Chapter 14

Orthogonality

Definition 13.19 introduced the angle between two vectors in an inner product space, for vector spaces all of whose scalars are real numbers.

In particular, two (non-zero) vectors, $\mathbf{u}$ and $\mathbf{v}$, are perpendicular to each other, or *orthogonal* if the angle between them is a right angle. Since $\cos\frac{\pi}{2} = 0$, this is equivalent to $\langle\!\langle \mathbf{u}, \mathbf{v} \rangle\!\rangle = 0$.

Observe that this is expressed purely in terms of the inner product, without appeal to the notion of angle, so we have no need to restrict ourselves to real scalars. Hence we can define orthogonality in any inner product space.

Definition 14.1. Let $\langle\!\langle\ ,\ \rangle\!\rangle$ be an inner product on the vector space V over the subfield $\mathbb{F}$ of $\mathbb{C}$. The vectors $\mathbf{u}, \mathbf{v} \in V$ are said to be *orthogonal* if and only if $\langle\!\langle \mathbf{u}, \mathbf{v} \rangle\!\rangle = 0$.

Before investigating orthogonality in any detail, we provide a geometric application. Orthogonality generalises the notion of a right angle,which is central to Pythagoras' Theorem in geometry. We prove a generalised Pythagoras' Theorem.

Theorem 14.2 (**Pythagoras' Theorem**). *Let $\langle\!\langle\ ,\ \rangle\!\rangle$ be an inner product on V and $\|\ \ \|$ the norm it induces. If $\mathbf{u}, \mathbf{v} \in V$ are orthogonal, then*

$$\|\mathbf{u} + \mathbf{v}\|^2 = \|\mathbf{u}\|^2 + \|\mathbf{v}\|^2.$$

Proof. Since $\mathbf{u}, \mathbf{v}$ are orthogonal, $\langle\!\langle \mathbf{u}, \mathbf{v} \rangle\!\rangle = 0$. Thus

$$\begin{aligned} \|\mathbf{u} + \mathbf{v}\|^2 &:= \langle\!\langle \mathbf{u} + \mathbf{v}, \mathbf{u} + \mathbf{v} \rangle\!\rangle \\ &= \langle\!\langle \mathbf{u}, \mathbf{u} \rangle\!\rangle + \langle\!\langle \mathbf{u}, \mathbf{v} \rangle\!\rangle + \langle\!\langle \mathbf{v}, \mathbf{u} \rangle\!\rangle + \langle\!\langle \mathbf{v}, \mathbf{v} \rangle\!\rangle \\ &= \|\mathbf{u}\|^2 + \|\mathbf{v}\|^2 && \text{by orthogonality} \end{aligned}$$

□

The next lemma is an immediate consequence of the axioms for inner products.

Lemma 14.3. *Let $(V, \langle\!\langle\ ,\ \rangle\!\rangle)$ be an inner product space. Then $\mathbf{0}_V$ is orthogonal to every $\mathbf{v} \in V$.*

Another elementary consequence is that non-zero orthogonal vectors must be linearly independent.

Theorem 14.4. *Let $(V, \langle\!\langle\ ,\ \rangle\!\rangle)$ be an inner product space. Take $\{\mathbf{v}_i \mid i \in I\} \subseteq V$ such that $\mathbf{v}_i \neq \mathbf{0}_V$ for all $i \in I$ and $\langle\!\langle \mathbf{v}_i, \mathbf{v}_j \rangle\!\rangle = 0$ unless $i = j$. Then $\{\mathbf{v}_i \mid i \in I\}$ is a set of linearly independent vectors.*

Proof. Take $\alpha_i \in \mathbb{F}$ $(i \in I)$ such that

$$\sum_{i\in I} \alpha_i \mathbf{v}_i = \mathbf{0}_V.$$

T For $j \in I$

$$\begin{aligned} 0 &= \langle\!\langle \sum_{i\in I} \alpha_i \mathbf{v}_i, \mathbf{v}_j \rangle\!\rangle \\ &= \sum_{i\in I} \alpha_i \langle\!\langle \mathbf{v}_i, \mathbf{v}_j \rangle\!\rangle \\ &= \alpha_j \langle\!\langle \mathbf{v}_j, \mathbf{v}_j \rangle\!\rangle && \text{as } \langle\!\langle \mathbf{v}_i, \mathbf{v}_j \rangle\!\rangle = 0 \text{ unless } i = j. \end{aligned}$$

Since $\mathbf{v}_j \neq \mathbf{0}_V$, $\langle\!\langle \mathbf{v}_j, \mathbf{v}_j \rangle\!\rangle \neq 0$, and so $\alpha_j = 0$. □

Unit vectors, that is, vectors whose length (norm) is 1, play a special rôle, especially when they are mutually orthogonal.

