Further Mathematical Methods (Linear Algebra)

Solutions For The 2002 Examination

Question 1

(a) To be an inner product on the real vector space V, a function $\langle \mathbf{x}, \mathbf{y} \rangle$ which maps vectors $\mathbf{x}, \mathbf{y} \in V$ to $\mathbb{R}$ must be such that:

i. **Positivity:** $\langle \mathbf{x}, \mathbf{x} \rangle \geq 0$ and, $\langle \mathbf{x}, \mathbf{x} \rangle = 0$ if and only if $\mathbf{x} = \mathbf{0}$.

ii. **Symmetry:** $\langle \mathbf{x}, \mathbf{y} \rangle = \langle \mathbf{y}, \mathbf{x} \rangle$.

iii. **Linearity:** $\langle \alpha\mathbf{x} + \beta\mathbf{y}, \mathbf{z} \rangle = \alpha\langle \mathbf{x}, \mathbf{z} \rangle + \beta\langle \mathbf{y}, \mathbf{z} \rangle$.

for all vectors $\mathbf{x}, \mathbf{y}, \mathbf{z} \in V$ and all scalars $\alpha, \beta \in \mathbb{R}$.

(b) We need to show that the function defined by

$$\langle \mathbf{x}, \mathbf{y} \rangle = \mathbf{x}^t \mathsf{A}^t \mathsf{A} \mathbf{y},$$

where A is an invertible 3×3 matrix and $\mathbf{x}, \mathbf{y}$ are any vectors in $\mathbb{R}^3$ is an inner product on $\mathbb{R}^3$.[1] To do this, we show that this formula satisfies all of the conditions given in part **(a)**. Thus, taking any three vectors $\mathbf{x}$, $\mathbf{y}$ and $\mathbf{z}$ in $\mathbb{R}^3$ and any two scalars α and β in $\mathbb{R}$ we have:

i. $\langle \mathbf{x}, \mathbf{x} \rangle = \mathbf{x}^t \mathsf{A}^t \mathsf{A} \mathbf{x} = (\mathsf{A}\mathbf{x})^t (\mathsf{A}\mathbf{x})$, and since $\mathsf{A}\mathbf{x}$ is itself a vector in $\mathbb{R}^3$, say $[a_1, a_2, a_3]^t$ with $a_1, a_2, a_3 \in \mathbb{R}$, we have

$$\langle \mathbf{x}, \mathbf{x} \rangle = \begin{bmatrix} a_1 & a_2 & a_3 \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \\ a_3 \end{bmatrix} = a_1^2 + a_2^2 + a_3^2,$$

which is the sum of the squares of three real numbers and as such it is real and non-negative. Further, to show that $\langle \mathbf{x}, \mathbf{x} \rangle = 0$ if and only if $\mathbf{x} = \mathbf{0}$, we note that:

- **LTR:** If $\langle \mathbf{x}, \mathbf{x} \rangle = 0$, then $a_1^2 + a_2^2 + a_3^2 = 0$. But, this is the sum of the squares of three real numbers and so it must be the case that $a_1 = a_2 = a_3 = 0$. Thus, $\mathsf{A}\mathbf{x} = \mathbf{0}$, and since A is invertible, this gives us $\mathbf{x} = \mathbf{0}$ as the only solution.
- **RTL:** If $\mathbf{x} = \mathbf{0}$, then $\mathsf{A}\mathbf{0} = \mathbf{0}$ and so clearly, $\langle \mathbf{0}, \mathbf{0} \rangle = \mathbf{0}^t \mathbf{0} = 0$.

(as required).

ii. As $\langle \mathbf{x}, \mathbf{y} \rangle = \mathbf{x}^t \mathsf{A}^t \mathsf{A} \mathbf{y}$ is a real number it is unaffected by transposition and so we have:

$$\langle \mathbf{x}, \mathbf{y} \rangle = \langle \mathbf{x}, \mathbf{y} \rangle^t = (\mathbf{x}^t \mathsf{A}^t \mathsf{A} \mathbf{y})^t = \mathbf{y}^t \mathsf{A}^t \mathsf{A} \mathbf{x} = \langle \mathbf{y}, \mathbf{x} \rangle,$$

iii. Taking the vector $\alpha\mathbf{x} + \beta\mathbf{y} \in \mathbb{R}$ we have:

$$\langle \alpha\mathbf{x} + \beta\mathbf{y}, \mathbf{z} \rangle = (\alpha\mathbf{x} + \beta\mathbf{y})^t \mathsf{A}^t \mathsf{A} \mathbf{z} = (\alpha\mathbf{x}^t + \beta\mathbf{y}^t) \mathsf{A}^t \mathsf{A} \mathbf{z} = \alpha\mathbf{x}^t \mathsf{A}^t \mathsf{A} \mathbf{z} + \beta\mathbf{y}^t \mathsf{A}^t \mathsf{A} \mathbf{z} = \alpha\langle \mathbf{x}, \mathbf{z} \rangle + \beta\langle \mathbf{y}, \mathbf{z} \rangle.$$

Consequently, the formula given above does define an inner product on $\mathbb{R}^3$ (as required).

(c) We are given that A is the matrix

$$\mathsf{A} = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 1 & 0 & 1 \end{bmatrix}.$$

[1] Of course, strictly speaking, the quantity given by $\mathbf{x}^t \mathsf{A}^t \mathsf{A} \mathbf{y}$ is a 1×1 matrix. However, it is obvious that here we can treat it as a real number by stipulating that its single entry is the required real number.

So, taking an arbitrary vector $\mathbf{x} = [x_1, x_2, x_3]^t \in \mathbb{R}^3$, we can see that,

$$\mathsf{A}\mathbf{x} = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 1 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} x_1 + x_2 \\ x_2 + x_3 \\ x_1 + x_3 \end{bmatrix},$$

and so using the inner product defined in **(b)**, its norm will be given by

$$\|\mathbf{x}\|^2 = \langle \mathbf{x}, \mathbf{x} \rangle = (\mathsf{A}\mathbf{x})^t(\mathsf{A}\mathbf{x}) = (x_1 + x_2)^2 + (x_2 + x_3)^2 + (x_1 + x_3)^2.$$

Thus, a condition that must be satisfied by the components of this vector if it is to have unit norm using the given inner product is,

$$(x_1 + x_2)^2 + (x_2 + x_3)^2 + (x_1 + x_3)^2 = 1.$$

Hence, to find a symmetric matrix B which expresses this condition in the form $\mathbf{x}^t\mathsf{B}\mathbf{x} = 1$, we expand this out to get:

$$2x_1^2 + 2x_2^2 + 2x_3^2 + 2x_1x_2 + 2x_2x_3 + 2x_1x_3 = 1,$$

which can be written in matrix form as

$$\mathbf{x}^t\mathsf{B}\mathbf{x} = \begin{bmatrix} x_1 & x_2 & x_3 \end{bmatrix} \begin{bmatrix} 2 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = 1,$$

where the required symmetric matrix is

$$\mathsf{B} = \begin{bmatrix} 2 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 2 \end{bmatrix}.$$

Otherwise: $\mathsf{B} = \mathsf{A}^t\mathsf{A}$.

