Section 6

Inner product spaces

We now head off in a different direction from the subject of representing linear transformations by matrices. We shall consider the topic of inner product spaces. These are vector spaces endowed with an "inner product" (essentially a generalisation of the dot product of vectors in $\mathbb{R}^3$) and are extremely important. If time allows (which it probably won't!), we shall see a link to the topic of diagonalisation.

Throughout this section (and the rest of the course), our base field F will be either $\mathbb{R}$ or $\mathbb{C}$. Recall that if $z = x + iy \in \mathbb{C}$, the *complex conjugate* of z is given by

$$\bar{z} = x - iy.$$

To save space and time, we shall use the complex conjugate even when $F = \mathbb{R}$. Thus, when $F = \mathbb{R}$ and $\bar{\alpha}$ appears, it means $\bar{\alpha} = \alpha$ for a scalar $\alpha \in \mathbb{R}$.

Definition 6.1 Let $F = \mathbb{R}$ or $\mathbb{C}$. An *inner product space* is a vector space V over F together with an *inner product*

$$\begin{aligned} V \times V &\to F \\ (v, w) &\mapsto \langle v, w \rangle \end{aligned}$$

such that

(i) $\langle u + v, w \rangle = \langle u, w \rangle + \langle v, w \rangle$ for all $u, v, w \in V$,

(ii) $\langle \alpha v, w \rangle = \alpha \langle v, w \rangle$ for all $v, w \in V$ and $\alpha \in F$,

(iii) $\langle v, w \rangle = \overline{\langle w, v \rangle}$ for all $v, w \in V$,

(iv) $\langle v, v \rangle$ is a real number satisfying $\langle v, v \rangle \geqslant 0$ for all $v \in V$,

(v) $\langle v, v \rangle = 0$ if and only if $v = \mathbf{0}$.

Thus, in the case **when $\boldsymbol{F} = \mathbb{R}$**, our inner product is symmetric in the sense that Condition (iii) then becomes

$$\langle v, w \rangle = \langle w, v \rangle \qquad \text{for all } v, w \in V.$$

Example 6.2 (i) The vector space $\mathbb{R}^n$ of column vectors of real numbers is an inner product space with respect to the usual *dot product*:

$$\left\langle \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}, \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} \right\rangle = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \cdot \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} = \sum_{i=1}^{n} x_i y_i.$$

Note that if $\boldsymbol{v} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$, then

$$\langle \boldsymbol{v}, \boldsymbol{v} \rangle = \sum_{i=1}^{n} x_i^2$$

and from this Condition (iv) follows immediately.

(ii) We can endow $\mathbb{C}^n$ with an inner product by introducing the complex conjugate:

$$\left\langle \begin{pmatrix} z_1 \\ z_2 \\ \vdots \\ z_n \end{pmatrix}, \begin{pmatrix} w_1 \\ w_2 \\ \vdots \\ w_n \end{pmatrix} \right\rangle = \sum_{i=1}^{n} z_i \bar{w}_i.$$

Note that if $\boldsymbol{v} = \begin{pmatrix} z_1 \\ z_2 \\ \vdots \\ z_n \end{pmatrix}$, then

$$\langle \boldsymbol{v}, \boldsymbol{v} \rangle = \sum_{i=1}^{n} z_i \bar{z}_i = \sum_{i=1}^{n} |z_i|^2.$$

(iii) If $a < b$, the set $C[a, b]$ of continuous functions $f \colon [a, b] \to \mathbb{R}$ is a real vector space when we define

$$(f + g)(x) = f(x) + g(x)$$
$$(\alpha f)(x) = \alpha \cdot f(x)$$

for $f, g \in C[a,b]$ and $\alpha \in \mathbb{R}$. In fact, $C[a,b]$ is an inner product space when we define

$$\langle f, g \rangle = \int_a^b f(x)g(x)\,\mathrm{d}x.$$

Since $f(x)^2 \geqslant 0$ for all x, we have

$$\langle f, f \rangle = \int_a^b f(x)^2\,\mathrm{d}x \geqslant 0.$$

(iv) The space $\mathcal{P}_n$ of real polynomials of degree at most n is a real vector space of dimension $n+1$. It becomes an inner product space by inheriting the inner product from above, for example:

$$\langle f, g \rangle = \int_0^1 f(x)g(x)\,\mathrm{d}x$$

for real polynomials $f(x), g(x) \in \mathcal{P}_n$.

We can also generalise these last two examples to complex-valued functions. For example, the complex vector space of polynomials

$$f(x) = \alpha_n x^n + \alpha_{n-1}x^{n-1} + \cdots + \alpha_1 x + \alpha_0$$

where $\alpha_0, \alpha_1, \dots, \alpha_n \in \mathbb{C}$ becomes an inner product space when we define

$$\langle f, g \rangle = \int_0^1 f(x)\overline{g(x)}\,\mathrm{d}x$$

where

$$\overline{f(x)} = \bar{\alpha}_n x^n + \bar{\alpha}_{n-1}x^{n-1} + \cdots + \bar{\alpha}_1 x + \bar{\alpha}_0.$$

Definition 6.3 Let V be an inner product space with inner product $\langle \cdot, \cdot \rangle$. The *norm* is the function $\| \cdot \| \colon V \to \mathbb{R}$ defined by

$$\|v\| = \sqrt{\langle v, v \rangle}.$$

(This makes sense since $\langle v, v \rangle \geqslant 0$ for all $v \in V$.)

Lemma 6.4 *Let V be an inner product space with inner product $\langle \cdot, \cdot \rangle$. Then*

(i) $\langle v, \alpha w \rangle = \bar{\alpha}\langle v, w \rangle$ *for all $v, w \in V$ and $\alpha \in F$;*

(ii) $\|\alpha v\| = |\alpha| \cdot \|v\|$ *for all $v \in V$ and $\alpha \in F$;*

(iii) $\|v\| > 0$ *whenever $v \neq \mathbf{0}$.*

Proof: (i)

$$\langle v, \alpha w\rangle = \overline{\langle \alpha w, v\rangle} = \overline{\alpha\langle w, v\rangle} = \bar{\alpha}\overline{\langle w, v\rangle} = \bar{\alpha}\langle v, w\rangle.$$

(ii)

$$\|\alpha v\|^2 = \langle \alpha v, \alpha v\rangle = \alpha\langle v, \alpha v\rangle = \alpha\bar{\alpha}\langle v, v\rangle = |\alpha|^2\|v\|^2$$

and taking square roots gives the result.

