Math 108B	Professor: Padraic Bartlett

Lecture 3: Orthonormal Bases

Week 3 | *UCSB 2014*

In our last class, we introduced the concept of "changing bases," and talked about writing vectors and linear transformations in other bases. In the homework and in class, we saw that in several situations this idea of "changing basis" could make a linear transformation much easier to work with; in several cases, we saw that linear transformations under a certain basis would become diagonal, which made tasks like raising them to large powers far easier than these problems would be in the standard basis.

But how do we find these "nice" bases? What does it mean for a basis to be"nice?" In this set of lectures, we will study one potential answer to this question: the concept of an **orthonormal basis**.

1 Orthogonality

To start, we should define the notion of **orthogonality**. First, recall/remember the definition of the **dot product**:

Definition. Take two vectors $(x_1, \ldots x_n), (y_1, \ldots y_n) \in \mathbb{R}^n$. Their **dot product** is simply the sum

$$x_1y_1 + x_2y_2 + \ldots x_ny_n.$$

Many of you have seen an alternate, geometric definition of the dot product:

Definition. Take two vectors $(x_1, \ldots x_n), (y_1, \ldots y_n) \in \mathbb{R}^n$. Their **dot product** is the product

$$||\vec{x}|| \cdot ||\vec{y}|| \cos(\theta),$$

where θ is the angle between $\vec{x}$ and $\vec{y}$, and $||\vec{x}||$ denotes the length of the vector $\vec{x}$, i.e. the distance from $(x_1, \ldots x_n)$ to $(0, \ldots 0)$.

These two definitions are equivalent:

Theorem. Let $\vec{x} = (x_1, x_2, x_3), \vec{y} = (y_1, y_2, y_3)$ be a pair of vectors in $\mathbb{R}^3$. Then the algebraic interpretation of $\vec{x} \cdot \vec{y}$, given by

$$x_1y_1 + x_2y_2 + \ldots x_ny_n,$$

is equal to the geometric interpretation

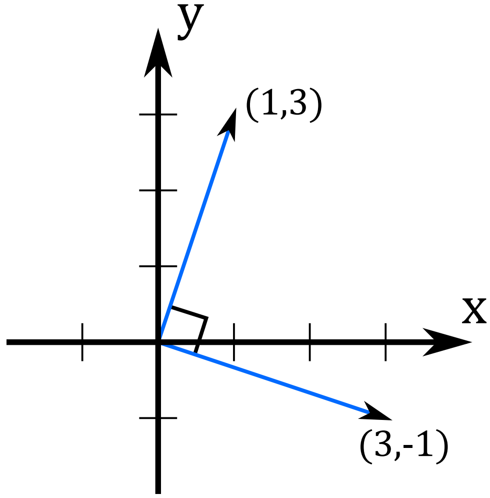
$$||\vec{x}|| \cdot ||\vec{y}|| \cos(\theta),$$

where θ is the angle between $\vec{x}$ and $\vec{y}$.

This is not a hard theorem to prove; it's essentially a consequence of some basic trigonometry, and you should do it if you're interested! Instead, I mostly want to mention it because it lets us discuss the concept of **orthogonality**:

Definition. Two vectors $\vec{v}, \vec{w}$ are called **orthogonal** if their dot product is 0.

Geometrically, we can interpret this as saying that two vectors are orthogonal whenever the angle θ between them is such that $\cos(\theta) = 0$; i.e. that these two vectors meet at either a $90°$ or $270°$-degree (i.e. $\pm\pi/2$ in radians) angle!



Two orthogonal vectors. Note that their algebraic dot product $1 \cdot 3 + 3 \cdot (-1)$ is zero, and that the angle between them is $\pi/2$.

