LINEAR ALGEBRA

W W L CHEN

Chapter 9

REAL INNER PRODUCT SPACES

9.1. Euclidean Inner Products

In this section, we consider vectors of the form $\mathbf{u} = (u_1, \ldots, u_n)$ in the euclidean space $\mathbb{R}^n$. In particular, we shall generalize the concept of dot product, norm and distance, first developed for $\mathbb{R}^2$ and $\mathbb{R}^3$ in Chapter 4.

Definition. Suppose that $\mathbf{u} = (u_1, \ldots, u_n)$ and $\mathbf{v} = (v_1, \ldots, v_n)$ are vectors in $\mathbb{R}^n$. The euclidean dot product of $\mathbf{u}$ and $\mathbf{v}$ is defined by

$$\mathbf{u} \cdot \mathbf{v} = u_1 v_1 + \ldots + u_n v_n,$$

the euclidean norm of $\mathbf{u}$ is defined by

$$\|\mathbf{u}\| = (\mathbf{u} \cdot \mathbf{u})^{1/2} = (u_1^2 + \ldots + u_n^2)^{1/2},$$

and the euclidean distance between $\mathbf{u}$ and $\mathbf{v}$ is defined by

$$d(\mathbf{u}, \mathbf{v}) = \|\mathbf{u} - \mathbf{v}\| = ((u_1 - v_1)^2 + \ldots + (u_n - v_n)^2)^{1/2}.$$

PROPOSITION 9A. *Suppose that $\mathbf{u}, \mathbf{v}, \mathbf{w} \in \mathbb{R}^n$ and $c \in \mathbb{R}$. Then*
(a) $\mathbf{u} \cdot \mathbf{v} = \mathbf{v} \cdot \mathbf{u}$;
(b) $\mathbf{u} \cdot (\mathbf{v} + \mathbf{w}) = (\mathbf{u} \cdot \mathbf{v}) + (\mathbf{u} \cdot \mathbf{w})$;
(c) $c(\mathbf{u} \cdot \mathbf{v}) = (c\mathbf{u}) \cdot \mathbf{v}$; *and*
(d) $\mathbf{u} \cdot \mathbf{u} \geq 0$, *and* $\mathbf{u} \cdot \mathbf{u} = 0$ *if and only if* $\mathbf{u} = \mathbf{0}$.

PROPOSITION 9B. (CAUCHY-SCHWARZ INEQUALITY) *Suppose that* $\mathbf{u}, \mathbf{v} \in \mathbb{R}^n$. *Then*

$$|\mathbf{u} \cdot \mathbf{v}| \le \|\mathbf{u}\| \, \|\mathbf{v}\|.$$

In other words,

$$|u_1 v_1 + \ldots + u_n v_n| \le (u_1^2 + \ldots + u_n^2)^{1/2} (v_1^2 + \ldots + v_n^2)^{1/2}.$$

PROPOSITION 9C. *Suppose that* $\mathbf{u}, \mathbf{v} \in \mathbb{R}^n$ *and* $c \in \mathbb{R}$. *Then*
(a) $\|\mathbf{u}\| \ge 0$;
(b) $\|\mathbf{u}\| = 0$ *if and only if* $\mathbf{u} = \mathbf{0}$;
(c) $\|c\mathbf{u}\| = |c| \, \|\mathbf{u}\|$; *and*
(d) $\|\mathbf{u} + \mathbf{v}\| \le \|\mathbf{u}\| + \|\mathbf{v}\|$.

PROPOSITION 9D. *Suppose that* $\mathbf{u}, \mathbf{v}, \mathbf{w} \in \mathbb{R}^n$. *Then*
(a) $d(\mathbf{u}, \mathbf{v}) \ge 0$;
(b) $d(\mathbf{u}, \mathbf{v}) = 0$ *if and only if* $\mathbf{u} = \mathbf{v}$;
(c) $d(\mathbf{u}, \mathbf{v}) = d(\mathbf{v}, \mathbf{u})$; *and*
(d) $d(\mathbf{u}, \mathbf{v}) \le d(\mathbf{u}, \mathbf{w}) + d(\mathbf{w}, \mathbf{v})$.

Remark. Parts (d) of Propositions 9C and 9D are known as the Triangle inequality.

In $\mathbb{R}^2$ and $\mathbb{R}^3$, we say that two non-zero vectors are perpendicular if their dot product is zero. We now generalize this idea to vectors in $\mathbb{R}^n$.

Definition. Two vectors $\mathbf{u}, \mathbf{v} \in \mathbb{R}^n$ are said to be orthogonal if $\mathbf{u} \cdot \mathbf{v} = 0$.

Example 9.1.1. Suppose that $\mathbf{u}, \mathbf{v} \in \mathbb{R}^n$ are orthogonal. Then

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

This is an extension of Pythagoras's theorem.

Remarks. (1) Suppose that we write $\mathbf{u}, \mathbf{v} \in \mathbb{R}^n$ as column matrices. Then

$$\mathbf{u} \cdot \mathbf{v} = \mathbf{v}^t \mathbf{u},$$

where we use matrix multiplication on the right hand side.

(2) Matrix multiplication can be described in terms of dot product. Suppose that A is an $m \times n$ matrix and B is an $n \times p$ matrix. If we let $\mathbf{r}_1, \ldots, \mathbf{r}_m$ denote the vectors formed from the rows of A, and let $\mathbf{c}_1, \ldots, \mathbf{c}_p$ denote the vectors formed from the columns of B, then

$$AB = \begin{pmatrix} \mathbf{r}_1 \cdot \mathbf{c}_1 & \ldots & \mathbf{r}_1 \cdot \mathbf{c}_p \\ \vdots & & \vdots \\ \mathbf{r}_m \cdot \mathbf{c}_1 & \ldots & \mathbf{r}_m \cdot \mathbf{c}_p \end{pmatrix}.$$

9.2. Real Inner Products

The purpose of this section and the next is to extend our discussion to define inner products in real vector spaces. We begin by giving a reminder of the basics of real vector spaces or vector spaces over $\mathbb{R}$.

DEFINITION. A real vector space V is a set of objects, known as vectors, together with vector addition $+$ and multiplication of vectors by elements of $\mathbb{R}$, and satisfying the following properties:

(VA1) For every $\mathbf{u}, \mathbf{v} \in V$, we have $\mathbf{u} + \mathbf{v} \in V$.
(VA2) For every $\mathbf{u}, \mathbf{v}, \mathbf{w} \in V$, we have $\mathbf{u} + (\mathbf{v} + \mathbf{w}) = (\mathbf{u} + \mathbf{v}) + \mathbf{w}$.
(VA3) There exists an element $\mathbf{0} \in V$ such that for every $\mathbf{u} \in V$, we have $\mathbf{u} + \mathbf{0} = \mathbf{0} + \mathbf{u} = \mathbf{u}$.
(VA4) For every $\mathbf{u} \in V$, there exists $-\mathbf{u} \in V$ such that $\mathbf{u} + (-\mathbf{u}) = \mathbf{0}$.
(VA5) For every $\mathbf{u}, \mathbf{v} \in V$, we have $\mathbf{u} + \mathbf{v} = \mathbf{v} + \mathbf{u}$.
(SM1) For every $c \in \mathbb{R}$ and $\mathbf{u} \in V$, we have $c\mathbf{u} \in V$.
(SM2) For every $c \in \mathbb{R}$ and $\mathbf{u}, \mathbf{v} \in V$, we have $c(\mathbf{u} + \mathbf{v}) = c\mathbf{u} + c\mathbf{v}$.
(SM3) For every $a, b \in \mathbb{R}$ and $\mathbf{u} \in V$, we have $(a + b)\mathbf{u} = a\mathbf{u} + b\mathbf{u}$.
(SM4) For every $a, b \in \mathbb{R}$ and $\mathbf{u} \in V$, we have $(ab)\mathbf{u} = a(b\mathbf{u})$.
(SM5) For every $\mathbf{u} \in V$, we have $1\mathbf{u} = \mathbf{u}$.

REMARK. The elements $a, b, c \in \mathbb{R}$ discussed in (SM1)–(SM5) are known as scalars. Multiplication of vectors by elements of $\mathbb{R}$ is sometimes known as scalar multiplication.

