Semester 1, 2015-16

1. (i) $\begin{pmatrix} 0 & 0 & 1 & 1 & 1 & 0 \\ -1 & 1 & 0 & 0 & 0 & 0 \\ -1 & 1 & 1 & 2 & 3 & 0 \\ 0 & 0 & 1 & 3 & 5 & 0 \end{pmatrix} \xrightarrow{R_1 \leftrightarrow R_2} \begin{pmatrix} -1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 1 & 0 \\ -1 & 1 & 1 & 2 & 3 & 0 \\ 0 & 0 & 1 & 3 & 5 & 0 \end{pmatrix} \xrightarrow{R_3 - R_1}$

$$\begin{pmatrix} -1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 1 & 0 \\ 0 & 0 & 1 & 2 & 3 & 0 \\ 0 & 0 & 1 & 3 & 5 & 0 \end{pmatrix} \xrightarrow[R_4 - R_2]{R_3 - R_2} \begin{pmatrix} -1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 & 2 & 0 \\ 0 & 0 & 0 & 2 & 4 & 0 \end{pmatrix} \xrightarrow{R_4 - 2R_3}$$

$$\begin{pmatrix} -1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 & 2 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix} \xrightarrow{-R_1} \begin{pmatrix} 1 & -1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & -1 & 0 \\ 0 & 0 & 0 & 1 & 2 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix} \xrightarrow{R_2 - R_3}$$

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

(ii) By Part (i), the general solution of $T(\boldsymbol{x}) = \boldsymbol{A}\boldsymbol{x} = \mathbf{0}$ is

$$\boldsymbol{x} = \begin{pmatrix} r \\ r \\ s \\ -2s \\ s \\ t \end{pmatrix} = r\begin{pmatrix} 1 \\ 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} + s\begin{pmatrix} 0 \\ 0 \\ 1 \\ -2 \\ 1 \\ 0 \end{pmatrix} + t\begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 1 \end{pmatrix} \quad \text{for } r, s, t \in \mathbb{R}.$$

So $\{ (1,1,0,0,0,0)^{\mathrm{T}},\ (0,0,1,-2,1,0)^{\mathrm{T}},\ (0,0,0,0,0,1)^{\mathrm{T}} \}$ is a basis for the kernel of T.

Since the range of T is equal to the column space of $\boldsymbol{A}$, by Part (i), $\{ (0,-1,-1,0)^{\mathrm{T}},\ (1,0,1,1)^{\mathrm{T}},\ (1,0,2,3)^{\mathrm{T}} \}$ is a basis for the range of T.

2. (a) $\left(\begin{array}{ccc|ccc} 1 & 3 & 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 & 1 & 0 \\ -1 & 0 & 1 & 0 & 0 & 1 \end{array}\right) \xrightarrow{R_3 + R_1} \left(\begin{array}{ccc|ccc} 1 & 3 & 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 & 1 & 0 \\ 0 & 3 & 2 & 1 & 0 & 1 \end{array}\right) \xrightarrow{R_3 - 3R_2}$

$$\left(\begin{array}{ccc|ccc} 1 & 3 & 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 & 1 & 0 \\ 0 & 0 & -1 & 1 & -3 & 1 \end{array}\right) \xrightarrow{-R_3} \left(\begin{array}{ccc|ccc} 1 & 3 & 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 & 1 & 0 \\ 0 & 0 & 1 & -1 & 3 & -1 \end{array}\right) \xrightarrow[R_2 - R_3]{R_1 - R_3}$$

$$\left(\begin{array}{ccc|ccc} 1 & 3 & 1 & 2 & -3 & 1 \\ 0 & 1 & 1 & 1 & -2 & 1 \\ 0 & 0 & 1 & -1 & 3 & -1 \end{array}\right) \xrightarrow{R_1 - 3R_2} \left(\begin{array}{ccc|ccc} 1 & 3 & 1 & -1 & 3 & -2 \\ 0 & 1 & 1 & 1 & -2 & 1 \\ 0 & 0 & 1 & -1 & 3 & -1 \end{array}\right)$$

Thus the transition matrix from T to S is $\boldsymbol{P}^{-1} = \begin{pmatrix} -1 & 3 & -2 \\ 1 & -2 & 1 \\ -1 & 3 & -1 \end{pmatrix}$.

(b) Since the transition matrix from T to S is $\big([\boldsymbol{v_1}]_S \;\; [\boldsymbol{v_2}]_S \;\; [\boldsymbol{v_3}]_S\big)$, by Part (a),

$$[\boldsymbol{v_1}]_S = \begin{pmatrix} -1 \\ 1 \\ -1 \end{pmatrix}, \quad [\boldsymbol{v_2}]_S = \begin{pmatrix} 3 \\ -2 \\ 3 \end{pmatrix}, \quad [\boldsymbol{v_3}]_S = \begin{pmatrix} -2 \\ 1 \\ -1 \end{pmatrix}.$$

So

$$\begin{aligned} \boldsymbol{v_1} &= -\boldsymbol{u_1} + \boldsymbol{u_2} - \boldsymbol{u_3} \\ &= -(1,1,1) + (0,1,1) - (0,0,1) \\ &= (-1,\ 0,\ -1), \end{aligned}$$

$$\begin{aligned} \boldsymbol{v_2} &= 3\boldsymbol{u_1} - 2\boldsymbol{u_2} + 3\boldsymbol{u_3} \\ &= 3(1,1,1) - 2(0,1,1) + 3(0,0,1) \\ &= (3,\ 1,\ 4), \end{aligned}$$

$$\begin{aligned} \boldsymbol{v_3} &= -2\boldsymbol{u_1} + \boldsymbol{u_2} - \boldsymbol{u_3} \\ &= -2(1,1,1) + (0,1,1) - (0,0,1) \\ &= (-2,\ -1,\ -2). \end{aligned}$$

3. (i) The characteristic polynomial of $\boldsymbol{A}$ is

$$\det(\lambda \boldsymbol{I} - \boldsymbol{A}) = \begin{vmatrix} \lambda-2 & 0 & -1 \\ -1 & \lambda-1 & -a \\ 0 & 0 & \lambda-1 \end{vmatrix} = (\lambda-1)\begin{vmatrix} \lambda-2 & 0 \\ -1 & \lambda-1 \end{vmatrix} = (\lambda-1)^2(\lambda-2).$$

So the eigenvalues of $\boldsymbol{A}$ are 1 and 2.

(ii) $\lambda = 1$:

- When $a = 1$, $\quad \boldsymbol{I} - \boldsymbol{A} = \begin{pmatrix} -1 & 0 & -1 \\ -1 & 0 & -1 \\ 0 & 0 & 0 \end{pmatrix} \xrightarrow[\text{Elimination}]{\text{Gauss-Jordan}} \begin{pmatrix} 1 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$.

  Then the general solution to $(\boldsymbol{I} - \boldsymbol{A})\boldsymbol{x} = \boldsymbol{0}$ is

$$\boldsymbol{x} = \begin{pmatrix} -t \\ s \\ t \end{pmatrix} = s\begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} + t\begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix} \quad \text{for } s, t \in \mathbb{R}.$$

  So $\{\, (0,1,0)^{\mathrm{T}},\ (-1,0,1)^{\mathrm{T}} \,\}$ is a basis for E_1.

- When $a \neq 1$, $\quad \boldsymbol{I} - \boldsymbol{A} = \begin{pmatrix} -1 & 0 & -1 \\ -1 & 0 & -a \\ 0 & 0 & 0 \end{pmatrix} \underset{\text{Elimination}}{\overset{\text{Gauss-Jordan}}{\longrightarrow}} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}$.

  Then the general solution to $(\boldsymbol{I} - \boldsymbol{A})\boldsymbol{x} = \boldsymbol{0}$ is

$$\boldsymbol{x} = \begin{pmatrix} 0 \\ s \\ 0 \end{pmatrix} = s \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \quad \text{for } s \in \mathbb{R}.$$

  So $\{ (0,1,0)^{\mathrm{T}} \}$ is a basis for E_1.

$\lambda = 2$**:** $\quad 2\boldsymbol{I} - \boldsymbol{A} = \begin{pmatrix} 0 & 0 & -1 \\ -1 & 1 & -a \\ 0 & 0 & 1 \end{pmatrix} \underset{\text{Elimination}}{\overset{\text{Gauss-Jordan}}{\longrightarrow}} \begin{pmatrix} 1 & -1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}$.

Then the general solution to $(2\boldsymbol{I} - \boldsymbol{A})\boldsymbol{x} = \boldsymbol{0}$ is

$$\boldsymbol{x} = \begin{pmatrix} t \\ t \\ 0 \end{pmatrix} = t \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} \quad \text{for } t \in \mathbb{R}.$$

So $\{ (1,1,0)^{\mathrm{T}} \}$ is a basis for E_2.

(iii) If $a \neq 1$, then we only have two linearly independent eigenvectors and hence $\boldsymbol{A}$ is not diagonalizable.

If $a = 1$, then we have three linearly independent eigenvectors and hence $\boldsymbol{A}$ is diagonalizable.

(iv) Let $\boldsymbol{P} = \begin{pmatrix} 0 & -1 & 1 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}$ and $\boldsymbol{D} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{pmatrix}$.

When $a = 1$, we have $\boldsymbol{D} = \boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}$.

