Lecture 2a

Orthonormal Bases

(pages 321-323)

Let's stop and think about the features we really want in a basis. By its definition, it is a linearly independent spanning set, but the reason we wanted those features is that we wanted to be able to uniquely write every vector as a linear combination of the basis vectors. That is, we wanted to be able to assign coordinates based on our basis. Once we have coordinates, we can use all the results from $\mathbb{R}^n$, including finding a matrix to represent any linear transformation. So, while any basis will lead to coordinates, it would be nice if those coordinates were easy to find! That's the great thing about our standard bases. But sometimes the standard basis doesn't fit with our situation. (Various geometrical transformations are better defined with respect to other bases, for example.) So, let's see if we can figure out a more general class of basis, whose coordinates are easy to find, but that aren't necessarily the standard basis (or just some rearrangement of it).

First, I want to point out that from this point forward, we will again focus our attention purely on $\mathbb{R}^n$, since we now know that we can extend these results to any finite dimensional vector space. And once we are focused on $\mathbb{R}^n$, we can turn our attention to the dot product. Recall the definition from Section 1.3:

<u>Definition</u>: Let $\vec{x} = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$ and $\vec{y} = \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix}$ be vectors in $\mathbb{R}^n$. Then the **dot product** of $\vec{x}$ and $\vec{y}$ is

$$\vec{x} \cdot \vec{y} = x_1y_1 + \cdots + x_ny_n$$

We also want to recall that two vectors $\vec{x}$ and $\vec{y}$ are orthogonal if $\vec{x} \cdot \vec{y} = 0$. We now wish to extend this notion of orthogonality to sets of vectors.

<u>Definition</u>: A set of vectors $\{\vec{v}_1, \ldots, \vec{v}_k\}$ in $\mathbb{R}^n$ is **orthogonal** if $\vec{v}_i \cdot \vec{v}_j = 0$ whenever $i \neq j$.

Example: The set $\left\{ \begin{bmatrix} 2 \\ -3 \end{bmatrix}, \begin{bmatrix} 6 \\ 4 \end{bmatrix} \right\}$ is orthogonal, because $\begin{bmatrix} 2 \\ -3 \end{bmatrix} \cdot \begin{bmatrix} 6 \\ 4 \end{bmatrix} = (2)(6) + (-3)(4) = 12 - 12 = 0$.

The set $\left\{ \begin{bmatrix} 7 \\ 4 \end{bmatrix}, \begin{bmatrix} -1 \\ 2 \end{bmatrix} \right\}$ is not orthogonal, because $\begin{bmatrix} 7 \\ 4 \end{bmatrix} \cdot \begin{bmatrix} -1 \\ 2 \end{bmatrix} = (7)(-1) + (4)(2) = -7 + 8 = -1 \neq 0$.

The set $\left\{ \begin{bmatrix} 3 \\ 1 \\ 4 \end{bmatrix}, \begin{bmatrix} -1 \\ -1 \\ 1 \end{bmatrix}, \begin{bmatrix} 5 \\ -7 \\ -2 \end{bmatrix} \right\}$ is orthogonal, because

$$\begin{bmatrix} 3 \\ 1 \\ 4 \end{bmatrix} \cdot \begin{bmatrix} -1 \\ -1 \\ 1 \end{bmatrix} = -3 - 1 + 4 = 0$$
$$\begin{bmatrix} 3 \\ 1 \\ 4 \end{bmatrix} \cdot \begin{bmatrix} 5 \\ -7 \\ -2 \end{bmatrix} = 15 - 7 - 8 = 0$$
$$\begin{bmatrix} -1 \\ -1 \\ 1 \end{bmatrix} \cdot \begin{bmatrix} 5 \\ -7 \\ -2 \end{bmatrix} = -5 + 7 - 2 = 0$$

The set $\left\{ \begin{bmatrix} 1 \\ 2 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 5 \\ -2 \\ 4 \\ -1 \end{bmatrix}, \begin{bmatrix} 8 \\ -3 \\ -12 \\ -2 \end{bmatrix}, \begin{bmatrix} 6 \\ -5 \\ -9 \\ 4 \end{bmatrix} \right\}$ is not orthogonal, because

$$\begin{bmatrix} 8 \\ -3 \\ -12 \\ -2 \end{bmatrix} \cdot \begin{bmatrix} 6 \\ -5 \\ -9 \\ 4 \end{bmatrix} = 48 + 15 + 108 - 8 = 163 \neq 0.$$

Example: The standard basis $\{\vec{e}_1, \ldots, \vec{e}_n\}$ for $\mathbb{R}^n$ is orthogonal, since $\vec{e}_i \cdot \vec{e}_j = 0$ if $i \neq j$.

Why are orthogonal sets so great? Well, here's the first thing:

Theorem 7.1.1: If $\{\vec{v}_1, \ldots, \vec{v}_k\}$ is an orthogonal set of non-zero vectors in $\mathbb{R}^n$, it is linearly independent.