Definition 14.5. Let $\langle\!\langle\ ,\ \rangle\!\rangle$ be an inner product on the vector space V. Then u_i $(i \in I$ are *orthonormal* if and only if $\langle\!\langle \mathbf{u}_i, \mathbf{u}_j \rangle\!\rangle = \delta_{ij}$, where is the *Kronecker delta*, defined by

$$\delta ij := \begin{cases} 1 & \text{if } i = j \\ 0 & \text{otherwise.} \end{cases}$$

The basis $\mathcal{B} = \{ve_i \mid i \in I\}$ is an *orthonormal basis* if and only if the vectors in $\mathcal{B}$ are orthonormal.

Example 14.6. $V := \{f : [0, 2\pi] \to \mathbb{R} \mid f \text{ is continuous}\}$ is a real vector space and

$$\langle\!\langle\ ,\ \rangle\!\rangle : V \times V \longrightarrow \mathbb{R}, \quad (f, g) \longmapsto \frac{1}{\sqrt{\pi}} \int_0^{2\pi} f(t)g(t)dt$$

defines an inner product on V.

For $n \in \mathbb{N} \setminus \{0\}$ define

$$\begin{aligned} c_n &: [0, 2\pi] \longrightarrow \mathbb{R}, \quad x \longmapsto \cos(nx) \\ s_n &: [0, 2\pi] \longrightarrow \mathbb{R}, \quad x \longmapsto \sin(nx) \end{aligned}$$

Then $\{c_n, s_n \mid n = 1, 2, \ldots\}$ is a set of orthonormal vectors in V.

The verification is a routine exercise in integrating trigonometric functions and left to the reader.

Orthonormal bases are particularly convenient for numerous purposes. For example, the coordinates of any vector with respect to an orthonormal basis can be computed directly, using only the inner product.

Theorem 14.7. *Let* $\{\mathbf{e}_i \mid i \in I\}$ *be an orthonormal basis for the inner product space* V*. Take* $\mathbf{v} \in V$*. Then*

$$\mathbf{v} = \sum_{i\in I} \langle\!\langle \mathbf{v}, \mathbf{e}_i \rangle\!\rangle \mathbf{e}_i.$$

Proof. Take $\mathbf{v} \in V$.

Since $\{\mathbf{e}_i \mid i \in I\}$ is a basis for V, there are uniquely determined $\alpha_i \in \mathbb{F}$ $(i \in I)$ with

$$\mathbf{v} = \sum \alpha_i \mathbf{e}_i.$$

Take $j \in I$. Then

$$\begin{aligned}
\langle\!\langle \mathbf{v}, \mathbf{e}_j \rangle\!\rangle &= \langle\!\langle \sum_{i\in I} \alpha_i \mathbf{e}_i, \mathbf{e}_j \rangle\!\rangle \\
&= \sum_{i\in I} \alpha_i \langle\!\langle \mathbf{e}_i, \mathbf{e}_j \rangle\!\rangle \\
&= \alpha_j && \text{as } \langle\!\langle \mathbf{e}_i, \mathbf{e}_j \rangle\!\rangle = \delta_{ij}.
\end{aligned}$$

□

Corollary 14.8. *If $\{\mathbf{e}_i \mid i \in I\}$ is an orthonormal basis for the inner product space V and $\mathbf{v} \in V$,*

$$\|\mathbf{v}\|^2_{\langle\!\langle\ ,\ \rangle\!\rangle} = \sum_{i\in I} |\langle\!\langle \mathbf{v}, \mathbf{e}_i \rangle\!\rangle|^2.$$

In particular, if $\mathbf{v} = \sum_{i\in I} \alpha_i \mathbf{e}_i$, then

$$\|\mathbf{v}\|^2_{\langle\!\langle\ ,\ \rangle\!\rangle} = \sum_{i\in I} |\alpha_i|^2.$$

Proof. Take $\mathbf{v} \in V$.

$$\begin{aligned}
\|\mathbf{v}\|^2_{\langle\!\langle\ ,\ \rangle\!\rangle} &= \langle\!\langle \mathbf{v}, \mathbf{v} \rangle\!\rangle \\
&= \langle\!\langle \sum_{i\in I} \langle\!\langle \mathbf{v}, \mathbf{e}_i \rangle\!\rangle \mathbf{e}_i, \sum_{j\in I} \langle\!\langle \mathbf{v}, \mathbf{e}_j \rangle\!\rangle \mathbf{e}_j \rangle\!\rangle && \text{by Theorem 14.7} \\
&= \sum_{i,j\in I} \langle\!\langle \mathbf{v}, \mathbf{e}_i \rangle\!\rangle \overline{\langle\!\langle \mathbf{v}, \mathbf{e}_j \rangle\!\rangle} \langle\!\langle \mathbf{e}_i, \mathbf{e}_j \rangle\!\rangle \\
&= \sum_{i\in I} \langle\!\langle \mathbf{v}, \mathbf{e}_i \rangle\!\rangle \overline{\langle\!\langle \mathbf{v}, \mathbf{e}_i \rangle\!\rangle} && \text{since } \langle\!\langle \mathbf{e}_i, \mathbf{e}_j \rangle\!\rangle = \delta_{ij} \\
&= \sum_{i\in I} |\langle\!\langle \mathbf{v}, \mathbf{e}_i \rangle\!\rangle|^2.
\end{aligned}$$

