Chapter 13

Orthonormal Bases

You probably recall the dot product from Calculus class, but a brief review won't hurt.

Definition. Let $\boldsymbol{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$, $\boldsymbol{y} = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} \in \mathbb{R}^n$. Then the ***dot product*** of $\boldsymbol{x}$ and $\boldsymbol{y}$ is

$$\boldsymbol{x} \cdot \boldsymbol{y} = x_1 y_1 + x_2 y_2 + \cdots + x_n y_n.$$

□

Example. Let $\boldsymbol{x} = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}$ and $\boldsymbol{y} = \begin{pmatrix} 2 \\ 1 \\ 4 \end{pmatrix}$. Then

$$\boldsymbol{x} \cdot \boldsymbol{y} = 1 \times 2 + 2 \times 1 + 3 \times 4 = 16.$$

☑

Note that we can also write the dot product as a matrix product; given $\boldsymbol{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$ and $\boldsymbol{y} = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}$, the matrix product

$$\begin{aligned} \boldsymbol{x}^T \boldsymbol{y} &= \begin{pmatrix} x_1 & x_2 & \dots & x_n \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} \\ &= x_1 y_1 + x_2 y_2 + \cdots + x_n y_n, \end{aligned}$$

and so the matrix product $\boldsymbol{x}^T\boldsymbol{y}$ is simply the dot product.

The dot product is often written as $\langle \boldsymbol{x}, \boldsymbol{y} \rangle$. Since this notation no longer uses a dot, it is also called the ***inner product***. The inner product has the following properties.

Proposition. *For any $\boldsymbol{x}$, $\boldsymbol{y}$, $\boldsymbol{z} \in \mathbb{R}^n$, and any number a:*

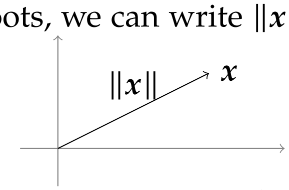
- $\langle \boldsymbol{x}, \boldsymbol{y} \rangle = \langle \boldsymbol{y}, \boldsymbol{x} \rangle$.
- $\langle \boldsymbol{x} + \boldsymbol{y}, \boldsymbol{z} \rangle = \langle \boldsymbol{x}, \boldsymbol{z} \rangle + \langle \boldsymbol{y}, \boldsymbol{z} \rangle$.
- $\langle a\boldsymbol{x}, \boldsymbol{y} \rangle = a\langle \boldsymbol{x}, \boldsymbol{y} \rangle$.
- $\langle \boldsymbol{x}, \boldsymbol{x} \rangle \geq 0$, *and* $\langle \boldsymbol{x}, \boldsymbol{x} \rangle = 0$ *only if* $\boldsymbol{x} = 0$. □

Remark. Any function $\langle \cdot, \cdot \rangle$ of two vectors which satisfies the properties of the previous proposition is called an inner product. We will stick to using the standard inner product, the one defined above, even though much of what follows would also be true if we used a different inner product. ⊠

In terms of the inner product, the ***length*** of a vector $\boldsymbol{x}$ is given as

$$\|x\| = \sqrt{\langle \boldsymbol{x}, \boldsymbol{x} \rangle}.$$

This is simply the distance from the point $\boldsymbol{x}$ to the origin, or put another way, the length of the arrow representing the vector $\boldsymbol{x}$. To avoid square roots, we can write $\|\boldsymbol{x}\|^2 = \langle \boldsymbol{x}, \boldsymbol{x} \rangle$.



Note that if $\boldsymbol{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$, then $\|\boldsymbol{x}\| = \sqrt{x_1^2 + x_2^2 + \cdots + x_n^2}$.

Vectors are often used to indicate directions in $\mathbb{R}^n$; in this case, the length of the vector isn't relevant. If a direction is to be indicated, it is often useful to use a vector with length 1. A vector with length 1 is called a ***unit vector***.

Definition. If $v \neq 0$ is a vector in $\mathbb{R}^n$, then

$$u = \frac{v}{\|v\|}$$

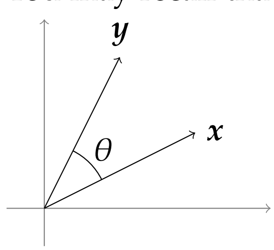
is the ***unit vector in the direction of*** v. $\square$

Example. Let $v = \begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix}$. Since $\|v\| = \sqrt{1^2 + 2^2 + 1^2} = \sqrt{6}$, the unit vector in the direction of v is

$$u = \frac{1}{\sqrt{6}} \begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix}.$$

☑

You may recall that if θ is the angle between vectors x and y,



then the dot product will be

$$x \cdot y = \|x\| \|y\| \cos(\theta).$$

When the angle between the vectors is a right angle, the $\cos(\theta)$ will be 0.

