Lecture 2e

Orthogonal Complement

(pages 333-334)

We have now seen that an orthonormal basis is a nice way to describe a subspace, but knowing that we want an orthonormal basis doesn't make one fall into our lap. In theory, the process for finding an orthonormal basis is easy. Start with one vector, add a vector that in the subspace that is orthogonal to your first vector, then add a vector in the subspace that is orthogonal to the first two vectors. The problem comes when you start to deal with very large spaces, and you are trying to find a vector that is orthogonal to the 37 vectors that came before it! (Especially when you consider that these vectors must have at least 38 entries.) On the face of it, we already have all the techniques we need, since the vector equation $\vec{x} \cdot \vec{v} = 0$ is simply a homogeneous linear equation whose variables are the entries of $\vec{x}$. So, if we needed to make $\vec{x}$ orthogonal to 37 vectors, we would be looking at solving a system of 37 linear equations. But that would be a *really* big matrix to row reduce (actually, it wouldn't–most people who use these techniques deal with way more variables and equations), so if we can find an easier way, we should!

Our first step on this journey will be to add some more definitions to our vocabulary. For, while we have already defined what it means for one vector to be orthogonal to another, we are now looking at having our vector be orthogonal to a SET of vectors.

Definition: Let $\mathbb{S}$ be a subspace of $\mathbb{R}^n$. We shall say that a vector $\vec{x}$ is **orthogonal** to $\mathbb{S}$ if

$$\vec{x} \cdot \vec{s} = 0 \quad \text{for all } \vec{s} \in \mathbb{S}$$

Example: Let $\mathbb{S} = \text{Span}\left\{ \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix} \right\}$. Then $\begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix}$ is orthogonal to $\mathbb{S}$, because given any element $a \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} + b \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} a \\ 0 \\ 0 \\ b \end{bmatrix}$ of Span $\left\{ \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix} \right\}$, we see that $\begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix} \cdot \begin{bmatrix} a \\ 0 \\ 0 \\ b \end{bmatrix} = 0$. We also see that $\begin{bmatrix} 0 \\ 3 \\ 0 \\ 0 \end{bmatrix}$ is orthogonal to $\mathbb{S}$, since

$$\begin{bmatrix} 0 \\ 3 \\ 0 \\ 0 \end{bmatrix} \cdot \begin{bmatrix} a \\ 0 \\ 0 \\ b \end{bmatrix} = 3 \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix} \cdot \begin{bmatrix} a \\ 0 \\ 0 \\ b \end{bmatrix} = 3(0) = 0$$

It is also easy to notice that $\begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \end{bmatrix}$ is orthogonal to $\mathbb{S}$, since we also have that $\begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \end{bmatrix} \cdot \begin{bmatrix} a \\ 0 \\ 0 \\ b \end{bmatrix} = 0$. But $\begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}$ is not orthogonal to $\mathbb{S}$, since $\begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}$ is an element of $\mathbb{S}$, but $\begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} \cdot \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} = 1 \neq 0$. (As usual, the best way to show that $\vec{x}$ is *not* orthogonal to a subspace $\mathbb{S}$ is to find a specific element $\vec{s}$ of $\mathbb{S}$ such that $\vec{x} \cdot \vec{s} \neq 0$.)

Note that in order to show that $\vec{x}$ is orthogonal to a subspace $\mathbb{S}$, we only need to show that $\vec{x}$ is orthogonal to the basis vectors for $\mathbb{S}$. In fact, we can show that $\vec{x}$ is orthogonal to any spanning set for $\mathbb{S}$–we don't need to worry about the independence of the vectors for this property. For, if $\mathcal{A} = \{\vec{a}_1, \ldots, \vec{a}_k\}$ is a spanning for $\mathbb{S}$, then every element $\vec{s}$ of $\mathbb{S}$ can be written as $s_1\vec{a}_1 + \cdots + s_k\vec{a}_k$ for some $s_1, \ldots, s_k \in \mathbb{R}$. And if $\vec{x}$ is orthogonal every $\vec{a}_i$ for $1 \leq i \leq k$, then we have the following:

$$\begin{aligned} \vec{x} \cdot \vec{s} &= \vec{x} \cdot (s_1\vec{a}_1 + \cdots + s_k\vec{a}_k) \\ &= (\vec{x} \cdot s_1\vec{a}_1) + \cdots + (\vec{x} \cdot s_k\vec{a}_k) \\ &= s_1(\vec{x} \cdot \vec{a}_1) + \cdots + s_k(\vec{x} \cdot \vec{a}_k) \\ &= s_1(0) + \cdots + s_k(0) \\ &= 0 \end{aligned}$$

