

## LECTURE 6

# Inner product spaces (2)

### 6.1 Orthogonal bases

Let  $V$  be an inner product space and let  $B$  denote a basis for  $V$ . Choosing the elements of  $B$  as orthogonal vectors has many benefits and is common in many applications. We consider such a choice below, including some examples.

**Definition 6.1 — Orthogonal set.** Let  $(V, \langle \cdot, \cdot \rangle)$  be an inner product space. Let  $S = \{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k\}$  be a set of nonzero vectors in  $V$ . Then  $S$  is said to be orthogonal if for all  $i \neq j$

$$\langle \mathbf{v}_i, \mathbf{v}_j \rangle = 0.$$

Note that, by definition, orthogonal sets do not include the zero vector. The advantage of working with such sets can be derived immediately.

**Proposition 6.1** Orthogonal sets are linearly independent.

*Proof.* Let  $S = \{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k\}$  be an orthogonal set of nonzero vectors in an inner product space  $(V, \langle \cdot, \cdot \rangle)$ . Consider the linear combination

$$\mathbf{0} = a_1 \mathbf{v}_1 + a_2 \mathbf{v}_2 + \cdots + a_k \mathbf{v}_k.$$

Then for any  $j = 1, 2, \dots, k$ ,

$$0 = \langle \mathbf{0}, \mathbf{v}_j \rangle = \langle a_1 \mathbf{v}_1 + a_2 \mathbf{v}_2 + \cdots + a_k \mathbf{v}_k, \mathbf{v}_j \rangle = a_j \|\mathbf{v}_j\|^2 \implies a_j = 0.$$

Hence, the zero vector can only be written as the trivial linear combination of vectors in  $S$  and therefore  $S$  is linearly independent. ■

Let now  $V$  be an inner product space with dimension  $\dim V = n$ . Choosing a set of  $n$  orthogonal vectors results in a basis for  $V$ .

**Definition 6.2 — Orthogonal basis.** Let  $(V, \langle \cdot, \cdot \rangle)$  be an inner product space. A basis set is said to be an orthogonal basis for  $V$  if it is an orthogonal set.

We immediately derive the following property of orthogonal sets.

**Proposition 6.2** Let  $(V, \langle \cdot, \cdot \rangle)$  be an inner product space, with  $\dim V = n$ . Let  $S$  be an orthogonal set of vectors from  $V$  with  $|S| = n$ . Then  $S$  is an orthogonal basis for  $V$ .

*Proof.* By Proposition 6.1,  $S$  is linearly independent. The result follows by applying Proposition 4.7. ■

Orthogonal bases allow for the coordinates of a generic vector  $\mathbf{v} \in V$  to be computed via evaluations of inner products. This is a key advantage over computing the coordinates in the usual way, by solving a (possibly large) linear system of equations.

**Proposition 6.3** Let  $(V, \langle \cdot, \cdot \rangle)$  be an inner product space. Let  $B = \{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n\}$  denote an orthogonal basis for  $V$ . Then the coordinates  $a_i$  of any vector  $\mathbf{v} \in V$  are given by

$$a_i = \frac{\langle \mathbf{v}, \mathbf{v}_i \rangle}{\langle \mathbf{v}_i, \mathbf{v}_i \rangle}.$$

*Proof.* Let  $\mathbf{v} = a_1\mathbf{v}_1 + a_2\mathbf{v}_2 + \cdots + a_n\mathbf{v}_n$ . Then, using orthogonality,

$$\langle \mathbf{v}, \mathbf{v}_i \rangle = \langle a_1\mathbf{v}_1 + a_2\mathbf{v}_2 + \cdots + a_n\mathbf{v}_n, \mathbf{v}_i \rangle = a_i \langle \mathbf{v}_i, \mathbf{v}_i \rangle,$$

and the result follows. ■



The coordinates  $a_i$  in Proposition 6.3 are known as the **Fourier coefficients** of  $\mathbf{v}$  in the basis  $B$ .

