MA2101 AY18/19 SEM 1 SOLUTIONS

JOSHUA KIM KWAN AND THANG PANG ERN

Question 1. Consider a 3×3 real matrix

$$A = \begin{pmatrix} 1 & 0 & 0 \\ 3 & -1 & -2 \\ -5 & 4 & 5 \end{pmatrix}.$$

- **(a)** Find the characteristic polynomial of A and show that the eigenvalues of A are 1 and 3.
- **(b)** Find a basis for the eigenspace $E_1(A)$.
- **(c)** Is A diagonalizable?
- **(d)** Find an invertible matrix P such that $P^{-1}AP$ is an upper triangular matrix.
  Hint: If you computation in (b) is correct, you will find that $(0,-1,1)^\top$ is an eigenvector of A.
- **(e)** Write down a Jordan canonical form for A. (You do not need to explain your answer.)

Solution.

- **(a)** The characteristic polynomial of A is given by

$$p_A(x) = \det \begin{pmatrix} x-1 & 0 & 0 \\ -3 & x+1 & 2 \\ 5 & -4 & x-5 \end{pmatrix} = (x-1)\det \begin{pmatrix} x+1 & 2 \\ -4 & x-5 \end{pmatrix} = (x-1)^2(x-3).$$

  The eigenvalues of A are precisely the roots of $p_A(x)$. When $p_A(x) = 0$, then $x = 1$ or $x = 3$. Hence the eigenvalues of A are 1 and 3.
- **(b)** When $x = 1$, consider the matrix equation $(I_3 - A)\,y = 0$ where $y \in \mathbb{R}^2$. Then

$$I_3 - A = \begin{pmatrix} 0 & 0 & 0 \\ -3 & 2 & 2 \\ 5 & -4 & -4 \end{pmatrix} \sim \begin{pmatrix} 1 & -4/5 & -4/5 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \end{pmatrix}$$

  so that $y = (0,a,-a)^\top$ for $a \in \mathbb{R}$. Therefore a basis for $E_1(A)$ is $\left\{(0,1,-1)^\top\right\}$.
- **(c)** No, A is not diagonalizable. Since

$$\dim_{\mathbb{R}} E_1(A) = 1 < 2 = \text{ algebraic multiplicity of 1 in } p_A(x)$$

  then A is not diagonalizable.

(d) Let

$$S = \begin{pmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 1 & 0 & 1 \end{pmatrix}.$$

Then

$$S^{-1}AS = \begin{pmatrix} 1 & -3 & 2 \\ 0 & 1 & 0 \\ 0 & -2 & 3 \end{pmatrix}.)$$

Let

$$P = S\begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}.$$

Then

$$P^{-1}AP = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}^{-1} S^{-1}AS \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}\begin{pmatrix} 1 & -3 & 2 \\ 0 & 1 & 0 \\ 0 & -2 & 3 \end{pmatrix}\begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 2 & -3 \\ 0 & 3 & -2 \\ 0 & 0 & 1 \end{pmatrix}$$

which is upper triangular.

(e) A Jordan canonical form for A is given by

$$J_A = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 3 \end{pmatrix}.$$

□

Question 2. Let V be a real vector space with a basis $B = \{v_1, v_2, v_3, v_4\}$ and let T be a linear operator on V such that

$$[T]_B = \begin{pmatrix} 2 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 2 & -2 & 0 & 0 \\ -1 & 1 & 1 & 2 \end{pmatrix}.$$

(a) Find a basis for $\text{Ker}(T)$ and a basis for $R(T)$.

(b) Find $\text{rank}(T)$ and $\text{nullity}(T)$.

(c) Show that $T^2 = 2T$.

(d) Write down the minimal polynomial of T. Is T diagonalizable? (You do not need to explain your answers.)

Solution.

(a) We first find a basis for $\operatorname{Ker}([T]_B)$. Since

$$[T]_B = \begin{pmatrix} 2 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 2 & -2 & 0 & 0 \\ -1 & 1 & 1 & 2 \end{pmatrix} \sim \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

then a basis for $\operatorname{Ker}([T]_B)$ is given by $\left\{(0,0,-2,1)^\top\right\}$. Hence, a basis for $\operatorname{Ker}(T)$ is $\{-2v_3+v_4\}$.

