Linear Algebra from Scratch: Orthogonality

Instructor Anthony

"It does not matter how slowly you go as long as you do not stop." - Confucius

Udemy Open Course

Table of Contents

udemy

Table of Contents

udemy

The Scalar Product in $\mathbb{R}^n$

Definition (The Scalar Product)

Suppose $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$, then the **Scalar Product** between $\mathbf{x}$ and $\mathbf{y}$ denoted $\mathbf{x}^T\mathbf{y} = \langle \mathbf{x}, \mathbf{y} \rangle$ is:

$$\mathbf{x}^T\mathbf{y} \stackrel{\triangle}{\equiv} \mathbf{x} \cdot \mathbf{y} = x_1 y_1 + x_2 y_2 + \cdots + x_j y_j + \cdots + x_n y_n$$

udemy

The Scalar Product in $\mathbb{R}^n$

Definition (The Scalar Product)

Suppose $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$, then the **Scalar Product** between $\mathbf{x}$ and $\mathbf{y}$ denoted $\mathbf{x}^T\mathbf{y} = \langle \mathbf{x}, \mathbf{y} \rangle$ is:

$$\mathbf{x}^T\mathbf{y} \overset{\triangle}{\equiv} \mathbf{x} \cdot \mathbf{y} = x_1 y_1 + x_2 y_2 + \cdots + x_j y_j + \cdots + x_n y_n$$

Remark: For any $\alpha, \beta \in \mathbb{R}$ and $\mathbf{x}, \mathbf{y}, \mathbf{z} \in \mathbb{R}^n$, one has:

$$\langle \alpha\mathbf{x} + \beta\mathbf{y}, \mathbf{z} \rangle = \alpha\langle \mathbf{x}, \mathbf{z} \rangle + \beta\langle \mathbf{y}, \mathbf{z} \rangle$$

Example: Scalar Product in $\mathbb{R}^3$

Consider the vectors: $\mathbf{x} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$ and $\mathbf{y} = \begin{bmatrix} 2 \\ 4 \\ 6 \end{bmatrix}$ Then, the Scalar Product between $\mathbf{x}$ and $\mathbf{y}$ is:

$$\mathbf{x}^T\mathbf{y} = (1, 2, 3) * \begin{bmatrix} 2 \\ 4 \\ 6 \end{bmatrix} = 1(2) + 2(4) + 3(6) = 28$$

udemy

Reference: Leon, 2014 [2]

Geometrical Interpretation: Scalar Product

Definition (Euclidean Length)

Consider a vector $\mathbf{x} \in \mathbb{R}^n$, then the **Euclidean Length** is defined as the magnitude of the vector:

$$||\mathbf{x}||_2 \triangleq \sqrt{\langle \mathbf{x}, \mathbf{x} \rangle} = \begin{cases} \sqrt{x_1^2 + x_2^2 + \cdots + x_j^2 + \cdots + x_n^2} \,\Big|\, \mathbf{x} \in \mathbb{R}^n \\ \sqrt{x_1^2 + x_2^2} \,\Big|\, \mathbf{x} \in \mathbb{R}^2 \\ \sqrt{x_1^2 + x_2^2 + x_3^2} \,\Big|\, \mathbf{x} \in \mathbb{R}^3 \end{cases}$$

In the special case where $n = 2, 3$ the magnitude is the length of the vector.

udemy

Geometrical Interpretation: Scalar Product

Definition (Euclidean Length)

Consider a vector $\mathbf{x} \in \mathbb{R}^n$, then the **Euclidean Length** is defined as the magnitude of the vector:

$$||\mathbf{x}||_2 \stackrel{\triangle}{\equiv} \sqrt{\langle \mathbf{x}, \mathbf{x} \rangle} = \begin{cases} \sqrt{x_1^2 + x_2^2 + \cdots + x_j^2 + \cdots + x_n^2} \,\Big|\, \mathbf{x} \in \mathbb{R}^n \\ \sqrt{x_1^2 + x_2^2} \,\Big|\, \mathbf{x} \in \mathbb{R}^2 \\ \sqrt{x_1^2 + x_2^2 + x_3^2} \,\Big|\, \mathbf{x} \in \mathbb{R}^3 \end{cases}$$

In the special case where $n = 2, 3$ the magnitude is the length of the vector.

Definition (Distance Between Vectors)

Suppose $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$, then the **Distance** between $\mathbf{x}$ and $\mathbf{y}$ is length of the vector obtained by joining the Terminal Point of each vector.
Mathematically, one has:

$$\delta(\mathbf{x}, \mathbf{y}) = ||\mathbf{x} - \mathbf{y}||_2 = \sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2 + \cdots + (x_j - y_j)^2 + \cdots + (x_n - y_n)^2}$$

Geometrical Interpretation: Scalar Product

Definition (Euclidean Length)

Consider a vector $\mathbf{x} \in \mathbb{R}^n$, then the **Euclidean Length** is defined as the magnitude of the vector:

$$||\mathbf{x}||_2 \stackrel{\triangle}{\equiv} \sqrt{\langle \mathbf{x}, \mathbf{x} \rangle} = \begin{cases} \sqrt{x_1^2 + x_2^2 + \cdots + x_j^2 + \cdots + x_n^2} \, \Big| \, \mathbf{x} \in \mathbb{R}^n \\ \sqrt{x_1^2 + x_2^2} \, \Big| \, \mathbf{x} \in \mathbb{R}^2 \\ \sqrt{x_1^2 + x_2^2 + x_3^2} \, \Big| \, \mathbf{x} \in \mathbb{R}^3 \end{cases}$$

In the special case where $n = 2, 3$ the magnitude is the length of the vector.

Definition (Distance Between Vectors)

Suppose $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$, then the **Distance** between $\mathbf{x}$ and $\mathbf{y}$ is length of the vector obtained by joining the Terminal Point of each vector.
Mathematically, one has:

$$\delta(\mathbf{x}, \mathbf{y}) = ||\mathbf{x} - \mathbf{y}||_2 = \sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2 + \cdots + (x_j - y_j)^2 + \cdots + (x_n - y_n)^2}$$

Generally, if $\langle \cdot \rangle$ is some inner product defined for $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$, then the distance between $\mathbf{x}, \mathbf{y}$ is: $\sqrt{\langle \mathbf{x} - \mathbf{y}, \mathbf{x} - \mathbf{y} \rangle}$.

Representation: Distance

Distance between x and y $\in \mathbb{R}^2$.

Suppose: $\mathbf{x} = \begin{bmatrix} -4 \\ 2 \end{bmatrix}$ and $\mathbf{y} = \begin{bmatrix} -5 \\ -2 \end{bmatrix}$,

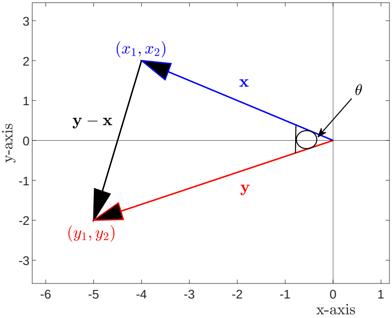
then: $\mathbf{y} - \mathbf{x} = \begin{bmatrix} -1 \\ -4 \end{bmatrix}$ and

$$\delta(\mathbf{x}, \mathbf{y}) = \sqrt{1 + 16} = \sqrt{17}$$

$$||\mathbf{x}||_2 = \sqrt{16 + 4} = 2\sqrt{5}$$

$$||\mathbf{y}||_2 = \sqrt{25 + 4} = \sqrt{29}$$

□



udemy

The Unit Vector

Definition (Unit Vector)

Suppose $\mathbf{x} \in \mathbb{R}^n$ where $\mathbf{x} \neq \mathbf{0}$, then the **Unit Vector** of $\mathbf{x}$ is $\mathbf{u_x}$, which is obtained by diving the vector by its Euclidean Length:

$$\mathbf{u_x} = \frac{\mathbf{x}}{||\mathbf{x}||_2}$$

udemy

The Unit Vector

Definition (Unit Vector)

Suppose $\mathbf{x} \in \mathbb{R}^n$ where $\mathbf{x} \neq \mathbf{0}$, then the **Unit Vector** of $\mathbf{x}$ is $\mathbf{u_x}$, which is obtained by diving the vector by its Euclidean Length:

$$\mathbf{u_x} = \frac{\mathbf{x}}{||\mathbf{x}||_2}$$

Remark: The **Unit Vector** of $\mathbf{x}$ can be interpreted as a vector with unitary magnitude in the direction of $\mathbf{x}$.

Scalar Product in $\mathbb{R}^n$

Theorem ($\langle \cdot \rangle$ for $\mathbb{R}^n$)

Suppose $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ *and* θ *is the Angle Between* $\mathbf{x}$ *and* $\mathbf{y}$*, then the* **Scalar Product** *between* $\mathbf{x}$ *and* $\mathbf{y}$ *is:*

$$\mathbf{x}^T\mathbf{y} = ||\mathbf{x}|| * ||\mathbf{y}|| * Cos(\theta)$$

Scalar Product in $\mathbb{R}^n$

Theorem ($\langle \cdot \rangle$ for $\mathbb{R}^n$)

Suppose $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ *and* θ *is the Angle Between* $\mathbf{x}$ *and* $\mathbf{y}$*, then the* **Scalar Product** *between* $\mathbf{x}$ *and* $\mathbf{y}$ *is:*

$$\mathbf{x}^T\mathbf{y} = ||\mathbf{x}|| * ||\mathbf{y}|| * Cos(\theta)$$

Lemma (Inner Product Between Unit Vectors in $\mathbb{R}^n$)

Observe that if $||\mathbf{x}|| = 1 = ||\mathbf{y}||$*, then* $\mathbf{x}$ *and* $\mathbf{y}$ *are Unit Vectors denoted* $\mathbf{u_x}$ *and* $\mathbf{u_y}$ *respectively, and one has:*

$$\mathbf{u_x}^T\mathbf{u_y} = Cos(\theta)$$

Scalar Product in $\mathbb{R}^n$

Theorem ($\langle \cdot \rangle$ for $\mathbb{R}^n$)

Suppose $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ *and* θ *is the Angle Between* $\mathbf{x}$ *and* $\mathbf{y}$, *then the* **Scalar Product** *between* $\mathbf{x}$ *and* $\mathbf{y}$ *is:*

$$\mathbf{x}^T \mathbf{y} = ||\mathbf{x}|| * ||\mathbf{y}|| * Cos(\theta)$$

Lemma (Inner Product Between Unit Vectors in $\mathbb{R}^n$)

Observe that if $||\mathbf{x}|| = 1 = ||\mathbf{y}||$, *then* $\mathbf{x}$ *and* $\mathbf{y}$ *are Unit Vectors denoted* $\mathbf{u_x}$ *and* $\mathbf{u_y}$ *respectively, and one has:*

$$\mathbf{u_x}^T \mathbf{u_y} = Cos(\theta)$$

Which states that the Inner Product between Unit Vectors can be expressed as the $Cos(\theta)$ *where* θ *is the angle between the unit vectors.*

Scalar Product in $\mathbb{R}^n$

Theorem ($\langle \cdot \rangle$ for $\mathbb{R}^n$)

Suppose $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ *and* θ *is the Angle Between* $\mathbf{x}$ *and* $\mathbf{y}$*, then the* **Scalar Product** *between* $\mathbf{x}$ *and* $\mathbf{y}$ *is:*

$$\mathbf{x}^T\mathbf{y} = ||\mathbf{x}|| * ||\mathbf{y}|| * Cos(\theta)$$

Lemma (Inner Product Between Unit Vectors in $\mathbb{R}^n$)

Observe that if $||\mathbf{x}|| = 1 = ||\mathbf{y}||$*, then* $\mathbf{x}$ *and* $\mathbf{y}$ *are Unit Vectors denoted* $\mathbf{u_x}$ *and* $\mathbf{u_y}$ *respectively, and one has:*

$$\mathbf{u_x}^T\mathbf{u_y} = Cos(\theta)$$

Which states that the Inner Product between Unit Vectors can be expressed as the $Cos(\theta)$ *where* θ *is the angle between the unit vectors.*

Remark*: These results can be generalized to any Inner Product* $\langle \cdot \rangle$ *that is well defined.*

Reference: Leon, 2014 [2]

The Angle Between Vectors

Definition (The Angle Between Vectors)

Suppose that $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$, then the **Angle** Between $\mathbf{x}$ and $\mathbf{y}$ denoted θ is:

$$\theta = Cos^{-1}\left(\frac{\mathbf{x}^T\mathbf{y}}{||\mathbf{x}||_2 * ||\mathbf{y}||_2}\right) = Cos^{-1}(\mathbf{u_x}^T\mathbf{u_y})$$

Example:

udemy

The Angle Between Vectors

Definition (The Angle Between Vectors)

Suppose that $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$, then the **Angle** Between $\mathbf{x}$ and $\mathbf{y}$ denoted θ is:

$$\theta = Cos^{-1}\left(\frac{\mathbf{x}^T\mathbf{y}}{||\mathbf{x}||_2 * ||\mathbf{y}||_2}\right) = Cos^{-1}(\mathbf{u_x}^T\mathbf{u_y})$$

Remark: The results can be generalized to any space with a well defined Inner Product by replacing $\mathbf{x}^T\mathbf{y}$ with $\langle \mathbf{x}, \mathbf{y} \rangle$

Example:

udemy

The Angle Between Vectors

Definition (The Angle Between Vectors)

Suppose that $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$, then the **Angle** Between $\mathbf{x}$ and $\mathbf{y}$ denoted θ is:

$$\theta = Cos^{-1}\left(\frac{\mathbf{x}^T\mathbf{y}}{||\mathbf{x}||_2 * ||\mathbf{y}||_2}\right) = Cos^{-1}(\mathbf{u_x}^T\mathbf{u_y})$$

Remark: The results can be generalized to any space with a well defined Inner Product by replacing $\mathbf{x}^T\mathbf{y}$ with $\langle \mathbf{x}, \mathbf{y} \rangle$

Example:
Suppose that $\mathbf{x} = (3, 4)^T$ and $\mathbf{y} = (-1, 1)$, then:

udemy

The Angle Between Vectors

Definition (The Angle Between Vectors)

Suppose that $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$, then the **Angle** Between $\mathbf{x}$ and $\mathbf{y}$ denoted θ is:

$$\theta = Cos^{-1}\left(\frac{\mathbf{x}^T\mathbf{y}}{||\mathbf{x}||_2 * ||\mathbf{y}||_2}\right) = Cos^{-1}(\mathbf{u_x}^T\mathbf{u_y})$$

Remark: The results can be generalized to any space with a well defined Inner Product by replacing $\mathbf{x}^T\mathbf{y}$ with $\langle\mathbf{x}, \mathbf{y}\rangle$

Example:

Suppose that $\mathbf{x} = (3, 4)^T$ and $\mathbf{y} = (-1, 1)$, then:

$$\mathbf{u_x} = \frac{1}{\sqrt{9+16}} * \begin{bmatrix} 3 \\ 4 \end{bmatrix} = \begin{bmatrix} 3/5 \\ 4/5 \end{bmatrix} \;\&\; \mathbf{u_y} = \frac{1}{\sqrt{2}} * \begin{bmatrix} -1 \\ 1 \end{bmatrix} = \begin{bmatrix} -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{bmatrix}$$

$$\theta = Cos^{-1}(\mathbf{u_x}^T\mathbf{u_y}) = 1.43 = 1.43 * \frac{180}{\pi} = 81.87^\circ$$

udemy

Representation: Angle

Angle between $\mathbf{x}$ and $\mathbf{y} \in \mathbb{R}^2$.

