be exactly one such α. These intuitive ideas are correct for finite-dimensional subspaces and for some, but not all, infinite-dimensional subspaces. Since the precise situation is too complicated to treat here, we shall prove only the following result.

Theorem 4. *Let* W *be a subspace of an inner product space* V *and let* β *be a vector in* V.

(i) *The vector* α *in* W *is a best approximation to* β *by vectors in* W *if and only if* $\beta - \alpha$ *is orthogonal to every vector in* W.

(ii) *If a best approximation to* β *by vectors in* W *exists, it is unique.*

(iii) *If* W *is finite-dimensional and* $\{\alpha_1, \ldots, \alpha_n\}$ *is any orthonormal basis for* W, *then the vector*

$$\alpha = \sum_k \frac{(\beta|\alpha_k)}{||\alpha_k||^2}\alpha_k$$

is the (unique) best approximation to β *by vectors in* W.

Proof. First note that if γ is any vector in V, then $\beta - \gamma = (\beta - \alpha) + (\alpha - \gamma)$, and

$$||\beta - \gamma||^2 = ||\beta - \alpha||^2 + 2\,\mathrm{Re}\,(\beta - \alpha|\alpha - \gamma) + ||\alpha - \gamma||^2.$$

Now suppose $\beta - \alpha$ is orthogonal to every vector in W, that γ is in W and that $\gamma \neq \alpha$. Then, since $\alpha - \gamma$ is in W, it follows that

$$\begin{aligned} ||\beta - \gamma||^2 &= ||\beta - \alpha||^2 + ||\alpha - \gamma||^2 \\ &> ||\beta - \alpha||^2. \end{aligned}$$

Conversely, suppose that $||\beta - \gamma|| \geq ||\beta - \alpha||$ for every γ in W. Then from the first equation above it follows that

$$2\,\mathrm{Re}\,(\beta - \alpha|\alpha - \gamma) + ||\alpha - \gamma||^2 \geq 0$$

for all γ in W. Since every vector in W may be expressed in the form $\alpha - \gamma$ with γ in W, we see that

$$2\,\mathrm{Re}\,(\beta - \alpha|\tau) + ||\tau||^2 \geq 0$$

for every τ in W. In particular, if γ is in W and $\gamma \neq \alpha$, we may take

$$\tau = -\frac{(\beta - \alpha|\alpha - \gamma)}{||\alpha - \gamma||^2}(\alpha - \gamma).$$

Then the inequality reduces to the statement

$$-2\frac{|(\beta - \alpha|\alpha - \gamma)|^2}{||\alpha - \gamma||^2} + \frac{|(\beta - \alpha|\alpha - \gamma)|^2}{||\alpha - \gamma||^2} \geq 0.$$

This holds if and only if $(\beta - \alpha|\alpha - \gamma) = 0$. Therefore, $\beta - \alpha$ is orthogonal to every vector in W. This completes the proof of the equivalence of the two conditions on α given in (i). The orthogonality condition is evidently satisfied by at most one vector in W, which proves (ii).

Now suppose that W is a finite-dimensional subspace of V. Then we know, as a corollary of Theorem 3, that W has an orthogonal basis. Let $\{\alpha_1, \ldots, \alpha_n\}$ be any orthogonal basis for W and define α by (8-11). Then, by the computation in the proof of Theorem 3, $\beta - \alpha$ is orthogonal to each of the vectors α_k ($\beta - \alpha$ is the vector obtained at the last stage when the orthogonalization process is applied to $\alpha_1, \ldots, \alpha_n, \beta$). Thus $\beta - \alpha$ is orthogonal to every linear combination of $\alpha_1, \ldots, \alpha_n$, i.e., to every vector in W. If γ is in W and $\gamma \neq \alpha$, it follows that $||\beta - \gamma|| > ||\beta - \alpha||$. Therefore, α is the best approximation to β that lies in W. ∎

Definition. *Let* V *be an inner product space and* S *any set of vectors in* V. *The* **orthogonal complement** *of* S *is the set* $S^\perp$ *of all vectors in* V *which are orthogonal to every vector in* S.

