Examples sheet 2 – Solutions – Linear Algebra

INNER PRODUCTS. NORMS.

1. Recall that inner products are (induced by) symmetric and positive-definite bilinear forms. Thus, we need to check the following properties:

i. symmetry: $\mathcal{B}(\mathbf{v},\mathbf{w}) = \mathcal{B}(\mathbf{w},\mathbf{v})$;

ii. linearity: $\mathcal{B}(a\mathbf{u}+b\mathbf{v},\mathbf{w}) = a\mathcal{B}(\mathbf{u},\mathbf{w}) + b\mathcal{B}(\mathbf{v},\mathbf{w})$;

iii. non-negativity: $\mathcal{B}(\mathbf{v},\mathbf{v}) \geq 0, \quad \mathcal{B}(\mathbf{v},\mathbf{v}) = 0 \iff \mathbf{v} = \mathbf{0}$.

(a) Let $V = \mathcal{P}_n(\mathbb{R})$ and let $\mathcal{B}(\cdot,\cdot) : V \times V \to \mathbb{R}$ be given by

$$\mathcal{B}(p,q) := \int_{-1}^{1} p'(x)q'(x)\mathrm{d}x + p(0)q(0).$$

For any $p, q, r \in \mathcal{P}_n(\mathbb{R})$,

i. symmetry holds:

$$\mathcal{B}(p,q) = \int_{-1}^{1} p'(x)q'(x)\mathrm{d}x + p(0)q(0) = \int_{-1}^{1} q'(x)p'(x)\mathrm{d}x + q(0)p(0) = \mathcal{B}(q,p).$$

ii. linearity holds:

$$\begin{aligned}
\mathcal{B}(ap+bq,r) &= \int_{-1}^{1} (ap'(x)+bq'(x))r'(x)\mathrm{d}x + (ap(0)+bq(0))r(0) \\
&= a\left(\int_{-1}^{1} p'(x)r'(x)\mathrm{d}x + p(0)r(0)\right) + b\left(\int_{-1}^{1} q'(x)r'(x)\mathrm{d}x + q(0)r(0)\right) \\
&= a\mathcal{B}(p,r) + b\mathcal{B}(q,r).
\end{aligned}$$

iii. non-negativity holds:

$$\mathcal{B}(p,p) = \int_{-1}^{1} \left(p'(x)\right)^2 \mathrm{d}x + p(0)^2 \geq 0$$

with

$$\mathcal{B}(p,p) = 0 \iff \int_{-1}^{1} \left(p'(x)\right)^2 \mathrm{d}x + p(0)^2 = 0 \iff \begin{cases} p'(x) & = 0 \\ p(0) & = 0 \end{cases} \iff \begin{cases} p(x) & = c \\ p(0) & = 0 \end{cases} \iff p(x) = 0.$$

Hence, $\mathcal{B}(\cdot,\cdot)$ defines an inner product on $\mathcal{P}_n(\mathbb{R}) \times \mathcal{P}_n(\mathbb{R})$.

Remark. If the second term is omitted in the definition of $\mathcal{B}(\cdot,\cdot)$, then we would not be able to establish the last property, as $\mathcal{B}(c,c) = 0$ for any real constant.

(b) Let $V = \mathbb{Z}_2^3(\mathbb{Z}_2)$ and let $\mathcal{B}(\cdot,\cdot) : V \times V \to \mathbb{R}$ be given by

$$\mathcal{B}(\mathbf{v},\mathbf{w}) := (v_1w_1 + v_2w_2 + v_3w_3) \pmod 2.$$

For any $\mathbf{u},\mathbf{v},\mathbf{w} \in V$ and $a, b \in \mathbb{Z}_2$

i. symmetry holds:

$$\mathcal{B}(\mathbf{v},\mathbf{w}) = (v_1w_1 + v_2w_2 + v_3w_3) \pmod 2 = (w_1v_1 + w_2v_2 + w_3v_3) \pmod 2 = \mathcal{B}(\mathbf{w},\mathbf{v});$$

ii. linearity holds:

$$\begin{aligned}
\mathcal{B}(a\mathbf{u}+b\mathbf{v},\mathbf{w}) &= [(au_1+bv_1)w_1 + (au_2+bv_2)w_2 + (au_3+bv_3)w_3] \pmod 2 \\
&= (au_1w_1 + bv_1w_1 + au_2w_2 + bv_2w_2 + au_3w_3 + bv_3w_3) \pmod 2 \\
&= a\,(u_1w_1 + u_2w_2 + u_3w_3) \pmod 2 + b\,(v_1w_1 + v_2w_2 + v_3w_3) \pmod 2 \\
&= a\mathcal{B}(\mathbf{u},\mathbf{w}) + b\mathcal{B}(\mathbf{v},\mathbf{w});
\end{aligned}$$

iii. non-negativity fails: setting $\mathbf{v} := (1,1,0) \neq \mathbf{0}$, we get

$$\mathcal{B}(\mathbf{v},\mathbf{v}) = (1\cdot 1 + 1\cdot 1 + 0\cdot 0) \pmod 2 = 0.$$

Hence, $\mathcal{B}(\cdot,\cdot)$ does not define an inner product on $\mathbb{Z}_2^3(\mathbb{Z}_2)$.