(d) Since we are given the eigenvectors we can calculate the eigenvalues, using $\mathsf{B}\mathbf{x} = \lambda\mathbf{x}$, as follows:

- For $[0, 1, -1]^t$, we have

$$\begin{bmatrix} 2 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 2 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \\ -1 \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \\ -1 \end{bmatrix},$$

  and so 1 is the eigenvalue corresponding to this eigenvector.

- For $[2, -1, -1]^t$, we have

$$\begin{bmatrix} 2 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 2 \end{bmatrix} \begin{bmatrix} 2 \\ -1 \\ -1 \end{bmatrix} = \begin{bmatrix} 2 \\ -1 \\ -1 \end{bmatrix},$$

  and so 1 is the eigenvalue corresponding to this eigenvector.

- For $[1, 1, 1]^t$, we have

$$\begin{bmatrix} 2 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 2 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 4 \\ 4 \\ 4 \end{bmatrix} = 4 \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix},$$

  and so 4 is the eigenvalue corresponding to this eigenvector.

So, for an orthogonal matrix P, we need an orthonormal set of eigenvectors. But, the eigenvectors are already mutually orthogonal, and so we only need to normalise them. Doing this, we find that:

$$\mathsf{P} = \frac{1}{\sqrt{6}} \begin{bmatrix} 0 & 2 & \sqrt{2} \\ \sqrt{3} & -1 & \sqrt{2} \\ -\sqrt{3} & -1 & \sqrt{2} \end{bmatrix},$$

and the corresponding diagonal matrix D is

$$\mathsf{D} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 4 \end{bmatrix},$$

where $\mathsf{P}^t\mathsf{B}\mathsf{P} = \mathsf{D}$.

(e) To find a basis S of $\mathbb{R}^3$ such that the coordinate vector of $\mathbf{x}$ with respect to this basis, i.e. $[\mathbf{x}]_S$ is such that

$$[\mathbf{x}]_S^t\mathsf{C}^t\mathsf{C}[\mathbf{x}]_S = 1,$$

and the associated diagonal matrix C we note that, since $\mathsf{P}^t\mathsf{B}\mathsf{P} = \mathsf{D}$ we have $\mathsf{B} = \mathsf{P}\mathsf{D}\mathsf{P}^t$ and so

$$\mathbf{x}^t\mathsf{B}\mathbf{x} = 1 \implies \mathbf{x}^t\mathsf{P}\mathsf{D}\mathsf{P}^t\mathbf{x} = 1 \implies (\mathsf{P}^t\mathbf{x})^t\mathsf{D}(\mathsf{P}^t\mathbf{x}) = 1.$$

Now, $\mathbf{x}$ can be written as

$$\mathbf{x} = y_1\mathbf{v}_1 + y_2\mathbf{v}_2 + y_3\mathbf{v}_3 = \mathsf{P}\begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix}_S = \mathsf{P}[\mathbf{x}]_S,$$

where the vectors in the set $S = \{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ are the column vectors of the matrix P. As such, since P is orthogonal, we have

$$\mathsf{P}^t\mathbf{x} = [\mathbf{x}]_S,$$

i.e. $\mathsf{P}^t\mathbf{x}$ is the coordinate vector of $\mathbf{x}$ with respect to the basis given by the column vectors of P. That is, if we let S be the basis of $\mathbb{R}^3$ given by the set of vectors

$$S = \left\{ \frac{1}{\sqrt{2}}\begin{bmatrix} 0 \\ 1 \\ -1 \end{bmatrix}, \frac{1}{\sqrt{6}}\begin{bmatrix} 2 \\ -1 \\ -1 \end{bmatrix}, \frac{1}{\sqrt{3}}\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \right\},$$

we have

$$[\mathbf{x}]_S^t\mathsf{D}[\mathbf{x}]_S = 1.$$

Now, D is the diagonal matrix above, and so if we choose the diagonal matrix C to be

$$\mathsf{C} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{bmatrix},$$

then we have

$$[\mathbf{x}]_S^t\mathsf{C}^t\mathsf{C}[\mathbf{x}]_S = 1,$$

as required.

Question 2.

(a) The Leslie matrix for such a population involving three age classes (i.e. C_1, C_2, C_3) is:

$$\mathsf{L} = \begin{bmatrix} a_1 & a_2 & a_3 \\ b_1 & 0 & 0 \\ 0 & b_2 & 0 \end{bmatrix},$$

Here the demographic parameters a_1, a_2, a_3, b_1 and b_2 have the following meanings:

- for $i = 1, 2, 3$, a_i will be the average number of daughters born to a female in the ith age class.
- for $i = 1, 2$, b_i will be the fraction of females in age class C_i expected to survive for the next twenty years and hence enter age class C_{i+1}.

It should be clear that since the number of females entering into the next age class is determined by the relevant b_i, in successive time periods, we have

- $x_2^{(k)} = b_1 x_1^{(k-1)}$ gives the number of females surviving long enough to go from C_1 to C_2, and as such the remaining $(1 - b_1)x_1^{(k-1)}$ females in C_1 die.
- $x_3^{(k)} = b_2 x_2^{(k-1)}$ gives the number of females surviving long enough to go from C_2 to C_3, and as such the remaining $(1 - b_2)x_2^{(k-1)}$ females in C_2 die.

whereas the $x_3^{(k-1)}$ females in C_3 all die by the end of the $(k-1)$th time period. As such, the number of different DNA samples that will be available for cloning at the end of the the $(k-1)$th time period will be

$$(1 - b_1)x_1^{(k-1)} + (1 - b_2)x_2^{(k-1)} + x_3^{(k-1)}.$$

Thus, since the 're-birth' afforded by cloning creates new members of the first age class together with those created by natural means, we have

$$\begin{aligned} x_1^{(k)} &= \underbrace{a_1 x_1^{(k-1)} + a_2 x_2^{(k-1)} + a_3 x_3^{(k-1)}}_{\text{by 'birth'}} + \underbrace{(1 - b_1)x_1^{(k-1)} + (1 - b_2)x_2^{(k-1)} + x_3^{(k-1)}}_{\text{by 'cloning'}} \\ &= (1 + a_1 - b_1)x_1^{(k-1)} + (1 + a_2 - b_2)x_2^{(k-1)} + (1 + a_3)x_3^{(k-1)} \end{aligned}$$

females in C_1 by the end of the kth time period. Consequently, the Leslie matrix L, i.e. the matrix such that $\mathbf{x}^{(k)} = \mathsf{L}\mathbf{x}^{(k-1)}$, is given by

$$\mathsf{L} = \begin{bmatrix} 1 + a_1 - b_1 & 1 + a_2 - b_2 & 1 + a_3 \\ b_1 & 0 & 0 \\ 0 & b_2 & 0 \end{bmatrix}.$$

as required.

