LECTURE 5

# Inner product spaces (1)

One of the binary operations involving Euclidean vectors is the dot product. This operation induces various geometrical concepts such as angle between vectors, length and orthogonality. In this lecture, we will generalise the concept of dot product to vectors other than Euclidean. Vector spaces equipped with this type of operation will afford additional structure by extending some familiar concepts in geometry.

## 5.1 Inner products

The dot (or scalar) product is an operation between two vectors which returns a scalar. In the case of Euclidean vectors, this scalar is a real number. We recall it below:

$$\mathbf{A} \cdot \mathbf{B} = |\mathbf{A}|\,|\mathbf{B}| \cos\alpha,$$

where $\alpha$ is the non-reflex angle (i.e., $\alpha \in [0, \pi]$) between the oriented direction lines of $\mathbf{A}$ and $\mathbf{B}$.

It is not immediately clear how this definition can be extended to general vector spaces $V$: what is the angle between two vectors in $\mathbb{R}^4$ or that between two polynomials in $\mathcal{P}_n$? Instead of using the above definition in its form, we aim to extract some of the features it encapsulates.

We note the following three properties associated with the dot product:

- commutativity (or symmetry): $\mathbf{A} \cdot \mathbf{B} = \mathbf{B} \cdot \mathbf{A}$;
- non-negativity: $\mathbf{A} \cdot \mathbf{A} \geq 0$, with $\mathbf{A} \cdot \mathbf{A} = 0$ if and only if $\mathbf{A} = \mathbf{0}$;
- linearity: $(a\mathbf{A} + b\mathbf{B}) \cdot \mathbf{C} = a\mathbf{A} \cdot \mathbf{C} + b\mathbf{B} \cdot \mathbf{C}$.

Note that while the first two can be immediately verified, the third is best checked using the representations of $\mathbf{A}, \mathbf{B}, \mathbf{C}$ in the orthonormal basis $\{\mathbf{i}, \mathbf{j}, \mathbf{k}\}$.

We extend the definition of this operation to the case of a real finite-dimensional vector space by using these three properties. Before we give the definition, we note that the dot product can be viewed as a function of two arguments which returns real values:

$$\mathcal{B}(\mathbf{A}, \mathbf{B}) = |\mathbf{A}|\,|\mathbf{B}| \cos\alpha.$$

Moreover, we note that $\mathcal{B}(\cdot, \cdot)$ is linear in each argument. These observations are formalised in the following definition.

**Definition 5.1** Let $U, V, W$ denote vector spaces over a field $\mathbb{F}$. A **bilinear function** is a function of two arguments $\mathcal{B}(\cdot,\cdot) : V \times W \to U$, which is linear in each argument:

- $\mathcal{B}(a\mathbf{u} + b\mathbf{v}, \mathbf{w}) = a\mathcal{B}(\mathbf{u}, \mathbf{w}) + b\mathcal{B}(\mathbf{v}, \mathbf{w})$,
- $\mathcal{B}(\mathbf{u}, a\mathbf{v} + b\mathbf{w}) = a\mathcal{B}(\mathbf{u}, \mathbf{v}) + b\mathcal{B}(\mathbf{u}, \mathbf{w})$.

We say $\mathcal{B}(\cdot,\cdot)$ is a **bilinear form** if $U = \mathbb{F}$.

The special case $W = V$ is commonly used in applications, with the following terminology applying:

- $\mathcal{B}(\cdot,\cdot) : V \times V$ is **symmetric** if $\mathcal{B}(\mathbf{v}, \mathbf{w}) = \mathcal{B}(\mathbf{w}, \mathbf{v})$;
- $\mathcal{B}(\cdot,\cdot) : V \times V$ is **anti-symmetric** if $\mathcal{B}(\mathbf{v}, \mathbf{w}) = -\mathcal{B}(\mathbf{w}, \mathbf{v})$;

Using the above terminology for the choice $V = W = \mathbb{E}^3, U = \mathbb{R}$, one can identify the dot product as a symmetric bilinear form. Note that we replace the generic notation $\mathcal{B}(\cdot,\cdot)$ with $\langle\cdot,\cdot\rangle$.