4. (a) (i) Apply Gram-Schmidt Process to $\{\boldsymbol{u_1}, \boldsymbol{u_2}, \boldsymbol{u_3}\}$:

Let

$$\begin{aligned}
\boldsymbol{v_1} &= (1,1,0,0), \\
\boldsymbol{v_2} &= (0,2,1,1) - \tfrac{(0,2,1,1)\cdot(1,1,0,0)}{(1,1,0,0)\cdot(1,1,0,0)}(1,1,0,0) \\
&= (0,2,1,1) - (1,1,0,0) \\
&= (-1,1,1,1), \\
\boldsymbol{v_3} &= (1,1,3,1) - \tfrac{(1,1,3,1)\cdot(1,1,0,0)}{(1,1,0,0)\cdot(1,1,0,0)}(1,1,0,0) - \tfrac{(1,1,3,1)\cdot(-1,1,1,1)}{(-1,1,1,1)\cdot(-1,1,1,1)}(-1,1,1,1) \\
&= (1,1,3,1) - (1,1,0,0) - (-1,1,1,1) \\
&= (1,-1,2,0).
\end{aligned}$$

Then

$$S = \left\{ \tfrac{1}{||\boldsymbol{v_1}||}\boldsymbol{v_1},\ \tfrac{1}{||\boldsymbol{v_2}||}\boldsymbol{v_2},\ \tfrac{1}{||\boldsymbol{v_3}||}\boldsymbol{v_3} \right\}$$

$$= \left\{ \left(\tfrac{1}{\sqrt{2}},\ \tfrac{1}{\sqrt{2}},\ 0,\ 0\right),\ \left(-\tfrac{1}{2},\ \tfrac{1}{2},\ \tfrac{1}{2},\ \tfrac{1}{2}\right),\ \left(\tfrac{1}{\sqrt{6}},\ -\tfrac{1}{\sqrt{6}},\ \tfrac{2}{\sqrt{6}},\ 0\right) \right\}$$

is an orthonormal basis for V.

(ii) **Method 1:**

The projection of $\boldsymbol{w}$ onto V

$$= \left[(1,0,0,1)\cdot\left(\tfrac{1}{\sqrt{2}},\ \tfrac{1}{\sqrt{2}},\ 0,\ 0\right)\right]\left(\tfrac{1}{\sqrt{2}},\ \tfrac{1}{\sqrt{2}},\ 0,\ 0\right)$$
$$+\left[(1,0,0,1)\cdot\left(-\tfrac{1}{2},\ \tfrac{1}{2},\ \tfrac{1}{2},\ \tfrac{1}{2}\right)\right]\left(-\tfrac{1}{2},\ \tfrac{1}{2},\ \tfrac{1}{2},\ \tfrac{1}{2}\right)$$
$$+\left[(1,0,0,1)\cdot\left(\tfrac{1}{\sqrt{6}},\ -\tfrac{1}{\sqrt{6}},\ \tfrac{2}{\sqrt{6}},\ 0\right)\right]\left(\tfrac{1}{\sqrt{6}},\ -\tfrac{1}{\sqrt{6}},\ \tfrac{2}{\sqrt{6}},\ 0\right)$$
$$= \tfrac{1}{\sqrt{2}}\left(\tfrac{1}{\sqrt{2}},\ \tfrac{1}{\sqrt{2}},\ 0,\ 0\right) + 0\left(-\tfrac{1}{2},\ \tfrac{1}{2},\ \tfrac{1}{2},\ \tfrac{1}{2}\right) + \tfrac{1}{\sqrt{6}}\left(\tfrac{1}{\sqrt{6}},\ -\tfrac{1}{\sqrt{6}},\ \tfrac{2}{\sqrt{6}},\ 0\right)$$
$$= \left(\tfrac{2}{3},\ \tfrac{1}{3},\ \tfrac{1}{3},\ 0\right).$$

Method 2:

Let $\boldsymbol{A} = \begin{pmatrix} 1 & 0 & 1 \\ 1 & 2 & 1 \\ 0 & 1 & 3 \\ 0 & 1 & 1 \end{pmatrix}$.

Solving $\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x} = \boldsymbol{A}^{\mathrm{T}}\boldsymbol{w}$, we obtain $\boldsymbol{x} = \begin{pmatrix} \frac{1}{2} \\ -\frac{1}{6} \\ \frac{1}{6} \end{pmatrix}$ which is the least square solution to $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{w}$.

As $\begin{pmatrix} 1 & 0 & 1 \\ 1 & 2 & 1 \\ 0 & 1 & 3 \\ 0 & 1 & 1 \end{pmatrix}\begin{pmatrix} \frac{1}{2} \\ -\frac{1}{6} \\ \frac{1}{6} \end{pmatrix} = \begin{pmatrix} \frac{2}{3} \\ \frac{1}{3} \\ \frac{1}{3} \\ 0 \end{pmatrix}$, the projection of w onto V is $\left(\frac{2}{3},\ \frac{1}{3},\ \frac{1}{3},\ 0\right)$.

(b) **Method 1:**

Let $\{\boldsymbol{w_1}, \boldsymbol{w_2}, \ldots, \boldsymbol{w_k}\}$ be a basis for W.

$$\boldsymbol{u} \in W^{\perp} \quad\Leftrightarrow\quad \begin{cases} \boldsymbol{w_1}\cdot\boldsymbol{u} = 0 \\ \boldsymbol{w_2}\cdot\boldsymbol{u} = 0 \\ \quad\vdots \\ \boldsymbol{w_k}\cdot\boldsymbol{u} = 0 \end{cases} \quad\Leftrightarrow\quad \begin{cases} \boldsymbol{w_1}\boldsymbol{u}^{\mathrm{T}} = 0 \\ \boldsymbol{w_2}\boldsymbol{u}^{\mathrm{T}} = 0 \\ \quad\vdots \\ \boldsymbol{w_k}\boldsymbol{u}^{\mathrm{T}} = 0 \end{cases} \quad\Leftrightarrow\quad \begin{pmatrix} \boldsymbol{w_1} \\ \boldsymbol{w_2} \\ \vdots \\ \boldsymbol{w_k} \end{pmatrix}\boldsymbol{u}^{\mathrm{T}} = \boldsymbol{0}$$

Let $\boldsymbol{A} = \begin{pmatrix} \boldsymbol{w_1} \\ \boldsymbol{w_2} \\ \vdots \\ \boldsymbol{w_k} \end{pmatrix}$ which is a $k \times n$ matrix.

Then W is the row space of $\boldsymbol{A}$ and $W^{\perp}$ is the nullspace of $\boldsymbol{A}$.

By the Dimension Theorem for Matrices,

$$\begin{aligned}
\dim(W) + \dim(W^{\perp}) &= \dim(\text{the row space of } \boldsymbol{A}) + \dim(\text{the nullspace of } \boldsymbol{A}) \\
&= \operatorname{rank}(\boldsymbol{A}) + \operatorname{nullity}(\boldsymbol{A}) \\
&= \text{the number of columns of } \boldsymbol{A} \\
&= n.
\end{aligned}$$

Method 2:

Let $\{\boldsymbol{w_1}, \boldsymbol{w_2}, \ldots, \boldsymbol{w_k}\}$ be a basis for W and $\{\boldsymbol{v_1}, \boldsymbol{v_2}, \ldots, \boldsymbol{v_m}\}$ a basis for $W^{\perp}$, where $k = \dim(W)$ and $m = \dim(W^{\perp})$.

We first claim that $S = \{\boldsymbol{w_1}, \boldsymbol{w_2}, \ldots, \boldsymbol{w_k}, \boldsymbol{v_1}, \boldsymbol{v_2}, \ldots, \boldsymbol{v_m}\}$ is a basis for $\mathbb{R}^n$.

Proof.

- For any $\boldsymbol{u} \in \mathbb{R}^n$, we can write $\boldsymbol{u} = \boldsymbol{p} + \boldsymbol{n}$ where $\boldsymbol{p} \in W$ and $\boldsymbol{n} \in W^{\perp}$. As $\boldsymbol{p} \in \operatorname{span}\{\boldsymbol{w_1}, \boldsymbol{w_2}, \ldots, \boldsymbol{w_k}\}$ and $\boldsymbol{n} \in \operatorname{span}\{\boldsymbol{v_1}, \boldsymbol{v_2}, \ldots, \boldsymbol{v_m}\}$,

  $$\boldsymbol{u} = \boldsymbol{p} + \boldsymbol{n} \in \operatorname{span}\{\boldsymbol{w_1}, \boldsymbol{w_2}, \ldots, \boldsymbol{w_k}, \boldsymbol{v_1}, \boldsymbol{v_2}, \ldots, \boldsymbol{v_m}\} = \operatorname{span}(S).$$

  So $\operatorname{span}(S) = \mathbb{R}^n$.

- Consider the vector equation

  $$c_1\boldsymbol{w_1} + c_2\boldsymbol{w_2} + \cdots + c_k\boldsymbol{w_k} + d_1\boldsymbol{v_1} + d_2\boldsymbol{v_2} + \cdots + d_m\boldsymbol{v_m} = \boldsymbol{0}. \qquad (\sharp)$$

  Let $\boldsymbol{p'} = c_1\boldsymbol{w_1} + c_2\boldsymbol{w_2} + \cdots + c_k\boldsymbol{w_k}$ and $\boldsymbol{n'} = d_1\boldsymbol{v_1} + d_2\boldsymbol{v_2} + \cdots + d_m\boldsymbol{v_m}$.

  Since $\boldsymbol{p'} \in W$ and $\boldsymbol{n'} \in W^{\perp}$, $\boldsymbol{p'} \cdot \boldsymbol{n'} = 0$.