That's orthogonality! In our introduction to these notes, we said that the idea of **orthogonality** was motivated by a desire to understand what makes a basis "useful." In the next section, we look at how orthogonality and bases interact:

2 An Orthonormal Basis: Definitions

Consider the following definition:

Definition. Suppose we have a basis $B = \{\vec{b_1}, \ldots \vec{b_n}\}$ for some space F^n, where F is either $\mathbb{R}$ or $\mathbb{C}$. We will say that this basis is **orthonormal** if it satisfies the following two properties:

1. **Orthogonality**: if we take any $\vec{b_i}, \vec{b_j}$ with $i \neq j$, then $\vec{b_i}$ and $\vec{b_j}$ are orthogonal: i.e. their dot product is 0.

2. **Unit length**: the length $||\vec{b_i}||$ of any vector $\vec{b_i}$ is 1.

This is a kind of basis! Basically, it's the best kind of basis, for the following reason: suppose that $B = \{\vec{b_1}, \ldots \vec{b_n}\}$ is an orthonormal basis for F^n. As discussed last week, we know that

$$A = \begin{bmatrix} \vdots & & \vdots \\ \vec{b_1} & \ldots & \vec{b_n} \\ \vdots & & \vdots \end{bmatrix}$$

is the matrix that converts vectors written in the basis B into vectors written in the standard basis, and

$$A^{-1} = \begin{bmatrix} \vdots & & \vdots \\ \vec{b_1} & \dots & \vec{b_n} \\ \vdots & & \vdots \end{bmatrix}^{-1}$$

is the matrix that converts vectors in the standard basis to vectors in the basis B. Then we have the following result:

Proposition. A^{-1} is precisely A^T, the transpose[1] of A! In other words,

$$A^{-1} = \begin{bmatrix} \dots & \vec{b_1} & \dots \\ & \vdots & \\ \dots & \vec{b_n} & \dots \end{bmatrix},$$

the matrix whose rows are given by the vectors of B.

Proof. Just check the answer by multiplying these two matrices together! In other words, just look at

$$A^T \cdot A = \begin{bmatrix} \dots & \vec{b_1} & \dots \\ & \vdots & \\ \dots & \vec{b_n} & \dots \end{bmatrix} \cdot \begin{bmatrix} \vdots & & \vdots \\ \vec{b_1} & \dots & \vec{b_n} \\ \vdots & & \vdots \end{bmatrix}.$$

To get the entries of the product of these two matrices, we just take the dot product of a row of the first matrix with a column of the second matrix. For the two matrices at hand, this means that the entry in (i, j) of the product is just

$$\vec{b_i} \cdot \vec{b_j}.$$

But we know these dot products! In particular, we know that the basis B is orthonormal: therefore, whenever $i \neq j$, this dot product is 0, because these two vectors are orthogonal. As well, when $i = j$, our dot product is just

$$\vec{b_i} \cdot \vec{b_i} = b_{i,1} \cdot b_{i,1} + b_{i,2} \cdot b_{i,2} + \dots + b_{i,n} \cdot b_{i,n} = \left(\sqrt{b_{i,1}^2 + \dots + b_{i,n}^2}\right)^2 = 1^2 = 1,$$

because the length $\sqrt{b_{i,1}^2 + \dots + b_{i,n}^2}$ of each vector $\vec{b_i}$ is precisely 1.

In other words, we've just shown that the product of A^T and A is the matrix where $(i, j) = 0$ whenever $i \neq j$ and $(i, j) = 1$ if $i = j$; i.e. it's the identity matrix! In other words, by definition we have that A^T is the inverse of A, as multiplying A by A^T yields the identity matrix. □

[1] Given a $m \times n$ matrix A, the transpose A^T of A is the $n \times m$ matrix that you get by "flipping" A over its main diagonal: i.e. by setting entry (i, j) of A^T to be $a_{j,i}$. For example, the transpose of $\begin{bmatrix} 2 & 3 \\ 1 & 4 \end{bmatrix}$ is $\begin{bmatrix} 2 & 1 \\ 3 & 4 \end{bmatrix}$.

Why do we care about this result? Well: it lets us effectively calculate the matrix that transforms standard-basis vectors into basis-B vectors for free! Normally, we have to explicitly find the inverse of this matrix, which is computationally expensive and difficult. For this nice basis, however, you just have to find the transpose of $\begin{bmatrix} \vdots & & \vdots \\ \vec{b_1} & \dots & \vec{b_n} \\ \vdots & & \vdots \end{bmatrix}$, which is really easy!