A basis for $R([T]_B)$ is given by $\left\{(2,0,2,-1)^\top,(0,2,-2,1)^\top,(0,0,0,1)^\top\right\}$. Hence, a basis for $R(T)$ is given by $\{2v_1+2v_3-v_4, 2v_2-2v_3+v_4, v_4\}$.

(b) We have

$$\operatorname{rank}(T) = \dim(R(T)) = 3$$
$$\operatorname{nullity}(T) = \dim(\operatorname{Ker}(T)) = 1.$$

(c) It suffices to show that $\left[T^2\right]_B = [2T]_B$. We have

$$\begin{aligned} \left[T^2\right]_B = [T]_B^2 &= \begin{pmatrix} 2 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 2 & -2 & 0 & 0 \\ -1 & 1 & 1 & 2 \end{pmatrix} \begin{pmatrix} 2 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 2 & -2 & 0 & 0 \\ -1 & 1 & 1 & 2 \end{pmatrix} \\ &= \begin{pmatrix} 4 & 0 & 0 & 0 \\ 0 & 4 & 0 & 0 \\ 4 & -4 & 0 & 0 \\ -2 & 2 & 2 & 4 \end{pmatrix} = 2\begin{pmatrix} 2 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 2 & -2 & 0 & 0 \\ -1 & 1 & 1 & 2 \end{pmatrix} = [2T]_B \end{aligned}$$

so that $T^2 = 2T$.

(d) The minimal polynomial is given by $m_T(x) = x^2-2x = x(x-2)$. Since the minimal polynomial splits into distinct linear factors, then T is diagonalizable. □

Question 3. Let $V = M_{n\times n}(\mathbb{F})$, where $\mathbb{F}$ is a field, and let $P \in V$ be a symmetric matrix, i.e. $P^\top = P$. Define $W = \left\{B \in V : BP \text{ is symmetric (i.e. } BP = (BP)^\top = P^\top B^\top = PB^\top\right\}$.
Warning: BP symmetric does not imply that B is symmetric. Also, P may not be invertible.)

(a) Show that W is a subspace of V.

(b) For $B \in W$, show that $B^k \in W$ for all positive integer k. (Hint: $\left(B^\top\right)^k = \left(B^k\right)^\top$.)

(c) For $B \in W$, if B is invertible, show that $B^{-1} \in W$.

(d) Let $n = 2$ and $P = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$. Find a basis for W.

Solution.

(a) Let $B_1, B_2 \in W$ and $\alpha_1, \alpha_2 \in \mathbb{F}$. Then

$$\begin{aligned}((\alpha_1 B_1 + \alpha_2 B_2)P)^\top &= (\alpha_1 B_1 P)^\top + (\alpha_2 B_2 P)^\top \\ &= \alpha_1 P^\top B_1^\top + \alpha_2 P^\top B_2^\top \\ &= \alpha_1 B_1 P + \alpha_2 B_2 P \\ &= (\alpha_1 B_1 + \alpha_2 B_2)P.\end{aligned}$$

Hence, W is a subspace of V.

(b) Let $k \in \mathbb{Z}^+$ and $B \in W$. Then

$$\begin{aligned}\left(B^k P\right)^\top &= P^\top \left(B^k\right)^\top \\ &= P^\top \left(B^\top\right)^k \\ &= P \left(B^\top\right)^k \\ &= P \underbrace{B^\top \cdots B^\top}_{k \text{ times}} \\ &= BP \underbrace{B^\top \cdots B^\top}_{k-1 \text{ times}} \\ &= \ldots \\ &= B^k P\end{aligned}$$

where we used the associativity of multiplication for matrices. Therefore, $B^k \in W$.

(c) Let $B \in W$ and suppose that B is invertible. Then $BP = PB^\top$ which implies that $B^{-1}P = P\left(B^\top\right)^{-1} = P\left(B^{-1}\right)^\top = \left(B^{-1}P\right)^\top$. Hence, $B^{-1} \in W$.

