Lecture 2b

Coordinates with respect to an Orthonormal Basis

(pages 323-325)

I said that our goal was to create a class of basis that was not necessarily the standard basis, but whose coordinates were still easy to calculate. As you may have guessed, an orthonormal basis is such a basis. Note first that, since every orthonormal set is linearly independent, once we have a set of n vectors from $\mathbb{R}^n$ in an orthonormal set, we automatically know that it is a basis by the two-out-of-three rule. The following theorem shows how to find the coordinates with respect to an orthonormal basis.

Theorem 7.1.2: If $\mathcal{B} = \{\vec{v}_1, \dots, \vec{v}_n\}$ is an orthonormal basis for $\mathbb{R}^n$, then the i-th coordinate of a vector $\vec{x} \in \mathbb{R}^n$ with respect to $\mathcal{B}$ is

$$b_i = \vec{x} \cdot \vec{v}_i$$

It follows that $\vec{x}$ can be written as

$$\vec{x} = (\vec{x} \cdot \vec{v}_1)\vec{v}_1 + \cdots + (\vec{x} \cdot \vec{v}_n)\vec{v}_n$$

Proof of Theorem 7.1.2: The proof of this is similar to the proof that orthogonal sets are linearly independent, except that instead of looking at a linear combination that is equal to the zero vector, we now look at a linear combination that is equal to $\vec{x}$. Since $\mathcal{B}$ is a basis for $\mathbb{R}^n$, we know that there are scalars $b_1, \dots, b_n$ such that

$$b_1\vec{v}_1 + \cdots + b_n\vec{v}_n = \vec{x}$$

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

$$\begin{aligned}
(b_1\vec{v}_1 + \cdots + b_n\vec{v}_n) \cdot \vec{v}_i &= \vec{x} \cdot \vec{v}_i \\
(b_1\vec{v}_1) \cdot \vec{v}_i + \cdots + (b_n\vec{v}_n) \cdot \vec{v}_i &= \vec{x} \cdot \vec{v}_i \\
b_1(\vec{v}_1 \cdot \vec{v}_i) + \cdots + b_i(\vec{v}_i \cdot \vec{v}_i) + \cdots + b_n(\vec{v}_n \cdot \vec{v}_i) &= \vec{x} \cdot \vec{v}_i \\
b_1(0) + \cdots + b_i(||\vec{v}_i||^2) + \cdots + b_n(0) &= \vec{x} \cdot \vec{v}_i \\
b_i &= \vec{x} \cdot \vec{v}_i
\end{aligned}$$

Example: To find the coordinates of $\vec{x} = \begin{bmatrix} -4 \\ -7 \end{bmatrix}$ with respect to the orthonormal basis $\mathcal{B}_1 = \left\{ \begin{bmatrix} 2/\sqrt{13} \\ -3/\sqrt{13} \end{bmatrix}, \begin{bmatrix} 6/\sqrt{52} \\ 4/\sqrt{52} \end{bmatrix} \right\}$ for $\mathbb{R}^2$, we calculate them individually:

$$b_1 = \begin{bmatrix} -4 \\ -7 \end{bmatrix} \cdot \begin{bmatrix} 2/\sqrt{13} \\ -3/\sqrt{13} \end{bmatrix} = \frac{-8}{\sqrt{13}} + \frac{21}{\sqrt{13}} = \frac{13}{\sqrt{13}}$$

$$b_2 = \begin{bmatrix} -4 \\ -7 \end{bmatrix} \cdot \begin{bmatrix} 6/\sqrt{52} \\ 4/\sqrt{52} \end{bmatrix} = \frac{-24}{\sqrt{52}} - \frac{28}{\sqrt{52}} = -\frac{52}{\sqrt{52}}$$

And so see that $[\vec{x}]_{\mathcal{B}_1} = \begin{bmatrix} 13/\sqrt{13} \\ -52/\sqrt{52} \end{bmatrix}$

To find the coordinates of $\vec{y} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$ with respect to the orthonormal basis $\mathcal{B}_2 = \left\{ \begin{bmatrix} 3/\sqrt{26} \\ 1/\sqrt{26} \\ 4/\sqrt{26} \end{bmatrix}, \begin{bmatrix} -1/\sqrt{3} \\ -1/\sqrt{3} \\ 1/\sqrt{3} \end{bmatrix}, \begin{bmatrix} 5/\sqrt{78} \\ -7/\sqrt{78} \\ -2/\sqrt{78} \end{bmatrix} \right\}$ for $\mathbb{R}^3$, we again calculate them individually:

$$b_1 = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} \cdot \begin{bmatrix} 3/\sqrt{26} \\ 1/\sqrt{26} \\ 4/\sqrt{26} \end{bmatrix} = \frac{2}{\sqrt{26}} + \frac{2}{\sqrt{26}} + \frac{12}{\sqrt{26}} = \frac{17}{\sqrt{26}}$$

$$b_2 = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} \cdot \begin{bmatrix} -1/\sqrt{3} \\ -1/\sqrt{3} \\ 1/\sqrt{3} \end{bmatrix} = -\frac{1}{\sqrt{3}} - \frac{2}{\sqrt{3}} + \frac{3}{\sqrt{3}} = 0$$

$$b_3 = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} \cdot \begin{bmatrix} 5/\sqrt{78} \\ -7/\sqrt{78} \\ -2/\sqrt{78} \end{bmatrix} = \frac{5}{\sqrt{78}} - \frac{14}{\sqrt{78}} - \frac{6}{\sqrt{78}} = -\frac{15}{\sqrt{78}}$$

And so see that $[\vec{y}]_{\mathcal{B}_2} = \begin{bmatrix} 17/\sqrt{26} \\ 0 \\ -15/\sqrt{78} \end{bmatrix}$

Looking at these strange coordinates you may be wondering if orthonormal bases are really worth the trouble, but in forcing the vectors in our basis to have length one, it turns out that coordinates now preserve length. That is, $||\vec{x}|| = ||[\vec{x}]_{\mathcal{B}}||$. In fact, we see that coordinates with respect to an orthonormal basis B will also preserve dot products.

Theorem 7.1.A: Let $\vec{x}$ and $\vec{y}$ be vectors in $\mathbb{R}^n$, let $\mathcal{B} = \{\vec{v}_1, \ldots, \vec{v}_n\}$ be an orthonormal basis for $\mathbb{R}^n$, and let $[\vec{x}]_{\mathcal{B}} = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$ and $[\vec{y}]_{\mathcal{B}} = \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix}$. Then

(1) $\vec{x} \cdot \vec{y} = [\vec{x}]_{\mathcal{B}} \cdot [\vec{y}]_{\mathcal{B}}$
(2) $||\vec{x}|| = ||[\vec{x}]_{\mathcal{B}}||$

Proof of Theorem 7.1.A: We see that

$$\begin{aligned}
\vec{x} \cdot \vec{y} = \; & (x_1\vec{v}_1 + \cdots + x_n\vec{v}_n) \cdot (y_1\vec{v}_1 + \cdots + y_n\vec{v}_n) \\
= \; & x_1\vec{v}_1 \cdot (y_1\vec{v}_1 + \cdots + y_n\vec{v}_n) + \cdots + x_n\vec{v}_n \cdot (y_1\vec{v}_1 + \cdots + y_n\vec{v}_n) \\
= \; & (x_1\vec{v}_1) \cdot (y_1\vec{v}_1) + \cdots + (x_1\vec{v}_1) \cdot (y_n\vec{v}_n) \\
& + \cdots + \\
& +(x_n\vec{v}_n) \cdot (y_1\vec{v}_1) + \cdots + (x_n\vec{v}_n) \cdot (y_n\vec{v}_n) \\
= \; & x_1y_1(\vec{v}_1 \cdot \vec{v}_1) + \cdots + (x_1y_n)(\vec{v}_1 \cdot \vec{v}_n) + x_2y_1(\vec{v}_2 \cdot \vec{v}_1) + \cdots + x_2y_n(\vec{v}_2 \cdot \vec{v}_n) \\
& + \cdots + \\
& +x_ny_1(\vec{v}_n \cdot \vec{v}_1) + \cdots + x_ny_n(\vec{v}_n \cdot \vec{v}_n) \\
& = x_1y_1 + x_2y_2 + \cdots + x_ny_n
\end{aligned}$$

since $\vec{v}_i \cdot \vec{v}_i = ||\vec{v}_i||^2 = 1^2 = 1$ and $\vec{v}_i \cdot \vec{v}_j = 0$ when $i \neq j$. As so we have that $\vec{x} \cdot \vec{y} = x_1y_1 + x_2y_2 + \cdots + x_ny_n = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} \cdot \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix} = [\vec{x}]_{\mathcal{B}} \cdot [\vec{y}]_{\mathcal{B}}$.

This proves part (1), and part (2) follows easily from (1), since $||x||^2 = \vec{x} \cdot \vec{x} = [\vec{x}]_{\mathcal{B}} \cdot [\vec{x}]_{\mathcal{B}} = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} \cdot \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = x_1^2 + \cdots + x_n^2$. And so we have $||x|| = \sqrt{x_1^2 + \cdots + x_n^2} = ||[\vec{x}]_{\mathcal{B}}||$.