3 An Orthonormal Basis: Examples

Before we do more theory, we first give a quick example of two orthonormal bases, along with their change-of-basis matrices.

Example. One trivial example of an orthonormal basis is the **standard basis**! In other words, look at the basis $E = \{\vec{e_1} = (1, 0, 0, \dots 0), \dots \vec{e_n} = (0, \dots 0, 1)\}$. Everything in this basis trivially has length 1 and is orthogonal.

Example. Here are four vectors that form an orthonormal basis for $\mathbb{R}^4$:

$$H = \left\{ \left(\frac{1}{2}, \frac{1}{2}, \frac{1}{2}, \frac{1}{2} \right), \left(\frac{1}{2}, \frac{1}{2}, -\frac{1}{2}, -\frac{1}{2} \right), \left(\frac{1}{2}, -\frac{1}{2}, -\frac{1}{2}, \frac{1}{2} \right), \left(\frac{1}{2}, -\frac{1}{2}, \frac{1}{2}, -\frac{1}{2} \right), \right\}.$$

To see that they are all mutually orthogonal, it helps to factor out a $\frac{1}{2}$ from these four vectors:

$$H = \left\{ \frac{1}{2}(1, 1, 1, 1), \frac{1}{2}(1, -1, 1, -1), \frac{1}{2}(1, -1, -1, 1), \frac{1}{2}(1, 1, -1, -1) \right\}.$$

You are invited to check that these four vectors are all orthogonal and have unit length, and thus form an orthonormal basis! The vector that converts vectors from the basis H to the standard basis is just the matrix with these vectors as its columns:

$$\begin{bmatrix} \frac{1}{2} & \frac{1}{2} & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & -\frac{1}{2} & -\frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} & -\frac{1}{2} & -\frac{1}{2} \\ \frac{1}{2} & -\frac{1}{2} & \frac{1}{2} & -\frac{1}{2} \end{bmatrix}.$$

By using our proposition, we can write the matrix that converts vectors from the standard basis into the basis H as just the transpose of the above matrix:

$$\begin{bmatrix} \frac{1}{2} & \frac{1}{2} & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & -\frac{1}{2} & \frac{1}{2} & -\frac{1}{2} \\ \frac{1}{2} & -\frac{1}{2} & -\frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} & -\frac{1}{2} & -\frac{1}{2} \end{bmatrix}.$$

(This orthonormal basis has some nice properties, that students from last quarter will recognize: it's a collection of vectors such that (1) every entry in every vector has the same absolute value, and (2) they're all orthogonal! In general, finding such sets of vectors is an open problem in mathematics, with a number of applications. It is conjectured that for any n, we can find a basis for $\mathbb{R}^{4n}$ consisting of such vectors; thus far, we only know this up to $n = 167$! So, um, fun extra-extra credit problem: find one of these bases?)

Excellent! We have now carefully defined the concept of an orthonormal basis, looked at some examples of orthonormal bases, and proven a theorem that explains part of why they are useful to us!

This leaves one immediately obvious question to study: how can we make more?

4 Turning Bases Into Orthonormal Bases: Projection

Specifically, let's consider the following question:

Question 1. *Suppose that we have some k-dimensional vector space V; for example, we could have $V = \mathbb{R}^k$, or $V = \{(x_1, \ldots x_{k+1}) : x_1 + \ldots x_{k+1}\}$, which you can show is a k-dimensional subspace of $\mathbb{R}^{k+1}$. V can be lots of things.*

Can we always find an orthonormal basis O_B for V, no matter what V is?

