Lecture 2f

Projections Onto A Subspace

(pages 334-336)

Now that we have explored what it means to be orthogonal to a set, we can return to our original question of how to make an orthonormal basis. We will construct such a basis one vector at a time, so for now let us assume that we have an orthonormal set $\{\vec{v}_1, \dots, \vec{v}_k\}$, and we want to find a vector $\vec{v}_{k+1}$ such that the set $\{\vec{v}_1, \dots, \vec{v}_k, \vec{v}_{k+1}\}$ is orthonormal. Well, if we can find any vector $\vec{x}$ such that $\vec{x} \notin \text{Span}\{\vec{v}_1, \dots, \vec{v}_k\}$, then we can split $\vec{x}$ into two pieces: the part of $\vec{x}$ that is in $\text{Span}\{\vec{v}_1, \dots, \vec{v}_k\}$, and the part of $\vec{x}$ that is orthogonal to $\text{Span}\{\vec{v}_1, \dots, \vec{v}_k\}$. We did something similar to this in Math 106, when we looked at the values $\text{proj}_{\vec{y}}\vec{x}$ and $\text{perp}_{\vec{y}}\vec{x}$. So what we want to do is expand these definitions to now look at the projection of a vector $\vec{x}$ onto a subspace $\mathbb{S}$, instead of just a vector. Recalling that $\text{proj}_{\vec{y}}\vec{x}$ is a scalar multiple of $\vec{y}$, we will now define $\text{proj}_{\mathbb{S}}\vec{x}$ to be a linear combination of the basis vectors for $\mathbb{S}$, as this would be the generalization of a scalar multiple of one vector.

But $\text{proj}_{\vec{y}}\vec{x}$ wasn't just any scalar multiple of $\vec{y}$. The scalar was $\frac{\vec{x}\cdot\vec{y}}{||\vec{y}||^2}$. So, to make sure that our definitions coincide in the case that $\mathbb{S}$ is the span of a single vector, we will take the scalars in our expanded definition to be $\frac{\vec{x}\cdot\vec{v}_i}{||\vec{v}_i||^2}$.

<u>Definition</u>: Let $\mathbb{S}$ be a k-dimensional subspace of $\mathbb{R}^n$ and let $\mathcal{B} = \{\vec{v}_1, \dots, \vec{v}_k\}$ be an orthogonal basis of $\mathbb{S}$. If $\vec{x}$ is any vector in $\mathbb{R}^n$, the **projection** of $\vec{x}$ onto $\mathbb{S}$ is defined to be

$$\text{proj}_{\mathbb{S}}\vec{x} = \frac{\vec{x}\cdot\vec{v}_1}{||\vec{v}_1||^2}\vec{v}_1 + \frac{\vec{x}\cdot\vec{v}_2}{||\vec{v}_2||^2}\vec{v}_2 + \cdots + \frac{\vec{x}\cdot\vec{v}_k}{||\vec{v}_k||^2}\vec{v}_k$$

Note that this definition only works if $\mathcal{B}$ is an orthogonal basis–we will not consider projections based on arbitrary bases.

Continuing to parallel our original construction of $\text{proj}_{\vec{y}}$ and $\text{perp}_{\vec{y}}$, we now define $\text{perp}_{\mathbb{S}}$ as follows:

<u>Definition</u>: The **projection of $\vec{x}$ perpendicular to** $\mathbb{S}$ is defined to be

$$\text{perp}_{\mathbb{S}}\vec{x} = \vec{x} - \text{proj}_{\mathbb{S}}\vec{x}$$

Now, our hope is that $\text{perp}_{\mathbb{S}}\vec{x}$ will in fact an element of $\mathbb{S}^\perp$. And it turns out that it is. To verify this, we will show that $\text{perp}_{\mathbb{S}}\vec{x}\cdot\vec{v}_i = 0$ for all vectors in our basis $\mathcal{B} = \{\vec{v}_1, \dots, \vec{v}_k\}$ for $\mathbb{S}$.

