Further Mathematical Methods (Linear Algebra)

Solutions For The 2001 Examination

Question 1

(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:** $\alpha\mathbf{x} \in W$.

To be an inner product on 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) Clearly, for any vector $\mathbf{u} \in \mathbb{P}_2^{[0,1]}$, we have

$$\mathbf{f} = \sum_{i=0}^{2} a_i \mathbf{x}^{\mathbf{i}} = \sum_{i=0}^{n} a_i \mathbf{x}^{\mathbf{i}},$$

where, for $3 \leq i \leq n$, we have $a_i = 0$. Thus, $\mathbf{u} \in \mathbb{P}_n^{[0,1]}$ too and so $\mathbb{P}_2^{[0,1]}$ is a sub*set* of $\mathbb{P}_2^{[0,1]}$. To show that it is a sub*space*, we take any two vectors in $\mathbb{P}_2^{[0,1]}$, say

$$\mathbf{f} = \sum_{i=0}^{2} a_i \mathbf{x}^{\mathbf{i}} \quad \text{and} \quad \mathbf{g} = \sum_{i=0}^{2} b_i \mathbf{x}^{\mathbf{i}},$$

and any scalar $\alpha \in \mathbb{R}$ and note that $\mathbb{P}_2^{[0,1]}$ is closed under:

- **vector addition** since

$$\mathbf{f} + \mathbf{g} = \sum_{i=0}^{2} a_i \mathbf{x}^{\mathbf{i}} + \sum_{i=0}^{2} b_i \mathbf{x}^{\mathbf{i}} = \sum_{i=0}^{2} (a_i + b_i) \mathbf{x}^{\mathbf{i}},$$

  and so $\mathbf{f} + \mathbf{g} \in \mathbb{P}_2^{[0,1]}$ too since $a_i + b_i \in \mathbb{R}$ for $0 \leq i \leq 2$.

- **scalar multiplication** since

$$\alpha\mathbf{f} = \sum_{i=0}^{2} \alpha a_i \mathbf{x}^{\mathbf{i}},$$

  and so $\alpha\mathbf{g} \in \mathbb{P}_2^{[0,1]}$ too since $\alpha a_i \in \mathbb{R}$ for $0 \leq i \leq 2$.

as required.

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

$$\langle \mathbf{f}, \mathbf{g} \rangle = a_0 b_0 + a_1 b_1 + a_2 b_2,$$

defines an inner product on $\mathbb{P}_2^{[0,1]}$. To do this, we show that this formula satisfies all of the conditions given in part **(a)**. Thus, taking any three vectors $\mathbf{f}$, $\mathbf{g}$ and $\mathbf{h}$ in $\mathbb{P}_2^{[0,1]}$ and any two scalars α and β in $\mathbb{R}$ we have:

i. $\langle \mathbf{f}, \mathbf{f} \rangle = a_0^2 + a_1^2 + a_2^2$ which is the sum of the squares of three real numbers and as such it is real and non-negative. Further, to show that $\langle \mathbf{f}, \mathbf{f} \rangle = 0$ if and only if $\mathbf{f} = \mathbf{0}$ (where here, $\mathbf{0}$ is the zero polynomial), we note that:

- **LTR:** If $\langle \mathbf{f}, \mathbf{f} \rangle = 0$, then $a_0^2 + a_1^2 + a_2^2 = 0$. But, this is the sum of the squares of three real numbers and so it must be the case that $a_0 = a_1 = a_2 = 0$. Thus, $\mathbf{f} = \mathbf{0}$.
- **RTL:** If $\mathbf{f} = \mathbf{0}$, then $a_0 = a_1 = a_2 = 0$. Thus, $\langle \mathbf{f}, \mathbf{f} \rangle = 0$.

(as required).

ii. Obviously, $\langle \mathbf{f}, \mathbf{g} \rangle = a_0 b_0 + a_1 b_1 + a_2 b_2 = b_0 a_0 + b_1 a_1 + b_2 a_2 = \langle \mathbf{f}, \mathbf{g} \rangle$.

iii. We note that the vector $\alpha \mathbf{f} + \beta \mathbf{g}$ is just another quadratic and so:

$$\begin{aligned} \langle \alpha \mathbf{f} + \beta \mathbf{g}, \mathbf{h} \rangle &= (\alpha a_0 + \beta b_0) c_0 + (\alpha a_1 + \beta b_1) c_1 + (\alpha a_2 + \beta b_2) c_2 \\ &= \alpha (a_0 c_0 + a_1 c_1 + a_2 c_2) + \beta (b_0 c_0 + b_1 c_1 + b_2 c_2) \\ \therefore \quad \langle \alpha \mathbf{f} + \beta \mathbf{g}, \mathbf{h} \rangle &= \alpha \langle \mathbf{f}, \mathbf{h} \rangle + \beta \langle \mathbf{g}, \mathbf{h} \rangle \end{aligned}$$

where $\mathbf{h} : x \to c_0 + c_1 x + c_2 x^2$ for all $x \in \mathbb{R}$.

Consequently, the formula given above does define an inner product on $\mathbb{P}_2^{[0,1]}$ (as required).

(d) To show that a set of vectors is a basis for $\mathbb{P}_2^{[0,1]}$ we have to show that it spans this space and that it is linearly independent. Thus,

- For $S = \{\mathbf{1}, \mathbf{x}, \mathbf{x^2}\}$: Firstly, these vectors span the space as,

$$\text{Lin}(S) = \{a_0 \mathbf{1} + a_1 \mathbf{x} + a_2 \mathbf{x^2} \mid \text{ for all } a_0, a_1, a_2 \in \mathbb{R}\} = \mathbb{P}_2^{[0,1]},$$

They are also linearly independent since calculating the Wronskian for these vectors we have,

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

and this is non-zero for all $x \in [0, 1]$.