□

Example 14.9. Take $\mathbb{R}^n$ with the Eucldean inner product

$$\langle\!\langle\ ,\ \rangle\!\rangle_{\mathbb{R}^n} : \mathbb{R}^n \times \mathbb{R}^n \longrightarrow \mathbb{R}, \quad ((x_1, \dots, x_n), (y_1, \dots, y_n)) \longmapsto \sum_{j=1}^{n} x_j y_j$$

As the reader no dout knows from Cartesian geomtery, that the *standard basis vectors*

$$\mathbf{e}_n = (0, \dots, 0, 1)$$

form an orthonormal basis for $\mathbb{R}^n$,

$$\langle\!\langle (x_1, \dots, x_n), \mathbf{e}_j \rangle\!\rangle = x_j$$

and

$$\|(x_1, \dots, x_n)\|^2 = \sum_{j=1}^{n} x_j^2$$

This example shows that the above results generalise familiar ones from Cartesian geometry.

Since orthonormal bases are so useful and important, it is particularly satisfying that they can always be constructed. Given any basis whatsoever for an inner product space, there is an algorithm for constructing an orthonormal basis from it.

Theorem 14.10 (Gram-Schmidt Orthonormalisation). *Every finitely generated inner product space admits an orthonormal basis.*[1]

Proof. Let $\langle\!\langle\ ,\ \rangle\!\rangle$ be a inner product on the finitely generated vector space V over $\mathbb{F}$. Given a basis $\{\mathbf{u}_1, \ldots, \mathbf{u}_m\}$ of V, we construct an orthonormal basis $\{\mathbf{e}_1, \ldots, \mathbf{e}_m\}$ by means of a recursive procedure (algorithm), the *Gram-Schmidt procedure.*

Since $\{\mathbf{u}_1, \ldots, \mathbf{u}_m\}$ is a basis, $\mathbf{u}_1 \neq \mathbf{0}_V$. Put

$$\mathbf{e}_1 := \frac{\mathbf{u}_1}{\|\mathbf{u}_1\|}.$$

Suppose that orthonormal vectors $\mathbf{e}_1, \ldots \mathbf{e}_j$ have been constructed for $1 \leq j < m$ such that $\langle \mathbf{e}_1, \ldots, \mathbf{e}_j \rangle = \langle \mathbf{u}_1, \ldots, \mathbf{u}_j \rangle$.
Put

$$\mathbf{v}_{j+1} := \mathbf{u}_{j+1} - \sum_{i=1}^{j} \langle\!\langle \mathbf{u}_{j+1}, \mathbf{e}_i \rangle\!\rangle \mathbf{e}_i.$$

For each $k \leq j$,

$$\begin{aligned}\langle\!\langle \mathbf{v}_{j+1}, \mathbf{e}_k \rangle\!\rangle &= \langle\!\langle \mathbf{u}_{j+1} - \sum_{i=1}^{j} \langle\!\langle \mathbf{u}_{j+1}, \mathbf{e}_i \rangle\!\rangle \mathbf{e}_i, \mathbf{e}_k \rangle\!\rangle \\ &= \langle\!\langle \mathbf{u}_{j+1}, \mathbf{e}_k \rangle\!\rangle - \sum_{i=1}^{j} \langle\!\langle \mathbf{u}_{j+1}, \mathbf{e}_i \rangle\!\rangle \langle\!\langle \mathbf{e}_i, \mathbf{e}_k \rangle\!\rangle \\ &= \langle\!\langle \mathbf{u}_{j+1}, \mathbf{e}_k \rangle\!\rangle - \langle\!\langle \mathbf{u}_{j+1}, \mathbf{e}_k \rangle\!\rangle = 0\end{aligned}$$

Hence the vectors $\mathbf{e}_1, \ldots, \mathbf{e}_j, \mathbf{v}_{j+1}$ are mutually orthogonal.