(d) Let $B \in W$ and write

$$B = \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix}.$$

Then

$$\begin{pmatrix} b_{12} & b_{11} \\ b_{22} & b_{21} \end{pmatrix} = \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} = BP = (BP)^\top = \left(\begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \right)^\top = \begin{pmatrix} b_{12} & b_{22} \\ b_{11} & b_{21} \end{pmatrix}.$$

This implies that $b_{11} = b_{22}$ and we can write

$$B = b_{12}\begin{pmatrix}1 & 0\\ 0 & 0\end{pmatrix} + b_{11}\begin{pmatrix}0 & 1\\ 1 & 0\end{pmatrix} + b_{21}\begin{pmatrix}0 & 0\\ 0 & 1\end{pmatrix}.$$

Therefore, a basis for W is

$$\left\{\begin{pmatrix}1 & 0\\ 0 & 0\end{pmatrix}, \begin{pmatrix}0 & 1\\ 1 & 0\end{pmatrix}, \begin{pmatrix}0 & 0\\ 0 & 1\end{pmatrix}\right\}.$$

$\square$

Question 4. Let $\mathcal{P}_2(\mathbb{R})$ be equipped with an inner product such that

$$\langle p(x), q(x)\rangle = \frac{1}{2}\int_{-1}^{1} p(t)q(t)\,dt$$

for $p(x), q(x) \in \mathcal{P}_2(\mathbb{R})$ and let T be a linear operator on $\mathcal{P}_2(\mathbb{R})$ such that

$$T(p(x)) = \frac{dp(x)}{dx}$$

for $p(x) \in \mathcal{P}_2(\mathbb{R})$.

(a) Let $B = \left\{1, \sqrt{3}x, \frac{\sqrt{5}}{2}\left(3x^2-1\right)\right\}$ which is an orthonormal basis for $\mathcal{P}_2(\mathbb{R})$. Find $[T]_B$ and $[T^*]_B$.

(b) For $a+bx+cx^2 \in \mathcal{P}_2(\mathbb{R})$, write down a formula for $T^*\left(a+bx+cx^2\right)$.

Solution.

(a) We have

$$\begin{aligned}
T(1) &= \frac{d}{dx}(1) = 0\\
T\left(\sqrt{3}x\right) &= \frac{d}{dx}\left(\sqrt{3}x\right) = \sqrt{3}\\
T\left(\frac{\sqrt{5}}{2}\left(3x^2-1\right)\right) &= \frac{d}{dx}\left(\frac{\sqrt{5}}{2}\left(3x^2-1\right)\right) = 3\sqrt{5}x = \frac{3\sqrt{5}}{\sqrt{3}}\left(\sqrt{3}x\right)
\end{aligned}$$

so that

$$\begin{aligned}
[T]_B &= \begin{pmatrix}0 & \sqrt{3} & 0\\ 0 & 0 & \frac{3\sqrt{5}}{\sqrt{3}}\\ 0 & 0 & 0\end{pmatrix}\\
[T^*]_B = ([T]_B)^* &= \begin{pmatrix}0 & 0 & 0\\ \sqrt{3} & 0 & 0\\ 0 & \frac{3\sqrt{5}}{\sqrt{3}} & 0\end{pmatrix}.
\end{aligned}$$

(b) We have

$$
\begin{aligned}
T^*\left(a+bx+cx^2\right) &= aT^*(1)+bT^*(x)+cT^*\left(x^2\right) \\
&= a\left(\sqrt{3}\cdot\sqrt{3}x\right)+b\left(\frac{3\sqrt{5}}{\sqrt{3}}\cdot\frac{\sqrt{5}}{2}\left(3x^2-1\right)\right) \\
&= 3ax+\sqrt{15}b\left(\frac{\sqrt{5}}{2}\left(3x^2-1\right)\right) \\
&= -\frac{5\sqrt{3}}{2}b+3ax+\frac{15\sqrt{3}}{2}bx^2.
\end{aligned}
$$

$\square$

Question 5. Let A be an $m\times n$ matrix over a field $\mathbb{F}$. Define

$$W_1=\left\{u\in\mathbb{F}^n : Au^\top=0^\top\right\} \text{ and } W_2=\{vA : v\in\mathbb{F}^m\}.$$

(In here, the vectors in $\mathbb{F}^m$ and $\mathbb{F}^n$ are written as row vectors.)