- For $S' = \{\mathbf{1} + \mathbf{x}, \mathbf{1} - \mathbf{x}, \mathbf{x} + \mathbf{x^2}\}$: Firstly, these vectors span the space as,

$$\begin{aligned} \text{Lin}(S') &= \{a_0(\mathbf{1} + \mathbf{x}) + a_1(\mathbf{1} - \mathbf{x}) + a_2(\mathbf{x} + \mathbf{x^2}) \mid \text{ for all } a_0, a_1, a_2 \in \mathbb{R}\} \\ &= \{(a_0 + a_1)\mathbf{1} + (a_0 - a_1 + a_2)\mathbf{x} + a_2 \mathbf{x^2} \mid \text{ for all } a_0, a_1, a_2 \in \mathbb{R}\} \\ &= \{b_0 \mathbf{1} + b_1 \mathbf{x} + b_2 \mathbf{x^2} \mid \text{ for all } b_0, b_1, b_2 \in \mathbb{R}\} \\ &= \mathbb{P}_2^{[0,1]}, \end{aligned}$$

They are also linearly independent since calculating the Wronskian for these vectors we have, [expanding along the bottom row]

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

and this is non-zero for all $x \in [0, 1]$.

Consequently, the sets S and S' are both bases of $\mathbb{P}_2^{[0,1]}$.

(e) To find a matrix A such that

$$[\mathbf{f}]_S = \mathsf{A}[\mathbf{f}]_{S'},$$

where $[\mathbf{f}]_S$ and $[\mathbf{f}]_{S'}$ are the coordinate vectors of $\mathbf{f} \in \mathbb{P}_2^{[0,1]}$ relative to the bases S and S' respectively we use the definition of coordinate vector. That is, we use the fact that the equality

$$a\mathbf{1} + b\mathbf{x} + c\mathbf{x^2} = a'(\mathbf{1} + \mathbf{x}) + b'(\mathbf{1} - \mathbf{x}) + c'(\mathbf{x} + \mathbf{x^2}),$$

holds if and only if

$$\begin{bmatrix} a \\ b \\ c \end{bmatrix}_S = \mathsf{A} \begin{bmatrix} a' \\ b' \\ c' \end{bmatrix}_{S'},$$

where

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

is the required matrix.

Question 2.

(a) The Leslie matrix for the unicorn population is given by:

$$\mathsf{L} = \begin{bmatrix} 0 & 1 & 2 \\ 1/18 & 0 & 0 \\ 0 & 1/6 & 0 \end{bmatrix},$$

and to verify that its unique real positive eigenvalue is 1/3, we use the result given in the lectures to find an eigenvector corresponding to this *purported* eigenvalue, i.e.

$$\mathbf{v}_1 = \begin{bmatrix} 1 \\ b_1/\lambda_1 \\ b_1 b_2/\lambda_1^2 \end{bmatrix} = \begin{bmatrix} 1 \\ 1/6 \\ 1/12 \end{bmatrix} \quad \text{and so we take} \quad \mathbf{v}_1 = \begin{bmatrix} 12 \\ 2 \\ 1 \end{bmatrix}.$$

Then, as

$$\mathsf{L}\mathbf{v}_1 = \begin{bmatrix} 0 & 1 & 2 \\ 1/18 & 0 & 0 \\ 0 & 1/6 & 0 \end{bmatrix} \begin{bmatrix} 12 \\ 2 \\ 1 \end{bmatrix} = \begin{bmatrix} 4 \\ 2/3 \\ 1/3 \end{bmatrix} = \frac{1}{3}\mathbf{v}_1,$$

1/3 is indeed an eigenvalue of L. Thus, noting that:

i. the population distribution vector $\mathbf{x}^{(k)}$ behaves as

$$\mathbf{x}^{(k)} \simeq c\lambda_1^k \mathbf{v}_1 = \frac{c}{3^k} \begin{bmatrix} 12 \\ 2 \\ 1 \end{bmatrix},$$

(for some constant c) in the long-term.

ii. the proportion of the population in each of the three age classes becomes constant in the ratio 12:2:1 in the long-term.

iii. the growth rate of the population in each age class is 1/3, i.e. the population in each age class decreases by $66\frac{2}{3}\%$ every time period (i.e. every ten years), in the long-term.

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

$$\dot{y}_1 = 3y_1 - y_1^2 - 6y_1 y_2$$
$$\dot{y}_2 = 3y_2 - y_2^2 - 2y_1 y_2$$

are given by the solutions of the simultaneous equations

$$y_1(3 - y_1 - 6y_2) = 0$$
$$y_2(3 - y_2 - 2y_1) = 0$$

i.e. by $(y_1, y_2) = (0,0), (0,3), (3,0)$ and $(\frac{15}{11}, \frac{3}{11})$.

To assess the stability of the steady state given by $(y_1, y_2) = (0, 3)$, we evaluate the Jacobian for this system at $(0, 3)$, i.e.

$$\mathsf{DF}[(y_1, y_2)] = \begin{bmatrix} 3 - 2y_1 - 6y_2 & -6y_1 \\ -2y_2 & 3 - 2y_2 - 2y_1 \end{bmatrix} \implies \mathsf{DF}[(0,3)] = \begin{bmatrix} -15 & 0 \\ -6 & -3 \end{bmatrix},$$

and find the eigenvalues of this matrix. So, solving

$$\begin{vmatrix} -15 - \lambda & 0 \\ -6 & -3 - \lambda \end{vmatrix} = 0 \implies (15 + \lambda)(3 + \lambda) = 0,$$

we find that the eigenvalues are $\lambda = -15$ and $\lambda = -3$. According to a result given in the lectures, since these are both real and negative, the steady state $(y_1, y_2) = (0, 3)$ is asymptotically stable.