Moreover, $\mathbf{v}_{j+1} \neq \mathbf{0}_V$. For otherwise, by Equation **??**, $\mathbf{u}_{j+1} = \sum_{i=1}^{j} \langle\!\langle \mathbf{u}_{j+1}, \mathbf{e}_{ii} \rangle\!\rangle \mathbf{e}_i$, contradicting the linear independence of $\mathbf{u}_1, \ldots, \mathbf{u}_{j+1}$.
We may therefore put

$$\mathbf{e}_{j+1} := \frac{1}{\|\mathbf{v}_{j+1}\|} \mathbf{v}_{j+1}.$$

This clearly renders $\mathbf{e}_1, \ldots \mathbf{e}_j$ orthonormal, and hence, by Theorem (14.4), linearly independent. Thus, by Theorem (8.6), $\langle \mathbf{e}_1, \ldots, \mathbf{e}_{j+1} \rangle = \langle \mathbf{u}_1, \ldots \mathbf{u}_{j+1} \rangle$. In particular, $\{\mathbf{e}_1, \ldots \mathbf{e}_m\}$ is an orthonormal basis for V. □

Example 14.11. Consider $\mathbb{R}^2$ with the inner product

$$\langle\!\langle\ ,\ \rangle\!\rangle \colon \mathbb{R}^2 \times \mathbb{R}^2 \longrightarrow \mathbb{R}, \quad \big((x, y), (u, v)\big) \longmapsto xu + 2xv + 2yu + 5yv.$$

and basis $\mathcal{B} = \{(1, 1), (0, 1)\}$.

[1] There is an extension of this to inner product spaces which are not finitely generated. You will meet such matters in measure and integration theory, and in functional analysis, for example. We do not pursue such matters further here.

We construct an orthonormal basis by applying the Gram-Schmid procedure to $\mathcal{B}$.

Before doing so, we verify that $\mathcal{B}$ is a basis for $\mathbb{R}^2$ and that $\langle\!\langle\ ,\ \rangle\!\rangle$ is an inner product.

Since $\mathcal{B}$ comprises two vectors and $\dim_{\mathbb{R}}(\mathbb{R}^2) = 2$, it follows from Theorem 8.6 that $\mathcal{B}$ is a basis for $\mathbb{R}^2$ if and only if $(1,1)$ and $(0,1)$ are linearly independent.

Choosing a basis for $\mathbb{R}^2$ is precisely choosing an isomorphism $\mathbb{R}^2 \to \mathbb{R}_{(2)}$, and so $(1,1)$ and $(0,1)$ are linearly independent if and only if $\mathbf{c}_1$ and $\mathbf{c}_2$, their co-ordinate vectors with respect to the chosen basis, are linearly independent.

This is true if and only if the matrix $\underline{\mathbf{A}}$, whose columns are $\mathbf{c}_1$ and $\mathbf{c}_2$, has rank 2, that is $\det(\underline{\mathbf{A}}) \neq 0$.

Since this does not depend on the choice of basis, we choose the standard basis for $\mathbb{R}^2$, $\{(1,0),(0,1)\}$.

The co-ordinate vectors are then $\begin{bmatrix}1\\1\end{bmatrix}$ and $\begin{bmatrix}0\\1\end{bmatrix}$ respectively, and $\det(\underline{\mathbf{A}}) = \det\left(\begin{bmatrix}1 & 0\\1 & 1\end{bmatrix}\right) = 1 \neq 0$.

We turn to verifying that $\langle\!\langle\ ,\ \rangle\!\rangle$ is a real inner product.

$$\langle\!\langle (x,y),(x,y)\rangle\!\rangle = x^2 + 4xy + 5y^2 = (x+2y)^2 + y^2 \geq 0 \quad (x,y) \in \mathbb{R}^2.$$

Moreover, $\langle\!\langle (x,y),(x,y)\rangle\!\rangle = 0$ if and only if $(x+2y)^2 + y^2 = 0$ if and only if $x = -2y$ and $y = 0$, that is, $(x,y) = (0,0)$.

$$\langle\!\langle (u,v),(x,y)\rangle\!\rangle = ux + 2uy + 2vx + 5vy = xu + 2xv + 2yu + 5yv = \langle\!\langle (x,y),(u,v)\rangle\!\rangle.$$

$$\begin{aligned}\langle\!\langle \alpha(x,y) + \beta(r,s),(u,v)\rangle\!\rangle &= \langle\!\langle (\alpha x + \beta r, \alpha y + \beta s),(u,v)\rangle\!\rangle\\ &= (\alpha x + \beta r)u + 2\alpha x + \beta r)v + 2(\alpha y + \beta s)u + 5(\alpha y + \beta s)v\\ &= \alpha(xu + 2xv + 2yu + 5yv) + \beta(ru + 2rv + 2su + 5sv)\\ &= \alpha\langle\!\langle (x,y),(u,v)\rangle\!\rangle + \beta\langle\!\langle (r,s),(u.v)\rangle\!\rangle\end{aligned}$$

We now apply the Gram-Schmid procedure.

