15. Lecture 15: 5.1 Orthonormal Bases and Orthogonal Projection

The **dot product** between two vectors $\mathbf{x}=\begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$ and $\mathbf{y}=\begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix}$ is $\mathbf{x}\cdot\mathbf{y} = x_1y_1 + \cdots + x_ny_n$.
The **length** of the vector $\mathbf{x}$ is $\|\mathbf{x}\| = \sqrt{\mathbf{x}\cdot\mathbf{x}} = \sqrt{x_1^2 + \cdots + x_n^2}$.
Two vectors $\mathbf{x}$ and $\mathbf{y}$ are said to be **orthogonal** if $\mathbf{x}\cdot\mathbf{y} = 0$.
If θ is the angle between $\mathbf{x}$ and $\mathbf{y}$ then $\mathbf{x}\cdot\mathbf{y} = \|\mathbf{x}\|\,\|\mathbf{y}\|\cos\theta$.
Cauchy-Schwarz inequality: $|\mathbf{x}\cdot\mathbf{y}| \le \|\mathbf{x}\|\,\|\mathbf{y}\|$.
The **distance** between $\mathbf{x}$ and $\mathbf{y}$ is $\|\mathbf{y}-\mathbf{x}\|$.
The **Pythagorean law:** says that $\|\mathbf{x}+\mathbf{y}\|^2 = \|\mathbf{x}\|^2 + \|\mathbf{y}\|^2$ if and only if $\mathbf{x}\cdot\mathbf{y} = 0$. This follows since $\|\mathbf{x}+\mathbf{y}\|^2 = (\mathbf{x}+\mathbf{y})\cdot(\mathbf{x}+\mathbf{y}) = \mathbf{x}\cdot\mathbf{x}+\mathbf{y}\cdot\mathbf{y}+\mathbf{x}\cdot\mathbf{y}+\mathbf{y}\cdot\mathbf{x} = \|\mathbf{x}\|^2+\|\mathbf{y}\|^2+2\mathbf{x}\cdot\mathbf{y}$.

$\mathbf{z}$ is said to be **orthogonal to** a subspace W if it is orthogonal to every vector in W.
The set of all vectors $\mathbf{z}$ that are orthogonal to a subspace $W \subset \mathbf{R}^n$ is called the **orthogonal complement** of W and is denoted by $W^\perp$.

$$W^\perp = \{\mathbf{z}\in\mathbf{R}^n;\ \mathbf{z}\cdot\mathbf{y} = 0, \quad \text{for every} \quad \mathbf{y}\in W\}$$

Ex If W is plane through the origin in $\mathbf{R}^3$ and L is the line through the origin perpendicular to W, then $W^\perp = L$. In fact, clearly $L\subset W^\perp$ since L is perpendicular to W and any vector not in L is not perpendicular to W. Similarly $L^\perp = W$.
Ex If $V = \{\mathbf{x}\in\mathbf{R}^3;\ \mathbf{x} = \alpha(1,1,1),\ \text{for some } \alpha\}$ then $V^\perp = \{\mathbf{y}\in\mathbf{R}^3;\ \alpha(1,1,1)\cdot\mathbf{y} = 0,\ \text{for every } \alpha\} = \{\mathbf{y}\in\mathbf{R}^3;\ y_1+y_2+y_3 = 0\}$.
(1) If $W = \text{Span}\{\mathbf{v}_1,\dots,\mathbf{v}_k\}$ then $\mathbf{z}\in W^\perp$ if and only if $\mathbf{z}\cdot\mathbf{v}_1 = \cdots = \mathbf{z}\cdot\mathbf{v}_k = 0$.
(2) $W^\perp$ is a subspace.
(3) Every vector $\mathbf{x}\in\mathbf{R}^3$ can be uniquely written as $\mathbf{x} = \mathbf{y}+\mathbf{z}$, where $\mathbf{y}\in W$ and $\mathbf{z}\in W^\perp$.

