Orthogonal and orthonormal bases

What you need to know already:	*What you can learn here:*
➤ What a basis for a subspace is. ➤ When two vectors are orthogonal.	➤ Two special types of bases that have many convenient properties.

We have seen earlier that any subspace is the span of some set of vectors in it. We have also seen that some spanning sets are too big, in the sense that they contain vectors that are not needed. This has led us to the idea of a *basis*, that is, a minimal spanning set. Still, not all bases are created equal. In this section we shall explore two features that make some bases, in the immortal words of George Orwell, "*more equal than others*".

For instance, you are used to the standard basis of $\mathbb{R}^3$ that consists of $\{\mathbf{i}, \mathbf{j}, \mathbf{k}\}$, but have you wondered what is so special about it?

Yes: it is simple, since each of those vectors only has a 1 and two 0's as components.

Yes, but that itself is not what makes this basis better. As a basis, its real advantage should be in how other vectors are related to it, that is, in terms of coordinates. Notice that the coordinates of any vector in this basis are exactly the components of the vector itself.

Now the question is: why are the coordinates so simple in this basis? You may think that this is because we built all basic mathematics on this choice, but that is a cop out. The real reason is what you shall see in this section. Let's start from the most basic property.

Definition

A basis $B = \{\mathbf{v}_1, \cdots, \mathbf{v}_k\}$ of a subspace $\mathbb{S}$ of $\mathbb{R}^n$ (including the whole space) is said to be ***orthogonal*** if any two vectors in it are orthogonal to each other.

Notice that $\{\mathbf{i}, \mathbf{j}, \mathbf{k}\}$ is an orthogonal basis and that is one reason why it is better.

And why is this property good to have for a basis?

The main advantage of orthogonal bases is indeed related to the coordinates of any vector in that basis. We know that any vector can be written in a unique way as a linear combination of the elements of a basis, but how do we get such coordinates and what is their meaning? We know that we can get those coordinates by using a routine system, but the meaning of the coordinates may be rather obscure.

For orthogonal bases we have a clear and useful answer.

Technical fact

If $B=\{\mathbf{v}_1,\cdots,\mathbf{v}_k\}$ is an orthogonal basis of a subspace $\mathbb{S}$ of $\mathbb{R}^n$, then any vector $\mathbf{w}$ of $\mathbb{S}$ is the sum of its projections on each basis vector:

$$\mathbf{w}=\frac{\mathbf{v}_1\bullet\mathbf{w}}{\mathbf{v}_1\bullet\mathbf{v}_1}\mathbf{v}_1+\frac{\mathbf{v}_2\bullet\mathbf{w}}{\mathbf{v}_2\bullet\mathbf{v}_2}\mathbf{v}_2+\ldots+\frac{\mathbf{v}_k\bullet\mathbf{w}}{\mathbf{v}_k\bullet\mathbf{v}_k}\mathbf{v}_k$$

Proof

We know that each vector **w** has a unique set of coordinates in the given basis:

$$\mathbf{w}=w_1\mathbf{v}_1+w_2\mathbf{v}_2+\cdots+w_k\mathbf{v}_k$$

To see what these coordinates are, we can take the dot product of **w** with $\mathbf{v}_1$ and use the fact that the basis vectors are all orthogonal to each other:

$$\mathbf{w}\bullet\mathbf{v}_1=w_1\mathbf{v}_1\bullet\mathbf{v}_1+w_2\mathbf{v}_2\bullet\mathbf{v}_1+\cdots+w_k\mathbf{v}_k\bullet\mathbf{v}_1=w_1\mathbf{v}_1\bullet\mathbf{v}_1$$

But this implies that $w_1=\dfrac{\mathbf{w}\bullet\mathbf{v}_1}{\mathbf{v}_1\bullet\mathbf{v}_1}$, which is what was claimed. Of course we can do the same for all other vectors, thus proving the above fact.

In other words, for an orthogonal basis, the linear combination produced by the coordinates splits the vector into the sum of its projections, just as we know to happen in the standard $\{\mathbf{i},\mathbf{j},\mathbf{k}\}$ basis.