(a) Suppose $\mathbb{F}=\mathbb{R}$.
- **(i)** Prove that $\mathbb{R}^n=W_1\oplus W_2$.
  (Hint: There are many ways to do this question. One possible way is to show $W_1=W_2^\perp$ by using the usual inner product on $\mathbb{R}^n$.)
- **(ii)** Prove that $\left\{u\in\mathbb{R}^n : A^\top Au^\top=0^\top\right\}=W_1$.

(b) Suppose that $\mathbb{F}=\mathbb{F}_2$, the field of 2 elements.
- **(i)** Is $\mathbb{F}_2^n=W_1+W_2$? Justify your answer.
- **(ii)** Is $\left\{u\in\mathbb{F}_2^n : A^\top Au^\top=0^\top\right\}=W_1$? Justify your answer.

Solution.

(a) **(i)** We first show that $W_1=W_2^\perp$ under the usual inner product $\langle,\rangle$ of $\mathbb{R}^n$. Let $w_i\in W_i$ for $i=1,2$. Then $\langle w_1,w_2\rangle=\langle w_1,v_2A\rangle$ for some $v_2\in\mathbb{F}^m$.
- **(ii)** Is $\left\{vA^\top A : v\in\mathbb{R}^n\right\}=W_2$? Justify your answer.
- **(iii)** Is $\left\{u\in\mathbb{R}^m : AA^\top u^\top=0^\top\right\}=W_1$? Justify your answer.

$$\langle w_1,w_2\rangle=\langle w_1,v_2A\rangle=v_2Aw_1^\top=0$$

so that $W_1=W_2^\perp$.

Now, let $u \in \mathbb{R}^n$. Since W_1 is a subspace of $\mathbb{R}^n$, which is finite dimensional, then we may let $B_2 = \{v_1, \ldots, v_s\}$ be an orthonormal basis for W_2 (by the Gram-Schmidt process). Write

$$u = \sum_{i=1}^{s} \langle u, v_i \rangle v_i + \left(u - \sum_{i=1}^{s} \langle u, v_i \rangle v_i \right).$$

It suffices to show that $u - \sum_{i=1}^{s} \langle u, v_i \rangle v_i \in W_1 = W_2^{\perp}$. We have

$$\left\langle u - \sum_{i=1}^{s} \langle u, v_i \rangle v_i, v_j \right\rangle = \langle u, v_j \rangle - \sum_{i=1}^{s} \langle u, v_i \rangle \langle v_i, v_j \rangle = 0$$

since $\{v_1, \ldots, v_s\}$ forms an orthonormal basis for W_2 by definition. Hence, $u \in W_1 + W_2$. It is clear that $\mathbb{R}^n \supseteq W_1 + W_2$. Since $W_1 = W_2^{\perp}$, then $W_1 \cap W_2 = \{0\}$. Therefore, $\mathbb{R}^n = W_1 \oplus W_2$.

(iv) Let $v \in \left\{u \in \mathbb{R}^n : A^\top A u^\top = 0^\top\right\}$. Then $A^\top A v^\top = 0^\top$. This implies that $vA^\top A v^\top = 0$, so that $\left(Av^\top\right)^\top \left(Av^\top\right) = 0$ in $\mathbb{R}$. Hence, $Av^\top = 0$. Therefore, $v \in W_1$.
On the other hand, let $u \in W_1$. Then $Au^\top = 0^\top$ and so $A^\top A u^\top = 0^\top$. Therefore, $u \in \left\{u \in \mathbb{R}^n : A^\top A u^\top = 0^\top\right\}$.
Hence, $\left\{u \in \mathbb{R}^n : A^\top A u^\top = 0^\top\right\} = W_1$.