When the elements of an orthogonal set are unit vectors, the manipulations and expressions arising are further simplified. In this case, we use the convention that the generic  $i$ th unit vector is denoted by  $\mathbf{e}_i$ .

**Definition 6.3 — Orthonormal set.** Let  $(V, \langle \cdot, \cdot \rangle)$  be an inner product space. A set  $S$  of nonzero vectors  $\mathbf{e}_i$  in  $V$  is said to be orthonormal if

$$\langle \mathbf{e}_i, \mathbf{e}_j \rangle = \delta_{ij} := \begin{cases} 0 & i \neq j, \\ 1 & i = j, \end{cases}$$

for all  $i, j = 1, 2, \dots, |S|$ .

Correspondingly, we have the following definition.

**Definition 6.4 — Orthonormal basis.** Let  $(V, \langle \cdot, \cdot \rangle)$  be an inner product space. A set is said to be an orthonormal basis for  $V$  if

- it is an orthonormal set in  $V$ ;
- it is a basis set for  $V$ .

The representation of the Fourier coefficients of a vector  $\mathbf{v}$  is further simplified when working with an orthonormal basis:

$$a_i = \langle \mathbf{v}, \mathbf{e}_i \rangle.$$

In the context of coordinates, there is one more significant result concerning the evaluation of inner products.

**Proposition 6.4** Let  $(V, \langle \cdot, \cdot \rangle)$  be an inner product space with dimension  $n$ . Let  $B = \{\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n\}$  denote an orthonormal basis for  $V$ . Let  $\mathbf{u}, \mathbf{v}$  have respective coordinates  $a_i, b_j$ , with  $i, j = 1, \dots, n$ . Then

$$\langle \mathbf{u}, \mathbf{v} \rangle = a_1b_1 + a_2b_2 + \cdots + a_nb_n.$$

*Proof.* We have, using orthogonality,

$$\langle \mathbf{u}, \mathbf{v} \rangle = \langle a_1\mathbf{e}_1 + a_2\mathbf{e}_2 + \cdots + a_n\mathbf{e}_n, b_1\mathbf{e}_1 + b_2\mathbf{e}_2 + \cdots + b_n\mathbf{e}_n \rangle = a_1b_1 + a_2b_2 + \cdots + a_nb_n.$$

■



The result of Proposition 6.4 is related to one of our previous observations, namely, that there is a one-to-one correspondence between any finite-dimensional vector space  $V$  and  $\mathbb{R}^n$ . In this case we witness the correspondence between an inner product of two vectors in  $V$  and the Euclidean inner product of their respective coordinates with respect to some orthonormal basis. This correspondence will be discussed in Part II.

## 6.2 Projections

The concept of orthogonality in a vector space allows for the extension of the geometric concept of orthogonal projection. In turn, orthogonal projections will allow us to decompose a vector into a sum of projections: a so-called orthogonal decomposition. These results will enable us to devise a procedure for constructing an orthonormal basis for any inner product space in the next lecture.

Let us start by aiming to write a vector as a sum of two vectors: one parallel to a fixed direction  $\mathbf{u}$  and the other orthogonal to it. We have

$$\mathbf{v} = \mathbf{v}^{\parallel} + \mathbf{v}^{\perp},$$

where  $\mathbf{v}^{\parallel} = a\mathbf{u}$  and  $\langle \mathbf{v}^{\perp}, \mathbf{v}^{\parallel} \rangle = 0$ . Note that if we identify  $a$ , then the component  $\mathbf{v}^{\parallel}$  is known and so is  $\mathbf{v}^{\perp}$ . Taking inner products with  $\mathbf{u}$  on both sides we get

$$\langle \mathbf{v}, \mathbf{u} \rangle = \langle \mathbf{v}^{\parallel} + \mathbf{v}^{\perp}, \mathbf{u} \rangle = a \langle \mathbf{u}, \mathbf{u} \rangle + \frac{1}{a} \langle \mathbf{v}^{\perp}, \mathbf{v}^{\parallel} \rangle = a \langle \mathbf{u}, \mathbf{u} \rangle + 0 \implies a = \frac{\langle \mathbf{v}, \mathbf{u} \rangle}{\langle \mathbf{u}, \mathbf{u} \rangle}.$$