DEFINITION. Suppose that V is a real vector space, and that W is a subset of V. Then we say that W is a subspace of V if W forms a real vector space under the vector addition and scalar multiplication defined in V.

REMARK. Suppose that V is a real vector space, and that W is a non-empty subset of V. Then W is a subspace of V if the following conditions are satisfied:

(SP1) For every $\mathbf{u}, \mathbf{v} \in W$, we have $\mathbf{u} + \mathbf{v} \in W$.
(SP2) For every $c \in \mathbb{R}$ and $\mathbf{u} \in W$, we have $c\mathbf{u} \in W$.

The reader may refer to Chapter 5 for more details and examples.

We are now in a position to define an inner product on a real vector space V. The following definition is motivated by Proposition 9A concerning the properties of the euclidean dot product in $\mathbb{R}^n$.

DEFINITION. Suppose that V is a real vector space. By a real inner product on V, we mean a function $\langle\ ,\ \rangle : V \times V \to \mathbb{R}$ which satisfies the following conditions:

(IP1) For every $\mathbf{u}, \mathbf{v} \in V$, we have $\langle \mathbf{u}, \mathbf{v} \rangle = \langle \mathbf{v}, \mathbf{u} \rangle$.
(IP2) For every $\mathbf{u}, \mathbf{v}, \mathbf{w} \in V$, we have $\langle \mathbf{u}, \mathbf{v} + \mathbf{w} \rangle = \langle \mathbf{u}, \mathbf{v} \rangle + \langle \mathbf{u}, \mathbf{w} \rangle$.
(IP3) For every $\mathbf{u}, \mathbf{v} \in V$ and $c \in \mathbb{R}$, we have $c\langle \mathbf{u}, \mathbf{v} \rangle = \langle c\mathbf{u}, \mathbf{v} \rangle$.
(IP4) For every $\mathbf{u} \in V$, we have $\langle \mathbf{u}, \mathbf{u} \rangle \geq 0$, and $\langle \mathbf{u}, \mathbf{u} \rangle = 0$ if and only if $\mathbf{u} = \mathbf{0}$.

REMARKS. (1) The properties (IP1)–(IP4) describe respectively symmetry, additivity, homogeneity and positivity.

(2) We sometimes simply refer to an inner product if we know that V is a real vector space.

DEFINITION. A real vector space with an inner product is called a real inner product space.

Our next definition is a natural extension of the idea of euclidean norm and euclidean distance.

DEFINITION. Suppose that $\mathbf{u}$ and $\mathbf{v}$ are vectors in a real inner pruduct space V. Then the norm of $\mathbf{u}$ is defined by

$$\|\mathbf{u}\| = \langle \mathbf{u}, \mathbf{u} \rangle^{1/2},$$

and the distance between $\mathbf{u}$ and $\mathbf{v}$ is defined by

$$d(\mathbf{u}, \mathbf{v}) = \|\mathbf{u} - \mathbf{v}\|.$$

Example 9.2.1. For $\mathbf{u}, \mathbf{v} \in \mathbb{R}^n$, let $\langle \mathbf{u}, \mathbf{v} \rangle = \mathbf{u} \cdot \mathbf{v}$, the euclidean dot product discussed in the last section. This satisfies Proposition 9A and hence conditions (IP1)–(IP4). The inner product is known as the euclidean inner product in $\mathbb{R}^n$.

Example 9.2.2. Let $w_1, \ldots, w_n$ be positive real numbers. For $\mathbf{u} = (u_1, \ldots, u_n)$ and $\mathbf{v} = (v_1, \ldots, v_n)$ in $\mathbb{R}^n$, let

$$\langle \mathbf{u}, \mathbf{v} \rangle = w_1 u_1 v_1 + \ldots + w_n u_n v_n.$$

It is easy to check that conditions (IP1)–(IP4) are satisfied. This inner product is called a weighted euclidean inner product in $\mathbb{R}^n$, and the positive real numbers $w_1, \ldots, w_n$ are known as weights. The unit circle with respect to this inner product is given by

$$\{\mathbf{u} \in \mathbb{R}^n : \|\mathbf{u}\| = 1\} = \{\mathbf{u} \in \mathbb{R}^n : \langle \mathbf{u}, \mathbf{u} \rangle = 1\} = \{\mathbf{u} \in \mathbb{R}^n : w_1 u_1^2 + \ldots + w_n u_n^2 = 1\}.$$

Example 9.2.3. Let A be a fixed invertible $n \times n$ matrix with real entries. For $\mathbf{u}, \mathbf{v} \in \mathbb{R}^n$, interpreted as column matrices, let

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

the euclidean dot product of the vectors $A\mathbf{u}$ and $A\mathbf{v}$. It can be checked that conditions (IP1)–(IP4) are satisfied. This inner product is called the inner product generated by the matrix A. To check conditions (IP1)–(IP4), it is useful to note that

$$\langle \mathbf{u}, \mathbf{v} \rangle = (A\mathbf{v})^t A\mathbf{u} = \mathbf{v}^t A^t A\mathbf{u}.$$

Example 9.2.4. Consider the vector space $\mathcal{M}_{2,2}(\mathbb{R})$ of all 2×2 matrices with real entries. For matrices

$$U = \begin{pmatrix} u_{11} & u_{12} \\ u_{21} & u_{22} \end{pmatrix} \quad \text{and} \quad V = \begin{pmatrix} v_{11} & v_{12} \\ v_{21} & v_{22} \end{pmatrix}$$

in $\mathcal{M}_{2,2}(\mathbb{R})$, let

$$\langle U, V \rangle = u_{11}v_{11} + u_{12}v_{12} + u_{21}v_{21} + u_{22}v_{22}.$$

It is easy to check that conditions (IP1)–(IP4) are satisfied.

Example 9.2.5. Consider the vector space P_2 of all polynomials with real coefficients and of degree at most 2. For polynomials

$$p = p(x) = p_0 + p_1 x + p_2 x^2 \quad \text{and} \quad q = q(x) = q_0 + q_1 x + q_2 x^2$$

in P_2, let

$$\langle p, q \rangle = p_0 q_0 + p_1 q_1 + p_2 q_2.$$

It can be checked that conditions (IP1)–(IP4) are satisfied.

Example 9.2.6. It is not difficult to show that $C[a, b]$, the collection of all real valued functions continuous in the closed interval $[a, b]$, forms a real vector space. We also know from the theory of real valued functions that functions continuous over a closed interval $[a, b]$ are integrable over $[a, b]$. For $f, g \in C[a, b]$, let

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

It can be checked that conditions (IP1)–(IP4) are satisfied.

9.3. Angles and Orthogonality

Recall that in $\mathbb{R}^2$ and $\mathbb{R}^3$, we can actually define the euclidean dot product of two vectors $\mathbf{u}$ and $\mathbf{v}$ by the formula

$$\mathbf{u} \cdot \mathbf{v} = \|\mathbf{u}\| \, \|\mathbf{v}\| \cos\theta, \tag{1}$$

where θ is the angle between $\mathbf{u}$ and $\mathbf{v}$. Indeed, this is the approach taken in Chapter 4, and the Cauchy-Schwarz inequality, as stated in Proposition 9B, follows immediately from (1), since $|\cos\theta| \leq 1$.

The picture is not so clear in the euclidean space $\mathbb{R}^n$ when $n > 3$, although the Cauchy-Schwarz inequality, as given by Proposition 9B, does allow us to recover a formula of the type (1). But then the number θ does not have a geometric interpretation.

We now study the case of a real inner product space. Our first task is to establish a generalized version of Proposition 9B.