(iii) $\langle v, v\rangle > 0$ whenever $v \neq \mathbf{0}$. □

Theorem 6.5 (Cauchy–Schwarz Inequality) *Let V be an inner product space with inner product $\langle \cdot, \cdot \rangle$. Then*

$$|\langle u, v\rangle| \leqslant \|u\| \cdot \|v\|$$

for all $u, v \in V$.

Proof: If $v = \mathbf{0}$, then we see

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

Hence

$$|\langle u, v\rangle| = 0 = \|u\| \cdot \|v\|$$

as $\|v\| = 0$.

In the remainder of the proof we assume $v \neq \mathbf{0}$. Let α be a scalar, put $w = u + \alpha v$ and expand $\langle w, w\rangle$:

$$\begin{aligned} 0 \leqslant \langle w, w\rangle &= \langle u + \alpha v, u + \alpha v\rangle \\ &= \langle u, u\rangle + \alpha\langle v, u\rangle + \bar{\alpha}\langle u, v\rangle + \alpha\bar{\alpha}\langle v, v\rangle \\ &= \|u\|^2 + \alpha\overline{\langle u, v\rangle} + \bar{\alpha}\langle u, v\rangle + |\alpha|^2\|v\|^2. \end{aligned}$$

Now take $\alpha = -\langle u, v\rangle/\|v\|^2$. We deduce

$$\begin{aligned} 0 &\leqslant \|u\|^2 - \frac{\langle u, v\rangle \cdot \overline{\langle u, v\rangle}}{\|v\|^2} - \frac{\overline{\langle u, v\rangle}\langle u, v\rangle}{\|v\|^2} + \frac{|\langle u, v\rangle|^2}{\|v\|^4}\|v\|^2 \\ &= \|u\|^2 - \frac{|\langle u, v\rangle|^2}{\|v\|^2}, \end{aligned}$$

so

$$|\langle u, v\rangle|^2 \leqslant \|u\|^2\|v\|^2$$

and taking square roots gives the result. □

Corollary 6.6 (Triangle Inequality) *Let V be an inner product space. Then*

$$\|u + v\| \leqslant \|u\| + \|v\|$$

for all $u, v \in V$.

PROOF:

$$
\begin{aligned}
\|u+v\|^2 &= \langle u+v, u+v\rangle \\
&= \langle u,u\rangle + \langle u,v\rangle + \langle v,u\rangle + \langle v,v\rangle \\
&= \|u\|^2 + \langle u,v\rangle + \overline{\langle u,v\rangle} + \|v\|^2 \\
&= \|u\|^2 + 2\,\mathrm{Re}\langle u,v\rangle + \|v\|^2 \\
&\leqslant \|u\|^2 + 2|\langle u,v\rangle| + \|v\|^2 \\
&\leqslant \|u\|^2 + 2\|u\|\cdot\|v\| + \|v\|^2 && \text{(by Cauchy–Schwarz)} \\
&= (\|u\| + \|v\|)^2
\end{aligned}
$$

and taking square roots gives the result. □

The triangle inequality is a fundamental observation that tells us we can use the norm to measure distance on an inner product space in the same way that modulus $|x|$ is used to measure distance on $\mathbb{R}$ or $\mathbb{C}$. We can then perform analysis and speak of continuity and convergence. This topic is addressed in greater detail in the study of Functional Analysis.

Orthogonality and orthonormal bases

Definition 6.7 Let V be an inner product space.

(i) Two vectors v and w are said to be *orthogonal* if $\langle v, w\rangle = 0$.

(ii) A set $\mathscr{A}$ of vectors is *orthogonal* if every pair of vectors within it are orthogonal.

(iii) A set $\mathscr{A}$ of vectors is *orthonormal* if it is orthogonal and every vector in $\mathscr{A}$ has unit norm.

Thus the set $\mathscr{A} = \{v_1, v_2, \dots, v_k\}$ is orthonormal if

$$\langle v_i, v_j\rangle = \delta_{ij} = \begin{cases} 0 & \text{if } i \neq j \\ 1 & \text{if } i = j. \end{cases}$$

An *orthonormal basis* for an inner product space V is a basis which is itself an orthonormal set.

Example 6.8 (i) The standard basis $\mathscr{E} = \{\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n\}$ is an orthonormal basis for $\mathbb{R}^n$:

$$\langle \mathbf{e}_i, \mathbf{e}_j\rangle = \mathbf{e}_i \cdot \mathbf{e}_j = \begin{cases} 0 & \text{if } i \neq j \\ 1 & \text{if } i = j. \end{cases}$$

(ii) Consider the inner product space $C[-\pi, \pi]$, consisting of all continous functions $f \colon [-\pi, \pi] \to \mathbb{R}$, with inner product

$$\langle f, g \rangle = \int_{-\pi}^{\pi} f(x) g(x) \, \mathrm{d}x.$$

Define

$$\begin{aligned} e_0(x) &= \frac{1}{\sqrt{2\pi}} \\ e_n(x) &= \frac{1}{\sqrt{\pi}} \cos nx \\ f_n(x) &= \frac{1}{\sqrt{\pi}} \sin nx \end{aligned}$$

for $n = 1, 2, \dots$. These functions (without the scaling) were studied in MT2001. We have the following facts

$$\begin{aligned} \langle e_m, e_n \rangle &= 0 && \text{if } m \neq n, \\ \langle f_m, f_n \rangle &= 0 && \text{if } m \neq n, \\ \langle e_m, f_n \rangle &= 0 && \text{for all } m, n \end{aligned}$$

and

$$\|e_n\| = \|f_n\| = 1 \qquad \text{for all } n.$$

(The reason for the scaling factors is to achieve unit norm for each function.) The topic of Fourier series relates to expressing functions as linear combinations of the orthonormal set

