Lecture 20: Orthogonality

Definition

Let $\mathbf{x} = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$ and $\mathbf{y} = \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix}$ be vectors in $\mathbb{R}^n$.

1.The **dot product** $\mathbf{x}$ and $\mathbf{y}$ is the real number $x_1y_1 + x_2y_2 + \cdots + x_ny_n$. In terms of matrices, we can also write the dot product as:

$$\begin{bmatrix} x_1 & x_2 & \cdots & x_n \end{bmatrix} \cdot \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} = \mathbf{x}^t \cdot \mathbf{y}.$$

Since we write vectors in $\mathbb{R}^n$ as column vectors, and, strictly speaking, we cannot form a column product $\mathbf{x} \cdot \mathbf{y}$, we will write $\mathbf{x} * \mathbf{y}$ for the dot product of $\mathbf{x}$ and $\mathbf{y}$.

2. The length of $\mathbf{x}$ is the non-negative real number

$$||\mathbf{x}|| = \sqrt{x_1^2 + \cdots + x_n^2} = \sqrt{\mathbf{x} * \mathbf{x}}.$$

Properties of the Dot Product and Length

Let $\mathbf{x}$, $\mathbf{y}$ and $\mathbf{z}$ belong to $\mathbb{R}^n$.

(i) $\mathbf{x} * \mathbf{y} = \mathbf{y} * \mathbf{x}$.

(ii) $\mathbf{x} * (\mathbf{y} + \mathbf{z}) = \mathbf{x} * \mathbf{y} + \mathbf{x} * \mathbf{z}$.

(iii) $(\lambda \mathbf{x}) * \mathbf{y} = x * (\lambda \mathbf{y}) = \lambda(\mathbf{x} * \mathbf{y})$, for all $\lambda \in \mathbb{R}$.

(iv) $||\mathbf{x}|| \geq 0$ and $||\mathbf{x}|| = 0$ if and only if $\mathbf{x} = \mathbf{0}$.

(v) $||\lambda \mathbf{x}|| = |\lambda| \cdot ||\mathbf{x}||$.

(vi) $||\mathbf{x}||^2 = \mathbf{x} * \mathbf{x}$.

Illustrations of the Properties of Dot Product and Length. Suppose

$$\mathbf{x} = \begin{bmatrix} 3 \\ -2 \\ 1 \end{bmatrix}, \mathbf{y} = \begin{bmatrix} 7 \\ -3 \\ 2 \end{bmatrix}, \mathbf{z} = \begin{bmatrix} 3 \\ 3 \\ 2 \end{bmatrix}, \text{ and } \lambda = -5.$$

1. $\mathbf{x} * \mathbf{y} = 3 \cdot 7 + (-2) \cdot (-3) + 1 \cdot 2 = 29.$

   $\mathbf{y} * \mathbf{x} = 7 \cdot 3 + (-3) \cdot (-2) + 2 \cdot 1 = 29.$

2. $\mathbf{x} * (\mathbf{y} + \mathbf{z}) = \mathbf{x} * \begin{bmatrix} 10 \\ 0 \\ 4 \end{bmatrix} = 3 \cdot 10 + (-2) \cdot 0 + 1 \cdot 4 = 34.$

   $\mathbf{x} * \mathbf{y} + \mathbf{x} * \mathbf{z} = (21 + 6 + 2) + (9 - 6 + 2) = 30 + 4 = 34.$

3. $(-5\mathbf{x}) * y = (-15) \cdot 7 + (10) \cdot (-3) + -5 \cdot 2 = -105 - 30 - 10 = -145.$

   $\mathbf{x} * (-5\mathbf{y}) = 3 \cdot (-35) + (-2) \cdot (15) + 1 \cdot (-10) = -105 - 30 - 10 = -145.$

   $-5 \cdot (\mathbf{x} * \mathbf{y}) = -5 \cdot (3 \cdot 7 + (-2) \cdot (-3) + 1 \cdot 2) = -5 \cdot 29 = -145.$