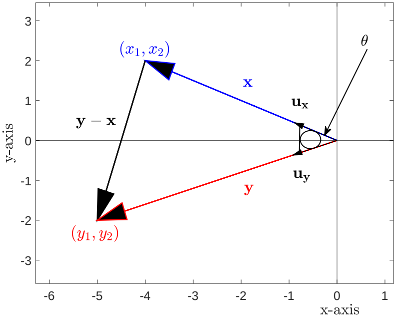
Suppose: $\mathbf{x} = \begin{bmatrix} -4 \\ 2 \end{bmatrix}$ and $\mathbf{y} = \begin{bmatrix} -5 \\ -2 \end{bmatrix}$,
then: the Angle Between $\mathbf{x}$ and $\mathbf{y}$ is θ
where:

$$\mathbf{u_x} = \frac{1}{\sqrt{20}} * \begin{bmatrix} -4 \\ 2 \end{bmatrix} = \begin{bmatrix} -0.8944 \\ 0.4472 \end{bmatrix}$$

$$\mathbf{u_y} = \frac{1}{\sqrt{29}} * \begin{bmatrix} -5 \\ -2 \end{bmatrix} = \begin{bmatrix} -0.9285 \\ -0.3714 \end{bmatrix}$$

$$\theta = Cos^{-1}(\mathbf{u_x}^T \mathbf{u_y}) = 0.84 = 48.37^\circ$$

□



Cauchy-Schwartz Inequality

Theorem (Cauchy-Schwartz Inequality)

Suppose $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$, *then the magnitude of the Scalar Product is always less than or equal than the product of the Norms of the vectors.*

$$\left|\mathbf{x}^T\mathbf{y}\right| \leq ||\mathbf{x}|| * ||\mathbf{y}||$$

udemy

Cauchy-Schwartz Inequality

Theorem (Cauchy-Schwartz Inequality)

Suppose $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$, *then the magnitude of the Scalar Product is always less than or equal than the product of the Norms of the vectors.*

$$\left|\mathbf{x}^T\mathbf{y}\right| \leq ||\mathbf{x}|| * ||\mathbf{y}||$$

Remark: *Equality holds* $\longleftrightarrow$ $Cos(\theta) = 1 \leftrightarrow \theta = 0$ *Or at least one of the vectors is* $\mathbf{0}$, *e.g., one of the vectors is a scalar multiple of the other.*

udemy

Cauchy-Schwartz Inequality

Theorem (Cauchy-Schwartz Inequality)

Suppose $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$, *then the magnitude of the Scalar Product is always less than or equal than the product of the Norms of the vectors.*

$$\left|\mathbf{x}^T \mathbf{y}\right| \leq ||\mathbf{x}|| * ||\mathbf{y}||$$

Remark: *Equality holds* $\longleftrightarrow$ *Cos*$(\theta) = 1 \leftrightarrow \theta = 0$ *Or at least one of the vectors is* $\mathbf{0}$, *e.g., one of the vectors is a scalar multiple of the other.*

Example: Suppose that $\mathbf{x} = (3, 4)^T$ and $\mathbf{y} = (-1, 1)^T$, then:

udemy

Cauchy-Schwartz Inequality

Theorem (Cauchy-Schwartz Inequality)

Suppose $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$, *then the magnitude of the Scalar Product is always less than or equal than the product of the Norms of the vectors.*

$$\left|\mathbf{x}^T\mathbf{y}\right| \leq ||\mathbf{x}|| * ||\mathbf{y}||$$

Remark: *Equality holds* $\longleftrightarrow$ $Cos(\theta) = 1 \leftrightarrow \theta = 0$ *Or at least one of the vectors is* $\mathbf{0}$, *e.g., one of the vectors is a scalar multiple of the other.*

Example: Suppose that $\mathbf{x} = (3,4)^T$ and $\mathbf{y} = (-1,1)^T$, then:

$$|\mathbf{x}^T\mathbf{y}| = |3(-1) + 4(1)| = 1 \leq$$
$$\leq ||\mathbf{x}|| * ||\mathbf{y}|| = 5 * \sqrt{2}$$

Orthogonal Vectors

Definition (Orthogonality)

Suppose $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$, then $\mathbf{x}$ and $\mathbf{y}$ are **Orthogonal** iff their inner product is 0:

$$\mathbf{x}^T \mathbf{y} = 0$$

Example:

udemy

Orthogonal Vectors

Definition (Orthogonality)

Suppose $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$, then $\mathbf{x}$ and $\mathbf{y}$ are **Orthogonal** iff their inner product is 0:

$$\mathbf{x}^T \mathbf{y} = 0$$

Remark: The results also hold for any **Inner Product**.

Example:

udemy

Orthogonal Vectors

Definition (Orthogonality)

Suppose $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$, then $\mathbf{x}$ and $\mathbf{y}$ are **Orthogonal** iff their inner product is 0:

$$\mathbf{x}^T \mathbf{y} = 0$$

Remark: The results also hold for any **Inner Product**.

Example:

1 Show that $\mathbf{x} = (2, -3, 1)^T$ and $\mathbf{y} = (1, 1, 1)^T$ are **Orthogonal**.

udemy

Orthogonal Vectors

Definition (Orthogonality)

Suppose $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$, then $\mathbf{x}$ and $\mathbf{y}$ are **Orthogonal** iff their inner product is 0:

$$\mathbf{x}^T \mathbf{y} = 0$$

Remark: The results also hold for any **Inner Product**.

Example:

1 Show that $\mathbf{x} = (2, -3, 1)^T$ and $\mathbf{y} = (1, 1, 1)^T$ are **Orthogonal**.

Observe that:

$$\mathbf{x}^T \mathbf{y} = 2(1) - 3(1) + 1(1) = 0$$

udemy

Projection & Orthogonality

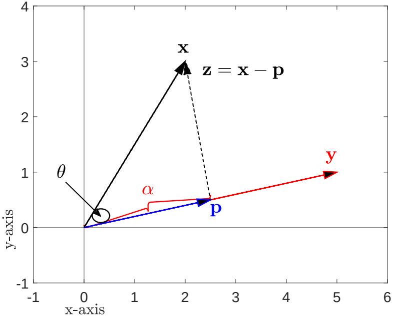
Definition (Scalar and Vector Projection)

Suppose $\mathbf{x}, \mathbf{y} \in \mathbb{R}^2$ and $\theta = Cos^{-1}(\mathbf{u_x}^T \mathbf{u_y})$, then the Scalar Projection of $\mathbf{x}$ onto $\mathbf{y}$ is:

$$\alpha = \frac{\mathbf{x}^T \mathbf{y}}{||\mathbf{y}||} = ||\mathbf{x}|| * Cos(\theta)$$

The Vector Projection of $\mathbf{x}$ onto $\mathbf{y}$ is:

$$\mathbf{p} = \alpha * \mathbf{u_y}$$



Example: Projection

Suppose that $\mathbf{x} = (2,3)^T$ and $\mathbf{y} = (1,-1)^T$, find the Scalar and Vector projection of $\mathbf{x}$ onto $\mathbf{y}$:

udemy

Example: Projection

Suppose that $\mathbf{x} = (2,3)^T$ and $\mathbf{y} = (1,-1)^T$, find the Scalar and Vector projection of $\mathbf{x}$ onto $\mathbf{y}$:

$$\alpha = \frac{\mathbf{x}^T\mathbf{y}}{||\mathbf{y}||} = -\frac{1}{\sqrt{2}}$$

udemy

Example: Projection

Suppose that $\mathbf{x} = (2, 3)^T$ and $\mathbf{y} = (1, -1)^T$, find the Scalar and Vector projection of $\mathbf{x}$ onto $\mathbf{y}$:

$$\alpha = \frac{\mathbf{x}^T\mathbf{y}}{||\mathbf{y}||} = -\frac{1}{\sqrt{2}}$$

$$\mathbf{p} = \alpha \mathbf{u_y} = -\frac{1}{\sqrt{2}} * \begin{bmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \end{bmatrix} = \begin{bmatrix} -1/2 \\ 1/2 \end{bmatrix}$$

udemy

The Cross Product in $\mathbb{R}^3$

Definition (The Cross Product in $\mathbb{R}^3$)

Suppose that $\mathbf{x}, \mathbf{y} \in \mathbb{R}^3$ and let $\hat{i}, \hat{j}, \hat{k}$ represent the columns of I_3, then the **Cross Product** between $\mathbf{x}$ and $\mathbf{y}$ denoted $\mathbf{x} \times \mathbf{y}$ is:

$$\begin{aligned}
\mathbf{x} \times \mathbf{y} &= \begin{vmatrix} \hat{i} & \hat{j} & \hat{k} \\ x_1 & x_2 & x_3 \\ y_1 & y_2 & y_3 \end{vmatrix} \\
&= (x_2y_3 - x_3y_2)\hat{i} - (x_1y_3 - x_3y_1)\hat{j} + (x_1y_2 - x_2y_1)\hat{k} \\
&= \begin{bmatrix} x_2y_3 - x_3y_2 \\ x_3y_1 - x_1y_3 \\ x_1y_2 - x_2y_1 \end{bmatrix}
\end{aligned}$$

udemy

The Cross Product in $\mathbb{R}^3$

Definition (The Cross Product in $\mathbb{R}^3$)

Suppose that $\mathbf{x}, \mathbf{y} \in \mathbb{R}^3$ and let $\hat{i}, \hat{j}, \hat{k}$ represent the columns of I_3, then the **Cross Product** between $\mathbf{x}$ and $\mathbf{y}$ denoted $\mathbf{x} \times \mathbf{y}$ is:

$$\begin{aligned}
\mathbf{x} \times \mathbf{y} &= \begin{vmatrix} \hat{i} & \hat{j} & \hat{k} \\ x_1 & x_2 & x_3 \\ y_1 & y_2 & y_3 \end{vmatrix} \\
&= (x_2 y_3 - x_3 y_2)\hat{i} - (x_1 y_3 - x_3 y_1)\hat{j} + (x_1 y_2 - x_2 y_1)\hat{k} \\
&= \begin{bmatrix} x_2 y_3 - x_3 y_2 \\ x_3 y_1 - x_1 y_3 \\ x_1 y_2 - x_2 y_1 \end{bmatrix}
\end{aligned}$$

Remark: The $\mathbf{x} \times \mathbf{y}$ is orthogonal to both $\mathbf{x}$ and $\mathbf{y}$.

Normal Vector: $\mathbb{R}^3$

Definition (The Normal Vector: $\mathbf{N} \in \mathbb{R}^3$)

Suppose that $\mathbf{x}, \mathbf{y} \in \mathbb{R}^3$, then the **Normal Vector** N is the vector that is orthogonal to $\mathbf{x}$ and $\mathbf{y}$; and hence is orthogonal to any linear combination of $\mathbf{x}$ and $\mathbf{y}$:

udemy

Normal Vector: $\mathbb{R}^3$

Definition (The Normal Vector: $\mathbf{N} \in \mathbb{R}^3$)

Suppose that $\mathbf{x}, \mathbf{y} \in \mathbb{R}^3$, then the **Normal Vector** N is the vector that is orthogonal to $\mathbf{x}$ and $\mathbf{y}$; and hence is orthogonal to any linear combination of $\mathbf{x}$ and $\mathbf{y}$:

$$
\begin{aligned}
\mathbf{N} &= Det(J) \\
&= \begin{vmatrix} \hat{i} & \hat{j} & \hat{k} \\ x_1 & x_2 & x_3 \\ y_1 & y_2 & y_3 \end{vmatrix}
\end{aligned}
$$

udemy

Normal Vector: $\mathbb{R}^3$

Definition (The Normal Vector: $\mathbf{N} \in \mathbb{R}^3$)

Suppose that $\mathbf{x}, \mathbf{y} \in \mathbb{R}^3$, then the **Normal Vector** N is the vector that is orthogonal to $\mathbf{x}$ and $\mathbf{y}$; and hence is orthogonal to any linear combination of $\mathbf{x}$ and $\mathbf{y}$:

$$
\begin{aligned}
\mathbf{N} &= Det(J) \\
&= \begin{vmatrix} \hat{i} & \hat{j} & \hat{k} \\ x_1 & x_2 & x_3 \\ y_1 & y_2 & y_3 \end{vmatrix} \\
&= (x_2 y_3 - x_3 y_2)\hat{i} - (x_1 y_3 - x_3 y_1)\hat{j} + (x_1 y_2 - x_2 y_1)\hat{k} \\
&= \begin{bmatrix} x_2 y_3 - x_3 y_2 \\ x_3 y_1 - x_1 y_3 \\ x_1 y_2 - x_2 y_1 \end{bmatrix} \quad \textit{Where}: \ \langle \mathbf{N}, \alpha\mathbf{x} + \beta\mathbf{y} \rangle = \mathbf{0} \ \textit{For each} \ \alpha, \beta \in \mathbb{R}
\end{aligned}
$$

udemy

Equation of the Plane in $\mathbb{R}^3$

Lemma (Equation of the Plane passing through Point P_0 and is orthogonal to $\mathbf{N}$)

Suppose $P_0 = (x_0, y_0, z_0)^T$ is a point in $\mathbb{R}^3$ and S is the set of all points in $\mathbb{R}^3$ containing the point P_0 and is orthogonal to $\mathbf{N}$, then any point $P = (x, y, z)$ belongs to the Plane Π provided:

$$(\overrightarrow{P_0 P})^T \mathbf{N} = 0 \longleftrightarrow A(x - x_0) + B(y - y_0) + C(z - z_0) = 0$$

$$\textit{Where: } \mathbf{N} = (A, B, C)^T \ \& \ P_0 = (x_0, y_0, z_0)^T.$$

udemy

Example: Normal Vector in $\mathbb{R}^3$

Find the equation of the plane passing through the points: $P_1 = (1, 1, 2)$, $P_2 = (2, 3, 3)$, and $P_3 = (3, -3, 3)$.

Example: Normal Vector in $\mathbb{R}^3$

Find the equation of the plane passing through the points:
$P_1 = (1, 1, 2)$, $P_2 = (2, 3, 3)$, and $P_3 = (3, -3, 3)$.

Proof.

Observe that: $\mathbf{x} = \overrightarrow{P_1P_2} = (1, 2, 1)^T$ and $\mathbf{y} = \overrightarrow{P_1P_3} = (2, -4, 1)^T$. The Normal Vector is obtained by taking the cross product between $\mathbf{x}$ and $\mathbf{y}$:

$\square$

Example: Normal Vector in $\mathbb{R}^3$

Find the equation of the plane passing through the points:
$P_1 = (1,1,2)$, $P_2 = (2,3,3)$, and $P_3 = (3,-3,3)$.

