
&= (x-2)\det\begin{pmatrix} x-3 & -1 & 2 \\ 1 & x-1 & -2 \\ 0 & 0 & x-2 \end{pmatrix} \\
&= (x-2)^2 \det\begin{pmatrix} x-3 & -1 \\ 1 & x-1 \end{pmatrix} \\
&= (x-2)^2[(x-3)(x-1) - (-1)(1)] = (x-2)^4.
\end{aligned}$$

(b) We have

$$2\mathbf{I}_4 - \mathbf{A} = \begin{pmatrix} 2-3 & -1 & 2 & -1 \\ 1 & 2-1 & -2 & -1 \\ 0 & 0 & 2-2 & -1 \\ 0 & 0 & 0 & 2-2 \end{pmatrix} = \begin{pmatrix} -1 & -1 & 2 & -1 \\ 1 & 1 & -2 & -1 \\ 0 & 0 & 0 & -1 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

Now, it is easy to see that $\{\begin{pmatrix}-1 & -1 & 2 & -1\end{pmatrix}, \begin{pmatrix}1 & 1 & -2 & -1\end{pmatrix}\}$ is a linearly independent set, since $\begin{pmatrix}-1 & -1 & 2 & -1\end{pmatrix}$ is not a multiple of $\begin{pmatrix}1 & 1 & -2 & -1\end{pmatrix}$. Also, we see that $\begin{pmatrix}0 & 0 & 0 & -1\end{pmatrix} = \frac{1}{2}\left[\begin{pmatrix}-1 & -1 & 2 & -1\end{pmatrix} + \begin{pmatrix}1 & 1 & -2 & -1\end{pmatrix}\right]$, so it follows that the maximal number of linearly independent row vectors of $2\mathbf{I}_4 - \mathbf{A}$ is equal to 2.
Hence, one has $\text{rank}(2\mathbf{I}_4 - \mathbf{A}) = 2$, so it follows from the Rank-Nullity Theorem that $\dim(E_2) = \dim(\text{Ker}(2\mathbf{I}_4 - \mathbf{A})) = \text{nullity}(2\mathbf{I}_4 - \mathbf{A}) = 4 - \text{rank}(2\mathbf{I}_4 - \mathbf{A}) = 2$, where E_2 denotes the eigenspace of $\mathbf{A}$ that is associated with the eigenvalue 2.

(c) Firstly, we shall compute the minimal polynomial of $\mathbf{A}$, $m_{\mathbf{A}}(x)$. Note that $m_{\mathbf{A}}(x)|\chi_{\mathbf{A}}(x)$, so we must have $m_{\mathbf{A}}(x) = (x-2)^k$, where $1 \le k \le 4$.

Clearly, $k \neq 1$ since $\mathbf{A} - 2\mathbf{I}_4 \neq \mathbf{0}_4$. Now consider $(\mathbf{A} - 2\mathbf{I}_4)^2$. We have

$$(\mathbf{A} - 2\mathbf{I}_4)^2 = \begin{pmatrix} 1 & 1 & -2 & 1 \\ -1 & -1 & 2 & 1 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} 1 & 1 & -2 & 1 \\ -1 & -1 & 2 & 1 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

This implies that $\mathbf{A}$ satisfies the polynomial $(x-2)^2$. So the minimal polynomial of $\mathbf{A}$, $m_{\mathbf{A}}(x)$, is $m_{\mathbf{A}}(x) = (x-2)^2$.

Now, we note that the sum of the total sizes of the Jordan blocks must be equal to 4, since the degree of the characteristic polynomial of $\mathbf{A}$ is equal to 4.
Moreover, since the eigenspace of $\mathbf{A}$ that is associated with the eigenvalue 2 has dimension 2, it follows that the Jordan Canonical Form of $\mathbf{A}$ will have 2 Jordan blocks corresponding to the eigenvalue 2.
Finally, since $m_{\mathbf{A}}(x) = (x-2)^2$, it follows that the maximal size of each of the Jordan blocks with eigenvalue 2 must be 2, and there must be at least one Jordan block with eigenvalue 2, of size 2.