PROPOSITION 9E. (CAUCHY-SCHWARZ INEQUALITY) *Suppose that* $\mathbf{u}$ *and* $\mathbf{v}$ *are vectors in a real inner product space* V*. Then*

$$|\langle \mathbf{u}, \mathbf{v} \rangle| \leq \|\mathbf{u}\| \, \|\mathbf{v}\|. \tag{2}$$

Proof. Our proof here looks like a trick, but it works. Suppose that $\mathbf{u}$ and $\mathbf{v}$ are vectors in a real inner product space V. If $\mathbf{u} = \mathbf{0}$, then since $0\mathbf{u} = \mathbf{0}$, it follows that

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

so that (2) is clearly satisfied. We may suppose therefore that $\mathbf{u} \neq \mathbf{0}$, so that $\langle \mathbf{u}, \mathbf{u} \rangle \neq 0$. For every real number t, it follows from (IP4) that $\langle t\mathbf{u} + \mathbf{v}, t\mathbf{u} + \mathbf{v} \rangle \geq 0$. Hence

$$0 \leq \langle t\mathbf{u} + \mathbf{v}, t\mathbf{u} + \mathbf{v} \rangle = t^2\langle \mathbf{u}, \mathbf{u} \rangle + 2t\langle \mathbf{u}, \mathbf{v} \rangle + \langle \mathbf{v}, \mathbf{v} \rangle.$$

Since $\langle \mathbf{u}, \mathbf{u} \rangle \neq 0$, the right hand side is a quadratic polynomial in t. Since the inequality holds for every real number t, it follows that the quadratic polynomial

$$t^2\langle \mathbf{u}, \mathbf{u} \rangle + 2t\langle \mathbf{u}, \mathbf{v} \rangle + \langle \mathbf{v}, \mathbf{v} \rangle$$

has either repeated roots or no real root, and so the discriminant is non-positive. In other words, we must have

$$0 \geq (2\langle \mathbf{u}, \mathbf{v} \rangle)^2 - 4\langle \mathbf{u}, \mathbf{u} \rangle \langle \mathbf{v}, \mathbf{v} \rangle = 4\langle \mathbf{u}, \mathbf{v} \rangle^2 - 4\|\mathbf{u}\|^2\|\mathbf{v}\|^2.$$

The inequality (2) follows once again. ◯

Example 9.3.1. Note that Proposition 9B is a special case of Proposition 9E. In fact, Proposition 9B represents the Cauchy-Schwarz inequality for finite sums, that for $u_1, \ldots, u_n, v_1, \ldots, v_n \in \mathbb{R}$, we have

$$\left| \sum_{i=1}^n u_i v_i \right| \leq \left(\sum_{i=1}^n u_i^2 \right)^{1/2} \left(\sum_{i=1}^n v_i^2 \right)^{1/2}.$$

Example 9.3.2. Applying Proposition 9E to the inner product in the vector space $C[a,b]$ studied in Example 9.2.6, we obtain the Cauchy-Schwarz inequality for integrals, that for $f, g \in C[a,b]$, we have

$$\left| \int_a^b f(x)g(x) \, \mathrm{d}x \right| \leq \left(\int_a^b f^2(x) \, \mathrm{d}x \right)^{1/2} \left(\int_a^b g^2(x) \, \mathrm{d}x \right)^{1/2}.$$

Next, we investigate norm and distance. We generalize Propositions 9C and 9D.

PROPOSITION 9F. *Suppose that* $\mathbf{u}$ *and* $\mathbf{v}$ *are vectors in a real inner product space, and that* $c \in \mathbb{R}$. *Then*
(a) $\|\mathbf{u}\| \geq 0$;
(b) $\|\mathbf{u}\| = 0$ *if and only if* $\mathbf{u} = \mathbf{0}$;
(c) $\|c\mathbf{u}\| = |c|\,\|\mathbf{u}\|$; *and*
(d) $\|\mathbf{u}+\mathbf{v}\| \leq \|\mathbf{u}\| + \|\mathbf{v}\|$.

PROPOSITION 9G. *Suppose that* $\mathbf{u}$, $\mathbf{v}$ *and* $\mathbf{w}$ *are vectors in a real inner product space. Then*
(a) $d(\mathbf{u},\mathbf{v}) \geq 0$;
(b) $d(\mathbf{u},\mathbf{v}) = 0$ *if and only if* $\mathbf{u} = \mathbf{v}$;
(c) $d(\mathbf{u},\mathbf{v}) = d(\mathbf{v},\mathbf{u})$; *and*
(d) $d(\mathbf{u},\mathbf{v}) \leq d(\mathbf{u},\mathbf{w}) + d(\mathbf{w},\mathbf{v})$.

The proofs are left as exercises.

The Cauchy-Schwarz inequality, as given by Proposition 9E, allows us to recover a formula of the type

$$\langle \mathbf{u},\mathbf{v}\rangle = \|\mathbf{u}\|\,\|\mathbf{v}\|\cos\theta. \tag{3}$$

Although the number θ does not have a geometric interpretation, we can nevertheless interpret it as the angle between the two vectors $\mathbf{u}$ and $\mathbf{v}$ under the inner product $\langle\ ,\ \rangle$. Of particular interest is the case when $\cos\theta = 0$; in other words, when $\langle \mathbf{u},\mathbf{v}\rangle = 0$.

Definition. Suppose that $\mathbf{u}$ and $\mathbf{v}$ are non-zero vectors in a real inner product space V. Then the unique real number $\theta \in [0,\pi]$ satisfying (3) is called the angle between $\mathbf{u}$ and $\mathbf{v}$ with respect to the inner product $\langle\ ,\ \rangle$ in V.

Definition. Two vectors $\mathbf{u}$ and $\mathbf{v}$ in a real inner product space are said to be orthogonal if $\langle \mathbf{u},\mathbf{v}\rangle = 0$.

Definition. Suppose that W is a subspace of a real inner product space V. A vector $\mathbf{u} \in V$ is said to be orthogonal to W if $\langle \mathbf{u},\mathbf{w}\rangle = 0$ for every $\mathbf{w} \in W$. The set of all vectors $\mathbf{u} \in V$ which are orthogonal to W is called the orthogonal complement of W, and denoted by $W^\perp$; in other words,

$$W^\perp = \{\mathbf{u} \in V : \langle \mathbf{u},\mathbf{w}\rangle = 0 \text{ for every } \mathbf{w} \in W\}.$$

Example 9.3.3. In $\mathbb{R}^3$, the non-trivial subspaces are lines and planes through the origin. Under the euclidean inner product, two non-zero vectors are orthogonal if and only if they are perpendicular. It follows that if W is a line through the origin, then $W^\perp$ is the plane through the origin and perpendicular to the line W. Also, if W is a plane through the origin, then $W^\perp$ is the line through the origin and perpendicular to the plane W.

Example 9.3.4. In $\mathbb{R}^4$, let us consider the two vectors $\mathbf{u} = (1,1,1,0)$ and $\mathbf{v} = (1,0,1,1)$. Under the euclidean inner product, we have

$$\|\mathbf{u}\| = \|\mathbf{v}\| = \sqrt{3} \qquad \text{and} \qquad \langle \mathbf{u},\mathbf{v}\rangle = 2.$$

This verifies the Cauchy-Schwarz inequality. On the other hand, if $\theta \in [0,\pi]$ represents the angle between $\mathbf{u}$ and $\mathbf{v}$ with respect to the euclidean inner product, then (3) holds, and we obtain $\cos\theta = 2/3$, so that $\theta = \cos^{-1}(2/3)$.

EXAMPLE 9.3.5. In $\mathbb{R}^4$, it can be shown that

$$W = \{(w_1, w_2, 0, 0) : w_1, w_2 \in \mathbb{R}\}$$

is a subspace. Consider now the euclidean inner product, and let

$$A = \{(0, 0, u_3, u_4) : u_3, u_4 \in \mathbb{R}\}.$$

We shall show that $A \subseteq W^\perp$ and $W^\perp \subseteq A$, so that $W^\perp = A$. To show that $A \subseteq W^\perp$, note that for every $(0, 0, u_3, u_4) \in A$, we have

$$\langle (0, 0, u_3, u_4), (w_1, w_2, 0, 0) \rangle = (0, 0, u_3, u_4) \cdot (w_1, w_2, 0, 0) = 0$$

for every $(w_1, w_2, 0, 0) \in W$, so that $(0, 0, u_3, u_4) \in W^\perp$. To show that $W^\perp \subseteq A$, note that for every $(u_1, u_2, u_3, u_4) \in W^\perp$, we need to have

$$\langle (u_1, u_2, u_3, u_4), (w_1, w_2, 0, 0) \rangle = (u_1, u_2, u_3, u_4) \cdot (w_1, w_2, 0, 0) = u_1 w_1 + u_2 w_2 = 0$$

for every $(w_1, w_2, 0, 0) \in W$. The choice $(w_1, w_2, 0, 0) = (1, 0, 0, 0)$ requires us to have $u_1 = 0$, while the choice $(w_1, w_2, 0, 0) = (0, 1, 0, 0)$ requires us to have $u_2 = 0$. Hence we must have $u_1 = u_2 = 0$, so that $(u_1, u_2, u_3, u_4) \in A$.

