4 Inner product spaces

Chapter contents

4.1 Dot product on $\mathbb{R}^n$

The *Euclidean inner product* (or *dot product*) of $\mathbf{u} = \begin{bmatrix} u_1 \\ \vdots \\ u_n \end{bmatrix} \in \mathbf{R}^n$ and $\mathbf{v} = \begin{bmatrix} v_1 \\ \vdots \\ v_n \end{bmatrix} \in \mathbf{R}^n$ is

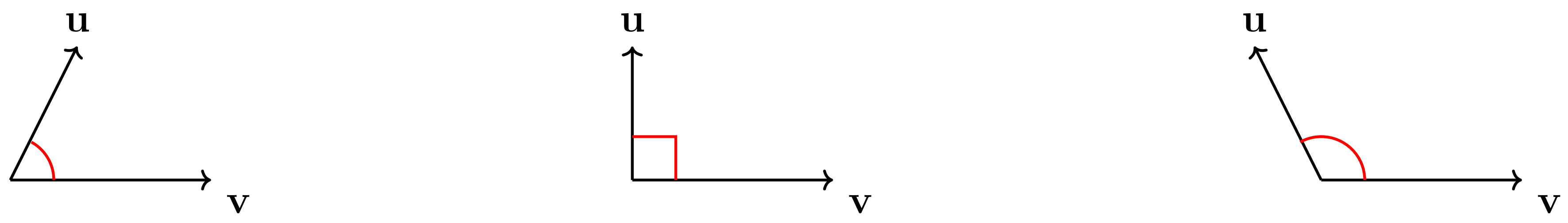
$$\mathbf{u} \cdot \mathbf{v} = u_1 v_1 + \cdots + u_n v_n.$$

The *length* (or *norm* or *magnitude*) of $\mathbf{u}$ is

$$\|\mathbf{u}\| = \sqrt{u_1^2 + \cdots + u_n^2} = \sqrt{\mathbf{u} \cdot \mathbf{u}}.$$

A *unit vector* is a vector $\mathbf{u}$ with $\|\mathbf{u}\| = 1$.

Angles Geometrically, two non-parallel vectors $\mathbf{u}, \mathbf{v} \in \mathbf{R}^n$ form a plane W in $\mathbf{R}^n$. We define the *angle between* $\mathbf{u}$ and $\mathbf{v}$ to be the angle between these vectors as line segments in the plane W (where this notion is already familiar).



For any $\mathbf{u}, \mathbf{v} \in \mathbf{R}^n$ we have

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

where θ is the angle between $\mathbf{u}$ and $\mathbf{v}$.

In particular: two nonzero vectors $\mathbf{u}, \mathbf{v} \in \mathbf{R}^n$ are perpendicular if and only if $\mathbf{u} \cdot \mathbf{v} = 0$.

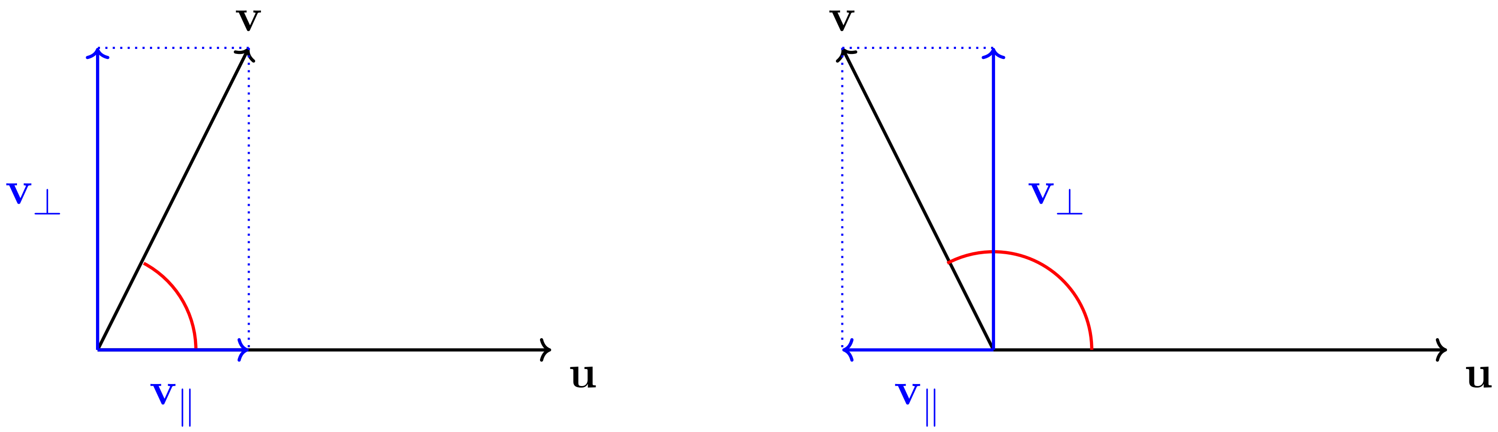
Orthogonal projection Given $\mathbf{u}, \mathbf{v} \in \mathbf{R}^n$ with $\mathbf{u} \neq \mathbf{0}$, the *orthogonal projection of* $\mathbf{v}$ *onto* $\mathbf{u}$ is