As $\langle\!\langle (1,1),(1,1)\rangle\!\rangle = 10$, we put

$$\mathbf{e}_1 := \frac{1}{\sqrt{10}}(1,1).$$

As $\langle\!\langle \mathbf{e}_1,(0,1)\rangle\!\rangle = \frac{1}{\sqrt{10}}(2+5) = \frac{7}{\sqrt{10}}$, we put

$$\mathbf{e}_2^* := (0,1) - \langle\!\langle (0,1),\mathbf{e}_1\rangle\!\rangle\mathbf{e}_1 = (0,1) - \frac{7}{10}(1,1) = \frac{1}{10}(-7,3)$$

As $\langle\!\langle -7,3),(-7,3)\rangle\!\rangle = (-7+6)^2 + 3^2 = 10$ we put

$$\mathbf{e}_2 := \frac{1}{\sqrt{10}}(-7,3)$$

Our orthonormal basis is thus

$$\left\{\frac{1}{\sqrt{10}}(1,1), \frac{1}{\sqrt{10}}(-7,3)\right\}$$

Comment: We went to the trouble of verifying that $\mathcal{B}$ is a basis as illustration of how the theory developed earlier makes arguments simpler and avoids tedious, repetitive computations.

14.1 Orthogonal Complements

Definition 14.12. Let $\langle\!\langle\ ,\ \rangle\!\rangle$ be an inner product on the vector space V. The *orthogonal complement*, $S^\perp$, of the subset, S, of V, is the set of all vectors in V, that are orthogonal to every vector in S.

$$S^\perp := \{\mathbf{v} \in V \mid \langle\!\langle \mathbf{v}, \mathbf{x} \rangle\!\rangle = 0 \text{ for all } \mathbf{x} \in S\}$$

Theorem 14.13. *Let S be a subset of the inner product space $(V, \langle\!\langle\ ,\ \rangle\!\rangle)$. Then*

(i) $S^\perp$ *is a vector subspace of* V.

(ii) If $S \subseteq T$, *then* $T^\perp \subseteq S^\perp$

(iii) $S^\perp = \langle S \rangle^\perp$

(iv) $\langle S \rangle \leq (S^\perp)^\perp$

If, in addition, V is finitely generated,

(v) $V = \langle S \rangle \oplus \langle S \rangle^\perp$

(vi) $(S^\perp)^\perp = \langle S \rangle$.

Proof. (i) Take $\mathbf{u}, \mathbf{v} \in S^\perp, \alpha, \beta \in \mathbb{F}$ and $\mathbf{x} \in S$. Then

$$\langle\!\langle \alpha\mathbf{u} + \beta\mathbf{v}, \mathbf{x} \rangle\!\rangle = \alpha\langle\!\langle \mathbf{u}, \mathbf{x} \rangle\!\rangle + \beta\langle\!\langle \mathbf{v}, \mathbf{x} \rangle\!\rangle = 0,$$

whence $\alpha\mathbf{u} + \beta\mathbf{v} \in S^\perp$.

(ii) Take $\mathbf{v} \in T^\perp$ and $\mathbf{x} \in S$.

Since $S \subseteq T$, $\mathbf{x} \in T$, and so $\langle\!\langle \mathbf{v}, \mathbf{x} \rangle\!\rangle = 0$, whence $\mathbf{v} \in S^\perp$.

(iii) Since $S \subseteq \langle S \rangle$, it follows from (ii) that $\langle S \rangle^\perp \subseteq S^\perp$.

For the reverse inclusion, take $\mathbf{v} \in S^\perp$ and $\mathbf{x} \in \langle S \rangle$.

Then $\mathbf{x} = \alpha_1\mathbf{x}_1 + \cdots + \alpha_k\mathbf{x}_k$ for some $\alpha_1, \ldots, \alpha_k \in \mathbb{F}$ and $\mathbf{x}_1, \ldots, \mathbf{x}_k \in S$. Thus

$$\begin{aligned}
\langle\!\langle \mathbf{x}, \mathbf{v} \rangle\!\rangle &= \langle\!\langle \sum_{j=1}^{k} \alpha_j\mathbf{x}_j, \mathbf{v} \rangle\!\rangle \\
&= \sum_{j=1}^{k} \alpha_j \langle\!\langle \mathbf{x}_j, \mathbf{v} \rangle\!\rangle \\
&= \sum_{j=1}^{k} \alpha_j 0 \\
&= 0,
\end{aligned}$$

(iv) Take $\mathbf{x} \in \langle S \rangle$ and $\mathbf{v} \in S^\perp = \langle S \rangle^\perp$. Then $\langle\!\langle \mathbf{x}, \mathbf{v} \rangle\!\rangle = \langle\!\langle \mathbf{v}, \mathbf{x} \rangle\!\rangle = 0$, whence $\mathbf{x} \in (S^\perp)^\perp$.

Suppose that V is finitely generated.

