3
Scalar Products

20 September 2021

Although vector products on $\mathbb{R}^n$ are rare, every coordinate space $\mathbb{R}^n$ is equipped with an operation that sends a pair of vectors to a scalar. This chapter explores this scalar product. We highlight its applications to inequalities, orthogonal projections, and hyperplanes.

3.0 The Dot Product

How do we combine two vectors to obtain a scalar? The dot product may be defined algebraically or geometrically.

The name "dot product" and the notation were introduced in 1881 by J.W. Gibbs.

3.0.0 Definition. For any two vector $\vec{v}$ and $\vec{w}$ in $\mathbb{R}^n$, the following two definitions of the ***dot product*** $\vec{v} \cdot \vec{w} \in \mathbb{R}$ are equivalent.

(geometric) When $0 \leqslant \theta \leqslant \pi$ is the angle between the vectors $\vec{v}$ and $\vec{w}$, we set $\vec{v} \cdot \vec{w} := \|\vec{v}\| \, \|\vec{w}\| \cos(\theta)$.

(algebraic) Assuming that $\vec{v} := v_1 \vec{e}_1 + v_2 \vec{e}_2 + \cdots + v_n \vec{e}_n$ and $\vec{w} := w_1 \vec{e}_1 + w_2 \vec{e}_2 + \cdots + w_n \vec{e}_n$, we set

$$\vec{v} \cdot \vec{w} := v_1 w_1 + v_2 w_2 + \cdots + v_n w_n \,.$$

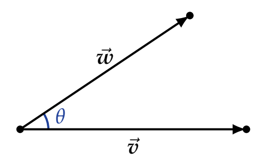

Figure 3.0: The angle θ between the vectors $\vec{v}$ and $\vec{w}$

3.0.1 Problem. For all $1 \leqslant j \leqslant k \leqslant n$, demonstrate that $\vec{e}_j \cdot \vec{e}_k = \delta_{j,k}$.

The ***Kronecker delta*** is the function defined by

$$\delta_{j,k} := \begin{cases} 1 & \text{if } j = k, \\ 0 & \text{if } j \neq k. \end{cases}$$

Geometric solution. Since the standard basis $\vec{e}_1, \vec{e}_2, \dots, \vec{e}_n$ consists of pairwise perpendicular unit vectors, the geometric definition of the dot product implies that $\vec{e}_j \cdot \vec{e}_k = \|\vec{e}_j\| \, \|\vec{e}_k\| \cos\left(\frac{\pi}{2}\right) = (1)(1)(0) = 0$, for all $j \neq k$, and $\vec{e}_j \cdot \vec{e}_j = \|\vec{e}_j\| \, \|\vec{e}_j\| \cos(0) = (1)(1)(1) = 1$. □

Algebraic solution. Since the vector $\vec{e}_j$ has 1 in the j-th entry and zero elsewhere, the algebraic definition of the dot product gives

$$\vec{e}_j \cdot \vec{e}_k = (0)(0) + (0)(0) + \cdots + \underbrace{(1)(0)}_{j\text{-th summand}} + \cdots + \underbrace{(0)(1)}_{k\text{-th summand}} + \cdots + (0)(0) = 0 \,,$$

$$\vec{e}_j \cdot \vec{e}_j = (0)(0) + (0)(0) + \cdots + \underbrace{(1)(1)}_{j\text{-th summand}} + \cdots + (0)(0) = 1 \,,$$

and we conclude that $\vec{e}_j \cdot \vec{e}_k = \delta_{j,k}$. □

3.0.2 Proposition (Properties of the dot product)**.** *For all $\vec{u}, \vec{v}, \vec{w} \in \mathbb{R}^n$ and all $c \in \mathbb{R}$, the dot product has the following five properties.*

(commutativity)	$\vec{v} \cdot \vec{w} = \vec{w} \cdot \vec{v}$
(compatibility with scalar multiplication)	$\vec{v} \cdot (c\,\vec{w}) = c(\vec{v} \cdot \vec{w}) = (c\,\vec{v}) \cdot \vec{w}$
(distributivity)	$\vec{u} \cdot (\vec{v} + \vec{w}) = \vec{u} \cdot \vec{v} + \vec{u} \cdot \vec{w}$

(nonnegativity) $\vec{v}\cdot\vec{v} \geqslant 0$

(positivity) $\vec{v}\cdot\vec{v} = 0$ *if and only if* $\vec{v} = \vec{\mathbf{0}}$

Geometric proof. Let $0 \leqslant \theta \leqslant \pi$ be the angle between $\vec{v}$ and $\vec{w}$. Since multiplication in $\mathbb{R}$ is commutative, it follows that

$$\vec{v}\cdot\vec{w} = \|\vec{v}\|\,\|\vec{w}\|\cos(\theta) = \|\vec{w}\|\,\|\vec{v}\|\cos(\theta) = \vec{w}\cdot\vec{v}\,,$$

proving the commutativity of the dot product.

Scalar multiplication by a nonnegative number c rescales the magnitude without changing the direction, so we have

$$\begin{aligned}(c\,\vec{v})\cdot\vec{w} &= \|c\,\vec{v}\|\,\|\vec{w}\|\cos(\theta) = c\big(\|\vec{v}\|\,\|\vec{w}\|\cos(\theta)\big) = c(\vec{v}\cdot\vec{w})\\ &= \|\vec{v}\|\,\|c\,\vec{w}\|\cos(\theta) = \vec{v}\cdot(c\,\vec{w})\,.\end{aligned}$$

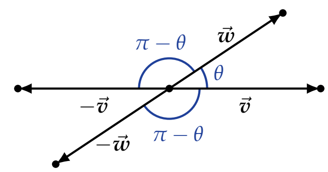

Figure 3.1: Angles between $\pm\vec{v}$ and $\pm\vec{w}$

In contrast, scalar multiplication by negative number c gives a vector in the opposite direction and rescales the magnitude by $|c|$, so

$$\begin{aligned}(c\,\vec{v})\cdot\vec{w} &= \|c\,\vec{v}\|\,\|\vec{w}\|\cos(\pi-\theta) = -\,|c|\,\big(\|\vec{v}\|\,\|\vec{w}\|\cos(\theta)\big) = c(\vec{v}\cdot\vec{w})\\ &= \|\vec{v}\|\,\|c\,\vec{w}\|\cos(\pi-\theta) = \vec{v}\cdot(c\,\vec{w})\,.\end{aligned}$$

Hence, the dot product is compatible with scalar multiplication.