The Leslie matrix for the population in question is

$$\mathsf{L} = \begin{bmatrix} 0 & 0 & 3/2 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix},$$

we can see that

$$\mathsf{L}^2 = \begin{bmatrix} 0 & 0 & 3/2 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} 0 & 0 & 3/2 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 3/2 & 0 \\ 0 & 0 & 3/2 \\ 1 & 0 & 0 \end{bmatrix},$$

and,

$$\mathsf{L}^3 = \begin{bmatrix} 0 & 0 & 3/2 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} 0 & 3/2 & 0 \\ 0 & 0 & 3/2 \\ 1 & 0 & 0 \end{bmatrix} = \begin{bmatrix} 3/2 & 0 & 0 \\ 0 & 3/2 & 0 \\ 0 & 0 & 3/2 \end{bmatrix} = \frac{3}{2} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

As such, we can see that after j sixty year periods (i.e. where $k = 3j$) the population distribution vector will be given by

$$\mathbf{x}^{(3j)} = \left(\frac{3}{2}\right)^j \mathbf{x}^{(0)}.$$

as such we can see that:

- Every sixty years the proportion of the population in each age class is the same as it was at the beginning.
- Every sixty years the number of females in each age class changes by a factor of 3/2 (i.e. increases by 50%).

as such we can see that overall the population of females is increasing in a cyclic manner with a period of sixty years.

(b) The steady states of the coupled non-linear differential equations

$$\begin{aligned}\dot{y}_1 &= y_1 - 2y_1^2 - 3y_1y_2 \\ \dot{y}_2 &= 4y_2 - 2y_2^2 - y_1y_2\end{aligned}$$

are given by the solutions of the simultaneous equations

$$\begin{aligned}y_1 - 2y_1^2 - 3y_1y_2 &= 0 \\ 4y_2 - 2y_2^2 - y_1y_2 &= 0\end{aligned}$$

i.e. by $(y_1, y_2) = (0, 0), (0, 2), (1/2, 0)$ and $(-10, 7)$, where the latter steady state is found by solving the linear simultaneous equations

$$\begin{aligned}1 - 2y_1 - 3y_2 &= 0 \\ 4 - 2y_2 - y_1 &= 0\end{aligned}$$

for y_1 and y_2.

The steady state whose asymptotic stability we have to establish is clearly $(0, 2)$. In order to establish it, we work find the Jacobian matrix for this system of differential equations, i.e.

$$\mathsf{DF}[\mathbf{y}] = \begin{bmatrix} 1 - 4y_1 - 3y_2 & -3y_1 \\ -y_2 & 4 - 4y_2 - y_1 \end{bmatrix},$$

and evaluating this at the relevant steady state gives

$$\mathsf{DF}[(0, 2)] = \begin{bmatrix} -5 & 0 \\ -2 & -4 \end{bmatrix}.$$

The eigenvalues of this matrix are -4 and -5 and, since these are both real and negative, this steady state is asymptotically stable.

Question 3.

(a) For a non-empty subset W of V to be a subspace of V we require that, for all vectors $\mathbf{x}, \mathbf{y} \in W$ and all scalars $\alpha \in \mathbb{R}$:

i. **Closure under vector addition:** $\mathbf{x} + \mathbf{y} \in W$.

ii. **Closure under scalar multiplication:** $\alpha\mathbf{x} \in W$.

(b) The *sum* of two subspaces Y and Z of a vector space V, denoted by $Y + Z$, is defined to be the set

$$Y + Z = \{\mathbf{y} + \mathbf{z} | \mathbf{y} \in Y \text{ and } \mathbf{z} \in Z\}.$$

To show that $Y + Z$ is a subspace of V, we note that:

- As Y and Z are both subspaces of V, the additive identity of V (say, the vector $\mathbf{0}$) is in both Y and Z. As such the vector $\mathbf{0} + \mathbf{0} = \mathbf{0} \in Y + Z$ and so this set is non-empty.

As such, referring to part **(a)**, we consider two general vectors in $Y + Z$, say

$$\mathbf{x}_1 = \mathbf{y}_1 + \mathbf{z}_1 \quad \text{where} \quad \mathbf{y}_1 \in Y \quad \text{and} \quad \mathbf{z}_1 \in Z,$$

$$\mathbf{x}_2 = \mathbf{y}_2 + \mathbf{z}_2 \quad \text{where} \quad \mathbf{y}_1 \in Y \quad \text{and} \quad \mathbf{z}_2 \in Z,$$

and any scalar $\alpha \in \mathbb{R}$, to see that:

- $Y + Z$ is closed under vector addition since:

$$\mathbf{x}_1 + \mathbf{x}_2 = (\mathbf{y}_1 + \mathbf{z}_1) + (\mathbf{y}_2 + \mathbf{z}_2) = (\mathbf{y}_1 + \mathbf{y}_2) + (\mathbf{z}_1 + \mathbf{z}_2),$$

  which is a vector in $Y + Z$ since $\mathbf{y}_1 + \mathbf{y}_2 \in Y$ and $\mathbf{z}_1 + \mathbf{z}_2 \in Z$ (as Y and Z are closed under vector addition).

- $Y + Z$ is closed under scalar multiplication since:

$$\alpha\mathbf{x}_1 = \alpha(\mathbf{y}_1 + \mathbf{z}_1) = \alpha\mathbf{y}_1 + \alpha\mathbf{z}_1,$$

  which is a vector in $Y + Z$ since $\alpha\mathbf{y}_1 \in Y$ and $\alpha\mathbf{z}_1 \in Z$ (as Y and Z are closed under scalar multiplication).