**Definition 5.2 — Inner product.** Let $V(\mathbb{R})$ be a real vector space. A real inner product is a function $\langle\cdot,\cdot\rangle : V \times V \to \mathbb{R}$ which satisfies the following properties:

i. symmetry: $\langle \mathbf{u}, \mathbf{v}\rangle = \langle \mathbf{v}, \mathbf{u}\rangle$;
ii. linearity: $\langle a\mathbf{u} + b\mathbf{v}, \mathbf{w}\rangle = a\langle \mathbf{u}, \mathbf{w}\rangle + b\langle \mathbf{v}, \mathbf{w}\rangle$;
iii. non-negativity: $\langle \mathbf{v}, \mathbf{v}\rangle \geq 0$, with $\langle \mathbf{v}, \mathbf{v}\rangle = 0$ if and only if $\mathbf{v} = \mathbf{0}$.

Property ii. requires linearity in the first argument of the inner product. However, by symmetry, we have linearity also in the second argument, as the following exercise shows.

**Exercise 5.1** Let $V$ be a real vector space equipped with a real inner product. Show that for any $\mathbf{u}, \mathbf{v}, \mathbf{w} \in V$ and $a, b \in \mathbb{R}$, properties i. and ii. imply that $\langle\cdot,\cdot\rangle$ is linear in its second argument:

$$\langle \mathbf{u}, a\mathbf{v} + b\mathbf{w}\rangle = a\langle \mathbf{u}, \mathbf{v}\rangle + b\langle \mathbf{u}, \mathbf{w}\rangle .$$

R Property iii. is also referred to as **positive-definiteness**. This terminology is also extended to bilinear forms. Thus, inner products are symmetric and positive-definite bilinear forms.

**Example 5.1** We define the **standard/Euclidean inner product** on $\mathbb{R}^n$ via

$$\langle \mathbf{u}, \mathbf{v}\rangle = u_1v_1 + u_2v_2 + \cdots + u_nv_n.$$

It is straightforward to verify the three properties in the definition.

**Example 5.2** We define the inner product on $\mathcal{P}_n([a, b])$ via

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

Properties i. and ii. are straightforward to verify; property iii. is verified using the standard properties of integrals.

R Note that the notation $\mathcal{P}_n([a, b])$ indicates the set of polynomials defined on the closed interval $[a, b]$. This choice ensures that the integral is a real number, which is an implicit requirement in the definition of the inner product.

The previous example indicates that property iii. may be the usual focus when establishing when a bilinear function is an inner product. Indeed, here is an example where it fails to be satisfied.

**Example 5.3** Let $V = \mathcal{P}_n([a,b])$ and define $\mathcal{B} : V \times V \to \mathbb{R}$ via

$$\mathcal{B}(p,q) = \int_a^b p'(x)q'(x)\mathrm{d}x.$$

While $\mathcal{B}$ is a non-negative symmetric bilinear form, it fails to be zero only at the zero vector since the choice $p(x) = 1$ yields

$$\mathcal{B}(p,p) = \int_a^b [p'(x)]^2 \mathrm{d}x = \int_a^b 0\mathrm{d}x = 0.$$

Hence, $\mathcal{B}$ is not an inner product.

**Exercise 5.2** Let $V = C^1([a,b])$ and consider the following candidate for an inner product:

$$\mathcal{B}(f,g) = \int_a^b \left[f(x)g(x) + f'(x)g'(x)\right] \mathrm{d}x.$$

Show that $\mathcal{B}$ is an inner product.

Definition 5.2 is restricted to the case of vector spaces over the reals. Can we have fields $\mathbb{F}$ other than the reals, i.e., can we have $\langle \cdot, \cdot \rangle : V(\mathbb{F}) \times V(\mathbb{F}) \to \mathbb{F}$? The answer depends on the validity of property iii. The inequality $\langle \mathbf{v}, \mathbf{v} \rangle \geq 0$ implies that $\langle \mathbf{v}, \mathbf{v} \rangle$ takes a value in an ordered field. This cannot be $\mathbb{C}$, nor any finite field. However, if we make the choice $\mathbb{F} = \mathbb{C}$, we can redefine property i., so that $\langle \mathbf{v}, \mathbf{v} \rangle$ is a real number and property iii. makes sense.

**Definition 5.3 — Complex inner product.** Let $V(\mathbb{C})$ be a complex vector space. A complex inner product is a function $\langle \cdot, \cdot \rangle : V \times V \to \mathbb{C}$ which satisfies the following properties:

i. conjugate symmetry: $\langle \mathbf{u}, \mathbf{v} \rangle = \overline{\langle \mathbf{v}, \mathbf{u} \rangle}$;
ii. linearity: $\langle a\mathbf{u} + b\mathbf{v}, \mathbf{w} \rangle = a \langle \mathbf{u}, \mathbf{w} \rangle + b \langle \mathbf{v}, \mathbf{w} \rangle$;
iii. non-negativity: $\langle \mathbf{v}, \mathbf{v} \rangle \geq 0$, with $\langle \mathbf{v}, \mathbf{v} \rangle = 0$ if and only if $\mathbf{v} = \mathbf{0}$.