(v) For any subspace W of V, if $\mathbf{v} \in W \cap W^\perp$, then $\langle\!\langle \mathbf{v}, \mathbf{v} \rangle\!\rangle = 0$, whence $\mathbf{v} = \mathbf{0}_V$.

Hence $W \cap W^\perp = \{\mathbf{0}_V\}$.

It is therefore sufficient to show that $V = W + W^\perp$ and then take $W := \langle S \rangle$.

Let $\{\mathbf{e}, \dots, \mathbf{e}_k\}$ be an orthonormal basis for W and take $\mathbf{v} \in V$.

Put $\mathbf{x} := \langle\!\langle \mathbf{v}, \mathbf{e}_1 \rangle\!\rangle \mathbf{e}_1 + \cdots + \langle\!\langle \mathbf{v}, \mathbf{e}_k \rangle\!\rangle \mathbf{e}_k$ and $\mathbf{y} := \mathbf{v} - \mathbf{x}$.

Clearly, $\mathbf{x} \in W$ and $\mathbf{v} = \mathbf{x} + \mathbf{y}$.

It remains only to show that $\mathbf{y} \in W^{\perp}$.

$$
\begin{aligned}
\langle\!\langle \mathbf{y}, \mathbf{e}_i \rangle\!\rangle &= \langle\!\langle \mathbf{v}, \mathbf{e}_i \rangle\!\rangle - \langle\!\langle \mathbf{x}, \mathbf{e}_i \rangle\!\rangle \\
&= \langle\!\langle \mathbf{v}, \mathbf{e}_i \rangle\!\rangle - \langle\!\langle \sum_{j=1}^{k} \langle\!\langle \mathbf{v}, \mathbf{e}_j \rangle\!\rangle \mathbf{e}_j, \mathbf{e}_i \rangle\!\rangle \\
&= \langle\!\langle \mathbf{v}, \mathbf{e}_i \rangle\!\rangle - \sum_{j=1}^{k} \langle\!\langle \mathbf{v}, \mathbf{e}_j \rangle\!\rangle \langle\!\langle \mathbf{e}_j, \mathbf{e}_i \rangle\!\rangle && \text{by bi-linearity} \\
&= \langle\!\langle \mathbf{v}, \mathbf{e}_i \rangle\!\rangle - \langle\!\langle \mathbf{v}, \mathbf{e}_i \rangle\!\rangle = 0 && \text{by orthonormality}
\end{aligned}
$$

Thus, $\mathbf{y} \in W^{\perp}$.

(vi) Using (v) twice,

$$V = W \oplus W^{\perp} = W^{\perp} \oplus (W^{\perp})^{\perp}.$$

Since V is finitely generated, $W \cong (W^{\perp})^{\perp}$.

By (iv), $W \leq (W^{\perp})^{\perp}$.

Thus, $W = (W^{\perp})^{\perp}$ since V (and therefore also $(W^{\perp})^{\perp}$) is finitely generated. □

We show by example that (iv) and (v) need not hold if the vector space in question is not finitely generated.

Example 14.14. Take $V := \mathbb{R}[t]$, the real vector space of all real polynomials in one indeterminate. This is not finitely generated, since $\{t^n \mid n \in \mathbb{N}\}$ is an infinite set of linearly independent vectors. We use the inner product

$$\langle\!\langle\ ,\ \rangle\!\rangle \colon \mathbb{R}[t] \times \mathbb{R}[t] \longrightarrow \mathbb{R}, \quad (p, q) \longmapsto \int_0^1 p(x)q(x)\, dx$$

As subspace we take

$$W := \{p \mid p(0) = 0\}$$

Take any $h \in W^{\perp}$. Then

$$
\begin{aligned}
\|th\|^2 &= \int_0^1 \big(xh(x)\big)^2\, dx \\
&= \int_0^1 h(x)x^2h(x)\, dx \\
&= \langle\!\langle h, t^2h \rangle\!\rangle \\
&= 0
\end{aligned}
$$

since $t^2h \in W$ and $h \in W^{\perp}$.

Thus th is the 0 polynomial, whence h must be the zero polynomial.

Consequently $W^{\perp} = \{\mathbf{0}_V\}$.

It follows that $(W^{\perp})^{\perp} = V \neq W$ and also that $W + W^{\perp} = W \neq V$.

14.2 Orthogonal Transformations

When we studied vector spaces without considering any additional structure, linear transformations provide the means for comparing them, since linear transformations are precisely those functions between vector spaces which respect the vector space operations.

We have now specialised to subfields $\mathbb{F}$ of $\mathbb{C}$ in order to be able to introduce the notion of an inner product, which, as we have already seen, allows us to speak of angles and distances, thereby allowing us to "do" geometry.

It therefore behooves us to restrict ourselves to those linear transformations between inner product spaces, that respect the additional structure.