  By $(\sharp)$, we have

  $$\begin{aligned}
  & \boldsymbol{p'} + \boldsymbol{n'} = 0 \\
  \Rightarrow\ & \boldsymbol{p'} \cdot \boldsymbol{n'} + \boldsymbol{n'} \cdot \boldsymbol{n'} = \boldsymbol{0} \cdot \boldsymbol{n'} \\
  \Rightarrow\ & \boldsymbol{n'} \cdot \boldsymbol{n'} = 0. \\
  \Rightarrow\ & \boldsymbol{n'} = \boldsymbol{0} \\
  \Rightarrow\ & d_1\boldsymbol{v_1} + d_2\boldsymbol{v_2} + \cdots + d_m\boldsymbol{v_m} = \boldsymbol{0}. \qquad (\dagger)
  \end{aligned}$$

  As $\{\boldsymbol{v_1}, \boldsymbol{v_2}, \ldots, \boldsymbol{v_m}\}$ is linearly independent, the equation $(\dagger)$ has only the trivial solution, i.e. $d_1 = 0$, $d_2 = 0$, $\ldots$, $d_m = 0$.

  Substituting $(\dagger)$ into $(\sharp)$, we get

  $$c_1\boldsymbol{w_1} + c_2\boldsymbol{w_2} + \cdots + c_k\boldsymbol{w_k} = \boldsymbol{0}. \qquad (\ddagger)$$

  As $\{\boldsymbol{w_1}, \boldsymbol{w_2}, \ldots, \boldsymbol{w_k}\}$ is linearly independent, the equation $(\ddagger)$ has only the trivial solution, i.e. $c_1 = 0$, $c_2 = 0$, $\ldots$, $c_k = 0$.

  Thus the equation $(\sharp)$ has only the trivial solution and hence S is linearly independent.

From the arguments above, we have shown that S is a basis for $\mathbb{R}^n$.

Hence $\dim(W) + \dim(W^{\perp}) = k + m = |S| = \dim(\mathbb{R}^n) = n$.

5. (a) For example, $\boldsymbol{A} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$.

(b) Observe that $\boldsymbol{A}^{n+m} = \boldsymbol{A}^n \boldsymbol{A}^m = \boldsymbol{0}\boldsymbol{A}^m = \boldsymbol{0}$ for all integer $m \geq 0$.

Consider the vector equation

$$c_1\boldsymbol{v} + c_2\boldsymbol{A}\boldsymbol{v} + \cdots + c_n\boldsymbol{A}^{n-1}\boldsymbol{v} = \boldsymbol{0}. \tag{$*$}$$

Pre-multiplying $\boldsymbol{A}^{n-1}$ to both sides of $(*)$, we have

$$\begin{aligned} & \boldsymbol{A}^{n-1}(c_1\boldsymbol{v} + c_2\boldsymbol{A}\boldsymbol{v} + \cdots + c_n\boldsymbol{A}^{n-1}\boldsymbol{v}) = \boldsymbol{A}^{n-1}\boldsymbol{0} \\ \Rightarrow\quad & c_1\boldsymbol{A}^{n-1}\boldsymbol{v} + c_2\boldsymbol{A}^{n}\boldsymbol{v} + \cdots + c_n\boldsymbol{A}^{2n-2}\boldsymbol{v} = \boldsymbol{0} \\ \Rightarrow\quad & c_1\boldsymbol{A}^{n-1}\boldsymbol{v} = \boldsymbol{0}. \end{aligned}$$

Since $\boldsymbol{A}^{n-1}\boldsymbol{v}$ is a nonzero vector, $c_1 = 0$.

Suppose we have already shown that $c_1 = 0$, $c_2 = 0$, $\ldots$, $c_i = 0$ for some i with $1 \leq i < n$.
Pre-multiplying $\boldsymbol{A}^{n-i-1}$ to both sides of $(*)$, we have

$$\begin{aligned} & \boldsymbol{A}^{n-i-1}(c_{i+1}\boldsymbol{A}^{i}\boldsymbol{v} + c_{i+2}\boldsymbol{A}^{i+1}\boldsymbol{v} + \cdots + c_n\boldsymbol{A}^{n-1}\boldsymbol{v}) = \boldsymbol{A}^{n-i-1}\boldsymbol{0} \\ \Rightarrow\quad & c_{i+1}\boldsymbol{A}^{n-1}\boldsymbol{v} + c_{i+2}\boldsymbol{A}^{n}\boldsymbol{v} + \cdots + c_n\boldsymbol{A}^{2n-i-2}\boldsymbol{v} = \boldsymbol{0} \\ \Rightarrow\quad & c_{i+1}\boldsymbol{A}^{n-1}\boldsymbol{v} = \boldsymbol{0}. \end{aligned}$$

Since $\boldsymbol{A}^{n-1}\boldsymbol{v}$ is a nonzero vector, $c_{i+1} = 0$.

By mathematical induction, we have shown that $c_1 = 0$, $c_2 = 0$, $\ldots$, $c_n = 0$, i.e. $(*)$ has only the trivial solution.

So $\boldsymbol{v}$, $\boldsymbol{A}\boldsymbol{v}$, $\ldots$, $\boldsymbol{A}^{n-1}\boldsymbol{v}$ are linearly independent.

Since $\dim(\mathbb{R}^n) = n$, the n linearly independent vectors form a basis for $\mathbb{R}^n$.

(c) Note that

$$\begin{aligned} \boldsymbol{A}\boldsymbol{P} &= \boldsymbol{A}\begin{pmatrix}\boldsymbol{A}^{n-1}\boldsymbol{v} & \boldsymbol{A}^{n-2}\boldsymbol{v} & \cdots & \boldsymbol{A}\boldsymbol{v} & \boldsymbol{v}\end{pmatrix} \\ &= \begin{pmatrix}\boldsymbol{A}\boldsymbol{A}^{n-1}\boldsymbol{v} & \boldsymbol{A}\boldsymbol{A}^{n-2}\boldsymbol{v} & \cdots & \boldsymbol{A}\boldsymbol{A}\boldsymbol{v} & \boldsymbol{A}\boldsymbol{v}\end{pmatrix} \\ &= \begin{pmatrix}\boldsymbol{A}^{n}\boldsymbol{v} & \boldsymbol{A}^{n-1}\boldsymbol{v} & \cdots & \boldsymbol{A}^2\boldsymbol{v} & \boldsymbol{A}\boldsymbol{v}\end{pmatrix} \\ &= \begin{pmatrix}\boldsymbol{0} & \boldsymbol{A}^{n-1}\boldsymbol{v} & \cdots & \boldsymbol{A}^2\boldsymbol{v} & \boldsymbol{A}\boldsymbol{v}\end{pmatrix}. \end{aligned}$$

On the other hand,

$$\boldsymbol{P}\begin{pmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & 0 \\ \vdots & \vdots & & \ddots & 1 \\ 0 & 0 & \cdots & \cdots & 0 \end{pmatrix} = \boldsymbol{P}\begin{pmatrix} \boldsymbol{0} & \boldsymbol{e_1} & \cdots & \boldsymbol{e_{n-2}} & \boldsymbol{e_{n-1}} \end{pmatrix}$$
$$= \begin{pmatrix} \boldsymbol{P0} & \boldsymbol{Pe_1} & \cdots & \boldsymbol{Pe_{n-2}} & \boldsymbol{Pe_{n-1}} \end{pmatrix}$$
$$= \begin{pmatrix} \boldsymbol{0} & \boldsymbol{A}^{n-1}\boldsymbol{v} & \cdots & \boldsymbol{A}^2\boldsymbol{v} & \boldsymbol{Av} \end{pmatrix}.$$

So $\boldsymbol{AP} = \boldsymbol{P}\begin{pmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & 0 \\ \vdots & \vdots & & \ddots & 1 \\ 0 & 0 & \cdots & \cdots & 0 \end{pmatrix}$.

Since $\boldsymbol{P}$ is invertible,

$$\boldsymbol{P}^{-1}\boldsymbol{AP} = \boldsymbol{P}^{-1}\boldsymbol{P}\begin{pmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & 0 \\ \vdots & \vdots & & \ddots & 1 \\ 0 & 0 & \cdots & \cdots & 0 \end{pmatrix} = \begin{pmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & 0 \\ \vdots & \vdots & & \ddots & 1 \\ 0 & 0 & \cdots & \cdots & 0 \end{pmatrix}.$$

6. (a) Observe that $\boldsymbol{Ae_i} = \boldsymbol{a_i}$ for $i = 1, 2, \ldots, n$. For $i \neq j$,

$$\begin{aligned} & \boldsymbol{e_i} \text{ is orthogonal to } \boldsymbol{e_j} \\ \Rightarrow\ & \boldsymbol{Ae_i} \text{ is orthogonal to } \boldsymbol{Ae_j} \\ \Rightarrow\ & \boldsymbol{a_i} \text{ is orthogonal to } \boldsymbol{a_j}. \end{aligned}$$

So $\{\boldsymbol{a_1}, \boldsymbol{a_2}, \ldots, \boldsymbol{a_n}\}$ is an orthogonal set.

Since $\dim(\mathbb{R}^n) = n$, $\{\boldsymbol{a_1}, \boldsymbol{a_2}, \ldots, \boldsymbol{a_n}\}$ is an orthogonal basis for $\mathbb{R}^n$.