We also see that, if $\vec{x}$ is orthogonal to $\mathbb{S}$, then $t\vec{x}$ is orthogonal to $\mathbb{S}$ for all scalars $t \in \mathbb{R}$, since $(t\vec{x}) \cdot \vec{s} = t(\vec{x} \cdot \vec{s}) = t(0) = 0$. Also, if both $\vec{x}$ and $\vec{y}$ are orthogonal to $\mathbb{S}$, then so is $\vec{x} + \vec{y}$, since $(\vec{x} + \vec{y}) \cdot \vec{s} = (\vec{x} \cdot \vec{s}) + (\vec{y} \cdot \vec{s}) = 0 + 0 = 0$. If we add to these facts the fact that $\vec{0}$ is orthogonal to $\mathbb{S}$, since $\vec{0} \cdot \vec{v} = 0$ for *any* vector $\vec{v}$, and thus definitely for the ones in $\mathbb{S}$, then we have shown that the set of all vectors orthogonal to $\mathbb{S}$ is never the empty set. And this means that we have shown that the set of all vectors orthogonal to $\mathbb{S}$ is itself a subspace of $\mathbb{R}^n$, since it is a non-empty subset of $\mathbb{R}^n$ that is closed under addition and scalar multiplication. We give this subset the following name:

<u>Definition</u>: We call the set of all vectors orthogonal to $\mathbb{S}$ the **orthogonal complement** of $\mathbb{S}$ and denote it $\mathbb{S}^{\perp}$. That is

$$\mathbb{S}^{\perp} = \{\vec{x} \in \mathbb{R}^n \mid \vec{x} \cdot \vec{s} = 0 \text{ for all } \vec{s} \in \mathbb{S}\}$$

Example: Find a basis for $\mathbb{S}^{\perp}$, where $\mathbb{S} = \text{Span}\left\{ \begin{bmatrix} 1 \\ 2 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 3 \\ 6 \\ 1 \\ 4 \end{bmatrix} \right\}$.

Again, a vector is orthogonal to $\mathbb{S}$ if it is orthogonal to the vectors in its spanning set, so we are looking for vectors $\vec{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix}$ such that $\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} \cdot \begin{bmatrix} 1 \\ 2 \\ 0 \\ 1 \end{bmatrix} = 0$ and $\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} \cdot \begin{bmatrix} 3 \\ 6 \\ 1 \\ 4 \end{bmatrix} = 0$. This is the same as looking for solutions to the following system:

$$\begin{array}{rrrrl} x_1 & +2x_2 & & +x_4 & = 0 \\ 3x_1 & +6x_2 & +x_3 & +4x_4 & = 0 \end{array}$$

To solve this homogeneous system, we row reduce its coefficient matrix:

$$\begin{bmatrix} 1 & 2 & 0 & 1 \\ 3 & 6 & 1 & 4 \end{bmatrix} \sim \begin{bmatrix} 1 & 2 & 0 & 1 \\ 0 & 0 & 1 & 1 \end{bmatrix}$$

From our RREF matrix, we see that our system is equivalent to

$$\begin{array}{rrrrl} x_1 & +2x_2 & & +x_4 & = 0 \\ & & x_3 & +x_4 & = 0 \end{array}$$

Replacing the variable x_2 with the parameter s and the variable x_4 with the parameter t, we get that the general solution to this system is

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} -2s - t \\ s \\ -t \\ t \end{bmatrix} = s \begin{bmatrix} -2 \\ 1 \\ 0 \\ 0 \end{bmatrix} + t \begin{bmatrix} -1 \\ 0 \\ -1 \\ 1 \end{bmatrix}$$

The general solution to our system is a list of all the vectors $\vec{x}$ that are orthogonal to $\mathbb{S}$, so we see that $\left\{ \begin{bmatrix} -2 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} -1 \\ 0 \\ -1 \\ 1 \end{bmatrix} \right\}$ is a spanning set for $\mathbb{S}^{\perp}$. Moreover, these vectors are not a scalar multiple of each other, and thus are linearly independent, so we have that $\left\{ \begin{bmatrix} -2 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} -1 \\ 0 \\ -1 \\ 1 \end{bmatrix} \right\}$ is a basis for $\mathbb{S}^\perp$.