(v) Yes. We show that $\mathrm{Row}\left(A^\top A\right) = \mathrm{Row}(A)$. Let $x \in \mathrm{Row}\left(A^\top A\right)$. Then $x = vA^\top A$ for some $v \in \mathbb{R}^n$. Since $vA^\top \in \mathbb{R}^m$, then $vA^\top A \in \mathrm{Row}(A)$, so that $x \in \mathrm{Row}(A)$. This implies that $\mathrm{Row}\left(A^\top A\right) \subseteq \mathrm{Row}(A)$. Since $\mathrm{rank}\left(A^\top\right) = \mathrm{rank}(A)$, then $\mathrm{Row}\left(A^\top A\right) = \mathrm{Row}(A) = W_2$. Therefore, $\left\{vA^\top A : v \in \mathbb{R}^n\right\} = W_2$.

(vi) No. This is because the left hand side is a subset of $\mathbb{R}^m$ while W_1 is a subset of $\mathbb{R}^n$.

(b) **(i)** No. Let $A = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$. There are only four vectors in $\mathbb{F}_2$, namely,

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

Note that

$$\begin{pmatrix} 1 & 0 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 1 \end{pmatrix} = \begin{pmatrix} 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$$

$$\begin{pmatrix} 1 & 1 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} = \begin{pmatrix} 0 & 0 \end{pmatrix}$$

so that $W_2 = \left\{ \begin{pmatrix} 0 & 0 \end{pmatrix}, \begin{pmatrix} 1 & 1 \end{pmatrix} \right\}$. On the other hand,

$$\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$
$$\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \end{pmatrix} \neq \begin{pmatrix} 0 \\ 0 \end{pmatrix} \neq \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

so that $W_1 = W_2 \subsetneq \mathbb{F}_2^2$. Therefore, $\mathbb{F}_2^2 \neq W_1 + W_2$.

(ii) No. Let $A = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$, so that $A^\top A = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$. This implies that

$$\left\{ u \in \mathbb{F}_2^n : A^\top A u^\top = 0^\top \right\} = \mathbb{F}_2^2.$$

But $W_1 = \left\{ \begin{pmatrix} 0 & 0 \end{pmatrix}, \begin{pmatrix} 1 & 1 \end{pmatrix} \right\} \subsetneq \mathbb{F}_2^2$. □

Question 6. Let V be a finite dimensional (complex) vector space with an inner product $\langle , \rangle$. and let T be an invertible linear operator on V.

(a) Show that $T^* \circ T$ is unitarily diagonalizable and all its eigenvalues are nonzero real positive numbers.

(b) For $u, v \in V$, define $[u, v] = \langle T(u), T(v) \rangle$.

(i) Show that $[,]$ is also an inner product on V.

(ii) Suppose that for all $u, v \in V$,

$$\frac{[u,v]}{\sqrt{[u,u][v,v]}} = \frac{\langle u,v \rangle}{\sqrt{\langle u,u \rangle \langle v,v \rangle}}.$$

(If V is over $\mathbb{R}$, this means that the angle between u and v computed using the new inner product is the same as the angle computed using the original inner product.)

Prove that T is a scalar multiple of a unitary operator on V.

(Hint: By (a), there exists an orthonormal basis $B = \{v_1, v_2, \ldots, v_n\}$ such that for each i, $(T^* \circ T)(v_i) = \lambda_i v_i$ where λ_i is a nonzero real positive number.)

Solution.

(a) Let $S = T^* \circ T$. Then $S^* = (T^* \circ T)^* = T^* \circ T$, so that S is self-adjoint. Now,

$$S \circ S^* = S^* \circ S$$

since S is self-adjoint, and so S is normal. Therefore, S is unitarily diagonalizable.

Next, let μ be an eigenvalue of S with corresponding eigenvector $x \in V$. Then

$$\overline{\mu}\langle x,x\rangle = \langle x,\mu x\rangle = \langle x,S(x)\rangle = \langle S^*(x),x\rangle = \langle S(x),x\rangle = \langle \mu x,x\rangle = \mu\langle x,x\rangle$$

so that $\overline{\mu} = \mu$. Hence, all its eigenvalues are real numbers. It remains to show that $\mu > 0$. We have

$$0 < \langle T(x),T(x)\rangle = \langle x,T^* \circ T(x)\rangle = \mu\langle x,x\rangle$$

so that $\mu > 0$ as $\langle x,x\rangle > 0$.

