Math 313 Lecture #26
§6.2: Orthogonal Sets

Orthogonal vectors have many properties that make them great for computations.

A set $\{\vec{v}_1, \ldots, \vec{v}_n\}$ in $\mathbb{R}^n$ is said to be an **orthogonal** set if

$$\vec{v}_i \cdot \vec{v}_j = 0 \text{ when } i \neq j.$$

Example. In $\mathbb{R}^4$, the vectors

$$\vec{u}_1 = \begin{bmatrix} 1 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \ \vec{u}_2 = \begin{bmatrix} 2 \\ -2 \\ 1 \\ -1 \end{bmatrix}, \ \vec{u}_3 = \begin{bmatrix} 0 \\ 0 \\ 3 \\ 3 \end{bmatrix},$$

form an orthogonal set.

Theorem 4. If $\{\vec{v}_1, \ldots, \vec{v}_n\}$ is an orthogonal set of nonzero vectors in $\mathbb{R}^n$, then the vectors $\vec{v}_1, \ldots, \vec{v}_n$ are linearly independent.

Proof. Suppose that $\{\vec{v}_1, \ldots, \vec{v}_n\}$ is an orthogonal set, i.e., $\vec{v}_i \cdot \vec{v}_j = 0$ when $i \neq j$, and set

$$c_1\vec{v}_1 + \cdots + c_n\vec{v}_n = \vec{0}.$$

For $1 \leq j \leq n$, taking the inner product of both sides of the linear combination equation with $\vec{v}_j$ gives

$$c_1\vec{v}_1 \cdot \vec{v}_j + \cdots + c_j\vec{v}_j \cdot \vec{v}_j + \cdots + c_n\vec{v}_n \cdot \vec{v}_j = \vec{0} \cdot \vec{v}_j = 0.$$

All of the inner products here except $\vec{v}_j \cdot \vec{v}_j$ are zero by orthogonality, and so we get

$$c_j\vec{v}_j \cdot \vec{v}_j = 0.$$

Since $\vec{v}_j \neq 0$, then $0 \neq \|\vec{v}_j\|^2 = \vec{v}_j \cdot \vec{v}_j$, and so $c_j = 0$.

This is true for any $1 \leq j \leq n$, and so the vectors $\vec{v}_1, \ldots, \vec{v}_n$ are linearly independent. □

An orthogonal set $\{\vec{v}_1, \ldots, \vec{v}_n\}$ in $\mathbb{R}^n$ is said to be an **orthonormal set** if

$$\|\vec{v}_i\| = 1 \text{ for all } i = 1, \ldots, n.$$

By Theorem 4, if $\mathcal{B} = \{\vec{u}_1, \ldots, \vec{u}_n\}$ is an orthonormal set in $\mathbb{R}^n$, then $\mathcal{B}$ is a basis for the subspace $W = \text{Span}\{\vec{u}_1, \ldots, \vec{u}_n\}$ of $\mathbb{R}^n$.

In this case, we say that B is an **orthonormal basis** for W.

Example (Continued). In $\mathbb{R}^4$, the set of orthogonal vectors

$$\vec{u}_1 = \begin{bmatrix} 1 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \ \vec{u}_2 = \begin{bmatrix} 2 \\ -2 \\ 1 \\ -1 \end{bmatrix}, \ \vec{u}_3 = \begin{bmatrix} 0 \\ 0 \\ 3 \\ 3 \end{bmatrix},$$

can be made into an orthonormal set by scaling each vector by the reciprocal of it length:

$$\vec{v}_1 = \frac{\vec{u}_1}{\|\vec{u}_1\|} = \begin{bmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \\ 0 \\ 0 \end{bmatrix}, \ \vec{v}_2 = \frac{\vec{u}_2}{\|\vec{u}_2\|} = \begin{bmatrix} 2/\sqrt{10} \\ -2/\sqrt{10} \\ 1/\sqrt{10} \\ -1/\sqrt{10} \end{bmatrix}, \ \vec{v}_3 = \frac{\vec{u}_3}{\|\vec{u}_3\|} = \begin{bmatrix} 0 \\ 0 \\ 3/\sqrt{18} \\ -3/\sqrt{18} \end{bmatrix}.$$

Then $\{\vec{v}_1, \vec{v}_2, \vec{v}_3\}$ is an orthonormal basis for $\text{Span}\{\vec{u}_1, \vec{u}_2, \vec{u}_3\}$.