Consequently, as $Y + Z$ is non-empty and closed under vector addition and scalar multiplication it is a subspace of V (as required).

(c) For a real vector space V, we say that V is the *direct sum* of two of its subspaces (say, Y and Z), denoted by $V = Y \oplus Z$, if $V = Y + Z$ and every vector $\mathbf{v} \in V$ can be written *uniquely* in terms of vectors in Y and vectors in Z, i.e. we can write

$$\mathbf{v} = \mathbf{y} + \mathbf{z},$$

for vectors $\mathbf{y} \in Y$ and $\mathbf{z} \in Z$ in exactly one way.
To prove that

> The sum of Y and Z is direct iff $Y \cap Z = \{\mathbf{0}\}$.

we note that:

- **LTR:** We are given that the sum of Y and Z is direct. Now, it is clear that
  - We can write $\mathbf{0} \in Y + Z$ as $\mathbf{0} = \mathbf{0} + \mathbf{0}$ since $Y + Z$, Y and Z are all subspaces.
  - Given any vector $\mathbf{u} \in Y \cap Z$, we have $\mathbf{u} \in Y$, $\mathbf{u} \in Z$ and $-\mathbf{u} \in Z$ (as Z is a subspace). As such, we can write $\mathbf{0} \in Y + Z$ as $\mathbf{0} = \mathbf{u} + (-\mathbf{u})$.

But, as $Y+Z$ is direct, by uniqueness, we can write $\mathbf{0} \in Y+Z$ in only one way using vectors Y and vectors in Z, i.e. we must have $\mathbf{u} = \mathbf{0}$. Consequently, we have $Y \cap Z = \{\mathbf{0}\}$ (as required).

- **RTL:** We are given $Y \cap Z = \{\mathbf{0}\}$ and we note that any vector $\mathbf{u} \in Y + Z$ can be written as $\mathbf{u} = \mathbf{y} + \mathbf{z}$ in terms of a vector $\mathbf{y} \in Y$ and a vector $\mathbf{z} \in Z$. In order to establish that this sum is direct, we have to show that this representation of $\mathbf{x}$ is unique.

  To do this, suppose that there are two ways of writing such a vector in terms of a vector in Y and a vector in Z, i.e.

  $$\mathbf{x} = \mathbf{y} + \mathbf{z} = \mathbf{y}' + \mathbf{z}' \text{ for } \mathbf{y}, \mathbf{y}' \in Y \text{ and } \mathbf{z}, \mathbf{z}' \in Z,$$

  as such, we have

  $$\underbrace{\mathbf{y} - \mathbf{y}'}_{\in Y} = \underbrace{\mathbf{z}' - \mathbf{z}}_{\in Z},$$

  that is, using the assumption above we have

  $$\mathbf{y} - \mathbf{y}' = \mathbf{z}' - \mathbf{z} \in Y \cap Z = \{\mathbf{0}\},$$

  Consequently, we can see that $\mathbf{y} - \mathbf{y}' = \mathbf{z}' - \mathbf{z} = \mathbf{0}$, i.e. $\mathbf{y} = \mathbf{y}'$ and $\mathbf{z} = \mathbf{z}'$, and as such the representation is unique (as required).

(d) We are given that Y and Z are subspaces of the vector space V with bases given by $\{\mathbf{y}_1, \ldots, \mathbf{y}_r\}$ and $\{\mathbf{z}_1, \ldots, \mathbf{z}_s\}$ respectively. So, in order to prove that

$V = Y \oplus Z$ iff $\{\mathbf{y}_1, \ldots, \mathbf{y}_r, \mathbf{z}_1, \ldots, \mathbf{z}_s\}$ is a basis of V.

we have to prove it both ways, i.e.

- **LTR:** Given that $V = Y \oplus Z$, we know that every vector in V can be written uniquely in terms of a vector in Y and a vector in Z. But, every vector in Y can be written uniquely as a linear combination of the vectors in $\{\mathbf{y}_1, \ldots, \mathbf{y}_r\}$ (as it is a basis of Y) and every vector in Z can be written uniquely as a linear combination of the vectors in $\{\mathbf{z}_1, \ldots, \mathbf{z}_s\}$ (as it is a basis of Z). Thus, every vector in V can be written uniquely as a linear combination of vectors in the set $\{\mathbf{y}_1, \ldots, \mathbf{y}_r, \mathbf{z}_1, \ldots, \mathbf{z}_s\}$, i.e. this set is a basis of V (as required).

- **RTL:** Given that $\{\mathbf{y}_1, \ldots, \mathbf{y}_r, \mathbf{z}_1, \ldots, \mathbf{z}_s\}$ is a basis of V we need to establish that $V = Y \oplus Z$, and to do this, we start by noting that

  $$\begin{aligned} V &= \text{Lin}\{\mathbf{y}_1, \ldots, \mathbf{y}_r, \mathbf{z}_1, \ldots, \mathbf{z}_s\} \\ &= \{\alpha_1 \mathbf{y}_1 + \cdots + \alpha_r \mathbf{y}_r + \beta_1 \mathbf{z}_1 + \cdots + \beta_s \mathbf{z}_s | \alpha_1, \ldots, \alpha_r, \beta_1, \ldots, \beta_s \in \mathbb{R}\} \\ &= \{\mathbf{y} + \mathbf{z} | \mathbf{y} \in Y \text{ and } \mathbf{z} \in Z\} \\ \therefore \quad V &= Y + Z \end{aligned}$$

  As such, we only need to establish that the sum is direct and there are two ways of doing this:

  - **Method 1:** (Short) Using part **(c)** we show that every vector in V can be written uniquely in terms of vectors in Y and vectors in Z.

    Clearly, as $\{\mathbf{y}_1, \ldots, \mathbf{y}_r, \mathbf{z}_1, \ldots, \mathbf{z}_s\}$ is a basis of V, every vector in V can be written uniquely as a linear combination of these vectors. As such, every vector in V can be written uniquely as the sum of a vector in Y (a unique linear combination of the vectors in $\{\mathbf{y}_1, \ldots, \mathbf{y}_r\}$, a basis of Y) and a vector in Z (a unique linear combination of the vectors in $\{\mathbf{z}_1, \ldots, \mathbf{z}_s\}$, a basis of Z), as required.

  - **Method 2:** (Longer) Using part **(c)** we show that $Y \cap Z = \{\mathbf{0}\}$.

