Lecture 21: Gram-Schmidt Orthogonalization

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

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

Definition

3. Vectors $\mathbf{x}$ and $\mathbf{y}$ are **orthogonal** exactly when $\mathbf{x} * \mathbf{y} = 0$.

4. 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,

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

Definition

5. 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 $i \neq j$.

6. 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 $\mathbf{e_i} * \mathbf{e_j} = \mathbf{0}$.

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

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

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

Class Example. Suppose $w_1 = \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix}$ and $w_2 = \begin{bmatrix} 0 \\ 4 \\ 6 \end{bmatrix}$ 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 $w_1 * w_2$.

$w_1 * w_1 = 1^2 + 0^2 + 1^2 = 2$. $w_1 * w_2 = 0 + 0 - 6 = -6$. Thus,

$$\frac{w_1 * w_2}{w_1 * w_1} = -3.$$

Therefore $w_2' = w_2 - (-3)w_1 = \begin{bmatrix} 0 \\ 4 \\ 6 \end{bmatrix} + 3 \cdot \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix} = \begin{bmatrix} 3 \\ 4 \\ 3 \end{bmatrix}$.

$w_1' = \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix}, w_2' = \begin{bmatrix} 3 \\ 4 \\ 3 \end{bmatrix}$ is an orthogonal basis for W.

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

Remark. Though there are infinitely many vectors in $\mathbb{R}^3$ orthogonal to $w_1 = w_1'$, this process picks a vector orthogonal to w_1 so that w_1 and the new vector form a basis for the given subspace.

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 $w_1 = \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix}$ and $w_2 = \begin{bmatrix} 0 \\ 4 \\ 6 \end{bmatrix}$, and derived an orthogonal basis $w_1' = \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix}$ and $w_2' = \begin{bmatrix} 3 \\ 4 \\ 3 \end{bmatrix}$.

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

$||w_1'|| = \sqrt{2}$ and $||w_2'|| = \sqrt{34}$, thus

$u_1 = \begin{bmatrix} \frac{1}{\sqrt{2}} \\ \frac{0}{\sqrt{2}} \\ \frac{-1}{\sqrt{2}} \end{bmatrix}$ and $u_2 = \begin{bmatrix} \frac{3}{\sqrt{34}} \\ \frac{4}{\sqrt{34}} \\ \frac{3}{\sqrt{34}} \end{bmatrix}$, 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

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

Easy Example 1. Consider $\mathbf{x} = \begin{bmatrix} -4 \\ 2 \end{bmatrix}$ and the orthonormal basis $\mathbf{e}_1$ and $\mathbf{e}_2$. Then $\mathbf{x} = 4\mathbf{e}_1 - 3\mathbf{e}_2$. In terms of the dot product:

$$\mathbf{x} * \mathbf{e}_1 = (4\mathbf{e}_1 - 3\mathbf{e}_2) * \mathbf{e}_1 = (4\mathbf{e}_1 * \mathbf{e}_1) - (3\mathbf{e}_2 * \mathbf{e}_1) = 4 - 0 = 4.$$

$$\mathbf{x} * \mathbf{e}_2 = (4\mathbf{e}_1 - 3\mathbf{e}_2) * \mathbf{e}_2 = (4\mathbf{e}_1 * \mathbf{e}_2) - (3\mathbf{e}_2 * \mathbf{e}_2) = 0 - 3 = -3.$$

Example 2. Consider the orthonormal basis $u_1 = \begin{bmatrix} -\frac{\sqrt{2}}{2} \\ \frac{\sqrt{2}}{2} \end{bmatrix}$ and $u_2 = \begin{bmatrix} \frac{\sqrt{2}}{2} \\ \frac{\sqrt{2}}{2} \end{bmatrix}$. Let's write $\mathbf{x} = \begin{bmatrix} 2 \\ 4 \end{bmatrix}$ in terms of this basis.

$\mathbf{x} * u_1 = 2 \cdot (-\frac{\sqrt{2}}{2}) + 4 \cdot \frac{\sqrt{2}}{2} = -\sqrt{2} + 2\sqrt{2} = \sqrt{2}$

$\mathbf{x} * u_2 = 2 \cdot (\frac{\sqrt{2}}{2}) + 4 \cdot \frac{\sqrt{2}}{2} = \sqrt{2} + 2\sqrt{2} = 3\sqrt{2}$

Therefore, $\mathbf{x} = \sqrt{2}u_1 + 3\sqrt{2}u_2$.