$$\{\, e_0, e_n, f_n \mid n = 1, 2, 3, \dots \,\}.$$

Theorem 6.9 *An orthogonal set of non-zero vectors is linearly independent.*

PROOF: Let $\mathscr{A} = \{v_1, v_2, \dots, v_k\}$ be an orthogonal set of non-zero vectors. Suppose that

$$\sum_{i=1}^{k} \alpha_i v_i = \mathbf{0}.$$

Then, by linearity of the inner product in the first entry, for $j = 1, 2, \dots, k$ we have

$$0 = \left\langle \sum_{i=1}^{k} \alpha_i v_i, v_j \right\rangle = \sum_{i=1}^{k} \alpha_i \langle v_i, v_j \rangle = \alpha_j \|v_j\|^2,$$

since by assumption $\langle v_i, v_j \rangle = 0$ for $i \neq j$. Now $v_j \neq \mathbf{0}$, so $\|v_j\| \neq 0$. Hence we must have

$$\alpha_j = 0 \qquad \text{for all } j.$$

Thus $\mathscr{A}$ is linearly independent. $\square$

Problem: Given a (finite-dimensional) inner product space V, how do we find an orthonormal basis?

Theorem 6.10 (Gram–Schmidt Process) *Suppose that V is a finite-dimensional inner product space with basis $\{v_1, v_2, \dots, v_n\}$. The following procedure constructs an orthonormal basis $\{e_1, e_2, \dots, e_n\}$ for V.*

Step 1: *Define $e_1 = \frac{1}{\|v_1\|} v_1$.*

Step k: *Suppose $\{e_1, e_2, \dots, e_{k-1}\}$ has been constructed. Define*

$$w_k = v_k - \sum_{i=1}^{k-1} \langle v_k, e_i \rangle e_i$$

and

$$e_k = \frac{1}{\|w_k\|} w_k.$$

Proof: We claim that $\{e_1, e_2, \dots, e_k\}$ is always an orthonormal set contained in $\mathrm{Span}(v_1, v_2, \dots, v_k)$.

Step 1: v_1 is a non-zero vector, so $\|v_1\| \neq 0$ and hence $e_1 = \frac{1}{\|v_1\|} v_1$ is defined. Now

$$\|e_1\| = \left\| \frac{1}{\|v_1\|} v_1 \right\| = \frac{1}{\|v_1\|} \cdot \|v_1\| = 1.$$

Hence $\{e_1\}$ is an orthonormal set (there are no orthogonality conditions to check) and by definition $e_1 \in \mathrm{Span}(v_1)$.

Step k: Suppose that we have shown $\{e_1, e_2, \dots, e_{k-1}\}$ is an orthonormal set contained in $\mathrm{Span}(v_1, v_2, \dots, v_{k-1})$. Consider

$$w_k = v_k - \sum_{i=1}^{k-1} \langle v_k, e_i \rangle.$$

We claim that $w_k \neq \mathbf{0}$. Indeed, if $w_k = \mathbf{0}$, then

$$\begin{aligned} v_k = \sum_{i=1}^{k-1} \langle v_k, e_i \rangle e_i &\in \mathrm{Span}(e_1, \dots, e_{k-1}) \\ &\subseteq \mathrm{Span}(v_1, \dots, v_{k-1}). \end{aligned}$$

But this contradicts $\{v_1, v_2, \dots, v_n\}$ being linearly independent. Thus $w_k \neq \mathbf{0}$ and hence $e_k = \frac{1}{\|w_k\|} w_k$ is defined.

By construction $\|e_k\| = 1$ and

$$\begin{aligned} e_k &= \frac{1}{\|w_k\|}\left(v_k - \sum_{i=1}^{k-1} \langle v_k, e_i \rangle e_i\right) \\ &\in \operatorname{Span}(e_1, \dots, e_{k-1}, v_k) \\ &\subseteq \operatorname{Span}(v_1, \dots, v_{k-1}, v_k). \end{aligned}$$

It remains to check that e_k is orthogonal to e_j for $j = 1, 2, \dots, k-1$. We calculate

$$\begin{aligned} \langle w_k, e_j \rangle &= \left\langle v_k - \sum_{i=1}^{k-1} \langle v_k, e_i \rangle e_i, e_j \right\rangle \\ &= \langle v_k, e_j \rangle - \sum_{i=1}^{k-1} \langle v_k, e_i \rangle \langle e_i, e_j \rangle \\ &= \langle v_k, e_j \rangle - \langle v_k, e_j \rangle \|e_j\|^2 \qquad\qquad \text{(by inductive hypothesis)} \\ &= \langle v_k, e_j \rangle - \langle v_k, e_j \rangle = 0. \end{aligned}$$

Hence

$$\langle e_k, e_j \rangle = \left\langle \frac{1}{\|w_k\|} w_k, e_j \right\rangle = \frac{1}{\|w_k\|} \langle w_k, e_j \rangle = 0$$

for $j = 1, 2, \dots, k-1$.

This completes the induction. We conclude that, at the final stage, $\{e_1, e_2, \dots, e_n\}$ is an orthonormal set. Theorem 6.9 tells us this set is linearly independent and hence a basis for V (since $\dim V = n$). □

Example 6.11 *Consider $\mathbb{R}^3$ with the usual inner product. Find an orthonormal basis for the subspace U spanned by the vectors*

$$\boldsymbol{v}_1 = \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix} \qquad \textit{and} \qquad \boldsymbol{v}_2 = \begin{pmatrix} 2 \\ 3 \\ 1 \end{pmatrix}.$$

SOLUTION: We apply the Gram–Schmidt Process to $\{\boldsymbol{v}_1, \boldsymbol{v}_2\}$.

$$\|\boldsymbol{v}_1\|^2 = \left\langle \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix} \right\rangle = 1^2 + (-1)^2 = 2.$$

Take

$$\boldsymbol{e}_1 = \frac{1}{\|\boldsymbol{v}_1\|} \boldsymbol{v}_1 = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}.$$

