Lecture 9. Orthogonal Bases
November 2022

Orthogonal Sets and the Expansion Theorem

DEFINITION 9.1 We say that two vectors $\mathbf{x}$ and $\mathbf{y}$ in a vector space V are **orthogonal** if $\langle \mathbf{x}, \mathbf{y} \rangle = 0$. More generally, a set $\{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_k\}$ of vectors in V is called an **orthogonal set** if

$$\forall i \neq j \quad \langle \mathbf{x}_i, \mathbf{x}_j \rangle = 0 \qquad \text{and} \qquad \forall i \quad \mathbf{x}_i \neq \mathbf{0}$$

A set $\{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_k\}$ of vectors in V is called **orthonormal** if it is orthogonal and, in addition, each $\mathbf{x}_i$ is a unit vector:

$$\forall i \neq j \quad \langle \mathbf{x}_i, \mathbf{x}_j \rangle = 0 \qquad \text{and} \qquad \forall i \quad \|\mathbf{x}_i\| = 1$$

EXAMPLE 9.2 The standard basis $\{\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n\}$ is an orthonormal set in $\mathbb{R}^n$.

DEFINITION 9.3 If $\{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_k\}$ is an orthogonal set, then

$$\left\{ \frac{1}{\|\mathbf{x}_1\|}\mathbf{x}_1, \frac{1}{\|\mathbf{x}_2\|}\mathbf{x}_2, \ldots, \frac{1}{\|\mathbf{x}_k\|}\mathbf{x}_k \right\}$$

is an orthonormal set, and we say that it is the result of **normalizing** the orthogonal set $\{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_k\}$.

The most important result about orthogonality is Pythagoras' theorem. Given orthogonal vectors $\mathbf{v}$ and $\mathbf{w}$ in $\mathbb{R}^3$, it asserts that $\|\mathbf{v} + \mathbf{w}\|^2 = \|\mathbf{v}\|^2 + \|\mathbf{w}\|^2$. In this form the result holds for any orthogonal set in a vector space V.

THEOREM 9.4 (Pythagoras' Theorem) If $\{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_k\}$ is an orthogonal set in V, then

$$\|\mathbf{x}_1 + \mathbf{x}_2 + \cdots + \mathbf{x}_k\|^2 = \|\mathbf{x}_1\|^2 + \|\mathbf{x}_2\|^2 + \cdots + \|\mathbf{x}_k\|^2$$

Proof. The fact that $\langle \mathbf{x}_i, \mathbf{x}_j \rangle = 0$ whenever $i \neq j$ gives

$$\|\mathbf{x}_1 + \mathbf{x}_2 + \cdots + \mathbf{x}_k\|^2 = \langle \mathbf{x}_1 + \mathbf{x}_2 + \cdots + \mathbf{x}_k, \mathbf{x}_1 + \mathbf{x}_2 + \cdots + \mathbf{x}_k \rangle$$
$$= \langle \mathbf{x}_1, \mathbf{x}_1 \rangle + \langle \mathbf{x}_2, \mathbf{x}_2 \rangle + \cdots + \langle \mathbf{x}_k, \mathbf{x}_k \rangle + \sum_{i \neq j} \langle \mathbf{x}_i, \mathbf{x}_j \rangle = \|\mathbf{x}_1\|^2 + \|\mathbf{x}_2\|^2 + \cdots + \|\mathbf{x}_k\|^2 + 0.$$

This is what we wanted. $\square$

If $\mathbf{v}$ and $\mathbf{w}$ are orthogonal, nonzero vectors in $\mathbb{R}^3$, then they are certainly not parallel, and so are linearly independent. The next theorem gives a far-reaching extension of this observation.

THEOREM 9.5 Every orthogonal set in a vector space V is linearly independent.

Proof. Let $\{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_k\}$ be an orthogonal set in V and suppose a linear combination vanishes:

$$t_1\mathbf{x}_1 + t_2\mathbf{x}_2 + \cdots + t_k\mathbf{x}_k = \mathbf{0}.$$

Then

$$0 = \langle \mathbf{x}_1, \mathbf{0} \rangle = \langle \mathbf{x}_1, t_1\mathbf{x}_1 + t_2\mathbf{x}_2 + \cdots + t_k\mathbf{x}_k \rangle$$
$$= t_1\langle \mathbf{x}_1, \mathbf{x}_1 \rangle + t_2\langle \mathbf{x}_1, \mathbf{x}_2 \rangle + \cdots + t_k\langle \mathbf{x}_1, \mathbf{x}_k \rangle = t_1\|\mathbf{x}_1\|^2$$

Since $\|\mathbf{x}_1\|^2 \neq 0$, this implies that $t_1 = 0$. Similarly $t_i = 0$ for each i. □

Theorem 9.5 suggests considering orthogonal bases for a vector space V, that is orthogonal sets that span V. These turn out to be the best bases in the sense that, when expanding a vector as a linear combination of the basis vectors, there are explicit formulas for the coefficients.