Proof.

Observe that: $\mathbf{x} = \overrightarrow{P_1P_2} = (1,2,1)^T$ and $\mathbf{y} = \overrightarrow{P_1P_3} = (2,-4,1)^T$. The Normal Vector is obtained by taking the cross product between $\mathbf{x}$ and $\mathbf{y}$:

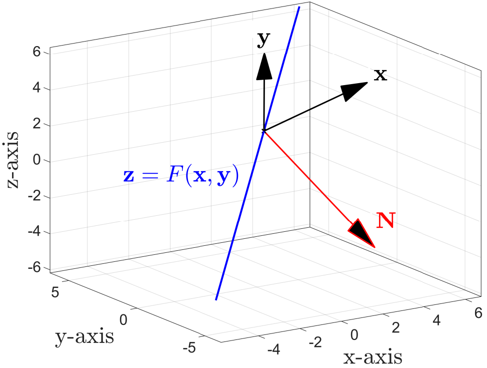
$$\mathbf{N} = \begin{bmatrix} 2(1) - (-4)1 \\ 1(2) - (1)1 \\ 1(-4) - 2(2) \end{bmatrix} = \begin{bmatrix} 6 \\ 1 \\ -8 \end{bmatrix}$$

Hence, the equation of the plane through point P_1 is:

$$6(x-1) + 1(y-1) - 8(z-2) = 0$$

$\square$

Example: Normal Vector in $\mathbb{R}^3$



udemy

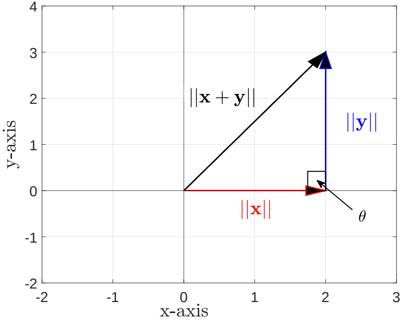
Reference: Leon, 2014 [2]

Pythagorean Law Generalizes Pythagorean Identity

Lemma (Pythagorean Law)

Suppose that $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ *such that* $\mathbf{x}^T\mathbf{y} = 0$, *then:*

$$||\mathbf{x} + \mathbf{y}||_2^2 = ||\mathbf{x}||_2^2 + ||\mathbf{y}||_2^2$$



udemy

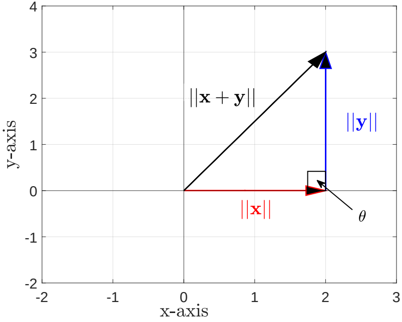
Pythagorean Law Generalizes Pythagorean Identity

Lemma (Pythagorean Law)

Suppose that $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ *such that* $\mathbf{x}^T\mathbf{y} = 0$, *then:*

$$||\mathbf{x} + \mathbf{y}||_2^2 = ||\mathbf{x}||_2^2 + ||\mathbf{y}||_2^2$$

Remark: *The results will hold for any well defined inner product as well. Taking:*
$a = ||\mathbf{x}||_2, b = ||\mathbf{y}||_2, c = ||\mathbf{x}+\mathbf{y}||_2$
then: $c^2 = a^2 + b^2$.
$\mathbf{x}, \mathbf{y} \in \mathbb{R}^2$



Application: Search Database

1. Suppose $A \in \mathbb{Z}^{mxn}$ and $i \sim i_{th}$ *Keyword* and $j \sim j_{th}$ *Module*, then: a_{ij} is the frequency (count) of the i_{th} keyword belonging to j_{th} module.

Application: Search Database

1. Suppose $A \in \mathbb{Z}^{mxn}$ and $i \sim i_{th}$ *Keyword* and $j \sim j_{th}$ *Module*, then: a_{ij} is the frequency (count) of the i_{th} keyword belonging to j_{th} module.
2. We are interested in finding the best search using k of the m keywords belonging to the modules. The key question we are interested in is: which modules are the best searches for the set of keywords in our search?

Application: Search Database

1. Suppose $A \in \mathbb{Z}^{mxn}$ and $i \sim i_{th}$ *Keyword* and $j \sim j_{th}$ *Module*, then: a_{ij} is the frequency (count) of the i_{th} keyword belonging to j_{th} module.
2. We are interested in finding the best search using k of the m keywords belonging to the modules. The key question we are interested in is: which modules are the best searches for the set of keywords in our search?
3. To tackle this issue, one requires techniques to identify the modules that best match the search words, yet need not match identically.

Application: Search Database

1. Suppose $A \in \mathbb{Z}^{mxn}$ and $i \sim i_{th}$ *Keyword* and $j \sim j_{th}$ *Module*, then: a_{ij} is the frequency (count) of the i_{th} keyword belonging to j_{th} module.
2. We are interested in finding the best search using k of the m keywords belonging to the modules. The key question we are interested in is: which modules are the best searches for the set of keywords in our search?
3. To tackle this issue, one requires techniques to identify the modules that best match the search words, yet need not match identically.

Remark: $\mathbb{Z}^{mxn}$ corresponds to all matrices with m rows and n columns where each entry is Integer: $z_{ij} \in \mathbb{Z}$.

Application: Search Database

Suppose we are given the following **Database Table**:

	C1	C2	C3	C4	C5	C6	C7	C8
Determinants	0	6	3	0	1	0	1	1
Eigenvalues	0	0	0	0	0	5	3	2
Linear	5	4	4	5	4	0	3	3
Matrices	6	5	3	3	4	4	3	2
Numerical	0	0	0	0	3	0	4	3
Orthogonality	0	0	0	0	4	6	0	2
Spaces	0	0	5	2	3	3	0	1
Systems	5	3	3	2	4	2	1	1
Transformations	0	0	0	5	1	3	1	0
Vector	0	4	4	3	4	1	0	3

The intersection of **Module**: $C6$ and **Keyword**: Orthogonality is 6 telling us the number of occurrences of Orthogonality in $C6$ appears 6 times.

udemy

Application: Search Database

$$Let: \quad A = \begin{bmatrix} 0 & 6 & 3 & 0 & 1 & 0 & 1 & 1 \\ \hline 0 & 0 & 0 & 0 & 0 & 5 & 3 & 2 \\ \hline 5 & 4 & 4 & 5 & 4 & 0 & 3 & 3 \\ \hline 6 & 5 & 3 & 3 & 4 & 4 & 3 & 2 \\ \hline 0 & 0 & 0 & 0 & 3 & 0 & 4 & 3 \\ \hline 0 & 0 & 0 & 0 & 4 & \color{blue}{6} & 0 & 2 \\ \hline 0 & 0 & 5 & 2 & 3 & 3 & 0 & 1 \\ \hline 5 & 3 & 3 & 2 & 4 & 2 & 1 & 1 \\ \hline 0 & 0 & 0 & 5 & 1 & 3 & 1 & 0 \\ \hline 0 & 4 & 4 & 3 & 4 & 1 & 0 & 3 \end{bmatrix} \longrightarrow A \in \mathbb{R}^{10x8}$$

Take: $\mathbf{q_j} = \frac{\mathbf{a_j}}{||\mathbf{a_j}||_2} \sim Unit\ Vector$

q_{ij} is the Relative (Module) Frequency of Keyword i belonging to Module j.

udemy

Application: Search Database

$$Q = \begin{bmatrix} 0 & 0.594 & 0.327 & 0 & 0.1 & 0 & 0.147 & 0.154 \\ \hline 0 & 0 & 0 & 0 & 0 & 0.5 & 0.442 & 0.309 \\ \hline 0.539 & 0.396 & 0.436 & 0.574 & 0.4 & 0 & 0.442 & 0.463 \\ \hline 0.647 & 0.495 & 0.327 & 0.344 & 0.4 & 0.4 & 0.442 & 0.309 \\ \hline 0 & 0 & 0 & 0 & 0.3 & 0 & 0.59 & 0.463 \\ \hline 0 & 0 & 0 & 0 & 0.4 & 0.6 & 0 & 0.309 \\ \hline 0 & 0 & 0.546 & 0.229 & 0.3 & 0.3 & 0 & 0.154 \\ \hline 0.539 & 0.297 & 0.327 & 0.229 & 0.4 & 0.2 & 0.147 & 0.154 \\ \hline 0 & 0 & 0 & 0.574 & 0.1 & 0.3 & 0.147 & 0 \\ \hline 0 & 0.396 & 0.436 & 0.344 & 0.4 & 0.1 & 0 & 0.463 \\ \hline \end{bmatrix}$$

Suppose we are interested in the $k = 3$ Keywords:
$\{Determinants, \; Linear, \; Transformations\}$

Take $\mathbf{x} \in \mathbb{R}^{10}$ such that $x_i = 1/\sqrt{k}$ for the rows corresponding to the Keywords and 0 otherwise, then: $\mathbf{x} \sim Unit \; Vector$ and is referred to as a **Search Vector**.

Application: Search Database

Taking $y_i = \mathbf{q_i}^T \mathbf{x} = Cos(\theta_i)$ where θ_i is the angle between Unit Vectors will be used as the Optimal Search criteria within the Database. Then, one obtains:

$$\mathbf{x} = \begin{bmatrix} 0.577 \\ \hline 0 \\ \hline 0.577 \\ \hline 0 \\ \hline 0 \\ \hline 0 \\ \hline 0 \\ \hline 0 \\ \hline 0.577 \\ \hline 0 \end{bmatrix} \quad \& \quad \mathbf{y} = Q^T \mathbf{x} = \begin{bmatrix} 0.311 \\ \hline 0.57123 \\ \hline 0.44025 \\ \hline 0.6624 \\ \hline 0.3462 \\ \hline 0.1731 \\ \hline 0.42467 \\ \hline 0.35601 \end{bmatrix}$$

The **Best** search comes from Module $y_4 = C4$ and the **Worst** search comes from Module $y_6 = C6$.

Remark: If $y_i = 0 \longleftrightarrow \mathbf{q_i}^T \mathbf{x} = 0$, which implies that the search vector is orthogonal to the i_{th} column vector of the Database Matrix.

udemy

Application: Statistics

(1) Suppose we are interested in how Assignments, Exams, and Finals are related for a group of 7 students where 200 represents a perfect score.

(2) Consider the raw data set:

	Assignments	Exams	Final
S1	198	200	196
S2	160	165	165
S3	158	158	133
S4	150	165	91
S5	175	182	151
S6	134	135	101
S7	152	136	80
Average	161	163	131
Variance	417.6667	546	1814.3333

We will take $A \in \mathbb{R}^{7x3}$ as the array containing the data of the 7 students for the 3 categories.

udemy

Application: Statistics

After Taking the Deviation from the Average (Mean), we will obtain $X \in \mathbb{R}^{7x3}$ where $x_{ij} = \left(a_{ij} - \sum_{i=1}^{7} \frac{a_{ij}}{7}\right)$:

Application: Statistics

After Taking the Deviation from the Average (Mean), we will obtain $X \in \mathbb{R}^{7x3}$ where $x_{ij} = \left(a_{ij} - \sum_{i=1}^{7} \frac{a_{ij}}{7}\right)$:

$$X = \begin{bmatrix} 37 & 37 & 65 \\ \hline -1 & 2 & 34 \\ \hline -3 & -5 & 2 \\ \hline -11 & 2 & -40 \\ \hline 14 & 19 & 20 \\ \hline -27 & -28 & -30 \\ \hline -9 & -27 & -51 \end{bmatrix}$$

Application: Statistics

After Taking the Deviation from the Average (Mean), we will obtain $X \in \mathbb{R}^{7x3}$ where $x_{ij} = \left(a_{ij} - \sum_{i=1}^{7} \frac{a_{ij}}{7}\right)$:

$$X = \begin{bmatrix} 37 & 37 & 65 \\ \hline -1 & 2 & 34 \\ \hline -3 & -5 & 2 \\ \hline -11 & 2 & -40 \\ \hline 14 & 19 & 20 \\ \hline -27 & -28 & -30 \\ \hline -9 & -27 & -51 \end{bmatrix}$$

We are interested in how the grades are related to one another, that is how does the performance in a student's Assignments (Column 1) impact their Exam (Column 2) and Final (Column 3) scores?

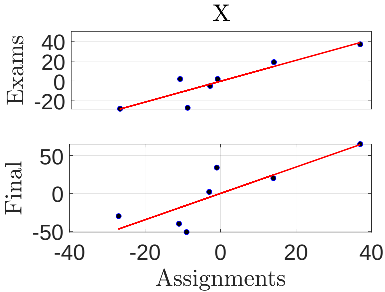
udemy

Application: Statistics

Definition (Deviation Matrix)

Suppose $A \in \mathbb{R}^{mxn}$ and column average $\overline{x}_j = \sum_{i=1}^{m} \frac{a_{ij}}{m}$, then The Deviation Matrix $X \in \mathbb{R}^{mxn}$ where $x_{ij} = a_{ij} - \overline{x}_j$. In our example:

$$X = \begin{bmatrix} 37 & 37 & 65 \\ \hline -1 & 2 & 34 \\ \hline -3 & -5 & 2 \\ \hline -11 & 2 & -40 \\ \hline 14 & 19 & 20 \\ \hline -27 & -28 & -30 \\ \hline -9 & -27 & -51 \end{bmatrix}$$



udemy

Application: Statistics

Definition (Standardized Matrix)

Suppose that $X \in \mathbb{R}^{mxn}$ is The Deviation Matrix, consider $s_j^2 = \sum_{i=1}^{m} \frac{(a_{ij} - \bar{x}_j)^2}{m-1}$ as the Sample Variance of the j_{th} column of A, then The Standardized Matrix is: $Z \in \mathbb{R}^{mxn}$ where $z_{ij} = \frac{a_{ij} - \bar{x}_j}{s_j}$.

udemy

Application: Statistics

Definition (Standardized Matrix)

Suppose that $X \in \mathbb{R}^{mxn}$ is The Deviation Matrix, consider $s_j^2 = \sum_{i=1}^{m} \frac{(a_{ij} - \overline{x}_j)^2}{m-1}$ as the Sample Variance of the j_{th} column of A, then The Standardized Matrix is: $Z \in \mathbb{R}^{mxn}$ where $z_{ij} = \frac{a_{ij} - \overline{x}_j}{s_j}$.

Remark: If the sample size $m \geq 30$ or the data from the sample is approximately normally distributed, then z_{ij} is the z-score associated to the Normal (Bell or Gaussian) curve, which tells us the number of standard deviations from the mean at the data point.

Application: Statistics

Definition (Standardized Matrix)

Suppose that $X \in \mathbb{R}^{mxn}$ is The Deviation Matrix, consider $s_j^2 = \sum_{i=1}^{m} \frac{(a_{ij} - \overline{x}_j)^2}{m-1}$ as the Sample Variance of the j_{th} column of A, then The Standardized Matrix is: $Z \in \mathbb{R}^{mxn}$ where $z_{ij} = \frac{a_{ij} - \overline{x}_j}{s_j}$.