Now

$$\langle \boldsymbol{v}_2, \boldsymbol{e}_1 \rangle = \left\langle \begin{pmatrix} 2 \\ 3 \\ 1 \end{pmatrix}, \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix} \right\rangle = \frac{1}{\sqrt{2}}(2-1) = \frac{1}{\sqrt{2}}.$$

Put

$$\begin{aligned} \boldsymbol{w}_2 &= \boldsymbol{v}_2 - \langle \boldsymbol{v}_2, \boldsymbol{e}_1 \rangle \boldsymbol{e}_1 \\ &= \begin{pmatrix} 2 \\ 3 \\ 1 \end{pmatrix} - \frac{1}{\sqrt{2}} \cdot \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix} \\ &= \begin{pmatrix} 2 \\ 3 \\ 1 \end{pmatrix} - \begin{pmatrix} 1/2 \\ 0 \\ -1/2 \end{pmatrix} = \begin{pmatrix} 3/2 \\ 3 \\ 3/2 \end{pmatrix}. \end{aligned}$$

So

$$\|\boldsymbol{w}_2\|^2 = (3/2)^2 + 3^2 + (3/2)^2 = \frac{27}{2}$$

and

$$\|\boldsymbol{w}_2\| = \frac{3\sqrt{3}}{\sqrt{2}}.$$

Take

$$\boldsymbol{e}_2 = \frac{1}{\|\boldsymbol{w}_2\|} \boldsymbol{w}_2 = \sqrt{\frac{2}{3}} \begin{pmatrix} 1/2 \\ 1 \\ 1/2 \end{pmatrix} = \frac{1}{\sqrt{6}} \begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix}.$$

Thus

$$\left\{ \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}, \frac{1}{\sqrt{6}} \begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix} \right\}$$

is an orthonormal basis for U. □

Example 6.12 (Laguerre polynomials) We can define an inner product on the space $\mathcal{P}$ of real polynomials $f(x)$ by

$$\langle f, g \rangle = \int_0^\infty f(x) g(x) \mathrm{e}^{-x} \, \mathrm{d}x.$$

The *Laguerre polynomials* form the orthonormal basis for $\mathcal{P}$ that is produced when we apply the Gram–Schmidt process to the standard basis

$$\{1, x, x^2, x^3, \dots\}$$

of monomials.

Determine the first three Laguerre polynomials.

SOLUTION: We apply the Gram–Schmidt process to the basis $\{1, x, x^2\}$ for the inner product space $\mathcal{P}_2$, of polynomials of degree at most 2, with inner product as above. We shall make use of the fact (determined by induction and integration by parts) that

$$\int_0^\infty x^n \mathrm{e}^{-x} \, \mathrm{d}x = n!$$

Define $f_i(x) = x^i$ for $i = 0, 1, 2$. Then

$$\|f_0\|^2 = \int_0^\infty f_0(x)^2 \mathrm{e}^{-x} \, \mathrm{d}x = \int_0^\infty \mathrm{e}^{-x} \, \mathrm{d}x = 1,$$

so

$$L_0(x) = \frac{1}{\|f_0\|} f_0(x) = 1.$$

We now calculate L_1. First

$$\langle f_1, L_0 \rangle = \int_0^\infty f_1(x) L_0(x) \mathrm{e}^{-x} \, \mathrm{d}x = \int_0^\infty x \mathrm{e}^{-x} \, \mathrm{d}x = 1.$$

The Gram-Schmidt process says we first put

$$w_1(x) = f_1(x) - \langle f_1, L_0 \rangle L_0(x) = x - 1.$$

Now

$$\begin{aligned} \|w_1\|^2 &= \int_0^\infty w_1(x)^2 \mathrm{e}^{-x} \, \mathrm{d}x \\ &= \int_0^\infty (x^2 \mathrm{e}^{-x} - 2x\mathrm{e}^{-x} + \mathrm{e}^{-x}) \, \mathrm{d}x \\ &= 2 - 2 + 1 = 1. \end{aligned}$$

Hence

$$L_1(x) = \frac{1}{\|w_1\|} w_1(x) = x - 1.$$

In the next step of the Gram–Schmidt process, we calculate

$$\langle f_2, L_0 \rangle = \int_0^\infty x^2 \mathrm{e}^{-x} \, \mathrm{d}x = 2$$

and

$$\begin{aligned} \langle f_2, L_1 \rangle &= \int_0^\infty x^2 (x-1) \mathrm{e}^{-x} \, \mathrm{d}x \\ &= \int_0^\infty (x^3 \mathrm{e}^{-x} - x^2 \mathrm{e}^{-x}) \, \mathrm{d}x \\ &= 3! - 2! = 6 - 2 = 4. \end{aligned}$$

So we put

$$\begin{aligned} w_2(x) &= f_2(x) - \langle f_2, L_0 \rangle L_0(x) - \langle f_2, L_1 \rangle L_1(x) \\ &= x^2 - 4(x-1) - 2 \\ &= x^2 - 4x + 2. \end{aligned}$$

Now

$$\begin{aligned} \|w_2\|^2 &= \int_0^\infty w_2(x)^2 \mathrm{e}^{-x} \, \mathrm{d}x \\ &= \int_0^\infty (x^4 - 8x^3 + 20x^2 - 16x + 4)\mathrm{e}^{-x} \, \mathrm{d}x \\ &= 4! - 8 \cdot 3! + 20 \cdot 2! - 16 + 4 \\ &= 4. \end{aligned}$$

Hence we take

$$L_2(x) = \frac{1}{\|w_2\|} w_2(x) = \tfrac{1}{2}(x^2 - 4x + 2).$$

Similar calculations can be performed to determine L_3, L_4, …, but they become increasingly more complicated (and consequently less suitable for presenting on a whiteboard!). □

Example 6.13 Define an inner product on the space $\mathcal{P}$ of real polynomials by

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

Applying the Gram–Schmidt process to the monomials $\{1, x, x^2, x^3, \dots\}$ produces an orthonormal basis (with respect to this inner product). The polynomials produced are scalar multiples of the *Legendre polynomials*:

$$\begin{aligned} P_0(x) &= 1 \\ P_1(x) &= x \\ P_2(x) &= \tfrac{1}{2}(3x^2 - 1) \\ &\vdots \end{aligned}$$

The set $\{ P_n(x) \mid n = 0, 1, 2, \dots \}$ of Legendre polynomials is *orthogonal*, but *not* orthonormal. This is the reason why the Gram–Schmidt process only produces a scalar multiple of them. The scalars appearing are determined by the norms of the P_n with respect to this inner product.