(c) We are asked to find the general solution of the coupled linear differential equations given by

$$\begin{bmatrix} \dot{h}_1 \\ \dot{h}_2 \end{bmatrix} = \mathsf{DF}[(0,3)] \begin{bmatrix} h_1 \\ h_2 \end{bmatrix},$$

where $\mathsf{DF}[(0,3)]$ is the Jacobian of the system in **(b)** evaluated at the steady state $(y_1, y_2) = (0, 3)$. That is, we just have to solve the coupled linear differential equations given by

$$\begin{bmatrix} \dot{h}_1 \\ \dot{h}_2 \end{bmatrix} = \begin{bmatrix} -15 & 0 \\ -6 & -3 \end{bmatrix} \begin{bmatrix} h_1 \\ h_2 \end{bmatrix}.$$

To do this, we know that the eigenvalues of the matrix are -15 and -3, and we can easily see that the corresponding eigenvectors are $[2, 1]^t$ and $[0, 1]^t$. Thus, setting

$$\mathsf{P} = \begin{bmatrix} 2 & 0 \\ 1 & 1 \end{bmatrix} \quad \text{and} \quad \mathsf{D} = \begin{bmatrix} -15 & 0 \\ 0 & -3 \end{bmatrix},$$

we set $\mathbf{z} = \mathsf{P}^{-1}\mathbf{h}$ so that

$$\dot{\mathbf{h}} = \mathsf{DF}[(0,3)]\mathbf{h} = \mathsf{PDP}^{-1}\mathbf{h} \implies \dot{\mathbf{z}} = \mathsf{D}\mathbf{z},$$

since $\mathsf{DF}[(0,3)] = \mathsf{PDP}^{-1}$. So, we now have to solve the uncoupled linear differential equation given by

$$\dot{z}_1 = -15z_1 \implies \int \frac{dz_1}{z_1} = -15 \int dt \implies \ln z_1 = -15t + c \implies z_1 = Ae^{-15t},$$

for some constants c and A such that $A = e^c$, and similarly,

$$\dot{z}_2 = -3z_2 \ \text{ gives } \ z_2 = Be^{-3t},$$

for some constant B. Thus, the required general solution is

$$\mathbf{h}(t) = \begin{bmatrix} h_1(t) \\ h_2(t) \end{bmatrix} = \begin{bmatrix} 2 & 0 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} Ae^{-15t} \\ Be^{-3t} \end{bmatrix} = \begin{bmatrix} 2Ae^{-15t} \\ Ae^{-15t} + Be^{-3t} \end{bmatrix},$$

since $\mathbf{h} = \mathsf{P}\mathbf{z}$.

So, given that $\mathbf{h}(t)$ is related to $\mathbf{y}(t)$ in **(b)** by

$$\mathbf{h}(t) = \mathbf{y}(t) - \begin{bmatrix} 0 \\ 3 \end{bmatrix} \quad \text{we have} \quad \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 3 \end{bmatrix} + \begin{bmatrix} 2Ae^{-15t} \\ Ae^{-15t} + Be^{-3t} \end{bmatrix},$$

and so a particular solution to this system of coupled linear differential equations using the initial conditions given for $\mathbf{y}(t)$, i.e. $y_1(0) = 1$ and $y_2(0) = 4$, can be found by noting that at $t = 0$,

$$\begin{bmatrix} 1 \\ 4 \end{bmatrix} = \begin{bmatrix} 0 \\ 3 \end{bmatrix} + \begin{bmatrix} 2A \\ A + B \end{bmatrix} \implies \begin{bmatrix} 2A \\ A + B \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \end{bmatrix} \implies \begin{bmatrix} A \\ B \end{bmatrix} = \begin{bmatrix} 1/2 \\ 1/2 \end{bmatrix},$$

That is, the sought after particular solution is

$$\mathbf{h}(t) = \begin{bmatrix} h_1(t) \\ h_2(t) \end{bmatrix} = \frac{1}{2} \begin{bmatrix} 2e^{-15t} \\ e^{-15t} + e^{-3t} \end{bmatrix}.$$

Clearly, in the long term, this means that $\mathbf{h}(t) \to \mathbf{0}$ and so we find that $\mathbf{y}(t) \to [0, 3]^t$, i.e. the unicorn population dies out and the amount of virus in a cubic millimetre of blood approaches 3 according to this model.

Question 3.

(a) For two subspaces Y and Z of $\mathbb{R}^n$, we know that

i. $\mathbb{R}^n$ is the *direct sum* of Y and Z, denoted by $\mathbb{R}^n = Y \oplus Z$, if every vector $\mathbf{v} \in \mathbb{R}^n$ 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.

ii. the matrix P is a *projection* of $\mathbb{R}^n$ onto Y parallel to Z, if $\mathbb{R}^n = Y \oplus Z$ and for every $\mathbf{v} \in \mathbb{R}^n$ we have $\mathsf{P}\mathbf{v} = \mathbf{y} \in Y$.

Further, if the matrix P is a projection of $\mathbb{R}^n$ onto Y parallel to Z, we know that for any $\mathbf{v} \in \mathbb{R}^n$, we can write $\mathbf{v} = \mathbf{y} + \mathbf{z}$ uniquely in terms of vectors $\mathbf{y} \in Y$ and $\mathbf{z} \in Z$ as $\mathbb{R}^n = Y \oplus Z$. So, by definition, for any $\mathbf{v} \in \mathbb{R}^n$, we have $\mathsf{P}\mathbf{v} = \mathbf{y}$ and so

$$\mathsf{P}\mathbf{v} = \mathbf{v} - \mathbf{z} \implies (\mathsf{I} - \mathsf{P})\mathbf{v} = \mathbf{z} \in Z,$$

Thus, $\mathsf{I} - \mathsf{P}$ is a matrix which represents a projection of $\mathbb{R}^n$ onto Z parallel to Y.

(b) Given that Y and Z are subspaces such that $\mathbb{R}^n = Y \oplus Z$, we are asked to prove the following:

i. If P is a projection of $\mathbb{R}^n$ onto Y parallel to Z, then $Y = R(\mathsf{P})$ and $Z = N(\mathsf{P})$.