Example: $B_1=\left\{\begin{bmatrix}1\\0\\0\end{bmatrix},\begin{bmatrix}1\\1\\0\end{bmatrix},\begin{bmatrix}1\\1\\1\end{bmatrix}\right\}$, $B_2=\left\{\begin{bmatrix}1\\1\\0\end{bmatrix},\begin{bmatrix}-1\\1\\0\end{bmatrix},\begin{bmatrix}0\\0\\3\end{bmatrix}\right\}$

Consider these two bases for $\mathbb{R}^3$. Both are bases and, from a certain point of view, B_1 is better, as its vectors only contain 0's and 1's. But let's look at it from a coordinates point of view.

If we consider the vector $\mathbf{w}=[1\ -2\ 3]$, its coordinate in the first base are $(3,-5,3)$:

$$3\begin{bmatrix}1\\0\\0\end{bmatrix}-5\begin{bmatrix}1\\1\\0\end{bmatrix}+3\begin{bmatrix}1\\1\\1\end{bmatrix}=\begin{bmatrix}1\\-2\\3\end{bmatrix}$$

To find these coordinates, we solve a linear system, which is not a difficult task, given the simplicity of the numbers. But what do those coordinates tell us, beyond their meaning from the definition?

The coordinates in the second base are $(-0.5,-1.5,1)$ and we can also get these through a linear system, but their meaning is more concrete. The projection of **w** on the first vector is half the first vector, in the opposite direction:

$$proj_{\mathbf{v}_1}\mathbf{w}=\frac{\mathbf{v}_1\bullet\mathbf{w}}{\mathbf{v}_1\bullet\mathbf{v}_1}\mathbf{v}_1=\frac{[1\ 1\ 0]\bullet[1\ -2\ 3]}{[1\ 1\ 0]\bullet[1\ 1\ 0]}\mathbf{v}_1=\frac{-1}{2}\mathbf{v}_1$$

Similarly for the other two (check it!).

So, the advantage of an orthogonal basis is in the interpretation of the coordinates in that basis?

Yes, although there are other advantages to be found down the road. And there is another type of bases that goes even further, in the same spirit.

Definition

A basis $B = \{\mathbf{v}_1, \cdots, \mathbf{v}_k\}$ of a subspace $\mathbb{S}$ of $\mathbb{R}^n$ (including the whole space) is said to be ***orthonormal*** if it is ***orthogonal*** and all its vectors are ***unit vectors***.

And what is the additional advantage of this one?

Well, what is the dot product of a unit vector with itself?

As for all vectors, it is equal to square of its length, which is therefore … 1!

Exactly, which leads to the following, even simpler property:

Technical fact

If $B = \{\mathbf{v}_1, \cdots, \mathbf{v}_k\}$ is an orthonormal basis of a subspace $\mathbb{S}$ of $\mathbb{R}^n$, then any vector **w** of $\mathbb{S}$ can be written as:

$$\mathbf{w} = (\mathbf{v}_1 \bullet \mathbf{w})\mathbf{v}_1 + (\mathbf{v}_2 \bullet \mathbf{w})\mathbf{v}_2 + \ldots + (\mathbf{v}_k \bullet \mathbf{w})\mathbf{v}_k$$

Therefore, the coordinates in this basis are given simply by the respective dot products and still represent the multiple of the basis vector that is used to build up the given vector.

Example: $B^* = \left\{ \begin{bmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \\ 0 \end{bmatrix}, \begin{bmatrix} -1/\sqrt{2} \\ 1/\sqrt{2} \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \right\}$

If we divide each vector of the earlier basis B_2 by its length, we obtain the orthonormal basis we are considering here.

The coordinate of the vector $\mathbf{w} = [1 \ -2 \ 3]$ in this basis are given by the dot products:

$$\begin{bmatrix} 1 \\ -2 \\ 3 \end{bmatrix} \bullet \begin{bmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \\ 0 \end{bmatrix} = -\frac{1}{\sqrt{2}} \quad ; \quad \begin{bmatrix} 1 \\ -2 \\ 3 \end{bmatrix} \bullet \begin{bmatrix} -1/\sqrt{2} \\ 1/\sqrt{2} \\ 0 \end{bmatrix} = -\frac{3}{\sqrt{2}}$$

$$\begin{bmatrix} 1 \\ -2 \\ 3 \end{bmatrix} \bullet \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} = 3$$

Indeed we can see that:

$$-\frac{1}{\sqrt{2}} \begin{bmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \\ 0 \end{bmatrix} - \frac{3}{\sqrt{2}} \begin{bmatrix} -1/\sqrt{2} \\ 1/\sqrt{2} \\ 0 \end{bmatrix} + 3 \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ -2 \\ 3 \end{bmatrix}$$

So, are we dealing with just a matter of convenience?