(b) For $i \neq j$,

$$\begin{aligned} \frac{1}{\sqrt{2}} &= \frac{||\boldsymbol{e_i}||}{||\boldsymbol{e_i} + \boldsymbol{e_j}||} \\ &= \frac{\boldsymbol{e_i} \cdot (\boldsymbol{e_i} + \boldsymbol{e_j})}{||\boldsymbol{e_i}||\,||\boldsymbol{e_i} + \boldsymbol{e_j}||} \\ &= \cos(\text{the angle between } \boldsymbol{e_i} \text{ and } \boldsymbol{e_i} + \boldsymbol{e_j}) \\ &= \cos(\text{the angle between } \boldsymbol{Ae_i} = \boldsymbol{a_i} \text{ and } A(\boldsymbol{e_i} + \boldsymbol{e_j}) = \boldsymbol{a_i} + \boldsymbol{a_j}) \\ &= \frac{\boldsymbol{a_i} \cdot (\boldsymbol{a_i} + \boldsymbol{a_j})}{||\boldsymbol{a_i}||\,||\boldsymbol{a_i} + \boldsymbol{a_j}||} = \frac{||\boldsymbol{a_i}||}{||\boldsymbol{a_i} + \boldsymbol{a_j}||} \end{aligned}$$

and hence

$$\frac{||\boldsymbol{a_i}||}{||\boldsymbol{a_i}+\boldsymbol{a_j}||} = \frac{1}{\sqrt{2}} = \frac{||\boldsymbol{a_j}||}{||\boldsymbol{a_i}+\boldsymbol{a_j}||} \quad \Rightarrow \quad ||\boldsymbol{a_i}|| = ||\boldsymbol{a_j}||.$$

Let $c = ||\boldsymbol{a_1}|| = ||\boldsymbol{a_2}|| = \cdots = ||\boldsymbol{a_n}||$. Then $\{\,\frac{1}{c}\boldsymbol{a_1},\ \frac{1}{c}\boldsymbol{a_2},\ \ldots,\ \frac{1}{c}\boldsymbol{a_n}\,\}$ is an orthonormal basis for $\mathbb{R}^n$.

Thus $\frac{1}{c}\boldsymbol{A} = \begin{pmatrix} \frac{1}{c}\boldsymbol{a_1} & \frac{1}{c}\boldsymbol{a_2} & \cdots & \frac{1}{c}\boldsymbol{a_n} \end{pmatrix}$ is an orthogonal matrix.

Remark: This was the most difficult part in the exam paper and very few students could answer correctly.

Semester 1, 2016-17

1. (a) $\begin{pmatrix} 1 & 0 & 1 & -1 \\ 0 & k & 1 & k \\ k & k & 2 & 0 \\ k & 0 & k & 0 \end{pmatrix} \xrightarrow[R_4 - kR_1]{R_3 - kR_1} \begin{pmatrix} 1 & 0 & 1 & -1 \\ 0 & k & 1 & k \\ 0 & k & 2-k & k \\ 0 & 0 & 0 & k \end{pmatrix} \xrightarrow{R_3 - R_2} \begin{pmatrix} 1 & 0 & 1 & -1 \\ 0 & k & 1 & k \\ 0 & 0 & 1-k & 0 \\ 0 & 0 & 0 & k \end{pmatrix}$

Case 1: $k = 0$.

$$\begin{pmatrix} 1 & 0 & 1 & -1 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \xrightarrow[R_1 - R_2]{R_3 - R_2} \begin{pmatrix} 1 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

Case 2: $k = 1$.

$$\begin{pmatrix} 1 & 0 & 1 & -1 \\ 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \xrightarrow{R_3 \leftrightarrow R_4} \begin{pmatrix} 1 & 0 & 1 & -1 \\ 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix} \xrightarrow[R_2 - R_3]{R_1 + R_3} \begin{pmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

Case 3: $k \neq 0$ and $k \neq 1$.

$$\begin{pmatrix} 1 & 0 & 1 & -1 \\ 0 & k & 1 & k \\ 0 & 0 & 1-k & 0 \\ 0 & 0 & 0 & k \end{pmatrix} \xrightarrow[\frac{1}{k}R_4]{\substack{\frac{1}{k}R_2 \\ \frac{1}{1-k}R_3}} \begin{pmatrix} 1 & 0 & 1 & -1 \\ 0 & 1 & \frac{1}{k} & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \xrightarrow[R_2 - R_4]{R_1 + R_4} \begin{pmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & \frac{1}{k} & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

$$\xrightarrow[R_2 - \frac{1}{k}R_3]{R_1 - R_3} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

Alternatively, for this case, since $\boldsymbol{A}$ is invertible, the reduced row-echelon form is an identity matrix.

(b) **Case 1:** $k = 0$.

The general solution for $\boldsymbol{Ax} = \boldsymbol{0}$ is

$$\boldsymbol{x} = \begin{pmatrix} t \\ s \\ 0 \\ t \end{pmatrix} = s\begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix} + t\begin{pmatrix} 1 \\ 0 \\ 0 \\ 1 \end{pmatrix} \quad \text{for } s, t \in \mathbb{R}.$$

Thus $\{\, (0,1,0,0)^{\mathrm{T}},\ (1,0,0,1)^{\mathrm{T}} \,\}$ is a basis for the nullspace of $\boldsymbol{A}$.

Case 2: $k = 1$.

The general solution for $\boldsymbol{Ax} = \mathbf{0}$ is

$$\boldsymbol{x} = \begin{pmatrix} -t \\ -t \\ t \\ 0 \end{pmatrix} = t \begin{pmatrix} -1 \\ -1 \\ 1 \\ 0 \end{pmatrix} \quad \text{for } s, t \in \mathbb{R}.$$

Thus $\{ (-1, -1, 1, 0)^{\mathrm{T}} \}$ is a basis for the nullspace of $\boldsymbol{A}$.

Case 3: $k \neq 0$ and $k \neq 1$.

The linear system $\boldsymbol{Ax} = \mathbf{0}$ has only the trivial solution, i.e. the solution space is the zero space.

The basis for the nullspace of $\boldsymbol{A}$ is the empty set, $\emptyset$.

2. (a) Since

$$\begin{aligned} &(a + b - 2c,\ 2b - c,\ 3c + d,\ a + 3b + d) \\ = {}& a(1, 0, 0, 1) + b(1, 2, 0, 3) + c(-2, -1, 3, 0) + d(0, 0, 1, 1), \end{aligned}$$

$V = \text{span}\{ (1, 0, 0, 1),\ (1, 2, 0, 3),\ (-2, -1, 3, 0),\ (0, 0, 1, 1) \}$ and hence is a subspace of $\mathbb{R}^4$.

(b) **Method 1:**

$$\begin{pmatrix} 1 & 0 & 0 & 1 \\ 1 & 2 & 0 & 3 \\ -2 & -1 & 3 & 0 \\ 0 & 0 & 1 & 1 \end{pmatrix} \xrightarrow[\text{Elimination}]{\text{Guassian}} \begin{pmatrix} 1 & 0 & 0 & 1 \\ 0 & 2 & 0 & 2 \\ 0 & 0 & 3 & 3 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

$\{ (1, 0, 0, 1),\ (0, 2, 0, 2),\ (0, 0, 3, 3) \}$ is a basis for V and $\dim(V) = 3$.

Method 2:

$$\begin{pmatrix} 1 & 1 & -2 & 0 \\ 0 & 2 & -1 & 0 \\ 0 & 0 & 3 & 1 \\ 1 & 3 & 0 & 1 \end{pmatrix} \xrightarrow[\text{Elimination}]{\text{Guassian}} \begin{pmatrix} 1 & 1 & -2 & 0 \\ 0 & 2 & -1 & 0 \\ 0 & 0 & 3 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

$\{ (1, 0, 0, 1),\ (1, 2, 0, 3),\ (-2, -1, 3, 0) \}$ is a basis for V and $\dim(V) = 3$.

(c) (i) Yes.

$$(1 + a + b - 2c,\ 2b - c,\ 3c + d,\ 1 + a + 3b + d) = (0, 0, 0, 0)$$

$$\Leftrightarrow \quad \begin{cases} a + \ \ b - 2c \qquad\ \ = -1 \\ \qquad 2b - \ \ c \qquad\ \ = \ \ 0 \\ \qquad\qquad\ \ 3c + d = \ \ 0 \\ a + 3b \qquad\ \ + d = -1. \end{cases}$$

$$\begin{pmatrix} 1 & 1 & -2 & 0 & -1 \\ 0 & 2 & -1 & 0 & 0 \\ 0 & 0 & 3 & 1 & 0 \\ 1 & 3 & 0 & 1 & -1 \end{pmatrix} \xrightarrow[\text{Elimination}]{\text{Guassian}} \begin{pmatrix} 1 & 1 & -2 & 0 & -1 \\ 0 & 2 & -1 & 0 & 0 \\ 0 & 0 & 3 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

The linear system is consistent, i.e. there exists real numbers a, b, c, d such that $(1 + a + b - 2c,\ 2b - c,\ 3c + d,\ 1 + a + 3b + d) = (0, 0, 0, 0)$.

So the zero vector is contained in V.

(ii) Yes.

Remark: Actually, $W = V$.

3. (a) $\det(\lambda \boldsymbol{I} - \boldsymbol{B}) = \begin{vmatrix} \lambda - 4 & 0 & -2 & 2 \\ 0 & \lambda - 4 & 2 & -2 \\ 0 & 0 & \lambda - 2 & -2 \\ 0 & 0 & -2 & \lambda - 2 \end{vmatrix}$

$$= (\lambda - 4)^2 \begin{vmatrix} \lambda - 2 & -2 \\ -2 & \lambda - 2 \end{vmatrix}$$

$$= \lambda\,(\lambda - 4)^3.$$

The eigenvalues of $\boldsymbol{B}$ are 0 and 4.