Definition 14.15. Let $(V, \langle\!\langle\ ,\ \rangle\!\rangle_V)$ and $(W, \langle\!\langle\ ,\ \rangle\!\rangle_W)$ be inner product spaces over $\mathbb{F}$.

The linear transformation $T : V \to W$ *preserves the inner product* if and only if for all $\mathbf{u}, \mathbf{v} \in V$

$$\langle\!\langle T(\mathbf{u}), T(\mathbf{v})\rangle\!\rangle_W = \langle\!\langle \mathbf{u}, \mathbf{v}\rangle\!\rangle_V.$$

Theorem 14.16. *Let $T\colon V \to W$ be a linear transformation between finite dimensional inner product spaces $(V, \langle\!\langle\ ,\ \rangle\!\rangle_V), (W, \langle\!\langle\ ,\ \rangle\!\rangle_W)$. Then the following are equivalent.*

(a) T preserves the inner product.

(b) $\|T(\mathbf{u})\|_W = \|\mathbf{u}\|_V$ for all $\mathbf{u} \in V$.

(c) *If $\{\mathbf{e}_1, \ldots, \mathbf{e}_n\}$ is an orthonormal basis for V, then $\{T(\mathbf{e}_1), \ldots, T(\mathbf{e}_n)\}$ is an orthonormal basis for im(T).*

Proof. (a) $\Rightarrow$ (b): Take $\mathbf{u} \in V$. Then

$$\|T(\mathbf{u})\|_W^2 = \langle\!\langle T(\mathbf{u}), T(\mathbf{u})\rangle\!\rangle_W = \langle\!\langle \mathbf{u}, \mathbf{u}\rangle\!\rangle_V = \|\mathbf{u}\|_V^2.$$

(b) $\Rightarrow$ (a): Take $\mathbf{u}, \mathbf{v} \in V$. Then

$$\begin{aligned}\|\mathbf{u}\|_V^2 + \langle\!\langle T(\mathbf{u}), T(\mathbf{v})\rangle\!\rangle_W + \langle\!\langle T(\mathbf{v}), T(\mathbf{u})\rangle\!\rangle_W + \|\mathbf{v}\|_V^2 &= \|T(\mathbf{u})\|_W^2 + \langle\!\langle T(\mathbf{u}), T(\mathbf{v})\rangle\!\rangle_W + \langle\!\langle T(\mathbf{v}), T(\mathbf{u})\rangle\!\rangle_W + \|T(\mathbf{v})\|^2 \\ &= \|T(\mathbf{u}) + T(\mathbf{v})\|_W^2 \\ &= \|T(\mathbf{u} + \mathbf{v})\|_W^2 \\ &= \|\mathbf{u} + \mathbf{v}\|_V^2 \\ &= \|\mathbf{u}\|_V^2 + \langle\!\langle \mathbf{u}, \mathbf{v}\rangle\!\rangle_V + \langle\!\langle \mathbf{v}, \mathbf{u}\rangle\!\rangle + \|\mathbf{v}\|_V^2,\end{aligned}$$

Thus, the real parts of $\langle\!\langle T(\mathbf{u}), T(\mathbf{v})\rangle\!\rangle_W$ and $\langle\!\langle \mathbf{u}, \mathbf{v}\rangle\!\rangle_V$ agree.

$$\begin{aligned}\|\mathbf{u}\|_V^2 - i\langle\!\langle T(\mathbf{u}), T(\mathbf{v})\rangle\!\rangle_W + iIPT(\mathbf{v})T(\mathbf{u})_W + \|\mathbf{v}\|_V^2 &= \|T(\mathbf{u})\|_W^2 - i\langle\!\langle T(\mathbf{u}), T(\mathbf{v})\rangle\!\rangle_W + i\langle\!\langle T(\mathbf{v}), T(\mathbf{u})\rangle\!\rangle_W + \|T(\mathbf{v}) \\ &= \|T(\mathbf{u}) + iT(\mathbf{v})\|_W^2 \\ &= \|T(\mathbf{u} + i\mathbf{v})\|_W^2 \\ &= \|\mathbf{u} + ivv\|_V^2 \\ &= \|\mathbf{u}\|_V^2 - i\langle\!\langle \mathbf{u}, \mathbf{v}\rangle\!\rangle_V + i\langle\!\langle \mathbf{v}, \mathbf{u}\rangle\!\rangle + \|\mathbf{v}\|_V^2.\end{aligned}$$

Thus, the imaginary parts of $\langle\!\langle T(\mathbf{u}), T(\mathbf{v})\rangle\!\rangle_W$ and $\langle\!\langle \mathbf{u}, \mathbf{v}\rangle\!\rangle_V$ also agree.

Hence $\langle\!\langle T(\mathbf{u}), T(\mathbf{v})\rangle\!\rangle_W = \langle\!\langle \mathbf{u}, \mathbf{v}\rangle\!\rangle_V$.