**Exercise 5.3** Let $V(\mathbb{C})$ be a vector space equipped with a complex inner product. Show that for any vector $\mathbf{v} \in V$ there holds $\langle \mathbf{v}, \mathbf{v} \rangle \in \mathbb{R}$.

In the following, we will focus our attention mostly on real vector spaces, although we will occasionally comment also on the complex case.

## 5.2 Inner product spaces

Inner products provide additional structure to vector spaces. We will focus in the remainder of this course on vector spaces equipped with a real inner product.

**Definition 5.4 — Inner product spaces.** A vector space $V(\mathbb{F})$ equipped with an inner product $\langle \cdot, \cdot \rangle : V \times V \to \mathbb{F}$ is called an inner product space.

To indicate that $V(\mathbb{F})$ is an inner product space, we sometimes write $(V(\mathbb{F}), \langle \cdot, \cdot \rangle)$ or $(V, \langle \cdot, \cdot \rangle)$, although often the inner product will be evident from the context.

R One can define more than one inner product on a vector space. The choice usually is provided and/or justified by applications.

Given an inner product space, we can immediately introduce two concepts analogous to those used for Euclidean spaces, namely, length (or norm) of a vector and angle between vectors.

### 5.2.1 Length

**Definition 5.5 — Length/norm of a vector.** Let $V(\mathbb{F})$ be an inner product space equipped with an inner product $\langle\cdot,\cdot\rangle$. For any vector $\mathbf{v} \in V$ we denote its length or norm by $\|\mathbf{v}\|$ and define it via

$$\|\mathbf{v}\| = \sqrt{\langle \mathbf{v}, \mathbf{v}\rangle}.$$

Note that the definition makes sense (i.e., the length is defined to be a positive real number, unless the vector is the zero vector) due to property iii. of inner products. This means that every inner product space has automatically a norm function defined on it: $\|\cdot\| : V \to \mathbb{R}_+ \cup \{0\}$. This makes $V$ a **normed vector space**; the norm is called the **induced norm**. To indicate this relationship, the inner-product on $V$ and the corresponding induced norm are sometimes denoted by $\langle\cdot,\cdot\rangle_V$, $\|\cdot\|_V$, respectively.
This definition allows us to establish the following results and properties involving norms:

- Cauchy-Schwarz inequality;
- triangle inequality;
- length of a scaled vector.

We derive each result below.

**Proposition 5.1 — Cauchy-Schwarz inequality.** Let $\|\cdot\|$ denote the norm induced by a real inner product on a vector space $V(\mathbb{R})$. Then

$$|\langle \mathbf{u}, \mathbf{v}\rangle| \leq \|\mathbf{u}\| \, \|\mathbf{v}\| \quad \text{for any } \mathbf{u}, \mathbf{v} \in V.$$

*Proof.* Let $\mathbf{u}, \mathbf{v} \in V$ and let $a \in \mathbb{R}$. Then, using the properties of inner-products,

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

This is a quadratic inequality in $a$, which holds provided the discriminant associated with the quadratic on the right is non-positive:

$$\Delta := 4\langle \mathbf{u}, \mathbf{v}\rangle^2 - 4\|\mathbf{u}\|^2\|\mathbf{v}\|^2 \leq 0 \Longrightarrow |\langle \mathbf{u}, \mathbf{v}\rangle| \leq \|\mathbf{u}\| \, \|\mathbf{v}\| .$$

■

The triangle inequality can be viewed as corollary of the Cauchy-Schwarz inequality.