Remark: If the sample size $m \geq 30$ or the data from the sample is approximately normally distributed, then z_{ij} is the z-score associated to the Normal (Bell or Gaussian) curve, which tells us the number of standard deviations from the mean at the data point.

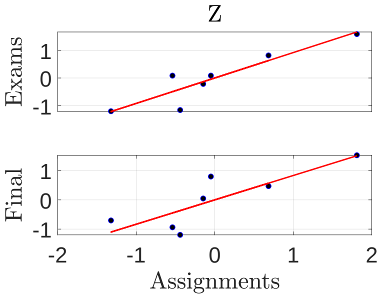
In our example:

$$Z = \begin{bmatrix} 1.8105 & 1.5835 & 1.526 \\ -0.048931 & 0.085592 & 0.79822 \\ -0.14679 & -0.21398 & 0.046954 \\ -0.53824 & 0.085592 & -0.93908 \\ 0.68504 & 0.81312 & 0.46954 \\ -1.3211 & -1.1983 & -0.70431 \\ -0.44038 & -1.1555 & -1.1973 \end{bmatrix}$$

Application: Statistics

$$Z = \left[\begin{array}{ccc} 1.81 & 1.58 & 1.53 \\ \hline -0.049 & 0.09 & 0.80 \\ \hline -0.15 & -0.21 & 0.05 \\ \hline -0.54 & 0.09 & -0.94 \\ \hline 0.69 & \color{red}{0.81} & 0.47 \\ \hline -1.32 & -1.20 & -0.70 \\ \hline -0.44 & -1.16 & -1.20 \end{array}\right]$$

Remark: The intersection of the 5_{th} row and 2_{nd} column is 0.81, which tells us that The Exam score for Student 5 is 0.81 Standard Deviations above the mean.



Application: Statistics

Definition (Correlation Matrix)

Suppose $X \in \mathbb{R}^{mxn}$ is **The Deviation Matrix**, take $U \in \mathbb{R}^{mxn}$ whose j_{th} column is: $\mathbf{u_j} = \frac{\mathbf{x_j}}{||\mathbf{x_j}||_2}$ for $j = \overline{1, n}$, then the **Correlation Matrix** is:

$$C = U'U \ \ni: c_{ij} = \begin{cases} > 0 \longleftrightarrow \textit{Positively Correlated} \\ = 0 \longleftrightarrow \textit{Uncorrelated} \\ < 0 \longleftrightarrow \textit{Negatively Correlated} \end{cases}$$

udemy

Application: Statistics

Definition (Correlation Matrix)

Suppose $X \in \mathbb{R}^{mxn}$ is **The Deviation Matrix**, take $U \in \mathbb{R}^{mxn}$ whose j_{th} column is: $\mathbf{u_j} = \frac{\mathbf{x_j}}{||\mathbf{x_j}||_2}$ for $j = \overline{1, n}$, then the **Correlation Matrix** is:

$$C = U'U \ \ni: c_{ij} = \begin{cases} > 0 \longleftrightarrow \textit{Positively Correlated} \\ = 0 \longleftrightarrow \textit{Uncorrelated} \\ < 0 \longleftrightarrow \textit{Negatively Correlated} \end{cases}$$

Remark: $|c_{ij}| \leq 1$ tells us correlation between the i_{th} and j_{th} columns of X.

udemy

Application: Statistics

Definition (Correlation Matrix)

Suppose $X \in \mathbb{R}^{mxn}$ is **The Deviation Matrix**, take $U \in \mathbb{R}^{mxn}$ whose j_{th} column is: $\mathbf{u_j} = \frac{\mathbf{x_j}}{||\mathbf{x_j}||_2}$ for $j = \overline{1, n}$, then the **Correlation Matrix** is:

$$C = U'U \;\ni: c_{ij} = \begin{cases} > 0 \longleftrightarrow \textit{Positively Correlated} \\ = 0 \longleftrightarrow \textit{Uncorrelated} \\ < 0 \longleftrightarrow \textit{Negatively Correlated} \end{cases}$$

Remark: $|c_{ij}| \leq 1$ tells us correlation between the i_{th} and j_{th} columns of X.

In our example:

$$C = \begin{bmatrix} 1 & 0.91615 & 0.83361 \\ 0.91615 & 1 & 0.83392 \\ 0.83361 & 0.83392 & 1 \end{bmatrix}$$

udemy

Application: Statistics

Definition (Covariance Matrix)

Suppose $X \in \mathbb{R}^{mxn}$ is the Deviation Matrix where $x_{ij} = \left(a_{ij} - \sum_{i=1}^{m} \frac{a_{ij}}{m}\right)$, then the Covariance Matrix is $C_v = \frac{1}{m-1} X'X$ is the Covariance between x_i and x_j denoted $c_{v_{ij}} = Cov(\mathbf{x_i}, \mathbf{x_j}) = \frac{\mathbf{x_i}^T \mathbf{x_j}}{m-1}$

udemy

Application: Statistics

Definition (Covariance Matrix)

Suppose $X \in \mathbb{R}^{mxn}$ is the Deviation Matrix where $x_{ij} = \left(a_{ij} - \sum_{i=1}^{m} \frac{a_{ij}}{m}\right)$, then the Covariance Matrix is $C_v = \frac{1}{m-1} X'X$ is the Covariance between x_i and x_j denoted $c_{v_{ij}} = Cov(\mathbf{x_i}, \mathbf{x_j}) = \frac{\mathbf{x_i}^T \mathbf{x_j}}{m-1}$

Remark: The diagonal elements of C_v is $c_{v_{jj}} = Cov(\mathbf{x_j}, \mathbf{x_j}) = Var(\mathbf{x_j})$. That is the diagonal elements correspond to the Variance of the column entries of X.

udemy

Application: Statistics

Definition (Covariance Matrix)

Suppose $X \in \mathbb{R}^{mxn}$ is the Deviation Matrix where $x_{ij} = \left(a_{ij} - \sum_{i=1}^{m} \frac{a_{ij}}{m}\right)$, then the Covariance Matrix is $C_v = \frac{1}{m-1} X'X$ is the Covariance between x_i and x_j denoted $c_{v_{ij}} = Cov(\mathbf{x_i}, \mathbf{x_j}) = \frac{\mathbf{x_i}^T \mathbf{x_j}}{m-1}$

Remark: The diagonal elements of C_v is $c_{v_{jj}} = Cov(\mathbf{x_j}, \mathbf{x_j}) = Var(\mathbf{x_j})$. That is the diagonal elements correspond to the Variance of the column entries of X.

In our example:

$$C_v = \frac{1}{m-1} X'X = \begin{bmatrix} 417.6667 & 437.5 & 725.6667 \\ 437.5 & 546 & 830 \\ 725.6667 & 830 & 1814.3333 \end{bmatrix}$$

udemy

Application: Psychology

1. In 1904, Psychologist Charles Spearman conducted a project where he gave a collection of exams to 23 students to determine the relational learning between different subjects.

The results are summarized as follows:

Application: Psychology

1. In 1904, Psychologist Charles Spearman conducted a project where he gave a collection of exams to 23 students to determine the relational learning between different subjects.
The results are summarized as follows:

	Classics	French	English	Math	Music
Classics	1	0.83	0.78	0.7	0.63
French	0.83	1	0.67	0.67	0.57
English	0.78	0.67	1	0.64	0.51
Math	0.7	0.67	0.64	1	0.51
Music	0.63	0.57	0.51	0.51	1

udemy

Application: Psychology

1. In 1904, Psychologist Charles Spearman conducted a project where he gave a collection of exams to 23 students to determine the relational learning between different subjects.

The results are summarized as follows:

	Classics	French	English	Math	Music
Classics	1	0.83	0.78	0.7	0.63
French	0.83	1	0.67	0.67	0.57
English	0.78	0.67	1	0.64	0.51
Math	0.7	0.67	0.64	1	0.51
Music	0.63	0.57	0.51	0.51	1

Remark: We can take $C \in \mathbb{R}^{5x5}$ as the Correlation Matrix where $c_{ij} = Cov(x_i, x_j) = \frac{\mathbf{x_i}^T \mathbf{x_j}}{n-1}$.

udemy

Table of Contents

udemy

Orthogonal Subspaces

Recall: Given $A \in \mathbb{R}^{mxn}$, $\mathbf{x} \in Nul(A) \longleftrightarrow Ax = \mathbf{a_1}x_1 + \mathbf{a_2}x_2 + \cdots + \mathbf{a_n}x_n = \mathbf{0}$, which implies that: $\mathbf{x}$ is Orthogonal to each column of A^T, so that $\mathbf{x}$ is Orthogonal to the Span of the columns of A^T, e.g., if $\mathbf{y} \in C(A^T) \longrightarrow \mathbf{y}^T\mathbf{x} = 0$

udemy

Orthogonal Subspaces

Recall: Given $A \in \mathbb{R}^{mxn}$, $\mathbf{x} \in Nul(A) \longleftrightarrow Ax = \mathbf{a_1}x_1 + \mathbf{a_2}x_2 + \cdots + \mathbf{a_n}x_n = \mathbf{0}$, which implies that: $\mathbf{x}$ is Orthogonal to each column of A^T, so that $\mathbf{x}$ is Orthogonal to the Span of the columns of A^T, e.g., if $\mathbf{y} \in C(A^T) \longrightarrow \mathbf{y}^T\mathbf{x} = 0$

Definition (Orthogonal Subspaces: $U \perp V$)

Let U, $V \sim$ *Vector Spaces*, and consider $X \subset U$ and $Y \subset V$, then X and Y are **Orthogonal Subspaces**: $X \perp Y \longleftrightarrow$ for each $\mathbf{x} \in X$ and $\mathbf{y} \in Y$, $\langle \mathbf{x}, \mathbf{y} \rangle = 0$

udemy

Orthogonal Subspaces

Recall: Given $A \in \mathbb{R}^{mxn}$, $\mathbf{x} \in Nul(A) \longleftrightarrow Ax = \mathbf{a_1}x_1 + \mathbf{a_2}x_2 + \cdots + \mathbf{a_n}x_n = \mathbf{0}$, which implies that: $\mathbf{x}$ is Orthogonal to each column of A^T, so that $\mathbf{x}$ is Orthogonal to the Span of the columns of A^T, e.g., if $\mathbf{y} \in C(A^T) \longrightarrow \mathbf{y}^T\mathbf{x} = 0$

Definition (Orthogonal Subspaces: $U \perp V$)

Let U, $V \sim$ *Vector Spaces*, and consider $X \subset U$ and $Y \subset V$, then X and Y are **Orthogonal Subspaces**: $X \perp Y \longleftrightarrow$ for each $\mathbf{x} \in X$ and $\mathbf{y} \in Y$, $\langle \mathbf{x}, \mathbf{y} \rangle = 0$

Remark: $\mathbf{x}$ is Orthogonal to Y ($\mathbf{x} \perp Y$) $\longleftrightarrow$ for each $\mathbf{y} \in Y$, $\langle \mathbf{x}, \mathbf{y} \rangle = 0$.

For the scalar product, one often has: $U = \mathbb{R}^n = V$ and $\langle \mathbf{x}, \mathbf{y} \rangle = \mathbf{x}^T\mathbf{y}$

Orthogonal Subspaces

Recall: Given $A \in \mathbb{R}^{mxn}$, $\mathbf{x} \in Nul(A) \longleftrightarrow Ax = \mathbf{a_1}x_1 + \mathbf{a_2}x_2 + \cdots + \mathbf{a_n}x_n = \mathbf{0}$, which implies that: $\mathbf{x}$ is Orthogonal to each column of A^T, so that $\mathbf{x}$ is Orthogonal to the Span of the columns of A^T, e.g., if $\mathbf{y} \in C(A^T) \longrightarrow \mathbf{y}^T\mathbf{x} = 0$

Definition (Orthogonal Subspaces: $U \perp V$)

Let U, $V \sim$ *Vector Spaces*, and consider $X \subset U$ and $Y \subset V$, then X and Y are **Orthogonal Subspaces**: $X \perp Y \longleftrightarrow$ for each $\mathbf{x} \in X$ and $\mathbf{y} \in Y$, $\langle \mathbf{x}, \mathbf{y} \rangle = 0$
Remark: $\mathbf{x}$ is Orthogonal to Y $(\mathbf{x} \perp Y) \longleftrightarrow$ for each $\mathbf{y} \in Y$, $\langle \mathbf{x}, \mathbf{y} \rangle = 0$.
For the scalar product, one often has: $U = \mathbb{R}^n = V$ and $\langle \mathbf{x}, \mathbf{y} \rangle = \mathbf{x}^T\mathbf{y}$

Lemma (Orthogonality of a Basis)

Suppose $U, V \sim$ Vector Space and $S = \{\mathbf{u_1}, \mathbf{u_2}, \ldots, \mathbf{u_n}\}$ is a basis on U, then $\mathbf{y} \in V$ is Orthogonal to $U \longleftrightarrow \mathbf{y}$ is Orthogonal to each Basis element in S: $\langle \mathbf{y}, \mathbf{u_j} \rangle = 0$

udemy

Example: Orthogonal Subspaces

Suppose $X = \{\alpha \mathbf{e_1} \,|\, \mathbf{e_1} \in \mathbb{R}^3, \alpha \in \mathbb{R}\}$ and $Y = \{\beta \mathbf{e_2} \,|\, \mathbf{e_2} \in \mathbb{R}^3, \beta \in \mathbb{R}\}$, then $X, Y \subset \mathbb{R}^3$.

Show that $X \perp Y$

udemy

Example: Orthogonal Subspaces

Suppose $X = \{\alpha \mathbf{e_1} \mid \mathbf{e_1} \in \mathbb{R}^3, \alpha \in \mathbb{R}\}$ and $Y = \{\beta \mathbf{e_2} \mid \mathbf{e_2} \in \mathbb{R}^3, \beta \in \mathbb{R}\}$, then $X, Y \subset \mathbb{R}^3$.

Show that $X \perp Y$

Proof.

For any $\mathbf{x} \in X$, one has: $\mathbf{x} = \begin{bmatrix} x_1 \\ 0 \\ 0 \end{bmatrix}$ and for each $\mathbf{y} \in Y$, one has: $\mathbf{y} = \begin{bmatrix} 0 \\ y_2 \\ 0 \end{bmatrix}$

Then, it follows that:

In fact, Any vector of the form $y^* = \begin{bmatrix} 0 \\ y_2 \\ y_3 \end{bmatrix}$ is Orthogonal to X, we will refer to

$Span(\mathbf{e_2}, \mathbf{e_3})$ as The **Orthogonal Complement** of X denoted $X^{\perp}$

□

udemy

Example: Orthogonal Subspaces

Suppose $X = \{\alpha \mathbf{e_1} \mid \mathbf{e_1} \in \mathbb{R}^3, \alpha \in \mathbb{R}\}$ and $Y = \{\beta \mathbf{e_2} \mid \mathbf{e_2} \in \mathbb{R}^3, \beta \in \mathbb{R}\}$, then $X, Y \subset \mathbb{R}^3$.

Show that $X \perp Y$

Proof.