The orthogonal complement of V is the zero subspace, and conversely $\{0\}^\perp = V$. If S is any subset of V, its orthogonal complement $S^\perp$ (S perp) is always a subspace of V. For S is non-empty, since it contains 0; and whenever α and β are in $S^\perp$ and c is any scalar,

$$\begin{aligned}(c\alpha + \beta|\gamma) &= c(\alpha|\gamma) + (\beta|\gamma)\\ &= c0 + 0\\ &= 0\end{aligned}$$

for every γ in S, thus $c\alpha + \beta$ also lies in S. In Theorem 4 the characteristic property of the vector α is that it is the only vector in W such that $\beta - \alpha$ belongs to $W^\perp$.

Definition. *Whenever the vector* α *in Theorem 4 exists it is called the* **orthogonal projection of** β **on** W. *If every vector in* V *has an orthogonal projection on* W, *the mapping that assigns to each vector in* V *its orthogonal projection on* W *is called the* **orthogonal projection of** V **on** W.

By Theorem 4, the orthogonal projection of an inner product space on a finite-dimensional subspace always exists. But Theorem 4 also implies the following result.

Corollary. *Let* V *be an inner product space,* W *a finite-dimensional subspace, and* E *the orthogonal projection of* V *on* W. *Then the mapping*

$$\beta \rightarrow \beta - E\beta$$

is the orthogonal projection of V *on* $W^\perp$.

Proof. Let β be an arbitrary vector in V. Then $\beta - E\beta$ is in $W^\perp$, and for any γ in $W^\perp$, $\beta - \gamma = E\beta + (\beta - E\beta - \gamma)$. Since $E\beta$ is in W and $\beta - E\beta - \gamma$ is in $W^\perp$, it follows that

$$\begin{aligned}||\beta - \gamma||^2 &= ||E\beta||^2 + ||\beta - E\beta - \gamma||^2 \\ &\geq ||\beta - (\beta - E\beta)||^2\end{aligned}$$

with strict inequality when $\gamma \neq \beta - E\beta$. Therefore, $\beta - E\beta$ is the best approximation to β by vectors in $W^\perp$. ∎

EXAMPLE 14. Give R^3 the standard inner product. Then the orthogonal projection of $(-10, 2, 8)$ on the subspace W that is spanned by $(3, 12, -1)$ is the vector

$$\begin{aligned}\alpha &= \frac{((-10, 2, 8)|(3, 12, -1))}{9 + 144 + 1}(3, 12, -1) \\ &= \frac{-14}{154}(3, 12, -1).\end{aligned}$$

The orthogonal projection of R^3 on W is the linear transformation E defined by

$$(x_1, x_2, x_3) \rightarrow \left(\frac{3x_1 + 12x_2 - x_3}{154}\right)(3, 12, -1).$$

The rank of E is clearly 1; hence its nullity is 2. On the other hand,

$$E(x_1, x_2, x_3) = (0, 0, 0)$$

if and only if $3x_1 + 12x_2 - x_3 = 0$. This is the case if and only if (x_1, x_2, x_3) is in $W^\perp$. Therefore, $W^\perp$ is the null space of E, and $\dim (W^\perp) = 2$. Computing

$$(x_1, x_2, x_3) - \left(\frac{3x_1 + 12x_2 - x_3}{154}\right)(3, 12, -1)$$

we see that the orthogonal projection of R^3 on $W^\perp$ is the linear transformation $I - E$ that maps the vector (x_1, x_2, x_3) onto the vector

$$\frac{1}{154}(145x_1 - 36x_2 + 3x_3, -36x_1 + 10x_2 + 12x_3, 3x_1 + 12x_2 + 153x_3).$$

The observations made in Example 14 generalize in the following fashion.