Proof: We are given that P is a projection of $\mathbb{R}^n$ onto Y parallel to Z and so $\mathbb{R}^n = Y \oplus Z$. To prove that $Y = R(\mathsf{P})$ we note that:

- For any $\mathbf{y} \in Y$ we have $\mathbf{y} = \mathbf{y} + \mathbf{0}$ and so, by the definition of P, $\mathsf{P}\mathbf{y} = \mathbf{y}$. Thus, $\mathbf{y} \in R(\mathsf{P})$ and so $Y \subseteq R(\mathsf{P})$.
- For any $\mathbf{y} \in R(\mathsf{P})$, there exists an $\mathbf{x} \in \mathbb{R}^n$ such that $\mathsf{P}\mathbf{x} = \mathbf{y}$ and so, by the definition of P, $\mathbf{y} \in Y$. Thus, $R(\mathsf{P}) \subseteq Y$.

and for $Z = N(\mathsf{P})$ we note that:

- For any $\mathbf{z} \in Z$ we have $\mathbf{z} = \mathbf{0} + \mathbf{z}$ and so, by the definition of P, $\mathsf{P}\mathbf{z} = \mathbf{0}$. Thus, $\mathbf{z} \in N(\mathsf{P})$ and so $Z \subseteq N(\mathsf{P})$.
- For any $\mathbf{z} \in N(\mathsf{P})$, we have $\mathsf{P}\mathbf{z} = \mathbf{0}$ and so,

$$\mathsf{P}\mathbf{z} = \mathbf{z} - \mathbf{z} \implies \mathbf{z} = (\mathsf{I} - \mathsf{P})\mathbf{z}.$$

But, as we saw in **(a)**, $(\mathsf{I} - \mathsf{P})\mathbf{x} \in Z$ for every $\mathbf{x} \in \mathbb{R}^n$. Thus, $\mathbf{z} \in Z$, and so $N(\mathsf{P}) \subseteq Z$.

Consequently, $Y = R(\mathsf{P})$ and $Z = N(\mathsf{P})$, as required.

ii. P is a projection if and only if P is idempotent.

Proof: This is an 'if and only if' statement and so it has to be proved 'both ways':

LTR: Suppose that P is a projection of X onto Y and so, by **(i)**, $Y = R(\mathsf{P})$. So, for any $\mathbf{y} \in Y$, we have:

$$\mathbf{y} = \mathbf{y} + \mathbf{0} \implies \mathsf{P}\mathbf{y} = \mathbf{y}.$$

Now, for any $\mathbf{x} \in X$, $\mathsf{P}\mathbf{x} \in Y$ and so,

$$\mathsf{P}(\mathsf{P}\mathbf{x}) = \mathsf{P}\mathbf{x} \implies \mathsf{P}^2\mathbf{x} = \mathsf{P}\mathbf{x}.$$

Thus, as this must hold for all $\mathbf{x} \in X$, we have $\mathsf{P}^2 = \mathsf{P}$, i.e. P is idempotent (as required).

RTL: Suppose that P is idempotent, i.e. $\mathsf{P}^2 = \mathsf{P}$. We need to prove that P is a projection, that is, we need to establish that, for some subspaces Y and Z of X such that $X = Y \oplus Z$, P

will map any vector $\mathbf{x} \in X$ to a vector in Y. So, noting the result in **(i)**, we show that $R(\mathsf{P})$ and $N(\mathsf{P})$ are subspaces of X such that $X = R(\mathsf{P}) \oplus N(\mathsf{P})$ and that P will map every vector $\mathbf{x} \in X$ to a vector in $R(\mathsf{P})$.

We know that if $R(\mathsf{P})$ and $N(\mathsf{P})$ are subsets of a vector space X, then they are subspaces of X. Thus, to establish that $X = R(\mathsf{P}) \oplus N(\mathsf{P})$, we use the fact that

$$X = R(\mathsf{P}) \oplus N(\mathsf{P}) \text{ if and only if } X = R(\mathsf{P}) + N(\mathsf{P}) \text{ and } R(\mathsf{P}) \cap N(\mathsf{P}) = \{\mathbf{0}\},$$

So, noting that:

- For any vector $\mathbf{x} \in X$, we can write

  $$\mathbf{x} = \mathsf{P}\mathbf{x} + (\mathbf{x} - \mathsf{P}\mathbf{x}).$$

  Clearly, the vector $\mathsf{P}\mathbf{x} \in R(\mathsf{P})$ and the vector $\mathbf{x} - \mathsf{P}\mathbf{x} \in N(\mathsf{P})$ since

  $$\mathsf{P}(\mathbf{x} - \mathsf{P}\mathbf{x}) = \mathsf{P}\mathbf{x} - \mathsf{P}^2\mathbf{x} = \mathsf{P}\mathbf{x} - \mathsf{P}\mathbf{x} = \mathbf{0},$$

  using the fact that P is idempotent. Thus, $X \subseteq R(\mathsf{P}) + N(\mathsf{P})$. Consequently, as $R(\mathsf{P}) + N(\mathsf{P}) \subseteq X$ too (by the definition of 'sum'), we have $X = R(\mathsf{P}) + N(\mathsf{P})$.
- Let $\mathbf{u}$ be any vector in $R(\mathsf{P}) \cap N(\mathsf{P})$, i.e. $\mathbf{u} \in R(\mathsf{P})$ and $\mathbf{u} \in N(\mathsf{P})$. Now, this means that there is a vector $\mathbf{v} \in X$ such that $\mathsf{P}\mathbf{v} = \mathbf{u}$ and that $\mathsf{P}\mathbf{u} = \mathbf{0}$, so

  $$\mathsf{P}\mathbf{u} = \mathbf{0} \implies \mathsf{P}(\mathsf{P}\mathbf{v}) = \mathbf{0} \implies \mathsf{P}^2\mathbf{v} = \mathbf{0} \implies \mathsf{P}\mathbf{v} = \mathbf{0},$$

  since P is idempotent. Thus, $\mathbf{u} = \mathsf{P}\mathbf{v} = \mathbf{0}$ and so, $R(\mathsf{P}) \cap N(\mathsf{P}) = \{\mathbf{0}\}$.

we can see that $X = R(\mathsf{P}) \oplus N(\mathsf{P})$.