Thus, there exists a unique choice of  $a$  for which this decomposition of a generic vector  $\mathbf{v}$  holds. The resulting component  $\mathbf{v}^{\parallel}$  is known as the orthogonal projection of  $\mathbf{v}$  onto  $\mathbf{u}$ . We include its form in the following definition.

**Definition 6.5 — Orthogonal projection (onto a vector).** Let  $(V, \langle \cdot, \cdot \rangle)$  be an inner product space and let  $\mathbf{u} \in V \setminus \{\mathbf{0}\}$ . The orthogonal projection of  $\mathbf{v} \in V$  onto  $\mathbf{u}$ , denoted by  $\mathbf{v}_{\mathbf{u}}^{\parallel}$ , is the vector

$$\mathbf{v}_{\mathbf{u}}^{\parallel} := \frac{\langle \mathbf{v}, \mathbf{u} \rangle}{\langle \mathbf{u}, \mathbf{u} \rangle} \mathbf{u}.$$

Note that we changed the notation from  $\mathbf{v}^{\parallel}$  to  $\mathbf{v}_{\mathbf{u}}^{\parallel}$ , as we want to emphasise that the orthogonal projection depends on the vector we project on. Using this notation, the decomposition of  $\mathbf{v}$  is

$$\mathbf{v} = \mathbf{v}_{\mathbf{u}}^{\parallel} + \mathbf{v}_{\mathbf{u}}^{\perp},$$

where  $\mathbf{v}_{\mathbf{u}}^{\perp} \perp \mathbf{v}_{\mathbf{u}}^{\parallel}$  is given by

$$\mathbf{v}_{\mathbf{u}}^{\perp} = \mathbf{v} - \frac{\langle \mathbf{v}, \mathbf{u} \rangle}{\langle \mathbf{u}, \mathbf{u} \rangle} \mathbf{u}.$$

Note that we can also write the orthogonal projection using the angle  $\alpha$  between  $\mathbf{v}$  and  $\mathbf{u}$ :

$$\mathbf{v}_{\mathbf{u}}^{\parallel} := \frac{\|\mathbf{v}\| \cos \alpha}{\|\mathbf{u}\|} \mathbf{u} = \|\mathbf{v}\| \cos \alpha \hat{\mathbf{u}}.$$

Definition 6.5 is essentially the statement provided for the Euclidean space  $\mathbb{E}^3$  in Lecture 1, although we note that it allows also for the projection of the vector  $\mathbf{0}$  onto any vector  $\mathbf{u}$ , which is  $\mathbf{0}$ . A further comparison to the Euclidean space is included in the next result: in  $\mathbb{E}^3$ , the segment perpendicular to the direction line of  $\mathbf{u}$  had the least length among all segments drawn between  $\mathbf{v}$  and the direction line of  $\mathbf{u}$ . This is the case also in inner product spaces.

**Proposition 6.5** Let  $(V, \langle \cdot, \cdot \rangle)$  be an inner product space and let  $\mathbf{u} \in V \setminus \{\mathbf{0}\}$ . Let  $\mathbf{v} = \mathbf{v}_{\mathbf{u}}^{\parallel} + \mathbf{v}_{\mathbf{u}}^{\perp}$ , where  $\mathbf{v}_{\mathbf{u}}^{\parallel}$  is the projection of  $\mathbf{v}$  onto  $\mathbf{u}$ . Then

$$\|\mathbf{v}_{\mathbf{u}}^{\perp}\| = \|\mathbf{v} - \mathbf{v}_{\mathbf{u}}^{\parallel}\| \leq \|\mathbf{v} - \mathbf{z}\| \quad \text{for all } \mathbf{z} \in U := \text{span}\{\mathbf{u}\}.$$