If the angle between $\vec{u}$ and $\vec{v}+\vec{w}$ is ϑ, the angle between $\vec{u}$ and $\vec{v}$ is ϕ, and the angle between $\vec{u}$ and $\vec{w}$ is ψ, then trigonometry and vector addition imply that $\|\vec{v}+\vec{w}\|\cos(\vartheta) = \|\vec{v}\|\cos(\phi) + \|\vec{w}\|\cos(\psi)$. We deduce that

$$\begin{aligned}\vec{u}\cdot(\vec{v}+\vec{w}) &= \|\vec{u}\|\,\|\vec{v}+\vec{w}\|\cos(\vartheta)\\ &= \|\vec{u}\|\,\big(\|\vec{v}\|\cos(\phi) + \|\vec{w}\|\cos(\psi)\big) = \vec{u}\cdot\vec{v} + \vec{u}\cdot\vec{w}\,,\end{aligned}$$

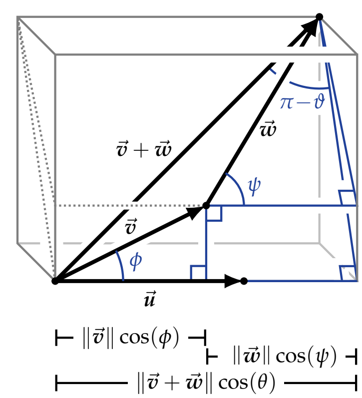

Figure 3.2: Angles between the vector $\vec{u}$ and the vectors $\vec{v}$, $\vec{w}$, $\vec{v}+\vec{w}$

which shows that the dot product is distributive.

We have $\vec{v}\cdot\vec{v} = \|\vec{v}\|\,\|\vec{v}\|\cos(0) = \|\vec{v}\|^2 \geqslant 0$ because the square of any real number is nonnegative. Since the number 0 has a unique square root and the zero vector is the unique vector with magnitude equal to 0, we have $\vec{v}\cdot\vec{v} = 0$ if and only if $\vec{v} = \vec{\mathbf{0}}$. □

Algebraic proof. Since multiplication in $\mathbb{R}$ is commutative, we have

$$\begin{aligned}\vec{v}\cdot\vec{w} &= v_1w_1 + v_2w_2 + \cdots + v_nw_n = w_1v_1 + w_2v_2 + \cdots + w_nv_n = \vec{w}\cdot\vec{v}\,,\\ (c\,\vec{v})\cdot\vec{w} &= (cv_1)w_1 + (cv_2)w_2 + \cdots + (cv_n)w_n = c(v_1w_1 + v_2w_2 + \cdots + v_nw_n) = c(\vec{v}\cdot\vec{w})\\ &= v_1(cw_1) + v_2(cw_2) + \cdots + v_n(cw_n) = \vec{v}\cdot(c\,\vec{w})\,.\end{aligned}$$

proving commutativity and compatibility with scalar multiplication. Similarly, multiplication in $\mathbb{R}$ is distributive, so we obtain

$$\begin{aligned}\vec{u}\cdot(\vec{v}+\vec{w}) &= u_1(v_1+w_1) + u_2(v_2+w_2) + \cdots u_n(v_n+w_n)\\ &= (u_1v_1 + u_2v_2 + \cdots + u_nv_n) + (u_1w_1 + u_2w_2 + \cdots + u_nw_n) = \vec{u}\cdot\vec{v} + \vec{u}\cdot\vec{w}\,.\end{aligned}$$

Since $\vec{v}\cdot\vec{v} = v_1^2 + v_2^2 + \cdots + v_n^2$, the dot product is nonnegative because the square of any real number is nonnegative. A sum of nonnegative numbers equals zero if and only if each summand is zero, so we conclude that $\vec{v}\cdot\vec{v}$ if and only if $\vec{v} = \vec{\mathbf{0}}$. □

WHY DO THE TWO DEFINITIONS OF THE DOT PRODUCT AGREE? For all $\vec{v} := v_1\,\vec{e}_1 + v_2\,\vec{e}_2 + \cdots + v_n\,\vec{e}_n$ and $\vec{w} := w_1\,\vec{e}_1 + w_2\,\vec{e}_2 + \cdots + w_n\,\vec{e}_n$, the properties of dot product establish that

$$\begin{aligned}\vec{v}\cdot\vec{w} &= (v_1\,\vec{e}_1 + v_2\,\vec{e}_2 + \cdots + v_n\,\vec{e}_n)\cdot(w_1\,\vec{e}_1 + w_2\,\vec{e}_2 + \cdots + w_n\,\vec{e}_n)\\ &= \begin{array}{cccc} v_1w_1(\vec{e}_1\cdot\vec{e}_1) + & v_1w_2(\vec{e}_1\cdot\vec{e}_2) + & \cdots + & v_1w_n(\vec{e}_1\cdot\vec{e}_n)\\ +v_2w_1(\vec{e}_1\cdot\vec{e}_2) + & v_2w_2(\vec{e}_2\cdot\vec{e}_2) + & \cdots + & v_2w_n(\vec{e}_2\cdot\vec{e}_n)\\ \vdots & \vdots & \ddots & \vdots\\ +v_nw_1(\vec{e}_1\cdot\vec{e}_n) + & v_nw_2(\vec{e}_2\cdot\vec{e}_n) + & \cdots + & v_nw_n(\vec{e}_n\cdot\vec{e}_n)\,. \end{array}\end{aligned}$$

Hence, it suffices to know that geometric and algebraic definitions of the dot products agree on $\vec{e}_j\cdot\vec{e}_k$ where $1 \leqslant j \leqslant k \leqslant n$; see Problem 3.0.1.

Exercises

3.0.3 Problem. Determine which of the following statements are true. If a statement is false, then provide a counterexample.

i. The dot product of two vectors is another vector.
ii. In $\mathbb{R}^2$, the dot product and complex multiplication are equal.
iii. In $\mathbb{R}^3$, the dot product and the cross product are equal.
iv. The dot product of two nonzero vector equals 0 if and only if the angle between the two vectors is $\pi/2$.
v. The dot product is anti-commutative.
vi. The dot product is associative.

3.0.4 Problem. If $\vec{v} = \vec{e}_1$ and $\vec{w} = 2\vec{e}_1 + 2\vec{e}_2$, then compute $\vec{v}\cdot\vec{w}$ both geometrically and algebraically.