For any $\mathbf{x} \in X$, one has: $\mathbf{x} = \begin{bmatrix} x_1 \\ 0 \\ 0 \end{bmatrix}$ and for each $\mathbf{y} \in Y$, one has: $\mathbf{y} = \begin{bmatrix} 0 \\ y_2 \\ 0 \end{bmatrix}$

Then, it follows that:

$$\mathbf{x}^T\mathbf{y} = x_1(0) + 0(y_2) + 0(0) = 0$$

Therefore, $X \perp Y$

In fact, Any vector of the form $y^* = \begin{bmatrix} 0 \\ y_2 \\ y_3 \end{bmatrix}$ is Orthogonal to X, we will refer to

$Span(\mathbf{e_2}, \mathbf{e_3})$ as The **Orthogonal Complement** of X denoted $X^\perp$

$\square$

Orthogonal Complement

Definition (Orthogonal Complement)

Let $Y, U \sim$ *Vector Spaces* and suppose $Y \subset U$, then The Orthogonal Complement of Y relative to U are all the vectors belonging to U that are Orthogonal to Y:

$$\begin{aligned} Y^{\perp} &= \{\mathbf{x} \in U \,|\, \mathbf{x} \perp Y\} \\ &= \{\mathbf{x} \in U \,|\, \langle \mathbf{x}, \mathbf{y} \rangle = 0 \textit{ for each } \mathbf{y} \in Y\} \end{aligned}$$

Orthogonal Complement

Definition (Orthogonal Complement)

Let $Y, U \sim$ *Vector Spaces* and suppose $Y \subset U$, then The Orthogonal Complement of Y relative to U are all the vectors belonging to U that are Orthogonal to Y:

$$\begin{aligned} Y^{\perp} &= \{\mathbf{x} \in U \,|\, \mathbf{x} \perp Y\} \\ &= \{\mathbf{x} \in U \,|\, \langle \mathbf{x}, \mathbf{y} \rangle = 0 \textit{ for each } \mathbf{y} \in Y\} \end{aligned}$$

Remarks: 1 We generally will take: $U = \mathbb{R}^n$ and $\langle \mathbf{x}, \mathbf{y} \rangle = \mathbf{x}^T \mathbf{y}$

udemy

Orthogonal Complement

Definition (Orthogonal Complement)

Let $Y, U \sim$ *Vector Spaces* and suppose $Y \subset U$, then The Orthogonal Complement of Y relative to U are all the vectors belonging to U that are Orthogonal to Y:

$$\begin{aligned} Y^{\perp} &= \{\mathbf{x} \in U \mid \mathbf{x} \perp Y\} \\ &= \{\mathbf{x} \in U \mid \langle \mathbf{x}, \mathbf{y} \rangle = 0 \textit{ for each } \mathbf{y} \in Y\} \end{aligned}$$

Remarks: 1 We generally will take: $U = \mathbb{R}^n$ and $\langle \mathbf{x}, \mathbf{y} \rangle = \mathbf{x}^T \mathbf{y}$

2 If $X, Y \subset \mathbb{R}^n$ where $X \perp Y$, then: $X \cap Y = \{\mathbf{0}\}$

udemy

Orthogonal Complement

Definition (Orthogonal Complement)

Let $Y, U \sim$ *Vector Spaces* and suppose $Y \subset U$, then The Orthogonal Complement of Y relative to U are all the vectors belonging to U that are Orthogonal to Y:

$$\begin{aligned} Y^{\perp} &= \{\mathbf{x} \in U \mid \mathbf{x} \perp Y\} \\ &= \{\mathbf{x} \in U \mid \langle \mathbf{x}, \mathbf{y} \rangle = 0 \textit{ for each } \mathbf{y} \in Y\} \end{aligned}$$

Remarks: 1 We generally will take: $U = \mathbb{R}^n$ and $\langle \mathbf{x}, \mathbf{y} \rangle = \mathbf{x}^T \mathbf{y}$

2 If $X, Y \subset \mathbb{R}^n$ where $X \perp Y$, then: $X \cap Y = \{\mathbf{0}\}$

3 If $Y \subset \mathbb{R}^n \longrightarrow Y^{\perp} \subset \mathbb{R}^n$

udemy

Orthogonal Complement

Definition (Orthogonal Complement)

Let $Y, U \sim$ *Vector Spaces* and suppose $Y \subset U$, then The Orthogonal Complement of Y relative to U are all the vectors belonging to U that are Orthogonal to Y:

$$\begin{aligned} Y^{\perp} &= \{\mathbf{x} \in U \mid \mathbf{x} \perp Y\} \\ &= \{\mathbf{x} \in U \mid \langle \mathbf{x}, \mathbf{y} \rangle = 0 \textit{ for each } \mathbf{y} \in Y\} \end{aligned}$$

Remarks: 1 We generally will take: $U = \mathbb{R}^n$ and $\langle \mathbf{x}, \mathbf{y} \rangle = \mathbf{x}^T \mathbf{y}$

2 If $X, Y \subset \mathbb{R}^n$ where $X \perp Y$, then: $X \cap Y = \{\mathbf{0}\}$

3 If $Y \subset \mathbb{R}^n \longrightarrow Y^{\perp} \subset \mathbb{R}^n$

Example: Take $Y = Span(\mathbf{e}_2) \subset \mathbb{R}^3$, then: The **Orthogonal Complement** of Y relative to $\mathbb{R}^3$ is: $Y^{\perp} = Span(\mathbf{e}_1, \mathbf{e}_3)$

udemy

The Fundamental Subspace Theorem: FST

Theorem (Fundamental Subspace Theorem)

Suppose that $A \in \mathbb{R}^{m x n}$*, then*

$$N(A) \perp C(A) \longleftrightarrow N(A) \perp R(A^T) \longleftrightarrow N(A) = R(A^T)^{\perp} \text{ and } N(A^T) = R(A)^{\perp}$$

udemy

The Fundamental Subspace Theorem: FST

Theorem (Fundamental Subspace Theorem)

Suppose that $A \in \mathbb{R}^{mxn}$, *then*

$$N(A) \perp C(A) \longleftrightarrow N(A) \perp R(A^T) \longleftrightarrow N(A) = R(A^T)^{\perp} \text{ and } N(A^T) = R(A)^{\perp}$$

Proof.

Since Nul(A) $\perp$ Col(A) as $Nul(A), Col(A) \subset \mathbb{R}^n$ and $Nullity(A) + Rank(A) = n$, then $Nul(A) \subset Col(A)$

$\square$

The Fundamental Subspace Theorem: FST

Theorem (Fundamental Subspace Theorem)

Suppose that $A \in \mathbb{R}^{mxn}$, then

$$N(A) \perp C(A) \longleftrightarrow N(A) \perp R(A^T) \longleftrightarrow N(A) = R(A^T)^\perp \text{ and } N(A^T) = R(A)^\perp$$

Proof.

Since Nul(A) $\perp$ Col(A) as $Nul(A), Col(A) \subset \mathbb{R}^n$ and $Nullity(A) + Rank(A) = n$, then $Nul(A) \subset Col(A)$

Conversely: Suppose $\mathbf{x} \in R(A^T)^\perp$, then $\mathbf{x}$ is Orthogonal to each Row vector of A^T and hence is Orthogonal to each Column vector of A. It follows that:

$$A\mathbf{x} = \mathbf{0} \rightarrow \mathbf{x} \in Nul(A)$$

$\square$

The Fundamental Subspace Theorem: FST

Theorem (Fundamental Subspace Theorem)

Suppose that $A \in \mathbb{R}^{mxn}$, *then*

$$N(A) \perp C(A) \longleftrightarrow N(A) \perp R(A^T) \longleftrightarrow N(A) = R(A^T)^\perp \text{ and } N(A^T) = R(A)^\perp$$

Proof.

Since Nul(A) $\perp$ Col(A) as $Nul(A), Col(A) \subset \mathbb{R}^n$ and $Nullity(A) + Rank(A) = n$, then $Nul(A) \subset Col(A)$

Conversely: Suppose $\mathbf{x} \in R(A^T)^\perp$, then $\mathbf{x}$ is Orthogonal to each Row vector of A^T and hence is Orthogonal to each Column vector of A. It follows that:

$$A\mathbf{x} = \mathbf{0} \to \mathbf{x} \in Nul(A)$$

Remarks: 1 The second part of the proof can be shown by substituting $B^T = A$

$\square$

The Fundamental Subspace Theorem: FST

Theorem (Fundamental Subspace Theorem)

Suppose that $A \in \mathbb{R}^{mxn}$*, then*

$$N(A) \perp C(A) \longleftrightarrow N(A) \perp R(A^T) \longleftrightarrow N(A) = R(A^T)^{\perp} \text{ and } N(A^T) = R(A)^{\perp}$$

Proof.

Since Nul(A) $\perp$ Col(A) as $Nul(A), Col(A) \subset \mathbb{R}^n$ and $Nullity(A) + Rank(A) = n$, then $Nul(A) \subset Col(A)$

Conversely: Suppose $\mathbf{x} \in R(A^T)^{\perp}$, then $\mathbf{x}$ is Orthogonal to each Row vector of A^T and hence is Orthogonal to each Column vector of A. It follows that:

$$A\mathbf{x} = \mathbf{0} \to \mathbf{x} \in Nul(A)$$

Remarks: 1 The second part of the proof can be shown by substituting $B^T = A$

2 **Important**: The second part $N(A^T) = R(A)^{\perp}$ plays a significant role when finding the "Closest" solution to Inconsistent Systems: $Ax = b$ with **no solution**.

$\square$

Example: FST

Show that $Nul(A^T) \perp Col(A)$

Let $A = \begin{bmatrix} 1 & 0 \\ 2 & 0 \end{bmatrix}$ and $A^T = \begin{bmatrix} 1 & 2 \\ 0 & 0 \end{bmatrix}$ Then $Col(A) = \left\{ \alpha * \begin{bmatrix} 1 \\ 2 \end{bmatrix} \;\middle|\; \alpha \in \mathbb{R} \right\}$

udemy

Example: FST

Show that $Nul(A^T) \perp Col(A)$

Let $A = \begin{bmatrix} 1 & 0 \\ 2 & 0 \end{bmatrix}$ and $A^T = \begin{bmatrix} 1 & 2 \\ 0 & 0 \end{bmatrix}$ Then $Col(A) = \left\{ \alpha * \begin{bmatrix} 1 \\ 2 \end{bmatrix} \middle| \alpha \in \mathbb{R} \right\}$

Observe $Nul(A^T) = \left\{ \beta * \begin{bmatrix} -2 \\ 1 \end{bmatrix} \middle| \beta \in \mathbb{R} \right\}$

udemy

Example: FST

Show that $Nul(A^T) \perp Col(A)$

Let $A = \begin{bmatrix} 1 & 0 \\ 2 & 0 \end{bmatrix}$ and $A^T = \begin{bmatrix} 1 & 2 \\ 0 & 0 \end{bmatrix}$ Then $Col(A) = \left\{ \alpha * \begin{bmatrix} 1 \\ 2 \end{bmatrix} \;\middle|\; \alpha \in \mathbb{R} \right\}$

Observe $Nul(A^T) = \left\{ \beta * \begin{bmatrix} -2 \\ 1 \end{bmatrix} \;\middle|\; \beta \in \mathbb{R} \right\}$

Observe: $\langle (-2, 1)^T, (1, 2)^T \rangle = 0$ and $Nul(A^T) \perp Col(A)$

udemy

Basis for Disjoint Orthogonal Subspaces

Theorem (Basis of Disjoint Subspaces)

Let $S \subset \mathbb{R}^n$ and $S^\perp$ is the Orthogonal Subspace of S relative to $\mathbb{R}^n$. Suppose we are given $\mathcal{B}_1 = \{\mathbf{x_1}, \mathbf{x_2}, \ldots, \mathbf{x_r}\}$ is a Basis on S and $\mathcal{B}_2 = \{\mathbf{x_{r+1}}, \ldots, \mathbf{x_n}\}$ is a Basis on $S^\perp$ such that $\mathcal{B}_1 \cap \mathcal{B}_2 = \{\mathbf{0}\}$ Then $\mathcal{B} = \{\mathbf{x_1}, \mathbf{x_2}, \ldots, \mathbf{x_r}, \mathbf{x_{r+1}}, \ldots, \mathbf{x_n}\} = \mathcal{B}_1 \cup \mathcal{B}_2 = \mathbb{R}^n$.

udemy

Basis for Disjoint Orthogonal Subspaces

Theorem (Basis of Disjoint Subspaces)

Let $S \subset \mathbb{R}^n$ and $S^\perp$ is the Orthogonal Subspace of S relative to $\mathbb{R}^n$.
Suppose we are given $\mathcal{B}_1 = \{\mathbf{x_1}, \mathbf{x_2}, \ldots, \mathbf{x_r}\}$ is a Basis on S and
$\mathcal{B}_2 = \{\mathbf{x_{r+1}}, \ldots, \mathbf{x_n}\}$ is a Basis on $S^\perp$ such that $\mathcal{B}_1 \cap \mathcal{B}_2 = \{\mathbf{0}\}$
Then $\mathcal{B} = \{\mathbf{x_1}, \mathbf{x_2}, \ldots, \mathbf{x_r}, \mathbf{x_{r+1}}, \ldots, \mathbf{x_n}\} = \mathcal{B}_1 \cup \mathcal{B}_2 = \mathbb{R}^n$.
Furthermore: $dim(S) + dim(S^\perp) = r + (n - r) = n$

Basis for Disjoint Orthogonal Subspaces

Theorem (Basis of Disjoint Subspaces)

Let $S \subset \mathbb{R}^n$ and $S^{\perp}$ is the Orthogonal Subspace of S relative to $\mathbb{R}^n$. Suppose we are given $\mathcal{B}_1 = \{\mathbf{x_1}, \mathbf{x_2}, \ldots, \mathbf{x_r}\}$ is a Basis on S and $\mathcal{B}_2 = \{\mathbf{x_{r+1}}, \ldots, \mathbf{x_n}\}$ is a Basis on $S^{\perp}$ such that $\mathcal{B}_1 \cap \mathcal{B}_2 = \{\mathbf{0}\}$

Then $\mathcal{B} = \{\mathbf{x_1}, \mathbf{x_2}, \ldots, \mathbf{x_r}, \mathbf{x_{r+1}}, \ldots, \mathbf{x_n}\} = \mathcal{B}_1 \cup \mathcal{B}_2 = \mathbb{R}^n$.