Orthogonal Sets

A set of vectors $\{\mathbf{u}_1, \ldots, \mathbf{u}_p\}$ in called an **orthogonal set** if $\mathbf{u}_i \cdot \mathbf{u}_j = 0$ when $i \neq j$.

Ex Show that $\mathbf{u}_1 = \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}$, $\mathbf{u}_2 = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}$, $\mathbf{u}_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$ is an orthogonal set? **Sol:**

$\mathbf{u}_1 \cdot \mathbf{u}_2 = 1\cdot 1 + -1\cdot 1 + 0\cdot 0 = 0$, $\quad \mathbf{u}_1 \cdot \mathbf{u}_3 = 1\cdot 0 + -1\cdot 0 + 0\cdot 1 = 0$, $\quad \mathbf{u}_2 \cdot \mathbf{u}_3 = 1\cdot 0 + 1\cdot 0 + 0\cdot 1 = 0$

Th Suppose that $S = \{\mathbf{u}_1, \ldots, \mathbf{u}_p\}$ is an orthogonal set of nonzero vectors in $\mathbf{R}^n$ and $W = \text{Span}\{\mathbf{u}_1, \ldots, \mathbf{u}_p\}$. Then S is a linearly independent set and a basis for W.

Pf Suppose that

$$c_1\mathbf{u}_1 + c_2\mathbf{u}_2 + \cdots + c_p\mathbf{u}_p = \mathbf{0},$$

then

$$\big(c_1\mathbf{u}_1 + c_2\mathbf{u}_2 + \cdots + c_p\mathbf{u}_p\big) \cdot \mathbf{u}_1 = 0,$$
$$c_1\mathbf{u}_1 \cdot \mathbf{u}_1 + c_2\mathbf{u}_2 \cdot \mathbf{u}_1 + \cdots + c_p\mathbf{u}_p \cdot \mathbf{u}_1 = 0,$$
$$c_1\mathbf{u}_1 \cdot \mathbf{u}_1 = 0,$$

$c_1 = 0$ since $\mathbf{u}_1 \cdot \mathbf{u}_1 > 0$. Similarly $c_2 = \cdots = c_p = 0$, so S is a linearly independent set.
A set of vectors $\{\mathbf{u}_1, \ldots, \mathbf{u}_p\}$ is called an orthogonal basis if it is orthogonal and a basis.

Th If $S = \{\mathbf{u}_1, \ldots, \mathbf{u}_p\}$ is an orthogonal basis for a subspace W and $\mathbf{y} \in W$, then

$$\mathbf{x} = c_1\mathbf{u}_1 + c_2\mathbf{u}_2 + \cdots + c_p\mathbf{u}_p, \qquad \text{where} \quad c_i = \frac{\mathbf{x} \cdot \mathbf{u}_i}{\mathbf{u}_i \cdot \mathbf{u}_i}$$

Pf

$$\mathbf{x} \cdot \mathbf{u}_1 = \big(c_1\mathbf{u}_1 + c_2\mathbf{u}_2 + \cdots + c_p\mathbf{u}_p\big) \cdot \mathbf{u}_1 = c_1\mathbf{u}_1 \cdot \mathbf{u}_1 + c_2\mathbf{u}_2 \cdot \mathbf{u}_1 + \cdots + c_p\mathbf{u}_p \cdot \mathbf{u}_1 = c_1\mathbf{u}_1 \cdot \mathbf{u}_1$$

Hence $c_1 = \dfrac{\mathbf{u}_1 \cdot \mathbf{x}}{\mathbf{u}_1 \cdot \mathbf{u}_1}$ and similarly $c_2 = \dfrac{\mathbf{u}_2 \cdot \mathbf{x}}{\mathbf{u}_2 \cdot \mathbf{u}_2}$, $\quad \cdots \quad c_p = \dfrac{\mathbf{u}_p \cdot \mathbf{x}}{\mathbf{u}_p \cdot \mathbf{u}_p}$

Ex Write $\mathbf{x} = \begin{bmatrix} 3 \\ 7 \\ 4 \end{bmatrix}$ as a linear combination of $\mathbf{u}_1 = \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}$, $\mathbf{u}_2 = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}$, $\mathbf{u}_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$.

