23. Orthogonal and orthonormal bases

Lemma 23.1. *If $\vec{v}_1, \vec{v}_2, \dots, \vec{v}_p$ are non-zero pairwise orthogonal vectors in $\mathbb{R}^n$ then $\vec{v}_1, \vec{v}_2, \dots, \vec{v}_p$ are linearly independent.*

Proof. Suppose that there are scalars $x_1, x_2, \dots, x_p$ such that

$$x_1\vec{v}_1 + x_2\vec{v}_2 + \cdots + x_p\vec{v}_p = \vec{0}.$$

Dot both sides with respect to $\vec{v}_1$:

$$\vec{v}_1 \cdot (x_1\vec{v}_1 + x_2\vec{v}_2 + \cdots + x_p\vec{v}_p) = \vec{v}_1 \cdot \vec{0}.$$

Distributing the parentheses we have

$$x_1\vec{v}_1 \cdot \vec{v}_1 + x_2\vec{v}_1 \cdot \vec{v}_2 + \cdots + x_p\vec{v}_1 \cdot \vec{v}_p = 0.$$

By assumption

$$\vec{v}_1 \cdot \vec{v}_2 = \vec{v}_1 \cdot \vec{v}_3 = \cdots = \vec{v}_1 \cdot \vec{v}_p = 0.$$

So we get

$$x_1(\vec{v}_1 \cdot \vec{v}_1) = 0.$$

$\vec{v}_1$ is not the zero vector and so the length of $\vec{v}_1$ is non-zero. So

$$x_1 = 0.$$

Similarly for all of the other scalars

$$x_2 = x_3 = \cdots = x_p = 0.$$

Thus $\vec{v}_1, \vec{v}_2, \dots, \vec{v}_p$ are linearly independent. □

Definition 23.2. *$\vec{v}_1, \vec{v}_2, \dots, \vec{v}_n$ is an orthogonal basis if $\vec{v}_1, \vec{v}_2, \dots, \vec{v}_n$ is a basis and $\vec{v}_1, \vec{v}_2, \dots, \vec{v}_n$ are pairwise orthogonal.*

Example 23.3. *Are $\vec{v}_1 = (1, 1, 1)$, $\vec{v}_2 = (-2, 1, 1)$, and $\vec{v}_3 = (0, 1, -1)$ an orthogonal basis of $\mathbb{R}^3$?*

We check that these vectors are pairwise orthogonal:

$$\begin{aligned}
\vec{v}_1 \cdot \vec{v}_2 &= (1, 1, 1) \cdot (-2, 1, 1) = -2 + 1 + 1 = 0 \\
\vec{v}_1 \cdot \vec{v}_3 &= (1, 1, 1) \cdot (0, 1, -1) = 1 - 1 = 0 \\
\vec{v}_2 \cdot \vec{v}_3 &= (-2, 1, 1) \cdot (0, 1, -1) = 1 - 1 = 0.
\end{aligned}$$

Thus $\vec{v}_1$, $\vec{v}_2$ and $\vec{v}_3$ are pairwise orthogonal. By (23.1) they are linearly independent. As we have three independent vectors in $\mathbb{R}^3$ they are a basis. So they are an orthogonal basis.

If $\vec{b}$ is any vector in $\mathbb{R}^3$ then we can write $\vec{b}$ as a linear combination of $\vec{v}_1$, $\vec{v}_2$ and $\vec{v}_3$:

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

In general to find the scalars c_1, c_2 and c_3 there is nothing for it but to solve some linear equations. However it is must easier if we use the fact that $\vec{v}_1$, $\vec{v}_2$ and $\vec{v}_3$ are orthogonal.

For example, suppose that $\vec{b} = (5, 3, 1)$. If we dot both sides with $\vec{v}_1$ we get:

$$\begin{aligned}\vec{v}_1 \cdot \vec{b} &= c_1\vec{v}_1 \cdot \vec{v}_1 + c_2\vec{v}_1 \cdot \vec{v}_2 + c_3\vec{v}_1 \cdot \vec{v}_3 \\ &= c_1\vec{v}_1 \cdot \vec{v}_1 \\ &= 3c_1.\end{aligned}$$

Thus

$$c_1 = \frac{(1,1,1) \cdot (5,3,1)}{3} = \frac{9}{3} = 3.$$

Similarly

$$c_2 = \frac{(-2,1,1) \cdot (5,3,1)}{(-2,1,1) \cdot (-2,1,1)} = \frac{-6}{6} = -1.$$

Finally

$$c_3 = \frac{(0,1,-1) \cdot (5,3,1)}{(0,1,-1) \cdot (0,1,-1)} = \frac{2}{2} = 1.$$

In general if we are given an orthogonal basis $\vec{v}_1, \vec{v}_2, \dots, \vec{v}_n$, a vector $\vec{b}$ and we write

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

then

$$c_i = \frac{\vec{v}_i \cdot \vec{b}}{\vec{v}_i \cdot \vec{v}_i}.$$

To see this dot both sides of the previous equation with respect to $\vec{v}_i$.