The answer to this is ... not obvious at all! So, we do what mathematicians always do when presented with a difficult problem: look at special cases! Specifically, let's consider the simpler case where $k = 2$, i.e we're dealing with a two-dimensional space V:

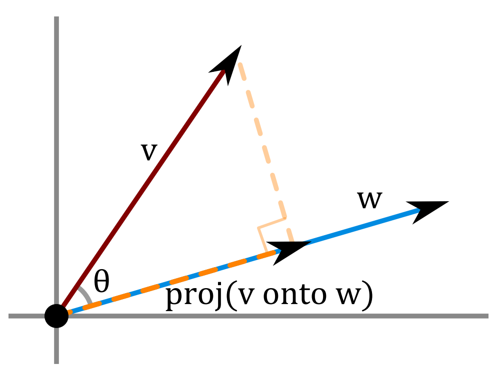
Question 2. *Suppose that we have vector space V that's two-dimensional: i.e. it has a basis $B = \{\vec{b_1}, \vec{b_2}\}$ for a space V. Can we find a new,* ***orthonormal*** *basis O_B for V?*

Answer. First, make the following definition:

Definition. Let $\vec{v}, \vec{w}$ be a pair of vectors in $\mathbb{R}^n$. The **projection** of $\vec{v}$ onto $\vec{w}$, denoted proj($\vec{v}$ onto $\vec{w}$), is the following vector:

- Take the vector $\vec{w}$.
- Draw a line perpendicular to the vector $\vec{w}$, that goes through the point $\vec{v}$ and intersects the line spanned by the vector $\vec{w}$.
- proj($\vec{v}$ onto $\vec{w}$) is precisely the point at which this perpendicular line intersects $\vec{w}$.

We illustrate this below:



In particular, it bears noting that this vector is a multiple of $\vec{w}$.

A formula for this vector is the following:

$$\text{proj}(\vec{v} \text{ onto } \vec{w}) = \frac{\vec{v} \cdot \vec{w}}{||\vec{w}||^2} \cdot \vec{w}.$$

To see why, simply note that the vector we want is, by looking at the above picture, something of length $\cos(\theta) \cdot ||\vec{v}||$, in the direction of $\vec{w}$. In other words,

$$\text{proj}(\vec{v} \text{ onto } \vec{w}) = \cos(\theta) \cdot ||\vec{v}|| \cdot \frac{\vec{w}}{||\vec{w}||}.$$

Now, use the angle form of the dot product to see that because $\vec{w} \cdot \vec{v} = ||\vec{w}||||\vec{v}|| \cos(\theta)$, we have

$$\text{proj}(\vec{v} \text{ onto } \vec{w}) = \frac{\vec{v} \cdot \vec{w}}{||\vec{v}||||\vec{w}||}||\vec{v}||\frac{\vec{w}}{||\vec{w}||}.$$

Canceling the $||\vec{v}||$'s gives us the desired formula.

Using this, you can define the "orthogonal part" of $\vec{v}$ over $\vec{w}$ in a similar fashion:

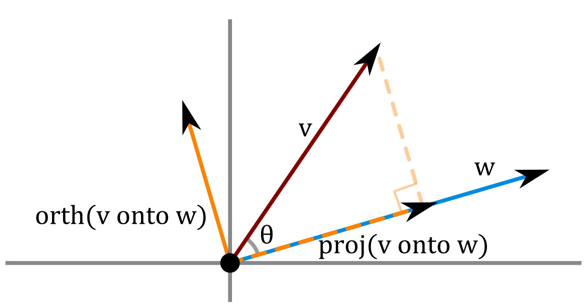
Definition. Let $\vec{v}, \vec{w}$ be a pair of vectors in $\mathbb{R}^n$. The **orthogonal part** of $\vec{v}$ over $\vec{w}$, denoted orth($\vec{v}$ onto $\vec{w}$), is the following vector:

$$\text{orth}(\vec{v} \text{ over } \vec{w}) = \vec{v} - \text{proj}(\vec{v} \text{ onto } \vec{w})$$