Illustrations of the Properties of Dot Product and Length cont. Suppose

$$\mathbf{x} = \begin{bmatrix} 3 \\ -2 \\ 1 \end{bmatrix}, \mathbf{y} = \begin{bmatrix} 7 \\ -3 \\ 2 \end{bmatrix}, \mathbf{z} = \begin{bmatrix} 3 \\ 3 \\ 2 \end{bmatrix}, \text{ and } \lambda = -5.$$

4. $||\mathbf{x}|| = \sqrt{3^2 + (-2)^2 + 1^2} = \sqrt{9 + 4 + 1} = \sqrt{14} \geq 0.$

5. $||(-5)\mathbf{x}|| = || \begin{bmatrix} -15 \\ 10 \\ -5 \end{bmatrix} || = \sqrt{(-15)^2 + 10^2 + (-5)^2} = \sqrt{350}.$

   $|-5| \cdot ||\mathbf{x}|| = 5 \cdot \sqrt{3^2 + (-2)^2 + 1^2} = 5 \cdot \sqrt{14} = \sqrt{25 \cdot 14} = \sqrt{350}.$

6. $\mathbf{x} * \mathbf{x} = 3^2 + (-2)^2 + 1^2 = 14.$

   $||\mathbf{x}||^2 = (\sqrt{3^2 + (-2)^2 + 1^2})^2 = 3^3 + (-1)^2 + 1^2 = 14.$

Theorem

Let $\mathbf{x}, \mathbf{y}$ be non-zero vectors in $\mathbb{R}^n$. Then $\mathbf{x} \cdot \mathbf{y} = ||\mathbf{x}|| \cdot ||\mathbf{y}|| \cos(\theta)$, where θ is the angle between $\mathbf{x}$ and $\mathbf{y}$.

Remark. 1. The theorem above can be visualized in $\mathbb{R}^2$ or $\mathbb{R}^3$. For higher dimensions, we can just define θ to be $\cos^{-1}(\frac{\mathbf{x}*\mathbf{y}}{||\mathbf{x}||\cdot||\mathbf{y}||})$.

2. Since $\cos(\frac{\pi}{2}) = 0$, it follows that $\mathbf{x}$ and $\mathbf{y}$ are **orthogonal** exactly when $\mathbf{x} * \mathbf{y} = 0$.

3. Vectors $\mathbf{x}$ and $\mathbf{y}$ are **orthonomal** if they are orthogonal and have length one. For example, $\mathbf{x} = \begin{bmatrix} \frac{\sqrt{2}}{2} \\ \frac{\sqrt{2}}{2} \end{bmatrix}$ and $\mathbf{y} = \begin{bmatrix} -\frac{\sqrt{2}}{2} \\ \frac{\sqrt{2}}{2} \end{bmatrix}$, are orthonormal:

$$||\mathbf{x}|| = \sqrt{(\frac{\sqrt{2}}{2})^2 + (\frac{\sqrt{2}}{2})^2} = \sqrt{\frac{2}{4} + \frac{2}{4}} = 1,$$

and similarly, $||\mathbf{y}|| = 1$. In addition,

$$\mathbf{x} * \mathbf{y} = \frac{\sqrt{2}}{2} \cdot (-\frac{\sqrt{2}}{2}) + \frac{\sqrt{2}}{2} \cdot \frac{\sqrt{2}}{2} = -\frac{2}{4} + \frac{2}{4} = 0.$$

Remark continued.

4. A set of vectors $u_1, \ldots, u_r$ is an **orthonormal system** if each vector u_i has length one and any two vectors u_i and u_j are orthogonal.

In other words: $||u_i|| = 1$, for all i and $u_i * u_j = 0$, for all $i \neq j$.

Equivalently: $u_i * u_i = 1$ for all i and $u_i * u_j = 0$, for all $i \neq j$.