EXAMPLE 9.3.6. Let us consider the inner product on $\mathcal{M}_{2,2}(\mathbb{R})$ discussed in Example 9.2.4. Let

$$U = \begin{pmatrix} 1 & 0 \\ 3 & 4 \end{pmatrix} \qquad \text{and} \qquad V = \begin{pmatrix} 4 & 2 \\ 0 & -1 \end{pmatrix}.$$

Then $\langle U, V \rangle = 0$, so that the two matrices are orthogonal.

EXAMPLE 9.3.7. Let us consider the inner product on P_2 discussed in Example 9.2.5. Let

$$p = p(x) = 1 + 2x + 3x^2 \qquad \text{and} \qquad q = q(x) = 4 + x - 2x^2.$$

Then $\langle p, q \rangle = 0$, so that the two polynomials are orthogonal.

EXAMPLE 9.3.8. Let us consider the inner product on $C[a, b]$ discussed in Example 9.2.6. In particular, let $[a, b] = [0, \pi/2]$. Suppose that

$$f(x) = \sin x - \cos x \qquad \text{and} \qquad g(x) = \sin x + \cos x.$$

Then

$$\langle f, g \rangle = \int_0^{\pi/2} f(x)g(x)\,\mathrm{d}x = \int_0^{\pi/2} (\sin x - \cos x)(\sin x + \cos x)\,\mathrm{d}x = \int_0^{\pi/2} (\sin^2 x - \cos^2 x)\,\mathrm{d}x = 0,$$

so that the two functions are orthogonal.

EXAMPLE 9.3.9. Suppose that A is an $m \times n$ matrix with real entries. Recall that if we let $\mathbf{r}_1, \ldots, \mathbf{r}_m$ denote the vectors formed from the rows of A, then the row space of A is given by

$$\{c_1\mathbf{r}_1 + \ldots + c_m\mathbf{r}_m : c_1, \ldots, c_m \in \mathbb{R}\},$$

and is a subspace of $\mathbb{R}^n$. On the other hand, the set

$$\{\mathbf{x} \in \mathbb{R}^n : A\mathbf{x} = \mathbf{0}\}$$

is called the nullspace of A, and is also a subspace of $\mathbb{R}^n$. Clearly, if $\mathbf{x}$ belongs to the nullspace of A, then $\mathbf{r}_i \cdot \mathbf{x} = 0$ for every $i = 1, \ldots, m$. In fact, the row space of A and the nullspace of A are orthogonal complements of each other under the euclidean inner product in $\mathbb{R}^n$. On the other hand, the column space of A is the row space of A^t. It follows that the column space of A and the nullspace of A^t are orthogonal complements of each other under the euclidean inner product in $\mathbb{R}^m$.

EXAMPLE 9.3.10. Suppose that $\mathbf{u}$ and $\mathbf{v}$ are orthogonal vectors in an inner product space. Then

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

This is a generalized version of Pythagoras's theorem.

REMARK. We emphasize here that orthogonality depends on the choice of the inner product. Very often, a real vector space has more than one inner product. Vectors orthogonal with respect to one may not be orthogonal with respect to another. For example, the vectors $\mathbf{u} = (1,1)$ and $\mathbf{v} = (1,-1)$ in $\mathbb{R}^2$ are orthogonal with respect to the euclidean inner product

$$\langle \mathbf{u}, \mathbf{v} \rangle = u_1 v_1 + u_2 v_2,$$

but not orthogonal with respect to the weighted euclidean inner product

$$\langle \mathbf{u}, \mathbf{v} \rangle = 2u_1 v_1 + u_2 v_2.$$

9.4. Orthogonal and Orthonormal Bases

Suppose that $\mathbf{v}_1, \ldots, \mathbf{v}_r$ are vectors in a real vector space V. We often consider linear combinations of the type $c_1\mathbf{v}_1 + \ldots + c_r\mathbf{v}_r$, where $c_1, \ldots, c_r \in \mathbb{R}$. The set

$$\mathrm{span}\{\mathbf{v}_1, \ldots, \mathbf{v}_r\} = \{c_1\mathbf{v}_1 + \ldots + c_r\mathbf{v}_r : c_1, \ldots, c_r \in \mathbb{R}\}$$

of all such linear combinations is called the span of the vectors $\mathbf{v}_1, \ldots, \mathbf{v}_r$. We also say that the vectors $\mathbf{v}_1, \ldots, \mathbf{v}_r$ span V if $\mathrm{span}\{\mathbf{v}_1, \ldots, \mathbf{v}_r\} = V$; in other words, if every vector in V can be expressed as a linear combination of the vectors $\mathbf{v}_1, \ldots, \mathbf{v}_r$.

It can be shown that $\mathrm{span}\{\mathbf{v}_1, \ldots, \mathbf{v}_r\}$ is a subspace of V. Suppose further that W is a subspace of V and $\mathbf{v}_1, \ldots, \mathbf{v}_r \in W$. Then $\mathrm{span}\{\mathbf{v}_1, \ldots, \mathbf{v}_r\} \subseteq W$.

On the other hand, the spanning set $\{\mathbf{v}_1, \ldots, \mathbf{v}_r\}$ may contain more vectors than are necessary to describe all the vectors in the span. This leads to the idea of linear independence.

DEFINITION. Suppose that $\mathbf{v}_1, \ldots, \mathbf{v}_r$ are vectors in a real vector space V.

(LD) We say that $\mathbf{v}_1, \ldots, \mathbf{v}_r$ are linearly dependent if there exist $c_1, \ldots, c_r \in \mathbb{R}$, not all zero, such that $c_1\mathbf{v}_1 + \ldots + c_r\mathbf{v}_r = \mathbf{0}$.

(LI) We say that $\mathbf{v}_1, \ldots, \mathbf{v}_r$ are linearly independent if they are not linearly dependent; in other words, if the only solution of $c_1\mathbf{v}_1 + \ldots + c_r\mathbf{v}_r = \mathbf{0}$ in $c_1, \ldots, c_r \in \mathbb{R}$ is given by $c_1 = \ldots = c_r = 0$.

DEFINITION. Suppose that $\mathbf{v}_1, \ldots, \mathbf{v}_r$ are vectors in a real vector space V. We say that $\{\mathbf{v}_1, \ldots, \mathbf{v}_r\}$ is a basis for V if the following two conditions are satisfied:

(B1) We have $\mathrm{span}\{\mathbf{v}_1, \ldots, \mathbf{v}_r\} = V$.

(B2) The vectors $\mathbf{v}_1, \ldots, \mathbf{v}_r$ are linearly independent.

Suppose that $\{\mathbf{v}_1, \ldots, \mathbf{v}_r\}$ is a basis for a real vector space V. Then it can be shown that every element $\mathbf{u} \in V$ can be expressed uniquely in the form $\mathbf{u} = c_1\mathbf{v}_1 + \ldots + c_r\mathbf{v}_r$, where $c_1, \ldots, c_r \in \mathbb{R}$.

We shall restrict our discussion to finite-dimensional real vector spaces. A real vector space V is said to be finite-dimensional if it has a basis containing only finitely many elements. Suppose that $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$

is such a basis. Then it can be shown that any collection of more than n vectors in V must be linearly dependent. It follows that any two bases for V must have the same number of elements. This common number is known as the dimension of V.

It can be shown that if V is a finite-dimensional real vector space, then any finite set of linearly independent vectors in V can be expanded, if necessary, to a basis for V. This establishes the existence of a basis for any finite-dimensional vector space. On the other hand, it can be shown that if the dimension of V is equal to n, then any set of n linearly independent vectors in V is a basis for V.

REMARK. The above is discussed in far greater detail, including examples and proofs, in Chapter 5.

The purpose of this section is to add the extra ingredient of orthogonality to the above discussion.