Theorem 5. *Let* W *be a finite-dimensional subspace of an inner product space* V *and let* E *be the orthogonal projection of* V *on* W. *Then* E *is an idempotent linear transformation of* V *onto* W, $W^\perp$ *is the null space of* E, *and*

$$V = W \oplus W^\perp.$$

Proof. Let β be an arbitrary vector in V. Then $E\beta$ is the best approximation to β that lies in W. In particular, $E\beta = \beta$ when β is in W. Therefore, $E(E\beta) = E\beta$ for every β in V; that is, E is idempotent: $E^2 = E$. To prove that E is a linear transformation, let α and β be any vectors in

V and c an arbitrary scalar. Then, by Theorem 4, $\alpha - E\alpha$ and $\beta - E\beta$ are each orthogonal to every vector in W. Hence the vector

$$c(\alpha - E\alpha) + (\beta - E\beta) = (c\alpha + \beta) - (cE\alpha + E\beta)$$

also belongs to $W^\perp$. Since $cE\alpha + E\beta$ is a vector in W, it follows from Theorem 4 that

$$E(c\alpha + \beta) = cE\alpha + E\beta.$$

Of course, one may also prove the linearity of E by using (8-11). Again let β be any vector in V. Then $E\beta$ is the unique vector in W such that $\beta - E\beta$ is in $W^\perp$. Thus $E\beta = 0$ when β is in $W^\perp$. Conversely, β is in $W^\perp$ when $E\beta = 0$. Thus $W^\perp$ is the null space of E. The equation

$$\beta = E\beta + \beta - E\beta$$

shows that $V = W + W^\perp$; moreover, $W \cap W^\perp = \{0\}$. For if α is a vector in $W \cap W^\perp$, then $(\alpha|\alpha) = 0$. Therefore, $\alpha = 0$, and V is the direct sum of W and $W^\perp$. ▌

Corollary. *Under the conditions of the theorem,* I − E *is the orthogonal projection of* V *on* $W^\perp$. *It is an idempotent linear transformation of* V *onto* $W^\perp$ *with null space* W.

Proof. We have already seen that the mapping $\beta \rightarrow \beta - E\beta$ is the orthogonal projection of V on $W^\perp$. Since E is a linear transformation, this projection on $W^\perp$ is the linear transformation $I - E$. From its geometric properties one sees that $I - E$ is an idempotent transformation of V onto W. This also follows from the computation

$$\begin{aligned}(I - E)(I - E) &= I - E - E + E^2 \\ &= I - E.\end{aligned}$$

Moreover, $(I - E)\beta = 0$ if and only if $\beta = E\beta$, and this is the case if and only if β is in W. Therefore W is the null space of $I - E$. ▌

The Gram-Schmidt process may now be described geometrically in the following way. Given an inner product space V and vectors $\beta_1, \ldots, \beta_n$ in V, let P_k $(k > 1)$ be the orthogonal projection of V on the orthogonal complement of the subspace spanned by $\beta_1, \ldots, \beta_{k-1}$, and set $P_1 = I$. Then the vectors one obtains by applying the orthogonalization process to $\beta_1, \ldots, \beta_n$ are defined by the equations

$$\text{(8-12)} \qquad \alpha_k = P_k\beta_k, \qquad 1 \le k \le n.$$

Theorem 5 implies another result known as **Bessel's inequality**.

Corollary. *Let* $\{\alpha_1, \ldots, \alpha_n\}$ *be an orthogonal set of non-zero vectors in an inner product space* V. *If* β *is any vector in* V, *then*

$$\sum_k \frac{|(\beta|\alpha_k)|^2}{||\alpha_k||^2} \le ||\beta||^2$$

and equality holds if and only if

$$\beta = \sum_k \frac{(\beta|\alpha_k)}{||\alpha_k||^2} \alpha_k.$$

Proof. Let $\gamma = \sum_k [(\beta|\alpha_k)/||\alpha_k||^2]\, \alpha_k$. Then $\beta = \gamma + \delta$ where $(\gamma|\delta) = 0$. Hence

$$||\beta||^2 = ||\gamma||^2 + ||\delta||^2.$$

It now suffices to prove that

$$||\gamma||^2 = \sum_k \frac{|(\beta|\alpha_k)|^2}{||\alpha_k||^2}.$$

This is straightforward computation in which one uses the fact that $(\alpha_j|\alpha_k) = 0$ for $j \neq k$. ▌

In the special case in which $\{\alpha_1, \ldots, \alpha_n\}$ is an orthonormal set, Bessel's inequality says that

$$\sum_k |(\beta|\alpha_k)|^2 \leq ||\beta||^2.$$

The corollary also tells us in this case that β is in the subspace spanned by $\alpha_1, \ldots, \alpha_n$ if and only if

$$\beta = \sum_k (\beta|\alpha_k)\, \alpha_k$$

or if and only if Bessel's inequality is actually an equality. Of course, in the event that V is finite dimensional and $\{\alpha_1, \ldots, \alpha_n\}$ is an orthogonal basis for V, the above formula holds for every vector β in V. In other words, if $\{\alpha_1, \ldots, \alpha_n\}$ is an orthonormal basis for V, the kth coordinate of β in the ordered basis $\{\alpha_1, \ldots, \alpha_n\}$ is $(\beta|\alpha_k)$.