For example,

$$\|P_0\|^2 = \int_{-1}^{1} P_0(x)^2 \, \mathrm{d}x = \int_{-1}^{1} \mathrm{d}x = 2,$$

so the polynomial of unit norm produced will be $\frac{1}{\sqrt{2}} P_0(x)$. Similar calculations (of increasing length) can be performed for the other polynomials.

The *Hermite polynomials* form an orthogonal set in the space $\mathscr{P}$ when we endow it with the following inner product

$$\langle f, g \rangle = \int_{-\infty}^{\infty} f(x) g(x) \mathrm{e}^{-x^2/2} \, \mathrm{d}x.$$

Again the orthonomal basis produced by applying the Gram–Schmidt process to the monomials are scalar multiples of the Hermite polynomials.

Orthogonal complements

Definition 6.14 Let V be an inner product space. If U is a subspace of V, the *orthogonal complement* to U is

$$U^{\perp} = \{ \, v \in V \mid \langle v, u \rangle = 0 \text{ for all } u \in U \, \}.$$

Thus $U^{\perp}$ consists of those vectors which are orthogonal to every single vector in U.

Lemma 6.15 *Let V be an inner product space and U be a subspace of V. Then*

(i) *$U^{\perp}$ is a subspace of V, and*

(ii) *$U \cap U^{\perp} = \{\mathbf{0}\}$.*

Proof: (i) First note $\langle \mathbf{0}, u \rangle = 0$ for all $u \in U$, so $\mathbf{0} \in U^{\perp}$. Now let $v, w \in U^{\perp}$ and $\alpha \in F$. Then

$$\langle v + w, u \rangle = \langle v, u \rangle + \langle w, u \rangle = 0 + 0 = 0$$

and

$$\langle \alpha v, u \rangle = \alpha \langle v, u \rangle = \alpha \cdot 0 = 0$$

for all $u \in U$. So we deduce $v + w \in U^{\perp}$ and $\alpha v \in U^{\perp}$. This shows that $U^{\perp}$ is a subspace.

(ii) Let $u \in U \cap U^{\perp}$. Then

$$\|u\|^2 = \langle u, u \rangle = 0$$

(since the element u is, in particular, orthogonal to itself). Hence $u = \mathbf{0}$. □

Theorem 6.16 *Let V be a finite-dimensional inner product space and U be a subspace of V. Then*

$$V = U \oplus U^{\perp}.$$

PROOF: We already know that $U \cap U^{\perp} = \{\mathbf{0}\}$, so it remains to show $V = U + U^{\perp}$.

Let $\{v_1, v_2, \dots, v_k\}$ be a basis for U. Extend it to a basis

$$\mathscr{B} = \{v_1, v_2, \dots, v_k, w_{k+1}, \dots, w_n\}$$

for V. Now apply the Gram–Schmidt process to $\mathscr{B}$ and hence produce an orthonormal basis $\mathscr{E} = \{e_1, e_2, \dots, e_n\}$ for V. By construction,

$$\{e_1, e_2, \dots, e_k\} \subseteq \operatorname{Span}(v_1, v_2, \dots, v_k) = U$$

and, since it is an orthonormal set, $\{e_1, e_2, \dots, e_k\}$ is a linearly independent set of size $k = \dim U$. Therefore $\{e_1, e_2, \dots, e_k\}$ is a basis for U.

Hence any vector $u \in U$ can be uniquely written as $u = \sum_{i=1}^{k} \alpha_i e_i$. Then for all such u

$$\langle u, e_j \rangle = \left\langle \sum_{i=1}^{k} \alpha_i e_i, e_j \right\rangle = \sum_{i=1}^{k} \alpha_i \langle e_i, e_j \rangle = 0$$

for $j = k+1$, $k+2$, $\dots$, n. That is,

$$e_{k+1}, e_{k+2}, \dots, e_n \in U^{\perp}.$$

Now if $v \in V$, we can write

$$v = \beta_1 e_1 + \cdots + \beta_k e_k + \beta_{k+1} e_{k+1} + \cdots + \beta_n e_n$$

for some scalars β_1, β_2, $\dots$, β_n and

$$\beta_1 e_1 + \cdots + \beta_k e_k \in U \qquad \text{and} \qquad \beta_{k+1} e_{k+1} + \cdots + \beta_n e_n \in U^{\perp}.$$

This shows that every vector in V is the sum of a vector in U and one in $U^{\perp}$, so

$$V = U + U^{\perp} = U \oplus U^{\perp},$$

as required to complete the proof. □

Once we have a direct sum, we can consider an associated projection map. In particular, we have the projection $P_U \colon V \to V$ onto U associated to the decomposition $V = U \oplus U^{\perp}$. This is given by

$$P_U(v) = u$$

where $v = u + w$ is the unique decomposition of v with $u \in U$ and $w \in U^{\perp}$.

Theorem 6.17 *Let V be a finite-dimensional inner product space and U be a subspace of V. Let $P_U \colon V \to V$ be the projection map onto U associated to the direct sum decomposition $V = U \oplus U^{\perp}$. If $v \in V$, then $P_U(v)$ is the vector in U that is closest to v.*

PROOF: Recall that the norm $\|\cdot\|$ determines the distance between two vectors, specifically $\|v-u\|$ is the distance from v to u. Write $v = u_0 + w_0$ where $u_0 \in U$ and $w_0 \in U^\perp$, so that $P_U(v) = u_0$. Then if u is any vector in U,

$$\begin{aligned}
\|v-u\|^2 &= \|v-u_0+(u_0-u)\|^2 \\
&= \|w_0+(u_0-u)\|^2 \\
&= \langle w_0+(u_0-u), w_0+(u_0-u)\rangle \\
&= \langle w_0, w_0\rangle + \langle w_0, u_0-u\rangle + \langle u_0-u, w_0\rangle + \langle u_0-u, u_0-u\rangle \\
&= \|w_0\|^2 + \|u_0-u\|^2 \qquad \text{(since } w_0 \text{ is orthogonal to } u_0-u \in U\text{)} \\
&\geqslant \|w_0\|^2 \qquad \text{(since } \|u_0-u\| \geqslant 0\text{)} \\
&= \|v-u_0\|^2 \\
&= \|v-P_U(v)\|^2.
\end{aligned}$$

Hence

$$\|v-u\| \geqslant \|v-P_U(v)\| \qquad \text{for all } u \in U.$$

This proves the theorem: $P_U(v)$ is closer to v than any other vector in U.