*Proof.* Let  $\mathbf{z} = \mathbf{v}_{\mathbf{u}}^{\parallel} + \mathbf{e}$ , for some  $\mathbf{e} \in U$ . Note that we can write  $\mathbf{z}$  in this form since  $\mathbf{v}_{\mathbf{u}}^{\parallel} \in U$  also. With this notation, we note that

$$\mathbf{v} - \mathbf{z} = \mathbf{v} - \mathbf{v}_{\mathbf{u}}^{\parallel} - \mathbf{e} = \mathbf{v}_{\mathbf{u}}^{\perp} - \mathbf{e},$$

where  $\mathbf{v}_{\mathbf{u}}^{\perp} \perp \mathbf{e}$ , since  $\mathbf{e} \in U$  is a multiple of  $\mathbf{u}$ . Therefore, we can apply the Pythagoras theorem:

$$\|\mathbf{v}_{\mathbf{u}}^{\perp} - \mathbf{e}\|^2 = \|\mathbf{v}_{\mathbf{u}}^{\perp}\|^2 + \|\mathbf{e}\|^2.$$

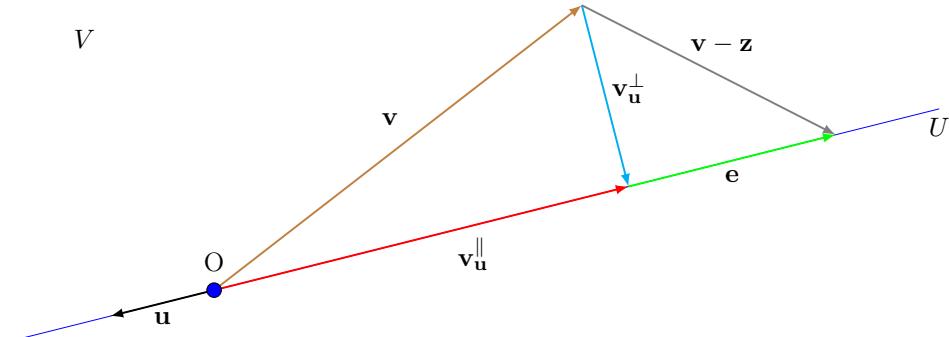
Hence,

$$\|\mathbf{v} - \mathbf{z}\|^2 = \|\mathbf{v}_{\mathbf{u}}^{\perp} - \mathbf{e}\|^2 = \|\mathbf{v}_{\mathbf{u}}^{\perp}\|^2 + \|\mathbf{e}\|^2 \geq \|\mathbf{v}_{\mathbf{u}}^{\perp}\|^2 = \|\mathbf{v} - \mathbf{v}_{\mathbf{u}}^{\parallel}\|^2,$$

with equality holding if and only if  $\mathbf{e} = \mathbf{0}$ , i.e., if and only if  $\mathbf{z} = \mathbf{v}_{\mathbf{u}}^{\parallel}$ . ■



The figure below is included for illustration: the subspace  $U$  is a line in the plane  $V$ . The proposition simply states that the length  $\|\mathbf{v}_{\mathbf{u}}^{\perp}\|$  of the perpendicular segment is shorter than the length  $\|\mathbf{v} - \mathbf{z}\|$  of any other segment drawn from the tip of  $\mathbf{v}$  to the line  $U$ . Note also that the proof highlights at the end the uniqueness of the segment of least length.



Orthogonal projection  $\mathbf{v}_{\mathbf{u}}$  of generic  $\mathbf{v}$  in (the plane)  $V$  on  $\mathbf{u} \in U$ .