Example 15. We shall apply the last corollary to the orthogonal sets described in Example 11. We find that

(a) $$\sum_{k=-n}^{n} \left| \int_0^1 f(t) e^{-2\pi i k t}\, dt \right|^2 \leq \int_0^1 |f(t)|^2\, dt$$

(b) $$\int_0^1 \left| \sum_{k=-n}^{n} c_k e^{2\pi i k t} \right|^2 dt = \sum_{k=-n}^{n} |c_k|^2$$

(c) $$\int_0^1 (\sqrt{2} \cos 2\pi t + \sqrt{2} \sin 4\pi t)^2\, dt = 1 + 1 = 2.$$

Exercises

1. Consider R^4 with the standard inner product. Let W be the subspace of R^4 consisting of all vectors which are orthogonal to both $\alpha = (1, 0, -1, 1)$ and $\beta = (2, 3, -1, 2)$. Find a basis for W.

2. Apply the Gram-Schmidt process to the vectors $\beta_1 = (1, 0, 1)$, $\beta_2 = (1, 0, -1)$, $\beta_3 = (0, 3, 4)$, to obtain an orthonormal basis for R^3 with the standard inner product.

3. Consider C^3, with the standard inner product. Find an orthonormal basis for the subspace spanned by $\beta_1 = (1, 0, i)$ and $\beta_2 = (2, 1, 1 + i)$.

4. Let V be an inner product space. The **distance** between two vectors α and β in V is defined by

$$d(\alpha, \beta) = ||\alpha - \beta||.$$

Show that

(a) $d(\alpha, \beta) \geq 0$;
(b) $d(\alpha, \beta) = 0$ if and only if $\alpha = \beta$;
(c) $d(\alpha, \beta) = d(\beta, \alpha)$;
(d) $d(\alpha, \beta) \leq d(\alpha, \gamma) + d(\gamma, \beta)$.

5. Let V be an inner product space, and let α, β be vectors in V. Show that $\alpha = \beta$ if and only if $(\alpha|\gamma) = (\beta|\gamma)$ for every γ in V.

6. Let W be the subspace of R^2 spanned by the vector $(3, 4)$. Using the standard inner product, let E be the orthogonal projection of R^2 onto W. Find

(a) a formula for $E(x_1, x_2)$;
(b) the matrix of E in the standard ordered basis;
(c) $W^{\perp}$;
(d) an orthonormal basis in which E is represented by the matrix

$$\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}.$$

7. Let V be the inner product space consisting of R^2 and the inner product whose quadratic form is defined by

$$||(x_1, x_2||^2 = (x_1 - x_2)^2 + 3x_2^2.$$

Let E be the orthogonal projection of V onto the subspace W spanned by the vector $(3, 4)$. Now answer the four questions of Exercise 6.

8. Find an inner product on R^2 such that $(\epsilon_1, \epsilon_2) = 2$.

9. Let V be the subspace of $R[x]$ of polynomials of degree at most 3. Equip V with the inner product

$$(f|g) = \int_0^1 f(t)g(t)\, dt.$$

(a) Find the orthogonal complement of the subspace of scalar polynomials.
(b) Apply the Gram-Schmidt process to the basis $\{1, x, x^2, x^3\}$.

10. Let V be the vector space of all $n \times n$ matrices over C, with the inner product $(A|B) = \operatorname{tr}(AB^*)$. Find the orthogonal complement of the subspace of diagonal matrices.

11. Let V be a finite-dimensional inner product space, and let $\{\alpha_1, \ldots, \alpha_n\}$ be an orthonormal basis for V. Show that for any vectors α, β in V

$$(\alpha|\beta) = \sum_{k=1}^{n} (\alpha|\alpha_k)\overline{(\beta|\alpha_k)}.$$