$\square$

Example 6.18 *Find the distance from the vector* $\boldsymbol{w}_0 = \begin{pmatrix} -1 \\ 5 \\ 1 \end{pmatrix}$ *in* $\mathbb{R}^3$ *to the subspace*

$$U = \operatorname{Span}\left(\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ -2 \end{pmatrix} \right).$$

SOLUTION: We need to find $U^\perp$, which must be a 1-dimensional subspace since $\mathbb{R}^3 = U \oplus U^\perp$. We solve the condition $\langle \boldsymbol{v}, \boldsymbol{u}\rangle = 0$ for all $\boldsymbol{u} \in U$:

$$\left\langle \begin{pmatrix} x \\ y \\ z \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \right\rangle = \begin{pmatrix} x \\ y \\ z \end{pmatrix} \cdot \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} = x+y+z$$

and

$$\left\langle \begin{pmatrix} x \\ y \\ z \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ -2 \end{pmatrix} \right\rangle = \begin{pmatrix} x \\ y \\ z \end{pmatrix} \cdot \begin{pmatrix} 0 \\ 1 \\ -2 \end{pmatrix} = y-2z.$$

Hence

$$x+y+z = y-2z = 0.$$

Given arbitrary z, we take $y = 2z$ and $x = -y-z = -3z$. Therefore

$$U^\perp = \left\{ \begin{pmatrix} -3z \\ 2z \\ z \end{pmatrix} \;\middle|\; z \in \mathbb{R} \right\} = \operatorname{Span}\left(\begin{pmatrix} -3 \\ 2 \\ 1 \end{pmatrix} \right).$$

The closest vector in U to $\boldsymbol{w}_0$ is $P_U(\boldsymbol{w}_0)$ where $P_U \colon \mathbb{R}^3 \to \mathbb{R}^3$ is the projection onto U associated to $\mathbb{R}^3 = U \oplus U^{\perp}$. To determine this we solve

$$\boldsymbol{w}_0 = \begin{pmatrix} -1 \\ 5 \\ 1 \end{pmatrix} = \alpha \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} + \beta \begin{pmatrix} 0 \\ 1 \\ -2 \end{pmatrix} + \gamma \begin{pmatrix} -3 \\ 2 \\ 1 \end{pmatrix},$$

so

$$\begin{aligned} \alpha \qquad - 3\gamma &= -1 && (6.1)\\ \alpha \quad + \beta + 2\gamma &= 5 && (6.2)\\ \alpha - 2\beta \quad + \gamma &= 1. && (6.3) \end{aligned}$$

Multiplying (6.2) by 2 and adding to (6.3) gives

$$3\alpha + 5\gamma = 11.$$

Then multiplying (6.1) by 3 and subtracting gives

$$14\gamma = 14.$$

Hence $\gamma = 1$, $\alpha = -1 + 3\gamma = 2$ and $\beta = 5 - \alpha - 2\gamma = 1$. We conclude

$$\begin{aligned} \boldsymbol{w}_0 &= 2 \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} + \begin{pmatrix} 0 \\ 1 \\ -2 \end{pmatrix} + \begin{pmatrix} -3 \\ 2 \\ 1 \end{pmatrix} \\ &= P_U(\boldsymbol{w}_0) + \begin{pmatrix} -3 \\ 2 \\ 1 \end{pmatrix}. \end{aligned}$$

We know $P_U(\boldsymbol{w}_0)$ is the nearest vector in U to $\boldsymbol{w}_0$, so the distance of $\boldsymbol{w}_0$ to U is

$$\|\boldsymbol{w}_0 - P_U(\boldsymbol{w}_0)\| = \left\| \begin{pmatrix} -3 \\ 2 \\ 1 \end{pmatrix} \right\| = \sqrt{(-3)^2 + 2^2 + 1^2} = \sqrt{14}.$$

$\square$

Example 6A (Exam Paper, January 2010) *Let $\langle \cdot, \cdot \rangle$ denote the usual inner product on $\mathbb{R}^4$, namely*

$$\langle \boldsymbol{u}, \boldsymbol{v} \rangle = \sum_{i=1}^{4} x_i y_i$$

for $\boldsymbol{u} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix}$ *and* $\boldsymbol{v} = \begin{pmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \end{pmatrix}$.

(i) *Apply the Gram–Schmidt Process to the set*

$$\mathscr{A} = \left\{ \begin{pmatrix} 1 \\ 1 \\ -1 \\ 1 \end{pmatrix}, \begin{pmatrix} 3 \\ 1 \\ -2 \\ 2 \end{pmatrix}, \begin{pmatrix} 2 \\ -4 \\ 3 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} \right\}$$

to produce an orthonormal basis for $\mathbb{R}^4$.

(ii) *Let* U *be the subspace of* $\mathbb{R}^4$ *spanned by*

$$\mathscr{B} = \left\{ \begin{pmatrix} 1 \\ 1 \\ -1 \\ 1 \end{pmatrix}, \begin{pmatrix} 3 \\ 1 \\ -2 \\ 2 \end{pmatrix} \right\}.$$

Find a basis for the orthogonal complement to U *in* $\mathbb{R}^4$.

(iii) *Find the vector in* U *that is nearest to* $\begin{pmatrix} 2 \\ -4 \\ 3 \\ 1 \end{pmatrix}$.