- For $\lambda = 0$, $\quad (\lambda \boldsymbol{I} - \boldsymbol{B})\boldsymbol{x} = \boldsymbol{0} \quad \Leftrightarrow \quad \begin{pmatrix} -4 & 0 & -2 & 2 \\ 0 & -4 & 2 & -2 \\ 0 & 0 & -2 & -2 \\ 0 & 0 & -2 & -2 \end{pmatrix} \boldsymbol{x} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}.$

$$\left(\begin{array}{cccc|c} -4 & 0 & -2 & 2 & 0 \\ 0 & -4 & 2 & -2 & 0 \\ 0 & 0 & -2 & -2 & 0 \\ 0 & 0 & -2 & -2 & 0 \end{array}\right) \xrightarrow[\text{Elimination}]{\text{Guass-Jordan}} \left(\begin{array}{cccc|c} 1 & 0 & 0 & -1 & 0 \\ 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right)$$

The general solution is

$$\boldsymbol{x} = \begin{pmatrix} t \\ -t \\ -t \\ t \end{pmatrix} = t \begin{pmatrix} 1 \\ -1 \\ -1 \\ 1 \end{pmatrix} \quad \text{for } t \in \mathbb{R}.$$

So $\{\,(1, -1, -1, 1)^{\mathrm{T}}\,\}$ is a basis for E_0.

- For $\lambda = 4$, $\quad (\lambda \boldsymbol{I} - \boldsymbol{B})\boldsymbol{x} = \boldsymbol{0} \quad \Leftrightarrow \quad \begin{pmatrix} 0 & 0 & -2 & 2 \\ 0 & 0 & 2 & -2 \\ 0 & 0 & 2 & -2 \\ 0 & 0 & 2 & -2 \end{pmatrix} \boldsymbol{x} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}$.

$$\left(\begin{array}{cccc|c} 0 & 0 & -2 & 2 & 0 \\ 0 & 0 & 2 & -2 & 0 \\ 0 & 0 & 2 & -2 & 0 \\ 0 & 0 & 2 & -2 & 0 \end{array}\right) \xrightarrow[\text{Elimination}]{\text{Guass-Jordan}} \left(\begin{array}{cccc|c} 0 & 0 & 1 & -1 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right)$$

The general solution is

$$\boldsymbol{x} = \begin{pmatrix} r \\ s \\ t \\ t \end{pmatrix} = r \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} + s \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix} + t \begin{pmatrix} 0 \\ 0 \\ 1 \\ 1 \end{pmatrix} \quad \text{for } r, s, t \in \mathbb{R}.$$

So $\{ (1, 0, 0, 0)^{\mathrm{T}},\ (0, 1, 0, 0)^{\mathrm{T}},\ (0, 0, 1, 1)^{\mathrm{T}} \}$ is a basis for E_4.

Let $\boldsymbol{P} = \begin{pmatrix} 1 & 1 & 0 & 0 \\ -1 & 0 & 1 & 0 \\ -1 & 0 & 0 & 1 \\ 1 & 0 & 0 & 1 \end{pmatrix}$ and $\boldsymbol{D} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 4 & 0 & 0 \\ 0 & 0 & 4 & 0 \\ 0 & 0 & 0 & 4 \end{pmatrix}$. Then $\boldsymbol{P}^{-1}\boldsymbol{B}\boldsymbol{P} = \boldsymbol{D}$.

(b) For example, let $\boldsymbol{C} = \boldsymbol{P} \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 2 \end{pmatrix} \boldsymbol{P}^{-1}$. Then $\boldsymbol{C}^2 = \boldsymbol{B}$.

4. (a) (i) $\begin{pmatrix} 1 & 0 & 0 & 1 & 0 & 0 \\ 1 & 1 & 1 & 0 & 1 & 0 \\ 1 & 1 & -1 & 0 & 0 & 1 \end{pmatrix} \xrightarrow[\text{Elimination}]{\text{Guass-Jordan}} \begin{pmatrix} 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & -1 & \frac{1}{2} & \frac{1}{2} \\ 0 & 0 & 1 & 0 & \frac{1}{2} & -\frac{1}{2} \end{pmatrix}$

(ii) $(\boldsymbol{e_1})_S = (1, -1, 0)$, $(\boldsymbol{e_2})_S = (0, \frac{1}{2}, \frac{1}{2})$, $(\boldsymbol{e_3})_S = (0, \frac{1}{2}, -\frac{1}{2})$.

(b) (i) $T(\boldsymbol{e_1}) = \boldsymbol{u_1} - \boldsymbol{u_2} = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} - \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$.

$T(\boldsymbol{e_2}) = 0\boldsymbol{u_1} + \frac{1}{2}\boldsymbol{u_2} = \frac{1}{2} \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ \frac{1}{2} \\ \frac{1}{2} \end{pmatrix}$.

$T(\boldsymbol{e_3}) = 0\boldsymbol{u_1} + \frac{1}{2}\boldsymbol{u_2} = \frac{1}{2} \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ \frac{1}{2} \\ \frac{1}{2} \end{pmatrix}$.

Thus the standard matrix for T is

$$\begin{pmatrix} T(\boldsymbol{e_1}) & T(\boldsymbol{e_2}) & T(\boldsymbol{e_3}) \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \frac{1}{2} & \frac{1}{2} \\ 0 & \frac{1}{2} & \frac{1}{2} \end{pmatrix}.$$

(ii) From the definition of T, $\mathrm{R}(T) = \mathrm{span}\{\boldsymbol{u_1}, \boldsymbol{u_2}\}$. It is obvious that $\{\boldsymbol{u_1}, \boldsymbol{u_2}\}$ is linearly independent and hence is a basis for $\mathrm{R}(T)$.

So $\mathrm{rank}(T) = \dim(\mathrm{R}(T)) = 2$

and $\mathrm{nullity}(T) = \dim(\mathbb{R}^3) - \mathrm{rank}(T) = 3 - 2 = 1$.

(iii) Since $\boldsymbol{u_3} \cdot \boldsymbol{u_1} = 0$ and $\boldsymbol{u_3} \cdot \boldsymbol{u_2} = 0$, $\boldsymbol{u_3}$ is orthogonal to V.

For any $\boldsymbol{x} = c_1\boldsymbol{u_1} + c_2\boldsymbol{u_2} + c_3\boldsymbol{u_3} \in \mathbb{R}^3$,

$$\boldsymbol{x} = c_1\boldsymbol{u_1} + c_2\boldsymbol{u_2} + c_3\boldsymbol{u_3} = T(\boldsymbol{x}) + c_3\boldsymbol{u_3}.$$

As $T(\boldsymbol{x}) = c_1\boldsymbol{u_1} + c_2\boldsymbol{u_2} \in V$ and $c_3\boldsymbol{u_3}$ is orthogonal to V, $T(\boldsymbol{x})$ is the orthogonal projection of $\boldsymbol{x}$ onto V.

5. (a) Let $\boldsymbol{A} = \begin{pmatrix} \boldsymbol{a_1} & \boldsymbol{a_2} & \cdots & \boldsymbol{a_n} \end{pmatrix}$ where $\boldsymbol{a_j}$ is the jth column of $\boldsymbol{A}$.

$$\boldsymbol{A}\boldsymbol{e_j} = \begin{pmatrix} \boldsymbol{a_1} & \boldsymbol{a_2} & \cdots & \boldsymbol{a_n} \end{pmatrix} \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} \begin{array}{l} \\ \\ \\ \leftarrow \text{the } j\text{th coordinate} \\ \\ \\ \\ \end{array}$$

$$= 0\boldsymbol{a_1} + \cdots 0\boldsymbol{a_{j-1}} + 1\boldsymbol{a_j} + 0\boldsymbol{a_{j+1}} + \cdots + 0\boldsymbol{a_n} = \boldsymbol{a_j}.$$

(b) (i) Since $\boldsymbol{A}^{m-1} \neq \boldsymbol{0}$, $\boldsymbol{A}^{m-1}$ has at least one nonzero column, say, the jth column is nonzero.

Let $\boldsymbol{u} = \boldsymbol{e_j}$. Then by (a), $\boldsymbol{A}^{m-1}\boldsymbol{e_j} =$ the jth column of $\boldsymbol{A}^{m-1} \neq \boldsymbol{0}$.

(ii) Observe that $\boldsymbol{A}^{m+k} = \boldsymbol{A}^m\boldsymbol{A}^k = \boldsymbol{0}\boldsymbol{A}^k = \boldsymbol{0}$ for all integer $k \geq 0$.

Consider the vector equation

$$c_1\boldsymbol{u} + c_2\boldsymbol{A}\boldsymbol{u} + \cdots + c_m\boldsymbol{A}^{m-1}\boldsymbol{u} = \boldsymbol{0}. \tag{$*$}$$

Pre-multiplying $\boldsymbol{A}^{m-1}$ to both sides of $(*)$, we have

$$\begin{aligned} & \boldsymbol{A}^{m-1}(c_1\boldsymbol{u} + c_2\boldsymbol{A}\boldsymbol{u} + \cdots + c_n\boldsymbol{A}^{m-1}\boldsymbol{u}) = \boldsymbol{A}^{m-1}\boldsymbol{0} \\ \Rightarrow\ & c_1\boldsymbol{A}^{m-1}\boldsymbol{u} + c_2\boldsymbol{A}^m\boldsymbol{u} + \cdots + c_n\boldsymbol{A}^{2m-2}\boldsymbol{u} = \boldsymbol{0} \\ \Rightarrow\ & c_1\boldsymbol{A}^{m-1}\boldsymbol{u} = \boldsymbol{0}. \end{aligned}$$

Since $\boldsymbol{A}^{m-1}\boldsymbol{u}$ is a nonzero vector, $c_1 = 0$.