Furthermore: $dim(S) + dim(S^{\perp}) = r + (n - r) = n$

Remark: *Special case is $Nul(A) = R(A^T)^{\perp} \rightarrow Nul(A) \perp C(A) \rightarrow$ $Nullity(A) + Rank(A) = (n - r) + r = n$*

The Direct Sum

Definition (Direct Sum: $U \oplus V$)

If $U, V \subset W$ where $W \sim$ *Vector Space* where each $\mathbf{w} \in W$: $\mathbf{w} = \mathbf{u} + \mathbf{v}$ for some $\mathbf{u} \in U$ and $\mathbf{v} \in V$, then: W is The Direct Sum of U and V denoted:

$$W = U \oplus V$$

udemy

The Direct Sum

Definition (Direct Sum: $U \oplus V$)

If $U, V \subset W$ where $W \sim$ *Vector Space* where each $\mathbf{w} \in W$: $\mathbf{w} = \mathbf{u} + \mathbf{v}$ for some $\mathbf{u} \in U$ and $\mathbf{v} \in V$, then: W is The Direct Sum of U and V denoted:

$$W = U \oplus V$$

Theorem (Direct Sum of an Arbitrary Subspace of $\mathbb{R}^n$)

Suppose $S \subset \mathbb{R}^n$ and $S^\perp$ is the Orthogonal Complement of S with respect to $\mathbb{R}^n$, then:

$$S \oplus S^\perp = \mathbb{R}^n$$

udemy

The Direct Sum

Definition (Direct Sum: $U \oplus V$)

If $U, V \subset W$ where $W \sim$ *Vector Space* where each $\mathbf{w} \in W$: $\mathbf{w} = \mathbf{u} + \mathbf{v}$ for some $\mathbf{u} \in U$ and $\mathbf{v} \in V$, then: W is The Direct Sum of U and V denoted:

$$W = U \oplus V$$

Theorem (Direct Sum of an Arbitrary Subspace of $\mathbb{R}^n$)

Suppose $S \subset \mathbb{R}^n$ and $S^\perp$ is the Orthogonal Complement of S with respect to $\mathbb{R}^n$, then:

$$S \oplus S^\perp = \mathbb{R}^n$$

Remark: *Taking* $S = Nul(A^T) \longrightarrow S^\perp = Nul(A^T)^\perp = R(A)$. *Hence,* $\mathbb{R}^n = Nul(A^T) \oplus R(A)$

The Direct Sum

Definition (Direct Sum: $U \oplus V$)

If $U, V \subset W$ where $W \sim$ *Vector Space* where each $\mathbf{w} \in W$: $\mathbf{w} = \mathbf{u} + \mathbf{v}$ for some $\mathbf{u} \in U$ and $\mathbf{v} \in V$, then: W is The Direct Sum of U and V denoted:

$$W = U \oplus V$$

Theorem (Direct Sum of an Arbitrary Subspace of $\mathbb{R}^n$)

Suppose $S \subset \mathbb{R}^n$ and $S^\perp$ is the Orthogonal Complement of S with respect to $\mathbb{R}^n$, then:

$$S \oplus S^\perp = \mathbb{R}^n$$

Remark: *Taking $S = Nul(A^T) \longrightarrow S^\perp = Nul(A^T)^\perp = R(A)$. Hence,*

$$\mathbb{R}^n = Nul(A^T) \oplus R(A)$$

Lemma (The Orthogonal Compliment of The Complement)

For any $S \subset \mathbb{R}^n$ one has: $(S^\perp)^\perp = S$

Fundamental Subspace for $Ax = b$

Corollary

Suppose $A \in \mathbb{R}^{mxn}$ *and* $\mathbf{b} \in \mathbb{R}^m$*, then there is* $\mathbf{x}$ *such that* $A\mathbf{x} = \mathbf{b}$ *or there is* $\mathbf{y} \in N(A^T)$ *where* $\langle \mathbf{y}, \mathbf{b} \rangle \neq \mathbf{0}$

udemy

Fundamental Subspace for $Ax = b$

Corollary

Suppose $A \in \mathbb{R}^{mxn}$ *and* $\mathbf{b} \in \mathbb{R}^m$*, then there is* $\mathbf{x}$ *such that* $A\mathbf{x} = \mathbf{b}$ *or there is* $\mathbf{y} \in N(A^T)$ *where* $\langle \mathbf{y}, \mathbf{b} \rangle \neq \mathbf{0}$

Remark: *In other words,* $A\mathbf{x} = \mathbf{b}$ *is consistent or there is a non-null* $\mathbf{y} \in Nul(A^T)$ *that is not orthogonal to* $\mathbf{b}$*.*

udemy

Example: Fundamental Subspace

$$\text{Given: } A = \begin{bmatrix} 1 & 1 & 2 \\ 0 & 1 & 1 \\ 1 & 3 & 4 \end{bmatrix} \rightarrow A^T = \begin{bmatrix} 1 & 0 & 1 \\ 1 & 1 & 3 \\ 2 & 1 & 4 \end{bmatrix}$$

Determine a basis for the Null Space of A^T and the Row Space of A:

Example: Fundamental Subspace

$$\text{Given: } A = \begin{bmatrix} 1 & 1 & 2 \\ 0 & 1 & 1 \\ 1 & 3 & 4 \end{bmatrix} \rightarrow A^T = \begin{bmatrix} 1 & 0 & 1 \\ 1 & 1 & 3 \\ 2 & 1 & 4 \end{bmatrix}$$

Determine a basis for the Null Space of A^T and the Row Space of A:

After Row Reducing A to RREF:

$$U = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 2 \\ 0 & 0 & 0 \end{bmatrix}$$

Example: Fundamental Subspace

$$\text{Given: } A = \begin{bmatrix} 1 & 1 & 2 \\ 0 & 1 & 1 \\ 1 & 3 & 4 \end{bmatrix} \rightarrow A^T = \begin{bmatrix} 1 & 0 & 1 \\ 1 & 1 & 3 \\ 2 & 1 & 4 \end{bmatrix}$$

Determine a basis for the Null Space of A^T and the Row Space of A:

After Row Reducing A to RREF:

$$U = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 2 \\ 0 & 0 & 0 \end{bmatrix}$$

It follows that $x_1 + x_3 = 0$ and $x_2 + 2x_3 = 0$. This implies that: $x_1 = -x_3$ and $x_2 = -2x_3$, taking $x_3 = \alpha$ then:

Example: Fundamental Subspace

$$\text{Given: } A = \begin{bmatrix} 1 & 1 & 2 \\ 0 & 1 & 1 \\ 1 & 3 & 4 \end{bmatrix} \rightarrow A^T = \begin{bmatrix} 1 & 0 & 1 \\ 1 & 1 & 3 \\ 2 & 1 & 4 \end{bmatrix}$$

Determine a basis for the Null Space of A^T and the Row Space of A:

After Row Reducing A to RREF:

$$U = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 2 \\ 0 & 0 & 0 \end{bmatrix}$$

It follows that $x_1 + x_3 = 0$ and $x_2 + 2x_3 = 0$. This implies that: $x_1 = -x_3$ and $x_2 = -2x_3$, taking $x_3 = \alpha$ then:

$$Nul(A^T) = Span\left\{ \begin{bmatrix} -1 \\ -2 \\ 1 \end{bmatrix} \right\} \ \& \ R(A) = Span\left\{ \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 2 \end{bmatrix} \right\}$$

Remark: $Rank(A) = Rank(A^T) = 2 = 3 - 1 = n - Nullity(A^T)$

General Solution of $A\mathbf{x} = \mathbf{b}$

Lemma (General Solution of $A\mathbf{x} = \mathbf{b}$)

Suppose $A \in \mathbb{R}^{mxn}$, $\mathbf{x} \in \mathbb{R}^n$, and $\mathbf{b} \in \mathbb{R}^m$, the General Solution of:

$$A\mathbf{x} = \mathbf{b} \; where: \; \mathbf{x} = \mathbf{x_0} + \mathbf{z} \; \ni: \begin{cases} \mathbf{x_0} \in R(A^T) \\ \mathbf{z} \in Nul(A) \end{cases}$$

General Solution of $A\mathbf{x} = \mathbf{b}$

Lemma (General Solution of $A\mathbf{x} = \mathbf{b}$)

Suppose $A \in \mathbb{R}^{mxn}$, $\mathbf{x} \in \mathbb{R}^n$, and $\mathbf{b} \in \mathbb{R}^m$, the General Solution of:

$$A\mathbf{x} = \mathbf{b} \ \textit{where}: \ \mathbf{x} = \mathbf{x_0} + \mathbf{z} \ \ni: \begin{cases} \mathbf{x_0} \in R(A^T) \\ \mathbf{z} \in Nul(A) \end{cases}$$

Remark: *Recalling that $N(A)^\perp = R(A^T)$ and $\mathbb{R}^n = Nul(A) + Nul(A)^\perp$*

udemy

Table of Contents

The Least Squares Problem

Definition (Least Squares Formulation)

Consider the over-determined system: $A\mathbf{x} = \mathbf{b}$ where $A \in \mathbb{R}^{mxn}, \mathbf{x} \in \mathbb{R}^n, \mathbf{b} \in \mathbb{R}^m$ and $m > n$.

Generally, the system is **inconsistent**, although we are interested in solving for the "closest" solution.

$$\textit{Take}: \ r(\mathbf{x}) = \mathbf{b} - A\mathbf{x} \sim \textit{Residual} \ni: \delta(b, \ A\mathbf{x}) = ||r(\mathbf{x})||_2 \sim \textit{Distance}$$

udemy

The Least Squares Problem

Definition (Least Squares Formulation)

Consider the over-determined system: $A\mathbf{x} = \mathbf{b}$ where $A \in \mathbb{R}^{mxn}, \mathbf{x} \in \mathbb{R}^n, \mathbf{b} \in \mathbb{R}^m$ and $m > n$.

Generally, the system is **inconsistent**, although we are interested in solving for the "closest" solution.

$$Take: \ r(\mathbf{x}) = \mathbf{b} - A\mathbf{x} \sim Residual \ni: \delta(b,\ A\mathbf{x}) = ||r(\mathbf{x})||_2 \sim Distance$$

Finding the closest solution is obtained by minimizing the Residual $r(\mathbf{x})$ wrt. $\mathbf{x}$:

$$Equivalently: \ Minimize: \ ||\mathbf{r}(\mathbf{x})||_2^2 \ for \ \mathbf{x} \in \mathbb{R}^n \ to \ obtain \ \hat{\mathbf{x}}$$

We will define the closest vector to $\mathbf{b}$ as $\mathbf{p} = A\hat{\mathbf{x}}$ the Projection Vector.

The Vector that Minimizes the Residual

Theorem (Projection Vector)

Suppose $S \subset \mathbb{R}^m$ *and for each* $\mathbf{b} \in \mathbb{R}^m$ *there is some* $\mathbf{p} \in S$ *that is "closest" to :*

$$||\mathbf{b} - \mathbf{y}||_2 > ||\mathbf{b} - \mathbf{p}||_2 \text{ for each } \mathbf{y} \in S$$

udemy

The Vector that Minimizes the Residual

Theorem (Projection Vector)

Suppose $S \subset \mathbb{R}^m$ *and for each* $\mathbf{b} \in \mathbb{R}^m$ *there is some* $\mathbf{p} \in S$ *that is "closest" to :*

$$||\mathbf{b} - \mathbf{y}||_2 > ||\mathbf{b} - \mathbf{p}||_2 \text{ for each } \mathbf{y} \in S$$

Remark: $\mathbf{p}$ *is the projection of* $\mathbf{b}$ *onto* S.

$\mathbf{p}$ *is "closest" to* $\mathbf{b} \longleftrightarrow \mathbf{b} - \mathbf{p} \perp S \longleftrightarrow \mathbf{b} - \mathbf{p} \in S^{\perp}$.

The Vector that Minimizes the Residual

Theorem (Projection Vector)

Suppose $S \subset \mathbb{R}^m$ and for each $\mathbf{b} \in \mathbb{R}^m$ there is some $\mathbf{p} \in S$ that is "closest" to :

$$||\mathbf{b} - \mathbf{y}||_2 > ||\mathbf{b} - \mathbf{p}||_2 \text{ for each } \mathbf{y} \in S$$

Remark: *$\mathbf{p}$ is the projection of $\mathbf{b}$ onto S.*

$\mathbf{p}$ *is "closest" to* $\mathbf{b} \longleftrightarrow \mathbf{b} - \mathbf{p} \perp S \longleftrightarrow \mathbf{b} - \mathbf{p} \in S^{\perp}$.

Proof.

Recall **The Direct Sum**: for each $\mathbf{b} \in \mathbf{R}^m$: $\mathbf{b} = \mathbf{p} + \mathbf{z}$ where $\mathbf{p} \in S$ and $\mathbf{z} \in S^{\perp}$ because $\mathbb{R}^m = S \oplus S^{\perp}$.

$\square$

The Vector that Minimizes the Residual

Theorem (Projection Vector)

Suppose $S \subset \mathbb{R}^m$ and for each $\mathbf{b} \in \mathbb{R}^m$ there is some $\mathbf{p} \in S$ that is "closest" to :

$$||\mathbf{b}-\mathbf{y}||_2 > ||\mathbf{b}-\mathbf{p}||_2 \text{ for each } \mathbf{y} \in S$$

Remark: *$\mathbf{p}$ is the projection of $\mathbf{b}$ onto S.*

$\mathbf{p}$ is "closest" to $\mathbf{b} \longleftrightarrow \mathbf{b}-\mathbf{p} \perp S \longleftrightarrow \mathbf{b}-\mathbf{p} \in S^{\perp}$.

Proof.

Recall **The Direct Sum**: for each $\mathbf{b} \in \mathbf{R}^m$: $\mathbf{b} = \mathbf{p} + \mathbf{z}$ where $\mathbf{p} \in S$ and $\mathbf{z} \in S^{\perp}$ because $\mathbb{R}^m = S \oplus S^{\perp}$.

$$\begin{aligned} ||\mathbf{b}-\mathbf{y}||_2^2 &= ||(\mathbf{b}-\mathbf{p}) + (\mathbf{p}-\mathbf{y})||_2^2 \\ &= ||\mathbf{b}-\mathbf{p}||_2^2 + ||\mathbf{p}-\mathbf{y}||_2^2 \\ &> ||\mathbf{b}-\mathbf{p}||_2^2 \ \textit{Provided } \mathbf{y} \neq \mathbf{p} \end{aligned}$$

Remark: The projection vector of **b** onto S is *unique*

$\square$

The Least Squares Solution

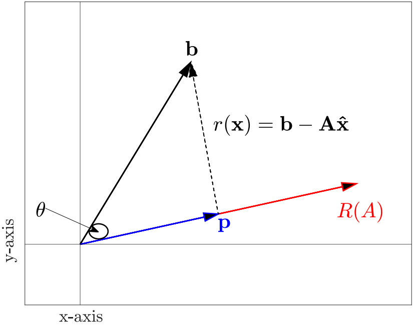
Finding the "closest" solution.

To minimize the distance between **b** and **p** or minimize: $r(\mathbf{x})$, we find the Projection of **b** onto $R(A)^{\perp}$, which is equivalent to solving for $Nul(A^T)$:

$$r(\hat{\mathbf{x}}) \in Nul(A^T)$$

Remark: the projection of **b** onto $R(A)$ is: $\mathbf{p} = A\hat{x}$

$\square$



udemy

The Normal Equations

Definition (Normal Equations for Least Squares Solution)

Suppose one solves the over-determined system $A\mathbf{x} = \mathbf{b}$ where $m > n$, $A \in \mathbb{R}^{mxn}$, $\mathbf{x} \in \mathbb{R}^n$, and $\mathbf{b} \in \mathbb{R}^m$.