It bears noting that this vector lives up to its name, and is in fact orthogonal to $\vec{w}$. This is not hard to see: just take the dot product of $\vec{w}$ with it! This yields

$$\begin{aligned}
\vec{w} \cdot (\vec{v} - \text{proj}(\vec{v} \text{ onto } \vec{w})) &= \vec{w} \cdot \vec{v} - \vec{w} \cdot \text{proj}(\vec{v} \text{ onto } \vec{w}) \\
&= \vec{w} \cdot \vec{v} - \vec{w} \cdot (\frac{\vec{v} \cdot \vec{w}}{||\vec{w}||^2} \cdot \vec{w}) \\
&= \vec{w} \cdot \vec{v} - \frac{\vec{v} \cdot \vec{w}}{||\vec{w}||^2}(\vec{w} \cdot \vec{w}) \\
&= \vec{w} \cdot \vec{v} - \frac{\vec{v} \cdot \vec{w}}{||\vec{w}||^2}(w_1^2 + \ldots w_n^2) \\
&= \vec{w} \cdot \vec{v} - \frac{\vec{v} \cdot \vec{w}}{||\vec{w}||^2}(||\vec{w}||^2) \\
&= 0.
\end{aligned}$$

As is often the case, actually drawing the vectors orth($\vec{v}$ over $\vec{w}$) and $\vec{w}$ makes this proof much more believable:



We can now answer the question that we asked at the start of this section! Specifically: suppose that $B = \{\vec{b_1}, \vec{b_2}\}$ is some basis for a two-dimensional space V.

We can make a new basis O_B for V as follows:

1. Define the first basis vector $\vec{r_1}'$ of O_B as just $\vec{b_1}$.

2. Now, set $\vec{r_2}'$ =orth($\vec{b_2}$ over $\vec{r_1}$). By our work above, this vector is orthogonal to $\vec{r_1}'$.

3. Finally, define $\vec{r_1} = \vec{r_1}'/||\vec{r_1}'||$, $\vec{r_2} = \vec{r_2}'/||\vec{r_2}'||$. By construction, these vectors are just the vectors from (1) and (2), except scaled to be length 1.

Therefore, the set O_B is made of orthonormal vectors! We just need to show that it is an orthonormal **basis for** V, and we're done! This is not hard to do: we just need to show that (1) O_B spans V, and (2) that O_B is not linearly dependent.

For (2): we will prove that in general, a set of k mutually orthogonal nonzero vectors $\{\vec{r_1}, \ldots \vec{r_k}\}$ is linearly independent. To see this, consider any linear combination of the orthogonal vectors that combines to $\vec{0}$:

$$a_1\vec{r_1} + \ldots a_k\vec{r_k} = \vec{0}.$$

Take any $\vec{r_i}$, and multiply the left- and right-hand-sides by $\vec{r_i}$. Because $\vec{r_i}, \vec{r_j}$ are orthogonal whenever $i \neq j$, the left-hand-side simplifies to $a_i\vec{r_i} \cdot \vec{r_i} = a_i \cdot ||\vec{r_i}||^2$, while the right-hand-side is just 0. We know that because $\vec{r_i}$ is nonzero, its length is also nonzero: therefore we can divide both sides by $||\vec{r_i}||^2$, and get that $a_i = 0$. This holds for every i!

Therefore, we have proven that the only way to combine the vectors $\{\vec{r_1}, \ldots \vec{r_k}\}$ to get $\vec{0}$ is to use the trivial linear combination where we multiply all of these vectors by 0. This is the definition of linear independence! So we've proven (2).

For (1): Look at the span of O_B, i.e. the set of all linear combinations of elements of O_B. On one hand, because both $\vec{r_1}, \vec{r_2}$ are made out of linear combinations of the vectors $\vec{b_1}, \vec{b_2}$, the span of O_B must be a subspace of V. On the other hand, the span of O_B is two-dimensional, because O_B contains two vectors! Therefore, the span of O_B is a two-dimensional thing living in a two-dimensional space V: in other words, it **is** V.

So we're done!

Here's a quick example of how the above works in practice:

Example. Examine the space V spanned by the two vectors $\{(1, 2, 2), (0, 3, 6)\}$. Find an orthonormal basis for V.