[There is a way to take any basis of a subspace and convert it into an orthonormal basis. This way is called the Gram-Schmidt process that we will learn in a future lecture.]

Orthogonality of a basis for a subspace means that computing the coordinates of any vector in that subspace can be done using only inner products.

Theorem 5. Let $\{\vec{u}_1, \vec{u}_2, \dots, \vec{u}_p\}$ be an orthogonal basis for a subspace W of $\mathbb{R}^n$. For each $\vec{y}$ in W, the unique weights in the linear combination

$$\vec{y} = c_1\vec{u}_1 + c_2\vec{u}_2 + \cdots + c_p\vec{u}_p$$

are given by

$$c_j = \frac{\vec{y} \cdot \vec{u}_j}{\vec{u}_j \cdot \vec{u}_j}, \ j = 1, 2, \dots, p.$$

Proof. We make use of the orthogonality of the basis for W when we compute

$$\vec{y} \cdot \vec{v}_j = c_1\vec{u}_1 \cdot \vec{v}_j + \cdots + c_j\vec{u}_j \cdot \vec{u}_j + \cdots + c_p\vec{u}_p \cdot \vec{v}_j = c_j\vec{u}_j \cdot \vec{u}_j.$$

Since $\{\vec{u}_1, \dots, \vec{u}_p\}$ is a basis, the vector $\vec{u}_j \neq 0$, so that $\vec{u}_j \cdot \vec{u}_j = \|\vec{u}_j\| \neq 0$.

Solving for c_j gives the desired formula. □

Example. Express

$$\vec{y} = \begin{bmatrix} 5 \\ -3 \\ 11 \\ 7 \end{bmatrix}$$

as a linear combination of the orthogonal set of vectors

$$\vec{u}_1 = \begin{bmatrix} 1 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \ \vec{u}_2 = \begin{bmatrix} 2 \\ -2 \\ 1 \\ -1 \end{bmatrix}, \ \vec{u}_3 = \begin{bmatrix} 0 \\ 0 \\ 3 \\ 3 \end{bmatrix}.$$

We compute the weights by inner products:

$$c_1 = \frac{\vec{y} \cdot \vec{u}_1}{\vec{u}_1 \cdot \vec{u}_1} = \frac{2}{2} = 1, \ c_2 = \frac{\vec{y} \cdot \vec{u}_2}{\vec{u}_2 \cdot \vec{u}_2} = \frac{20}{10} = 2, \ c_3 = \frac{\vec{y} \cdot \vec{u}_3}{\vec{u}_3 \cdot \vec{u}_3} = \frac{54}{18} = 3.$$

We can indeed verify this by computing $c_1\vec{u}_1 + c_2\vec{u}_2 + c_3\vec{u}_3$ to see if it equals $\vec{y}$.

An Orthogonal Projection. Let $\vec{u}$ be a nonzero vector in $\mathbb{R}^n$.

We would like to write a vector $\vec{y}$ in $\mathbb{R}^n$ as $\vec{y} = \vec{p} + \vec{z}$ where $\vec{p}$ is in the direction of $\vec{u}$ and $\vec{z}$ is orthogonal to $\vec{p}$.

The vector $\vec{p}$ is a scalar multiple of $\vec{u}$: there is a scalar α such that $\vec{p} = \alpha\vec{u}$.