Suppose we have already shown that $c_1 = 0$, $c_2 = 0$, $\ldots$, $c_i = 0$ for some i with $1 \le i < m$.
Pre-multiplying $\boldsymbol{A}^{m-i-1}$ to both sides of $(*)$, we have

$$\begin{aligned} & \boldsymbol{A}^{m-i-1}(c_{i+1}\boldsymbol{A}^{i}\boldsymbol{u} + c_{i+2}\boldsymbol{A}^{i+1}\boldsymbol{u} + \cdots + c_n\boldsymbol{A}^{m-1}\boldsymbol{u}) = \boldsymbol{A}^{m-i-1}\boldsymbol{0} \\ \Rightarrow\quad & c_{i+1}\boldsymbol{A}^{m-1}\boldsymbol{u} + c_{i+2}\boldsymbol{A}^{m}\boldsymbol{u} + \cdots + c_n\boldsymbol{A}^{2m-i-2}\boldsymbol{u} = \boldsymbol{0} \\ \Rightarrow\quad & c_{i+1}\boldsymbol{A}^{m-1}\boldsymbol{u} = \boldsymbol{0}. \end{aligned}$$

Since $\boldsymbol{A}^{m-1}\boldsymbol{u}$ is a nonzero vector, $c_{i+1} = 0$.

By mathematical induction, we have shown that $c_1 = 0$, $c_2 = 0$, $\ldots$, $c_m = 0$, i.e. $(*)$ has only the trivial solution.

So $\{\boldsymbol{u},\ \boldsymbol{A}\boldsymbol{u},\ \ldots,\ \boldsymbol{A}^{m-1}\boldsymbol{u}\}$ is linearly independent.

(c) (Prove by contradiction.)

Assume that $\boldsymbol{A}^n \ne \boldsymbol{0}$.

Applying (b) with $m = n+1$, there exists $\boldsymbol{u} \in \mathbb{R}^n$ such that $\{\boldsymbol{u},\ \boldsymbol{A}\boldsymbol{u},\ \ldots,\ \boldsymbol{A}^{n}\boldsymbol{u}\}$ is linearly independent.

But this contradicts the fact that a set of $n+1$ vectors in $\mathbb{R}^n$ cannot be linearly independent (by Theorem 3.2.7).

So we conclude that our assumption is wrong and hence $\boldsymbol{A}^n = \boldsymbol{0}$.

6. (a) (i) Since $\boldsymbol{B}^{\mathrm{T}} = \boldsymbol{B}$,

$$\boldsymbol{u} \cdot (\boldsymbol{B}\boldsymbol{v}) = \boldsymbol{u}^{\mathrm{T}}\boldsymbol{B}\boldsymbol{v} = \boldsymbol{u}^{\mathrm{T}}\boldsymbol{B}^{\mathrm{T}}\boldsymbol{v} = (\boldsymbol{B}\boldsymbol{u})^{\mathrm{T}}\boldsymbol{v} = (\boldsymbol{B}\boldsymbol{u}) \cdot \boldsymbol{v}.$$

By $\boldsymbol{B}\boldsymbol{u} = \lambda\boldsymbol{u}$ and $\boldsymbol{B}\boldsymbol{v} = \mu\boldsymbol{v}$,

$$\begin{aligned} & \lambda(\boldsymbol{u}\cdot\boldsymbol{v}) = (\lambda\boldsymbol{u})\cdot\boldsymbol{v} = (\boldsymbol{B}\boldsymbol{u})\cdot\boldsymbol{v} = \boldsymbol{u}\cdot(\boldsymbol{B}\boldsymbol{v}) = \boldsymbol{u}\cdot(\mu\boldsymbol{v}) = \mu(\boldsymbol{u}\cdot\boldsymbol{v}) \\ \Rightarrow\quad & (\lambda-\mu)(\boldsymbol{u}\cdot\boldsymbol{v}) = 0. \end{aligned}$$

As $\lambda \ne \mu$, $\boldsymbol{u}\cdot\boldsymbol{v} = 0$.

(ii) $\boldsymbol{B} = \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 0 & 3 \end{pmatrix}\begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{pmatrix}^{\mathrm{T}} = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}.$

(b) **Method 1:**

Since $\boldsymbol{C}$ is a symmetric matrix, there exists an orthonormal basis $\{\boldsymbol{u_1},\ \boldsymbol{u_2},\ \ldots,\ \boldsymbol{u_n}\}$ for $\mathbb{R}^n$ such that $\boldsymbol{C}\boldsymbol{u_i} = \lambda_i\boldsymbol{u_i}$ for $i = 1, 2, \ldots, n$.

For any nonzero $\boldsymbol{x} \in \mathbb{R}^n$, we can write $\boldsymbol{x} = c_1\boldsymbol{u_1} + c_2\boldsymbol{u_2} + \cdots + c_n\boldsymbol{u_n}$ for some $c_1, c_2, \ldots, c_n \in \mathbb{R}^n$.

Since $\{\boldsymbol{u_1},\, \boldsymbol{u_2},\, \ldots,\, \boldsymbol{u_n}\}$ is orthonormal, $\boldsymbol{u_i} \cdot \boldsymbol{u_i} = 1$ and $\boldsymbol{u_i} \cdot \boldsymbol{u_j} = 0$ for $i \neq j$. So

$$\begin{aligned}\boldsymbol{x}^{\mathrm{T}}\boldsymbol{x} &= \boldsymbol{x} \cdot \boldsymbol{x} \\ &= (c_1\boldsymbol{u_1} + c_2\boldsymbol{u_2} + \cdots + c_n\boldsymbol{u_n}) \cdot (c_1\boldsymbol{u_1} + c_2\boldsymbol{u_2} + \cdots + c_n\boldsymbol{u_n}) \\ &= c_1^2 + c_2^2 + \cdots + c_n^2.\end{aligned}$$

On the other hand,

$$\begin{aligned}\boldsymbol{x}^{\mathrm{T}}\boldsymbol{C}\boldsymbol{x} &= \boldsymbol{x} \cdot (\boldsymbol{C}\boldsymbol{x}) \\ &= (c_1\boldsymbol{u_1} + c_2\boldsymbol{u_2} + \cdots + c_n\boldsymbol{u_n}) \cdot (c_1\boldsymbol{C}\boldsymbol{u_1} + c_2\boldsymbol{C}\boldsymbol{u_2} + \cdots + c_n\boldsymbol{C}\boldsymbol{u_n}) \\ &= (c_1\boldsymbol{u_1} + c_2\boldsymbol{u_2} + \cdots + c_n\boldsymbol{u_n}) \cdot (\lambda_1 c_1\boldsymbol{u_1} + \lambda_2 c_2\boldsymbol{u_2} + \cdots + \lambda_n c_n\boldsymbol{u_n}) \\ &= \lambda_1 c_1^2 + \lambda_2 c_2^2 + \cdots + \lambda_n c_n^2.\end{aligned}$$

By $\lambda_1 \leq \lambda_2 \leq \cdots \leq \lambda_n$,

$$\lambda_1(c_1^2 + c_2^2 + \cdots + c_n^2) \leq \lambda_1 c_1^2 + \lambda_2 c_2^2 + \cdots + \lambda_n c_n^2 \leq \lambda_n(c_1^2 + c_2^2 + \cdots + c_n^2)$$

$$\Rightarrow \quad \lambda_1 \leq \frac{\lambda_1 c_1^2 + \lambda_2 c_2^2 + \cdots + \lambda_n c_n^2}{c_1^2 + c_2^2 + \cdots + c_n^2} \leq \lambda_n$$

$$\Rightarrow \quad \lambda_1 \leq \frac{\boldsymbol{x}^{\mathrm{T}}\boldsymbol{C}\boldsymbol{x}}{\boldsymbol{x}^{\mathrm{T}}\boldsymbol{x}} \leq \lambda_n.$$

Method 2:

Since $\boldsymbol{C}$ is a symmetric matrix, there exists an orthogonal matrix $\boldsymbol{P}$ such that

$$\boldsymbol{P}^{\mathrm{T}}\boldsymbol{C}\boldsymbol{P} = \begin{pmatrix} \lambda_1 & & & 0 \\ & \lambda_2 & & \\ & & \ddots & \\ 0 & & & \lambda_n \end{pmatrix}.$$

Define $\boldsymbol{y} = \boldsymbol{P}^T\boldsymbol{x}$, i.e. $\boldsymbol{x} = \boldsymbol{P}\boldsymbol{y}$. Write $\boldsymbol{y} = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}$. Then

$$\frac{\boldsymbol{x}^{\mathrm{T}}\boldsymbol{C}\boldsymbol{x}}{\boldsymbol{x}^{\mathrm{T}}\boldsymbol{x}} = \frac{\boldsymbol{y}^{\mathrm{T}}\boldsymbol{P}^{\mathrm{T}}\boldsymbol{C}\boldsymbol{P}\boldsymbol{y}}{\boldsymbol{y}^{\mathrm{T}}\boldsymbol{P}^{\mathrm{T}}\boldsymbol{P}\boldsymbol{y}} = \frac{\boldsymbol{y}^{\mathrm{T}}\boldsymbol{P}^{\mathrm{T}}\boldsymbol{C}\boldsymbol{P}\boldsymbol{y}}{\boldsymbol{y}^{\mathrm{T}}\boldsymbol{y}} = \frac{\lambda_1 y_1^2 + \lambda_2 y_2^2 + \cdots + \lambda_n y_n^2}{y_1^2 + y_2^2 + \cdots + y_n^2}.$$