Proof of Theorem 7.1.1 Note first that we need the vectors to be non-zero, since any set containing the zero vector is linearly dependent. And so, let's assume that $\{\vec{v}_1, \ldots, \vec{v}_k\}$ is an orthogonal set of non-zero vectors in $\mathbb{R}^n$. To see that $\{\vec{v}_1, \ldots, \vec{v}_k\}$ is linearly independent, let's assume we have scalars $c_1, \ldots, c_k$ such that

$$c_1\vec{v}_1 + \cdots + c_k\vec{v}_k = \vec{0}$$

Then, for every $1 \leq i \leq k$, we can take the dot product of $\vec{v}_i$ with both sides of this equation, getting

$$\begin{aligned}
(c_1\vec{v}_1 + \cdots + c_k\vec{v}_k)\cdot\vec{v}_i &= \vec{0}\cdot\vec{v}_i \\
(c_1\vec{v}_1)\cdot\vec{v}_i + \cdots + (c_k\vec{v}_k)\cdot\vec{v}_i &= 0 \\
c_1(\vec{v}_1\cdot\vec{v}_i) + \cdots + c_i(\vec{v}_i\cdot\vec{v}_i) + \cdots + c_k(\vec{v}_k\cdot\vec{v}_i) &= 0 \\
c_1(0) + \cdots + c_i(||\vec{v}_i||^2) + \cdots + c_k(0) &= 0 \\
c_i(||\vec{v}_i||^2) &= 0
\end{aligned}$$

Since we know $\vec{v}_i \neq \vec{0}$, we know that $||\vec{v}_i||^2 \neq 0$, which means we can divide by it and get that $c_i = 0$. And since this is true for all $1 \leq i \leq k$, we have shown that all $c_i = 0$ for $1 \leq i \leq k$, which means that $\{\vec{v}_1, \ldots, \vec{v}_k\}$ is linearly independent.

In addition to being orthogonal, the standard basis has one other nice property: the length (or norm) of each of the vectors is 1. If we add that requirement to a general orthogonal set, we say that the set is orthonormal.

Definition: A set $\{\vec{v}_1\ \ldots, \vec{v}_k\}$ of vectors in $\mathbb{R}^n$ is **orthonormal** if it is orthogonal and each vector $\vec{v}_i$ is a unit vector (that is, each vector is normalized).

Note that, since the zero vector does not have length 1, it can never be in an orthonormal set. So, using Theorem 7.1.1, we see that all orthonormal sets are linearly independent.

Example: The set $\left\{\begin{bmatrix} 2/\sqrt{13} \\ -3/\sqrt{13}\end{bmatrix}, \begin{bmatrix} 6/\sqrt{52} \\ 4/\sqrt{52}\end{bmatrix}\right\}$ is an orthonormal set, since $\begin{bmatrix} 2/\sqrt{13} \\ -3/\sqrt{13}\end{bmatrix}\cdot\begin{bmatrix} 6/\sqrt{52} \\ 4/\sqrt{52}\end{bmatrix} = 12/\sqrt{676} - 12/\sqrt{676} = 0$, $\left|\left|\begin{bmatrix} 2/\sqrt{13} \\ -3/\sqrt{13}\end{bmatrix}\right|\right| = \sqrt{4/13 + 9/13} = 1$, and $\left|\left|\begin{bmatrix} 6/\sqrt{52} \\ 4/\sqrt{52}\end{bmatrix}\right|\right| = \sqrt{36/52 + 16/52} = 1$.

The set $\left\{\begin{bmatrix} 3/\sqrt{26} \\ 1/\sqrt{26} \\ 4/\sqrt{26}\end{bmatrix}, \begin{bmatrix} -1/\sqrt{3} \\ -1/\sqrt{3} \\ 1/\sqrt{3}\end{bmatrix}, \begin{bmatrix} 5/\sqrt{78} \\ -7/\sqrt{78} \\ -2/\sqrt{78}\end{bmatrix}\right\}$ is orthonormal, since

$$\begin{bmatrix} 3/\sqrt{26} \\ 1/\sqrt{26} \\ 4/\sqrt{26}\end{bmatrix}\cdot\begin{bmatrix} -1/\sqrt{3} \\ -1/\sqrt{3} \\ 1/\sqrt{3}\end{bmatrix} = -\tfrac{3}{78} - \tfrac{1}{78} + \tfrac{4}{78} = 0$$
$$\begin{bmatrix} 3/\sqrt{26} \\ 1/\sqrt{26} \\ 4/\sqrt{26}\end{bmatrix}\cdot\begin{bmatrix} 5/\sqrt{78} \\ -7/\sqrt{78} \\ -2/\sqrt{78}\end{bmatrix} = \tfrac{15}{2028} - \tfrac{7}{2028} - \tfrac{8}{2028} = 0$$
$$\begin{bmatrix} -1/\sqrt{3} \\ -1/\sqrt{3} \\ 1/\sqrt{3}\end{bmatrix}\cdot\begin{bmatrix} 5/\sqrt{78} \\ -7/\sqrt{78} \\ -2/\sqrt{78}\end{bmatrix} = -\tfrac{5}{234} + \tfrac{7}{234} - \tfrac{2}{234} = 0$$

and

$$\left\|\begin{bmatrix} 3/\sqrt{26} \\ 1/\sqrt{26} \\ 4/\sqrt{26} \end{bmatrix}\right\| = \sqrt{\tfrac{9}{26} + \tfrac{1}{26} + \tfrac{16}{26}} = 1$$

$$\left\|\begin{bmatrix} -1/\sqrt{3} \\ -1/\sqrt{3} \\ 1/\sqrt{3} \end{bmatrix}\right\| = \sqrt{\tfrac{1}{3} + \tfrac{1}{3} + \tfrac{1}{3}} = 1$$

$$\left\|\begin{bmatrix} 5/\sqrt{78} \\ -7/\sqrt{78} \\ -2/\sqrt{78} \end{bmatrix}\right\| = \sqrt{\tfrac{25}{78} + \tfrac{49}{78} + \tfrac{4}{78}} = 1$$

Note that these examples were obtained from the orthogonal sets in our first example by simply finding the unit vectors that correspond to the original vectors. This is, in fact, the usual way that orthonormal sets are created: first worry about creating an orthogonal set, and then worry about making it orthonormal.