    Take any vector $\mathbf{u} \in Y \cap Z$, i.e. $\mathbf{u} \in Y$ and $\mathbf{u} \in Z$, and so we can write this vector as

    $$\mathbf{u} = \sum_{i=1}^{r} \alpha_i \mathbf{y}_i \text{ and } \mathbf{u} = \sum_{i=1}^{s} \beta_i \mathbf{z}_i.$$

But, since these two linear combinations of vectors represent the same vector, we can equate them. So doing this and rearranging, we get

$$\alpha_1 \mathbf{y}_1 + \cdots + \alpha_r \mathbf{y}_r - \beta_1 \mathbf{z}_1 - \cdots - \beta_s \mathbf{z}_s = \mathbf{0}.$$

But, since this vector equation involves the vectors which form a basis of V, these vectors are linearly independent and, as such, the *only* solution to this vector equation is the trivial solution, i.e. $\alpha_1 = \cdots = \alpha_r = \beta_1 = \cdots = \beta_s = 0$. Consequently, the vector $\mathbf{u}$ considered above must be $\mathbf{0}$, and so $Y \cap Z = \{\mathbf{0}\}$ (as required).

Hence, we are asked to establish that if $V = Y \oplus Z$, then $\dim(V) = \dim(Y) + \dim(Z)$. But, this is obvious since using the bases given above we have:

$$\dim(V) = r + s = \dim(Y) + \dim(Z),$$

as required.

Question 4.

(a) Let A be a real square matrix where $\mathbf{x}$ is an eigenvector of A with a corresponding eigenvalue of λ, i.e. $\mathsf{A}\mathbf{x} = \lambda\mathbf{x}$. We are asked to prove that:

i. $\mathbf{x}$ is an eigenvector of the identity matrix I with a corresponding eigenvalue of one.

Proof: Clearly, as $\mathbf{x}$ is an eigenvector, $\mathbf{x} \neq \mathbf{0}$, and so we have $\mathsf{I}\mathbf{x} = \mathbf{x} = 1 \cdot \mathbf{x}$. Thus, $\mathbf{x}$ is an eigenvector of the identity matrix I with a corresponding eigenvalue of one, as required.

ii. $\mathbf{x}$ is an eigenvector of $\mathsf{A} + \mathsf{I}$ with a corresponding eigenvalue of $\lambda + 1$.

Proof: Clearly, using the information about A and I given above, we have

$$(\mathsf{A} + \mathsf{I})\mathbf{x} = \mathsf{A}\mathbf{x} + \mathsf{I}\mathbf{x} = \lambda\mathbf{x} + 1\mathbf{x} = (\lambda + 1)\mathbf{x}.$$

Thus, $\mathbf{x}$ is an eigenvector of $\mathsf{A} + \mathsf{I}$ with a corresponding eigenvalue of $\lambda + 1$, as required.

iii. $\mathbf{x}$ is an eigenvector of A^2 with a corresponding eigenvalue of λ^2.

Proof: Clearly, using the information about A given above, we have

$$\mathsf{A}^2\mathbf{x} = \mathsf{A}(\lambda\mathbf{x}) = \lambda(\mathsf{A}\mathbf{x}) = \lambda(\lambda\mathbf{x}) = \lambda^2\mathbf{x}.$$

Thus, $\mathbf{x}$ is an eigenvector of A^2 with a corresponding eigenvalue of λ^2, as required.

As such, if the matrix A has eigenvalues λ and corresponding eigenvectors $\mathbf{x}$, for $k \in \mathbb{N}$, it should be clear that the matrix A^k has eigenvalues λ^k and corresponding eigenvectors $\mathbf{x}$.

(b) We are given that the $n \times n$ invertible matrix A has a spectral decomposition given by

$$\mathsf{A} = \sum_{i=1}^{n} \lambda_i \mathbf{x}_i \mathbf{x}_i^\dagger,$$

where, for $1 \leq i \leq n$, the $\mathbf{x}_i$ are an orthonormal set of eigenvectors corresponding to the eigenvalues λ_i of A. So, using **1(iii)**, it should be clear that

$$\mathsf{A}^2 = \sum_{i=1}^{n} \lambda_i^2 \mathbf{x}_i \mathbf{x}_i^\dagger,$$

and, for $k \in \mathbb{N}$, we can see that

$$\mathsf{A}^k = \sum_{i=1}^{n} \lambda_i^k \mathbf{x}_i \mathbf{x}_i^\dagger,$$

(c) We are given that the eigenvalues of the matrix A are all real and such that $-1 < \lambda_i \leq 1$ for $1 \leq i \leq n$. As such, we can re-write the definition of the natural logarithm of the matrix $\mathsf{I} + \mathsf{A}$ as follows:

$$\begin{aligned}
\ln(\mathsf{I} + \mathsf{A}) &= \sum_{k=1}^{\infty} \frac{(-1)^{k+1}}{k} \mathsf{A}^k \\
&= \sum_{k=1}^{\infty} \frac{(-1)^{k+1}}{k} \left[\sum_{i=1}^{n} \lambda_i^k \mathbf{x}_i \mathbf{x}_i^\dagger \right] \\
&= \sum_{i=1}^{n} \left[\sum_{k=1}^{\infty} \frac{(-1)^{k+1}}{k} \lambda_i^k \right] \mathbf{x}_i \mathbf{x}_i^\dagger \\
\therefore \ln(\mathsf{I} + \mathsf{A}) &= \sum_{i=1}^{n} \ln(1 + \lambda_i) \mathbf{x}_i \mathbf{x}_i^\dagger,
\end{aligned}$$

where we have used the Taylor series for $\ln(1+x)$ to establish that, for $1 \leq i \leq n$,

$$\sum_{k=1}^{\infty} \frac{(-1)^{k+1}}{k} \lambda_i^k = \ln(1+\lambda_i),$$

since it was stipulated that the eigenvalues are real and such that $-1 < \lambda_i \leq 1$.