$$
\begin{aligned}
\text{perp}_{\mathbb{S}}\vec{x}\cdot\vec{v}_i &= (\vec{x}-\text{proj}_{\mathbb{S}}\vec{x})\cdot\vec{v}_i \\
&= \vec{x}\cdot\vec{v}_i - \text{proj}_{\mathbb{S}}\vec{x}\cdot\vec{v}_i \\
&= \vec{x}\cdot\vec{v}_i - \left(\tfrac{\vec{x}\cdot\vec{v}_1}{||\vec{v}_1||^2}\vec{v}_1 + \cdots + \tfrac{\vec{x}\cdot\vec{v}_k}{||\vec{v}_k||^2}\vec{v}_k\right)\cdot\vec{v}_i \\
&= \vec{x}\cdot\vec{v}_i - \left(\tfrac{\vec{x}\cdot\vec{v}_1}{||\vec{v}_1||^2}\vec{v}_1\cdot\vec{v}_i + \cdots + \tfrac{\vec{x}\cdot\vec{v}_k}{||\vec{v}_k||^2}\vec{v}_k\cdot\vec{v}_i\right) \\
&= \vec{x}\cdot\vec{v}_i - (0+\cdots+0+\tfrac{\vec{x}\cdot\vec{v}_i}{||\vec{v}_i||^2}\vec{v}_i\cdot\vec{v}_i+0+\cdots+0) \\
&= \vec{x}\cdot\vec{v}_i - \tfrac{\vec{x}\cdot\vec{v}_i}{||\vec{v}_i||^2}\vec{v}_i\cdot\vec{v}_i \\
&= \vec{x}\cdot\vec{v}_i - \tfrac{\vec{x}\cdot\vec{v}_i}{||\vec{v}_i||^2}(||\vec{v}_i||^2) \\
&= \vec{x}\cdot\vec{v}_i - \vec{x}\cdot\vec{v}_i \\
&= 0
\end{aligned}
$$

Example: Let $\mathcal{B} = \left\{ \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ -1 \end{bmatrix} \right\}$ be an orthogonal basis for $\mathbb{S}$ and let $\vec{x} = \begin{bmatrix} 4 \\ 5 \\ -2 \end{bmatrix}$. To determine $\text{proj}_{\mathbb{S}}\vec{x}$ and $\text{perp}_{\mathbb{S}}\vec{x}$, we will first want to do the following calculations:

$$
\begin{bmatrix} 4 \\ 5 \\ -2 \end{bmatrix}\cdot\begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} = 4+10-6=8 \qquad \left|\left|\begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}\right|\right|^2 = 1^2+2^2+3^2=14
$$

$$
\begin{bmatrix} 4 \\ 5 \\ -2 \end{bmatrix}\cdot\begin{bmatrix} 1 \\ 1 \\ -1 \end{bmatrix} = 4+5+2=11 \qquad \left|\left|\begin{bmatrix} 1 \\ 1 \\ -1 \end{bmatrix}\right|\right|^2 = 1^2+1^2+(-1)^2=3
$$

Then we have that

$$
\begin{aligned}
\text{proj}_{\mathbb{S}}\vec{x} &= \tfrac{8}{14}\begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} + \tfrac{11}{3}\begin{bmatrix} 1 \\ 1 \\ -1 \end{bmatrix} \\
&= \tfrac{12}{21}\begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} + \tfrac{77}{21}\begin{bmatrix} 1 \\ 1 \\ -1 \end{bmatrix} \\
&= \begin{bmatrix} 89/21 \\ 101/21 \\ -41/21 \end{bmatrix}
\end{aligned}
$$

and this means that

$$\begin{aligned}
\text{perp}_{\mathbb{S}}\vec{x} &= \vec{x} - \text{proj}_{\mathbb{S}}\vec{x} \\
&= \begin{bmatrix} 4 \\ 5 \\ -2 \end{bmatrix} - \begin{bmatrix} 89/21 \\ 101/21 \\ -41/21 \end{bmatrix} \\
&= \begin{bmatrix} -5/21 \\ 4/21 \\ -1/21 \end{bmatrix}
\end{aligned}$$

We can verify our calculations by noticing that

$$\begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} \cdot \begin{bmatrix} -5/21 \\ 4/21 \\ -1/21 \end{bmatrix} = -\frac{5}{21} + \frac{8}{21} - \frac{3}{21} = 0 \quad \text{and} \quad \begin{bmatrix} 1 \\ 1 \\ -1 \end{bmatrix} \cdot \begin{bmatrix} -5/21 \\ 4/21 \\ -1/21 \end{bmatrix} = -\frac{5}{21} + \frac{4}{21} + \frac{1}{21} = 0$$

Before we finally move on to our algorithm for constructing an orthonormal basis, we want to notice one last feature of $\text{proj}_{\mathbb{S}}\vec{x}$, and that is that it is the vector in $\mathbb{S}$ that is closest to $\vec{x}$.

Theorem 7.2.2 (Approximation Theorem): Let $\mathbb{S}$ be a subspace of $\mathbb{R}^n$. Then, for any $\vec{x} \in \mathbb{R}^n$, the unique vector $\vec{s} \in \mathbb{S}$ that minimizes the distance $||\vec{x} - \vec{s}||$ is $\vec{s} = \text{proj}_{\mathbb{S}}\vec{x}$.