In the previous result, the vector  $\mathbf{v}_{\mathbf{u}}^{\perp}$  is orthogonal to  $\mathbf{u}$  and therefore to any multiple of  $\mathbf{u}$ , i.e.,  $\mathbf{v}_{\mathbf{u}}^{\perp}$  is orthogonal to any element in the subspace  $U = \text{span}\{\mathbf{u}\}$  of  $V$ . This suggests a more general definition of orthogonality.

**Definition 6.6 — Vector orthogonal to a subspace.** Let  $(V, \langle \cdot, \cdot \rangle)$  be an inner product space and let  $U$  be a subspace of  $V$ . We say  $\mathbf{w}$  is orthogonal to  $U$  if  $\mathbf{w} \perp \mathbf{u}$  for all  $\mathbf{u} \in U$ . We write  $\mathbf{w} \perp U$ .

**Proposition 6.6** Let  $(V, \langle \cdot, \cdot \rangle)$  be an inner product space and let  $U$  be a subspace of  $V$ . Let  $S$  be a spanning set for  $U$ . Then

$$\mathbf{v} \perp U \iff \mathbf{v} \perp S.$$

*Proof.* Exercise. ■

The choice of spanning set in the previous proposition can be a convenient one, for example, an orthogonal basis. Indeed, we will make this choice in the proof of the following result.

We are now ready to extend the concept of orthogonal projection introduced in Definition 6.5.

**Theorem 6.7** Let  $(V, \langle \cdot, \cdot \rangle)$  be an inner product space and let  $U$  be a subspace of  $V$ . Let  $\mathbf{v} \in V$ . Then there exists a unique vector  $\mathbf{v}_U^{\parallel} \in U$  such that  $\mathbf{v}_U^{\perp} := \mathbf{v} - \mathbf{v}_U^{\parallel} \perp U$ .

*Proof.* Let  $B = \{\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_k\}$  be an orthogonal basis for  $U$ . For any  $\mathbf{v} \in V$ , there exists a vector  $\mathbf{v}_U^{\parallel}$  in  $U$ , given by

$$\mathbf{v}_U^{\parallel} = a_1 \mathbf{u}_1 + a_2 \mathbf{u}_2 + \cdots + a_k \mathbf{u}_k, \quad \text{where } a_i := \frac{\langle \mathbf{v}, \mathbf{u}_i \rangle}{\langle \mathbf{u}_i, \mathbf{u}_i \rangle} \quad (i = 1, \dots, k),$$

such that  $\mathbf{v} - \mathbf{v}_U^{\parallel} = \mathbf{v}_U^{\perp} \perp U$ . To see this, we use Proposition 6.6 and check that  $\mathbf{v}_U^{\perp} \perp \mathbf{u}_j$  for all  $j$ :

$$\langle \mathbf{v}_U^{\perp}, \mathbf{u}_j \rangle = \langle \mathbf{v}, \mathbf{u}_j \rangle - \langle \mathbf{v}_U^{\parallel}, \mathbf{u}_j \rangle = \langle \mathbf{v}, \mathbf{u}_j \rangle - \sum_{i=1}^k a_i \langle \mathbf{u}_i, \mathbf{u}_j \rangle = \langle \mathbf{v}, \mathbf{u}_j \rangle - \frac{\langle \mathbf{v}, \mathbf{u}_j \rangle}{\langle \mathbf{u}_j, \mathbf{u}_j \rangle} \langle \mathbf{u}_j, \mathbf{u}_j \rangle = 0.$$

To prove uniqueness, assume there exists  $\tilde{\mathbf{v}}_U^{\parallel} \in U$  satisfying  $\mathbf{v} - \tilde{\mathbf{v}}_U^{\parallel} \perp U$ . Let  $\tilde{\mathbf{v}}_U^{\parallel} = \mathbf{v}_U^{\parallel} + \mathbf{e}$  for some  $\mathbf{e} \in U$  which we represent in the basis  $B$  as

$$\mathbf{e} = \sum_{i=1}^k c_i \mathbf{u}_i, \quad c_i = \frac{\langle \mathbf{e}, \mathbf{u}_i \rangle}{\langle \mathbf{u}_i, \mathbf{u}_i \rangle}.$$