3.0.5 Problem. A store sells computers, tablets, phones, and watches. The quantity vector $\vec{q}$ has components equal to the number of sales of each item. The price vector $\vec{p}$ has components equal to the price per unit of each item. What does the dot product $\vec{p}\cdot\vec{q}$ represent?

3.0.6 Problem. Let $\vec{v} \in \mathbb{R}^n$ have magnitude 2. If $\vec{u} \in \mathbb{R}^n$ has length 3, what are the maximum and minimum values of the dot product $\vec{u}\cdot\vec{v}$? What configurations lead to these extremal values?

3.0.7 Problem. Given $\vec{u} \in \mathbb{R}^n$ and $\vec{v} \in \mathbb{R}^n$ such that $\vec{u}\cdot\vec{w} = \vec{v}\cdot\vec{w}$ for all $\vec{w} \in \mathbb{R}^n$, prove that $\vec{u} = \vec{v}$.

3.0.8 Problem. For three vectors $\vec{u}, \vec{v}, \vec{w} \in \mathbb{R}^3$, the ***scalar triple product*** is $\vec{u}\cdot(\vec{v}\times\vec{w})$.

i. Prove that $|\vec{u}\cdot(\vec{v}\times\vec{w})|$ equals the volume of the parallelepiped formed by the vectors $\vec{u}, \vec{v}, \vec{w} \in \mathbb{R}^3$.
ii. Show that $\vec{u}\cdot(\vec{v}\times\vec{w}) = \vec{v}\cdot(\vec{w}\times\vec{u}) = \vec{w}\cdot(\vec{u}\times\vec{v})$.
iii. Demonstrate that the geometric definition of the cross product satisfies the distributivity property.

3.0.9 Problem. For any three vectors $\vec{u}, \vec{v}, \vec{w} \in \mathbb{R}^3$, prove the following identities:

i. $\vec{u} \times (\vec{v} \times \vec{w}) = (\vec{u} \cdot \vec{w})\, \vec{v} - (\vec{u} \cdot \vec{v})\, \vec{w}$,

ii. $\vec{u} \times (\vec{v} \times \vec{w}) + \vec{v} \times (\vec{w} \times \vec{u}) + \vec{w} \times (\vec{u} \times \vec{v}) = \vec{\mathbf{0}}$.

3.1 Essential Inequalities

HOW DOES THE DOT PRODUCT PRODUCE INEQUALITIES? The geometric definition of the dot product implies that nonzero vectors are perpendicular if and only if the angle between them is $\pi/2$. We typically use another term for this feature.

3.1.0 Definition. Two vectors $\vec{v}$ and $\vec{w}$ are ***orthogonal*** if $\vec{v} \cdot \vec{w} = 0$.

3.1.1 Problem. Which pairs among the three vectors

$$\vec{u} := \vec{e}_1 + \sqrt{3}\,\vec{e}_3\,, \qquad \vec{v} := \vec{e}_1 + \sqrt{3}\,\vec{e}_2\,, \qquad \vec{w} := \sqrt{3}\,\vec{e}_1 + \vec{e}_2 - \vec{e}_3\,,$$

are orthogonal?

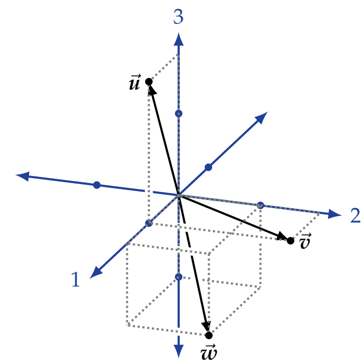

Figure 3.3: Three vectors in $\mathbb{R}^3$

Solution. Since we have

$$\begin{aligned} \vec{u} \cdot \vec{v} &= (1)(1) + (0)(\sqrt{3}) + (\sqrt{3})(0) = 1\,, \\ \vec{u} \cdot \vec{w} &= (1)(\sqrt{3}) + (0)(1) + (\sqrt{3})(-1) = 0\,, \\ \vec{v} \cdot \vec{w} &= (1)(\sqrt{3}) + (\sqrt{3})(1) + (0)(-1) = 2\sqrt{3}\,, \end{aligned}$$

only $\vec{u}$ and $\vec{w}$ are orthogonal. □

Adding two orthogonal vectors gives a right angled triangle. From the properties of the dot product, we easily obtain the celebrated relation among the three sides of a right angled triangle.

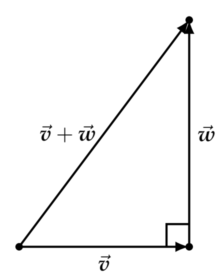

Figure 3.4: A right angled triangle

3.1.2 Proposition (Pythagorean theorem)**.** *For any pair of orthogonal vectors $\vec{v}$ and $\vec{w}$ in $\mathbb{R}^n$, we have $\|\vec{v} + \vec{w}\|^2 = \|\vec{v}\|^2 + \|\vec{w}\|^2$.*

Proof. Since $\vec{v}$ and $\vec{w}$ are orthogonal, we have $\vec{v} \cdot \vec{w} = 0$ and the properties of the dot product [3.0.2] give

$$\begin{aligned} \|\vec{v} + \vec{w}\|^2 &= (\vec{v} + \vec{w}) \cdot (\vec{v} + \vec{w}) \\ &= \|\vec{v}\|^2 + \vec{v} \cdot \vec{w} + \vec{w} \cdot \vec{v} + \|\vec{w}\|^2 = \|\vec{v}\|^2 + \|\vec{w}\|^2\,. \end{aligned}$$ □

3.1.3 Theorem (Cauchy–Schwarz inequality)**.** *For all $\vec{v}, \vec{w} \in \mathbb{R}^n$, we have*

$$|\vec{v} \cdot \vec{w}| \leqslant \|\vec{v}\|\,\|\vec{w}\| \qquad \text{or} \qquad \left(\sum_{i=1}^{n} v_i w_i \right)^2 \leqslant \left(\sum_{j=1}^{n} v_j^2 \right) \left(\sum_{k=1}^{n} w_k^2 \right).$$

Equality holds if and only if the vectors are parallel.

The inequality for sums was published in 1821 by A.-L. Cauchy.