(c) Let $V = \mathbb{R}^2$ and let $\mathcal{B}(\cdot,\cdot) : V \times V \to \mathbb{R}$ be given by

$$\mathcal{B}(\mathbf{v},\mathbf{w}) := v_1 a_{11} w_1 + v_1 a_{12} w_2 + v_2 a_{21} w_1 + v_2 a_{22} w_2, \qquad [a_{ij}] := \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} =: A.$$

For general matrices A, $\mathcal{B}(\cdot,\cdot)$ does not define an inner product. Let us find conditions on A such that it does. For any $\mathbf{u}, \mathbf{v}\,\mathbf{w} \in \mathbb{R}^2$ and $a, b \in \mathbb{R}$

i. symmetry holds provided

$$v_1 a_{11} w_1 + v_1 a_{12} w_2 + v_2 a_{21} w_1 + v_2 a_{22} w_2 = w_1 a_{11} v_1 + w_1 a_{12} v_2 + w_2 a_{21} v_1 + w_2 a_{22} v_2 \iff a_{12} = a_{21};$$

ii. linearity holds for any entries in A (straightforward to check);

iii. non-negativity holds provided

$$a_{11} v_1^2 + 2a_{12} v_1 v_2 + a_{22} v_2^2 \geq 0$$

with

$$a_{11} v_1^2 + 2a_{12} v_1 v_2 + a_{22} v_2^2 = 0 \iff v_1 = v_2 = 0.$$

First, note that if we take $\mathbf{v} = \mathbf{e}_1$ and $\mathbf{v} = \mathbf{e}_2$, we obtain $a_{11} \geq 0$ and $a_{22} \geq 0$, respectively. Note also that we cannot have $a_{11} = 0$, otherwise $\mathcal{B}(\mathbf{e}_1, \mathbf{e}_1) = 0$. A similar argument applies to a_{22}, so that we require $a_{11}, a_{22} > 0$.

Consider now the case $\mathbf{v} \neq \mathbf{0}$; without loss of generality, assume $v_2 \neq 0$. In this case, we require

$$a_{11} v_1^2 + 2a_{12} v_1 v_2 + a_{22} v_2^2 > 0 \iff a_{11}\left(\frac{v_1}{v_2}\right)^2 + 2a_{12}\frac{v_1}{v_2} + a_{22} > 0 \iff a_{11}x^2 + 2a_{12}x + a_{22} > 0$$

where we set $x = v_1/v_2$. This holds provided $\Delta := 4a_{12}^2 - 4a_{11}a_{22} < 0$, i.e., $\det A > 0$. Hence, we require the entries of A to satisfy the following properties

$$a_{12} = a_{21}, \quad a_{11}, a_{22} > 0, \quad \det A > 0.$$

We will see later that this is equivalent to requiring that A is a symmetric and positive definite matrix (i.e., a symmetric matrix with positive real eigenvalues).

(d) Let $V = \mathbb{R}^2$ and let $\mathcal{B}(\cdot,\cdot) : V \times V \to \mathbb{R}$ be given by

$$\mathcal{B}(\mathbf{v},\mathbf{w}) := v_1 w_2 + v_2 w_1.$$

For any $\mathbf{u}, \mathbf{v}, \mathbf{w} \in V$,

i. symmetry holds:

$$\mathcal{B}(\mathbf{v},\mathbf{w}) = v_1 w_2 + v_2 w_1 = w_1 v_2 + w_2 v_1 = \mathcal{B}(\mathbf{w},\mathbf{v});$$

ii. linearity holds:

$$\mathcal{B}(a\mathbf{u}+b\mathbf{v},\mathbf{w}) = (au_1+bv_1)w_2+(au_2+bv_2)w_1 = a(u_1w_2+u_2w_1)+b(v_1w_2+v_2w_1) = a\mathcal{B}(\mathbf{u},\mathbf{w})+b\mathcal{B}(\mathbf{v},\mathbf{w});$$

iii. non-negativity fails to hold:

$$\mathcal{B}(\mathbf{v},\mathbf{v}) = v_1 v_2 + v_2 v_1 = 2v_1 v_2 < 0 \quad \text{if, for example, } \mathbf{v} = \begin{bmatrix} 1 \\ -1 \end{bmatrix}.$$

Hence, $\mathcal{B}(\cdot,\cdot)$ does not define an inner product.

2. Let $(V, \langle\cdot,\cdot\rangle)$ be an inner product space. To show that $\langle \mathbf{0}, \mathbf{v}\rangle = 0$ for any $\mathbf{v} \in V$, consider

$$\langle \mathbf{0}, \mathbf{v}\rangle = \langle 0\cdot \mathbf{u}, \mathbf{v}\rangle = 0\cdot \langle \mathbf{u}, \mathbf{v}\rangle = 0.$$

Alternatively, for any $\mathbf{u}, \mathbf{v} \in V$,

$$0 = \langle \mathbf{u}, \mathbf{v}\rangle - \langle \mathbf{u}, \mathbf{v}\rangle = \langle \mathbf{u} - \mathbf{u}, \mathbf{v}\rangle = \langle \mathbf{0}, \mathbf{v}\rangle .$$

3. Let us check the norm properties.

i. Using the Cauchy-Schwarz inequality, namely, $\langle \mathbf{v}, \mathbf{w}\rangle \leq n(\mathbf{v})n(\mathbf{w})$, we find

$$n^2(\mathbf{v}+\mathbf{w}) = \langle \mathbf{v}+\mathbf{w}, \mathbf{v}+\mathbf{w}\rangle = \langle \mathbf{v}, \mathbf{v}\rangle + 2\langle \mathbf{v}, \mathbf{w}\rangle + \langle \mathbf{w}, \mathbf{w}\rangle \leq n^2(\mathbf{v}) + 2n(\mathbf{v})n(\mathbf{w}) + n^2(\mathbf{w}) = (n(\mathbf{v}) + n(\mathbf{w}))^2 ,$$

and the triangle inequality follows by taking square-roots.

ii. We have, using the non-negativity of the inner product,

$$n^2(a\mathbf{v}) = \langle a\mathbf{v}, a\mathbf{v}\rangle = a^2 \langle \mathbf{v}, \mathbf{v}\rangle \Longrightarrow n(a\mathbf{v}) = |a\sqrt{\langle \mathbf{v}, \mathbf{v}\rangle}| = |a||\sqrt{\langle \mathbf{v}, \mathbf{v}\rangle}| = |a| \sqrt{\langle \mathbf{v}, \mathbf{v}\rangle} = |a|\, n(\mathbf{v}).$$

iii. $n(\mathbf{v}) = \sqrt{\langle \mathbf{v}, \mathbf{v}\rangle} \geq 0$, by non-negativity of the inner product; moreover, $n(\mathbf{v}) = 0 \Longleftrightarrow \langle \mathbf{v}, \mathbf{v}\rangle = 0 \Longleftrightarrow \mathbf{v} = \mathbf{0}$.