5. The standard basis $\mathbf{e}_1, \mathbf{e}_2, \cdots, \mathbf{e}_n$ for $\mathbb{R}^n$ is an orthonormal system, in fact, an orthonormal basis.

For example: $\mathbf{e}_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \mathbf{e}_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \mathbf{e}_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$.

These vectors clearly have length one and $\mathrm{e_i} * \mathrm{e_j} = 0$.

Comment

Our goal is the following. Suppose W is a subspace of $\mathbb{R}^n$ with basis $w_1, \dots w_r$.

We will find a new basis $u_1, \dots, u_r$ for W forming an orthonormal system.

We call such a basis an orthonormal basis.

The strategy is as follows: We first replace $w_1, \dots, w_r$ with a new basis $w'_1, \dots, w'_r$ having the property that $w'_i * w'_j = 0$, i.e.,the new basis is an orthogonal basis.

Then we replace each w'_i by a unit vector u_i pointing in the same direction as w'_i.

Basic Principle. Let $\mathbf{x}$ be a vector in $\mathbb{R}^n$. Then $\mathbf{u} = \frac{1}{||\mathbf{x}||} \cdot \mathbf{x}$ has length one.

To see this, by Property (v):

$$||\frac{1}{||\mathbf{x}||} \cdot \mathbf{x}|| = |\frac{1}{||\mathbf{x}||}| \cdot ||\mathbf{x}|| = \frac{1}{||\mathbf{x}||} \cdot ||\mathbf{x}|| = 1.$$

For example, if $\mathbf{x} = \begin{bmatrix} 3 \\ 2 \end{bmatrix}$, then $||\mathbf{x}|| = \sqrt{9+4} = \sqrt{13}$.

$\frac{1}{||\mathbf{x}||} \cdot \mathbf{x} = \frac{1}{\sqrt{13}} \cdot \begin{bmatrix} 3 \\ 2 \end{bmatrix} = \begin{bmatrix} \frac{3}{\sqrt{13}} \\ \frac{2}{\sqrt{13}} \end{bmatrix}$.

$|| \begin{bmatrix} \frac{3}{\sqrt{13}} \\ \frac{2}{\sqrt{13}} \end{bmatrix} || = \sqrt{(\frac{3}{\sqrt{13}})^2 + (\frac{2}{\sqrt{13}})^2} = \sqrt{\frac{9}{13} + \frac{4}{13}} = \sqrt{\frac{13}{13}} = 1.$

How to orthogonalize.

Let us start with a set of two vectors w_1, w_2 that form a basis for the subspace W. We want to replace them by a new basis w_1', w_2' for W so that $w_1' * w_2' = 0$.

Step 1. Take $w_1' = w_1$.

Step 2. We seek a vector of the form $w_2' = w_2 - \lambda w_1'$ such that $w_1' * w_2' = 0$.

We solve the equation $w_1' * (w_2 - \lambda w_1') = 0$, for λ.

$$0 = w_1' * (w_2 - \lambda w_1') = w_1' * w_2 - w_1' * (\lambda w_1') = w_1' * w_2 - \lambda(w_1' * w_1'),$$

Therefore $\lambda(w_1' * w_1') = w_1' * w_2$, and thus $\lambda = \frac{w_1' * w_2}{w_1' * w_1'} = \frac{w_1 * w_2}{w_1 * w_1}$.

Conclusion. w_1, w_2' are orthogonal for $w_2' = w_2 - \frac{w_1 * w_2}{w_1 * w_1} \cdot w_1$.

Example

Convert the pair $w_1 = \begin{bmatrix} 1 \\ -1 \\ 2 \end{bmatrix}$, $w_2 = \begin{bmatrix} 4 \\ -4 \\ 5 \end{bmatrix}$ into an orthogonal pair.

Solution: Take $w_1' = w_1$. Now, $w_1 * w_1 = 1^2 + (-1)^2 + 2^2 = 6$ and $w_1 * w_2 = 4 + 4 + 10 = 18$. We take $\lambda = \frac{18}{6} = 3$.