Further, every vector $\mathbf{x} \in X$ is mapped to a vector in $R(\mathsf{P})$ by P since $\mathsf{P}\mathbf{x} \in R(\mathsf{P})$. Consequently, we can see that the idempotent matrix P is a projection (as required).

(c) The matrix P represents a projection of $\mathbb{R}^3$ onto Y parallel to Z where $\dim(Y) = 2$ and $\dim(Z) = 1$. By considering the matrix equation

$$(\mathsf{P} - \lambda\mathsf{I})\mathbf{x} = \mathbf{0},$$

we are asked to find the eigenvalues of the matrix P and the subspaces of $\mathbb{R}^3$ spanned by the eigenvectors corresponding to each of these eigenvalues. To do this, we note that:

- For any $\mathbf{y} \in Y$, we have

  $$\mathsf{P}\mathbf{y} = \mathbf{y} \implies (\mathsf{P} - \mathsf{I})\mathbf{y} = \mathbf{0} \implies (\mathsf{P} - 1\mathsf{I})\mathbf{y} = \mathbf{0},$$

  i.e. if $\mathbf{y} \neq \mathbf{0}$, $\mathbf{y}$ is an eigenvector of P corresponding to the eigenvalue 1. Clearly, the subspace Y is spanned by these eigenvectors and as $\dim(Y) = 2$, this eigenvalue will be of multiplicity 2.

- For any $\mathbf{z} \in Z$, we have

  $$\mathsf{P}\mathbf{z} = \mathbf{0} \implies \mathsf{P}\mathbf{z} - 0\mathbf{z} = \mathbf{0} \implies (\mathsf{P} - 0\mathsf{I})\mathbf{z} = \mathbf{0},$$

  i.e. if $\mathbf{z} \neq \mathbf{0}$, then $\mathbf{z}$ is an eigenvector of P corresponding to the eigenvalue 0. Clearly, the subspace Z is spanned by these eigenvectors and as $\dim(Z) = 1$, this eigenvalue will be of multiplicity 1.

Further, since $\mathbb{R}^3 = Y \oplus Z$, we have $3 = \dim(Y) + \dim(Z) = 1 + 2$ and so these are the only eigenvalues that we are going to find.

Question 4.

(a) A complex matrix A is:

i. *unitary* iff $\mathsf{AA}^\dagger = \mathsf{I}$.

ii. *normal* iff $\mathsf{AA}^\dagger = \mathsf{A}^\dagger\mathsf{A}$.

iii. *unitarily diagonalisable* iff there exists a unitary matrix P such that the matrix $\mathsf{P}^\dagger\mathsf{AP}$ is diagonal.

Also, a condition which will guarantee that a square matrix A has an inverse is $\det(\mathsf{A}) \neq 0$.

(b) Let A be a square complex matrix. We are asked to prove that:

i. A is invertible if and only if the eigenvalues of A are all non-zero.

Proof: Clearly, since λ is an eigenvalue of the matrix A if and only if $\det(\mathsf{A} - \lambda\mathsf{I}) = 0$, we have

$$\mathsf{A} \text{ is invertible iff } \det(\mathsf{A}) \neq 0 \text{ iff } \det(\mathsf{A} - 0\mathsf{I}) \neq 0 \text{ iff } \lambda = 0 \text{ is not an eigenvalue of } \mathsf{A},$$

as required.

ii. The eigenvalues of a unitary matrix A all have a modulus of one.

Proof: Let λ be any eigenvalue of A, and let $\mathbf{x}$ be an eigenvector of A corresponding to λ, i.e. $\mathsf{A}\mathbf{x} = \lambda\mathbf{x}$. As A is unitary, $\mathsf{A}^\dagger\mathsf{A} = \mathsf{I}$, and so

$$\mathbf{x}^\dagger\mathsf{A}^\dagger\mathsf{A}\mathbf{x} = \mathbf{x}^\dagger\mathsf{I}\mathbf{x} = \mathbf{x}^\dagger\mathbf{x}.$$

But, using the $(\mathsf{AB})^\dagger = \mathsf{B}^\dagger\mathsf{A}^\dagger$ rule, we can also see that

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

Equating these two expressions we find

$$|\lambda|^2\mathbf{x}^\dagger\mathbf{x} = \mathbf{x}^\dagger\mathbf{x} \implies (|\lambda|^2 - 1)\mathbf{x}^\dagger\mathbf{x} = 0.$$

But, as $\mathbf{x}$ is an eigenvector, $\mathbf{x}^\dagger\mathbf{x} = \|\mathbf{x}\|^2 \neq 0$, and so this gives $|\lambda|^2 = 1$. Thus, $|\lambda| = 1$ (as required).

(c) Let A be a square invertible matrix with eigenvalue λ and $\mathbf{x}$ as a corresponding eigenvector, i.e. $\mathsf{A}\mathbf{x} = \lambda\mathbf{x}$. We are asked to show that the matrix A^{-1} has λ^{-1} as an eigenvalue with $\mathbf{x}$ as a corresponding eigenvector. To do this, we note that as the matrix A is invertible, we know that the matrix A^{-1} exists and that $\lambda \neq 0$ by **(bi)**. Thus, since

$$\mathsf{A}\mathbf{x} = \lambda\mathbf{x} \implies \mathsf{A}^{-1}\mathsf{A}\mathbf{x} = \lambda\mathsf{A}^{-1}\mathbf{x} \implies \lambda^{-1}\mathbf{x} = \mathsf{A}^{-1}\mathbf{x},$$

the matrix A^{-1} has λ^{-1} as an eigenvalue with $\mathbf{x}$ as a corresponding eigenvector (as required).