4. We have, using the definition of induced norm,

$$\|\mathbf{v} \pm \mathbf{w}\|^2 = \|\mathbf{v}\|^2 \pm \langle \mathbf{v}, \mathbf{w}\rangle + \|\mathbf{w}\|^2 .$$

The result follows by adding the $+$ and $-$ identities above.

5. (a) We have

i. $\|\mathbf{v}+\mathbf{w}\|_1 = \sum_{i=1}^{n} |v_i + w_i| \leq \sum_{i=1}^{n} |v_i| + |w_i| = \sum_{i=1}^{n} |v_i| + \sum_{i=1}^{n} |w_i| = \|\mathbf{v}\|_1 + \|\mathbf{w}\|_1$.

ii. $\|a\mathbf{v}\|_1 = \sum_{i=1}^{n} |av_i| = \sum_{i=1}^{n} |a||v_i| = |a| \sum_{i=1}^{n} |v_i| = |a|\, \|\mathbf{v}\|_1$.

iii. $\|\mathbf{v}\|_1 = \sum_{i=1}^{n} |v_i| \geq 0$ with $\|\mathbf{v}\|_1 = 0$ if and only if $v_i = 0$ for $i = 1, \ldots, n$.

(b) We have, using the properties of the max and $|\cdot|$ functions,

i. $\|\mathbf{v}+\mathbf{w}\|_\infty = \max_{1\leq j\leq n} |v_j + w_j| \leq \max_{1\leq j\leq n} (|v_j| + |w_j|) = \max_{1\leq j\leq n} |v_j| + \max_{1\leq j\leq n} |w_j| = \|\mathbf{v}\|_\infty + \|\mathbf{w}\|_\infty$.

ii. $\|a\mathbf{v}\|_\infty = \max_{1\leq j\leq n} |a\mathbf{v}_j| = \max_{1\leq j\leq n} |a||\mathbf{v}_j| = |a| \max_{1\leq j\leq n} |\mathbf{v}_j| = |a|\, \|a\mathbf{v}\|_\infty$.

iii. $\|\mathbf{v}\|_\infty = \max_{1\leq j\leq n} |v_j| \geq 0$ with $\|\mathbf{v}\|_\infty = 0$ if and only if $v_j = 0$ for $j = 1, \ldots, n$.

6. Here is a counter-example: $\mathbf{v} = (2, 1), \mathbf{w} = (0, 2)$. Then

$$\|\mathbf{v}+\mathbf{w}\|_1^2 + \|\mathbf{v}-\mathbf{w}\|_1^2 = 25 + 9 = 34 \neq 26 = 2(\|\mathbf{v}\|_1^2 + \|\mathbf{w}\|_1^2).$$

The same counter-example applies to $\|\cdot\|_\infty$. Hence, $\|\cdot\|_1$ and $\|\cdot\|_\infty$ cannot be induced norms.

7. (a) We have

$$\|p\| = \sqrt{\langle p, p\rangle} = \left(\int_{-1}^{1} p^2(x)\mathrm{d}x\right)^{1/2} = \sqrt{2}.$$

(b) We find

$$\|p - q\| = \|2 - x\| = \sqrt{\langle 2 - x, 2 - x\rangle} = \left(\int_{-1}^{1} (2 - x)^2(x)\mathrm{d}x\right)^{1/2} = \sqrt{\frac{26}{3}}.$$

(c) We find

$$\|q\| = \|1-x\| = \sqrt{\langle 1-x, 1-x\rangle} = \left(\int_{-1}^{1}(1-x)^2 \mathrm{d}x\right)^{1/2} = \sqrt{\frac{8}{3}}$$

and

$$\langle p, q\rangle = \int_{-1}^{1} 1\cdot(x-1)\mathrm{d}x = -2,$$

so that the angle is obtained via

$$\cos\theta = \frac{\langle p, q\rangle}{\|p\|\,\|q\|} = \frac{-2}{\sqrt{2}\cdot\sqrt{8/3}} = -\frac{\sqrt{3}}{2} \Longrightarrow \theta = \frac{5\pi}{6}.$$

ORTHOGONALITY

8. (a) We have

$$\langle \mathbf{v}, \mathbf{w}\rangle = \langle \mathbf{v}, \mathbf{u}\rangle - \frac{\langle \mathbf{u}, \mathbf{v}\rangle}{\|\mathbf{v}\|^2}\langle \mathbf{v}, \mathbf{v}\rangle = 0.$$

(b) Since $\mathbf{u} = \mathbf{w} + \frac{\langle \mathbf{u}, \mathbf{v}\rangle}{\|\mathbf{v}\|^2}\mathbf{v}$, with the two vectors on the right being orthogonal, we can use Pythagoras' theorem to obtain

$$\|\mathbf{u}\|^2 = \|\mathbf{w}\|^2 + \left(\frac{\langle \mathbf{u}, \mathbf{v}\rangle}{\|\mathbf{v}\|^2}\right)^2 \|\mathbf{v}\|^2 \geq \left(\frac{\langle \mathbf{u}, \mathbf{v}\rangle}{\|\mathbf{v}\|^2}\right)^2 \|\mathbf{v}\|^2 = \frac{\langle \mathbf{u}, \mathbf{v}\rangle^2}{\|\mathbf{v}\|^2} \Longleftrightarrow \|\mathbf{u}\|^2\|\mathbf{v}\|^2 \geq \langle \mathbf{u}, \mathbf{v}\rangle^2,$$

and the result of the Cauchy-Schwarz inequality follows by taking the square-root.