Therefore:

$$w_2' = w_2 - 3w_1 = \begin{bmatrix} 4 \\ -4 \\ 5 \end{bmatrix} - 3 \cdot \begin{bmatrix} 1 \\ -1 \\ 2 \end{bmatrix} = \begin{bmatrix} 1 \\ -1 \\ -1 \end{bmatrix}.$$

Check: $w_1' * w_2' = w_1^t \cdot w_2' = \begin{bmatrix} 1 & -1 & 2 \end{bmatrix} \cdot \begin{bmatrix} 1 \\ -1 \\ -1 \end{bmatrix} = 1 + 1 - 2 = 0.$

QUESTION. If w_1, w_2 form a basis for the subspace W, do the new orthogonal vectors w_1', w_2' also form a basis for W? The new vectors are w_1 and $w_2 - \lambda w_1$.

Linear Independence. Suppose $\alpha w_1 + \beta(w_2 - \lambda w_1) = \mathbf{0}$.

Then, $(\alpha - \beta\lambda)w_1 + \beta w_2 = \mathbf{0}$. Since w_1, w_2 are linearly independent, $\alpha - \beta\lambda = 0$ and $\beta = 0$.

Thus: $\beta = 0 = \alpha$, so the new vectors w_1', w_2' are linearly independent.

Spanning. $w_1, w_2 - \lambda w_1$ are in $\text{span}\{w_1, w_2\}$, therefore

$$\text{span}\{w_1', w_2'\} \subseteq \text{span}\{w_1, w_2\}.$$

On the other hand, $w_2 = (w_2 - \lambda w_1) + \lambda \cdot w_1$, so w_1, w_2 are in $\text{span}\{w_1', w_2'\}$, therefore,

$$W = \text{span}\{w_1, w_2\} = \text{span}\{w_1', w_2'\},$$

so the new orthogonal vectors w_1', w_2' are a basis for W.

Theorem

Theorem (First Case of Gram-Schmidt Process). Let w_1, w_2 be a basis for the subspace $W \subseteq \mathbb{R}^n$. Then for

$$w_1' := w_1 \text{ and } w_2' := w_2 - \frac{w_1 * w_2}{w_1 * w_1} \cdot w_1,$$

w_1', w_2' is an orthogonal basis for W.

Class Example. Suppose $w_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$ and $w_2 = \begin{bmatrix} -2 \\ 5 \\ 3 \end{bmatrix}$ is a basis for the subspace W of $\mathbb{R}^3$. Find an orthogonal basis for W.

Solution. We need to calculate $w_1 * w_1$ and $w_1 * w_2$.

$w_1 * w_1 = 1 + 1 + 1 = 3$. $w_1 * w_2 = -2 + 5 + 3 = 6$. Thus, $\frac{w_1 * w_2}{w_1 * w_1} = 2$.

Therefore $w_2' = w_2 - 2w_1 = \begin{bmatrix} -2 \\ 5 \\ 3 \end{bmatrix} - 2 \cdot \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} -4 \\ 3 \\ 1 \end{bmatrix}$.

$w_1' = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, w_2' = \begin{bmatrix} -4 \\ 3 \\ 1 \end{bmatrix}$ is an orthogonal basis for W.

CHECK. $w_1' * w_2 = -4 + 3 + 1 = 0$.

Comment

Once we obtain an orthogonal basis for W, we can normalize these vectors to obtain an orthonormal basis.

Previous Example Revisited. We started with the basis $w_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$ and $w_2 = \begin{bmatrix} -2 \\ 5 \\ 3 \end{bmatrix}$, and derived an orthogonal basis $w_1' = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$ and $w_2' = \begin{bmatrix} -4 \\ 3 \\ 1 \end{bmatrix}$.

To get an orthonormal basis, we take $u_1 = \frac{1}{||w_1'||} \cdot w_1'$ and $u_2 = \frac{1}{||w_2'||} \cdot w_2'$.