Based on the facts above, we conclude that there must be 2 Jordan blocks of size 2 with eigenvalue 2, and hence the Jordan Canonical Form of $\mathbf{A}$ is

$$\begin{pmatrix} 2 & 1 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 2 & 1 \\ 0 & 0 & 0 & 2 \end{pmatrix}.$$

Question 4

(a) Let the adjoint of T be T^*, and let $\mathbf{X}, \mathbf{Y} \in \mathcal{M}_{n\times n}(\mathbb{C})$. Then we have

$$\langle T(\mathbf{X}), \mathbf{Y}\rangle = \langle \mathbf{BX}, \mathbf{Y}\rangle = \mathbf{Tr}(\mathbf{BXY}^*) = \mathbf{Tr}(\mathbf{XY}^*\mathbf{B}) = \mathbf{Tr}(\mathbf{X}(\mathbf{B^*Y})^*) = \langle \mathbf{X}, \mathbf{B}^*\mathbf{Y}\rangle = \langle \mathbf{X}, T^*(\mathbf{Y})\rangle.$$

So the adjoint of T, T^*, is the linear operator defined by $T^*(\mathbf{X}) = \mathbf{B}^*\mathbf{X}$ for all $\mathbf{X} \in \mathcal{M}_{n\times n}(\mathbb{C})$.
Note that an ordered basis for $\mathcal{M}_{n\times n}(\mathbb{C})$ is $\mathcal{B} = \{E_{11}, E_{21}, E_{31}, \cdots, E_{n1}, E_{12}, E_{22}, \cdots, E_{n2}, \cdots, E_{nn}\}$, where E_{ij} denote the $n \times n$ matrix in $\mathcal{M}_{n\times n}(\mathbb{C})$ whose (i,j)-th entry is equal to 1, and whose other entries are equal to 0.

Next, for typographical convenience, we shall define the matrix $\text{diag}(\mathbf{A})$ to the following $n^2 \times n^2$ matrix in $\mathcal{M}_{n^2\times n^2}(\mathbb{C})$

$$\text{diag}(\mathbf{A}) = \underbrace{\begin{pmatrix} \mathbf{A} & \mathbf{0}_{n\times n} & \cdots & \mathbf{0}_{n\times n} \\ \mathbf{0}_{n\times n} & \mathbf{A} & & \vdots \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0}_{n\times n} & \mathbf{0}_{n\times n} & \cdots & \mathbf{A} \end{pmatrix}}_{n \text{ times}}$$

for all $A \in \mathcal{M}_{n^2\times n^2}(\mathbb{C})$.

Now, let the (i,j)-th entry of $\mathbf{B}$ be b_{ij}. For any $\mathbf{X} = (x_{ij}) \in \mathcal{M}_{n\times n}(\mathbb{C})$, we see that $[\mathbf{X}]_{\mathcal{B}} = \begin{pmatrix} x_{11} & x_{21} & \cdots x_{nn}\end{pmatrix}^T$, and the (i,j)-th entry of $T(\mathbf{X}) = \mathbf{BX}$ is equal to $\sum_{k=1}^{n} b_{ik}x_{kj}$.

Thus, by noting that $[T(\mathbf{X})]_{\mathcal{B}} = [T]_{\mathcal{B}}[\mathbf{X}]_{\mathcal{B}}$, and by observing that

$$\begin{aligned}[T(\mathbf{X})]_{\mathcal{B}} &= \begin{pmatrix} \sum_{k=1}^{n} b_{1k}x_{k1} & \sum_{k=1}^{n} b_{2k}x_{k1} & \cdots & \sum_{k=1}^{n} b_{nk}x_{kn} \end{pmatrix}^T \\ &= \underbrace{\begin{pmatrix} \mathbf{B} & \mathbf{0}_{n\times n} & \cdots & \mathbf{0}_{n\times n} \\ \mathbf{0}_{n\times n} & \mathbf{B} & & \vdots \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0}_{n\times n} & \mathbf{0}_{n\times n} & \cdots & \mathbf{B} \end{pmatrix}}_{n \text{ times}} \begin{pmatrix} x_{11} \\ x_{21} \\ \vdots \\ x_{nn} \end{pmatrix} = \operatorname{diag}(\mathbf{B})[\mathbf{X}]_{\mathcal{B}},\end{aligned}$$