$$\mathbf{v}_{\parallel} = \operatorname{proj}_{\mathbf{u}} \mathbf{v} = \big(\mathbf{v} \cdot \hat{\mathbf{u}}\big)\hat{\mathbf{u}},$$

where $\hat{\mathbf{u}} = \frac{1}{\|\mathbf{u}\|}\mathbf{u}$ denotes the *unit vector in the direction of* $\mathbf{u}$.

The *component of* $\mathbf{v}$ *orthogonal to* $\mathbf{u}$ is

$$\mathbf{v}_{\perp} = \mathbf{v} - \mathbf{v}_{\parallel}.$$



Example 4.1.

(a) Find the orthogonal projection of $\begin{bmatrix}1\\2\\3\end{bmatrix}$ onto $\mathbf{e}_1$.

(b) Find the orthogonal projection of $\begin{bmatrix}1\\2\\3\end{bmatrix}$ onto $\begin{bmatrix}1\\1\\1\end{bmatrix}$.

(c) Find the component of $\begin{bmatrix}1\\2\\3\end{bmatrix}$ orthogonal to $\begin{bmatrix}1\\1\\1\end{bmatrix}$.

Properties of the dot product

1. $\mathbf{u} \cdot \mathbf{v} = \mathbf{v} \cdot \mathbf{u}$
2. $(\lambda \mathbf{u}) \cdot \mathbf{v} = \lambda(\mathbf{u} \cdot \mathbf{v})$
3. $(\mathbf{u} + \mathbf{v}) \cdot \mathbf{w} = \mathbf{u} \cdot \mathbf{w} + \mathbf{v} \cdot \mathbf{w}$
4. (a) $\mathbf{u} \cdot \mathbf{u} \geqslant 0$

   (b) $\mathbf{u} \cdot \mathbf{u} = 0 \Rightarrow \mathbf{u} = \mathbf{0}$.

To make our lives considerably easier, let's note that we can also think of dot product as a special case of matrix multiplication:

$$\mathbf{u} \cdot \mathbf{v} = \mathbf{u}^T \mathbf{v} = \mathbf{u}^T I \mathbf{v}.$$

This gives easy proofs of properties 1, 2, and 3, and invites generalisation: replace I by a suitable matrix.

4.2 Inner products

Let V be a vector space with field of scalars $\mathbf{R}$.

An *inner product on V* is a function

$$\langle\, ,\, \rangle : \; V \times V \longrightarrow \mathbf{R}$$

satisfying, for all $\mathbf{u}, \mathbf{v}, \mathbf{w} \in V$ and all $\lambda \in \mathbf{R}$:

1. $\langle \mathbf{u}, \mathbf{v} \rangle = \langle \mathbf{v}, \mathbf{u} \rangle$
2. $\langle \lambda\mathbf{u}, \mathbf{v} \rangle = \lambda\langle \mathbf{u}, \mathbf{v} \rangle$
3. $\langle \mathbf{u} + \mathbf{v}, \mathbf{w} \rangle = \langle \mathbf{u}, \mathbf{w} \rangle + \langle \mathbf{v}, \mathbf{w} \rangle$
4. (a) $\langle \mathbf{u}, \mathbf{u} \rangle \geqslant 0$

   (b) $\langle \mathbf{u}, \mathbf{u} \rangle = 0 \Rightarrow \mathbf{u} = \mathbf{0}$.

An *inner product space* is a vector space V together with a choice of inner product.

If W is a subspace of V, then W is itself an inner product space with respect to the inner product of V.