**Proposition 5.2 — Triangle inequality.** Let $V(\mathbb{R})$ be an inner product space. Then for any $\mathbf{u}, \mathbf{v} \in V$ there holds

$$\|\mathbf{u} + \mathbf{v}\| \leq \|\mathbf{u}\| + \|\mathbf{v}\| .$$

*Proof.* We have, using the Cauchy-Schwarz inequality,

$$\|\mathbf{u} + \mathbf{v}\|^2 = \langle \mathbf{u} + \mathbf{v}, \mathbf{u} + \mathbf{v}\rangle = \|\mathbf{u}\|^2 + 2\langle \mathbf{u}, \mathbf{v}\rangle + \|\mathbf{v}\|^2 \leq \|\mathbf{u}\|^2 + 2\|\mathbf{u}\| \, \|\mathbf{v}\| + \|\mathbf{v}\|^2 = (\|\mathbf{u}\| + \|\mathbf{v}\|)^2$$

and the result follows by taking square-roots on both sides. ■

**Proposition 5.3 — Length of scaled vector.** Let $V(\mathbb{R})$ be an inner product space. Then for any $\mathbf{v} \in V$ and $a \in \mathbb{R}$, there holds

$$\|a\mathbf{v}\| = |a| \, \|\mathbf{v}\| .$$

*Proof.* We have, using the properties of the inner product,

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

The result follows by taking square-roots on both sides. ■

R This result confirms that multiplying a vector by a general scalar $a$ results in a vector with length multiplied by $|a|$. Indeed, the sign of $a$ should not (and does not) play a role in the resulting length.

Before we consider some examples, we note that given any vector $\mathbf{v}$ in an normed space, we can always associate with it a **unit vector** defined via

$$\widehat{\mathbf{v}} := \frac{\mathbf{v}}{\|\mathbf{v}\|}.$$

This is simply a scaled vector with $a = 1/\|\mathbf{v}\|$, so that $\|\widehat{\mathbf{v}}\| = 1$.

Let us now consider some examples.

**Example 5.4 — Cauchy-Schwarz inequality on $\mathbb{R}^n$.** Let $\langle\cdot,\cdot\rangle$ denote the standard (Euclidean) inner product on $\mathbb{R}^n$. Then, by the Cauchy-Schwarz inequality, for any vectors $\mathbf{u}, \mathbf{v} \in \mathbb{R}^n$ there holds

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

**Example 5.5 — Cauchy-Schwarz inequality on $C^0([a,b])$.** Let $\langle\cdot,\cdot\rangle$ be defined via

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

Then, by the Cauchy-Schwarz inequality, for any continuous functions defined on $[a,b]$, there holds

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

### 5.2.2 Angles

The definition of angle between vectors can be extended directly from the Euclidean case to general inner product spaces

**Definition 5.6 — Angle between vectors.** Let $V(\mathbb{F})$ be an inner product space equipped with an inner product $\langle\cdot,\cdot\rangle$. The angle $\alpha$ between two non-zero vectors $\mathbf{u}, \mathbf{v} \in V$ is defined via

$$\cos\alpha = \frac{\langle\mathbf{u},\mathbf{v}\rangle}{\|\mathbf{u}\|\,\|\mathbf{v}\|}.$$

R The Cauchy-Schwarz inequality implies that, if $\mathbf{u}, \mathbf{v}$ are non-zero,

$$-|\langle\mathbf{u},\mathbf{v}\rangle| \leq \langle\mathbf{u},\mathbf{v}\rangle \leq |\langle\mathbf{u},\mathbf{v}\rangle| \iff -1 \leq \frac{\langle\mathbf{u},\mathbf{v}\rangle}{\|\mathbf{u}\|\,\|\mathbf{v}\|} \leq 1.$$

This confirms that the definition of the angle between vectors $\mathbf{u}$ and $\mathbf{v}$ is well-posed (given that for any angle $\alpha$ there holds $-1 \leq \cos\alpha \leq 1$).

The concept of angle between vectors leads naturally to the concept of orthogonal vectors.

**Definition 5.7 — Orthogonal vectors.** Let $V(\mathbb{R})$ denote an inner product space. Then the nonzero vectors $\mathbf{u}, \mathbf{v} \in V$ are said to be orthogonal if $\langle\mathbf{u},\mathbf{v}\rangle = 0$. We write $\mathbf{u} \perp \mathbf{v}$.

Orthogonal vectors in an inner product space satisfy the following characterisation.

**Proposition 5.4 — Pythagoras.** Let $V(\mathbb{R})$ denote an inner product space. Let $\mathbf{u}, \mathbf{v} \in V \setminus \{\mathbf{0}\}$. Then

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

*Proof.* Follows from the identity (worth remembering)

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

■

**Example 5.6** Let $r, s \in \mathbb{N}$ and $a > 0$. Define $p(x) = x^{2r}, q(x) = x^{2s-1}$. Then $p$ and $q$ are orthogonal in the inner product

$$\langle p, q\rangle := \int_{-a}^{a} p(x)q(x)\mathrm{d}x.$$