(a) $\Rightarrow$ (c): Let $\{\mathbf{e}_1, \ldots, \mathbf{e}_n\}$ be an orthonormal basis for V.

$$\langle\!\langle T(\mathbf{e}_i), T(\mathbf{e}_j)\rangle\!\rangle_W = \langle\!\langle \mathbf{e}_i, \mathbf{e}_j\rangle\!\rangle_V = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{if } i \neq j \end{cases}$$

Thus, $\{T(\mathbf{e}_1), \dots, T(\mathbf{e}_n)\}$ is orthonormal

(c) $\Rightarrow$ (b) Let $\{\mathbf{e}_1, \dots, \mathbf{e}_n\}$ be an orthonormal basis for V and take $\mathbf{v} \in V$.
Then $\mathbf{v} = \langle\!\langle \mathbf{v}, \mathbf{e}_1\rangle\!\rangle_V \mathbf{e}_1 + \cdots + \langle\!\langle \mathbf{v}, \mathbf{e}_n\rangle\!\rangle_V \mathbf{e}_n$, and so $T(\mathbf{v}) = \langle\!\langle \mathbf{v}, \mathbf{e}_1\rangle\!\rangle_V T(\mathbf{e}_1) + \cdots + \langle\!\langle \mathbf{v}, \mathbf{e}_n\rangle\!\rangle_V T(\mathbf{e}_n)$.
Since $T(\mathbf{e}_1), \dots, T(\mathbf{e}_n)$ is an orthonormal basis for V, Theorem 14.7 that for all i

$$\langle\!\langle T(\mathbf{v}), T(\mathbf{e}_i)\rangle\!\rangle_W = \langle\!\langle \mathbf{v}, \mathbf{e}_i\rangle\!\rangle_V.$$

By Corollary 14.8,

$$T(\mathbf{v}) = \langle\!\langle T(\mathbf{v}), T(\mathbf{e}_1)\rangle\!\rangle_V T(\mathbf{e}_1) + \cdots + \langle\!\langle T(\mathbf{v}), T(\mathbf{e}_n)\rangle\!\rangle_V T(\mathbf{e}_n).$$

Thus, $\|T(\mathbf{v})\|_W^2 = \|\mathbf{v}\|_V^2$. □

Corollary 14.17. *If the linear tranformations, $T : V \longrightarrow W$, between the inner product spaces $(V, \langle\!\langle\ ,\ \rangle\!\rangle_V), (W, \langle\!\langle\ ,\ \rangle\!\rangle_W)$ preserves the inner product, then T is injective.*

Proof. $T(\mathbf{v}) = \mathbf{0}_W$ if and only if $\|T(\mathbf{v})\|_W = 0$ if and only if $\|\mathbf{v}\|_V = 0$ if and only if $=\mathbf{0}_V$. □

The most important case, particularly from the point of view of applications, is when $W = V$ and $\langle\!\langle\ ,\ \rangle\!\rangle_W = \langle\!\langle\ ,\ \rangle\!\rangle_V$.

Definition 14.18. Let $(V, \langle\!\langle\ ,\ \rangle\!\rangle)$ be an inner product space. Then the linear transformation $T : V \longrightarrow V$ is said to be an *orthogonal transformation* with respect to $\langle\!\langle\ ,\ \rangle\!\rangle$ if and only if for all $\mathbf{u}, \mathbf{v} \in V$

$$\langle\!\langle T(\mathbf{u}), T(\mathbf{v})\rangle\!\rangle = \langle\!\langle \mathbf{u}, \mathbf{v}\rangle\!\rangle.$$

Observation 14.19. Traditionally, orthogonal endomorphisms of a complex inner product space are called *unitary*.

Lemma 14.20. *Let $(V, \langle\!\langle\ ,\ \rangle\!\rangle)$ be an inner product space. Then each orthogonal transformation $T : V \longrightarrow V$ is an isomorphism.*

Proof. If $\{\mathbf{e}_1, \dots, \mathbf{e}_n\}$ is an orthonormal basis for V, then $T(\mathbf{e}_1), \dots, T(\mathbf{e}_n)$ are orthonormal, hence linearly independent, and hence form a basis. □

14.3 Exercises

Exercise 14.1. For each of the following symmetric matrices, find an orthogonal matrix which diagonalises it.

(a) $\begin{bmatrix} 1 & 2 \\ 2 & 5 \end{bmatrix}$

(b) $\begin{bmatrix} 0 & -1 \\ -1 & 0 \end{bmatrix}$

(c) $\begin{bmatrix} 1 & -1 \\ -1 & 0 \end{bmatrix}$

(d) $\begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix}$

(e) $\begin{bmatrix} 1 & -1 \\ -1 & 2 \end{bmatrix}$

(f) $\begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix}$