Check:

$$\sqrt{2}u_1 + 3\sqrt{2}u_2 = \sqrt{2} \cdot \begin{bmatrix} -\frac{\sqrt{2}}{2} \\ \frac{\sqrt{2}}{2} \end{bmatrix} + 3\sqrt{2} \cdot \begin{bmatrix} \frac{\sqrt{2}}{2} \\ \frac{\sqrt{2}}{2} \end{bmatrix} = \begin{bmatrix} -1 \\ 1 \end{bmatrix} + \begin{bmatrix} 3 \\ 3 \end{bmatrix} = \begin{bmatrix} 2 \\ 4 \end{bmatrix} = \mathbf{x}.$$

Class Example

Find an orthonormal basis for the subspace of $\mathbb{R}^3$ with basis $w_1 = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}$ and $w_2 = \begin{bmatrix} 0 \\ 2 \\ 2 \end{bmatrix}$ and then use the dot product to write $v = \begin{bmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \\ \frac{2}{\sqrt{2}} \end{bmatrix}$ as a linear combination of those vectors.

First orthogonalize: $w_1 * w_1 = 2$ and $w_1 * w_2 = 2$. Thus $w_2' = w_2 - w_1 = \begin{bmatrix} -1 \\ 2 \\ 1 \end{bmatrix}$. $||w_1|| = \sqrt{2}$ and $||w_2'|| = \sqrt{6}$.

Therefore, $u_1 = \begin{bmatrix} \frac{1}{\sqrt{2}} \\ 0 \\ \frac{1}{\sqrt{2}} \end{bmatrix}$ and $u_2 = \begin{bmatrix} -\frac{1}{\sqrt{6}} \\ \frac{2}{\sqrt{6}} \\ \frac{1}{\sqrt{6}} \end{bmatrix}$ is the required orthonomal basis.

Class Example continued

$v * u_1 = \frac{1}{\sqrt{2}} \cdot \frac{1}{\sqrt{2}} + 0 + \frac{2}{\sqrt{2}} \cdot \frac{1}{\sqrt{2}} = \frac{1}{2} + \frac{2}{2} = \frac{3}{2}.$

$v * u_2 = \frac{1}{\sqrt{2}} \cdot (-\frac{1}{\sqrt{6}}) + \frac{1}{\sqrt{2}} \cdot \frac{2}{\sqrt{6}} + \frac{2}{\sqrt{2}} \cdot \frac{1}{\sqrt{6}} = \frac{3}{\sqrt{12}}.$

Thus, $v = \frac{3}{2} u_1 + \frac{3}{\sqrt{12}} u_2.$

Theorem

Theorem (Gram-Schmidt). Let $w_1, w_2, \ldots, w_r$ be a basis for the subspace W. Then an orthogonal basis for W is $w_1', w_2', \ldots, w_r'$, where:

(i) $w_1' = w_1$.

(ii) $w_2' = w_2 - \frac{w_1' * w_2}{w_1' * w_1'} \cdot w_1'$

(iii) $w_3' = w_3 - \frac{w_1' * w_3}{w_1' * w_1'} \cdot w_1' - \frac{w_2' * w_3}{w_2' * w_2'} \cdot w_2'$

$\vdots$

(r) $w_r' = w_r - \frac{w_1' * w_r}{w_1' * w_1'} \cdot w_1' - \frac{w_2' * w_r}{w_2' * w_2'} \cdot w_2' - \cdots - \frac{w_{r-1}' \cdot w_r}{w_{r-1}' \cdot w_{r-1}'} \cdot w_{r-1}'$pause

Example

Find an orthogonal basis for the subspace of $\mathbb{R}^4$ having basis $w_1 = \begin{bmatrix} 1 \\ 0 \\ 1 \\ 0 \end{bmatrix}$, $w_2 = \begin{bmatrix} 2 \\ 0 \\ 0 \\ 2 \end{bmatrix}$, $w_3 = \begin{bmatrix} 0 \\ 0 \\ 4 \\ 4 \end{bmatrix}$.

Solution. $w_1' = w_1$.

$$w_2' = w_2 - \frac{w_1 * w_2}{w_1 * w_1} \cdot w_1 = \begin{bmatrix} 2 \\ 0 \\ 0 \\ 2 \end{bmatrix} - \frac{2}{2} \begin{bmatrix} 1 \\ 0 \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ -1 \\ 2 \end{bmatrix}.$$

Example continued

$w_3' = w_3 - \frac{w_1 * w_3}{w_1 * w_1} \cdot w_1 - \frac{w_2' * w_3}{w_2' * w_2'} \cdot w_2' =$

$$\begin{bmatrix}0\\0\\4\\4\end{bmatrix} - \frac{4}{2} \cdot \begin{bmatrix}1\\0\\1\\0\end{bmatrix} - \frac{4}{6}\begin{bmatrix}1\\0\\-1\\2\end{bmatrix} = \begin{bmatrix}0\\0\\4\\4\end{bmatrix} - \begin{bmatrix}2\\0\\2\\0\end{bmatrix} + \begin{bmatrix}-\frac{2}{3}\\0\\\frac{2}{3}\\-\frac{4}{3}\end{bmatrix} = \begin{bmatrix}-\frac{8}{3}\\0\\\frac{8}{3}\\\frac{8}{3}\end{bmatrix}.$$

Thus, $w_1 = \begin{bmatrix}1\\0\\1\\0\end{bmatrix}, w_2' = \begin{bmatrix}1\\0\\-1\\2\end{bmatrix}. w_3' = \begin{bmatrix}-\frac{8}{3}\\0\\\frac{8}{3}\\\frac{8}{3}\end{bmatrix}$, is an orthogonal basis for the given subspace.