9. If $\langle \mathbf{v}, \mathbf{w}\rangle = 0$, we find

$$\|\mathbf{v} + a\mathbf{w}\|^2 = \|\mathbf{v}\|^2 + 2a\langle \mathbf{v}, \mathbf{w}\rangle + a^2\|\mathbf{w}\|^2 \geq \|\mathbf{v}\|^2.$$

If $\|\mathbf{v}\| \leq \|\mathbf{v} + a\mathbf{w}\|$, we find

$$\|\mathbf{v}\|^2 \leq \|\mathbf{v} + a\mathbf{w}\|^2 = \|\mathbf{v}\|^2 + 2a\langle \mathbf{v}, \mathbf{w}\rangle + a^2\|\mathbf{w}\|^2.$$

Hence, for all $a \in \mathbb{R}$,

$$a^2\|\mathbf{w}\|^2 \geq -2a\langle \mathbf{v}, \mathbf{w}\rangle$$

and setting $a = -\langle \mathbf{v}, \mathbf{w}\rangle / \|\mathbf{w}\|^2$ we find

$$\langle \mathbf{v}, \mathbf{w}\rangle^2 \geq 2\langle \mathbf{v}, \mathbf{w}\rangle^2 \Longrightarrow 0 \leq \langle \mathbf{v}, \mathbf{w}\rangle^2 \leq 0 \Longrightarrow \langle \mathbf{v}, \mathbf{w}\rangle = 0.$$

10. We show this by contradiction. Assume $\mathbf{0} \notin S$, so that all the elements of S are non-zero. We can then define $S' := \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ and note that this is an orthogonal set of non-zero vectors and therefore a linearly independent set with $|S'| = n$. Hence, S' is a basis for V and therefore we can express $\mathbf{v}_{n+1}$ as a linear combination of elements in S':

$$\mathbf{v}_{n+1} = a_1\mathbf{v}_1 + a_2\mathbf{v}_2 + \cdots + a_n\mathbf{v}_n,$$

where the (Fourier) coefficients are given by the expression

$$a_j = \frac{\langle \mathbf{v}_{n+1}, \mathbf{v}_j\rangle}{\langle \mathbf{v}_j, \mathbf{v}_j\rangle}.$$

By the orthogonality of S, $\langle \mathbf{v}_{n+1}, \mathbf{v}_j\rangle = 0$, so that $a_j = 0$ for all $j = 1, \ldots, n$ and hence $\mathbf{v}_{n+1} = \mathbf{0}$, which is a contradiction. Hence, at least one of the elements of S is the zero vector.

11. This question is simply confirming that the orthogonal projection onto a vector $\mathbf{u}$ is the same as the orthogonal projection on the one-dimensional vector space U spanned by $\mathbf{u}$. Since $U = \operatorname{span}\{\mathbf{u}\}$, a basis for U is $B = \{\mathbf{u}\}$. By Theorem 6.7 (see also Definition 6.7), the orthogonal projection onto U is the sum of the projections onto the basis elements, in this case, just one element:

$$\mathbf{v}_U^{\parallel} = \frac{\langle \mathbf{v}, \mathbf{u}\rangle}{\langle \mathbf{u}, \mathbf{u}\rangle}\mathbf{u} \Longrightarrow \mathbf{v}_U^{\perp} = \mathbf{v} - \frac{\langle \mathbf{v}, \mathbf{u}\rangle}{\langle \mathbf{u}, \mathbf{u}\rangle}\mathbf{u}.$$

12. Let $\mathbf{u} = \mathbf{v}_U^{\|} + \mathbf{e}$, where $\mathbf{e} \in U \setminus \{\mathbf{0}\}$. Note that, by the definition of orthogonal projection, $\mathbf{v}_U^{\perp} = \mathbf{v} - \mathbf{v}_U^{\|} \perp U$, i.e., $\langle \mathbf{v}_U^{\perp}, \mathbf{u} \rangle = 0$ for all $\mathbf{u} \in U$. In particular, this holds when $\mathbf{u} = \mathbf{e}$. Then

$$\|\mathbf{v} - \mathbf{u}\|^2 = \left\|\mathbf{v} - \mathbf{v}_U^{\|} - \mathbf{e}\right\|^2 = \left\|\mathbf{v} - \mathbf{v}_U^{\|}\right\|^2 - 2\left\langle \mathbf{v} - \mathbf{v}_U^{\|}, \mathbf{e}\right\rangle + \|\mathbf{e}\|^2 = \left\|\mathbf{v} - \mathbf{v}_U^{\|}\right\|^2 + \|\mathbf{e}\|^2 \geq \left\|\mathbf{v} - \mathbf{v}_U^{\|}\right\|^2,$$

since $\left\langle \mathbf{v} - \mathbf{v}_U^{\|}, \mathbf{e}\right\rangle = 0$ and $\|\mathbf{e}\| \geq 0$.