Geometric proof. Since $-1 \leqslant \cos(\theta) \leqslant 1$, we have $|\cos(\theta)| \leqslant 1$, and the geometric definition of the dot product [3.0.0] gives

$$|\vec{v} \cdot \vec{w}| = \|\vec{v}\| \, \|\vec{w}\| \, |\cos(\theta)| \leqslant \|\vec{v}\| \, \|\vec{w}\| \; .$$

Equality holds if and only if $\cos(\theta) = \pm 1$. Hence, either $\theta = 0$, which implies $\vec{v}$ and $\vec{w}$ point in the same direction, or $\theta = \pi$, which implies $\vec{v}$ and $\vec{w}$ point in the opposite directions. □

Algebraic proof. When $\vec{w} = \vec{0}$, both sides of the inequality equal 0, so we may assume $\vec{w} \neq \vec{0}$. Consider $\vec{u} := \vec{v} - \big(\frac{\vec{v}\cdot\vec{w}}{\vec{w}\cdot\vec{w}}\big)\vec{w} \in \mathbb{R}^n$, which means that $\vec{v} = \big(\frac{\vec{v}\cdot\vec{w}}{\vec{w}\cdot\vec{w}}\big)\vec{w} + \vec{u}$. Since $\vec{u} \cdot \vec{w} = \vec{v} \cdot \vec{w} - \big(\frac{\vec{v}\cdot\vec{w}}{\vec{w}\cdot\vec{w}}\big)\vec{w} \cdot \vec{w} = 0$, the vectors $\vec{u}$ and $\vec{w}$ are orthogonal. Hence, the Pythagorean theorem combined with the nonnegative of magnitude imply that

$$\|\vec{v}\|^2 = \left\| \left(\frac{\vec{v} \cdot \vec{w}}{\vec{w} \cdot \vec{w}} \right) \vec{w} + \vec{u} \right\| = \left\| \left(\frac{\vec{v} \cdot \vec{w}}{\vec{w} \cdot \vec{w}} \right) \vec{w} \right\|^2 + \|\vec{u}\|^2 = \frac{(\vec{v} \cdot \vec{w})^2}{\|\vec{w}\|^2} + \|\vec{u}\|^2 \geqslant \frac{(\vec{v} \cdot \vec{w})^2}{\|\vec{w}\|^2} \, .$$

Multiplying by $\|\vec{w}\|^2$ and taking square roots gives the inequality. We have equality if and only if $\vec{u} = \vec{0}$ which equivalent to saying that $\vec{v}$ is a scalar multiple of $\vec{w}$. □

3.1.4 Theorem (Triangle inequality)**.** *For all $\vec{v}, \vec{w} \in \mathbb{R}^n$, we have*

$$\|\vec{v} + \vec{w}\| \leqslant \|\vec{v}\| + \|\vec{w}\| \; .$$

Equality holds if and only if one vector is a nonnegative multiple of the other.

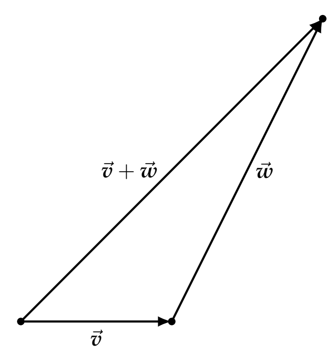

Figure 3.5: Triangle inequality

Geometrically, the triangle inequality asserts that the sum of the lengths of two sides in a triangle are at least the length of the other side. Equality occurs when the vertices of the triangle are collinear.

Proof. The properties of the dot product [3.0.0] give

$$\|\vec{v} + \vec{w}\|^2 = (\vec{v} + \vec{w}) \cdot (\vec{v} + \vec{w}) = \|\vec{v}\|^2 + \vec{v} \cdot \vec{w} + \vec{w} \cdot \vec{v} + \|\vec{w}\|^2 = \|\vec{v}\|^2 + 2\vec{v} \cdot \vec{w} + \|\vec{w}\|^2 \; .$$

For any $c \in \mathbb{R}$, we have $c \leqslant |c|$, so we obtain

$$\|\vec{v} + \vec{w}\|^2 = \|\vec{v}\|^2 + 2\vec{v} \cdot \vec{w} + \|\vec{w}\|^2 \leqslant \|\vec{v}\|^2 + 2 \, |\vec{v} \cdot \vec{w}| + \|\vec{w}\|^2 \; .$$

Applying the Cauchy-Schwarz inequality yields

$$\|\vec{v} + \vec{w}\|^2 \leqslant \|\vec{v}\|^2 + 2 \, |\vec{v} \cdot \vec{w}| + \|\vec{w}\|^2 \leqslant \|\vec{v}\|^2 + 2 \, \|\vec{v}\| \, \|\vec{w}\| + \|\vec{w}\|^2 = \big(\|\vec{v}\| + \|\vec{w}\|\big)^2 \, .$$

Taking a square root yields the desired inequality. Equality holds if and only if $\vec{v} \cdot \vec{w} = |\vec{v} \cdot \vec{w}|$ and $\vec{v}$ is parallel $\vec{w}$, which is equivalent to $\vec{v}$ being a nonnegative multiple of $\vec{w}$. □

Exercises

3.1.5 Problem. Determine which of the following statements are true. If a statement is false, then provide a counterexample.

i. The zero vector is orthogonal to every vector.
ii. The triangle inequality shows that one side of a triangle must be longer than the other two.

3.1.6 Problem. For any two vectors $\vec{v}, \vec{w} \in \mathbb{R}^n$, show that

$$\big| \|\vec{v}\| - \|\vec{w}\| \big| \leqslant \|\vec{v} - \vec{w}\| \, .$$

3.1.7 Problem. For $\vec{u}, \vec{v}, \vec{w} \in \mathbb{R}^n$, prove that

$$\left(\sum_{i=1}^{n} u_i v_i w_i \right)^2 \leqslant \left(\sum_{i=1}^{n} u_i^2 \right) \left(\sum_{j=1}^{n} v_j^2 \right) \left(\sum_{k=1}^{n} w_k^2 \right) .$$

3.1.8 Problem. For any $\vec{v}, \vec{w} \in \mathbb{R}^n$, consider the function $q \colon \mathbb{R} \to \mathbb{R}$ defined by $q(t) := (\vec{v} + t\vec{w}) \cdot (\vec{v} + t\vec{w})$. Explain why $q(t) \geqslant 0$ for all $t \in \mathbb{R}$. By interpreting $q(t)$ as a quadratic polynomial in t, show that $|\vec{v} \cdot \vec{w}| \leqslant \|\vec{v}\| \, \|\vec{w}\|$.