Solution: (i) Define

$$\boldsymbol{v}_1 = \begin{pmatrix} 1 \\ 1 \\ -1 \\ 1 \end{pmatrix}, \quad \boldsymbol{v}_2 = \begin{pmatrix} 3 \\ 1 \\ -2 \\ 2 \end{pmatrix}, \quad \boldsymbol{v}_3 = \begin{pmatrix} 2 \\ -4 \\ 3 \\ 1 \end{pmatrix}, \quad \boldsymbol{v}_4 = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}.$$

We perform the steps of the Gram–Schmidt Process:

Step 1:

$$\|\boldsymbol{v}_1\|^2 = 1^2 + 1^2 + (-1)^2 + 1^2 = 4,$$

so

$$\|\boldsymbol{v}_1\| = 2.$$

Take

$$\boldsymbol{e}_1 = \frac{1}{\|\boldsymbol{v}_1\|}\boldsymbol{v}_1 = \frac{1}{2}\begin{pmatrix} 1 \\ 1 \\ -1 \\ 1 \end{pmatrix}.$$

Step 2:

$$\langle \boldsymbol{v}_2, \boldsymbol{e}_1 \rangle = \frac{1}{2} \left\langle \begin{pmatrix} 3 \\ 1 \\ -2 \\ 2 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ -1 \\ 1 \end{pmatrix} \right\rangle = \tfrac{1}{2}(3+1+2+2) = 4.$$

Take

$$\boldsymbol{w}_2 = \boldsymbol{v}_2 - \langle \boldsymbol{v}_2, \boldsymbol{e}_1 \rangle \boldsymbol{e}_1 = \begin{pmatrix} 3 \\ 1 \\ -2 \\ 2 \end{pmatrix} - 2 \begin{pmatrix} 1 \\ 1 \\ -1 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ -1 \\ 0 \\ 0 \end{pmatrix}.$$

Then

$$\|\boldsymbol{w}_2\|^2 = 1^2 + (-1)^2 = 2,$$

so take

$$\boldsymbol{e}_2 = \frac{1}{\|\boldsymbol{w}_2\|} \boldsymbol{w}_2 = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -1 \\ 0 \\ 0 \end{pmatrix}.$$

Step 3:

$$\langle \boldsymbol{v}_3, \boldsymbol{e}_1 \rangle = \frac{1}{2} \left\langle \begin{pmatrix} 2 \\ -4 \\ 3 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ -1 \\ 1 \end{pmatrix} \right\rangle = \tfrac{1}{2}(2-4-3+1) = -2$$

$$\langle \boldsymbol{v}_3, \boldsymbol{e}_2 \rangle = \frac{1}{\sqrt{2}} \left\langle \begin{pmatrix} 2 \\ -4 \\ 3 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ -1 \\ 0 \\ 0 \end{pmatrix} \right\rangle = \tfrac{1}{\sqrt{2}}(2+4+0+0) = \frac{6}{\sqrt{2}} = 3\sqrt{2}.$$

Take

$$\begin{aligned} \boldsymbol{w}_3 &= \boldsymbol{v}_3 - \langle \boldsymbol{v}_3, \boldsymbol{e}_1 \rangle \boldsymbol{e}_1 - \langle \boldsymbol{v}_3, \boldsymbol{e}_2 \rangle \boldsymbol{e}_2 \\ &= \begin{pmatrix} 2 \\ -4 \\ 3 \\ 1 \end{pmatrix} + 2 \cdot \frac{1}{2} \begin{pmatrix} 1 \\ 1 \\ -1 \\ 1 \end{pmatrix} - 3\sqrt{2} \cdot \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -1 \\ 0 \\ 0 \end{pmatrix} \\ &= \begin{pmatrix} 2 \\ -4 \\ 3 \\ 1 \end{pmatrix} + \begin{pmatrix} 1 \\ 1 \\ -1 \\ 1 \end{pmatrix} - 3 \begin{pmatrix} 1 \\ -1 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 2 \\ 2 \end{pmatrix}. \end{aligned}$$

Then

$$\|\boldsymbol{w}_3\|^2 = 2^2 + 2^2 = 8,$$

so take

$$\boldsymbol{e}_3 = \frac{1}{\|\boldsymbol{w}_3\|}\boldsymbol{w}_3 = \frac{1}{2\sqrt{2}}\boldsymbol{w}_3 = \frac{1}{\sqrt{2}}\begin{pmatrix}0\\0\\1\\1\end{pmatrix}.$$

Step 4:

$$\langle \boldsymbol{v}_4, \boldsymbol{e}_1\rangle = \frac{1}{0}\left\langle \begin{pmatrix}1\\0\\0\\0\end{pmatrix}, \begin{pmatrix}1\\1\\-1\\1\end{pmatrix}\right\rangle = \tfrac{1}{2}$$

$$\langle \boldsymbol{v}_4, \boldsymbol{e}_2\rangle = \frac{1}{\sqrt{2}}\left\langle \begin{pmatrix}1\\0\\0\\0\end{pmatrix}, \begin{pmatrix}1\\-1\\0\\0\end{pmatrix}\right\rangle = \frac{1}{\sqrt{2}}$$

$$\langle \boldsymbol{v}_4, \boldsymbol{e}_3\rangle = \frac{1}{\sqrt{2}}\left\langle \begin{pmatrix}1\\0\\0\\0\end{pmatrix}, \begin{pmatrix}0\\0\\1\\1\end{pmatrix}\right\rangle = 0.$$

Take

$$\begin{aligned}
\boldsymbol{w}_4 &= \boldsymbol{v}_4 - \langle \boldsymbol{v}_4, \boldsymbol{e}_1\rangle\boldsymbol{e}_1 - \langle \boldsymbol{v}_4, \boldsymbol{e}_2\rangle\boldsymbol{e}_2 - \langle \boldsymbol{v}_4, \boldsymbol{e}_3\rangle\boldsymbol{e}_3 \\
&= \begin{pmatrix}1\\0\\0\\0\end{pmatrix} - \frac{1}{2}\cdot\frac{1}{2}\begin{pmatrix}1\\1\\-1\\1\end{pmatrix} - \frac{1}{\sqrt{2}}\cdot\frac{1}{\sqrt{2}}\begin{pmatrix}1\\-1\\0\\0\end{pmatrix} \\
&= \begin{pmatrix}1\\0\\0\\0\end{pmatrix} - \frac{1}{4}\begin{pmatrix}1\\1\\-1\\1\end{pmatrix} - \frac{1}{2}\begin{pmatrix}1\\-1\\0\\0\end{pmatrix} \\
&= \begin{pmatrix}1/4\\1/4\\1/4\\-1/4\end{pmatrix}.
\end{aligned}$$