(b) **(i)** We verify the properties of an inner product:

- For all $x_j, y_j, x, y \in V$ and $a_i, b_i \in \mathbb{C}$, we have

$$\begin{aligned}
[a_1x_1 + a_2x_2, y] &= \langle T(a_1x_1 + a_2x_2), T(y)\rangle \\
&= \langle a_1T(x_1) + a_2T(x_2), T(y)\rangle \\
&= a_1\langle T(x_1), T(y)\rangle + a_2\langle T(x_2), T(y)\rangle \\
&= a_1[x_1, y] + a_2[x_2, y] \\
[x, b_1y_1 + b_2y_2] &= \langle T(x), T(b_1y_1 + b_2y_2)\rangle \\
&= \langle T(x), b_1T(y_1) + b_2T(y_2)\rangle \\
&= \overline{b_1}\langle T(x), T(y_1)\rangle + \overline{b_2}\langle T(x), T(y_2)\rangle \\
&= \overline{b_1}[x, y_1] + \overline{b_2}[x, y_2].
\end{aligned}$$

- Let $x, y \in V$. Then we have

$$[y,x] = \langle T(y),T(x)\rangle = \overline{\langle T(x),T(y)\rangle} = \overline{[x,y]}.$$

- For all $0 \neq x \in V$,

$$[x,x] = \langle T(x),T(x)\rangle > 0.$$

Therefore, $[,]$ is an inner product.

(ii) By (a), there exists an orthonormal basis $B = \{v_1, v_2, \ldots, v_n\}$ such that for each i, $(T^* \circ T)(v_i) = \lambda_i v_i$ where λ_i is a nonzero real positive number. Observe that for any $i \neq j$, we have

$$\begin{aligned}\frac{1}{\sqrt{\langle v_i + v_j, v_i + v_j \rangle}} &= \frac{\sqrt{\langle v_i, v_i \rangle}}{\sqrt{\langle v_i + v_j, v_i + v_j \rangle}} \\ &= \frac{\langle v_i, v_i + v_j \rangle}{\sqrt{\langle v_i, v_i \rangle}\sqrt{\langle v_i + v_j, v_i + v_j \rangle}} \\ &= \frac{[v_i, v_i + v_j]}{\sqrt{[v_i, v_i]}\sqrt{[v_i + v_j, v_i + v_j]}} \\ &= \frac{\sqrt{[v_i, v_i]}}{\sqrt{[v_i + v_j, v_i + v_j]}}.\end{aligned}$$

This implies that (when swapping i and j),

$$\frac{\sqrt{[v_i, v_i]}}{\sqrt{[v_i + v_j, v_i + v_j]}} = \frac{1}{\sqrt{\langle v_i + v_j, v_i + v_j \rangle}} = \frac{\sqrt{[v_j, v_j]}}{\sqrt{[v_i + v_j, v_i + v_j]}}$$

so that $[v_i, v_i] = [v_j, v_j]$. Let $c = \sqrt{[v_i, v_i]}$. Note that

$$c = \sqrt{[v_i, v_i]} = \sqrt{\langle T(v_i), T(v_i) \rangle} = \sqrt{\langle v_i, T^*T(v_i) \rangle} = \sqrt{\lambda_i} \in \mathbb{R}.$$

We will show that $P = \frac{1}{c}T$ is a unitary operator. For any $v_i \in B$, we have

$$P^*P(v_i) = \frac{1}{\bar{c}c}T^*T(v_i) = \frac{\lambda_i v_i}{\lambda_i} = v_i.$$

Hence, $T = cP$ is a scalar multiple of a unitary linear operator. □

Question 7. Let T be a linear operator on a finite dimensional vector space V over a field $\mathbb{F}$.

(a) Let $p(x)$ and $q(x)$ be polynomials over $\mathbb{F}$ such that $p(x)$ and $q(x)$ do not have common factors. For any nonzero $v \in \operatorname{Ker}(p(T))$, show that $q(T)(v) \neq 0$.