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, clearly, the spectral decomposition of the matrix A^{-1} is

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

using the result that we have just proved.

(d) We are given that the complex matrix

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

and since

- $[1, i, 0]^t$ is an eigenvector of A, we have

$$\begin{bmatrix} -i & -i & 0 \\ i & -i & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ i \\ 0 \end{bmatrix} = \begin{bmatrix} -i+1 \\ i(1-i) \\ 0 \end{bmatrix} = (1-i) \begin{bmatrix} 1 \\ i \\ 0 \end{bmatrix},$$

i.e. this corresponds to an eigenvalue of $1 - i$.

- $[0, 0, 1]^t$ is an eigenvector of A, we have

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

i.e. this corresponds to an eigenvalue of 1.

- $-1 - i$ is an eigenvalue of A and so a corresponding eigenvector $[x, y, z]^t$ is given by

$$\begin{bmatrix} 1 & -i & 0 \\ i & 1 & 0 \\ 0 & 0 & 2+i \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \mathbf{0} \implies \begin{aligned} x - iy &= 0 \\ ix + y &= 0 \\ (2+i)z &= 0 \end{aligned} \implies \begin{aligned} ix + y &= 0 \\ ix + y &= 0 \\ z &= 0 \end{aligned},$$

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

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

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

(since the eigenvectors are already mutually orthogonal) and substitute all of this into the expression in **(c)**. Thus,

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

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

is the spectral decomposition of A.

Consequently, since the eigenvalues of A are all non-zero, by **(bi)**, this matrix is invertible and so by **(c)**, the spectral decomposition of its inverse is

$$\mathsf{A}^{-1} = \frac{1}{2(1-i)} \begin{bmatrix} 1 & -i & 0 \\ i & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix} - \frac{1}{2(1+i)} \begin{bmatrix} 1 & i & 0 \\ -i & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix} + 1 \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

And, since the eigenvalues of A are not all of modulus one (note that $|1 + i| = |-1 - i| = \sqrt{2}$), by **(bii)**, A is not unitary and so $\mathsf{A}^{-1} \neq \mathsf{A}^\dagger$. But clearly, taking the complex conjugate transpose of the spectral decomposition of A, we can see that

$$\mathsf{A}^\dagger = \frac{1+i}{2} \begin{bmatrix} 1 & -i & 0 \\ i & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix} - \frac{1-i}{2} \begin{bmatrix} 1 & i & 0 \\ -i & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix} + 1 \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix},$$

is the spectral decomposition of $\mathsf{A}^\dagger$.

Question 5.

(a) A *weak generalised inverse* of an $m \times n$ matrix A is any $n \times m$ matrix A^g which is such that $\mathsf{AA}^g\mathsf{A} = \mathsf{A}$.

(b) We are asked to prove that:

> The system of linear equations $\mathsf{A}\mathbf{x} = \mathbf{b}$ is consistent if and only if $\mathbf{b} = \mathsf{AA}^g\mathbf{b}$.
>
> **Proof:** This is an 'if and only if' statement and so we have to prove it 'both ways':
>
> **LTR:** If $\mathsf{A}\mathbf{x} = \mathbf{b}$ is consistent, then it must be the case that $\mathbf{b} \in R(\mathsf{A})$, i.e. there is an $\mathbf{x} \in \mathbb{R}^n$ such that $\mathsf{A}\mathbf{x} = \mathbf{b}$. Thus, as $\mathsf{AA}^g\mathsf{A} = \mathsf{A}$, this means that $\mathsf{AA}^g\mathsf{A}\mathbf{x} = \mathsf{AA}^g\mathbf{b}$ is the same as $\mathsf{A}\mathbf{x} = \mathsf{AA}^g\mathbf{b}$. Consequently, $\mathbf{b} = \mathsf{AA}^g\mathbf{b}$ (as required).
>
> **RTL:** If $\mathsf{AA}^g\mathbf{b} = \mathbf{b}$, then $\mathbf{x} = \mathsf{A}^g\mathbf{b}$ is clearly a solution of the matrix equation $\mathsf{A}\mathbf{x} = \mathbf{b}$. That is, this matrix equation has a solution and so it is consistent (as required).

Further, to show that:

> For any vector $\mathbf{w}$, the vector $\mathbf{x} = \mathsf{A}^g\mathbf{b} + (\mathsf{A}^g\mathsf{A} - \mathsf{I})\mathbf{w}$ is a solution of the consistent system of linear equations $\mathsf{A}\mathbf{x} = \mathbf{b}$.

we note that,

$$\mathsf{A}\mathbf{x} = \mathsf{A}\left[\mathsf{A}^g\mathbf{b} + (\mathsf{A}^g\mathsf{A} - \mathsf{I})\mathbf{w}\right] = \mathsf{AA}^g\mathbf{b} + (\mathsf{AA}^g\mathsf{A} - \mathsf{AI})\mathbf{w} = \mathbf{b} + (\mathsf{A} - \mathsf{A})\mathbf{w} = \mathbf{b},$$

where we have used the fact that $\mathsf{AA}^g\mathbf{b} = \mathbf{b}$ since the equations are assumed to be consistent.

(c) A *right inverse*, R, of a matrix A is any matrix R which is such that $\mathsf{AR} = \mathsf{I}$. We are then asked to show that:

i. Right inverses are weak generalised inverses:

   This is the case since
   $$\mathsf{ARA} = (\mathsf{AR})\mathsf{A} = \mathsf{IA} = \mathsf{A}.$$
   where we have used the fact that $\mathsf{AR} = \mathsf{I}$.

ii. If A is an $m \times n$ matrix of rank m, then the matrix $\mathsf{A}^t(\mathsf{AA}^t)^{-1}$ is a right inverse of A.