Answer. We proceed as in our example above. First, we set

$$\vec{r_1}' = (1, 2, 2).$$

Then, we define $\vec{r_2}'$ by

$$\begin{aligned}
\vec{r_2}' &= (0, 3, 6) - \text{proj}\left((0, 3, 6) \text{ onto } (1, 2, 2)\right) \\
&= (0, 3, 6) - \frac{(0, 3, 6) \cdot (1, 2, 2)}{(1, 2, 2) \cdot (1, 2, 2)}(1, 2, 2) \\
&= (0, 3, 6) - \frac{18}{9}(1, 2, 2) \\
&= (-2, -1, 2).
\end{aligned}$$

Finally, we scale each by their length:

$$\vec{r_1} = \frac{\vec{r_1}'}{||\vec{r_1}||} = \frac{1}{\sqrt{1^2+2^2+2^2}}(1,2,2) = \left(\frac{1}{3},\frac{2}{3},\frac{2}{3}\right),$$
$$\vec{r_2} = \frac{\vec{r_2}'}{||\vec{r_2}||} = \frac{1}{\sqrt{(-2)^2+(-1)^2+2^2}}(-2,-1,2) = \left(-\frac{2}{3},-\frac{1}{3},\frac{2}{3}\right).$$

An orthonormal basis for V!

You should verify for yourself that these two vectors are orthogonal, and that they can be combined to create both $(1,2,2)$ and $(0,3,6)$ (and therefore span V.)

One more example:

Example. Examine the space V spanned by the two vectors $\{(2,3,6),(1,0,2)\}$. Find an orthonormal basis for V.

Answer. We proceed as in our example above. First, we set

$$\vec{r_1}' = (2,3,6).$$

Then, we define $\vec{r_2}'$ by

$$\begin{aligned}\vec{r_2}' &= (1,0,2) - \text{proj}\left((1,0,2) \text{ onto } (2,3,6)\right)\\ &= (1,0,2) - \frac{(1,0,2)\cdot(2,3,6)}{(2,3,6)\cdot(2,3,6)}(2,3,6)\\ &= (1,0,2) - \frac{14}{49}(2,3,6)\\ &= (\frac{3}{7},-\frac{6}{7},\frac{2}{7}).\end{aligned}$$

Finally, we scale each by their length:

$$\vec{r_1} = \frac{\vec{r_1}'}{||\vec{r_1}||} = \frac{1}{\sqrt{2^2+3^2+6^2}}(2,3,6) = \left(\frac{2}{7},\frac{3}{7},\frac{6}{7}\right),$$
$$\vec{r_2} = \frac{\vec{r_2}'}{||\vec{r_2}||} = \frac{1}{\sqrt{(3/7)^2+(-6/7)^2+(2/7)^2}}(\frac{3}{7},-\frac{6}{7},\frac{2}{7}) = (\frac{3}{7},-\frac{6}{7},\frac{2}{7}).$$

An orthonormal basis for V!

By thinking about how we did the proof above, we can actually get a way to answer our generalized question! In other words, we can finally answer the following:

Question 3. *Suppose that we have some k-dimensional vector space V; for example, we could have $V = \mathbb{R}^k$, or $V = \{(x_1,\ldots x_{k+1}) : x_1 + \ldots x_{k+1}\}$, which you can show is a k-dimensional subspace of $\mathbb{R}^{k+1}$. V can be lots of things.*

Can we always find an orthonormal basis O_B for V, no matter what V is?

Answer. We just generalize the process that we used in our work above. Suppose that V is a k-dimensional space with a basis $B = \{\vec{b_1}, \dots \vec{b_k}\}$.

Consider the following process (called the Gram-Schmidt process, formally), to create a set of k nonzero orthogonal vectors $\{\vec{u_1}, \dots \vec{u_k}\}$ out of combinations of the B-elements:

- $\vec{u_1} = \vec{b_1}$.
- $\vec{u_2} = \vec{b_2} - \text{proj}(\vec{b_2} \text{ onto } \vec{u_1})$.
- $\vec{u_3} = \vec{b_3} - \text{proj}(\vec{b_3} \text{ onto } \vec{u_1}) - \text{proj}(\vec{b_3} \text{ onto } \vec{u_2})$.
- $\vec{u_4} = \vec{b_4} - \text{proj}(\vec{b_4} \text{ onto } \vec{u_1}) - \text{proj}(\vec{b_4} \text{ onto } \vec{u_2}) - \text{proj}(\vec{b_4} \text{ onto } \vec{u_3})$.

