# Inner Products and Orthonormal Decompositions

**Objective: To understand the notion of an inner product, with examples; to understand orthogonality and norm, and be able to express a vector in terms of an orthonormal basis.**

**Warm-up: Dot Product of Vectors:**

We work in three-dimensional real space, $\mathbb{R}^3$, with the usual dot product (Euclidean inner product):

$$(x_1, y_1, z_1) \cdot (x_2, y_2, z_2) = x_1 x_2 + y_1 y_2 + z_1 z_2).$$

1. Show that the three standard unit vectors $e_1 = (1, 0, 0)$, $e_2 = (0, 1, 0)$, and $e_3 = (0, 0, 1)$ satisfy the relations
$$e_i \cdot e_j = \begin{cases} 1: & i = j \\ 0: & i \neq j \end{cases}$$
2. Let $u$ be the vector $(3, -7, 4)$. Calculate $u \cdot e_i$ for $i = 1, \ldots, 3$, and express $u$ as a linear combination of $e_1$, $e_2$, and $e_3$.
3. Let $v$ be the vector $(v_1, v_2, v_3)$. Calculate $v \cdot e_i$ for $i = 1, \ldots, 3$, and express $v$ as a linear combination of $e_1$, $e_2$, and $e_3$.
4. Define three vectors:
$$f_1 = \left(\frac{1}{\sqrt{3}}, -\frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}\right)$$
$$f_2 = \left(2\sqrt{\frac{2}{21}}, -\frac{1}{\sqrt{42}}, -\frac{5}{\sqrt{42}}\right)$$
$$f_3 = \left(\sqrt{\frac{2}{7}}, \frac{3}{\sqrt{14}}, \frac{1}{\sqrt{14}}\right).$$
Show that
$$f_i \cdot f_j = \begin{cases} 1: & i = j \\ 0: & i \neq j \end{cases}$$
5. Let $u$, $v$ be as above. Express $u$ and $v$ as linear combinations of $f_1$, $f_2$, and $f_3$.

**Theory: Inner Products and Orthogonality:**

Let $V$ be a real vector space (*i.e.*, a collection of objects, called **vectors**, which can be added together and multiplied by real scalars. An **inner product** on $V$ is a function that takes two vectors $u$ and $v$ and returns a real number, often denoted $\langle u \mid v \rangle$, satisfying the following three conditions, for any vectors $u$, $v$, and $w$, and real scalars $\lambda$ and $\mu$:

$$\begin{aligned}
\langle \lambda u + \mu v \mid w \rangle &= \lambda \langle u \mid w \rangle + \mu \langle v \mid w \rangle && \text{linearity in 1}^{\text{st}} \text{ argument} \\
\langle u \mid v \rangle &= \langle v \mid u \rangle && \text{symmetry} \\
\langle u \mid u \rangle &> 0 \text{ if } u \neq 0 && \text{positive-definiteness.}
\end{aligned}$$

Show that $\langle 0 \mid 0 \rangle = 0$.

There is also a notion of inner product on a complex vector space (when scalars can be complex numbers, instead of just real). There, the symmetry condition must be replaced by **conjugate symmetry**:

$$\langle u \mid v \rangle = \overline{\langle v \mid u \rangle},$$

where bar denotes complex conjugation.

Given an inner product, we can define the **norm** of a vector $v$ to be

$$\|v\| = \sqrt{\langle v \mid v \rangle}.$$

By positive-definiteness, $\|v\| > 0$ whenever $v \neq 0$; moreover, $\|0\| = \sqrt{\langle 0 \mid 0 \rangle} = 0$. So a vector has norm 0 if and only if it is the zero vector. A **unit vector** is any vector of norm 1; given any non-zero vector $v$, we can scale $v$ to get a unit vector:

$$\hat{v} = \frac{v}{\|v\|}.$$

Show that for any $v \neq 0$,

$$\|\hat{v}\| = 1.$$

We say vectors $v_1, \ldots, v_n$ are **orthogonal** if $\langle v_i \mid v_j \rangle = 0$ whenever $i \neq j$. If also each $v_i$ is a unit vector, we say these vectors are **orthonormal**. We define the **Kronecker delta symbol** $\delta_{ij}$ by

$$\delta_{ij} = \begin{cases} 1: & i = j \\ 0: & i \neq j. \end{cases}$$

Then $v_1, \ldots, v_n$ are orthonormal if and only if

$$\langle v_i \mid v_j \rangle = \delta_{ij}.$$

**Practice:**

1. Show that the dot product of vectors is an inner product on $\mathbb{R}^n$.

2. Show that any inner product is linear in the **second** argument:
$$\langle u \mid \lambda v + \mu w \rangle = \lambda \langle u \mid v \rangle + \mu \langle u \mid w \rangle .$$