13. By **Q12**, setting $\mathbf{v} = \mathbf{1}$, we need to compute $\mathbf{u} = \mathbf{v}_U^{\|}$; this choice will ensure that $\|\mathbf{1} - \mathbf{u}\|$ takes the least value over U. Since our basis is already orthogonal, we can use the formula for $\mathbf{v}_U^{\|}$ given in the proof of Theorem 8.7:

$$\mathbf{u} := \mathbf{v}_U^{\|} = \frac{\langle \mathbf{1}, \mathbf{u}_1 \rangle}{\langle \mathbf{u}_1, \mathbf{u}_1 \rangle}\mathbf{u}_1 + \frac{\langle \mathbf{1}, \mathbf{u}_2 \rangle}{\langle \mathbf{u}_2, \mathbf{u}_2 \rangle}\mathbf{u}_2 = \frac{2}{2}\mathbf{u}_1 + \frac{1}{3}\mathbf{u}_2 = \begin{bmatrix} 1 \\ 1 \\ 0 \\ 0 \end{bmatrix} + \frac{1}{3}\begin{bmatrix} 1 \\ -1 \\ 0 \\ 1 \end{bmatrix} = \frac{1}{3}\begin{bmatrix} 4 \\ 2 \\ 0 \\ 1 \end{bmatrix}.$$

14. **(a)** By definition, $V^{\perp} = \{\mathbf{v} \in V : \langle \mathbf{v}, \mathbf{u} \rangle = 0 \text{ for all } \mathbf{u} \in V\}$. We verify the double inclusion $Z \subseteq V^{\perp}$ and $V^{\perp} \subseteq Z$.
Since $\langle \mathbf{0}, \mathbf{u} \rangle = 0$ for all $\mathbf{u} \in V$, we have $\mathbf{0} \in V^{\perp}$ and hence $Z \subseteq V^{\perp}$. On the other hand, $V^{\perp} \subseteq Z$ since if $\mathbf{w} \in V^{\perp}$, then $\mathbf{w} = \mathbf{0}$. Otherwise, if $\mathbf{w} \neq \mathbf{0}$ and $\langle \mathbf{w}, \mathbf{v} \rangle = 0$ for any $\mathbf{v} \in V$, then $\langle \mathbf{w}, \mathbf{v}_j \rangle = 0$ for $j = 1, 2, \ldots, n$ where $\mathbf{v}_j$ are the non-zero elements of some orthogonal basis B of V. This means that $S = B \cup \{\mathbf{w}\}$ is a linearly independent set with $|S| = n + 1$. By **Q10**, we must have $\mathbf{w} = \mathbf{0}$, a contradiction.

(b) By definition, $Z^{\perp} = \{\mathbf{v} \in V : \langle \mathbf{v}, \mathbf{0} \rangle = 0\}$. We verify the double inclusion $Z^{\perp} \subseteq V$ and $V \subseteq Z^{\perp}$.
By the above definition, $Z^{\perp} \subseteq V$. Let now $\mathbf{v} \in V$. Since $\langle \mathbf{v}, \mathbf{0} \rangle = 0$, $\mathbf{v} \in Z^{\perp}$. Hence, $V \subseteq Z^{\perp}$.

(c) By definition, $U^{\perp} = \{\mathbf{v} \in V : \langle \mathbf{v}, \mathbf{u} \rangle = 0 \text{ for all } \mathbf{u} \in U\}$. Let $\mathbf{v} \in U^{\perp}$. Then $\mathbf{v} \perp U$. But we also have $\mathbf{v} \in U$, so that $\mathbf{v} \perp \mathbf{v}$. Hence, $\langle \mathbf{v}, \mathbf{v} \rangle = 0$ and therefore $\mathbf{v} = \mathbf{0}$, by the non-negativity property of the inner product. Hence, $\mathbf{v} \in Z$ and therefore $U \cap U^{\perp} \subseteq Z$.
Note that we actually know from the direct sum property $V = U \oplus U^{\perp}$ that $U \cap U^{\perp} = Z$.

15. **(a)** This is done as in **Q27(a)**, Examples sheet 1.

(b) First, we find a basis for U. This is done using the approach in **Q27(b)**, Examples sheet 1.

$$x + 2y - z = 0 \overset{y=a, z=b}{\Longrightarrow} x = b - 2a,$$

so that

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} b - 2a \\ a \\ b \end{bmatrix} = a\begin{bmatrix} -2 \\ 1 \\ 0 \end{bmatrix} + b\begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} \in \operatorname{span}\left\{\begin{bmatrix} -2 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}\right\} =: \operatorname{span}\{\mathbf{u}_1, \mathbf{u}_2\} =: \operatorname{span} S.$$

Hence S is a spanning set for U. It is also linearly independent, since

$$a\begin{bmatrix} -2 \\ 1 \\ 0 \end{bmatrix} + b\begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \Longrightarrow a = b = 0.$$

Hence S is a basis for U. Since $\dim U = 2$, $\dim U^{\perp} = 1$; thus, we only need to find one vector $\mathbf{u} \perp U$, or equivalently, $\mathbf{u} \perp S$. This can be achieved using the Gram-Schmidt procedure. Let us choose $\mathbf{u}_3 \notin \operatorname{span} S$, e.g., $\mathbf{u}_3 = \mathbf{e}_1$. Consider now orthogonalising the set $\{\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3\}$, which is a basis for $V = \mathbb{R}^3$. This would achieve the following:

- replace the basis S of U with an orthogonal one;
- replace the vector $\mathbf{u}_3$ with a vector orthogonal to the other two, i.e., to S.