Sol Since $\{\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3\}$ is and orthogonal basis it follows from the previous theorem that $\mathbf{x} = c_1\mathbf{u}_1 + c_2\mathbf{u}_2 + c_3\mathbf{u}_3$, where

$$c_1 = \frac{\mathbf{u}_1 \cdot \mathbf{x}}{\mathbf{u}_1 \cdot \mathbf{u}_1} = -2, \qquad c_2 = \frac{\mathbf{u}_2 \cdot \mathbf{x}}{\mathbf{u}_2 \cdot \mathbf{u}_2} = 5, \qquad c_3 = \frac{\mathbf{u}_3 \cdot \mathbf{x}}{\mathbf{u}_3 \cdot \mathbf{u}_3} = 4$$

Hence $\mathbf{x} = -2\mathbf{u}_1 + 5\mathbf{u}_2 + 4\mathbf{u}_3$.

A set of vectors $\{\mathbf{u}_1, \ldots, \mathbf{u}_p\}$ is called **orthonormal** if it is a orthogonal set of unit vectors i.e.

$$\mathbf{u}_i \cdot \mathbf{u}_j = \delta_{ij} = \begin{cases} 0, & \text{if } i \neq j \\ 1, & \text{if } i = j \end{cases}$$

If $\{\mathbf{v}_1, \ldots, \mathbf{v}_p\}$ is an orthognal set then we get an orthonormal set $\mathbf{u}_i = \mathbf{v}_i / \|\mathbf{v}||_i\|$.
An **orthonormal basis** $\{\mathbf{u}_1, \ldots, \mathbf{u}_p\}$ for a subspace W is a basis that is also orthonormal.

Th If $S = \{\mathbf{u}_1, \ldots, \mathbf{u}_p\}$ is an orthonormal basis for a subspace W and $\mathbf{x} \in W$, then

$$\mathbf{x} = c_1\mathbf{u}_1 + c_2\mathbf{u}_2 + \cdots + c_p\mathbf{u}_p, \qquad \text{where} \quad c_i = \mathbf{x} \cdot \mathbf{u}_i$$

Orthogonal Projection

We will now calculate the **orthogonal projection** of $\mathbf{x}$ onto $\mathbf{u}$.
It is a vector $\mathbf{x}^{\|} = \alpha \mathbf{u}$ in the direction of $\mathbf{u}$, such that $\mathbf{x} - \mathbf{x}^{\|}$ is orthogonal to $\mathbf{u}$:

$$(\mathbf{x} - \alpha\mathbf{u}) \cdot \mathbf{v} = 0 \quad \Leftrightarrow \quad \mathbf{x} \cdot \mathbf{u} - \alpha\, \mathbf{u} \cdot \mathbf{u} = 0 \quad \Leftrightarrow \quad \alpha = \frac{\mathbf{x} \cdot \mathbf{u}}{\mathbf{u} \cdot \mathbf{u}}$$

The **orthogonal projection of x onto u** is the vector $\mathbf{x}^{\|} = \dfrac{\mathbf{x} \cdot \mathbf{u}}{\mathbf{u} \cdot \mathbf{u}} \mathbf{u}$. We can write

$$\mathbf{x} = \mathbf{x}^{\|} + \mathbf{x}^{\perp}$$

where $\mathbf{x}^{\perp}$ is called the **component orthogonal to u**.