3. Prove the important **Cauchy-Schwarz inequality**:
$$|\langle u \mid v \rangle| \leq \|u\| \times \|v\| .$$
Hint: consider
$$\left\| u - \frac{\langle u \mid v \rangle}{\|v\|^2} v \right\|^2 .$$

4. Let $F(\mathbb{R})$ be the set of (real-valued) functions on $\mathbb{R}$. Let $a \in \mathbb{R}$ be any fixed real number. Does the rule
$$\langle f \mid g \rangle = f(a)g(a)$$
defines an inner product on $F(\mathbb{R})$?

5. Let $L_2([0, 2\pi])$ be the set of square-integrable functions on the interval $[0, 2\pi]$—*i.e.*, functions $f$ from $[0, 2\pi]$ to $\mathbb{R}$ such that
$$\int_0^{2\pi} f(x)^2 \, \mathrm{d}x$$
exists and is finite. Show that the rule
$$\langle f \mid g \rangle = \frac{1}{\pi} \int_0^{2\pi} f(x)g(x) \, \mathrm{d}x$$
defines an inner product on $L_2([0, 2\pi])$.

6. Consider $L_2([0, 2\pi])$ with the inner product defined in question 4. Show that the functions $\cos(nx)$ for all different positive integer values of $n$ are orthonormal. That is, show that
$$\langle \cos(nx) \mid \cos(mx) \rangle = \delta_{nm}.$$

7. Let $V$ be a real vector space with an inner product. Suppose $v_1, \ldots, v_n$ are orthonormal, and
$$v = \sum_{i=1}^{n} \lambda_i v_i$$
for some scalars $\lambda_i \in \mathbb{R}$. Show that
$$\lambda_i = \langle v \mid v_i \rangle .$$

**Application: Orthonormal Approximations:**

We have seen that if $v_1, \ldots, v_n$ are orthonormal with respect to some inner product and $v$ is a **linear combination** of $v_1, \ldots, v_n$,

$$v = \sum_{i=1}^{n} \lambda_i v_i,$$

then the coefficients are given by inner products: $\lambda_i = \langle v \mid v_i \rangle$.

What if $v$ is *not* a linear combination of $v_1, \ldots, v_n$? Then we cannot hope to express $v$ exactly as a linear combination of $v_1, \ldots, v_n$, but we can still consider *approximating* $v$ by a linear combination of $v_1, \ldots, v_n$. One might expect that the same choice of coefficients will give the best approximation to $v$, and indeed we shall show this. But what do we mean by "best" approximation? We want to define a notion of distance, and then measure the error of an approximation $u$ by the distance from $u$ to $v$.

A **metric** (distance function) on a set is a function $d$ taking two inputs and giving a real output, $d(x, y) \in \mathbb{R}$, satisfying the following three conditions for any $x$, $y$, and $z$:

$$\begin{array}{ll} d(x,y) = d(y,x) & \text{symmetry} \\ d(x,y) \geq 0, \text{ with } > \text{ if } x \neq y & \text{positive-definiteness} \\ d(x,z) \leq d(x,y) + d(y,z) & \text{the triangle inequality.} \end{array}$$

Show that, on an inner product space, defining

$$d(u, v) = \|u - v\|$$

satisfies the conditions of a metric. Hint: you will need the Cauchy-Schwarz inequality (exercise 3 on the preceding page).

So if $u$ is an approximation of $v$, then the error of the approximation will be defined to be $\|v - u\|$. So we want to show that choosing $\lambda_i = \langle v \mid v_i \rangle$ minimises the error

$$\left\| v - \sum_{i=1}^{n} \lambda_i v_i \right\|.$$

**Application: Orthonormal Approximations (cont.):**

Let $v_1, \ldots, v_n$ be orthonormal and let $v$ be a function we wish to approximate. We prove that the linear combination

$$u = \sum_{i=1}^{n} \lambda_i v_i$$

which minimises $\|v - u\|$ (*i.e.* gives the best approximation) is the one where $\lambda_i = \langle v \mid v_i \rangle$ for each $i$.

1. First we show that

$$\left\| v - \sum_{i=1}^{n} \lambda_i v_i \right\|^2 = \|v\|^2 - 2 \sum_{i=1}^{n} \lambda_i \langle v \mid v_i \rangle + \sum_{i=1}^{n} \lambda_i^2. \qquad (\star)$$

   (a) First prove that for any vector $w$ and any integer $k$ between 1 and $n$:

   $$\|w - \lambda_k v_k\|^2 = \|w\|^2 - 2\lambda_k \langle w \mid v_k \rangle + \lambda_k^2.$$

   (b) Take $w = v$ and $k = 1$ to prove Equation $(\star)$ in the case $n = 1$.