Under our assumption, we must have for all  $j$

$$0 = \langle \mathbf{v} - \tilde{\mathbf{v}}_U^{\parallel}, \mathbf{u}_j \rangle = \langle \mathbf{v} - \mathbf{v}_U^{\parallel} - \mathbf{e}, \mathbf{u}_j \rangle = 0 - \langle \mathbf{e}, \mathbf{u}_j \rangle = -c_j \|\mathbf{u}_j\|^2 \implies c_j = 0 \implies \mathbf{e} = \mathbf{0}. \quad \blacksquare$$

The previous result suggests the following definition.

**Definition 6.7 — Orthogonal projection (onto a subspace).** Let  $(V, \langle \cdot, \cdot \rangle)$  be an inner product space and let  $U$  be a subspace of  $V$  spanned by an orthogonal basis set  $B = \{\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_k\}$ . Let  $\mathbf{v} \in V$ . The **orthogonal projection of  $\mathbf{v}$  onto  $U$**  is the vector

$$\mathbf{v}_U^{\parallel} = \sum_{i=1}^k \frac{\langle \mathbf{v}, \mathbf{u}_i \rangle}{\langle \mathbf{u}_i, \mathbf{u}_i \rangle} \mathbf{u}_i.$$



Note that by Theorem 6.7, the orthogonal projection onto  $U$  is a unique vector in  $U$ . Note also that if  $\dim U = k = 1$  (i.e.,  $U$  is a 'line'), we recover the expression for an orthogonal projection onto a vector given in Definition 6.5.

### 6.3 Orthogonal decompositions

The result of the previous proposition confirms that we can write any vector  $\mathbf{v}$  in an inner product space as the sum of two orthogonal vectors:  $\mathbf{v}_U^{\parallel}$  in  $U$  and  $\mathbf{v}_U^{\perp}$  in a set disjoint from  $U$ . We make this observation more precise via the following definition.

**Definition 6.8 — Orthogonal complement.** Let  $(V, \langle \cdot, \cdot \rangle)$  be an inner product space and let  $U$  be a subspace of  $V$ . The orthogonal complement of  $U$  in  $V$  is denoted by  $U^{\perp}$  and is defined to be the set of vectors in  $V$  perpendicular to  $U$

$$U^{\perp} := \{\mathbf{v} \in V : \mathbf{v} \perp U\}.$$

This definition implies that  $\mathbf{v}_U^{\perp} \in U^{\perp}$ . One can immediately establish the following properties of  $U^{\perp}$ .

**Proposition 6.8** Let  $(V, \langle \cdot, \cdot \rangle)$  be an inner product space and let  $U$  be a subspace of  $V$ . Then  $U^\perp$  is a subspace of  $V$  and  $U \cap U^\perp = \{\mathbf{0}\}$ .

*Proof.* The first statement follows from the Subspace criterion 1. The second statement follows by contradiction: if  $\mathbf{0} \neq \mathbf{w} \in U \cap U^\perp$ , then we must have  $\mathbf{w} \perp \mathbf{w}$ , i.e.,  $\langle \mathbf{w}, \mathbf{w} \rangle = 0$ , so that  $\mathbf{w} = \mathbf{0}$ . ■

We end with the following edifying result.

**Proposition 6.9** Let  $(V, \langle \cdot, \cdot \rangle)$  be an inner product space and let  $U$  be a subspace of  $V$ . Then  $V = U \oplus U^\perp$ .

*Proof.* The result follows by the Direct sum criterion 1, since  $U$  and  $U^\perp$  are two subspaces of  $V$  satisfying  $V = U + U^\perp$  (why?) and  $U \cap U^\perp = \{\mathbf{0}\}$ . ■

This is an example of an orthogonal decomposition of a vector space.