we deduce that $[T]_{\mathcal{B}} = \operatorname{diag}(\mathbf{B})$. Similarly, we have $[T^*]_{\mathcal{B}} = \operatorname{diag}(\mathbf{B}^*)$.
With the setup above, we shall proceed to prove the assertion given in the question.

$$\begin{aligned} T \text{ is unitarily diagonalizable,} &\Leftrightarrow [T]_{\mathcal{B}} \text{ is normal.} \\ &\Leftrightarrow [T]_{\mathcal{B}}[T]^*_{\mathcal{B}} = [T]^*_{\mathcal{B}}[T]_{\mathcal{B}} \\ &\Leftrightarrow \operatorname{diag}(\mathbf{B})\operatorname{diag}(\mathbf{B})^* = \operatorname{diag}(\mathbf{B})^*\operatorname{diag}(\mathbf{B}) \\ &\Leftrightarrow \operatorname{diag}(\mathbf{B}\mathbf{B}^*) = \operatorname{diag}(\mathbf{B}^*\mathbf{B}) \\ &\Leftrightarrow \mathbf{B}\mathbf{B}^* = \mathbf{B}^*\mathbf{B} \\ &\Leftrightarrow \mathbf{B} \text{ is normal} \end{aligned}$$

(b) We note that $\mathbf{B}_1^* = \mathbf{B}_1$, so $\mathbf{B}_1$ is Hermitian (and hence normal). By part (aii), T_1 is unitarily diagonalizable. Similarly, $\mathbf{B}_3^* = \mathbf{B}_3$, so T_3 is unitarily diagonalizable as well. Finally, since

$$\mathbf{B}_2\mathbf{B}_2^* = \begin{pmatrix} 1 & i \\ i & 0 \end{pmatrix}\begin{pmatrix} 1 & -i \\ -i & 0 \end{pmatrix} = \begin{pmatrix} 2 & -i \\ i & 1 \end{pmatrix} \neq \begin{pmatrix} 2 & i \\ -i & 1 \end{pmatrix} = \begin{pmatrix} 1 & -i \\ -i & 0 \end{pmatrix}\begin{pmatrix} 1 & i \\ i & 0 \end{pmatrix} = \mathbf{B}_2^*\mathbf{B}_2,$$

we see that T_2 is not unitarily diagonalizable.

Question 5

(a) We shall prove the contrapositive of the statement. If $\mathbf{v}_1, \cdots, \mathbf{v}_k$ are linearly dependent vectors in $\mathbb{R}^n$, then there exist $a_1, \cdots, a_k \in \mathbb{R}$ such that $a_1\mathbf{v}_1 + \cdots + a_k\mathbf{v}_k = \mathbf{0}_n$. This implies that $a_1(\mathbf{A}\mathbf{v}_1) + \cdots + a_k(\mathbf{A}\mathbf{v}_k) = \mathbf{A}(a_1\mathbf{v}_1 + \cdots + a_k\mathbf{v}_k) = \mathbf{A}\mathbf{0}_n = \mathbf{0}_m$, so $\mathbf{A}\mathbf{v}_1, \cdots, \mathbf{A}\mathbf{v}_k$ are linearly dependent vectors in $\mathbb{R}^m$.