3.2 Orthogonal Projections

How do we find the distance from a point to a line? Let $\vec{v} := \overrightarrow{PQ}$ be a nonzero vector and let ℓ denote the line through the points P and Q. Fix a point R and consider the vector $\vec{w} := \overrightarrow{RP}$. To determine the orthogonal distance from the point R to the line ℓ, it suffices to express the vector $\vec{w}$ as the sum of a vector parallel to $\vec{v}$ and a vector orthogonal to $\vec{v}$; $\vec{w} = \overrightarrow{RP} = \overrightarrow{RS} + \overrightarrow{SP}$. Any vector parallel to $\vec{v}$ has the form $\overrightarrow{SP} = c\,\vec{v}$ for some scalar $c \in \mathbb{R}$. If the difference $\overrightarrow{RS} = \overrightarrow{RP} - \overrightarrow{SP} = \vec{w} - c\,\vec{v}$ is orthogonal to $\vec{v}$, then the properties of the dot product [3.0.0] give

$$0 = \overrightarrow{RS} \cdot \vec{v} = (\vec{w} - c\,\vec{v}) \cdot \vec{v} = \vec{w} \cdot \vec{v} - c(\vec{v} \cdot \vec{v}) \, ,$$

so we deduce that $c = \frac{\vec{w} \cdot \vec{v}}{\vec{v} \cdot \vec{v}}$, $\overrightarrow{SP} = \left(\frac{\vec{w} \cdot \vec{v}}{\vec{v} \cdot \vec{v}} \right) \vec{v}$, and $\overrightarrow{RS} = \vec{w} - \left(\frac{\vec{w} \cdot \vec{v}}{\vec{v} \cdot \vec{v}} \right) \vec{v}$. Hence, the unique expression of $\vec{w}$ as the sum of a vector parallel to $\vec{v}$ and a vector orthogonal to $\vec{v}$ is

$$\vec{w} = \left(\frac{\vec{w} \cdot \vec{v}}{\vec{v} \cdot \vec{v}} \right) \vec{v} + \left[\vec{w} - \left(\frac{\vec{w} \cdot \vec{v}}{\vec{v} \cdot \vec{v}} \right) \vec{v} \right] .$$

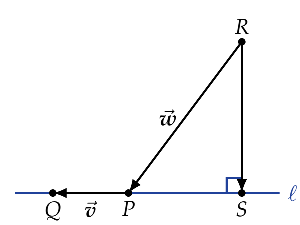

Figure 3.5: Projection onto a line

Thus, the orthogonal distance from R to ℓ is $\|\overrightarrow{RS}\| = \left\| \vec{w} - \left(\frac{\vec{w} \cdot \vec{v}}{\vec{v} \cdot \vec{v}} \right) \vec{v} \right\|$.

Inspired by this computation, we introduce the following function.

3.2.0 Definition. The ***orthogonal projection*** onto a nonzero vector $\vec{v}$ in $\mathbb{R}^n$ is the function $\operatorname{proj}_{\vec{v}} \colon \mathbb{R}^n \to \mathbb{R}^n$ defined, for all $\vec{w} \in \mathbb{R}^n$, by

$$\operatorname{proj}_{\vec{v}}(\vec{w}) := \left(\frac{\vec{w} \cdot \vec{v}}{\vec{v} \cdot \vec{v}} \right) \vec{v} \,.$$

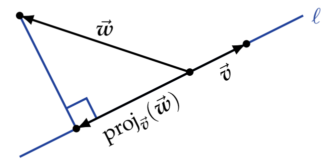

Figure 3.6: Visualizing projections

The image of a projection can be viewed as the shadow cast by the vector $\vec{w}$ on the line through the head and tail of $\vec{v}$. The ***orthogonal decomposition*** of a vector $\vec{w} \in \mathbb{R}^n$ with respect to $\vec{v}$ is the expression

$$\vec{w} = \underbrace{\operatorname{proj}_{\vec{v}}(\vec{w})}_{\text{parallel to } \vec{v}} + \underbrace{\big(\vec{w} - \operatorname{proj}_{\vec{v}}(\vec{w})\big)}_{\text{orthogonal to } \vec{v}} \,.$$

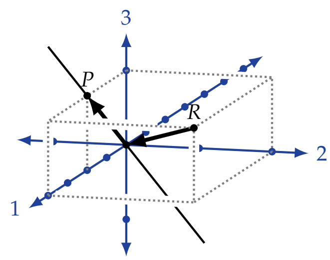

Figure 3.7: Orthogonal distance from a point to a line

3.2.1 Problem. Compute the orthogonal distance from the line through the points $O := (0,0,0)$ and $P := (2,0,1)$ to the point $R := (4,2,1)$.

Solution. As $\vec{v} := \overrightarrow{OP} = 2\,\vec{e}_1 + \vec{e}_3$ and $\vec{w} := \overrightarrow{RO} = -4\,\vec{e}_1 - 2\,\vec{e}_1 - \vec{e}_3$, we have $\operatorname{proj}_{\vec{v}}(\vec{w}) = \big(\frac{\vec{w}\cdot\vec{v}}{\vec{v}\cdot\vec{v}}\big)\vec{v} = \frac{-9}{5}(2\,\vec{e}_1 + \vec{e}_3)$, and the desired distance is

$$\begin{aligned} &\big\|(-4\,\vec{e}_1 - 2\,\vec{e}_1 - \vec{e}_3) + \tfrac{9}{5}(2\,\vec{e}_1 + \vec{e}_3)\big\| \\ &= \tfrac{1}{5}\,\|(-20+18)\,\vec{e}_1 - 10\,\vec{e}_2 + (-5+9)\,\vec{e}_3\| \\ &= \tfrac{1}{5}\sqrt{2^2 + (-10)^2 + (4)^2} = \tfrac{2}{5}\sqrt{30}\,. \end{aligned}$$

□

3.2.2 Proposition (Properties of orthogonal projections)**.** *For all vectors $\vec{u}$, $\vec{v}$, and $\vec{w}$ in $\mathbb{R}^n$ with $\vec{v} \neq \vec{0}$, we have the following.*

i. *If $\vec{w}$ is parallel to $\vec{v}$, then we have $\operatorname{proj}_{\vec{v}}(\vec{w}) = \vec{w}$.*
ii. *If $\vec{w}$ is orthogonal to $\vec{v}$, then we have $\operatorname{proj}_{\vec{v}}(\vec{w}) = \vec{0}$.*
iii. *The function $\operatorname{proj}_{\vec{v}}$ is idempotent: $\operatorname{proj}_{\vec{v}}\big(\operatorname{proj}_{\vec{v}}(\vec{w})\big) = \operatorname{proj}_{\vec{v}}(\vec{w})$.*
iv. *If $\vec{u}$ is parallel to $\vec{v}$, then we have $\big\|\vec{w} - \operatorname{proj}_{\vec{v}}(\vec{w})\big\| \leqslant \|\vec{w} - \vec{u}\|$, and equality holds if and only if $\vec{u} = \operatorname{proj}_{\vec{v}}(\vec{w})$.*