By $\lambda_1 \leq \lambda_2 \leq \cdots \leq \lambda_n$,

$$\lambda_1(y_1^2 + y_2^2 + \cdots + y_n^2) \leq \lambda_1 y_1^2 + \lambda_2 y_2^2 + \cdots + \lambda_n y_n^2 \leq \lambda_n(y_1^2 + y_2^2 + \cdots + y_n^2)$$

$$\Rightarrow \quad \lambda_1 \leq \frac{\lambda_1 y_1^2 + \lambda_2 y_2^2 + \cdots + \lambda_n y_n^2}{y_1^2 + y_2^2 + \cdots + y_n^2} \leq \lambda_n$$

$$\Rightarrow \quad \lambda_1 \leq \frac{\boldsymbol{x}^{\mathrm{T}}\boldsymbol{C}\boldsymbol{x}}{\boldsymbol{x}^{\mathrm{T}}\boldsymbol{x}} \leq \lambda_n.$$

Semester 1, 2017-18

1. (a) $\begin{pmatrix} 1 & 0 & 1 & 0 & 0 \\ -1 & 0 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 & 0 \end{pmatrix} \xrightarrow[R_4 - R_1]{R_2 + R_1} \begin{pmatrix} 1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 1 & 0 \end{pmatrix}$

$\xrightarrow{R_4 - R_2} \begin{pmatrix} 1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix} \xrightarrow{R_1 - R_2} \begin{pmatrix} 1 & 0 & 0 & -1 & 0 \\ 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$

(b) $\{ (1,0,0,-1,0),\ (0,0,1,1,0),\ (0,0,0,0,1) \}$ is a basis for the row space of $\boldsymbol{A}$.

(c) $\left\{ \begin{pmatrix} 1 \\ -1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix} \right\}$ is a basis for the column space of $\boldsymbol{A}$.

(d) The general solution to $\boldsymbol{Ax} = \boldsymbol{0}$ is

$$\boldsymbol{x} = \begin{pmatrix} t \\ s \\ -t \\ t \\ 0 \end{pmatrix} = s \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} + t \begin{pmatrix} 1 \\ 0 \\ -1 \\ 1 \\ 0 \end{pmatrix}$$

where s, t are arbitrary parameters.

Hence $\left\{ \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ -1 \\ 1 \\ 0 \end{pmatrix} \right\}$ is a basis for the nullspace of $\boldsymbol{A}$.

2. Let

$$\begin{aligned} \boldsymbol{v_1} &= (1,1,0,0), \\ \boldsymbol{v_2} &= (1,1,-1,-1) - \tfrac{2}{2}(1,1,0,0) = (0,0,-1,-1), \\ \boldsymbol{v_3} &= (1,a,1,a) - \tfrac{1+a}{2}(1,1,0,0) - \tfrac{-1-a}{2}(0,0,-1,-1) \\ &= \left(\tfrac{1-a}{2},\ \tfrac{a-1}{2},\ \tfrac{1-a}{2},\ \tfrac{a-1}{2} \right). \end{aligned}$$

- If $a \neq 1$, then

$$\begin{aligned}
&\left\{\tfrac{1}{||\boldsymbol{v_1}||}\boldsymbol{v_1},\ \tfrac{1}{||\boldsymbol{v_2}||}\boldsymbol{v_2},\ \tfrac{1}{||\boldsymbol{v_3}||}\boldsymbol{v_3}\right\}\\
=&\left\{\tfrac{1}{\sqrt{2}}(1,1,0,0),\ \tfrac{1}{\sqrt{2}}(0,0,-1,-1),\ \tfrac{1}{|1-a|}\left(\tfrac{1-a}{2},\ \tfrac{a-1}{2},\ \tfrac{1-a}{2},\ \tfrac{a-1}{2}\right)\right\}\\
=&\left\{\tfrac{1}{\sqrt{2}}(1,1,0,0),\ \tfrac{1}{\sqrt{2}}(0,0,-1,-1),\ \tfrac{1}{2}(1,-1,1,-1)\right\}\\
\text{or }&\left\{\tfrac{1}{\sqrt{2}}(1,1,0,0),\ \tfrac{1}{\sqrt{2}}(0,0,-1,-1),\ -\tfrac{1}{2}(1,-1,1,-1)\right\}
\end{aligned}$$

  is an orthonormal basis for V. (Actually, for the third vector, we can choose either $\frac{1}{2}(1,-1,1,-1)$ or $-\frac{1}{2}(1,-1,1,-1)$ since both cases give us orthonormal bases.)

- If $a = 1$, then

$$\left\{\tfrac{1}{||\boldsymbol{v_1}||}\boldsymbol{v_1},\ \tfrac{1}{||\boldsymbol{v_2}||}\boldsymbol{v_2}\right\} = \left\{\tfrac{1}{\sqrt{2}}(1,1,0,0),\ \tfrac{1}{\sqrt{2}}(0,0,-1,-1)\right\}$$

  is an orthonormal basis for V.

3. (a) Consider the vector equation

$$c_1\boldsymbol{w_1} + c_2\boldsymbol{w_2} + c_3\boldsymbol{w_3} = \boldsymbol{0}. \tag{†}$$

Substituting $\boldsymbol{w_1} = \boldsymbol{v_1} + 2\boldsymbol{v_2}$, $\boldsymbol{w_2} = \boldsymbol{v_2} + 2\boldsymbol{v_3}$ and $\boldsymbol{w_3} = \boldsymbol{v_3}$ into (†), we have

$$\begin{aligned}
& c_1(\boldsymbol{v_1} + 2\boldsymbol{v_2}) + c_2(\boldsymbol{v_2} + 2\boldsymbol{v_3}) + c_3\boldsymbol{v_3} = \boldsymbol{0}\\
\Rightarrow\quad & c_1\boldsymbol{v_1} + (2c_1 + c_2)\boldsymbol{v_2} + (2c_2 + c_3)\boldsymbol{v_1} = \boldsymbol{0}. \qquad (‡)
\end{aligned}$$

Since S is linearly independent, the coefficients in the equation (‡) must be all zero, i.e.

$$\begin{cases} c_1 = 0\\ 2c_1 + c_2 = 0\\ 2c_2 + c_3 = 0, \end{cases}$$

which implies $c_1 = 0$, $c_2 = 0$, $c_3 = 0$.

Since the equation (†) has only the trivial solution, T is linearly independent.

On the other hand, $|T| = |S| = \dim(W)$. We conclude that T is a basis for W.

(b) The transition matrix from T to S is

$$\begin{pmatrix}[\boldsymbol{w_1}]_S & [\boldsymbol{w_2}]_S & [\boldsymbol{w_3}]_S\end{pmatrix} = \begin{pmatrix}1 & 0 & 0\\ 2 & 1 & 0\\ 0 & 2 & 1\end{pmatrix}.$$

As

$$\left(\begin{array}{ccc|ccc} 1 & 0 & 0 & 1 & 0 & 0 \\ 2 & 1 & 0 & 0 & 1 & 0 \\ 0 & 2 & 1 & 0 & 0 & 1 \end{array}\right) \xrightarrow{R_2 - 2R_1} \left(\begin{array}{ccc|ccc} 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & -2 & 1 & 0 \\ 0 & 2 & 1 & 0 & 0 & 1 \end{array}\right)$$

$$\xrightarrow{R_3 - 2R_2} \left(\begin{array}{ccc|ccc} 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & -2 & 1 & 0 \\ 0 & 0 & 1 & 4 & -2 & 1 \end{array}\right),$$

the transition matrix from S to T is

$$\begin{pmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 0 & 2 & 1 \end{pmatrix}^{-1} = \begin{pmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ 4 & -2 & 1 \end{pmatrix}.$$

4. (a) The characteristic polynomial of $\boldsymbol{B}$ is

$$\begin{vmatrix} \lambda+2 & 0 & 2 & -1 \\ 1 & \lambda+1 & 2 & -1 \\ -1 & 0 & \lambda-1 & 1 \\ 0 & 0 & 0 & \lambda+1 \end{vmatrix} = (\lambda+1) \begin{vmatrix} \lambda+2 & 0 & 2 \\ 1 & \lambda+1 & 2 \\ -1 & 0 & \lambda-1 \end{vmatrix}$$
$$= (\lambda+1)[(\lambda+2)(\lambda+1)(\lambda-1) + 2(\lambda+1)]$$
$$= (\lambda+1)^3\lambda.$$

Hence the eigenvalues of $\boldsymbol{B}$ are -1 and 0.

(b) $$\left(\begin{array}{cccc|c} 1 & 0 & 2 & -1 & 0 \\ 1 & 0 & 2 & -1 & 0 \\ -1 & 0 & -2 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right) \underset{\text{Elimination}}{\xrightarrow{\text{Gauss-Jordan}}} \left(\begin{array}{cccc|c} 1 & 0 & 2 & -1 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right).$$

The general solution of $(\lambda\boldsymbol{I} - \boldsymbol{B})\boldsymbol{x} = \boldsymbol{0}$, with $\lambda = -1$, is

$$\boldsymbol{x} = \begin{pmatrix} -2s+t \\ r \\ s \\ t \end{pmatrix} = r\begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix} + s\begin{pmatrix} -2 \\ 0 \\ 1 \\ 0 \end{pmatrix} + t\begin{pmatrix} 1 \\ 0 \\ 0 \\ 1 \end{pmatrix}$$

where r, s, t are arbitrary parameters.