(b) Suppose $\mathbf{A}\mathbf{v}_1, \cdots, \mathbf{A}\mathbf{v}_k$ are linearly independent vectors in $\mathbb{R}^m$, and suppose there exist $a_1, \cdots, a_k \in \mathbb{R}$ such that $a_1(N + \mathbf{v}_1) + \cdots + a_k(N + \mathbf{v}_k) = N + (a_1\mathbf{v}_1 + \cdots + a_k\mathbf{v}_k) = \mathbf{0}_{\mathbb{R}/N} = N$. Then one has $a_1\mathbf{v}_1 + \cdots + a_k\mathbf{v}_k \in N$, and hence $\mathbf{A}(a_1\mathbf{v}_1 + \cdots + a_k\mathbf{v}_k) = a_1(\mathbf{A}\mathbf{v}_1) + \cdots + a_k(\mathbf{A}\mathbf{v}_k) = \mathbf{0}_m$, so we must have $a_i = 0$ for all $i = 1, \cdots k$. So $N + \mathbf{v}_1, \cdots, N + \mathbf{v}_k$ are linearly independent vectors in $\mathbb{R}/N$.

Conversely, suppose $N + \mathbf{v}_1, \cdots, N + \mathbf{v}_k$ are linearly independent vectors in $\mathbb{R}/N$, and there exist $b_1, \cdots, b_k \in \mathbb{R}$ such that $b_1(\mathbf{A}\mathbf{v}_1) + \cdots + b_k(\mathbf{A}\mathbf{v}_k) = \mathbf{A}(b_1\mathbf{v}_1 + \cdots + b_k\mathbf{v}_k) = \mathbf{0}_m$. This implies that $b_1\mathbf{v}_1 + \cdots + b_k\mathbf{v}_k \in N$, and thus $b_1(N + \mathbf{v}_1) + \cdots + b_k(N + \mathbf{v}_k) = N + (b_1\mathbf{v}_1 + \cdots + b_k\mathbf{v}_k) = N = \mathbf{0}_{\mathbb{R}/N}$. This implies that $b_i = 0$ for all $i = 1, \cdots k$, so $\mathbf{A}\mathbf{v}_1, \cdots, \mathbf{A}\mathbf{v}_k$ are linearly independent vectors in $\mathbb{R}^m$. We are done.

(c) Take $\mathbf{A} = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$, $\mathbf{v}_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$ and $\mathbf{v}_1 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$. Clearly, $\mathbf{v}_1$ and $\mathbf{v}_2$ are linearly independent vectors in $\mathbb{R}^2$, but $\mathbf{A}\mathbf{v}_1 = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}\begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}\begin{pmatrix} 0 \\ 1 \end{pmatrix} = \mathbf{A}\mathbf{v}_2$, so $\mathbf{A}\mathbf{v}_1$ and $\mathbf{A}\mathbf{v}_2$ are not linearly independent vectors in $\mathbb{R}^2$.

Since $\mathbf{A}\begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}\begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} a+b \\ a+b \end{pmatrix}$ for all $a, b \in \mathbb{R}$, we see that a basis for $\left\{\mathbf{A}\mathbf{u} | \mathbf{u} \in \mathbb{R}^2\right\}$ is $\left\{\begin{pmatrix} 1 \\ 1 \end{pmatrix} = \mathbf{A}\begin{pmatrix} 1 \\ 0 \end{pmatrix}\right\}$. Next, by part (b), we deduce that if $\left\{\mathbf{A}\begin{pmatrix} 1 \\ 0 \end{pmatrix}\right\}$ is a basis for $\left\{\mathbf{A}\mathbf{u} | \mathbf{u} \in \mathbb{R}^2\right\}$ then the basis for $\mathbb{R}^2/N$ is $\left\{N + \begin{pmatrix} 1 \\ 0 \end{pmatrix}\right\}$.

Question 6

(a) Since W is both S-invariant and T-invariant, it follows that $S(\mathbf{w}) \in W$ and $T(\mathbf{w}) \in W$ for all $\mathbf{w} \in W$. So one has $(S \circ T)(\mathbf{w}) = S(T(\mathbf{w})) \in W$, and thus W is $(S \circ T)$-invariant. It follows that $(S \circ T)|_W(\mathbf{w}) = (S \circ T)(\mathbf{w}) = S(T(\mathbf{w})) = S(T|_W(\mathbf{w})) = S|_W(T|_W(\mathbf{w})) = ((S|_W) \circ (T|_W))(\mathbf{w})$ for all $\mathbf{w} \in W$, so one has $(S \circ T)|_W = (S|_W) \circ (T|_W)$ as desired.