Definition. Two vectors x and y are (orthogonal) (or perpendicular) if $x \cdot y = 0$. $\square$

Example. The vectors $x = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$ and $y = \begin{pmatrix} 1 \\ -1 \end{pmatrix}$ are orthogonal, since $x \cdot y = 1 - 1 = 0$. ☑

Example. The vectors $x = \begin{pmatrix} 2 \\ 1 \end{pmatrix}$ and $y = \begin{pmatrix} 1 \\ -1 \end{pmatrix}$ are *not* orthogonal, since $x \cdot y = 2 - 1 \neq 0$. ☑

Definition. A set of vectors is an ***orthogonal set*** if every pair of vectors in the set are orthogonal. □

Example. Let $\boldsymbol{x} = \begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \end{pmatrix}$, $\boldsymbol{y} = \begin{pmatrix} -1 \\ 0 \\ 1 \\ 1 \end{pmatrix}$ and $\boldsymbol{z} = \begin{pmatrix} 1 \\ 0 \\ -1 \\ 2 \end{pmatrix}$. Then $\{\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z}\}$ is an orthogonal set since

$$\begin{aligned} \boldsymbol{x} \cdot \boldsymbol{y} &= 0 \\ \boldsymbol{x} \cdot \boldsymbol{z} &= 0 \\ \boldsymbol{y} \cdot \boldsymbol{z} &= 0 \end{aligned}$$

☑

If a vector $\boldsymbol{v}$ is a linear combination of elements of an orthogonal set, then it is easy to pick out the coefficients. If $\{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n\}$ is an orthogonal set and

$$\boldsymbol{v} = c_1 \boldsymbol{v}_1 + c_2 \boldsymbol{v}_2 + \cdots + c_n \boldsymbol{v}_n,$$

then by taking the dot product of both sides of the above equation with $\boldsymbol{v}_1$, we get

$$\begin{aligned} \boldsymbol{v} \cdot \boldsymbol{v}_1 &= (c_1 \boldsymbol{v}_1 + c_2 \boldsymbol{v}_2 + \cdots + c_n \boldsymbol{v}_n) \cdot \boldsymbol{v}_1 \\ &= c_1 \boldsymbol{v}_1 \cdot \boldsymbol{v}_1 + c_2 \boldsymbol{v}_2 \cdot \boldsymbol{v}_1 + \ldots c_n \boldsymbol{v}_n \cdot \boldsymbol{v}_1 \\ &= c_1 \|\boldsymbol{v}_1\|^2 + c_2 0 + \ldots c_n 0 \\ &= c_1 \|\boldsymbol{v}_1\|^2, \end{aligned}$$

and so $c_1 = \boldsymbol{v} \cdot \boldsymbol{v}_1 / \|\boldsymbol{v}_1\|^2$. Similarly, $c_2 = \boldsymbol{v} \cdot \boldsymbol{v}_2 / \|\boldsymbol{v}_2\|^2$, etc.

As an application of this, we get the following.

Proposition. *Suppose $\{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n\}$ is an orthogonal set of non-zero vectors. Then it is an independent set of vectors.* □

To see why this is true, suppose that $\{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n\}$ is an orthogonal set of non-zero vectors and

$$c_1 \boldsymbol{v}_1 + c_2 \boldsymbol{v}_2 + \cdots + c_n \boldsymbol{v}_n = 0.$$

Then $c_1 = 0 \cdot \boldsymbol{v}_1 / \|\boldsymbol{v}_1\|^2 = 0$, and similarly $c_2 = 0, \ldots$.

Definition. An ***orthogonal basis*** for $\mathbb{R}^n$ is an orthogonal set which is also a basis for $\mathbb{R}^n$. $\square$

Proposition. *Let $\{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n\}$ be an orthogonal basis for $\mathbb{R}^n$. For $\boldsymbol{x} \in \mathbb{R}^n$, we have*

$$\boldsymbol{x} = c_1 \boldsymbol{v}_1 + \cdots + c_n \boldsymbol{v}_n$$

for $c_1 = \frac{\boldsymbol{v} \cdot \boldsymbol{v}_1}{\|\boldsymbol{v}_1\|^2}$, ... $c_n = \frac{\boldsymbol{v} \cdot \boldsymbol{v}_n}{\|\boldsymbol{v}_n\|^2}$. $\square$

Example. Let $\boldsymbol{v}_1 = \begin{pmatrix} 2 \\ 2 \\ 1 \end{pmatrix}$, $\boldsymbol{v}_2 = \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}$ and $\boldsymbol{v}_3 = \begin{pmatrix} 1 \\ 1 \\ -4 \end{pmatrix}$. Then $\{\boldsymbol{v}_1, \boldsymbol{v}_2, \boldsymbol{v}_3\}$ is an orthogonal basis for $\mathbb{R}^3$.