Hence $\{ (0,1,0,0)^{\mathrm{T}},\ (-2,0,1,0)^{\mathrm{T}},\ (1,0,0,1)^{\mathrm{T}} \}$ is a basis for E_{-1}.

(c) $$\left(\begin{array}{cccc|c} 2 & 0 & 2 & -1 & 0 \\ 1 & 1 & 2 & -1 & 0 \\ -1 & 0 & -1 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 \end{array}\right) \underset{\text{Elimination}}{\xrightarrow{\text{Gauss-Jordan}}} \left(\begin{array}{cccc|c} 1 & 0 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right).$$

The general solution of $(\lambda \boldsymbol{I} - \boldsymbol{B})\boldsymbol{x} = \boldsymbol{0}$, with $\lambda = 0$, is

$$\boldsymbol{x} = \begin{pmatrix} -t \\ -t \\ t \\ 0 \end{pmatrix} = t \begin{pmatrix} -1 \\ -1 \\ 1 \\ 0 \end{pmatrix}$$

where t is an arbitrary parameter.

Hence $\{ (-1, -1, 1, 0)^{\mathrm{T}} \}$ is a basis for E_0.

(d) Let

$$\boldsymbol{P} = \begin{pmatrix} 0 & -2 & 1 & -1 \\ 1 & 0 & 0 & -1 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix} \quad \text{and} \quad \boldsymbol{D} = \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

Then $\boldsymbol{P}^{-1}\boldsymbol{B}\boldsymbol{P} = \boldsymbol{D}$.

(e) Note that

$$\boldsymbol{D}^{1101} = \begin{pmatrix} (-1)^{1101} & 0 & 0 & 0 \\ 0 & (-1)^{1101} & 0 & 0 \\ 0 & 0 & (-1)^{1101} & 0 \\ 0 & 0 & 0 & 0^{1101} \end{pmatrix} = \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} = \boldsymbol{D}.$$

So $\boldsymbol{B}^{1101} = (\boldsymbol{P}\boldsymbol{D}\boldsymbol{P}^{-1})^{1101} = \boldsymbol{P}\boldsymbol{D}^{1101}\boldsymbol{P}^{-1} = \boldsymbol{P}\boldsymbol{D}\boldsymbol{P}^{-1} = \boldsymbol{B}$.

5. (a) Take any $\boldsymbol{u} \in$ the nullspace of $\boldsymbol{C}$, i.e. $\boldsymbol{C}\boldsymbol{u} = \boldsymbol{0}$. Then

$$\boldsymbol{C}^2\boldsymbol{u} = \boldsymbol{C}\boldsymbol{C}\boldsymbol{u} = \boldsymbol{C}\boldsymbol{0} = \boldsymbol{0}$$

which implies $\boldsymbol{u} \in$ the nullspace of $\boldsymbol{C}^2$.

Hence we have shown that the nullspace of $\boldsymbol{C}$ is a subset of the nullspace of $\boldsymbol{C}^2$.

(b) From (a), the nullspace of $\boldsymbol{C}^2$ is a subspace of the nullspace of $\boldsymbol{C}$.

Suppose the order of $\boldsymbol{C}$ is n. Then

$$\begin{aligned} \dim(\text{the nullspace of } \boldsymbol{C}^2) &= \text{nullity}(\boldsymbol{C}^2) \\ &= n - \text{rank}(\boldsymbol{C}^2) \\ &= n - \text{rank}(\boldsymbol{C}) \\ &= \text{nullity}(\boldsymbol{C}) \\ &= \dim(\text{the nullspace of } \boldsymbol{C}). \end{aligned}$$

So the nullspace of $\boldsymbol{C}^2$ is equal to the nullspace of $\boldsymbol{C}$.

(c) Let $\boldsymbol{C} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$. Then $\boldsymbol{C}^2 = \boldsymbol{C}$ and hence $\text{rank}(\boldsymbol{C}^2) = \text{rank}(\boldsymbol{C})$.

(d) Let $\boldsymbol{C} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$. Then $\boldsymbol{C}^2 = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}^2 = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$.

Since $\text{rank}(\boldsymbol{C}) = 1$ and $\text{rank}(\boldsymbol{C}^2) = 0$, $\text{rank}(\boldsymbol{C}^2) < \text{rank}(\boldsymbol{C})$.

(e) By part (a), the nullspace of $\boldsymbol{C}$ is a subset of the nullspace of $\boldsymbol{C}^2$ which implies $\text{nullity}(\boldsymbol{C}^2) \geq \text{nullity}(\boldsymbol{C})$. Suppose the order of $\boldsymbol{C}$ is n. Then

$$\text{rank}(\boldsymbol{C}^2) = n - \text{nullity}(\boldsymbol{C}^2) \leq n - \text{nullity}(\boldsymbol{C}) = \text{rank}(\boldsymbol{C}).$$

Alternative Proof: We know that $\text{rank}(\boldsymbol{AB}) \leq \min\{\text{rank}(\boldsymbol{A}), \text{rank}(\boldsymbol{B})\}$ for any matrices $\boldsymbol{A}$ and $\boldsymbol{B}$ such that the product $\boldsymbol{AB}$ exists. So

$$\text{rank}(\boldsymbol{C}^2) \leq \min\{\text{rank}(\boldsymbol{C}), \text{rank}(\boldsymbol{C})\} = \text{rank}(\boldsymbol{C}).$$

6. (a) The standard matrix for T_λ is $\boldsymbol{A} - \lambda\boldsymbol{I}$.

(b)
$$\begin{aligned}(\boldsymbol{A} - \lambda\boldsymbol{I})(\boldsymbol{A} - \mu\boldsymbol{I}) &= \boldsymbol{A}(\boldsymbol{A} - \mu\boldsymbol{I}) - \lambda(\boldsymbol{A} - \mu\boldsymbol{I}) \\ &= \boldsymbol{A}^2 - \mu\boldsymbol{A} - \lambda\boldsymbol{A} + \lambda\mu\boldsymbol{I} \\ &= \boldsymbol{A}^2 - \lambda\boldsymbol{A} - \mu\boldsymbol{A} + \mu\lambda\boldsymbol{I} \\ &= \boldsymbol{A}(\boldsymbol{A} - \lambda\boldsymbol{I}) - \mu(\boldsymbol{A} - \lambda\boldsymbol{I}) = (\boldsymbol{A} - \mu\boldsymbol{I})(\boldsymbol{A} - \lambda\boldsymbol{I}).\end{aligned}$$

(c) (i) Since $(\boldsymbol{A} - \lambda_i\boldsymbol{I})\boldsymbol{v} = \boldsymbol{A}\boldsymbol{v} - \lambda_i\boldsymbol{v} = \lambda_i\boldsymbol{v} - \lambda_i\boldsymbol{v} = \mathbf{0}$,

$$\begin{aligned}&(\boldsymbol{A} - \lambda_1\boldsymbol{I})(\boldsymbol{A} - \lambda_2\boldsymbol{I})\cdots(\boldsymbol{A} - \lambda_k\boldsymbol{I})\boldsymbol{v} \\ &= (\boldsymbol{A} - \lambda_1\boldsymbol{I})\cdots(\boldsymbol{A} - \lambda_{i-1}\boldsymbol{I})(\boldsymbol{A} - \lambda_{i+1}\boldsymbol{I})\cdots(\boldsymbol{A} - \lambda_k\boldsymbol{I})(\boldsymbol{A} - \lambda_i\boldsymbol{I})\boldsymbol{v} \\ &= (\boldsymbol{A} - \lambda_1\boldsymbol{I})\cdots(\boldsymbol{A} - \lambda_{i-1}\boldsymbol{I})(\boldsymbol{A} - \lambda_{i+1}\boldsymbol{I})\cdots(\boldsymbol{A} - \lambda_k\boldsymbol{I})\mathbf{0} \\ &= \mathbf{0}.\end{aligned}$$

(ii) Since $\boldsymbol{A}$ is diagonalizable, $\boldsymbol{A}$ has n linearly independent eigenvectors, say, $\boldsymbol{v_1}$, $\boldsymbol{v_2}$, $\ldots$, $\boldsymbol{v_n}$ are n linearly independent eigenvectors.

By (i),

$$S(\boldsymbol{v_i}) = (\boldsymbol{A} - \lambda_1\boldsymbol{I})(\boldsymbol{A} - \lambda_2\boldsymbol{I})\cdots(\boldsymbol{A} - \lambda_k\boldsymbol{I})\boldsymbol{v_i} = \mathbf{0}$$

for all i.

Since $\dim(\mathbb{R}^n) = n$, $\{\boldsymbol{v_1}, \boldsymbol{v_2}, \ldots, \boldsymbol{v_n}\}$ is a basis for $\mathbb{R}^n$.

For any $\boldsymbol{u} \in \mathbb{R}^n$, there exists $c_1, c_2, \ldots, c_n \in \mathbb{R}^n$ such that $\boldsymbol{u} = c_1\boldsymbol{v_1} + c_2\boldsymbol{v_2} + \cdots + c_n\boldsymbol{v_n}$. Hence

$$\begin{aligned}S(\boldsymbol{u}) &= S(c_1\boldsymbol{v_1} + c_2\boldsymbol{v_2} + \cdots + c_n\boldsymbol{v_n}) \\ &= c_1S(\boldsymbol{v_1}) + c_2S(\boldsymbol{v_2}) + \cdots + c_nS(\boldsymbol{v_n}) \\ &= c_1\mathbf{0} + c_2\mathbf{0} + \cdots + c_n\mathbf{0} = \mathbf{0}.\end{aligned}$$

Hence S is the zero transformation.