DEFINITION. Suppose that V is a finite-dimensional real inner product space. A basis $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ of V is said to be an orthogonal basis of V if $\langle \mathbf{v}_i, \mathbf{v}_j \rangle = 0$ for every $i, j = 1, \ldots, n$ satisfying $i \neq j$. It is said to be an orthonormal basis if it satisfies the extra condition that $\|\mathbf{v}_i\| = 1$ for every $i = 1, \ldots, n$.

EXAMPLE 9.4.1. The usual basis $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ in $\mathbb{R}^n$, where

$$\mathbf{v}_i = (\underbrace{0, \ldots, 0}_{i-1}, 1, \underbrace{0, \ldots, 0}_{n-i})$$

for every $i = 1, \ldots, n$, is an orthonormal basis of $\mathbb{R}^n$ with respect to the euclidean inner product.

EXAMPLE 9.4.2. The vectors $\mathbf{v}_1 = (1, 1)$ and $\mathbf{v}_2 = (1, -1)$ are linearly independent in $\mathbb{R}^2$ and satisfy

$$\langle \mathbf{v}_1, \mathbf{v}_2 \rangle = \mathbf{v}_1 \cdot \mathbf{v}_2 = 0.$$

It follows that $\{\mathbf{v}_1, \mathbf{v}_2\}$ is an orthogonal basis of $\mathbb{R}^2$ with respect to the euclidean inner product. Can you find an orthonormal basis of $\mathbb{R}^2$ by normalizing $\mathbf{v}_1$ and $\mathbf{v}_2$?

It is theoretically very simple to express any vector as a linear combination of the elements of an orthogonal or orthonormal basis.

PROPOSITION 9H. *Suppose that V is a finite-dimensional real inner product space. If $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ is an orthogonal basis of V, then for every vector $\mathbf{u} \in V$, we have*

$$\mathbf{u} = \frac{\langle \mathbf{u}, \mathbf{v}_1 \rangle}{\|\mathbf{v}_1\|^2}\mathbf{v}_1 + \ldots + \frac{\langle \mathbf{u}, \mathbf{v}_n \rangle}{\|\mathbf{v}_n\|^2}\mathbf{v}_n.$$

Furthermore, if $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ is an orthonormal basis of V, then for every vector $\mathbf{u} \in V$, we have

$$\mathbf{u} = \langle \mathbf{u}, \mathbf{v}_1 \rangle \mathbf{v}_1 + \ldots + \langle \mathbf{u}, \mathbf{v}_n \rangle \mathbf{v}_n.$$

PROOF. Since $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ is a basis of V, there exist unique $c_1, \ldots, c_n \in \mathbb{R}$ such that

$$\mathbf{u} = c_1\mathbf{v}_1 + \ldots + c_n\mathbf{v}_n.$$

For every $i = 1, \ldots, n$, we have

$$\langle \mathbf{u}, \mathbf{v}_i \rangle = \langle c_1\mathbf{v}_1 + \ldots + c_n\mathbf{v}_n, \mathbf{v}_i \rangle = c_1\langle \mathbf{v}_1, \mathbf{v}_i \rangle + \ldots + c_n\langle \mathbf{v}_n, \mathbf{v}_i \rangle = c_i\langle \mathbf{v}_i, \mathbf{v}_i \rangle$$

since $\langle \mathbf{v}_j, \mathbf{v}_i \rangle = 0$ if $j \neq i$. Clearly $\mathbf{v}_i \neq \mathbf{0}$, so that $\langle \mathbf{v}_i, \mathbf{v}_i \rangle \neq 0$, and so

$$c_i = \frac{\langle \mathbf{u}, \mathbf{v}_i \rangle}{\langle \mathbf{v}_i, \mathbf{v}_i \rangle}$$

for every $i = 1, \ldots, n$. The first assertion follows immediately. For the second assertion, note that $\langle \mathbf{v}_i, \mathbf{v}_i \rangle = 1$ for every $i = 1, \ldots, n$. ◯

Collections of vectors that are orthogonal to each other are very useful in the study of vector spaces, as illustrated by the following important result.

PROPOSITION 9J. *Suppose that the non-zero vectors $\mathbf{v}_1, \ldots, \mathbf{v}_r$ in a finite-dimensional real inner product space are pairwise orthogonal. Then they are linearly independent.*

PROOF. Suppose that $c_1, \ldots, c_r \in \mathbb{R}$ and

$$c_1\mathbf{v}_1 + \ldots + c_r\mathbf{v}_r = \mathbf{0}.$$

Then for every $i = 1, \ldots, r$, we have

$$0 = \langle \mathbf{0}, \mathbf{v}_i \rangle = \langle c_1\mathbf{v}_1 + \ldots + c_r\mathbf{v}_r, \mathbf{v}_i \rangle = c_1\langle \mathbf{v}_1, \mathbf{v}_i \rangle + \ldots + c_r\langle \mathbf{v}_r, \mathbf{v}_i \rangle = c_i\langle \mathbf{v}_i, \mathbf{v}_i \rangle$$

since $\langle \mathbf{v}_j, \mathbf{v}_i \rangle = 0$ if $j \neq i$. Clearly $\mathbf{v}_i \neq \mathbf{0}$, so that $\langle \mathbf{v}_i, \mathbf{v}_i \rangle \neq 0$, and so we must have $c_i = 0$ for every $i = 1, \ldots, r$. It follows that $c_1 = \ldots = c_r = 0$. ◯

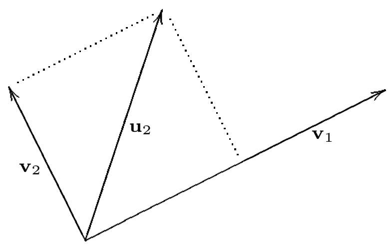
Of course, the above is based on the assumption that an orthogonal basis exists. Our next task is to show that this is indeed the case. Our proof is based on a technique which orthogonalizes any given basis of a vector space.

PROPOSITION 9K. *Every finite-dimensional real inner product space has an orthogonal basis, and hence also an orthonormal basis.*

REMARK. We shall prove Proposition 9K by using the Gram-Schmidt process. The central idea of this process, in its simplest form, can be described as follows. Suppose that $\mathbf{v}_1$ and $\mathbf{u}_2$ are two non-zero vectors in an inner product space, not necessarily orthogonal to each other. We shall attempt to remove some scalar multiple $\alpha_1\mathbf{v}_1$ from $\mathbf{u}_2$ so that $\mathbf{v}_2 = \mathbf{u}_2 - \alpha_1\mathbf{v}_1$ is orthogonal to $\mathbf{v}_1$; in other words, we wish to find a suitable real number α_1 such that

$$\langle \mathbf{v}_1, \mathbf{v}_2 \rangle = \langle \mathbf{v}_1, \mathbf{u}_2 - \alpha_1\mathbf{v}_1 \rangle = 0.$$

The idea is illustrated in the picture below.



We clearly need $\langle \mathbf{v}_1, \mathbf{u}_2 \rangle - \alpha_1\langle \mathbf{v}_1, \mathbf{v}_1 \rangle = 0$, and

$$\alpha_1 = \frac{\langle \mathbf{v}_1, \mathbf{u}_2 \rangle}{\langle \mathbf{v}_1, \mathbf{v}_1 \rangle} = \frac{\langle \mathbf{v}_1, \mathbf{u}_2 \rangle}{\|\mathbf{v}_1\|^2}$$

is a suitable choice, so that

$$\mathbf{v}_1 \qquad \text{and} \qquad \mathbf{v}_2 = \mathbf{u}_2 - \frac{\langle \mathbf{v}_1, \mathbf{u}_2 \rangle}{\|\mathbf{v}_1\|^2}\mathbf{v}_1 \tag{4}$$

are now orthogonal. Suppose in general that $\mathbf{v}_1, \ldots, \mathbf{v}_s$ and $\mathbf{u}_{s+1}$ are non-zero vectors in an inner product space, where $\mathbf{v}_1, \ldots, \mathbf{v}_s$ are pairwise orthogonal. We shall attempt to remove some linear combination