   (c) Now take

   $$w = v - \sum_{i=1}^{k-1} \lambda_i v_i$$

   and show that, for this $w$, $\langle w \mid v_i \rangle = \langle v \mid v_i \rangle$. Hence conclude that if Equation $(\star)$ is true for $n = k - 1$, then it is true for $n = k$.

   (d) Now we know that $(\star)$ is true for $n = 1$, and if true for $n = k - 1$, is also true for $n = k$. Why does this mean it is true for all $n$?

2. By Equation $(\star)$, in order to minimise

$$\left\| v - \sum_{i=1}^{n} \lambda_i v_i \right\|^2,$$

   it suffices to minimise $\lambda_i^2 - 2\lambda_i \langle v \mid v_i \rangle$ for each value of $i$. Why does minimising the square of the norm also minimise the norm itself?

3. Use Fermat's Method to show that $\lambda_i^2 - 2\lambda_i \langle v \mid v_i \rangle$ has its minimum value when $\lambda_i = \langle v \mid v_i \rangle$. This completes the proof.

**Practice:**

1. Consider 3D real space, $\mathbb{R}^3$, with the usual dot product.
   (a) Show that the vectors
   $$v_1 = \left(\frac{1}{\sqrt{2}}, 0, \frac{1}{\sqrt{2}}\right), \quad v_2 = \left(\frac{1}{\sqrt{2}}, 0, \frac{-1}{\sqrt{2}}\right)$$
   are orthonormal.
   (b) Let $v$ be the point $(-1, 7, 4)$. Find the best approximation to $v$ by a linear combination of $v_1$ and $v_2$, and the error in this approximation.

2. Consider $L_2([0, 2\pi])$, the space of square-integrable functions from $[0, 2\pi]$ to $\mathbb{R}$, with the inner product given by
   $$\langle f \mid g \rangle = \frac{1}{\pi} \int_0^{2\pi} f(x)g(x)\,\mathrm{d}x.$$
   (a) Show that the functions $\sin(nx)$ are orthonormal for different positive integer values of $n$. That is, show that
   $$\langle \sin(nx) \mid \sin(mx) \rangle = \delta_{nm}.$$
   (b) Let $f(x)$ be the function defined by
   $$f(x) = \begin{cases} 1: & 0 \le x < \pi \\ -1: & \pi \le x \le 2\pi. \end{cases}$$
   Find $\langle f(x) \mid \sin(nx) \rangle$ in terms of $n$.
   (c) Hence write down an expression for the best approximation to $f(x)$ by a linear combination of sine waves $\sin(nx)$ for $1 \le n \le N$. This is called the $N^{\text{th}}$ partial Fourier series of $f(x)$.

**Key Points to Remember:**

1. An **inner product** is a pairing which takes two vectors and gives a real number output, satisfying the following three conditions:

$$\begin{aligned} \langle \lambda u + \mu v \mid w \rangle &= a \langle u \mid w \rangle + b \langle v \mid w \rangle && \text{linearity in } 1^{\text{st}} \text{ argument} \\ \langle u \mid v \rangle &= \langle v \mid u \rangle && \text{symmetry} \\ \langle u \mid u \rangle &> 0 \text{ if } u \neq 0 && \text{positive-definiteness.} \end{aligned}$$

2. Given an inner product, the **norm** of a vector $v$ is $\|v\| = \sqrt{\langle v \mid v \rangle}$. This is 0 if $v = 0$, and otherwise is strictly positive.

3. The **distance** between two vectors $u$ and $v$ is the norm of their difference, $\|u - v\|$.

4. A **linear combination** of vectors $v_1, \ldots, v_n$ is any expression of the form

$$\sum_{i=1}^{n} \lambda_i v_i,$$

where the $\lambda_i$ are scalars.

5. We say vectors $v_1, \ldots, v_n$ are **orthogonal** if $\langle v_i \mid v_j \rangle = 0$ for all $i \neq j$. If also $\|v_i\| = 1$ for all $i$, we say they are **orthonormal**. We can concisely combine the two conditions for orthonormality by writing $\langle v_i \mid v_j \rangle = \delta_{ij}$.

6. Given $v_1, \ldots, v_n$ orthonormal, and another object $v$, we can find the best approximation to $v$ by a linear combination of the $v_i$; it is

$$\sum_{i=1}^{n} \langle v \mid v_i \rangle v_i.$$