Example 4.2. Consider $\mathbf{R}^2$ with

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

Let's note that this inner product can be rewritten as

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

Proposition 4.3. *Let* $A \in M_n(\mathbf{R})$ *and define, for* $\mathbf{u}, \mathbf{v} \in \mathbf{R}^n$*:*

$$\langle \mathbf{u}, \mathbf{v} \rangle = \mathbf{u}^T A \mathbf{v}.$$

Then $\langle \cdot, \cdot \rangle$ *satisfies conditions 2 and 3 of an inner product.*

Moreover, $\langle \cdot, \cdot \rangle$ *satisfies condition 1 of an inner product if and only if* A *is a symmetric matrix, that is* $A^T = A$*.*

Example 4.4. Consider $\mathbf{R}^3$ with

$$\langle \mathbf{u}, \mathbf{v} \rangle = u_1 v_1 - u_2 v_2 + u_3 v_3.$$

Example 4.5. Consider $\mathbf{R}^2$ with

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

Example 4.6. Consider the vector space of continuous functions $\mathcal{C}([0,1])$ with

$$\langle f, g \rangle = \int_0^1 f(x)g(x)\,dx.$$

Proposition 4.7. *Let V be an inner product space, then for any $\mathbf{v} \in V$ we have*

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

Proposition 4.8. *Let V be an inner product space and let S be a spanning set for V. If $\mathbf{u} \in V$ has the property that*

$$\langle \mathbf{u}, \mathbf{v} \rangle = 0 \quad \textit{for all } \mathbf{v} \in S,$$

then $\mathbf{u} = \mathbf{0}$.

4.3 The Cauchy–Schwarz inequality

Let V be an inner product space.

The *length* of a vector $\mathbf{u} \in V$ is

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

The *distance* between two vectors $\mathbf{u}, \mathbf{v} \in V$ is

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

To define the angle between two vectors, we need the Cauchy–Schwarz inequality:

Theorem 4.9. *Let* $\mathbf{u}$, $\mathbf{v}$ *be vectors in an inner product space* V*. Then*

$$|\langle \mathbf{u}, \mathbf{v} \rangle| \leqslant \|\mathbf{u}\| \, \|\mathbf{v}\|,$$

where equality holds if and only if $\mathbf{u}$ *and* $\mathbf{v}$ *are parallel.*

The *angle* θ between two vectors $\mathbf{u}, \mathbf{v} \in V$ is defined by the equation

$$\cos(\theta) = \frac{\langle \mathbf{u}, \mathbf{v} \rangle}{\|\mathbf{u}\| \, \|\mathbf{v}\|}.$$

Example 4.10. Consider $\mathbf{R}^2$ with the inner product

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

Compute $\|\mathbf{u}\|$, $d(\mathbf{u}, \mathbf{v})$, and the angle between $\mathbf{u}$ and $\mathbf{v}$ for $\mathbf{u} = \begin{bmatrix} 2 \\ 5 \end{bmatrix}$ and $\mathbf{v} = \begin{bmatrix} -1 \\ 3 \end{bmatrix}$.

4.4 Orthogonality

Let V be an inner product space.

We say that $\mathbf{u}, \mathbf{v} \in V$ are *orthogonal* if $\langle \mathbf{u}, \mathbf{v} \rangle = 0$.

We say that a subset $S \subseteq V$ is *orthogonal* if any two distinct vectors in S are orthogonal.

We say that a subset $S \subseteq V$ is *orthonormal* if it is orthogonal and every vector in S has length 1.

Example 4.11. In $\mathbf{R}^2$ with the dot product,

- the vectors $\begin{bmatrix} 1 \\ 1 \end{bmatrix}$ and $\begin{bmatrix} 1 \\ -1 \end{bmatrix}$ are

- the vectors $\begin{bmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{bmatrix}$ and $\begin{bmatrix} \frac{1}{\sqrt{2}} \\ \frac{-1}{\sqrt{2}} \end{bmatrix}$ are

Example 4.12. In $\mathcal{C}([-1,1])$ with inner product

$$\langle f, g\rangle = \int_{-1}^{1} f(x)g(x)\,dx,$$

the functions x^2 and x^3 are

Theorem 4.13. *Let $S = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ be an orthogonal set of nonzero vectors in an inner product space. Then S is linearly independent.*

(We call S an orthogonal basis for $\mathrm{Span}(S)$. Moreover, if S is an orthonormal set, we call it an orthonormal basis for $\mathrm{Span}(S)$.)

If $\{\mathbf{u}_1, \ldots, \mathbf{u}_n\}$ is an orthonormal basis for V, then every $\mathbf{v} \in V$ can be written

$$\mathbf{v} =$$

(We will soon see that every inner product space has orthonormal bases.)