Suppose the system is *inconsistent*, then one attempts to minimize the residual $||r(\mathbf{x})|| = ||\mathbf{b} - A\mathbf{x}||_2$ by solving for $Nul(A^T)$:

$$0 = A^T(r(\mathbf{x})) = A^T(\mathbf{b} - A\mathbf{x})$$

Or equivalently, by solving the **Normal Equations**:

$$A^T A\hat{\mathbf{x}} = A^T\mathbf{b}$$

udemy

The Normal Equations

Definition (Normal Equations for Least Squares Solution)

Suppose one solves the over-determined system $A\mathbf{x} = \mathbf{b}$ where $m > n$, $A \in \mathbb{R}^{mxn}$, $\mathbf{x} \in \mathbb{R}^n$, and $\mathbf{b} \in \mathbb{R}^m$.
Suppose the system is *inconsistent*, then one attempts to minimize the residual $||r(\mathbf{x})|| = ||\mathbf{b} - A\mathbf{x}||_2$ by solving for $Nul(A^T)$:

$$0 = A^T(r(\mathbf{x})) = A^T(\mathbf{b} - A\mathbf{x})$$

Or equivalently, by solving the **Normal Equations**:

$$A^T A\hat{\mathbf{x}} = A^T\mathbf{b}$$

Theorem (Unique Least Squares Solution)

Suppose one solves the **Inconsistent System of Equations***: $Ax = \mathbf{b}$ for $A \in \mathbb{R}^{mxn}$, $x \in \mathbf{R}^n$, and $\mathbf{b} \in \mathbf{R}^m$, provided A has* **full column rank***: $Rank(A) = n$, then $A^T A$ is nonsingular and there is a unique solution:*

$$\hat{\mathbf{x}} = (A^T A)^{-1} A^T\mathbf{b}$$

The Normal Equations

Definition (Normal Equations for Least Squares Solution)

Suppose one solves the over-determined system $A\mathbf{x} = \mathbf{b}$ where $m > n$, $A \in \mathbb{R}^{mxn}$, $\mathbf{x} \in \mathbb{R}^n$, and $\mathbf{b} \in \mathbb{R}^m$.

Suppose the system is *inconsistent*, then one attempts to minimize the residual $||r(\mathbf{x})|| = ||\mathbf{b} - A\mathbf{x}||_2$ by solving for $Nul(A^T)$:

$$0 = A^T(r(\mathbf{x})) = A^T(\mathbf{b} - A\mathbf{x})$$

Or equivalently, by solving the **Normal Equations**:

$$A^T A\hat{\mathbf{x}} = A^T\mathbf{b}$$

Theorem (Unique Least Squares Solution)

Suppose one solves the **Inconsistent System of Equations***: $Ax = \mathbf{b}$ for $A \in \mathbb{R}^{mxn}$, $x \in \mathbf{R}^n$, and $\mathbf{b} \in \mathbf{R}^m$, provided A has* **full column rank***: $Rank(A) = n$, then $A^T A$ is nonsingular and there is a unique solution:*

$$\hat{\mathbf{x}} = (A^T A)^{-1} A^T\mathbf{b}$$

Remark: *The projection of* $\mathbf{b}$ *onto* $R(A)$ *is:* $\mathbf{p} = A(A^T A)^{-1} A^T\mathbf{b}$.

Example: Least Squares Solution

1. Find the Solution of the System:

$$
\begin{aligned}
x_1 + x_2 &= 3 \\
-2x_1 + 3x_2 &= 1 \\
2x_1 - x_2 &= 2
\end{aligned}
$$

Example: Least Squares Solution

1 Find the Solution of the System:

$$
\begin{aligned}
x_1 + x_2 &= 3 \\
-2x_1 + 3x_2 &= 1 \\
2x_1 - x_2 &= 2
\end{aligned}
$$

$$
2 : [A \mid b] = \left[\begin{array}{cc|c} 1 & 1 & 3 \\ -2 & 3 & 1 \\ 2 & -1 & 2 \end{array}\right] \sim [U \mid \mathbf{b_u}] = \left[\begin{array}{cc|c} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{array}\right] \sim \textit{Inconsistent}
$$

udemy

Example: Least Squares Solution

1. Find the Solution of the System:

$$
\begin{aligned}
x_1 + x_2 &= 3 \\
-2x_1 + 3x_2 &= 1 \\
2x_1 - x_2 &= 2
\end{aligned}
$$

2. $: [A \mid b] = \left[\begin{array}{cc|c} 1 & 1 & 3 \\ -2 & 3 & 1 \\ 2 & -1 & 2 \end{array}\right] \sim [U \mid \mathbf{b_u}] = \left[\begin{array}{cc|c} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{array}\right] \sim \textit{Inconsistent}$

3. Since $Rank(A) = Rank(U) = 2$ and $A \in \mathbb{R}^{3x2}$, then A has Full Column Rank, $A^T A$ is invertible, and the Unique Solution is:

udemy

Example: Least Squares Solution

1 Find the Solution of the System:

$$\begin{aligned} x_1 + x_2 &= 3 \\ -2x_1 + 3x_2 &= 1 \\ 2x_1 - x_2 &= 2 \end{aligned}$$

$$2 : [A \mid b] = \left[\begin{array}{cc|c} 1 & 1 & 3 \\ -2 & 3 & 1 \\ 2 & -1 & 2 \end{array}\right] \sim [U \mid \mathbf{b_u}] = \left[\begin{array}{cc|c} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{array}\right] \sim \textit{Inconsistent}$$

3 Since $Rank(A) = Rank(U) = 2$ and $A \in \mathbb{R}^{3x2}$, then A has Full Column Rank, $A^T A$ is invertible, and the Unique Solution is:

$$\hat{\mathbf{x}} = (A^T A)^{-1} A^T \mathbf{b} = \left(\begin{bmatrix} 1 & -2 & 2 \\ 1 & 3 & -1 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ -2 & 3 \\ 2 & -1 \end{bmatrix} \right)^{-1} \begin{bmatrix} 1 & -2 & 2 \\ 1 & 3 & -1 \end{bmatrix} \begin{bmatrix} 3 \\ 1 \\ 2 \end{bmatrix} = \begin{bmatrix} 1.66 \\ 1.42 \end{bmatrix}$$

udemy

Choosing the Best Candidate

1. Suppose we wish to choose the **Best** candidate in the employment pool, the most "Ideal" candidate does not exist. Consider the following Doctor candidates for a Faculty position from the following:
{Gauss, Ipsen, O'Leary, Taussky}.

Choosing the Best Candidate

1. Suppose we wish to choose the **Best** candidate in the employment pool, the most "Ideal" candidate does not exist. Consider the following Doctor candidates for a Faculty position from the following:
{Gauss, Ipsen, O'Leary, Taussky}.
2. Furthermore, suppose the committee hires based on Research at rate 0.40, Teaching at 0.40, and Professional Activities at 0.20.

udemy

Choosing the Best Candidate

1. Suppose we wish to choose the **Best** candidate in the employment pool, the most "Ideal" candidate does not exist. Consider the following Doctor candidates for a Faculty position from the following: {Gauss, Ipsen, O'Leary, Taussky}.
2. Furthermore, suppose the committee hires based on Research at rate 0.40, Teaching at 0.40, and Professional Activities at 0.20.
3. Lastly, assume the *institution* has preference of **Quality over Quantity** of *Research* and chooses at rate at 0.60 for Quality and 0.40 for Quantity.

Choosing the Best Candidate

(1) Suppose we wish to choose the **Best** candidate in the employment pool, the most "Ideal" candidate does not exist. Consider the following Doctor candidates for a Faculty position from the following:
$\{$Gauss, Ipsen, O'Leary, Taussky$\}$.

(2) Furthermore, suppose the committee hires based on Research at rate 0.40, Teaching at 0.40, and Professional Activities at 0.20.

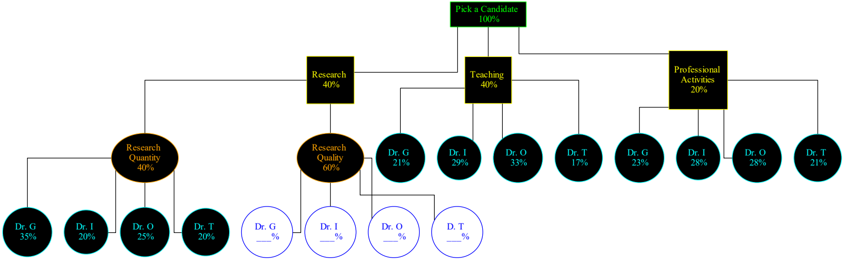
(3) Lastly, assume the *institution* has preference of **Quality over Quantity** of *Research* and chooses at rate at 0.60 for Quality and 0.40 for Quantity.

(4) Suppose the quantity of the research is determined by page number:

Candidate	Page Count	Relative Quantity	Quality
Dr. G	700	$700/2000 = 0.35$	?
Dr. I	400	$400/2000 = 0.20$	?
Dr. O	500	$500/2000 = 0.25$	?
Dr. T	400	$400/2000 = 0.20$	?
	2000	$2000/2000 = 1.00$	?

udemy

Choosing the Best Candidate



udemy

Reference: Leon, 2014 [2]

Application: Choosing the Best Candidate

Suppose the committee chooses the quality of publications with the following relation:

$$w_i = \beta w_j \sim \beta \in (1, 2]$$

After comparing each of the candidates closely to obtain:

$$\begin{array}{ll} 1: w_1 = 1.75 w_2 & 2: w_1 = 1.5 w_3 \\ 3: w_1 = 1.25 w_4 & 4: w_2 = 0.75 w_3 \\ 5: w_2 = 0.50 w_4 & 6: w_3 = 0.75 w_4 \end{array}$$

Subject to the following additional condition:

$$w_1 + w_2 + w_3 + w_4 = 1 \longleftrightarrow w_4 = 1 - (w_1 + w_2 + w_3)$$

Then, the system can be reduced to solving an over-determined 6x4 system of equations.

udemy

Application: Choosing the Best Candidate

One obtains the following system of equations:

$$\begin{aligned}
1w_1 - 1.75w_2 + 0w_3 &= 0 \\
1w_1 + 0w_2 - 1.5w_3 &= 0 \\
2.25w_1 + 1.25w_2 + 1.25w_3 &= 1.25 \\
0w_1 + 1w_2 - 0.75w_3 &= 0 \\
0.50w_1 + 1.5w_2 + 0.50w_3 &= 0.50 \\
0.75w_1 + 0.75w_2 + 1.75w_3 &= 0.75
\end{aligned}$$

After row reduction, one can check that A contains a row of zeros and a nonzero b_i on the right hand side, implying the system is *inconsistent*.

Since $Rank(A) = 3$, then there is a Unique Least Squares solution: $\hat{\mathbf{w}}$.

udemy

Application: Choosing the Best Candidate

After substitution of the equality: $w_4 = 1 - (w_1 + w_2 + w_3)$, one obtains:

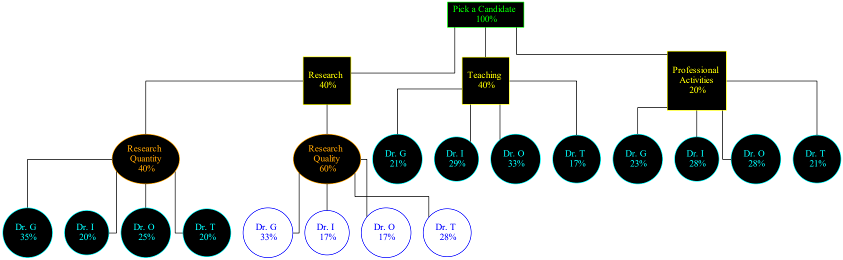
$$A = \begin{bmatrix} 1 & -1.75 & 0 \\ 1 & 0 & -1.5 \\ 2.25 & 1.25 & 1.25 \\ 0 & 1 & -0.75 \\ 0.50 & 1.50 & 0.50 \\ 0.75 & 0.75 & 1.75 \end{bmatrix} \quad \mathbf{b} = \begin{bmatrix} 0 \\ 0 \\ 1.2500 \\ 0 \\ 0.5000 \\ 0.7500 \end{bmatrix}$$

Since $Rank(A) = 3$, then $A^T A$ is nonsingular and the solution is:

$$\hat{\mathbf{w}} = (A^T A)^{-1} A^T \mathbf{b} = \begin{bmatrix} 0.33 \\ 0.17 \\ 0.22 \end{bmatrix} \text{ and } w_4 = 1 - (0.33 + 0.17 + 0.22) = 0.28$$

udemy

Choosing the Best Candidate



Application: Choosing the Best Candidate

Proof.

The Rating Vector r is determined by taking the weighted sum over all criteria and sub-criteria:

$$\mathbf{r} = 0.40\left\{0.40\begin{bmatrix}0.35\\0.20\\0.25\\0.20\end{bmatrix} + 0.60\begin{bmatrix}0.33\\0.17\\0.22\\0.28\end{bmatrix}\right\} + 0.40\left\{\begin{bmatrix}0.21\\0.29\\0.33\\0.17\end{bmatrix}\right\} + 0.20\left\{\begin{bmatrix}0.23\\0.28\\0.28\\0.21\end{bmatrix}\right\} = \begin{bmatrix}0.26\\0.25\\0.28\\0.21\end{bmatrix}$$

Dr. O is the best candidate!

$\square$

udemy

Table of Contents

The Inner Product Space

Definition (An Inner Product: $\langle \cdot \rangle$)

Suppose $V \sim \textit{Vector Space}$ and for each $\mathbf{x}, \mathbf{y} \in V$ one can derive a Real Number using the Inner Product operation $\langle \cdot \rangle$ provided the following conditions hold:

1. $\langle \mathbf{x}, \mathbf{x} \rangle \geq 0 \textit{ for each } \mathbf{x} \in V$

udemy

The Inner Product Space

Definition (An Inner Product: $\langle \cdot \rangle$)

Suppose $V \sim$ *Vector Space* and for each $\mathbf{x}, \mathbf{y} \in V$ one can derive a Real Number using the Inner Product operation $\langle \cdot \rangle$ provided the following conditions hold:

1. $\langle \mathbf{x}, \mathbf{x} \rangle \geq 0$ *for each* $\mathbf{x} \in V$
2. $\langle \mathbf{x}, \mathbf{y} \rangle = \langle \mathbf{y}, \mathbf{x} \rangle$ *for each* $\mathbf{x}, \mathbf{y} \in V$

udemy

The Inner Product Space

Definition (An Inner Product: $\langle \cdot \rangle$)

Suppose $V \sim$ *Vector Space* and for each $\mathbf{x}, \mathbf{y} \in V$ one can derive a Real Number using the Inner Product operation $\langle \cdot \rangle$ provided the following conditions hold:

1. $\langle \mathbf{x}, \mathbf{x} \rangle \geq 0$ *for each* $\mathbf{x} \in V$
2. $\langle \mathbf{x}, \mathbf{y} \rangle = \langle \mathbf{y}, \mathbf{x} \rangle$ *for each* $\mathbf{x}, \mathbf{y} \in V$
3. $\langle \alpha \mathbf{x} + \beta \mathbf{y}, \mathbf{z} \rangle = \alpha \langle \mathbf{x}, \mathbf{z} \rangle + \beta \langle \mathbf{y}, \mathbf{z} \rangle$ *for each* $\mathbf{x}, \mathbf{y}, \mathbf{z} \in V$ *and* $\alpha, \beta \in \mathbb{R}$

udemy

The Inner Product Space

Definition (An Inner Product: $\langle \cdot \rangle$)

Suppose $V \sim$ *Vector Space* and for each $\mathbf{x}, \mathbf{y} \in V$ one can derive a Real Number using the Inner Product operation $\langle \cdot \rangle$ provided the following conditions hold:

1. $\langle \mathbf{x}, \mathbf{x} \rangle \geq 0$ *for each* $\mathbf{x} \in V$
2. $\langle \mathbf{x}, \mathbf{y} \rangle = \langle \mathbf{y}, \mathbf{x} \rangle$ *for each* $\mathbf{x}, \mathbf{y} \in V$
3. $\langle \alpha\mathbf{x} + \beta\mathbf{y}, \mathbf{z} \rangle = \alpha\langle \mathbf{x}, \mathbf{z} \rangle + \beta\langle \mathbf{y}, \mathbf{z} \rangle$ *for each* $\mathbf{x}, \mathbf{y}, \mathbf{z} \in V$ *and* $\alpha, \beta \in \mathbb{R}$

Remark: If $V = \mathbb{R}^n$, then $\langle \mathbf{x}, \mathbf{y} \rangle = \mathbf{x}^T\mathbf{y}$.