(b) For any $\mathbf{v} \in E_\lambda$, we have $T(\mathbf{v}) = \lambda\mathbf{v}$, so one has $T(S(\mathbf{v})) = (T \circ S)(\mathbf{v}) = (S \circ T)(\mathbf{v}) = S(T(\mathbf{v})) = S(\lambda\mathbf{v}) = \lambda S(\mathbf{v})$. This shows that $S(\mathbf{v}) \in E_\lambda$ so E_λ is S-invariant as desired.

(c) We shall prove by strong induction on the dimension $m = \dim(V)$ of the vector space V, with the case $m = 1$ being trivial. Assume that the assertion holds for $m = 1, \cdots, k-1$, where $k > 1$. By induction hypothesis, there exists some ordered basis B for V such that $[T_1]_B, \cdots, [T_n]_B$ are diagonal.

Consider the case $m = k$. If we have $T_i(\mathbf{v}) = a_i\mathbf{v}$ for all $\mathbf{v} \in V$ for some $a_1, \cdots, a_n$, then clearly, any ordered basis would do since $[T_i]_B$ is a scalar (and hence diagonal) matrix for all $i = 1, \cdots, n$ and for any ordered basis B.
Else, let us assume WLOG that T_1 is not a scalar transformation. Then we see that T_1 must have at least two distinct eigenvalues (for otherwise if it has only one eigenvalue and it is diagonalizable, then there must exist some ordered basis B such that $[T_1]_B$ is a scalar matrix, a contradiction). Let the eigenvalues of T_1 be $\lambda_1, \cdots, \lambda_r$, with $\lambda_i \neq \lambda_j$ for all $i \neq j$ and $r > 1$. As T_1 is diagonalizable, we must have $V = \oplus_{i=1}^r E_{\lambda_i}$.

Note that by part (b), E_{λ_i} is a T_j-invariant subspace of V for all $i = 1, \cdots, r$ and $j = 1, \cdots, n$, so $T_j|_{E_{\lambda_i}}$ is a diagonalizable operator on E_{λ_i}. Moreover, by parts (a) and (b), we deduce that $(T_i|_{E_{\lambda_t}}) \circ (T_j|_{E_{\lambda_t}}) = (T_j|_{E_{\lambda_t}}) \circ (T_i|_{E_{\lambda_t}})$ for all $i, j \in \{1, \cdots n\}$ and $t = 1, \cdots, r$.

As $\dim(E_{\lambda_t}) < \dim(V)$ by assumption, it follows from the induction hypothesis that there exists an ordered basis B_t of E_{λ_t} for each t, such that $\left[T_i|_{E_{\lambda_t}}\right]_{B_t}$ is a diagonal matrix for all $i = 1, \cdots, n$ and $t = 1, \cdots, r$. Finally, by concatenating these ordered bases B_t to form an ordered basis $B = B_1 \cup \cdots \cup B_r$ for V, we see that $[T_1]_B, \cdots, [T_n]_B$ are diagonal matrices. So this completes the induction step and we are done.

Question 7

(a) Take any $\mathbf{v} \in V$. Then one has $P^2(\mathbf{v}) = P(\mathbf{v})$, so $P(\mathbf{v} - P(\mathbf{v})) = 0_V$. Hence $\mathbf{v} - P(\mathbf{v}) \in \mathrm{Ker}(R)$, so we have $\mathbf{v} = P(\mathbf{v}) + (\mathbf{v} - P(\mathbf{v})) \in \mathrm{R}(P) + \mathrm{Ker}(P)$ and hence $V = \mathrm{R}(P) + \mathrm{Ker}(P)$. Next, if $\mathbf{u} \in$

$\mathrm{R}(P) \cap \mathrm{Ker}(P)$, then one has $\mathbf{u} = P(\mathbf{w})$ for some $w \in V$. Hence $\mathbf{u} = P(\mathbf{w}) = P^2(\mathbf{w}) = P(\mathbf{u}) = \mathbf{0}_V$, so this implies that $\mathrm{R}(P) \cap \mathrm{Ker}(P) = \{\mathbf{0}_V\}$. Therefore we have $V = \mathrm{R}(P) \oplus \mathrm{Ker}(P)$ as desired.