Proposition 4.14. *Let $\mathbf{u}, \mathbf{v}$ be vectors in an inner product space V, with $\mathbf{u} \neq \mathbf{0}$. Then the vector*

$$\mathbf{v} - \frac{\langle \mathbf{v}, \mathbf{u} \rangle}{\langle \mathbf{u}, \mathbf{u} \rangle} \mathbf{u}$$

is orthogonal to $\mathbf{u}$.

This motivates the definition: the *orthogonal projection of* $\mathbf{v}$ *onto* $\mathbf{u}$ is

$$\operatorname{proj}_{\mathbf{u}}(\mathbf{v}) = \frac{\langle \mathbf{v}, \mathbf{u} \rangle}{\langle \mathbf{u}, \mathbf{u} \rangle}\, \mathbf{u}.$$

If $\mathbf{u}$ happens to be a unit vector, then the formula simplifies to

$$\operatorname{proj}_{\mathbf{u}}(\mathbf{v}) = \langle \mathbf{v}, \mathbf{u} \rangle\, \mathbf{u}.$$

Moreover, we can project onto a subspace W of V as follows: let $\{\mathbf{u}_1, \ldots, \mathbf{u}_m\}$ be an orthonormal basis for W. The *orthogonal projection of* $\mathbf{v}$ *onto* W is

$$\begin{aligned} \operatorname{proj}_{W}(\mathbf{v}) &= \operatorname{proj}_{\mathbf{u}_1}(\mathbf{v}) + \cdots + \operatorname{proj}_{\mathbf{u}_m}(\mathbf{v}) \\ &= \langle \mathbf{v}, \mathbf{u}_1 \rangle\, \mathbf{u}_1 + \cdots + \langle \mathbf{v}, \mathbf{u}_m \rangle\, \mathbf{u}_m. \end{aligned}$$

Note that this defines a linear transformation $\operatorname{proj}_W : V \longrightarrow V$ with image W.

4.5 The Gram–Schmidt orthonormalisation process

Let V be an inner product space.

There is a procedure that starts with an arbitrary spanning set $S = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ of V and returns an orthonormal basis $\mathcal{B} = \{\mathbf{u}_1, \ldots, \mathbf{u}_m\}$, where (of course) $m \leqslant n$.

Example 4.15. Let $W = \text{Span}(\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3) \subseteq \mathbf{R}^4$ with the dot product, where

$$\mathbf{v}_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} \qquad \mathbf{v}_2 = \begin{bmatrix} 2 \\ 4 \\ 2 \\ 4 \end{bmatrix} \qquad \mathbf{v}_3 = \begin{bmatrix} 1 \\ 5 \\ -1 \\ 3 \end{bmatrix}.$$

(a) Find an orthonormal basis for W.

(b) Find the point of W closest to the point $\mathbf{v} = \begin{bmatrix} 1 \\ 2 \\ 3 \\ 0 \end{bmatrix}$.

4.6 Orthogonal matrices

An $n \times n$ matrix Q is an *orthogonal matrix* if

$$Q^T Q = I.$$

In particular, Q is

Example 4.16. $Q = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}$

Proposition 4.17. *An* $n \times n$ *matrix is orthogonal if and only if its columns (equivalently, its rows) form an orthonormal basis of* $\mathbf{R}^n$.

Proposition 4.18. *If Q is an $n \times n$ orthogonal matrix then*

$$(Q\mathbf{u}) \cdot (Q\mathbf{v}) = \mathbf{u} \cdot \mathbf{v} \qquad \textit{for all } \mathbf{u}, \mathbf{v} \in \mathbf{R}^n.$$

So orthogonal matrices preserve the dot product of vectors. Therefore they also preserve lengths and angles.

Theorem 4.19. *If A is a symmetric $n \times n$ real matrix then*

(a) the characteristic polynomial of A has n real roots (counted with multiplicity) – so A has n (not necessarily distinct) real eigenvalues;

(b) A has n orthonormal eigenvectors.

This is a special case of the spectral theorem for linear operators on finite-dimensional inner product spaces, which is covered in MAST20022 **Group Theory and Linear Algebra**.

So if A is symmetric then there is an orthonormal basis of $\mathbf{R}^n$ consisting of eigenvectors of A.

Hence A is diagonalisable:

$$Q^{-1}AQ = D$$

by an orthogonal change of basis matrix Q.

(We say that A is *orthogonally diagonalisable*.)

Proposition 4.20. *Let A be a real symmetric matrix. Any two eigenvectors of A with distinct eigenvalues are orthogonal.*

Example 4.21. Given $A = \begin{bmatrix} 3 & 2 & 4 \\ 2 & 0 & 2 \\ 4 & 2 & 3 \end{bmatrix}$, orthogonally diagonalise A.