In our example, we were easily able to show that our spanning set was linearly independent because it only had two vectors in it. But what if our spanning set was bigger? Would we need to row reduce a matrix to prove that our spanning set is linearly independent? No! Because our basis for $\mathbb{S}^\perp$ can always be thought of as the nullspace for a matrix, and back in Math 106 we showed that the technique for finding the spanning set for the nullspace in fact results in a basis for the nullspace. And which matrix are we finding the nullspace of? It is the matrix whose ROWS are the vectors in the spanning set for our subset. NOT THE COLUMNS!!! I recommend going ahead and setting up the system of equations using the defining dot products ($\vec{x} \cdot \vec{s}_1$, $\vec{x} \cdot \vec{s}_2$, etc.) to avoid confusion on this matter. But we will use this fact when we prove the following theorem.

Theorem 7.2.1: Let $\mathbb{S}$ be a k-dimensional subspace of $\mathbb{R}^n$. Then

(1) $\mathbb{S} \cap \mathbb{S}^\perp = \{\vec{0}\}$
(2) $\dim(\mathbb{S}^\perp) = n - k$
(3) If $\{\vec{v}_1, \ldots, \vec{v}_k\}$ is an orthonormal basis for $\mathbb{S}$ and $\{\vec{v}_{k+1}, \ldots, \vec{v}_n\}$ is an orthonormal basis for $\mathbb{S}^\perp$, then $\{\vec{v}_1, \ldots, \vec{v}_k, \vec{v}_{k+1}, \ldots, \vec{v}_n\}$ is an orthonormal basis for $\mathbb{R}^n$.

Proof of Theorem 7.2.1: To see that $\mathbb{S} \cap \mathbb{S}^\perp = \{\vec{0}\}$, let $\vec{x} \in \mathbb{S} \cap \mathbb{S}^\perp$. Then $\vec{x}$ is an element of $\mathbb{S}^\perp$, so $\vec{x}$ is orthogonal to every element of $\mathbb{S}$. But we also have that $\vec{x}$ is an element of $\mathbb{S}$, so this means that $\vec{x}$ is orthogonal to itself. That is, $\vec{x} \cdot \vec{x} = 0$, which means that $\vec{x} = \vec{0}$. (See Theorem 1.3.1.)

Next, to see that $\dim(\mathbb{S}^\perp) = n - k$, let A be the matrix whose ROWS are the basis vectors of $\mathbb{S}$. Then A is a $k \times n$ matrix, and $\mathbb{S}$ is the rowspace of A. This means that the rank of A is the same as the dimension of $\mathbb{S}$, so $\operatorname{rank}(A) = k$. But we also have that $\mathbb{S}^\perp$ is the nullspace of A, and thus the dimension of $\mathbb{S}^\perp$ is the nullity of A. By the Rank Theorem, we know that $\operatorname{rank}(A) + \operatorname{nullity}(A) = n$, so the $\dim(\mathbb{S}^\perp) = \operatorname{nullity}(A) = n - \operatorname{rank}(A) = n - k$.

Finally, to see that $\{\vec{v}_1, \ldots, \vec{v}_k, \vec{v}_{k+1}, \ldots, \vec{v}_n\}$ is an orthonormal basis for $\mathbb{R}^n$, remember that we in fact only need to show that $\{\vec{v}_1, \ldots, \vec{v}_k, \vec{v}_{k+1}, \ldots, \vec{v}_n\}$ is an orthonormal set (as it will then automatically be a basis). That means we need to show that $\vec{v}_i \cdot \vec{v}_j = 0$ whenever $i \neq j$. We will break this into four different scenarios:

(a) $1 \leq i, j \leq k$. Then both $\vec{v}_i$ and $\vec{v}_j$ are in $\{\vec{v}_1, \ldots, \vec{v}_k\}$, which is an orthonormal set, so we know that $\vec{v}_i \cdot \vec{v}_j = 0$.

(b) $1 \leq i \leq k$ and $k+1 \leq j \leq n$. Then $\vec{v}_i \in \mathbb{S}$ and $\vec{v}_j \in \mathbb{S}^\perp$, so by the definition of $\mathbb{S}^\perp$ we know that $\vec{v}_i \cdot \vec{v}_j = 0$.

(c) $1 \leq j \leq k$ and $k+1 \leq i \leq n$. Then $\vec{v}_j \in \mathbb{S}$ and $\vec{v}_i \in \mathbb{S}^\perp$, so by the definition of $\mathbb{S}^\perp$ we know that $\vec{v}_i \cdot \vec{v}_j = 0$.

(d) $k+1 \leq i, j \leq n$. Then both $\vec{v}_i$ and $\vec{v}_j$ are in $\{\vec{v}_{k+1}, \ldots, \vec{v}_n\}$, which is an orthonormal set, so we know that $\vec{v}_i \cdot \vec{v}_j = 0$.