(Hint: You can use the fact that there exists polynomials $a(x)$ and $b(x)$ over $\mathbb{F}$ such that $a(x)p(x) + b(x)q(x) = 1$.)

(b) Let U be a T-cyclic subspace of V, i.e. $U = \operatorname{span}\{u, T(u), T^2(u), \ldots\}$ for some nonzero $u \in V$. Prove that $m_{T|_U}(x) = c_{T|_U}(x)$.

(Hint: $\deg\left(c_{T|_U}(x)\right) = \dim(U)$ which is the smallest positive integer k such that $T^k(u)$ is a linear combination of $u, T(u), \ldots, T^{k-1}(u)$.)

(c) Suppose

$$m_T(x) = c_T(x) = (x-\lambda_1)^{r_1}(x-\lambda_2)^{r_2}\cdots(x-\lambda_k)^{r_k}$$

where $\lambda_1, \lambda_2, \dots, \lambda_k$ are distinct eigenvalues of T. For $i = 1, 2, \dots, k$, choose $v_i \in \operatorname{Ker}((T-\lambda_i I_V)^{r_i})$ such that $v_i \notin \operatorname{Ker}\left((T-\lambda_i I_V)^{r_i-1}\right)$.

- **(i)** Let $s(x) = \frac{c_T(x)}{x-\lambda_i}$ for some $i \in \{1, 2, \dots, k\}$. Show that $s(T)\left(v_j\right) = 0$ for $j \neq i$ and $s(T)(v_i) \neq 0$.
- **(ii)** Define $w = v_1 + v_2 + \cdots + v_k$ and $W = \operatorname{span}\left\{w, T(w), T^2(w), \dots\right\}$. Prove that $W = V$.

  (Hint: $\deg(c_T(x)) = \dim(V)$ and $\deg\left(c_{T|_W}(x)\right) = \dim(W)$.)

Solution.

(a) Suppose that $q(T)(v) = 0$. Since $\gcd(p(x), q(x)) = 1$, then there exists $a(x)$ and $b(x)$ over $\mathbb{F}$ such that $a(x)p(x) + b(x)q(x) = 1$. Now,

$$v = I_V(v) = (a(T)p(T) + b(T)q(T))(v) = a(T)p(T)(v) + b(T)q(T)(v) = 0$$

which is a contradiction. Therefore, $q(T)(v) \neq 0$.

(b) Let $\deg\left(c_{T|_U}(x)\right) = \dim(U) = k$. By definition of minimal polynomial, we have $m_{T|_U}(x) | c_{T|_U}(x)$. This implies that $\deg\left(m_{T|_U}(x)\right) \leq \deg\left(c_{T|_U}(x)\right)$. If $\ell = \deg\left(m_{T|_U}(x)\right) < \deg\left(c_{T|_U}(x)\right)$, then

$$T|_U^{\ell} + c_{\ell-1}T|_U^{\ell-1} + \cdots + c_1 T|_U^1 + c_0 I_U = 0_U$$

which contradicts the minimality of k. Therefore, $\deg\left(m_{T|_U}(x)\right) = \deg\left(c_{T|_U}(x)\right)$ and so $m_{T|_U}(x) = c_{T|_U}(x)$.

(c) **(i)** Claim 1: For any $i \neq j$, we have $(T-\lambda_i I_V)^{r_i} \circ \left(T-\lambda_j I_V\right)^{r_j} = \left(T-\lambda_j I_V\right)^{r_j} \circ (T-\lambda_i I_V)^{r_i}$.

Proof of claim 1: Note that for any $v \in V$,

$$(T-\lambda_i I_V)\left(T-\lambda_j I_V\right) = T^2 - \lambda_i T - \lambda_j T + \lambda_i\lambda_j I_V = \left(T-\lambda_j I_V\right)(T-\lambda_i I_V).$$

This implies that $(T-\lambda_i I_V)^{r_i} \circ \left(T-\lambda_j I_V\right)^{r_j} = \left(T-\lambda_j I_V\right)^{r_j} \circ (T-\lambda_i I_V)^{r_i}$ by induction on r_j, fixing r_i.