Proof.

i. If $\vec{w}$ is parallel to $\vec{v}$, then there exists a scalar $c \in \mathbb{R}$ such that $\vec{w} = c\,\vec{v}$. It follows that

$$\operatorname{proj}_{\vec{v}}(\vec{w}) = \operatorname{proj}_{\vec{v}}(c\,\vec{v}) = \left(\frac{(c\,\vec{v}) \cdot \vec{v}}{\vec{v} \cdot \vec{v}} \right) \vec{v} = c \left(\frac{\vec{v} \cdot \vec{v}}{\vec{v} \cdot \vec{v}} \right) \vec{v} = c\,\vec{v} = \vec{w}\,.$$

ii. If $\vec{w}$ is orthogonal to $\vec{v}$, then we have $\vec{w} \cdot \vec{v} = 0$ which implies that $\operatorname{proj}_{\vec{v}}(\vec{w}) = \big(\frac{\vec{w}\cdot\vec{v}}{\vec{v}\cdot\vec{v}}\big)\vec{v} = 0\,\vec{v} = \vec{0}$.

iii. By definition, the vector $\operatorname{proj}_{\vec{v}}(\vec{w})$ is parallel to the vector $\vec{v}$ and part *i* implies that $\operatorname{proj}_{\vec{v}}\big(\operatorname{proj}_{\vec{v}}(\vec{w})\big) = \operatorname{proj}_{\vec{v}}(\vec{w})$.

iv. From the orthogonal decomposition of $\vec{w}$ with respect to $\vec{v}$, we see that the vector $\vec{w} - \operatorname{proj}_{\vec{v}}(\vec{w})$ is orthogonal to $\vec{v}$. Since $\vec{u}$ is parallel to $\vec{v}$, the vectors $\vec{w} - \operatorname{proj}_{\vec{v}}(\vec{w})$ and $\operatorname{proj}_{\vec{v}}(\vec{w}) - \vec{u}$ are also orthogonal, and the Pythagorean theorem [3.1.2] shows that

$$\big\|\vec{w} - \operatorname{proj}_{\vec{v}}(\vec{w})\big\|^2 + \big\|\operatorname{proj}_{\vec{v}}(\vec{w}) - \vec{u}\big\|^2 = \big\|\big(\vec{w} - \operatorname{proj}_{\vec{v}}(\vec{w})\big) + \big(\operatorname{proj}_{\vec{v}}(\vec{w}) - \vec{u}\big)\big\|^2 = \|\vec{w} - \vec{u}\|^2 \,.$$

The nonnegativity of magnitudes gives $\left\|\operatorname{proj}_{\vec{v}}(\vec{w}) - \vec{u}\right\|^2 \geqslant 0$, so

$$\left\|\vec{w} - \operatorname{proj}_{\vec{v}}(\vec{w})\right\|^2 \leqslant \left\|\vec{w} - \operatorname{proj}_{\vec{v}}(\vec{w})\right\|^2 + \left\|\operatorname{proj}_{\vec{v}}(\vec{w}) - \vec{u}\right\|^2 = \|\vec{w} - \vec{u}\|^2 .$$

Taking square roots establishes the desired inequality. We have equality if and only if $\left\|\operatorname{proj}_{\vec{v}}(\vec{w}) - \vec{u}\right\| = 0$ which is equivalent to $\vec{u} = \operatorname{proj}_{\vec{v}}(\vec{w})$. □

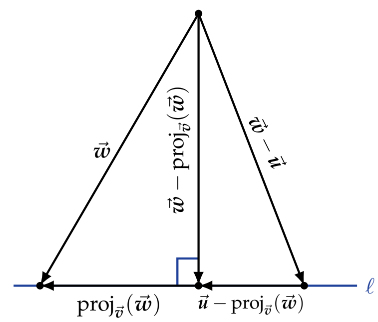

Figure 3.8: Minimizing distance from a point to a line

3.2.3 Remark. The fourth property of orthogonal projections can be geometrically rephrased as the minimum distance from a point to a line is the orthogonal distance.

3.2.4 Definition. An ***altitude*** of a triangle is a line passing through a vertex and orthogonal to the line containing the opposite side.

3.2.5 Problem. Prove that all the three altitudes of a triangle intersect at a common point.

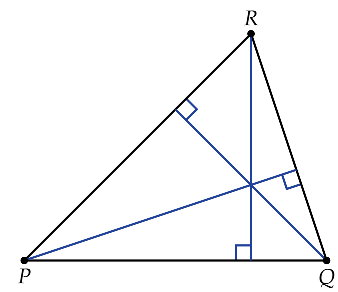

Figure 3.9: Altitudes in a triangle

Solution. Let P, Q, and R be the vertices of a triangle, and let O be the origin. A point X lies on the altitude through Q if and only if

$$\overrightarrow{QX} \cdot \overrightarrow{PR} = (\overrightarrow{OX} - \overrightarrow{OQ}) \cdot (\overrightarrow{OR} - \overrightarrow{OP}) = 0 .$$

Similarly, the point X lies on the altitude through R if and only if $(\overrightarrow{OX} - \overrightarrow{OR}) \cdot (\overrightarrow{OP} - \overrightarrow{OQ}) = 0$, and the point X lies on the altitude through P if and only if $(\overrightarrow{OX} - \overrightarrow{OP}) \cdot (\overrightarrow{OQ} - \overrightarrow{OR}) = 0$. The properties of the dot product [3.0.0] imply that, for any point X, we have

$$\begin{aligned} &(\overrightarrow{OX} - \overrightarrow{OQ}) \cdot (\overrightarrow{OR} - \overrightarrow{OP}) + (\overrightarrow{OX} - \overrightarrow{OR}) \cdot (\overrightarrow{OP} - \overrightarrow{OQ}) + (\overrightarrow{OX} - \overrightarrow{OP}) \cdot (\overrightarrow{OQ} - \overrightarrow{OR}) \\ &= (\overrightarrow{OX} \cdot \overrightarrow{OR} - \overrightarrow{OX} \cdot \overrightarrow{OP} - \overrightarrow{OQ} \cdot \overrightarrow{OR} + \overrightarrow{OP} \cdot \overrightarrow{OQ}) + (\overrightarrow{OX} \cdot \overrightarrow{OP} - \overrightarrow{OX} \cdot \overrightarrow{OQ} - \overrightarrow{OP} \cdot \overrightarrow{OR} + \overrightarrow{OQ} \cdot \overrightarrow{OR}) \\ &\qquad + (\overrightarrow{OX} \cdot \overrightarrow{OQ} - \overrightarrow{OX} \cdot \overrightarrow{OR} - \overrightarrow{OP} \cdot \overrightarrow{OQ} + \overrightarrow{OP} \cdot \overrightarrow{OR}) \\ &= 0 . \end{aligned}$$