(b) Define the map $P : \mathbb{R}^2 \to \mathbb{R}^2$ to be $P\left(\begin{pmatrix} a \\ b \end{pmatrix}\right) = \begin{pmatrix} a \\ 0 \end{pmatrix}$. Then it is clear that P is a linear operator on $\mathbb{R}^2$ and $P^2\left(\begin{pmatrix} a \\ b \end{pmatrix}\right) = P\left(\begin{pmatrix} a \\ 0 \end{pmatrix}\right) = \begin{pmatrix} a \\ 0 \end{pmatrix} = P\left(\begin{pmatrix} a \\ b \end{pmatrix}\right)$. However, it is clear that $P \neq O_V$ and $P \neq I_V$.
By observation, we see that $\mathrm{R}(P) = \left\{\begin{pmatrix} a \\ 0 \end{pmatrix} : a \in \mathbb{R}\right\}$ and $\mathrm{Ker}(P) = \left\{\begin{pmatrix} 0 \\ a \end{pmatrix} : a \in \mathbb{R}\right\}$.

(c) Take any $\mathbf{u} \in \mathrm{Ker}(P)$ and $\mathbf{v} \in V$. Then we see that $P(\mathbf{v}) \in \mathrm{R}(P)$. By assumption, we have $\langle P(\mathbf{u}+\mathbf{v}), (\mathbf{u}+\mathbf{v}) - P(\mathbf{u}+\mathbf{v})\rangle = 0$. This implies that

$$\begin{aligned}
\langle P(\mathbf{u}+\mathbf{v}), (\mathbf{u}+\mathbf{v}) - P(\mathbf{u}+\mathbf{v})\rangle &= \langle P(\mathbf{u}) + P(\mathbf{v}), \mathbf{u}+\mathbf{v} - P(\mathbf{u}) - P(\mathbf{v})\rangle \\
&= \langle P(\mathbf{v}), \mathbf{u}+\mathbf{v} - P(\mathbf{v})\rangle \\
&= \langle P(\mathbf{v}), \mathbf{u}\rangle + \langle P(\mathbf{v}), \mathbf{v} - P(\mathbf{v})\rangle \\
&= \langle P(\mathbf{v}), \mathbf{u}\rangle = 0.
\end{aligned}$$

Since this holds for all $\mathbf{v} \in V$, this would imply that $\mathbf{u} \in \mathrm{R}(P)^\perp$ so one has $\mathrm{Ker}(P) \subseteq \mathrm{R}(P)^\perp$.

Conversely, suppose $\mathbf{u} \in \mathrm{R}(P)^\perp$. Then one has $\langle P(\mathbf{v}), \mathbf{u}\rangle = 0$ for all $\mathbf{v} \in V$, and in particular, we have $\langle P(\mathbf{u}), \mathbf{u}\rangle = 0$.
This implies that $\langle P(\mathbf{u}), P(\mathbf{u})\rangle = 0 + \langle P(\mathbf{u}), P(\mathbf{u})\rangle = \langle P(\mathbf{u}), \mathbf{u} - P(\mathbf{u})\rangle + \langle P(\mathbf{u}), P(\mathbf{u})\rangle = \langle P(\mathbf{u}), \mathbf{u}\rangle = 0$, so we necessarily have $P(\mathbf{u}) = 0$. Hence $\mathbf{u} \in \mathrm{Ker}(P)$ so this implies that $\mathrm{R}(P)^\perp \subseteq \mathrm{Ker}(P)$.

Therefore, we have $\mathrm{Ker}(P) = \mathrm{R}(P)^\perp$ as desired.