$\alpha_1\mathbf{v}_1+\ldots+\alpha_s\mathbf{v}_s$ from $\mathbf{u}_{s+1}$ so that $\mathbf{v}_{s+1}=\mathbf{u}_{s+1}-\alpha_1\mathbf{v}_1-\ldots-\alpha_s\mathbf{v}_s$ is orthogonal to each of $\mathbf{v}_1,\ldots,\mathbf{v}_s$; in other words, we wish to find suitable real numbers $\alpha_1,\ldots,\alpha_s$ such that

$$\langle\mathbf{v}_i,\mathbf{v}_{s+1}\rangle=\langle\mathbf{v}_i,\mathbf{u}_{s+1}-\alpha_1\mathbf{v}_1-\ldots-\alpha_s\mathbf{v}_s\rangle=0$$

for every $i=1,\ldots,s$. We clearly need

$$\langle\mathbf{v}_i,\mathbf{u}_{s+1}\rangle-\alpha_1\langle\mathbf{v}_i,\mathbf{v}_1\rangle-\ldots-\alpha_s\langle\mathbf{v}_i,\mathbf{v}_s\rangle=\langle\mathbf{v}_i,\mathbf{u}_{s+1}\rangle-\alpha_i\langle\mathbf{v}_i,\mathbf{v}_i\rangle=0,$$

and

$$\alpha_i=\frac{\langle\mathbf{v}_i,\mathbf{u}_{s+1}\rangle}{\langle\mathbf{v}_i,\mathbf{v}_i\rangle}=\frac{\langle\mathbf{v}_i,\mathbf{u}_{s+1}\rangle}{\|\mathbf{v}_i\|^2}$$

is a suitable choice, so that

$$\mathbf{v}_1,\ldots,\mathbf{v}_s \qquad\text{and}\qquad \mathbf{v}_{s+1}=\mathbf{u}_{s+1}-\frac{\langle\mathbf{v}_1,\mathbf{u}_{s+1}\rangle}{\|\mathbf{v}_1\|^2}\mathbf{v}_1-\ldots-\frac{\langle\mathbf{v}_s,\mathbf{u}_{s+1}\rangle}{\|\mathbf{v}_s\|^2}\mathbf{v}_s \tag{5}$$

are now pairwise orthogonal.

EXAMPLE 9.4.3. The vectors

$$\mathbf{u}_1=(1,2,1,0),\qquad \mathbf{u}_2=(3,3,3,0),\qquad \mathbf{u}_3=(2,-10,0,0),\qquad \mathbf{u}_4=(-2,1,-6,2)$$

are linearly independent in $\mathbb{R}^4$, since

$$\det\begin{pmatrix}1&3&2&-2\\2&3&-10&1\\1&3&0&-6\\0&0&0&2\end{pmatrix}\neq 0.$$

Hence $\{\mathbf{u}_1,\mathbf{u}_2,\mathbf{u}_3,\mathbf{u}_4\}$ is a basis of $\mathbb{R}^4$. Let us consider $\mathbb{R}^4$ as a real inner product space with the euclidean inner product, and apply the Gram-Schmidt process to this basis. We have

$$\begin{aligned}
\mathbf{v}_1&=\mathbf{u}_1=(1,2,1,0),\\
\mathbf{v}_2&=\mathbf{u}_2-\frac{\langle\mathbf{v}_1,\mathbf{u}_2\rangle}{\|\mathbf{v}_1\|^2}\mathbf{v}_1=(3,3,3,0)-\frac{\langle(1,2,1,0),(3,3,3,0)\rangle}{\|(1,2,1,0)\|^2}(1,2,1,0)\\
&=(3,3,3,0)-\frac{12}{6}(1,2,1,0)=(3,3,3,0)+(-2,-4,-2,0)=(1,-1,1,0),\\
\mathbf{v}_3&=\mathbf{u}_3-\frac{\langle\mathbf{v}_1,\mathbf{u}_3\rangle}{\|\mathbf{v}_1\|^2}\mathbf{v}_1-\frac{\langle\mathbf{v}_2,\mathbf{u}_3\rangle}{\|\mathbf{v}_2\|^2}\mathbf{v}_2\\
&=(2,-10,0,0)-\frac{\langle(1,2,1,0),(2,-10,0,0)\rangle}{\|(1,2,1,0)\|^2}(1,2,1,0)-\frac{\langle(1,-1,1,0),(2,-10,0,0)\rangle}{\|(1,-1,1,0)\|^2}(1,-1,1,0)\\
&=(2,-10,0,0)+\frac{18}{6}(1,2,1,0)-\frac{12}{3}(1,-1,1,0)\\
&=(2,-10,0,0)+(3,6,3,0)+(-4,4,-4,0)=(1,0,-1,0),\\
\mathbf{v}_4&=\mathbf{u}_4-\frac{\langle\mathbf{v}_1,\mathbf{u}_4\rangle}{\|\mathbf{v}_1\|^2}\mathbf{v}_1-\frac{\langle\mathbf{v}_2,\mathbf{u}_4\rangle}{\|\mathbf{v}_2\|^2}\mathbf{v}_2-\frac{\langle\mathbf{v}_3,\mathbf{u}_4\rangle}{\|\mathbf{v}_3\|^2}\mathbf{v}_3\\
&=(-2,1,-6,2)-\frac{\langle(1,2,1,0),(-2,1,-6,2)\rangle}{\|(1,2,1,0)\|^2}(1,2,1,0)\\
&\quad-\frac{\langle(1,-1,1,0),(-2,1,-6,2)\rangle}{\|(1,-1,1,0)\|^2}(1,-1,1,0)-\frac{\langle(1,0,-1,0),(-2,1,-6,2)\rangle}{\|(1,0,-1,0)\|^2}(1,0,-1,0)\\
&=(-2,1,-6,2)+\frac{6}{6}(1,2,1,0)+\frac{9}{3}(1,-1,1,0)-\frac{4}{2}(1,0,-1,0)\\
&=(-2,1,-6,2)+(1,2,1,0)+(3,-3,3,0)+(-2,0,2,0)=(0,0,0,2).
\end{aligned}$$

It is easy to verify that the four vectors

$$\mathbf{v}_1 = (1,2,1,0), \qquad \mathbf{v}_2 = (1,-1,1,0), \qquad \mathbf{v}_3 = (1,0,-1,0), \qquad \mathbf{v}_4 = (0,0,0,2)$$

are pairwise orthogonal, so that $\{\mathbf{v}_1,\mathbf{v}_2,\mathbf{v}_3,\mathbf{v}_4\}$ is an orthogonal basis of $\mathbb{R}^4$. Normalizing each of these four vectors, we obtain the corresponding orthonormal basis

$$\left\{ \left(\frac{1}{\sqrt{6}},\frac{2}{\sqrt{6}},\frac{1}{\sqrt{6}},0\right), \left(\frac{1}{\sqrt{3}},-\frac{1}{\sqrt{3}},\frac{1}{\sqrt{3}},0\right), \left(\frac{1}{\sqrt{2}},0,-\frac{1}{\sqrt{2}},0\right), (0,0,0,1) \right\}.$$

Proof of Proposition 9K. Suppose that the vector space V has dimension of n. Then it has a basis of the type $\{\mathbf{u}_1,\ldots,\mathbf{u}_n\}$. We now let $\mathbf{v}_1 = \mathbf{u}_1$, and define $\mathbf{v}_2,\ldots,\mathbf{v}_n$ inductively by (4) and (5) to obtain a set of pairwise orthogonal vectors $\{\mathbf{v}_1,\ldots,\mathbf{v}_n\}$. Clearly none of these n vectors is zero, for if $\mathbf{v}_{s+1} = \mathbf{0}$, then it follows from (5) that $\mathbf{v}_1,\ldots,\mathbf{v}_s,\mathbf{u}_{s+1}$, and hence $\mathbf{u}_1,\ldots,\mathbf{u}_s,\mathbf{u}_{s+1}$, are linearly dependent, clearly a contradiction. It now follows from Proposition 9J that $\mathbf{v}_1,\ldots,\mathbf{v}_n$ are linearly independent, and so must form a basis of V. This proves the first assertion. To prove the second assertion, observe that each of the vectors

$$\frac{\mathbf{v}_1}{\|\mathbf{v}_1\|},\ldots,\frac{\mathbf{v}_n}{\|\mathbf{v}_n\|}$$

has norm 1. ○

Example 9.4.4. Consider the real inner product space P_2, where for polynomials

$$p = p(x) = p_0 + p_1x + p_2x^2 \qquad \text{and} \qquad q = q(x) = q_0 + q_1x + q_2x^2,$$

the inner product is defined by

$$\langle p,q\rangle = p_0q_0 + p_1q_1 + p_2q_2.$$

The polynomials

$$u_1 = 3+4x+5x^2, \qquad u_2 = 9+12x+5x^2, \qquad u_3 = 1-7x+25x^2$$

are linearly independent in P_2, since

$$\det\begin{pmatrix} 3 & 9 & 1 \\ 4 & 12 & -7 \\ 5 & 5 & 25 \end{pmatrix} \neq 0.$$