Let

$$q_i(T) = \prod_{\substack{j=1 \\ j\neq i}}^{k} \left(T-\lambda_j\right)^{r_j}$$

$$p_i(T) = (T-\lambda_i)^{r_i-1}$$

so that $p_i(T)$ and $q_i(T)$ are coprime. Let $j \neq i$. Then

$$\begin{aligned}
s(T)\left(v_j\right) &= (T-\lambda_1 I_V)^{r_1}(T-\lambda_2 I_V)^{r_2}\cdots(T-\lambda_k I_V)^{r_k} \\
&= (T-\lambda_1 I_V)^{r_1}\cdots\left(T-\lambda_{j-1}I_V\right)^{r_{j-1}}\left(T-\lambda_{j+1}I_V\right)^{r_{j+1}}\cdots(T-\lambda_k I_V)^{r_k}\left(T-\lambda_j I_V\right)^{r_j}\left(v_j\right) \\
&= 0 \\
s(T)(v_i) &= (T-\lambda_1)^{r_1}(T-\lambda_2)^{r_2}\cdots(T-\lambda_k)^{r_k} \\
&= (T-\lambda_1 I_V)^{r_1}\cdots(T-\lambda_{i-1}I_V)^{r_{i-1}}(T-\lambda_{i+1}I_V)^{r_{i+1}}\cdots(T-\lambda_k I_V)^{r_k}(T-\lambda_i I_V)^{r_i-1}(v_i) \\
&= q_i(T)\circ p_i(T)(v_i) \neq 0,
\end{aligned}$$

where we used

$$\left[a(T)\left(p_i(T)\right)^2 + b(T)q_i(T)p_i(T)\right](v_i) = p_i(T)(v_i) \neq 0.$$

(ii) Note that $c_{T|_W}(x) | c_T(x)$, so that $\deg\left(c_{T|_W}(x)\right) \leq \deg\left(c_T(x)\right)$. Write

$$c_{T|_W}(x) = \prod_{i=1}^{k}(x-\lambda_i)^{s_i}$$

where $s_i \leq r_i$ for all i. Suppose that there exists $i_1,\ldots,i_t \in \{s_1,\ldots,s_k\}$ such that $s_{i_j} < r_{i_j}$ for all $j \in \{1,\ldots,t\}$ and $s_{i_\ell} = r_{i_\ell}$ for all $\ell \notin \{1,\ldots,t\}$. Then we may write

$$\begin{aligned}
0 &= c_{T|_W}(w) \\
&= (T-\lambda_{i_1})^{s_{i_1}}\prod_{\substack{i=1\\ i\neq i_1}}^{k}(T-\lambda_i)^{s_i}(v_{i_1}) + \cdots + (T-\lambda_{i_t})^{s_{i_t}}\prod_{\substack{i=1\\ i\neq i_t}}^{k}(T-\lambda_i)^{s_i}(v_{i_t})
\end{aligned}$$

which implies that

$$(T-\lambda_{i_1})^{s_{i_1}}\prod_{\substack{i=1\\ i\neq i_1}}^{k}(T-\lambda_i)^{s_i}(v_{i_1}) \in \sum_{j=2}^{t}\operatorname{Ker}\left(\left(T-\lambda_{i_j}\right)^{r_{i_j}-s_{i_j}}\right) \subseteq \sum_{j=2}^{k}\operatorname{Ker}\left(\left(T-\lambda_j\right)^{r_j}\right).$$

But

$$(T-\lambda_{i_1})^{s_{i_1}}\prod_{\substack{i=1\\ i\neq i_1}}^{k}(T-\lambda_i)^{s_i}(v_{i_1}) \in \operatorname{Ker}\left((T-\lambda_1)^{r_1}\right)$$

so that

$$(T-\lambda_{i_1})^{s_{i_1}}\prod_{\substack{i=1\\ i\neq i_1}}^{k}(T-\lambda_i)^{s_i}(v_{i_1}) = 0.$$

This implies that $s(T)(v_i) = 0$, which is a contradiction. Therefore, we must have all $s_i = r_i$, and so $c_{T|_W}(x) = c_T(x)$. This means that $\dim(V) = \dim(W)$. Since W is also a subspace of V, then $W = V$. $\square$