For now you can see it that way, but we shall make other uses of both orthogonality and orthonormality. For now, just practice on the recognition of these properties.

Summary

- Orthogonal and orthonormal bases generate their vectors through projections and dot products respectively. In that sense, they are "nicer" than other bases.

Common errors to avoid

- Notice that while the interpretation of the coordinates in an orthogonal or orthonormal basis is clearer, the actual numbers one gets may be more convoluted. But when one has computers, who cares!

Learning questions for Section LA 9-1

Review questions:

1. Describe the main advantage of orthogonal and orthonormal bases.

Memory questions:

1. When is a basis *orthogonal*?

2. When is a basis *orthonormal*?

Computation questions:

1. Prove that the vectors $\mathbf{v}_1 = \begin{bmatrix} 1/3 \\ 2/3 \\ 2/3 \end{bmatrix}, \mathbf{v}_2 = \begin{bmatrix} 2/3 \\ -1/3 \\ 0 \end{bmatrix}, \mathbf{v}_3 = \begin{bmatrix} 1 \\ 2 \\ -5/2 \end{bmatrix}$ form an orthogonal basis for $\mathbb{R}^3$ and then use them to construct an orthonormal basis.

2. Check that $B = \left\{ \begin{bmatrix} -\frac{1}{2} & \frac{\sqrt{3}}{2} & 0 \end{bmatrix}, \begin{bmatrix} \frac{\sqrt{3}}{2} & \frac{1}{2} & 0 \end{bmatrix}, \begin{bmatrix} 0 & 1 & -1 \end{bmatrix} \right\}$ is a basis for $\mathbb{R}^3$ and then determine whether it is orthogonal, orthonormal or neither.

3. Show that if $\mathbf{u} = \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix}$ and $\mathbf{v} = \begin{bmatrix} 3 \\ -1 \\ -1 \end{bmatrix}$, then the set $B = \{\mathbf{u}, \mathbf{v}\}$ is an orthogonal basis for $span\{\mathbf{u}, \mathbf{v}\}$. Also, find the coordinates of the vector $\mathbf{w} = \begin{bmatrix} 1 \\ -5 \\ -3 \end{bmatrix}$ in $span\{\mathbf{u}, \mathbf{v}\}$ with respect to this basis.

4. Compute the projection of the vector $\begin{bmatrix} 3 \\ 1 \\ -2 \end{bmatrix}$ on the plane spanned by the vectors $\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$ and $\begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}$ by using the concept of orthogonality.

Theory questions:

1. What are two notations for the standard orthonormal basis vectors in $\mathbf{R}^3$?

2. If $\{\mathbf{u}, \mathbf{v}\}$ is a basis for $\mathbb{R}^2$, is it true that any 2D vector is equal to the sum of its projections on these two basis vectors?

3. If $B = \{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ is an orthogonal basis and $\mathbf{w} = c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + c_3\mathbf{v}_3$, what is the projection of $\mathbf{w}$ on $\mathbf{v}_1$?

4. What is the main advantage of orthogonal and orthonormal bases?

Proof questions:

1. If you are given two non-parallel vectors in $\mathbb{R}^4$ that are not orthogonal, prove that it is possible to use them to construct an orthogonal basis for their span. Then prove that it is also possible to use them to construct an orthonormal basis.

Templated questions:

1. Construct a set of 3 vectors in $\mathbb{R}^3$ and determine if it is a basis and, if it is, whether it is orthogonal, orthonormal or neither.

2. Select 2 non-parallel vectors in $\mathbb{R}^4$ and determine if they form an orthogonal or orthonormal basis for their span.

3. Check if any basis you are given is orthogonal, orthonormal or neither.

What questions do you have for your instructor?