Then

$$\|\boldsymbol{w}_4\|^2 = \left(\tfrac{1}{4}\right)^2 + \left(\tfrac{1}{4}\right)^2 + \left(\tfrac{1}{4}\right)^2 + \left(-\tfrac{1}{4}\right)^2 = \frac{1}{4},$$

so take

$$\boldsymbol{e}_4 = \frac{1}{\|\boldsymbol{w}_4\|}\boldsymbol{w}_4 = \frac{1}{2}\begin{pmatrix}1\\1\\1\\-1\end{pmatrix}.$$

Hence

$$\left\{ \frac{1}{2}\begin{pmatrix} 1 \\ 1 \\ -1 \\ 1 \end{pmatrix}, \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ -1 \\ 0 \\ 0 \end{pmatrix}, \frac{1}{\sqrt{2}}\begin{pmatrix} 0 \\ 0 \\ 1 \\ 1 \end{pmatrix}, \frac{1}{2}\begin{pmatrix} 1 \\ 1 \\ 1 \\ -1 \end{pmatrix} \right\}$$

is the orthonormal basis for $\mathbb{R}^4$ obtained by applying the Gram–Schmidt Process to $\mathscr{A}$.

(ii) In terms of the notation of (i), $U = \mathrm{Span}(\boldsymbol{v}_1, \boldsymbol{v}_2)$. However, the method of the Gram–Schmidt Process (see the proof of Theorem 6.10) shows that

$$\mathrm{Span}(\boldsymbol{e}_1, \boldsymbol{e}_2) = \mathrm{Span}(\boldsymbol{v}_1, \boldsymbol{v}_2) = U.$$

If $\boldsymbol{v} = \alpha\boldsymbol{e}_1 + \beta\boldsymbol{e}_2 + \gamma\boldsymbol{e}_3 + \delta\boldsymbol{e}_4$ is an arbitrary vector of $\mathbb{R}^4$ (expressed in terms of our orthonormal basis), then

$$\langle \boldsymbol{v}, \boldsymbol{e}_1 \rangle = \alpha \qquad \text{and} \qquad \langle \boldsymbol{v}, \boldsymbol{e}_2 \rangle = \beta.$$

Hence if $\boldsymbol{v} \in U^\perp$, then in particular $\alpha = \beta = 0$, so $U^\perp \subseteq \mathrm{Span}(\boldsymbol{e}_3, \boldsymbol{e}_4)$. Conversely, if $\boldsymbol{v} = \gamma\boldsymbol{e}_3 + \delta\boldsymbol{e}_4 \in \mathrm{Span}(\boldsymbol{e}_3, \boldsymbol{e}_4)$, then

$$\langle \zeta\boldsymbol{e}_1 + \eta\boldsymbol{e}_2, \gamma\boldsymbol{e}_3 + \delta\boldsymbol{e}_4 \rangle = 0$$

since $\langle \boldsymbol{e}_i, \boldsymbol{e}_j \rangle = 0$ for $i \neq j$. Hence every vector in $\mathrm{Span}(\boldsymbol{e}_3, \boldsymbol{e}_4)$ is orthogonal to every vector in U and we conclude

$$U^\perp = \mathrm{Span}(\boldsymbol{e}_3, \boldsymbol{e}_4).$$

Thus $\{\boldsymbol{e}_3, \boldsymbol{e}_4\}$ is a basis for $U^\perp$.

(iii) Let $P \colon V \to V$ be the projection onto U associated to the direct sum decomposition $V = U \oplus U^\perp$. Then $P(\boldsymbol{v})$ is the vector in U closest to v. Now in our application of the Gram–Schmidt Process,

$$\boldsymbol{w}_3 = \boldsymbol{v}_3 - \langle \boldsymbol{v}_3, \boldsymbol{e}_1 \rangle \boldsymbol{e}_1 - \langle \boldsymbol{v}_3, \boldsymbol{e}_2 \rangle \boldsymbol{e}_2,$$

so

$$P(\boldsymbol{w}_3) = P(\boldsymbol{v}_3) - \langle \boldsymbol{v}_3, \boldsymbol{e}_1 \rangle P(\boldsymbol{e}_1) - \langle \boldsymbol{v}_3, \boldsymbol{e}_1 \rangle P(\boldsymbol{e}_2).$$

Therefore

$$\boldsymbol{0} = P(\boldsymbol{v}_3) - \langle \boldsymbol{v}_3, \boldsymbol{e}_1 \rangle \boldsymbol{e}_1 - \langle \boldsymbol{v}_3, \boldsymbol{e}_2 \rangle \boldsymbol{e}_2,$$

since $\boldsymbol{w}_3 = \|\boldsymbol{w}_3\| \boldsymbol{e}_3 \in U^\perp$ and $\boldsymbol{e}_1, \boldsymbol{e}_2 \in U$. Hence the closest vector in U to $\boldsymbol{v}_3$ is

$$\begin{aligned} P(\boldsymbol{v}_3) &= \langle \boldsymbol{v}_3, \boldsymbol{e}_1 \rangle \boldsymbol{e}_1 + \langle \boldsymbol{v}_3, \boldsymbol{e}_2 \rangle \boldsymbol{e}_2 \\ &= (-2) \cdot \frac{1}{2}\begin{pmatrix} 1 \\ 1 \\ -1 \\ 1 \end{pmatrix} + 3\sqrt{2} \cdot \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ -1 \\ 0 \\ 0 \end{pmatrix} \end{aligned}$$

$$= -\begin{pmatrix} 1 \\ 1 \\ -1 \\ 1 \end{pmatrix} + 3\begin{pmatrix} 1 \\ -1 \\ 0 \\ 0 \end{pmatrix}$$

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

$\square$