   This is the case since if A is an $m \times n$ matrix of rank m, then AA^t is an $m \times m$ matrix where $\rho(\mathsf{AA}^t) = \rho(\mathsf{A}) = m$, i.e. the matrix AA^t is invertible and so the matrix $\mathsf{A}^t(\mathsf{AA}^t)^{-1}$ exists. Further, it is a right inverse since
   $$\mathsf{A}[\mathsf{A}^t(\mathsf{AA}^t)^{-1}] = (\mathsf{AA}^t)(\mathsf{AA}^t)^{-1} = \mathsf{I},$$
   as desired.

(d) To find a weak generalised inverse of the matrix

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

we note that this is a 2×3 matrix of rank 2 and so, by **(c)**, the matrix $\mathsf{A}^t(\mathsf{AA}^t)^{-1}$ is a weak generalised inverse of A. Thus, we have

$$\mathsf{AA}^t = \begin{bmatrix} -1 & 1 & 1 \\ 1 & 2 & 1 \end{bmatrix}\begin{bmatrix} -1 & 1 \\ 1 & 2 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} 3 & 2 \\ 2 & 6 \end{bmatrix} \implies (\mathsf{AA}^t)^{-1} = \frac{1}{14}\begin{bmatrix} 6 & -2 \\ -2 & 3 \end{bmatrix},$$

and so,

$$\mathsf{A}^g = \frac{1}{14}\begin{bmatrix} -1 & 1 \\ 1 & 2 \\ 1 & 1 \end{bmatrix}\begin{bmatrix} 6 & -2 \\ -2 & 3 \end{bmatrix} = \frac{1}{14}\begin{bmatrix} -8 & 5 \\ 2 & 4 \\ 4 & 1 \end{bmatrix},$$

is the required weak generalised inverse. So, to find all possible solutions to the system of linear equations given by

$$-x + y + z = -1$$
$$x + 2y + z = 1$$

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

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

Thus, as

$$\mathsf{A}^g\mathsf{A} = \frac{1}{14}\begin{bmatrix} -8 & 5 \\ 2 & 4 \\ 4 & 1 \end{bmatrix}\begin{bmatrix} -1 & 1 & 1 \\ 1 & 2 & 1 \end{bmatrix} = \frac{1}{14}\begin{bmatrix} 13 & 2 & -3 \\ 2 & 10 & 6 \\ -3 & 6 & 5 \end{bmatrix},$$

we have

$$\mathsf{A}^g\mathsf{A} - \mathsf{I} = \frac{1}{14}\left\{\begin{bmatrix} 13 & 2 & -3 \\ 2 & 10 & 6 \\ -3 & 6 & 5 \end{bmatrix} - \begin{bmatrix} 14 & 0 & 0 \\ 0 & 14 & 0 \\ 0 & 0 & 14 \end{bmatrix}\right\} = \frac{1}{14}\begin{bmatrix} -1 & 2 & -3 \\ 2 & -4 & 6 \\ -3 & 6 & -9 \end{bmatrix},$$

and

$$\mathsf{A}^g\mathbf{b} = \frac{1}{14}\begin{bmatrix} -8 & 5 \\ 2 & 4 \\ 4 & 1 \end{bmatrix}\begin{bmatrix} -1 \\ 1 \end{bmatrix} = \frac{1}{14}\begin{bmatrix} 13 \\ 2 \\ -3 \end{bmatrix}.$$

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

$$\mathbf{x} = \frac{1}{14}\left\{\begin{bmatrix} 13 \\ 2 \\ -3 \end{bmatrix} + \begin{bmatrix} -1 & 2 & -3 \\ 2 & -4 & 6 \\ -3 & 6 & -9 \end{bmatrix}\mathbf{w}\right\},$$

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

We know from **(b)** that the $(\mathsf{A}^g\mathsf{A} - IM)\mathbf{w}$ part of our solutions will yield a vector in $N(\mathsf{A})$ and so the solutions set is the translate of $N(\mathsf{A})$ by $\frac{1}{14}[13, 2, -3]^t$. 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

$$\mathbf{x} = \frac{1}{14}\begin{bmatrix} 13 \\ 2 \\ -3 \end{bmatrix} + \lambda\begin{bmatrix} -1 \\ 2 \\ -3 \end{bmatrix},$$

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

Question 6

(a) We are given an orthonormal set of vectors $\{\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n\}$ which span a vector space V. Clearly, for some $k < n$, an expression for $\mathsf{P}\mathbf{v}$, the orthogonal projection of a vector $\mathbf{v} \notin \text{Lin}(S)$ onto $\text{Lin}\{\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_k\}$ is

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

So, clearly, as any vector $\mathbf{v} \in V$ can be written as

$$\mathbf{v} = \sum_{i=1}^{n} \langle \mathbf{v}, \mathbf{e}_i \rangle \mathbf{e}_i,$$

we have

$$(\mathsf{I} - \mathsf{P})\mathbf{v} = \mathbf{v} - \mathsf{P}\mathbf{v} = \sum_{i=k+1}^{n} \langle \mathbf{v}, \mathbf{e}_i \rangle \mathbf{e}_i,$$

and so since the basis is orthonormal, we have

$$\langle \mathsf{P}\mathbf{v}, (\mathsf{I} - \mathsf{P})\mathbf{v} \rangle = 0.$$

So, noting that the mean square error associated with the vectors $\mathbf{v}$ and $\mathsf{P}\mathbf{v}$ is given by

$$\|\mathbf{v} - \mathsf{P}\mathbf{v}\|^2 = \langle \mathbf{v} - \mathsf{P}\mathbf{v}, \mathbf{v} - \mathsf{P}\mathbf{v} \rangle = \langle \mathbf{v}, \mathbf{v} - \mathsf{P}\mathbf{v} \rangle - \langle \mathsf{P}\mathbf{v}, \mathbf{v} - \mathsf{P}\mathbf{v} \rangle,$$

and since $\langle \mathsf{P}\mathbf{v}, \mathbf{v} - \mathsf{P}\mathbf{v} \rangle = 0$, we have

$$\|\mathbf{v} - \mathsf{P}\mathbf{v}\|^2 = \langle \mathbf{v}, \mathbf{v} - \mathsf{P}\mathbf{v} \rangle.$$