THEOREM 9.6 (Expansion Theorem) Let $\{\mathbf{f}_1, \mathbf{f}_2, \ldots, \mathbf{f}_m\}$ be an orthogonal basis of an inner product vector space V. If $\mathbf{x}$ is any vector in V, we have

$$\mathbf{x} = \frac{\langle \mathbf{x}, \mathbf{f}_1 \rangle}{\|\mathbf{f}_1\|^2}\mathbf{f}_1 + \frac{\langle \mathbf{x}, \mathbf{f}_2 \rangle}{\|\mathbf{f}_2\|^2}\mathbf{f}_2 + \cdots + \frac{\langle \mathbf{x}, \mathbf{f}_m \rangle}{\|\mathbf{f}_m\|^2}\mathbf{f}_m$$

Proof. Since $\{\mathbf{f}_1, \mathbf{f}_2, \ldots, \mathbf{f}_m\}$ spans V, we have $\mathbf{x} = t_1\mathbf{f}_1 + t_2\mathbf{f}_2 + \cdots + t_m\mathbf{f}_m$ where the t_i are scalars. To find t_1 we take the dot product of both sides with $\mathbf{f}_1$:

$$\langle \mathbf{x}, \mathbf{f}_1 \rangle = \langle t_1\mathbf{f}_1 + t_2\mathbf{f}_2 + \cdots + t_m\mathbf{f}_m, \mathbf{f}_1 \rangle = t_1\langle \mathbf{f}_1, \mathbf{f}_1 \rangle + t_2\langle \mathbf{f}_2, \mathbf{f}_1 \rangle + \cdots + t_m\langle \mathbf{f}_m, \mathbf{f}_1 \rangle = t_1\|\mathbf{f}_1\|^2$$

Since $\mathbf{f}_1 \neq \mathbf{0}$, this gives $t_1 = \dfrac{\langle \mathbf{x}, \mathbf{f}_1 \rangle}{\|\mathbf{f}_1\|^2}$. Similarly, $t_i = \dfrac{\langle \mathbf{x}, \mathbf{f}_i \rangle}{\|\mathbf{f}_i\|^2}$ for each i. □

LEMMA 9.7 (Orthogonal Lemma) Let $\{\mathbf{f}_1, \mathbf{f}_2, \ldots, \mathbf{f}_m\}$ be an orthogonal set in an inner product space V. If $\mathbf{x} \in \mathbf{V}$, but $\mathbf{x} \notin \operatorname{span}\{\mathbf{f_1}, \mathbf{f_2}, \ldots, \mathbf{f_m}\}$ and

$$\mathbf{f}_{m+1} = \mathbf{x} - \frac{\langle \mathbf{x}, \mathbf{f}_1 \rangle}{\|\mathbf{f}_1\|^2}\mathbf{f}_1 - \frac{\langle \mathbf{x}, \mathbf{f}_2 \rangle}{\|\mathbf{f}_2\|^2}\mathbf{f}_2 - \cdots - \frac{\langle \mathbf{x}, \mathbf{f}_m \rangle}{\|\mathbf{f}_m\|^2}\mathbf{f}_m$$

Then $\{\mathbf{f}_1, \ldots, \mathbf{f}_m, \mathbf{f}_{m+1}\}$ is an orthogonal set.

Proof. For convenience, write $t_i = \dfrac{\langle \mathbf{x}, \mathbf{f}_i \rangle}{\|\mathbf{f}_i\|^2}$ for each i. Given $1 \leqslant k \leqslant m$:

$$\langle \mathbf{f}_{m+1}, \mathbf{f}_k \rangle = \langle \mathbf{x} - t_1\mathbf{f}_1 - \cdots - t_k\mathbf{f}_k - \cdots - t_m\mathbf{f}_m, \mathbf{f}_k \rangle$$
$$= \langle \mathbf{x}, \mathbf{f}_k \rangle - t_1\langle \mathbf{f}_1, \mathbf{f}_k \rangle - \cdots - t_k\langle \mathbf{f}_k, \mathbf{f}_k \rangle - \cdots - t_m\langle \mathbf{f}_m, \mathbf{f}_k \rangle = \langle \mathbf{x}, \mathbf{f}_k \rangle - t_k\|\mathbf{f}_k\|^2 = 0$$

Obviously, $\mathbf{f}_{m+1} \neq \mathbf{0}$ if $\mathbf{x}$ is not in $\operatorname{span}\{\mathbf{f}_1, \mathbf{f}_2, \ldots, \mathbf{f}_m\}$. □

The orthogonal lemma has two important consequences. The first is an extension for orthogonal sets of the fundamental fact that any independent set is part of a basis.

THEOREM 9.8 Let V be a finite-dimensional inner vector space.

1. Every orthogonal subset $\{\mathbf{f}_1, \mathbf{f}_2, \ldots, \mathbf{f}_m\}$ in V is a subset of an orthogonal basis of V.
2. V has an orthogonal basis.