(d) We are given the matrix

$$\mathsf{A} = \frac{1}{4} \begin{bmatrix} -1 & 0 & -1 \\ 0 & 4 & 0 \\ -1 & 0 & -1 \end{bmatrix},$$

and to find its spectral decomposition, we need to find an orthonormal set of eigenvectors and the corresponding eigenvalues. For the eigenvalues, we solve the equation $\det(\mathsf{A} - \lambda \mathsf{I}) = 0$, i.e.

$$\begin{vmatrix} -1-4\lambda & 0 & -1 \\ 0 & 4-4\lambda & 0 \\ -1 & 0 & -1-4\lambda \end{vmatrix} = 0 \implies (4-4\lambda)[(1+4\lambda)^2 - 1] = 0 \implies (1-\lambda)\lambda(1+2\lambda) = 0,$$

and so the eigenvalues are $-1/2, 0, 1$. The eigenvectors that correspond to these eigenvalues are then given by solving the matrix equation $(\mathsf{A} - \lambda \mathsf{I})\mathbf{x} = 0$ for each eigenvalue:

- For $\lambda = -1/2$ we have:

$$\frac{1}{4} \begin{bmatrix} 1 & 0 & -1 \\ 0 & 6 & 0 \\ -1 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \mathbf{0} \implies \begin{aligned} x - z &= 0 \\ 6y &= 0 \\ -x + z &= 0 \end{aligned},$$

  i.e. $x = z$ for $z \in \mathbb{R}$ and $y = 0$. Thus, a corresponding eigenvector is $[1, 0, 1]^t$.

- For $\lambda = 0$ we have:

$$\frac{1}{4} \begin{bmatrix} -1 & 0 & -1 \\ 0 & 4 & 0 \\ -1 & 0 & -1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \mathbf{0} \implies \begin{aligned} -x - z &= 0 \\ 4y &= 0 \\ -x - z &= 0 \end{aligned},$$

  i.e. $x = -z$ for $z \in \mathbb{R}$ and $y = 0$. Thus, a corresponding eigenvector is $[-1, 0, 1]^t$.

- For $\lambda = 1$ we have:

$$\frac{1}{4} \begin{bmatrix} -5 & 0 & -1 \\ 0 & 0 & 0 \\ -1 & 0 & -5 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \mathbf{0} \implies \begin{aligned} -5x - z &= 0 \\ 0 &= 0 \\ -x - 5z &= 0 \end{aligned},$$

  i.e. $x = z = 0$ and $y \in \mathbb{R}$. Thus, a corresponding eigenvector is $[0, 1, 0]^t$.

So, to find the spectral decomposition of A, we need to find an orthonormal set of eigenvectors, i.e.

$$\left\{ \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \right\} \text{ becomes } \left\{ \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}, \frac{1}{\sqrt{2}} \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \right\},$$

(since the eigenvectors are already mutually orthogonal) and substitute all of this into the expression in **(c)**. Thus, since the $\lambda = 0$ term vanishes, we have:

$$\mathsf{A} = -\frac{1}{2} \begin{bmatrix} \frac{1}{\sqrt{2}} \\ 0 \\ \frac{1}{\sqrt{2}} \end{bmatrix} \begin{bmatrix} \frac{1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}} \end{bmatrix} + 1 \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \begin{bmatrix} 0 & 1 & 0 \end{bmatrix}$$

$$= -\frac{1}{2} \begin{bmatrix} \frac{1}{2} & 0 & \frac{1}{2} \\ 0 & 0 & 0 \\ \frac{1}{2} & 0 & \frac{1}{2} \end{bmatrix} + 1 \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$

is the spectral decomposition of A.

Consequently, using part **(c)**, we can see that $\ln(\mathsf{I}+\mathsf{A})$ is given by the matrix

$$\ln(\mathsf{I}+\mathsf{A}) = \ln\left(\frac{1}{2}\right)\begin{bmatrix} \frac{1}{2} & 0 & \frac{1}{2} \\ 0 & 0 & 0 \\ \frac{1}{2} & 0 & \frac{1}{2} \end{bmatrix} + \ln 2 \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix} = \frac{\ln 2}{2}\begin{bmatrix} -1 & 0 & -1 \\ 0 & 2 & 0 \\ -1 & 0 & -1 \end{bmatrix},$$

since $\ln(1/2) = -\ln 2$.

Question 5.

(a) A *strong generalised inverse* of an $m \times n$ matrix A is any $n \times m$ matrix A^G which is such that:

- $\mathsf{A}\mathsf{A}^G\mathsf{A} = \mathsf{A}$.
- $\mathsf{A}^G\mathsf{A}\mathsf{A}^G = \mathsf{A}^G$.
- $\mathsf{A}\mathsf{A}^G$ orthogonally projects $\mathbb{R}^m$ onto $R(\mathsf{A})$.
- $\mathsf{A}^G\mathsf{A}$ orthogonally projects $\mathbb{R}^n$ parallel to $N(\mathsf{A})$.

(b) Given that the matrix equation $\mathsf{A}\mathbf{x} = \mathbf{b}$ represents a set of m inconsistent equations in n variables, we can see that any vector of the form

$$\mathbf{x} = \mathsf{A}^G\mathbf{b} + (\mathsf{I} - \mathsf{A}^G\mathsf{A})\mathbf{w},$$

with $\mathbf{w} \in \mathbb{R}^n$ is a solution to the least squares fit problem associated with this set of linear equations since

$$\mathsf{A}\mathbf{x} = \mathsf{A}\mathsf{A}^G\mathbf{b} + (\mathsf{A} - \mathsf{A}\mathsf{A}^G\mathsf{A})\mathbf{w} = \mathsf{A}\mathsf{A}^G\mathbf{b} + (\mathsf{A} - \mathsf{A})\mathbf{w} = \mathsf{A}\mathsf{A}^G\mathbf{b},$$

using the first property of A^G given in **(a)**. As such, using the third property of A^G given in **(a)**, we can see that $\mathsf{A}\mathbf{x}$ is equal to $\mathsf{A}\mathsf{A}^G\mathbf{b}$, the orthogonal projection of $\mathbf{b}$ onto $R(\mathsf{A})$. But, by definition, a least squares analysis of an inconsistent set of equations $\mathsf{A}\mathbf{x} = \mathbf{b}$ minimises the distance (i.e. $\|\mathsf{A}\mathbf{x} - \mathbf{b}\|$) between the vector $\mathbf{b}$ and $R(\mathsf{A})$, which is exactly what this orthogonal projection does. Thus, the given vector is such a solution.