Thus, a point X lying on two altitudes also lies on the third. □

Exercises

3.2.6 Problem. Determine which of the following statements are true. If a statement is false, then provide a counterexample.

i. The orthogonal projection is defined for any nonzero vector in $\mathbb{R}^n$.
ii. The orthogonal projection onto a nonzero vector $\vec{v}$ is parallel to $\vec{v}$.
iii. For any vector $\vec{w}$, the vector $\vec{w} - \operatorname{proj}_{\vec{v}}(\vec{w})$ is orthogonal to $\vec{v}$.
iv. The orthogonal decomposition expresses any vector as the sum of two nonzero vectors.
v. The minimum distance from a point to a line is the orthogonal distance.

vi. In any triangle, the altitude and median through a given vertex coincide.

3.2.7 Problem. Write the vector $\vec{w} := 3\vec{e}_1 + 2\vec{e}_2 - 6\vec{e}_3$ as the sum of two vectors, one parallel and one orthogonal, to $\vec{v} := 2\vec{e}_1 - 4\vec{e}_2 + \vec{e}_3$.

3.2.8 Problem. Given $P := (1,2,3)$, $Q := (3,5,7)$, and $R := (2,5,3)$, find the distance from R to the line through P and Q.

3.2.9 Problem. Consider three distinct points P, Q, and R.
i. Choose R to be the origin and describe the position vectors $\vec{\ell}(t)$ corresponding to the points on the line through P and Q as a function of a parameter t.
ii. Show that the critical points of the function $\|\vec{\ell}(t)\|$ and $\|\vec{\ell}(t)\|^2$ coincide.
iii. Using techniques from calculus, minimize $\|\vec{\ell}(t)\|^2$ and prove that the minimum distance is the orthogonal distance.

3.3 Hyperplanes

What links lines in 2-space, planes in 3-space, and their analogs in n-space? These subsets have a uniform description.

We frequently use the variables (x, y, z) in $\mathbb{K}^3$ rather than (x_1, x_2, x_3), and use (x, y) in $\mathbb{K}^2$ rather than (x_1, x_2).

3.3.0 Definition. An ***affine hyperplane*** in $\mathbb{K}^n$ consists of all points $(x_1, x_2, \dots, x_n) \in \mathbb{K}^n$ that satisfy the equation

$$a_1 x_1 + a_2 x_2 + \cdots + a_n x_n = b\,,$$

for some scalars $a_1, a_2, \dots, a_n$, and b in $\mathbb{K}$.

Geometrically, a hyperplane is determined by a point and a vector. The ***normal vector*** of the hyperplane is

$$\vec{n} := a_1\,\vec{e}_1 + a_2\,\vec{e}_2 + \cdots + a_n\,\vec{e}_n \in \mathbb{K}^n\,.$$

If the point $P := (p_1, p_2, \dots, p_n) \in \mathbb{K}^n$ lies on this hyperplane meaning that $a_1 p_1 + a_2 p_2 + \cdots + a_n p_n = b$, then an arbitrary point $X := (x_1, x_2, \dots, x_n) \in \mathbb{K}^n$ lies on this hyperplane if and only if vector $\overrightarrow{PX}$ is orthogonal to the normal vector $\vec{n}$:

$$\begin{aligned} 0 = \vec{n} \cdot \overrightarrow{PX} &= a_1(x_1 - p_1) + a_2(x_2 - p_2) + \cdots + a_n(x_n - p_n) \\ &= a_1 x_1 + a_2 x_2 + \cdots + a_n x_n + (a_1 p_1 + a_2 p_2 + \cdots + a_n p_n) \\ \Leftrightarrow \quad b &= a_1 x_1 + a_2 x_2 + \cdots + a_n x_n. \end{aligned}$$

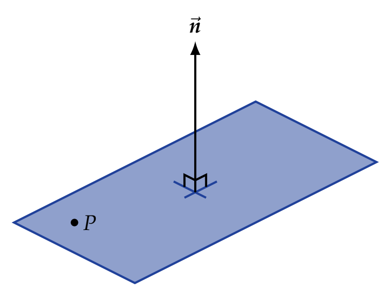

Figure 3.9: Normal to a hypersurface

Observe that the vector $-\vec{n}$ is also orthogonal to $\overrightarrow{PX}$.

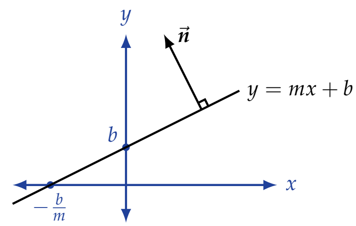

Figure 3.10: Normal to a hypersurface

3.3.1 Remark. In $\mathbb{K}^2$, the slope of a line encodes the normal vector. Specifically, the equation $y = mx + b$ is equivalent to $-mx + y = b$ which implies that $-m\,\vec{e}_1 + \vec{e}_2$ is the normal vector.

3.3.2 Problem. Find an equation for the hyperplane orthogonal to $8\,\vec{e}_1 - 3\,\vec{e}_2 - 7\,\vec{e}_4 + \vec{e}_5$ passing through the point $(1, 0, -7, 1, 3) \in \mathbb{R}^5$.