Let $\boldsymbol{v} = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}$. The we can write

$$\boldsymbol{v} = c_1 \boldsymbol{v}_1 + c_2 \boldsymbol{v}_2 + c_3 \boldsymbol{v}_3.$$

Since

$$\begin{aligned} \boldsymbol{v} \cdot \boldsymbol{v}_1 &= 9 \\ \|\boldsymbol{v}_1\|^2 &= 9 \\ \boldsymbol{v} \cdot \boldsymbol{v}_2 &= -1 \\ \|\boldsymbol{v}_2\|^2 &= 2 \\ \boldsymbol{v} \cdot \boldsymbol{v}_3 &= -9 \\ \|\boldsymbol{v}_3\|^2 &= 18, \end{aligned}$$

we get

$$\begin{aligned} c_1 &= 9/9 = 1 \\ c_2 &= -1/2 \\ c_3 &= -9/18 = -1/2. \end{aligned}$$

So

$$\boldsymbol{v} = 1\boldsymbol{v}_1 + \frac{-1}{2}\boldsymbol{v}_2 + \frac{-1}{2}\boldsymbol{v}_3.$$

If the elements of an orthogonal set have length 1, then formula for the coefficients in the expansion formula

$$\boldsymbol{v} = c_1\boldsymbol{v}_1 + c_2\boldsymbol{v}_2 + \cdots + c_n\boldsymbol{v}_n$$

takes on a slightly simpler form. For a general orthogonal basis, $c_i = \frac{\boldsymbol{v}\cdot\boldsymbol{v}_i}{\|\boldsymbol{v}_i\|^2}$. If the basis elements have length 1, then this formula becomes $c_i = \boldsymbol{v} \cdot \boldsymbol{v}_i$.

Definition. An ***orthonormal basis*** for $\mathbb{R}^n$ is an orthogonal basis for which all elements have length 1. □

Note that if $\{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n\}$ is an orthogonal basis for $\mathbb{R}^n$, then $\{\boldsymbol{u}_1, \ldots, \boldsymbol{u}_n\}$ is an orthonormal basis, where $\boldsymbol{u}_i = \boldsymbol{v}_i / \|\boldsymbol{v}_i\|$ for $i = 1, \ldots, n$.

Example. ***Let*** $\boldsymbol{v}_1 = \begin{pmatrix} 2 \\ 2 \\ 1 \end{pmatrix}$, $\boldsymbol{v}_2 = \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}$ ***and*** $\boldsymbol{v}_3 = \begin{pmatrix} 1 \\ 1 \\ -4 \end{pmatrix}$. ***Then*** $\{\boldsymbol{v}_1, \boldsymbol{v}_2, \boldsymbol{v}_3\}$ ***is an orthogonal basis for*** $\mathbb{R}^3$.

Find the corresponding orthonormal basis.

Since

$$\begin{aligned}
\|\boldsymbol{v}_1\| &= \sqrt{2^2 + 2^2 + 1^2} = 3 \\
\|\boldsymbol{v}_2\| &= \sqrt{1^2 + (-1)^2 + 0^2} = \sqrt{2} \\
\|\boldsymbol{v}_3\| &= \sqrt{1^2 + 1^2 + (-4)^2} = 3\sqrt{2}
\end{aligned}$$

our orthonormal basis is $\{\boldsymbol{u}_1, \boldsymbol{u}_2, \boldsymbol{u}_3\}$ where

$$\begin{aligned}
\boldsymbol{u}_1 &= \frac{1}{3}\boldsymbol{v}_1 = \frac{1}{3}\begin{pmatrix} 2 \\ 2 \\ 1 \end{pmatrix} \\
\boldsymbol{u}_2 &= \frac{1}{\sqrt{2}}\boldsymbol{v}_2 = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix} \\
\boldsymbol{u}_3 &= \frac{1}{3\sqrt{2}}\boldsymbol{v}_3 = \frac{1}{3\sqrt{2}}\begin{pmatrix} 1 \\ 1 \\ -4 \end{pmatrix}
\end{aligned}$$

is an orthogonal

takes on the simpler form If a vector $\boldsymbol{v}$ is a linear combination of elements of an orthogonal set, then it is easy to pick out the coefficients. If $\{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n\}$ is an orthogonal set and

$$\boldsymbol{v} = c_1\boldsymbol{v}_1 + c_2\boldsymbol{v}_2 + \cdots + c_n\boldsymbol{v}_n,$$

then by taking the dot product of both sides of the above equation with $\boldsymbol{v}_1$, we get