We have

$$\mathbf{u}_1' = \mathbf{u}_1 = \begin{bmatrix} -2 \\ 1 \\ 0 \end{bmatrix}, \quad \mathbf{u}_2' = \mathbf{u}_2 - \frac{\langle \mathbf{u}_2, \mathbf{u}_1' \rangle}{\langle \mathbf{u}_1', \mathbf{u}_1' \rangle} \mathbf{u}_1' = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} - \frac{(-2)}{5} \begin{bmatrix} -2 \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 1/5 \\ 2/5 \\ 1 \end{bmatrix},$$

$$\mathbf{u}_3' = \mathbf{u}_3 - \frac{\langle \mathbf{u}_3, \mathbf{u}_1' \rangle}{\langle \mathbf{u}_1', \mathbf{u}_1' \rangle} \mathbf{u}_1' - \frac{\langle \mathbf{u}_3, \mathbf{u}_2' \rangle}{\langle \mathbf{u}_2', \mathbf{u}_2' \rangle} \mathbf{u}_2' = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} - \frac{(-2)}{5} \begin{bmatrix} -2 \\ 1 \\ 0 \end{bmatrix} - \frac{1/5}{6/5} \begin{bmatrix} 1/5 \\ 2/5 \\ 1 \end{bmatrix} = \frac{1}{6} \begin{bmatrix} 1 \\ 2 \\ -1 \end{bmatrix}.$$

By the above construction, $\mathbf{u}_3' \perp S' = \{\mathbf{u}_1', \mathbf{u}_2'\}$. A spanning set for $U^\perp$ is therefore $\{\mathbf{u}_3'\}$.

ORTHOGONAL SETS. ORTHOGONALISATION.

16. Let $B = \{\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3\}$ denote the ordering of the vectors in the given basis set. We have

$$\mathbf{u}_1' = \mathbf{u}_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \quad \mathbf{u}_2' = \mathbf{u}_2 - \frac{\langle \mathbf{u}_2, \mathbf{u}_1' \rangle}{\langle \mathbf{u}_1', \mathbf{u}_1' \rangle} \mathbf{u}_1' = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} - \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix},$$

$$\mathbf{u}_3' = \mathbf{u}_3 - \frac{\langle \mathbf{u}_3, \mathbf{u}_1' \rangle}{\langle \mathbf{u}_1', \mathbf{u}_1' \rangle} \mathbf{u}_1' - \frac{\langle \mathbf{u}_3, \mathbf{u}_2' \rangle}{\langle \mathbf{u}_2', \mathbf{u}_2' \rangle} \mathbf{u}_2' = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} - \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} - \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}.$$

17. Since U contains two constraints in four variables, we can assign two generic values to two variables, in order to identify a basis, which we can then orthogonalise. Let $z = a, w = b$. Then

$$\begin{cases} x + y + z + w & = 0 \\ y + z & = 0 \end{cases} \xRightarrow{z=a,w=b} \begin{cases} x & = -b \\ y & = -a \end{cases}$$

so that

$$U = \left\{ \begin{bmatrix} x \\ y \\ z \\ w \end{bmatrix} : x = -b, y = -a, z = a, w = b, a, b \in \mathbb{R} \right\} = \left\{ \begin{bmatrix} -b \\ -a \\ a \\ b \end{bmatrix} : a, b, \in \mathbb{R} \right\} = \left\{ a \begin{bmatrix} 0 \\ -1 \\ 1 \\ 0 \end{bmatrix} + b \begin{bmatrix} -1 \\ 0 \\ 0 \\ 1 \end{bmatrix} : a, b, \in \mathbb{R} \right\}.$$

The above two vectors are already orthogonal, so they form the required basis set.

18. The Gram-Schmidt orthogonalisation procedure starts with an existing basis, say $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$, where $\mathbf{v}_i \in \mathbb{R}^4$. Let us place the vectors $\mathbf{v}_i$ in a matrix. This initial step can be written in matrix form as the identity

$$\begin{bmatrix} \times & \times & \times \\ \times & \times & \times \\ \times & \times & \times \\ \times & \times & \times \end{bmatrix} = \begin{bmatrix} \times & \times & \times \\ \times & \times & \times \\ \times & \times & \times \\ \times & \times & \times \end{bmatrix} \begin{bmatrix} 1 & & \\ & 1 & \\ & & 1 \end{bmatrix}.$$

The first step given in the question statement replaces $\mathbf{v}_1$ (highlighted in bold) with $\mathbf{v}_1' := \mathbf{v}_1 / \|\mathbf{v}_1\|$ (indicated by + entries):

$$\begin{bmatrix} \boldsymbol{\times} & \times & \times \\ \boldsymbol{\times} & \times & \times \\ \boldsymbol{\times} & \times & \times \\ \boldsymbol{\times} & \times & \times \end{bmatrix} = \begin{bmatrix} + & \times & \times \\ + & \times & \times \\ + & \times & \times \\ + & \times & \times \end{bmatrix} \begin{bmatrix} \oplus & & \\ & 1 & \\ & & 1 \end{bmatrix}$$

The second step of the process is

$$\mathbf{v}_2' = \mathbf{v}_2 - \frac{\langle \mathbf{v}_2, \mathbf{v}_1' \rangle}{\langle \mathbf{v}_1', \mathbf{v}_1' \rangle} \mathbf{v}_1'; \quad \mathbf{v}_2' = \mathbf{v}_2' / \|\mathbf{v}_2'\|,$$

which can be re-written as

$$\mathbf{v}_2 = \mathbf{v}_2' + \frac{\langle \mathbf{v}_2, \mathbf{v}_1' \rangle}{\langle \mathbf{v}_1', \mathbf{v}_1' \rangle}\mathbf{v}_1'.$$