Ex Suppose that $\{\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3\}$ is an orthogonal basis for $\mathbf{R}^3$ and let $W = \text{Span}\{\mathbf{u}_1, \mathbf{u}_2\}$. Write $\mathbf{x} = \mathbf{x}^{\|} + \mathbf{x}^{\perp}$, where $\mathbf{x} \in W$ and $\mathbf{x}^{\perp} \in W^{\perp}$, i.e. $\mathbf{x}^{\perp}$ is perpendicular to every vector in W, i.e $\mathbf{x}^{\perp} \cdot \mathbf{u}_1 = 0 = \mathbf{x}^{\perp} \cdot \mathbf{u}_2$.
Sol By a previous theorem we can write $\mathbf{x} = \dfrac{\mathbf{x} \cdot \mathbf{u}_1}{\mathbf{u}_1 \cdot \mathbf{u}_1}\mathbf{u}_1 + \dfrac{\mathbf{x} \cdot \mathbf{u}_2}{\mathbf{u}_2 \cdot \mathbf{u}_2}\mathbf{u}_2 + \dfrac{\mathbf{x} \cdot \mathbf{u}_3}{\mathbf{u}_3 \cdot \mathbf{u}_3}\mathbf{u}_3$ Let $\mathbf{x}^{\|} = \dfrac{\mathbf{x} \cdot \mathbf{u}_1}{\mathbf{u}_1 \cdot \mathbf{u}_1}\mathbf{u}_1 + \dfrac{\mathbf{x} \cdot \mathbf{u}_2}{\mathbf{u}_2 \cdot \mathbf{u}_2}\mathbf{u}_2$. Then $\mathbf{x}^{\|} \in W$ and $\mathbf{x}^{\perp} = \dfrac{\mathbf{x} \cdot \mathbf{u}_3}{\mathbf{u}_3 \cdot \mathbf{u}_3}\mathbf{u}_3$ is orthogonal to W, since $\mathbf{x}^{\perp} \cdot \mathbf{u}_1 = \mathbf{x}^{\perp} \cdot \mathbf{u}_2 = 0$.
$\mathbf{x}^{\|}$ is called the **projection of x onto** W

Ex Let $W = \text{Span}\{\mathbf{u}_1, \mathbf{u}_2\}$, where $\mathbf{u}_1 = \begin{bmatrix} 3 \\ 0 \\ 1 \end{bmatrix}$, $\mathbf{u}_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$, and let $\mathbf{x} = \begin{bmatrix} 0 \\ 3 \\ 10 \end{bmatrix}$.
Write $\mathbf{x} = \mathbf{x}^{\|} + \mathbf{x}^{\perp}$, where $\mathbf{x}^{\|} \in W$ and $\mathbf{x}^{\perp} \in W^{\perp}$.
Sol Since $\{\mathbf{u}_1, \mathbf{u}_2\}$ is an orthogonal basis it follows that

$$\mathbf{x}^{\|} = \frac{\mathbf{x} \cdot \mathbf{u}_1}{\mathbf{u}_1 \cdot \mathbf{u}_1} + \frac{\mathbf{x} \cdot \mathbf{u}_2}{\mathbf{u}_2 \cdot \mathbf{u}_2} = \frac{10}{10}\begin{bmatrix} 3 \\ 0 \\ 1 \end{bmatrix} + \frac{3}{1}\begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} -3 \\ 0 \\ 9 \end{bmatrix}.$$

$$\mathbf{x}^{\perp} = \mathbf{x} - \mathbf{x}^{\|} = \begin{bmatrix} 0 \\ 3 \\ 10 \end{bmatrix} - \begin{bmatrix} -3 \\ 0 \\ 9 \end{bmatrix} = \begin{bmatrix} 3 \\ 3 \\ 1 \end{bmatrix}$$