Hence $\{u_1,u_2,u_3\}$ is a basis of P_2. Let us apply the Gram-Schmidt process to this basis. We have

$$\begin{aligned}
v_1 &= u_1 = 3+4x+5x^2, \\
v_2 &= u_2 - \frac{\langle v_1,u_2\rangle}{\|v_1\|^2}v_1 = (9+12x+5x^2) - \frac{\langle 3+4x+5x^2, 9+12x+5x^2\rangle}{\|3+4x+5x^2\|^2}(3+4x+5x^2) \\
&= (9+12x+5x^2) - \frac{100}{50}(3+4x+5x^2) = (9+12x+5x^2) + (-6-8x-10x^2) = 3+4x-5x^2, \\
v_3 &= u_3 - \frac{\langle v_1,u_3\rangle}{\|v_1\|^2}v_1 - \frac{\langle v_2,u_3\rangle}{\|v_2\|^2}v_2 \\
&= (1-7x+25x^2) - \frac{\langle 3+4x+5x^2, 1-7x+25x^2\rangle}{\|3+4x+5x^2\|^2}(3+4x+5x^2) \\
&\quad - \frac{\langle 3+4x-5x^2, 1-7x+25x^2\rangle}{\|3+4x-5x^2\|^2}(3+4x-5x^2) \\
&= (1-7x+25x^2) - \frac{100}{50}(3+4x+5x^2) + \frac{150}{50}(3+4x-5x^2) \\
&= (1-7x+25x^2) + (-6-8x-10x^2) + (9+12x-15x^2) = 4-3x+0x^2.
\end{aligned}$$

It is easy to verify that the three polynomials

$$v_1 = 3 + 4x + 5x^2, \qquad v_2 = 3 + 4x - 5x^2, \qquad v_3 = 4 - 3x + 0x^2$$

are pairwise orthogonal, so that $\{v_1, v_2, v_3\}$ is an orthogonal basis of P_2. Normalizing each of these three polynomials, we obtain the corresponding orthonormal basis

$$\left\{ \frac{3}{\sqrt{50}} + \frac{4}{\sqrt{50}}x + \frac{5}{\sqrt{50}}x^2, \frac{3}{\sqrt{50}} + \frac{4}{\sqrt{50}}x - \frac{5}{\sqrt{50}}x^2, \frac{4}{5} - \frac{3}{5}x + 0x^2 \right\}.$$

9.5. Orthogonal Projections

The Gram-Schmidt process is an example of using orthogonal projections. The geometric interpretation of

$$\mathbf{v}_2 = \mathbf{u}_2 - \frac{\langle \mathbf{v}_1, \mathbf{u}_2 \rangle}{\|\mathbf{v}_1\|^2}\mathbf{v}_1$$

is that we have removed from $\mathbf{u}_2$ its orthogonal projection on $\mathbf{v}_1$; in other words, we have removed from $\mathbf{u}_2$ the component of $\mathbf{u}_2$ which is "parallel" to $\mathbf{v}_1$, so that the remaining part must be "perpendicular" to $\mathbf{v}_1$.

It is natural to consider the following question. Suppose that V is a finite-dimensional real inner product space, and that W is a subspace of V. Given any vector $\mathbf{u} \in V$, can we write

$$\mathbf{u} = \mathbf{w} + \mathbf{p},$$

where $\mathbf{w} \in W$ and $\mathbf{p} \in W^\perp$? If so, is this expression unique?

The following result answers these two questions in the affirmative.

PROPOSITION 9L. *Suppose that V is a finite-dimensional real inner product space, and that W is a subspace of V. Suppose further that $\{\mathbf{v}_1, \ldots, \mathbf{v}_r\}$ is an orthogonal basis of W. Then for any vector $\mathbf{u} \in V$,*

$$\mathbf{w} = \frac{\langle \mathbf{u}, \mathbf{v}_1 \rangle}{\|\mathbf{v}_1\|^2}\mathbf{v}_1 + \ldots + \frac{\langle \mathbf{u}, \mathbf{v}_r \rangle}{\|\mathbf{v}_r\|^2}\mathbf{v}_r$$

is the unique vector satisfying $\mathbf{w} \in W$ and $\mathbf{u} - \mathbf{w} \in W^\perp$.

PROOF. Note that the orthogonal basis $\{\mathbf{v}_1, \ldots, \mathbf{v}_r\}$ of W can be extended to a basis

$$\{\mathbf{v}_1, \ldots, \mathbf{v}_r, \mathbf{u}_{r+1}, \ldots, \mathbf{u}_n\}$$

of V which can then be orthogonalized by the Gram-Schmidt process to an orthogonal basis

$$\{\mathbf{v}_1, \ldots, \mathbf{v}_r, \mathbf{v}_{r+1}, \ldots, \mathbf{v}_n\}$$

of V. Clearly $\mathbf{v}_{r+1}, \ldots, \mathbf{v}_n \in W^\perp$. Suppose now that $\mathbf{u} \in V$. Then $\mathbf{u}$ can be expressed as a linear combination of $\mathbf{v}_1, \ldots, \mathbf{v}_n$ in a unique way. By Proposition 9H, this unique expression is given by

$$\mathbf{u} = \frac{\langle \mathbf{u}, \mathbf{v}_1 \rangle}{\|\mathbf{v}_1\|^2}\mathbf{v}_1 + \ldots + \frac{\langle \mathbf{u}, \mathbf{v}_n \rangle}{\|\mathbf{v}_n\|^2}\mathbf{v}_n = \mathbf{w} + \frac{\langle \mathbf{u}, \mathbf{v}_{r+1} \rangle}{\|\mathbf{v}_{r+1}\|^2}\mathbf{v}_{r+1} + \ldots + \frac{\langle \mathbf{u}, \mathbf{v}_n \rangle}{\|\mathbf{v}_n\|^2}\mathbf{v}_n.$$

Clearly $\mathbf{u} - \mathbf{w} \in W^\perp$. ◯

DEFINITION. The vector $\mathbf{w}$ in Proposition 9L is called the orthogonal projection of $\mathbf{u}$ on the subspace W, and denoted by $\text{proj}_W\mathbf{u}$. The vector $\mathbf{p} = \mathbf{u} - \mathbf{w}$ is called the component of $\mathbf{u}$ orthogonal to the subspace W.

EXAMPLE 9.5.1. Recall Example 9.4.3. Consider the subspace $W = \text{span}\{\mathbf{u}_1, \mathbf{u}_2\}$. Note that $\mathbf{v}_1$ and $\mathbf{v}_2$ can each be expressed as a linear combination of $\mathbf{u}_1$ and $\mathbf{u}_2$, and that $\mathbf{u}_1$ and $\mathbf{u}_2$ can each be expressed as a linear combination of $\mathbf{v}_1$ and $\mathbf{v}_2$. It follows that $\{\mathbf{v}_1, \mathbf{v}_2\}$ is an orthogonal basis of W. This basis can be extended to an orthogonal basis $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3, \mathbf{v}_4\}$ of $\mathbb{R}^4$. It follows that $W^\perp = \text{span}\{\mathbf{v}_3, \mathbf{v}_4\}$.