Solution. An equation is

$$8(x_1 - 1) - 3(x_2 - 0) - 0(x_3 + 7) - 7(x_4 - 1) + (x_5 - 3) = 0$$

or $8x_1 - 3x_2 - 7x_4 + x_5 = 4$. □

3.3.3 Problem. Find an equation for the plane passing through the origin and parallel to the plane $z = 4x - 3y + 8$.

Solution. Since a normal vector for both planes is $\vec{n} := 4\,\vec{e}_1 - 3\,\vec{e}_2 - \vec{e}_3$, the equation of the plane through the origin is $4x - 3y - z = 0$. □

3.3.4 Problem. Find the angle between the following two hyperplanes:

$$\begin{array}{rcccccl} 2(x_1 - 1) - (x_2 - 0) & + & (x_3 - 7) & + (x_4 + 4) & - (x_5 - 5) & + 2(x_6 + 2) & = 0 \\ -(x_1 + 1) + (x_2 + 8) & & & - (x_4 - 9) & & - \quad (x_6 - 4) & = 0\,. \end{array}$$

Solution. The angle between the hyperplanes is the angle between their normal vectors. If θ is the angle between these hyperplanes, then equivalent definitions of the dot product [3.0.0] give

$$\begin{aligned} -6 &= (2)(-1) + (-1)(1) + (1)(0) + (1)(-1) + (-1)(0) + (2)(-1) \\ &= \sqrt{12}\sqrt{4}\cos(\theta)\,, \end{aligned}$$

so we have $\cos(\theta) = \frac{-6}{4\sqrt{3}} = -\frac{\sqrt{3}}{2}$ and $\theta = \frac{5\pi}{6}$. □

A plane in $\mathbb{K}^3$ can also be determined by three points, assuming that they are not collinear.

3.3.5 Problem. Find an equation for the plane containing the three points $P := (1, 3, 0)$, $Q := (3, 4, -3)$, and $R := (3, 6, 2)$.

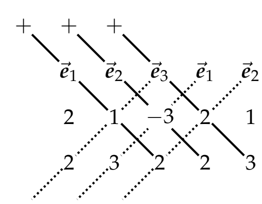

Figure 3.11: Computing the cross product

Solution. Since the points P, Q, and R lie in a plane, both of the vectors $\overrightarrow{PQ} = 2\,\vec{e}_1 + \vec{e}_2 - 3\,\vec{e}_3$ and $\overrightarrow{PR} = 2\,\vec{e}_1 + 3\,\vec{e}_2 + 2\,\vec{e}_3$ also lie in the plane. Thus, a normal vector $\vec{n}$ to the plane is given by

$$\begin{aligned} \vec{n} = \overrightarrow{PQ} \times \overrightarrow{PR} &= (2 + 9)\,\vec{e}_1 + (-6 - 4)\,\vec{e}_2 + (6 - 2)\,\vec{e}_3 \\ &= 11\,\vec{e}_1 - 10\,\vec{e}_2 + 4\,\vec{e}_3\,. \end{aligned}$$

Since the point P lies on the plane, we conclude that the equation is $11(x - 1) - 10(y - 3) + 4(z - 0) = 0$ or $10x - 10y + 4z = -19$. □

3.3.6 Problem. Find the orthogonal distance from the plane defined by $2x + 4y - z = -1$ to the point $P := (2, -1, 3)$.

Solution. From the given equation, we see that the normal vector to the plane is $\vec{n} := 2\,\vec{e}_1 + 4\,\vec{e}_2 - \vec{e}_3$. Since $2(0) + 4(0) - (1) = -1$, the point $Q := (0,0,1)$ lies in the plane and we have $\overrightarrow{QP} = 2\,\vec{e}_1 - \vec{e}_2 + 2\,\vec{e}_3$. Hence, the orthogonal distance from P to the plane is

$$\begin{aligned}\left\|\mathrm{proj}_{\vec{n}}(\overrightarrow{QP})\right\| &= \left\| \left(\frac{\overrightarrow{QP} \cdot \vec{n}}{\vec{n} \cdot \vec{n}} \right) \vec{n} \right\| = \frac{|\overrightarrow{QP} \cdot \vec{n}|}{\|\vec{n}\|} \\ &= \frac{|(2)(2) + (-1)(4) + (2)(-1)|}{\sqrt{(2)^2 + (4)^2 + (-1)^2}} = \frac{2}{\sqrt{21}}\,. \end{aligned}$$ □

3.3.7 Problem. Decide which of the three points $P := (-1,-1,1)$, $Q := (-1,-1,-1)$, and $R := (1,1,1)$ are on the same side of the plane $2x - 3y + 4z = 4$.

Solution. Since we have

$$\begin{aligned} 2(-1) - 3(-1) + 4(1) - 4 &= 1 > 0\,, \\ 2(-1) - 3(-1) + 4(-1) - 4 &= -7 < 0\,, \\ 2(1) - 3(1) + 4(1) - 4 &= -1 < 0\,, \end{aligned}$$

we see that Q and R lie on one side and P is on the other. □

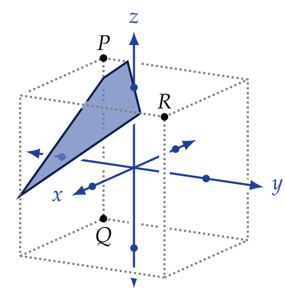

Figure 3.12: Points separated by a plane

Exercises

3.3.8 Problem. Determine which of the following statements are true. If a statement is false, then provide a counterexample.

i. Each hyperplane has a unique normal vector.
ii. A line in $\mathbb{R}^3$ is a hyperplane.
iii. A hyperplane in $\mathbb{K}^1$ consists of a single point.
iv. Lines in $\mathbb{K}^2$ with normal vector $\vec{e}_1$ have infinity slope.

3.3.9 Problem. Find a vector parallel to the line of intersection of the two planes $2x - 3y + 5 = 2$ and $4x + y - 3z = 7$.

3.3.10 Problem. Prove that the orthogonal distance between the hyperplane $a_1x_1 + a_2x_2 + \cdots + a_nx_n = b$ and the point $P := (p_1, p_2, \dots, p_n)$ is given by

$$\frac{|a_1p_1 + a_2p_2 + \cdots + a_np_n - b|}{\sqrt{a_1^2 + a_2^2 + \cdots + a_n^2}}\,.$$

3.3.11 Problem. Find the orthogonal distance between the following skew lines in $\mathbb{R}^3$. The first line passes through the points $O := (0,0,0)$ and $P := (-1,-1,1)$, and the second line passes through the points $Q := (0,-2,0)$ and $R := (2,0,5)$.

3.3.12 Problem. Find the orthogonal distance between the following skew lines in $\mathbb{R}^3$. The first line passes through the points $O := (0,0,0)$ and $P := (-1,-1,1)$, and the second line passes through the points $Q := (0,-2,0)$ and $R := (2,0,5)$.