The orthogonal decomposition theorem Let W be a subspace of $\mathbf{R}^n$ and suppose that $\{\mathbf{u}_1, \dots, \mathbf{u}_p\}$ is an orthogonal basis for W. Any $\mathbf{x} \in \mathbf{R}^n$ can be written uniquely as

$$\mathbf{x} = \mathbf{x}^{\|} + \mathbf{x}^{\perp},$$

where

$$\mathbf{x}^{\|} = \frac{\mathbf{x} \cdot \mathbf{u}_1}{\mathbf{u}_1 \cdot \mathbf{u}_1}\mathbf{u}_1 + \cdots + \frac{\mathbf{x} \cdot \mathbf{u}_p}{\mathbf{u}_p \cdot \mathbf{u}_p}\mathbf{u}_p$$

and $\mathbf{x}^{\perp} = \mathbf{x} - \mathbf{x}^{\|} \in W^{\perp}$, the orthogonal complement $W^{\perp} = \{\mathbf{z} \in \mathbf{R}^n;\ \mathbf{z} \cdot \mathbf{u}_1 = 0, \dots, \mathbf{z} \cdot \mathbf{u}_p = 0\}$.
$\mathbf{x}^{\|} = \text{proj}_W \mathbf{x}$ is called the **orthogonal projection of x onto** W.

Th Suppose that $\{\mathbf{u}_1, \dots, \mathbf{u}_p\}$ is an orthonormal basis for W, i.e. $\mathbf{u}_i \cdot \mathbf{u}_j = \delta_{ij}$. Then

$$\text{proj}_W \mathbf{x} = (\mathbf{x} \cdot \mathbf{u}_1)\mathbf{u}_1 + \cdots + (\mathbf{x} \cdot \mathbf{u}_p)\mathbf{u}_p$$

Summary

The **dot product** between two vectors $\mathbf{x}=\begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$ and $\mathbf{y}=\begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix}$ is $\mathbf{x}\cdot\mathbf{y} = x_1y_1 + \cdots + x_ny_n$.
The **length** of the vector $\mathbf{x}$ is $\|\mathbf{x}\| = \sqrt{\mathbf{x}\cdot\mathbf{x}} = \sqrt{x_1^2 + \cdots + x_n^2}$.
$\mathbf{x}$ and $\mathbf{y}$ are said to be **orthogonal** if $\mathbf{x}\cdot\mathbf{y} = 0$.

A set of vectors $\{\mathbf{u}_1, \dots, \mathbf{u}_p\}$ in called an **orthogonal set** if $\mathbf{u}_i\cdot\mathbf{u}_j = 0$ when $i\neq j$.

Th An orthogonal set of nonzero vectors is linearly independent.
Pf We need to show that $c_1\mathbf{u}_1 + \cdots + c_p\mathbf{u}_p = \mathbf{0}$ implies that $c_i = 0$. Dot product with $\mathbf{u}_i$ gives
$0 = \big(c_1\mathbf{u}_1 + \cdots + c_i\mathbf{u}_i + \cdots + c_p\mathbf{u}_p\big)\cdot\mathbf{u}_i = c_1\mathbf{u}_1\cdot\mathbf{u}_i + \cdots + c_i\mathbf{u}_i\cdot\mathbf{u}_i + \cdots + c_p\mathbf{u}_p\cdot\mathbf{u}_i = c_i\mathbf{u}_i\cdot\mathbf{u}_i = c_i$.

A set of vectors $\{\mathbf{u}_1, \dots, \mathbf{u}_p\}$ is called **orthonormal** if it is an orthogonal set of unit vectors i.e.

$$\mathbf{u}_i\cdot\mathbf{u}_j = \delta_{ij} = \begin{cases} 0, & \text{if } i\neq j \\ 1, & \text{if } i = j \end{cases}$$

If $\{\mathbf{v}_1, \dots, \mathbf{v}_p\}$ is an orthognal set then we get an orthonormal set by setting $\mathbf{u}_i = \mathbf{v}_i/\|\mathbf{v}_i\|$.
An **orthonormal basis** $\{\mathbf{u}_1, \dots, \mathbf{u}_p\}$ for a subspace W is a basis that is also orthonormal.