$\vdots$

- $\vec{u_k} = \vec{b_k} - \text{proj}(\vec{b_k} \text{ onto } \vec{u_1}) - \dots - \text{proj}(\vec{b_k} \text{ onto } \vec{u_{k-1}})$.

We claim that for any n, the vector $\vec{u_n}$ is orthogonal to any $\vec{u_m}$, with $n > m$, and prove this via strong induction. For $n = 1$, our claim is trivially true, and for $n = 2$ our claim is just the special case we studied earlier; so those are our base cases.

For the induction step: we assume that our claim holds for all $\vec{u_i}$, where $1 \leq i \leq n$. We seek to prove our claim for $n+1$: i.e. we want to show that

$$\vec{u_{n+1}} = \vec{b_{n+1}} - \text{proj}(\vec{b_{n+1}} \text{ onto } \vec{u_1}) - \dots - \text{proj}(\vec{b_{n+1}} \text{ onto } \vec{u_n})$$

is orthogonal to every $\vec{u_j}$, for $j < n+1$.

This is not hard to see — just pick any value of j, and calculate the dot product of $\vec{u_{n+1}}$ and $\vec{u_j}$! This is just

$$\begin{aligned}
\vec{u_{n+1}} \cdot \vec{u_j} &= \Big(\vec{b_{n+1}} - \text{proj}(\vec{b_{n+1}} \text{ onto } \vec{u_1}) - \dots - \text{proj}(\vec{b_{n+1}} \text{ onto } \vec{u_n})\Big) \cdot \vec{u_j} \\
&= \Big(\vec{b_{n+1}} - \text{proj}(\vec{b_{n+1}} \text{ onto } \vec{u_j})\Big) \cdot \vec{u_j} \quad \left.\begin{array}{r} - \Big(\text{proj}(\vec{b_{n+1}} \text{ onto } \vec{u_1})\Big) \cdot \vec{u_j} \\ - \Big(\text{proj}(\vec{b_{n+1}} \text{ onto } \vec{u_2})\Big) \cdot \vec{u_j} \\ \vdots \\ - \Big(\text{proj}(\vec{b_{n+1}} \text{ onto } \vec{u_n})\Big) \cdot \vec{u_j} \end{array}\right\} (\text{all } \vec{u_i} \neq \vec{u_j}'s).
\end{aligned}$$

On one hand, we know that the first term is just

$$\Big(\vec{b_{n+1}} - \text{proj}(\vec{b_{n+1}} \text{ onto } \vec{u_j})\Big) \cdot \vec{u_j} = \Big(\text{orth}(\vec{b_{n+1}} \text{ over } \vec{u_j}\Big) \cdot \vec{u_j},$$

which is 0 by construction (as these two vectors are orthogonal.)

On the other hand, each of the other terms is of the form

$$\Big(\text{proj}(\vec{b_{n+1}} \text{ onto } \vec{u_i})\Big) \cdot \vec{u_j} = (\text{some constant})\vec{u_i} \cdot \vec{u_j},$$

for $i \neq j$ and $i, j \leq n$. We know that because $i \neq j$, we either have $i > j$ or $j > i$, and both are bounded above by n. Therefore, in any case, our inductive assumption tells us that these two vectors $\vec{u_i}, \vec{u_j}$ are orthogonal! Therefore, their dot product is also 0.

Thus, the entire dot product $\vec{u_{n+1}} \cdot \vec{u_j}$ is zero, and we have proven our inductive claim. Consequently, we know that our claim holds for everything: i.e. the set $\{\vec{u_1}, \ldots \vec{u_k}\}$ is a collection of nonzero orthogonal vectors, made only out of the elements of B! If we divide each of these elements by their length, we get a collection of k unit-length mutually orthogonal vectors made out of the elements of B.