This means that $\mathbf{v}_2$ is a linear combination of $\mathbf{v}_1'$ and $\mathbf{v}_2'$. This is expressed below as follows: the second column on the left (indicated in bold) is a linear combination of the columns 1 and 2 on the right (indicated by $+$); the corresponding coefficients are indicated by $\oplus$ in the second matrix on the right:

$$\begin{bmatrix} \times & \boldsymbol{\times} & \times \\ \times & \boldsymbol{\times} & \times \\ \times & \boldsymbol{\times} & \times \\ \times & \boldsymbol{\times} & \times \end{bmatrix} = \begin{bmatrix} + & + & \times \\ + & + & \times \\ + & + & \times \\ + & + & \times \end{bmatrix} \begin{bmatrix} + & \oplus & \\ & \oplus & \\ & & 1 \end{bmatrix}$$

The final step is

$$\mathbf{v}_3' = \mathbf{v}_3 - \frac{\langle \mathbf{v}_3, \mathbf{v}_1' \rangle}{\langle \mathbf{v}_1', \mathbf{v}_1' \rangle}\mathbf{v}_1' - \frac{\langle \mathbf{v}_3, \mathbf{v}_2' \rangle}{\langle \mathbf{v}_2', \mathbf{v}_2' \rangle}\mathbf{v}_2'; \quad \mathbf{v}_3' = \mathbf{v}_3' / \left\|\mathbf{v}_3'\right\| \Longleftrightarrow \mathbf{v}_3 = \mathbf{v}_3' + \frac{\langle \mathbf{v}_3, \mathbf{v}_1' \rangle}{\langle \mathbf{v}_1', \mathbf{v}_1' \rangle}\mathbf{v}_1' + \frac{\langle \mathbf{v}_3, \mathbf{v}_2' \rangle}{\langle \mathbf{v}_2', \mathbf{v}_2' \rangle}\mathbf{v}_2'.$$

This is expressed below following the conventions described above: column 3 (indicated in bold) is a linear combination of the three columns on the right indicated by $+$, with the coefficients indicated by $\oplus$ in the second matrix.

$$\begin{bmatrix} \times & \times & \boldsymbol{\times} \\ \times & \times & \boldsymbol{\times} \\ \times & \times & \boldsymbol{\times} \\ \times & \times & \boldsymbol{\times} \end{bmatrix} = \begin{bmatrix} + & + & + \\ + & + & + \\ + & + & + \\ + & + & + \end{bmatrix} \begin{bmatrix} + & + & \oplus \\ & + & \oplus \\ & & \oplus \end{bmatrix}.$$

The complete process is shown on the front page of the 2LA Lecture Notes. This can be viewed as a so-called factorisation of a matrix: $A = QR$, where R is an upper triangular matrix (a matrix with zeros below the main diagonal) and a matrix Q with orthogonal columns (see also Lecture 10 in the notes for more details).

19. The purpose of this question is to illustrate how one can generate an orthogonal basis via a three-term recurrence, i.e., a relation where any element in the basis is obtained by using the previous two. In contrast, the generic Gram-Schmidt procedure is a k-term recurrence, as the kth term is obtained using the previous $k-1$ terms.Consider the polynomials defined by the three-term recurrence

$$q_{i+1}(x) = (x - \alpha_{i+1})q_i(x) - \beta_i q_{i-1}(x), \qquad (i = 0, \ldots, n-1),$$

where $q_{-1}(x) = 0, \quad q_0(x) = 1$ and

$$\alpha_{i+1} = \frac{\langle xq_i, q_i \rangle_\mu}{\langle q_i, q_i \rangle_\mu} \quad (i = 0, 1, \ldots), \qquad \beta_i = \frac{\langle q_i, q_i \rangle_\mu}{\langle q_{i-1}, q_{i-1} \rangle_\mu} \quad (i = 1, 2, \ldots).$$

(a) Let $\mu(x) = 1$ and consider the inner product

$$\langle p, q \rangle = \int_{-1}^{1} p(x)q(x)\mathrm{d}x.$$

The evaluations below use the properties of even/odd functions when integrated over an interval which is symmetric with respect to the origin.We first find q_1, q_2.

$$q_1(x) = (x - \alpha_1)q_0(x) - \beta_0 q_{-1}(x) = x - \alpha_1, \quad \text{where } \alpha_1 = \frac{\langle x, 1 \rangle}{\langle 1, 1 \rangle} = 0 \Longrightarrow q_1(x) = x.$$

$$q_2(x) = (x - \alpha_2)q_1(x) - \beta_1 q_0(x), \quad \text{where } \begin{cases} \alpha_2 = \dfrac{\langle x^2, x \rangle}{\langle x, x \rangle} = 0 \\ \beta_1 = \dfrac{\langle x, x \rangle}{\langle 1, 1 \rangle} = \dfrac{1}{3} \end{cases} \Longrightarrow q_2(x) = x^2 - \frac{1}{3}.$$

We now check orthogonality.

$$\langle q_0, q_1\rangle = \langle 1, x\rangle = 0, \quad \langle q_1, q_2\rangle = \left\langle x, x^2 - 1/3\right\rangle = 0,$$

$$\langle q_0, q_2\rangle = \left\langle 1, x^2 - 1/3\right\rangle = 2\int_0^1 \left(x^2 - \frac{1}{3}\right) \mathrm{d}x = \left[\frac{x^3}{3} - \frac{x}{3}\right]_0^1 = 0.$$

The polynomials generated using this μ-inner product are known as **Legendre polynomials**.