Th If $\{\mathbf{u}_1, \dots, \mathbf{u}_p\}$ is an orthonormal basis for a subspace W and $\mathbf{x}\in W$, then

$$\mathbf{x} = c_1\mathbf{u}_1 + \cdots + c_p\mathbf{u}_p, \qquad \text{where} \quad c_i = \mathbf{x}\cdot\mathbf{u}_i$$

Pf
$\mathbf{x}\cdot\mathbf{u}_i = \big(c_1\mathbf{u}_1 + \cdots + c_i\mathbf{u}_i + \cdots + c_p\mathbf{u}_p\big)\cdot\mathbf{u}_i = c_1\mathbf{u}_1\cdot\mathbf{u}_i + \cdots + c_i\mathbf{u}_i\cdot\mathbf{u}_i + \cdots + c_p\mathbf{u}_p\cdot\mathbf{u}_i = c_i\mathbf{u}_i\cdot\mathbf{u}_i = c_i$.

If W is a subspace, the **orthogonal complement** $W^\perp$ of W is the set of all vectors orthogonal to every vector in W i.e. $W^\perp = \{\mathbf{x}\in\mathbf{R}^n;\ \mathbf{x}\cdot\mathbf{w} = 0, \text{ for all } \mathbf{w}\in W\}$. $W^\perp$ is a subspace.
If $W = \text{Span}(\mathbf{u}_1, \dots, \mathbf{u}_p)$ then $W^\perp = \{\mathbf{x}\in\mathbf{R}^n;\ \mathbf{x}\cdot\mathbf{u}_1 = 0, \dots, \mathbf{x}\cdot\mathbf{u}_p = 0\}$.

The orthogonal decomposition theorem Let W be a subspace of $\mathbf{R}^n$ and suppose that $\{\mathbf{u}_1, \dots, \mathbf{u}_p\}$ is an orthonormal basis for W. Any $\mathbf{x}\in\mathbf{R}^n$ can be written uniquely as

$$\mathbf{x} = \mathbf{x}^\| + \mathbf{x}^\perp, \qquad \text{with} \quad \mathbf{x}^\| \in W, \quad \mathbf{x}^\perp \in W^\perp,$$

where

$$\mathbf{x}^\| = \text{proj}_W\mathbf{x} = c_1\mathbf{u}_1 + \cdots + c_p\mathbf{u}_p, \quad \text{where} \quad c_i = \mathbf{x}\cdot\mathbf{u}_i$$

$\mathbf{x}^\|$ is called the **orthogonal projection** of $\mathbf{x}$ onto W denoted by $\text{proj}_W\,\mathbf{x}$.
Pf That $\mathbf{x}^\|\in W$ is clear and that $\mathbf{x}^\perp\in W^\perp$ follows from that it is orthogonal to all the $\mathbf{u}_i$:

$$\mathbf{x}^\perp\cdot\mathbf{u}_i = (\mathbf{x} - \mathbf{x}^\|)\cdot\mathbf{u}_i = \mathbf{x}\cdot\mathbf{u}_i - \big(c_1\mathbf{u}_1 + \cdots + c_i\mathbf{u}_i + \cdots + c_p\mathbf{u}_p\big)\cdot\mathbf{u}_i = \mathbf{x}\cdot\mathbf{u}_i - c_i\mathbf{u}_i\cdot\mathbf{u}_i = 0.$$

The **Pythagorean law:** $\|\mathbf{x}+\mathbf{y}\|^2 = \|\mathbf{x}\|^2 + \|\mathbf{y}\|^2$ if and only if $\mathbf{x}\cdot\mathbf{y} = 0$.
Cauchy-Schwarz inequality: $|\mathbf{x}\cdot\mathbf{y}| \le \|\mathbf{x}\|\,\|\mathbf{y}\|$.