Definition 23.4. *$\vec{v}_1, \vec{v}_2, \dots, \vec{v}_n$ is an orthonormal basis if $\vec{v}_1, \vec{v}_2, \dots, \vec{v}_n$ is a basis, $\vec{v}_1, \vec{v}_2, \dots, \vec{v}_n$ are pairwise orthogonal and they have unit length.*

$\vec{e}_1, \vec{e}_2, \dots, \vec{e}_n$ is an orthonormal basis.

Example 23.5. *Is*

$$\vec{v}_1 = \frac{1}{\sqrt{2}}(1,1) \qquad \text{and} \qquad \vec{v}_2 = \frac{1}{\sqrt{2}}(-1,1)$$

an orthonormal basis of $\mathbb{R}^2$?

$$\vec{v}_1 \cdot \vec{v}_2 = 0,$$

so $\vec{v}_1$ and $\vec{v}_2$ are orthogonal. Therefore they are independent. But then they are an orthogonal basis of $\mathbb{R}^2$.

$$\vec{v}_1 \cdot \vec{v}_1 = \frac{1}{2}(1,1) \cdot (1,1) = 1.$$

So the length of $\vec{v}_1$ is one, as well. Similary $\vec{v}_2$ has unit length. Thus $\vec{v}_1$ and $\vec{v}_2$ are an orthonormal basis.

Let

$$A = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix}$$

be the matrix whose columns are the vectors $\vec{v}_1$ and $\vec{v}_2$. Consider

$$A^T = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix}.$$

Let's compute the product

$$A^T A = \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix} \frac{1}{2} \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix} = \frac{1}{2} \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix} = \frac{1}{2} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

It follows that the inverse of A is the transpose of A,

$$A^{-1} = A^T.$$

Definition 23.6. *Let A be a square matrix. We say that A is orthogonal if the columns of A are an orthonormal basis.*

Theorem 23.7. *Let A be a square matrix.*

Then A is orthogonal if and only if $A^{-1} = A^T$.

There isn't much to the proof of (23.7) it follows from the definition of an orthogonal matrix (23.6). It is probably best just to give an example.

Let's start with the vectors $\vec{v}_1 = (1, 1, 1)$, $\vec{v}_2 = (-2, 1, 1)$, and $\vec{v}_3 = (0, 1, 1)$. We already saw that these are an orthogonal basis. If we replace them by unit vectors we get an orthonormal basis.

The length of $\vec{v}_1$ is $\sqrt{3}$, the length of $\vec{v}_2$ is $\sqrt{6}$ and the length of $\vec{v}_3$ is $\sqrt{2}$. So

$$\hat{u}_1 = \frac{1}{\sqrt{3}}(1, 1, 1) \qquad \hat{u}_2 = \frac{1}{\sqrt{6}}(-2, 1, 1) \qquad \text{and} \qquad \hat{u}_3 = \frac{1}{\sqrt{2}}(0, 1, -1)$$

is an orthonormal basis.

Let

$$A = \frac{1}{\sqrt{6}} \begin{pmatrix} \sqrt{2} & -2 & 0 \\ \sqrt{2} & 1 & \sqrt{3} \\ \sqrt{2} & 1 & -\sqrt{3} \end{pmatrix}.$$

Then

$$A^T = \frac{1}{\sqrt{6}} \begin{pmatrix} \sqrt{2} & \sqrt{2} & \sqrt{2} \\ -2 & 1 & 1 \\ 0 & \sqrt{3} & -\sqrt{3} \end{pmatrix}.$$

We calculate $A^T A$:

$$\frac{1}{6}\begin{pmatrix}\sqrt{2} & \sqrt{2} & \sqrt{2}\\ -2 & 1 & 1\\ 0 & \sqrt{3} & -\sqrt{3}\end{pmatrix}\begin{pmatrix}\sqrt{2} & -2 & 0\\ \sqrt{2} & 1 & \sqrt{3}\\ \sqrt{2} & 1 & -\sqrt{3}\end{pmatrix} = \frac{1}{6}\begin{pmatrix}6 & 0 & 0\\ 0 & 6 & 0\\ 0 & 0 & 6\end{pmatrix} = I_3,$$

as expected.