(b) Let $\mu(x) = e^{-x}$ and consider the inner product

$$\langle p, q\rangle_\mu = \int_0^\infty e^{-x} p(x) q(x) \mathrm{d}x.$$

The evaluations below use the following definite integrals (which can be easily verified)

$$\int_0^\infty e^{-x} \mathrm{d}x = \int_0^\infty x e^{-x} \mathrm{d}x = 1, \quad \int_0^\infty e^{-x} x^2 \mathrm{d}x = 2, \quad \int_0^\infty e^{-x} x^3 \mathrm{d}x = 6.$$

We first find q_1, q_2.

$$q_1(x) = (x - \alpha_1) q_0(x) - \beta_0 q_{-1}(x) = x - \alpha_1, \quad \text{where } \alpha_1 = \frac{\langle x, 1\rangle_\mu}{\langle 1, 1\rangle_\mu} = 1 \Longrightarrow q_1(x) = x - 1.$$

$$q_2(x) = (x - \alpha_2) q_1(x) - \beta_1 q_0(x), \quad \text{where } \begin{cases} \alpha_2 = & \dfrac{\langle x(x-1), x-1\rangle_\mu}{\langle x-1, x-1\rangle_\mu} = 3 \\ \beta_1 = & \dfrac{\langle x-1, x-1\rangle_\mu}{\langle 1, 1\rangle_\mu} = 1 \end{cases} \Longrightarrow q_2(x) = x^2 - 4x + 2.$$

We now check orthogonality.

$$\langle q_0, q_1\rangle_\mu = \langle 1, x-1\rangle_\mu = 0, \quad \langle q_0, q_2\rangle_\mu = \left\langle 1, x^2 - 4x + 2\right\rangle_\mu = 2 - 4 + 2 = 0,$$

$$\langle q_1, q_2\rangle_\mu = \left\langle x - 1, x^2 - 4x + 2\right\rangle_\mu = \left\langle 1, x^3 - 4x^2 + 2x\right\rangle_\mu - \left\langle 1, x^2 - 4x + 2\right\rangle_\mu = 6 - 4\cdot 2 + 2 - 0 = 0.$$

The polynomials generated using this μ-inner product are known as **Laguerre polynomials**.

20. The purpose of this question is to highlight another family of polynomials that are orthogonal with respect to an inner product based on function evaluation, as opposed to function integration.

(a) Using the recurrence $C_{n+1}(x) = 2xC_n(x) - C_{n-1}(x)$, with $C_0(x) = 1, C_1(x) = x$ we find

$$C_2(x) = 2xC_1(x) - C_0(x) = 2x^2 - 1,$$

$$C_3(x) = 2xC_2(x) - C_1(x) = 4x^3 - 3x = 4x\left(x - \frac{\sqrt{3}}{2}\right)\left(x + \frac{\sqrt{3}}{2}\right).$$

(b) The expression for $\rho_{k,n}$ provided in the question becomes for $n = 3$

$$\rho_{k,3} = \cos\frac{(2k+1)\pi}{6}, \quad k = 0, 1, 2,$$

or

$$\rho_{0,3} = \cos\frac{\pi}{6} = \frac{\sqrt{3}}{2}, \quad \rho_{1,3} = \cos\frac{3\pi}{6} = 0, \quad \rho_{2,3} = \cos\frac{5\pi}{6} = -\frac{\sqrt{3}}{2},$$

which are indeed the roots of $C_3(x)$.

(c) Note first that $[\langle C_i, C_j\rangle]$ can be viewed as a matrix which is symmetric, due to the symmetry of the inner product, which is evident from its definition. Thus, the question amounts to checking that the matrix is diagonal; in turn, due to symmetry, this requires the evaluation of 3 entries only (above the main diagonal, corresponding to all i, j such that $i < j$). We find

$$\langle C_0, C_1\rangle = \sum_{k=0}^{2} C_0(\rho_{k,3})C_1(\rho_{k,3}) = \sum_{k=0}^{2} C_1(\rho_{k,3}) = \sum_{k=0}^{2} \rho_{k,3} = \frac{\sqrt{3}}{2} + 0 - \frac{\sqrt{3}}{2} = 0,$$

$$\langle C_0, C_2\rangle = \sum_{k=0}^{2} C_0(\rho_{k,3})C_2(\rho_{k,3}) = \sum_{k=0}^{2} C_2(\rho_{k,3}) = \left(2\cdot\frac{3}{4} - 1\right) + (2\cdot 0 - 1) + \left(2\cdot\frac{3}{4} - 1\right) = 0,$$

$$\langle C_1, C_2\rangle = \sum_{k=0}^{2} C_1(\rho_{k,3})C_2(\rho_{k,3}) = \sum_{k=0}^{2} \rho_{k,3}C_2(\rho_{k,3}) = \frac{\sqrt{3}}{2}\left(2\cdot\frac{3}{4} - 1\right) + 0 - \frac{\sqrt{3}}{2}\left(2\cdot\frac{3}{4} - 1\right) = 0.$$

Hence, the matrix $\langle C_i, C_j\rangle$ is diagonal, with diagonal entries

$$\langle C_0, C_0\rangle = \sum_{k=0}^{2} C_0(\rho_{k,3})C_0(\rho_{k,3}) = \sum_{k=0}^{2} 1^2 = 3,$$

$$\langle C_1, C_1\rangle = \sum_{k=0}^{2} C_1(\rho_{k,3})C_1(\rho_{k,3}) = \sum_{k=0}^{2} \rho_{k,3}^2 = \frac{3}{4} + 0 + \frac{3}{4} = \frac{3}{2},$$

$$\langle C_2, C_2\rangle = \sum_{k=0}^{2} C_2(\rho_{k,3})C_2(\rho_{k,3}) = \sum_{k=0}^{2} \left(2\rho_{k,3}^2 - 1\right)^2 = \left(2\cdot\frac{3}{4} - 1\right)^2 + (2\cdot 0 - 1)^2 + \left(2\cdot\frac{3}{4} - 1\right)^2 = \frac{3}{2}.$$

Note that the diagonal entries are all positive, due to the non-negativity of the inner product.