Problems for Chapter 9

1. In each of the following, determine whether $\langle\ ,\ \rangle$ is an inner product in the given vector space by checking whether conditions (IP1)–(IP4) hold:
   a) $\mathbb{R}^2$; $\langle \mathbf{u}, \mathbf{v} \rangle = 2u_1v_1 - u_2v_2$
   b) $\mathbb{R}^2$; $\langle \mathbf{u}, \mathbf{v} \rangle = u_1v_1 + 2u_1v_2 + u_2v_2$
   c) $\mathbb{R}^3$; $\langle \mathbf{u}, \mathbf{v} \rangle = u_1^2v_1^2 + u_2^2v_2^2 + u_3^2v_3^2$

2. Consider the vector space $\mathbb{R}^2$. Suppose that $\langle\ ,\ \rangle$ is the inner product generated by the matrix

$$A = \begin{pmatrix} 2 & 1 \\ 2 & 3 \end{pmatrix}.$$

   Evaluate each of the following:
   a) $\langle (1,2), (2,3) \rangle$
   b) $\|(1,2)\|$
   c) $d((1,2),(2,3))$

3. Suppose that the vectors $\mathbf{u}, \mathbf{v}, \mathbf{w}$ in an inner product space V satisfy $\langle \mathbf{u}, \mathbf{v} \rangle = 2$, $\langle \mathbf{v}, \mathbf{w} \rangle = -3$, $\langle \mathbf{u}, \mathbf{w} \rangle = 5$, $\|\mathbf{u}\| = 1$, $\|\mathbf{v}\| = 2$ and $\|\mathbf{w}\| = 7$. Evaluate each of the following:
   a) $\langle \mathbf{u}+\mathbf{v}, \mathbf{v}+\mathbf{w} \rangle$
   b) $\langle 2\mathbf{v}-\mathbf{w}, 3\mathbf{u}+2\mathbf{w} \rangle$
   c) $\langle \mathbf{u}-\mathbf{v}-2\mathbf{w}, 4\mathbf{u}+\mathbf{v} \rangle$
   d) $\|\mathbf{u}+\mathbf{v}\|$
   e) $\|2\mathbf{w}-\mathbf{v}\|$
   f) $\|\mathbf{u}-2\mathbf{v}+4\mathbf{w}\|$

4. Suppose that $\mathbf{u}$ and $\mathbf{v}$ are two non-zero vectors in the real vector space $\mathbb{R}^2$. Follow the steps below to establish the existence of a real inner product $\langle\ ,\ \rangle$ on $\mathbb{R}^2$ such that $\langle \mathbf{u}, \mathbf{v} \rangle \neq 0$.
   a) Explain, in terms of the euclidean inner product, why we may restrict our discussion to vectors of the form $\mathbf{u} = (x,y)$ and $\mathbf{v} = (ky, -kx)$, where $x, y, k \in \mathbb{R}$ satisfy $(x,y) \neq (0,0)$ and $k \neq 0$.
   b) Explain next why we may further restrict our discussion to vectors of the form $\mathbf{u} = (x,y)$ and $\mathbf{v} = (y, -x)$, where $x, y \in \mathbb{R}$ satisfy $(x,y) \neq (0,0)$.
   c) Let $\mathbf{u} = (x,y)$ and $\mathbf{v} = (y,-x)$, where $x, y \in \mathbb{R}$ and $(x,y) \neq (0,0)$. Consider the inner product on $\mathbb{R}^2$ generated by the real matrix

$$A = \begin{pmatrix} a & b \\ b & c \end{pmatrix},$$

   where $ac \neq b^2$. Show that $\langle \mathbf{u}, \mathbf{v} \rangle = (a^2 - c^2)xy + b(a+c)(y^2 - x^2)$.
   d) Suppose that $x^2 = y^2$. Show that the choice $a > c > b = 0$ will imply $\langle \mathbf{u}, \mathbf{v} \rangle \neq 0$.
   e) Suppose that $x^2 \neq y^2$. Show that the choice $c = a > b > 0$ will imply $\langle \mathbf{u}, \mathbf{v} \rangle \neq 0$.

5. Consider the real vector space $\mathbb{R}^2$.
   a) Find two distinct non-zero vectors $\mathbf{u}, \mathbf{v} \in \mathbb{R}^2$ such that $\langle \mathbf{u}, \mathbf{v} \rangle = 0$ for every weighted euclidean inner product on $\mathbb{R}^2$.
   b) Find two distinct non-zero vectors $\mathbf{u}, \mathbf{v} \in \mathbb{R}^2$ such that $\langle \mathbf{u}, \mathbf{v} \rangle \neq 0$ for any inner product on $\mathbb{R}^2$.

6. For each of the following inner product spaces and subspaces W, find $W^\perp$:
   a) $\mathbb{R}^2$ (euclidean inner product); $W = \{(x,y) \in \mathbb{R}^2 : x + 2y = 0\}$.
   b) $\mathcal{M}_{2,2}(\mathbb{R})$ (inner product discussed in Section 9.2);

$$W = \left\{ \begin{pmatrix} ta & 0 \\ 0 & tb \end{pmatrix} : t \in \mathbb{R} \right\},$$

   where a and b are non-zero.

7. Suppose that $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ is a basis for a real inner product space V. Does there exist $\mathbf{v} \in V$ which is orthogonal to every vector in this basis?

8. Use the Cauchy-Schwarz inequality to prove that $(a\cos\theta + b\sin\theta)^2 \leq a^2 + b^2$ for every $a, b, \theta \in \mathbb{R}$.
   [Hint: First find a suitable real inner product space.]

9. Prove Proposition 9F.

10. Show that $\langle \mathbf{u}, \mathbf{v} \rangle = \frac{1}{4}\|\mathbf{u}+\mathbf{v}\|^2 - \frac{1}{4}\|\mathbf{u}-\mathbf{v}\|^2$ for any $\mathbf{u}$ and $\mathbf{v}$ in a real inner product space.

11. Suppose that $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ is an orthonormal basis of a real inner product space V. Show that for every $\mathbf{u} \in V$, we have $\|\mathbf{u}\|^2 = \langle \mathbf{u}, \mathbf{v}_1 \rangle^2 + \ldots + \langle \mathbf{u}, \mathbf{v}_n \rangle^2$.

12. Show that if $\mathbf{v}_1, \ldots, \mathbf{v}_n$ are pairwise orthogonal in a real inner product space V, then
$$\|\mathbf{v}_1 + \ldots + \mathbf{v}_n\|^2 = \|\mathbf{v}_1\|^2 + \ldots + \|\mathbf{v}_n\|^2.$$

13. Show that $\mathbf{v}_1 = (2, -2, 1)$, $\mathbf{v}_2 = (2, 1, -2)$ and $\mathbf{v}_3 = (1, 2, 2)$ form an orthogonal basis of $\mathbb{R}^3$ under the euclidean inner product. Then write $\mathbf{u} = (-1, 0, 2)$ as a linear combination of $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$.

14. Let $\mathbf{u}_1 = (2, 2, -1)$, $\mathbf{u}_2 = (4, 1, 1)$ and $\mathbf{u}_3 = (1, 10, -5)$. Show that $\{\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3\}$ is a basis of $\mathbb{R}^3$, and apply the Gram-Schmidt process to this basis to find an orthonormal basis of $\mathbb{R}^3$.

15. Show that the vectors $\mathbf{u}_1 = (0, 2, 1, 0)$, $\mathbf{u}_2 = (1, -1, 0, 0)$, $\mathbf{u}_3 = (1, 2, 0, -1)$ and $\mathbf{u}_4 = (1, 0, 0, 1)$ form a basis of $\mathbb{R}^4$. Then apply the Gram-Schmidt process to find an orthogonal basis of $\mathbb{R}^4$. Find also the corresponding orthonormal basis of $\mathbb{R}^4$.

16. Consider the vector space P_2 with the inner product
$$\langle p, q \rangle = \int_0^1 p(x)q(x)\,\mathrm{d}x.$$
Apply the Gram-Schmidt process to the basis $\{1, x, x^2\}$ to find an orthogonal basis of P_2. Find also the corresponding orthonormal basis of P_2.

17. Suppose that we apply the Gram-Schmidt process to non-zero vectors $\mathbf{u}_1, \ldots, \mathbf{u}_n$ without first checking that these form a basis of the inner product space, and obtain $\mathbf{v}_s = \mathbf{0}$ for some $s = 1, \ldots, n$. What conclusion can we draw concerning the collection $\{\mathbf{u}_1, \ldots, \mathbf{u}_n\}$?