(c) To find a strong generalised inverse of the matrix

$$\mathsf{A} = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{bmatrix},$$

we use the given method. This is done by noting that the first column vector of A is linearly dependent on the other two since

$$\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} + \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix},$$

and so the matrix A is of rank 2 (as the other two column vectors are linearly independent). Thus, taking $k = 2$, the matrices B and C are given by:

$$\mathsf{B} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 0 \end{bmatrix} \quad \text{and} \quad \mathsf{C} = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix},$$

respectively. So to find the strong generalised inverse, we note that:

$$\mathsf{B}^t\mathsf{B} = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix} \implies (\mathsf{B}^t\mathsf{B})^{-1} = \frac{1}{2}\begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix},$$

and,

$$\mathsf{C}\mathsf{C}^t = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 1 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix} \implies (\mathsf{C}\mathsf{C}^t)^{-1} = \frac{1}{3}\begin{bmatrix} 2 & -1 \\ -1 & 2 \end{bmatrix}.$$

Thus, since

$$(\mathsf{B}^t\mathsf{B})^{-1}\mathsf{B}^t = \frac{1}{2}\begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix} \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix} = \frac{1}{2}\begin{bmatrix} 1 & 0 & 1 \\ 0 & 2 & 0 \end{bmatrix},$$

and,

$$\mathsf{C}^t(\mathsf{CC}^t)^{-1} = \frac{1}{3}\begin{bmatrix} 1 & 1 \\ 1 & 0 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} 2 & -1 \\ -1 & 2 \end{bmatrix} = \frac{1}{3}\begin{bmatrix} 1 & 1 \\ 2 & -1 \\ -1 & 2 \end{bmatrix},$$

we have,

$$\mathsf{A}^G = \mathsf{C}^t(\mathsf{CC}^t)^{-1}(\mathsf{B}^t\mathsf{B})^{-1}\mathsf{B}^t = \frac{1}{6}\begin{bmatrix} 1 & 1 \\ 2 & -1 \\ -1 & 2 \end{bmatrix}\begin{bmatrix} 1 & 0 & 1 \\ 0 & 2 & 0 \end{bmatrix} = \frac{1}{6}\begin{bmatrix} 1 & 2 & 1 \\ 2 & -2 & 2 \\ -1 & 4 & -1 \end{bmatrix},$$

which is the sought after strong generalised inverse of A. So to find the set of all solutions to the least squares fit problem associated with the system of linear equations given by

$$\begin{aligned} x + y &= 1 \\ x + z &= 2 \\ x + y &= -1 \end{aligned}$$

we note that these equations can be written in the form $\mathsf{A}\mathbf{x} = \mathbf{b}$ with

$$\mathsf{A} = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{bmatrix} \quad \mathbf{x} = \begin{bmatrix} x \\ y \\ z \end{bmatrix} \quad \text{and} \quad \mathbf{b} = \begin{bmatrix} 1 \\ 2 \\ -1 \end{bmatrix}.$$

Thus, as

$$\mathsf{A}^G\mathsf{A} = \frac{1}{6}\begin{bmatrix} 1 & 2 & 1 \\ 2 & -2 & 2 \\ -1 & 4 & -1 \end{bmatrix}\begin{bmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{bmatrix} = \frac{1}{6}\begin{bmatrix} 4 & 2 & 2 \\ 2 & 4 & -2 \\ 2 & -2 & 4 \end{bmatrix} = \frac{1}{3}\begin{bmatrix} 2 & 1 & 1 \\ 1 & 2 & -1 \\ 1 & -1 & 2 \end{bmatrix},$$

we have

$$\mathsf{A}^G\mathsf{A} - \mathsf{I} = \frac{1}{3}\left\{\begin{bmatrix} 2 & 1 & 1 \\ 1 & 2 & -1 \\ 1 & -1 & 2 \end{bmatrix} - \begin{bmatrix} 3 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 3 \end{bmatrix}\right\} = \frac{1}{3}\begin{bmatrix} -1 & 1 & 1 \\ 1 & -1 & -1 \\ 1 & -1 & -1 \end{bmatrix},$$

and

$$\mathsf{A}^G\mathbf{b} = \frac{1}{6}\begin{bmatrix} 1 & 2 & 1 \\ 2 & -2 & 2 \\ -1 & 4 & -1 \end{bmatrix}\begin{bmatrix} 1 \\ 2 \\ -1 \end{bmatrix} = \frac{1}{6}\begin{bmatrix} 4 \\ -4 \\ 8 \end{bmatrix} = \frac{1}{3}\begin{bmatrix} 2 \\ -2 \\ 4 \end{bmatrix}.$$

So, the solutions of this system of linear equations are given by

$$\mathbf{x} = \frac{1}{3}\left\{\begin{bmatrix} 2 \\ -2 \\ 4 \end{bmatrix} + \begin{bmatrix} -1 & 1 & 1 \\ 1 & -1 & -1 \\ 1 & -1 & -1 \end{bmatrix}\mathbf{w}\right\},$$

for any $\mathbf{w} \in \mathbb{R}^3$.

We know from parts **(a)** and **(b)** that the $(\mathsf{A}^g\mathsf{A} - \mathsf{I})\mathbf{w}$ part of our solutions will yield a vector in $N(\mathsf{A})$ and so the solution set is the translate of $N(\mathsf{A})$ by $\frac{1}{3}[2, -2, 4]^t$. Further, by the rank-nullity theorem, we have $\eta(\mathsf{A}) = 3 - \rho(\mathsf{A}) = 3 - 2 = 1$ and so $N(\mathsf{A})$ will be a line through the origin. Indeed, the vector equation of the line representing the solution set is

$$\mathbf{x} = \frac{2}{3}\begin{bmatrix} 1 \\ -1 \\ 2 \end{bmatrix} + \lambda\begin{bmatrix} -1 \\ 1 \\ 1 \end{bmatrix},$$

where $\lambda \in \mathbb{R}$.

Question 6

(a) To test the set of functions $\{1, x, x^2\}$ for linear independence we calculate the Wronskian as instructed, i.e.

$$W(x) = \begin{vmatrix} f_1 & f_2 & f_3 \\ f_1' & f_2' & f_3' \\ f_1'' & f_2'' & f_3'' \end{vmatrix} = \begin{vmatrix} 1 & x & x^2 \\ 0 & 1 & 2x \\ 0 & 0 & 2 \end{vmatrix} = 2,$$

and as $W(x) \neq 0$ the set is linearly independent (as required).

(b) We consider the inner product space formed by the vector space $\mathbb{P}_3^{[-2,2]}$ and the inner product

$$\langle f(x), g(x) \rangle = \int_{-1}^{1} f(x)g(x)dx.$$

To find an orthonormal basis of the space $\text{Lin}\{1, x, x^2\}$, we use the Gram-Schmidt procedure:

- We start with the vector 1, and note that

$$\|1\|^2 = \langle 1, 1 \rangle = \int_{-1}^{1} 1dx = [x]_{-1}^{1} = 2.$$

  Consequently, we set $\mathbf{e}_1 = 1/\sqrt{2}$.