Proof of the Approximation Theorem: Let $\{\vec{v}_1, \ldots, \vec{v}_k\}$ be an orthonormal basis for $\mathbb{S}$ and let $\{\vec{v}_{k+1}, \ldots, \vec{v}_n\}$ be an orthonormal basis for $\mathbb{S}^\perp$, so that $\mathcal{B} = \{\vec{v}_1, \ldots, \vec{v}_n\}$ is an orthonormal basis for $\mathbb{R}^n$. Then for any $\vec{x} \in \mathbb{R}^n$, there are scalars $x_1, \ldots, x_n$ such that

$$\vec{x} = x_1\vec{v}_1 + \cdots + x_k\vec{v}_k + x_{k+1}\vec{v}_{k+1} + \cdots + x_n\vec{v}_n$$

Let $\vec{s} = s_1\vec{v}_1 + \cdots + s_k\vec{v}_k$ be an element of $\mathbb{S}$. Then we also have that $\vec{s} = s_1\vec{v}_1 + \cdots + s_k\vec{v}_k + 0\vec{v}_{k+1} + \cdots + 0\vec{v}_n$, and so we can write $\vec{x} - \vec{s}$ as:

$$\begin{aligned}
\vec{x} - \vec{s} &= (x_1\vec{v}_1 + \cdots + x_k\vec{v}_k + x_{k+1}\vec{v}_{k+1} + \cdots + x_n\vec{v}_n) - (s_1\vec{v}_1 + \cdots + s_k\vec{v}_k + 0\vec{v}_{k+1} + \cdots + 0\vec{v}_n) \\
&= (x_1 - s_1)\vec{v}_1 + \cdots + (x_k - s_k)\vec{v}_k + x_{k+1}\vec{v}_{k+1} + \cdots + x_n\vec{v}_n
\end{aligned}$$

In order to minimize $||\vec{x} - \vec{s}||$, we will minimize the easier to calculate $||\vec{x} - \vec{s}||^2$. Recall that since $\mathcal{B}$ is an orthonormal basis, and as we have written $\vec{x} - \vec{s}$ in terms of $\mathcal{B}$ coordinates, we can still calculate $||\vec{x} - \vec{s}||^2$ the usual way–we sum the square of the coefficients. And so we see that

$$||\vec{x} - \vec{s}||^2 = (x_1 - s_1)^2 + \cdots + (x_k - s_k)^2 + x_{k+1}^2 + \cdots + x_n^2$$

And clearly this value is minimized by setting $s_i = x_i$, so that $x_i - s_i = 0$ for $i = 1, \ldots, k$. This means we have shown that the vector $\vec{s}$ in $\mathbb{S}$ that minimizes the distance $||\vec{x} - \vec{s}||$ is

$$\vec{s} = x_1\vec{v}_1 + \cdots + x_k\vec{v}_k$$

To see that this is equal to $\text{proj}_{\mathcal{S}}\vec{x}$, let's first recall that

$$\text{proj}_{\mathcal{S}}\vec{x} = (\vec{x} \cdot \vec{v}_1)\vec{v}_1 + \cdots + (\vec{x} \cdot \vec{v}_k)\vec{v}_k$$

since $\{\vec{v}_1, \ldots, \vec{v}_k\}$ is an orthonormal basis for $\mathbb{S}$. And then we notice that, for any $i = 1, \ldots, k$, we get

$$\begin{aligned} \vec{x} \cdot \vec{v}_i \quad &= (x_1\vec{v}_1 + \cdots + x_k\vec{v}_k + x_{k+1}\vec{v}_{k+1} + \cdots + x_n\vec{v}_n) \cdot \vec{v}_i \\ &= x_1(\vec{v}_1 \cdot \vec{v}_i) + \cdots + x_i(\vec{v}_i \cdot \vec{v}_i) + \cdots + x_k(\vec{v}_k \cdot \vec{v}_i) + x_{k+1}(\vec{v}_{k+1} \cdot \vec{v}_i) + \cdots + x_n(\vec{v}_n \cdot \vec{v}_i) \\ &= 0 + \cdots + 0 + x_i||\vec{v}_i||^2 + 0 + \cdots + 0 \\ &= x_i \end{aligned}$$

where we know that $\vec{v}_j \cdot \vec{v}_i = 0$ when $j \neq i$ because $\{\vec{v}_1, \ldots, \vec{v}_n\}$ is an orthonormal set, and $x_i||\vec{v}_i||^2 = x_i$ because $||\vec{v}_i|| = 1$. And so we see that

$$\begin{aligned} \text{proj}_{\mathbb{S}}\vec{x} \quad &= (\vec{x} \cdot \vec{v}_1)\vec{v}_1 + \cdots + (\vec{x} \cdot \vec{v}_k)\vec{v}_k \\ &= x_1\vec{v}_1 + \cdots + x_k\vec{v}_k \\ &= \vec{s} \text{ (our minimum distance vector calculated above)} \end{aligned}$$