Class Example

Find an orthogonal basis for the subspace W of $\mathbb{R}^4$ having basis

$$w_1 = \begin{bmatrix} 1 \\ 1 \\ -1 \\ -1 \end{bmatrix}, w_2 = \begin{bmatrix} 3 \\ 2 \\ 0 \\ 1 \end{bmatrix}, w_3 = \begin{bmatrix} 1 \\ 0 \\ 1 \\ 0 \end{bmatrix}.$$

Use the formulas:

(i) $w_1' = w_1$.

(ii) $w_2' = w_2 - \frac{w_1' * w_2}{w_1' * w_1'} \cdot w_1'$

(iii) $w_3' = w_3 - \frac{w_1' * w_3}{w_1' * w_1'} \cdot w_1' - \frac{w_2' * w_3}{w_2' * w_2'} \cdot w_2'$

Example continued

Solution. $w_1' = w_1$.

$$w_2' = w_2 - \frac{w_1 * w_2}{w_1 * w_1} \cdot w_1 = \begin{bmatrix} 3 \\ 2 \\ 0 \\ 1 \end{bmatrix} - \frac{4}{4} \cdot \begin{bmatrix} 1 \\ 1 \\ -1 \\ -1 \end{bmatrix} = \begin{bmatrix} 2 \\ 1 \\ 1 \\ 2 \end{bmatrix}.$$

$w_3' = w_3 - \frac{w_1 * w_3}{w_1 * w_1} \cdot w_1 - \frac{w_2' * w_3}{w_2' * w_2'} \cdot w_2' =$

$$\begin{bmatrix} 1 \\ 0 \\ 1 \\ 0 \end{bmatrix} - \frac{0}{4} \cdot \begin{bmatrix} 1 \\ 1 \\ -1 \\ -1 \end{bmatrix} - \frac{3}{10} \cdot \begin{bmatrix} 2 \\ 1 \\ 1 \\ 2 \end{bmatrix} = \frac{1}{10} \cdot \begin{bmatrix} 4 \\ -3 \\ 7 \\ -6 \end{bmatrix}.$$

Class Example continued

Thus, the subspace W has orthogonal basis,

$$w_1' = \begin{bmatrix} 1 \\ 1 \\ -1 \\ -1 \end{bmatrix}, w_2' = \begin{bmatrix} 2 \\ 1 \\ 1 \\ 2 \end{bmatrix}, w_3' = \frac{1}{10} \cdot \begin{bmatrix} 4 \\ -3 \\ 7 \\ -6 \end{bmatrix}.$$

Converting to an orthonormal basis yields:

$$u_1 = \frac{1}{2} \cdot \begin{bmatrix} 1 \\ 1 \\ -1 \\ -1 \end{bmatrix}, u_2 = \frac{1}{\sqrt{10}} \cdot \begin{bmatrix} 2 \\ 1 \\ 1 \\ 2 \end{bmatrix}, = \frac{1}{10\sqrt{110}} \cdot \begin{bmatrix} 4 \\ -3 \\ 7 \\ -6 \end{bmatrix}.$$

Value in using an orthonormal basis

Theorem. Let $u_1, u_2, \ldots, u_r$ be an orthonormal basis for $W \subseteq \mathbb{R}^n$. Then, for any vector $w \in W$,

$$w = (w * u_1) \cdot u_1 + (w * u_2) \cdot u_2 + \cdots + (w * u_r) \cdot u_r.$$

Example. In the example above, write $w = \begin{bmatrix} 5 \\ 3 \\ 0 \\ 0 \end{bmatrix}$ from W as a linear combination of u_1, u_2, u_3. Using the theorem:

$w * u_1 = \frac{1}{2} \cdot (5 + 3 + 0 + 0) = 4.$

$w * u_2 = \frac{1}{\sqrt{10}} \cdot (10 + 3 + 0 + 0) = \frac{13}{\sqrt{10}}.$

$w * u_3 = \frac{1}{10\sqrt{110}} \cdot (20 - 9 + 0 + 0) = \frac{11}{10\sqrt{110}}.$

Thus,

$$w = 4 \cdot u_1 + \frac{13}{\sqrt{10}} \cdot u_2 + \frac{11}{10\sqrt{110}} \cdot u_3.$$