- We need to find a vector $\mathbf{u}_2$ where

$$\mathbf{u}_2 = x - \left\langle x, \frac{1}{\sqrt{2}} \right\rangle \frac{1}{\sqrt{2}} = x - \frac{1}{2}\langle x, 1 \rangle,$$

  But, as

$$\langle x, 1 \rangle = \int_{-1}^{1} xdx = \left[\frac{x^2}{2}\right]_{-1}^{1} = 0,$$

  we have $\mathbf{u}_2 = x$. Now, as

$$\|x\|^2 = \langle x, x \rangle = \int_{-1}^{1} x^2dx = \left[\frac{x^3}{3}\right]_{-1}^{1} = \frac{2}{3},$$

  we set $\mathbf{e}_2 = \sqrt{\frac{3}{2}}\, x$.

- Lastly, we need to find a vector $\mathbf{u}_3$ where

$$\begin{aligned} \mathbf{u}_3 &= x^2 - \left\langle x^2, \sqrt{\frac{3}{2}}x \right\rangle \sqrt{\frac{3}{2}}x - \left\langle x^2, \frac{1}{\sqrt{2}} \right\rangle \frac{1}{\sqrt{2}} \\ &= x^2 - \frac{3}{2}\langle x^2, x \rangle x - \frac{1}{2}\langle x^2, 1 \rangle, \end{aligned}$$

  But, as

$$\langle x^2, x \rangle = \int_{-1}^{1} x^3dx = \left[\frac{x^4}{4}\right]_{-1}^{1} = 0,$$

  and,

$$\langle x^2, 1 \rangle = \int_{-1}^{1} x^2dx = \left[\frac{x^3}{3}\right]_{-1}^{1} = \frac{2}{3},$$

  we have $\mathbf{u}_3 = x^2 - 1/3 = (3x^2 - 1)/3$. Now, as

$$\begin{aligned} \|3x^2 - 1\|^2 = \langle 3x^2 - 1, 3x^2 - 1 \rangle &= \int_{-1}^{1} (3x^2 - 1)^2dx = \int_{-1}^{1} (9x^4 - 6x^2 + 1)dx \\ &= \left[\frac{9}{5}x^5 - 2x^3 + x\right]_{-1}^{1} = 2\left[\frac{9}{5} - 2 + 1\right] = 2\left[\frac{4}{5}\right] = \frac{8}{5} \end{aligned}$$

  we set $\mathbf{e}_3 = \sqrt{\frac{5}{8}}\,(3x^2 - 1)$.

Consequently, the set

$$\left\{ \frac{1}{\sqrt{2}}, \sqrt{\frac{3}{2}}\, x, \sqrt{\frac{5}{8}}\, (3x^2 - 1) \right\},$$

is an orthonormal basis for the space $\text{Lin}\{1, x, x^2\}$.

(c) We are given that $\{\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_k\}$ is an orthonormal basis of a subspace S of an inner product space V. Extending this to an orthonormal basis $\{\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_k, \mathbf{e}_{k+1}, \ldots, \mathbf{e}_n\}$ of V, we note that for any vector $\mathbf{x} \in V$,

$$\mathbf{x} = \sum_{i=1}^{n} \alpha_i \mathbf{e}_i.$$

Now, for any j (where $1 \leq j \leq n$) we have

$$\langle \mathbf{x}, \mathbf{e}_j \rangle = \left\langle \sum_{i=1}^{n} \alpha_i \mathbf{e}_i, \mathbf{e}_j \right\rangle = \sum_{i=1}^{n} \alpha_i \langle \mathbf{e}_i, \mathbf{e}_j \rangle = \alpha_j,$$

since we are using an orthonormal basis. Thus, we can write

$$\mathbf{x} = \sum_{i=1}^{n} \langle \mathbf{x}, \mathbf{e}_i \rangle \mathbf{e}_i = \underbrace{\sum_{i=1}^{k} \langle \mathbf{x}, \mathbf{e}_i \rangle \mathbf{e}_i}_{\text{in } S} + \underbrace{\sum_{i=k+1}^{n} \langle \mathbf{x}, \mathbf{e}_i \rangle \mathbf{e}_i}_{\text{in } S^{\perp}}.$$

and so, the orthogonal projection of V onto S [parallel to $S^{\perp}$] is given by

$$\mathsf{P}\mathbf{x} = \sum_{i=1}^{k} \langle \mathbf{x}, \mathbf{e}_i \rangle \mathbf{e}_i,$$

for any $\mathbf{x} \in V$ (as required).

(d) Using the result in **(c)**, it should be clear that a least squares approximation to x^4 in $\text{Lin}\{1, x, x^2\}$ will be given by $\mathsf{P}x^4$. So, using the inner product in **(b)** and the orthonormal basis for $\text{Lin}\{1, x, x^2\}$ which we found there, we have:

$$\begin{aligned}
\mathsf{P}x^4 &= \langle x^4, \mathbf{e}_1 \rangle \mathbf{e}_1 + \langle x^4, \mathbf{e}_2 \rangle \mathbf{e}_2 + \langle x^4, \mathbf{e}_3 \rangle \mathbf{e}_3 \\
&= \frac{1}{2}\langle x^4, 1 \rangle + \frac{3}{2}\langle x^4, x \rangle x + \frac{5}{8}\langle x^4, 3x^2 - 1 \rangle (3x^2 - 1) \\
&= \frac{1}{5} + \frac{2}{7}(3x^2 - 1) \\
\therefore\ \mathsf{P}x^4 &= \frac{3}{35}(10x^2 - 1)
\end{aligned}$$

since,

$$\begin{aligned}
\langle x^4, 1 \rangle &= \int_{-1}^{1} x^4 dx = \left[\frac{x^5}{5} \right]_{-1}^{1} = \frac{2}{5}, \\
\langle x^4, x \rangle &= \int_{-1}^{1} x^5 dx = \left[\frac{x^6}{6} \right]_{-1}^{1} = 0 \\
\langle x^4, 3x^2 - 1 \rangle &= \int_{-1}^{1} x^4 (3x^2 - 1) dx = \int_{-1}^{1} (3x^6 - x^4) dx = \left[\frac{3}{7}x^7 - \frac{x^5}{5} \right]_{-1}^{1} = 2\left[\frac{3}{7} - \frac{1}{5} \right] = 2\left[\frac{8}{35} \right] = \frac{16}{35}
\end{aligned}$$

Thus, our least squares approximation to x^4 is $\frac{3}{35}(10x^2 - 1)$.