$||w_1'|| = \sqrt{3}$ and $||w_2'|| = \sqrt{26}$, thus

$u_1 = \begin{bmatrix} \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{3}} \end{bmatrix}$ and $u_2 = \begin{bmatrix} \frac{-4}{\sqrt{26}} \\ \frac{3}{\sqrt{26}} \\ \frac{1}{\sqrt{26}} \end{bmatrix}$, is an orthonormal basis for W.

Value in using an orthonormal basis

Suppose u_1, u_2 is an orthonormal basis for the subspace $W \subseteq \mathbb{R}^n$. Let $w \in W$. Then

$$w = (w * u_1)u_1 + (w * u_2)u_2.$$

WHY: If we write $w = \alpha u_1 + \beta u_2$, then:

$$w * u_1 = (\alpha u_1 + \beta u_2) * u_1 = (\alpha u_1) * u_1 + (\beta u_2) * u_1 =$$

$$= \alpha(u_1 * u_1) + \beta(u_2 * u_1) = \alpha \cdot 1 + \beta \cdot 0 = \alpha.$$

and

$$w * u_2 = (\alpha u_1 + \beta u_2) * u_2 = (\alpha u_1) * u_2 + (\beta u_2) * u_2 =$$

$$= \alpha(u_1 * u_2) + \beta(u_2 * u_2) = \alpha \cdot 0 + \beta \cdot 1 = \beta.$$

Example continued

Previous Example Continued Further. Write

$w = w_1 + w_2 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} + \begin{bmatrix} -2 \\ 5 \\ 3 \end{bmatrix} = \begin{bmatrix} -1 \\ 6 \\ 4 \end{bmatrix}$ as a linear combination of u_1, u_2.

$w * u_1 = \frac{-1}{\sqrt{3}} + \frac{6}{\sqrt{3}} + \frac{4}{\sqrt{3}} = \frac{9}{\sqrt{3}}$.

$w * u_2 = \frac{4}{\sqrt{26}} + \frac{18}{\sqrt{26}} + \frac{4}{\sqrt{26}} = \frac{26}{\sqrt{26}}$

Thus: $w = \frac{9}{\sqrt{3}} \cdot u_1 + \frac{26}{\sqrt{26}} \cdot u_2$.

Notice: We can write w as a linear combination of u_1, u_2 without resorting to Gaussian elimination!

Class Example

Find an orthonormal basis for the subspace of $\mathbb{R}^3$ spanned by $w_1 = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}$ and $w_2 = \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}$ and then use the dot product to write $v = \begin{bmatrix} 4 \\ -1 \\ 3 \end{bmatrix}$ as a linear combination of those vectors.

Solution: $w_1' = w_1$ and $w_2' = w_2 - \lambda \cdot w_1$, where $\lambda = \frac{w_1 * w_2}{w_1 * w_1} = \frac{1}{2}$.

$$w_2' = \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix} - \frac{1}{2} \cdot \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} -\frac{1}{2} \\ 1 \\ \frac{1}{2} \end{bmatrix}.$$

$$u_1 = \frac{\sqrt{2}}{2} \cdot \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} \text{ and } u_2 = \frac{\sqrt{2}}{\sqrt{3}} \cdot \begin{bmatrix} -\frac{1}{2} \\ 1 \\ \frac{1}{2} \end{bmatrix}.$$

$v * u_1 = \frac{\sqrt{2}}{2}(4+0+3) = \frac{7\sqrt{2}}{2}$, $v * u_2 = \frac{\sqrt{2}}{\sqrt{3}}(-2-1+\frac{3}{2}) = -\frac{3\sqrt{2}}{2\sqrt{3}}$.
Therefore,

$$v = \frac{7\sqrt{2}}{2} \cdot u_1 - \frac{3\sqrt{2}}{2\sqrt{3}} \cdot u_2.$$