As proven earlier, this set is linearly independent and made out of elements of B and thus spans a k-dimensional subspace of V. But V is itself k-dimensional: so the span of this set must be V itself! So we've found an orthonormal basis for V, as desired.

We run one quick example here:

Example. Run the above process on the space spanned by the vectors $\{(1,1,0),(1,0,1),(0,1,1)\}$.

Proof. So: via the algorithm above, we define the vectors $\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3$ as follows:

$$\begin{aligned}
\vec{u}_1 &= (1,1,0),\\
\vec{u}_2 &= (1,0,1) - \text{proj}((1,0,1) \text{ onto } (1,1,0))\\
&= (1,0,1) - \frac{(1,1,0)\cdot(1,0,1)}{(1,1,0)\cdot(1,1,0)}(1,1,0)\\
&= (1,0,1) - \frac{1}{2}(1,1,0)\\
&= \left(\frac{1}{2}, -\frac{1}{2}, \frac{1}{2}\right)\\
\vec{u}_3 &= (0,1,1) - \text{proj}((0,1,1) \text{ onto } (1,1,0)) - \text{proj}\left((0,1,1) \text{ onto } \left(\frac{1}{2}, -\frac{1}{2}, \frac{1}{2}\right)\right)\\
&= (0,1,1) - \frac{(1,1,0)\cdot(0,1,1)}{(1,1,0)\cdot(1,1,0)}(1,1,0) - \frac{\left(\frac{1}{2}, -\frac{1}{2}, \frac{1}{2}\right)\cdot(0,1,1)}{\left(\frac{1}{2}, -\frac{1}{2}, \frac{1}{2}\right)\cdot\left(\frac{1}{2}, -\frac{1}{2}, \frac{1}{2}\right)}\left(\frac{1}{2}, -\frac{1}{2}, \frac{1}{2}\right)\\
&= (0,1,1) - \frac{1}{2}(1,1,0) - 0\left(\frac{1}{2}, -\frac{1}{2}, \frac{1}{2}\right)\\
&= (0,1,1) - \left(\frac{1}{2}, \frac{1}{2}, 0\right)\\
&= \left(-\frac{1}{2}, \frac{1}{2}, 1\right).
\end{aligned}$$

Fun fact: these are all now pairwise orthogonal! Our theorem tells us this directly, but we can double-check it to illustrate the idea:

$$(1,1,0)\cdot\left(\frac{1}{2},-\frac{1}{2},\frac{1}{2}\right)=-\frac{1}{2}+\frac{1}{2}=0.$$
$$(1,1,0)\cdot\left(-\frac{1}{2},\frac{1}{2},1\right)=-\frac{1}{2}+\frac{1}{2}=0.$$
$$\left(\frac{1}{2},-\frac{1}{2},\frac{1}{2}\right)\cdot\left(-\frac{1}{2},\frac{1}{2},1\right)=-\frac{1}{4}-\frac{1}{4}+\frac{1}{2}=0.$$

Now, to finish up, we just scale each vector by its length:

$$\vec{r_1}=\frac{\vec{u_1}}{||\vec{u_1}||}=\frac{1}{\sqrt{1^2+1^2+0^2}}(1,1,0)=\left(\frac{1}{\sqrt{2}},\frac{1}{\sqrt{2}},0\right),$$
$$\vec{r_2}=\frac{\vec{u_2}}{||\vec{u_2}||}=\frac{1}{\sqrt{(-1/2)^2+(1/2)^2+1^2}}\left(-\frac{1}{2},\frac{1}{2},1\right)=\left(-\frac{\sqrt{6}}{6},\frac{\sqrt{6}}{6},\frac{\sqrt{6}}{3}\right),$$
$$\vec{r_3}=\frac{\vec{u_3}}{||\vec{u_3}||}=\frac{1}{\sqrt{(1/2)^2+(-1/2)^2+(1/2)^2}}\left(\frac{1}{2},-\frac{1}{2},\frac{1}{2}\right)=\left(\frac{\sqrt{3}}{3},-\frac{\sqrt{3}}{3},\frac{\sqrt{3}}{3}\right).$$

Done! □