So, using our expression for $\mathsf{P}\mathbf{v}$ we get

$$\begin{aligned} \|\mathbf{v} - \mathsf{P}\mathbf{v}\|^2 &= \left\langle \mathbf{v}, \mathbf{v} - \sum_{i=1}^{k} \langle \mathbf{v}, \mathbf{e}_i \rangle \mathbf{e}_i \right\rangle \\ &= \langle \mathbf{v}, \mathbf{v} \rangle - \sum_{i=1}^{k} \langle \mathbf{v}, \mathbf{e}_i \rangle \langle \mathbf{v}, \mathbf{e}_i \rangle \\ &= \|\mathbf{v}\|^2 - \sum_{i=1}^{k} \langle \mathbf{v}, \mathbf{e}_i \rangle^2, \end{aligned}$$

as required.

(b) We are given the vector space which is the linear span of the functions e^x, 1 and e^{-x} of x defined over the interval $0 \leq x \leq 1$. So, with respect to the inner product given by

$$\langle \alpha e^x + \beta e^{-x} + \gamma, \alpha' e^x + \beta' e^{-x} + \gamma' \rangle = \alpha\alpha' + \beta\beta' + \gamma\gamma',$$

we can find the orthogonal projection of the unit function onto the subspace spanned by the functions e^x and e^{-x}. To do this, we note that since

$$\langle e^x, e^{-x} \rangle = (1)(0) + (0)(1) = 0,$$

the functions e^x and e^{-x} are orthogonal and as

$$\langle e^x, e^x \rangle = (1)(1) = 1 \quad \text{and} \quad \langle e^{-x}, e^{-x} \rangle = (1)(1) = 1,$$

they are unit too. Thus, the set $\{e^x, e^{-x}\}$ is already orthonormal. Thus, the required orthogonal projection is

$$\mathsf{P}1 = \langle 1, e^x \rangle e^x + \langle 1, e^{-x} \rangle e^{-x} = [(0)(1) + (1)(0)]e^x + [(0)(1) + (1)(0)]e^{-x} = 0.$$

As such, the mean square error associated with this orthogonal projection is given by

$$\|\mathsf{P}1 - 1\|^2 = \|1\|^2 = \langle 1, 1\rangle = (1)(1) = 1.$$

(c) We are asked to show that

$$\|\mathbf{x} + \mathbf{y}\| \leq \|\mathbf{x}\| + \|\mathbf{y}\|,$$

for all vectors $\mathbf{x}$ and $\mathbf{y}$ in a real inner product space. (Notice that this is just the Triangle Inequality.) To do this we note that:

$$\begin{aligned} \|\mathbf{x} + \mathbf{y}\|^2 &= \langle \mathbf{x} + \mathbf{y}, \mathbf{x} + \mathbf{y}\rangle \\ &= \langle \mathbf{x}, \mathbf{x}\rangle + \langle \mathbf{x}, \mathbf{y}\rangle + \langle \mathbf{y}, \mathbf{x}\rangle + \langle \mathbf{y}, \mathbf{y}\rangle \\ \implies \quad \|\mathbf{x} + \mathbf{y}\|^2 &= \|\mathbf{x}\|^2 + 2\langle \mathbf{x}, \mathbf{y}\rangle + \|\mathbf{y}\|^2, \end{aligned}$$

where we have used the symmetry property of real inner products. However, the Cauchy-Schwartz inequality tells us that

$$|\langle \mathbf{x}, \mathbf{y}\rangle| \leq \|\mathbf{x}\|\|\mathbf{y}\|,$$

and so

$$\|\mathbf{x} + \mathbf{y}\|^2 \leq \|\mathbf{x}\|^2 + 2\|\mathbf{x}\|\|\mathbf{y}\| + \|\mathbf{y}\|^2.$$

But, factorising the right-hand-side then gives,

$$\|\mathbf{x} + \mathbf{y}\|^2 \leq (\|\mathbf{x}\| + \|\mathbf{y}\|)^2,$$

and hence, since norms are non-negative, we have

$$\|\mathbf{x} + \mathbf{y}\| \leq \|\mathbf{x}\| + \|\mathbf{y}\|,$$

(as required)

(d) Considering the vector space $\mathbb{E}^{[0,t]}$ given by the linear span of the functions e^x and e^{-x} of x defined over the interval $0 \leq x \leq t$ for some $t > 0$ and using the inner product

$$\langle f(x), g(x)\rangle = \int_0^t f(x)g(x)\, dx,$$

defined on this vector space, we have

$$\begin{aligned} \|e^x + e^{-x}\|^2 &= \int_0^t (e^x + e^{-x})^2\, dx = \int_0^t (e^{2x} + 2 + e^{-2x})\, dx \\ &= \left[\frac{e^{2x}}{2} + 2x - \frac{e^{-2x}}{2}\right]_0^t \\ &= \left[\frac{e^{2t}}{2} + 2t - \frac{e^{-2t}}{2}\right] - \left[\frac{1}{2} + 0 - \frac{1}{2}\right] \\ &= \sinh(2t) + 2t, \end{aligned}$$

and,

$$\|e^x\|^2 = \int_0^t e^{2x}\, dx = \left[\frac{e^{2x}}{2}\right]_0^t = \frac{e^{2t}}{2} - \frac{1}{2},$$

and,

$$\|e^{-x}\|^2 = \int_0^t e^{-2x}\, dx = \left[-\frac{e^{-2x}}{2}\right]_0^t = -\frac{e^{-2t}}{2} + \frac{1}{2}.$$

So, using the result in **(c)**, we have

$$\sqrt{\sinh(2t) + 2t} \leq \sqrt{\frac{e^{2t}}{2} - \frac{1}{2}} + \sqrt{\frac{1}{2} - \frac{e^{-2t}}{2}},$$

as required.