Proof. 1 If $\operatorname{span}\{\mathbf{f}_1, \mathbf{f}_2, \ldots, \mathbf{f}_m\} = V$, it is already a basis. Otherwise, there exists $\mathbf{x}$ in V outside $\operatorname{span}\{\mathbf{f}_1, \mathbf{f}_2, \ldots, \mathbf{f}_m\}$. If $\mathbf{f}_{m+1}$ is as given in the orthogonal lemma, then $\mathbf{f}_{m+1}$ is in V and $\{\mathbf{f}_1, \ldots, \mathbf{f}_m, \mathbf{f}_{m+1}\}$ is orthogonal. If $\operatorname{span}\{\mathbf{f}_1, \ldots, \mathbf{f}_m, \mathbf{f}_{m+1}\} = V$, we are done. Otherwise, the process continues to create larger and larger orthogonal subsets of V. They are all independent by Theorem 9.5, so we have a basis when we reach a subset containing $\dim V$ vectors.

2 If $V = \{\mathbf{0}\}$, the empty basis is orthogonal. Otherwise, if $\mathbf{f} \neq \mathbf{0}$ is in V, then $\{\mathbf{f}\}$ is orthogonal, so (2) follows from (1). □

The second consequence of the orthogonal lemma is a procedure by which any basis $\{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_m\}$ of a V can be systematically modified to yield an orthogonal basis $\{\mathbf{f}_1, \mathbf{f}_2, \ldots, \mathbf{f}_m\}$ of V. The algorithm converts any basis of V itself into an orthogonal basis.

THEOREM 9.9 (Gram–Schmidt Orthogonalization Algorithm)

If $\{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_m\}$ is any basis of a finite-dimensional inner vector space V, construct $\mathbf{f}_1, \mathbf{f}_2, \ldots, \mathbf{f}_m$ in V successively as follows:

$$\mathbf{f}_1 = \mathbf{x}_1, \qquad \mathbf{f}_2 = \mathbf{x}_2 - \frac{\langle \mathbf{x}_2, \mathbf{f}_1 \rangle}{\|\mathbf{f}_1\|^2}\mathbf{f}_1, \qquad \mathbf{f}_3 = \mathbf{x}_3 - \frac{\langle \mathbf{x}_3, \mathbf{f}_1 \rangle}{\|\mathbf{f}_1\|^2}\mathbf{f}_1 - \frac{\langle \mathbf{x}_3, \mathbf{f}_2 \rangle}{\|\mathbf{f}_2\|^2}\mathbf{f}_2, \ldots$$
$$\mathbf{f}_m = \mathbf{x}_m - \frac{\langle \mathbf{x}_m, \mathbf{f}_1 \rangle}{\|\mathbf{f}_1\|^2}\mathbf{f}_1 - \cdots - \frac{\langle \mathbf{x}_m, \mathbf{f}_{k-1} \rangle}{\|\mathbf{f}_{m-1}\|^2}\mathbf{f}_{m-1}.$$

Then

1. $\{\mathbf{f}_1, \mathbf{f}_2, \ldots, \mathbf{f}_m\}$ is an orthogonal basis of V.
2. $\operatorname{span}\{\mathbf{f}_1, \mathbf{f}_2, \ldots, \mathbf{f}_k\} = \operatorname{span}\{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_k\}$ for each $k = 1, 2, \ldots, m$.

Proof. To start the process, take $\mathbf{f}_1 = \mathbf{x}_1$. Then $\mathbf{x}_2$ is not in $\operatorname{span}\{\mathbf{f}_1\}$ because $\{\mathbf{x}_1, \mathbf{x}_2\}$ is independent, so take

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

Thus $\{\mathbf{f}_1, \mathbf{f}_2\}$ is orthogonal by Orthogonal Lemma. Moreover, $\operatorname{span}\{\mathbf{f}_1, \mathbf{f}_2\} = \operatorname{span}\{\mathbf{x}_1, \mathbf{x}_2\}$, so $\mathbf{x}_3$ is not in $\operatorname{span}\{\mathbf{f}_1, \mathbf{f}_2\}$. Hence $\{\mathbf{f}_1, \mathbf{f}_2, \mathbf{f}_3\}$ is orthogonal where

$$\mathbf{f}_3 = \mathbf{x}_3 - \frac{\langle \mathbf{x}_3, \mathbf{f}_1 \rangle}{\|\mathbf{f}_1\|^2}\mathbf{f}_1 - \frac{\langle \mathbf{x}_3, \mathbf{f}_2 \rangle}{\|\mathbf{f}_2\|^2}\mathbf{f}_2.$$

Again, $\operatorname{span}\{\mathbf{f}_1, \mathbf{f}_2, \mathbf{f}_3\} = \operatorname{span}\{\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3\}$, so $\mathbf{x}_4$ is not in $\operatorname{span}\{\mathbf{f}_1, \mathbf{f}_2, \mathbf{f}_3\}$ and the process continues. At the mth iteration we construct an orthogonal set $\{\mathbf{f}_1, \mathbf{f}_2, \ldots, \mathbf{f}_m\}$ such that

$$\operatorname{span}\{\mathbf{f}_1, \mathbf{f}_2, \ldots, \mathbf{f}_m\} = \operatorname{span}\{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_m\} = V.$$

Hence $\{\mathbf{f}_1, \mathbf{f}_2, \ldots, \mathbf{f}_m\}$ is the desired orthogonal basis of V. □