With the vector $\vec{z} = \vec{y} - \vec{p} = \vec{y} - \alpha\vec{u}$ being orthogonal to $\vec{u}$, we have

$$0 = (\vec{y} - \alpha\vec{u}) \cdot \vec{u} = \vec{y} \cdot \vec{u} - \alpha\vec{u} \cdot \vec{u}.$$

Since $\vec{u} \neq 0$, we have $\vec{u} \cdot \vec{u} = \|\vec{u}\|^2 \neq 0$, so that we can solve for α getting

$$\alpha = \frac{\vec{y} \cdot \vec{u}}{\vec{u} \cdot \vec{u}}.$$

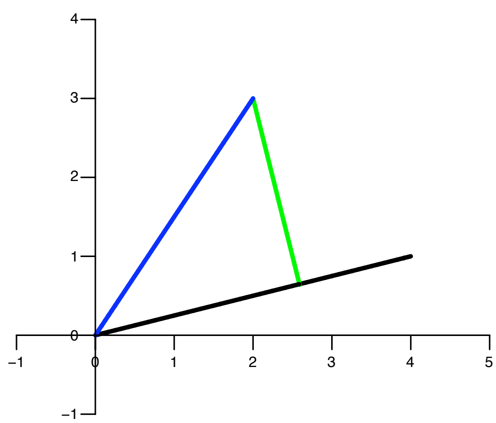
Having the value of α, we then have

$$\vec{p} = \alpha\vec{u} = \frac{\vec{y} \cdot \vec{u}}{\vec{u} \cdot \vec{u}}\vec{u}.$$

This vector $\vec{p}$ is the orthogonal projection of $\vec{y}$ onto the subspace L spanned by $\vec{u}$, and is denoted by

$$\text{proj}_L\vec{y} = \frac{\vec{y} \cdot \vec{u}}{\vec{u} \cdot \vec{u}}\vec{u}.$$

We illustrate this situation in $\mathbb{R}^2$: the black colored vector (the one pointing more right than up) is $\vec{u}$, the blue colored vector (the one pointing more up than right) is $\vec{y}$, and the the green colored on (pointing more up than left) is $\vec{z} = \vec{u} - \vec{p}$.



More on Orthonormal Sets. Notice that the formula for a projection simplifies if $\vec{u}$ is a unit vector, and notice also that the formula for the weights in Theorem 5 also simplify when the orthogonal set is an orthonormal set.

If we form a matrix U from columns vectors, we can detect if the columns form an orthonormal set by computing a certain matrix product.

Theorem 6. An $m \times n$ matrix U has orthonormal columns if and only if $U^T U = I$.

Proof. When we write $U = \begin{bmatrix} \vec{u}_1 & \vec{u}_2 & \cdots & \vec{u}_n \end{bmatrix}$ we have

$$U^T U = \begin{bmatrix} \vec{u}_1^T \\ \vec{u}_2^T \\ \vdots \\ \vec{u}_n^T \end{bmatrix} \begin{bmatrix} \vec{u}_1 & \vec{u}_2 & \cdots & \vec{u}_n \end{bmatrix} = \begin{bmatrix} \vec{u}_1^T \vec{u}_1 & \vec{u}_1^T \vec{u}_2 & \cdots & \vec{u}_1^T \vec{u}_n \\ \vec{u}_2^T \vec{u}_1 & \vec{u}_2^T \vec{u}_2 & \cdots & \vec{u}_2^T \vec{u}_n \\ \vdots & \vdots & \ddots & \vdots \\ \vec{u}_n^T \vec{u}_1 & \vec{u}_n^T \vec{u}_2 & \cdots & \vec{u}_n^T \vec{u}_n \end{bmatrix}.$$

We recognize that each entry of $U^T U$ is an inner product of two of the vector $\vec{u}_1, \vec{u}_2, \ldots, \vec{v}_n$.

When the set $\{\vec{u}_1, \vec{u}_2, \ldots, \vec{u}_n\}$ is an orthonormal set, then $U^T U = I$.

On the other hand, if $U^T U = I$, then $\{\vec{u}_1, \vec{u}_2, \ldots, \vec{u}_n\}$ is an orthonormal set. $\square$