Inner Products

Definition (Types of Inner Products)

1. (1) The Weighted Inner Product on $\mathbb{R}^n$:

$$\langle \mathbf{x}, \mathbf{y} \rangle = \sum_{j=1}^{n} x_j y_j w_j \ \textit{for} \ w_j \in \mathbb{R}$$

Inner Products

Definition (Types of Inner Products)

1. ① The Weighted Inner Product on $\mathbb{R}^n$:

$$\langle \mathbf{x}, \mathbf{y} \rangle = \sum_{j=1}^{n} x_j y_j w_j \textit{ for } w_j \in \mathbb{R}$$

2. ② The Inner Product on $\mathbb{R}^{mxn}$:

$$\langle A, B \rangle = \sum_{i=1}^{m} \sum_{j=1}^{n} a_{ij} * b_{ij}$$

Inner Products

Definition (Types of Inner Products)

1. ① The Weighted Inner Product on $\mathbb{R}^n$:

$$\langle \mathbf{x}, \mathbf{y} \rangle = \sum_{j=1}^{n} x_j y_j w_j \; \textit{for} \; w_j \in \mathbb{R}$$

2. ② The Inner Product on $\mathbb{R}^{mxn}$:

$$\langle A, B \rangle = \sum_{i=1}^{m} \sum_{j=1}^{n} a_{ij} * b_{ij}$$

3. ③ The Inner Product on P_n by summing the corresponding coefficients:

$$\langle p(\mathbf{x}), q(\mathbf{x}) \rangle = \sum_{j=1}^{n} p(x_i) * q(x_i)$$

Inner Products

Definition (Types of Inner Products)

1. (1) The Weighted Inner Product on $\mathbb{R}^n$:

$$\langle \mathbf{x}, \mathbf{y} \rangle = \sum_{j=1}^{n} x_j y_j w_j \ \ \textit{for} \ \ w_j \in \mathbb{R}$$

2. (2) The Inner Product on $\mathbb{R}^{mxn}$:

$$\langle A, B \rangle = \sum_{i=1}^{m} \sum_{j=1}^{n} a_{ij} * b_{ij}$$

3. (3) The Inner Product on P_n by summing the corresponding coefficients:

$$\langle p(\mathbf{x}), q(\mathbf{x}) \rangle = \sum_{j=1}^{n} p(x_i) * q(x_i)$$

4. (4) The Inner Product on $C[a, b]$: $\langle f, g \rangle = w \int_a^b f(x) * g(x) dx \ \sim \ w \in \mathbb{R}$

Examples: $\mathbb{R}^{mxn}$

Suppose $A, B \in \mathbb{R}^{3x3}$ where:

$$A = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 4 & 6 \\ 3 & 6 & 9 \end{bmatrix} \quad B = \begin{bmatrix} 2 & 4 & 6 \\ 0 & 2 & 4 \\ 1 & 3 & 5 \end{bmatrix}$$

Examples: $\mathbb{R}^{mxn}$

Suppose $A, B \in \mathbb{R}^{3x3}$ where:

$$A = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 4 & 6 \\ 3 & 6 & 9 \end{bmatrix} \quad B = \begin{bmatrix} 2 & 4 & 6 \\ 0 & 2 & 4 \\ 1 & 3 & 5 \end{bmatrix}$$

Then:

$$\langle A, B \rangle = 1(2) + 2(4) + 3(6) + 2(0) + 4(2) + 6(4) + 3(1) + 6(3) + 9(5) = 126$$

Since $\langle A, B \rangle \neq 0$ then A and B are **not** Orthogonal Matrices.

Example: Functions

1. Consider $C[-\pi, \pi]$ and let $f(\mathbf{x}) = Cos(x), g(\mathbf{x}) = Sin(x)$, and $w = 1/\pi$, show that f and g are Orthogonal Functions:

udemy

Example: Functions

1. Consider $C[-\pi,\pi]$ and let $f(\mathbf{x}) = Cos(x), g(\mathbf{x}) = Sin(x)$, and $w = 1/\pi$, show that f and g are Orthogonal Functions:

$$\langle f, g\rangle = \frac{1}{\pi}\int_{-\pi}^{\pi} Cos(x)Sin(x)dx = \frac{1}{\pi}\int_{-\pi}^{\pi} \frac{Sin(2x)}{2}dx = 0$$

Example: Functions

1. Consider $C[-\pi, \pi]$ and let $f(\mathbf{x}) = Cos(x), g(\mathbf{x}) = Sin(x)$, and $w = 1/\pi$, show that f and g are Orthogonal Functions:

$$\langle f, g \rangle = \frac{1}{\pi} \int_{-\pi}^{\pi} Cos(x)Sin(x)dx = \frac{1}{\pi} \int_{-\pi}^{\pi} \frac{Sin(2x)}{2} dx = 0$$

Remark: One can show that $\langle Sin(x), Sin(x) \rangle = 1 = \langle Cos(x), Cos(x) \rangle$ implying that $Cos(x)$ and $Sin(x)$ are unit vectors with respect to $\langle \cdot \rangle$.

udemy

Properties of The Inner Product

Definition (The Norm of an Inner Product)

Suppose $\mathbf{v} \in V$ where $V \sim$ *Inner Product Space*, then the **Norm** of $\mathbf{v}$ is:

$$||\mathbf{v}|| = \sqrt{\langle \mathbf{v}, \mathbf{v} \rangle}$$

udemy

Properties of The Inner Product

Definition (The Norm of an Inner Product)

Suppose $\mathbf{v} \in V$ where $V \sim$ *Inner Product Space*, then the **Norm** of $\mathbf{v}$ is:

$$||\mathbf{v}|| = \sqrt{\langle \mathbf{v}, \mathbf{v} \rangle}$$

Remark: If $V = \mathbb{R}^n$ and $\langle \mathbf{x}, \mathbf{y} \rangle = \mathbf{x}^T \mathbf{y}$, $||\mathbf{x}||_2^2$ is the Magnitude of $\mathbf{x}$.
Additionally, $\langle \mathbf{x}, \mathbf{y} \rangle = 0 \longleftrightarrow \mathbf{x}$ and $\mathbf{y}$ are Orthogonal.

udemy

Properties of The Inner Product

Definition (The Norm of an Inner Product)

Suppose $\mathbf{v} \in V$ where $V \sim \textit{Inner Product Space}$, then the **Norm** of $\mathbf{v}$ is:

$$||\mathbf{v}|| = \sqrt{\langle \mathbf{v}, \mathbf{v} \rangle}$$

Remark: If $V = \mathbb{R}^n$ and $\langle \mathbf{x}, \mathbf{y} \rangle = \mathbf{x}^T \mathbf{y}$, $||\mathbf{x}||_2^2$ is the Magnitude of $\mathbf{x}$.
Additionally, $\langle \mathbf{x}, \mathbf{y} \rangle = 0 \longleftrightarrow \mathbf{x}$ and $\mathbf{y}$ are Orthogonal.

Theorem (Pythagorean Law)

Suppose $\mathbf{u}, \mathbf{v} \in V$ *where* $V \sim \textit{Inner Product Space}$ *if* $\langle \mathbf{u}, \mathbf{v} \rangle = 0$, *then:*

$$||u + v||^2 = ||u||^2 + ||v||^2$$

The Projection: Revisited

Definition (The Projection of **u** onto **v**)

Suppose $V \sim$ *Inner Product Space* and $\mathbf{u}, \mathbf{v} \in V$ then the Scalar Projection of **u** onto **v** is:

$$\alpha = \frac{\langle \mathbf{u}, \mathbf{v} \rangle}{||\mathbf{v}||}$$

udemy

The Projection: Revisited

Definition (The Projection of **u** onto **v**)

Suppose $V \sim$ *Inner Product Space* and $\mathbf{u}, \mathbf{v} \in V$ then the Scalar Projection of **u** onto **v** is:

$$\alpha = \frac{\langle \mathbf{u}, \mathbf{v} \rangle}{||\mathbf{v}||}$$

The Vector Projection of **u** onto **v** is

$$\mathbf{p} = \alpha \left(\frac{v}{||\mathbf{v}||} \right)$$

udemy

The Projection: Revisited

Definition (The Projection of **u** onto **v**)

Suppose $V \sim$ *Inner Product Space* and $\mathbf{u}, \mathbf{v} \in V$ then the Scalar Projection of **u** onto **v** is:

$$\alpha = \frac{\langle \mathbf{u}, \mathbf{v} \rangle}{||\mathbf{v}||}$$

The Vector Projection of **u** onto **v** is

$$\mathbf{p} = \alpha \left(\frac{v}{||\mathbf{v}||} \right)$$

Remarks: 1 $\mathbf{u} = \mathbf{p} \longleftrightarrow \mathbf{u} \propto \mathbf{v}$

udemy

The Projection: Revisited

Definition (The Projection of **u** onto **v**)

Suppose $V \sim$ *Inner Product Space* and $\mathbf{u}, \mathbf{v} \in V$ then the Scalar Projection of **u** onto **v** is:

$$\alpha = \frac{\langle \mathbf{u}, \mathbf{v} \rangle}{||\mathbf{v}||}$$

The Vector Projection of **u** onto **v** is

$$\mathbf{p} = \alpha \left(\frac{v}{||\mathbf{v}||} \right)$$

Remarks: 1 $\mathbf{u} = \mathbf{p} \longleftrightarrow \mathbf{u} \propto \mathbf{v}$
2 $\mathbf{u} - \mathbf{p} \perp \mathbf{p}$

udemy

The Projection: Revisited

Definition (The Projection of $\mathbf{u}$ onto $\mathbf{v}$)

Suppose $V \sim$ *Inner Product Space* and $\mathbf{u}, \mathbf{v} \in V$ then the Scalar Projection of $\mathbf{u}$ onto $\mathbf{v}$ is:

$$\alpha = \frac{\langle \mathbf{u}, \mathbf{v} \rangle}{||\mathbf{v}||}$$

The Vector Projection of $\mathbf{u}$ onto $\mathbf{v}$ is

$$\mathbf{p} = \alpha \left(\frac{v}{||\mathbf{v}||} \right)$$

Remarks: 1 $\mathbf{u} = \mathbf{p} \longleftrightarrow \mathbf{u} \propto \mathbf{v}$

2 $\mathbf{u} - \mathbf{p} \perp \mathbf{p}$

Lemma (Cauchy-Schwartz Inequality)

Suppose $V \sim$ *Inner Product Space, then for each* $\mathbf{u}, \mathbf{v} \in V$*:*

$$|\langle \mathbf{u}, \mathbf{v} \rangle| \leq ||\mathbf{u}|| * ||\mathbf{v}|| \longleftrightarrow -1 \leq \frac{|\langle \mathbf{u}, \mathbf{v} \rangle|}{||\mathbf{u}|| * ||\mathbf{v}||} \leq 1$$

Normed Linear Space

Definition (The Normed Linear Space)

Suppose $V \sim$ Vector Space, then $V \sim$ **Normed Linear Space** if there is $||\mathbf{v}|| \in \mathbb{R}$ for each $\mathbf{v} \in V$ such that:

1. $||\mathbf{v}|| \geq 0$ *with equality* $\longleftrightarrow \mathbf{v} = 0$
2. $||\alpha\mathbf{v}|| = |\alpha| * ||\mathbf{v}||$ *for each* $\alpha \in \mathbb{R}$
3. $||\mathbf{v} + \mathbf{w}|| \leq ||\mathbf{v}|| + ||\mathbf{w}||$

Normed Linear Space

Definition (The Normed Linear Space)

Suppose $V \sim$ Vector Space, then $V \sim$ **Normed Linear Space** if there is $||\mathbf{v}|| \in \mathbb{R}$ for each $\mathbf{v} \in V$ such that:

1. $||\mathbf{v}|| \geq 0$ *with equality* $\longleftrightarrow \mathbf{v} = 0$
2. $||\alpha \mathbf{v}|| = |\alpha| * ||\mathbf{v}||$ *for each* $\alpha \in \mathbb{R}$
3. $||\mathbf{v} + \mathbf{w}|| \leq ||\mathbf{v}|| + ||\mathbf{w}||$

Theorem (Inner Product to Norm)

Suppose $V \sim$ *Inner Product Space, then a* **Norm** *can be defined as:*

$$||\mathbf{v}|| = \sqrt{\langle \mathbf{v}, \mathbf{v} \rangle}$$

Remark: $||\mathbf{v}|| = 1 \longleftrightarrow \mathbf{v} \sim$ *Unit Vector*

Different Types of Norms in $\mathbb{R}^n$

Lemma (Types of $|| \cdot ||$ in $\mathbb{R}^n$)

Consider the **Normed Linear Space***:* $\mathbb{R}^n$*. The following are Norms:*

1. $||\mathbf{x}||_1 = \sum_{j=1}^{n} |x_j| \longleftrightarrow L_1$ *Norm*
2. $||\mathbf{x}||_2 = \sqrt{\sum_{j=1}^{n} (x_j)^2} \longleftrightarrow L_2$ *Norm*
3. $||\mathbf{x}||_p = \left[\sum_{j=1}^{n} |x_j|^p \right]^{1/p} \longleftrightarrow L_p$ *Norm*
4. $||\mathbf{x}||_\infty = \max_{1 \le j \le n} |x_j| \longleftrightarrow L_\infty$ *Norm*

Definition (Distance in Normed Linear Space)

Suppose $V \sim$ **Normed Linear Space** with defined norm $||\mathbf{v}||$ for each $\mathbf{v} \in V$ and $\mathbf{x}, \mathbf{y} \in V$ then the Distance between $\mathbf{x}$ and $\mathbf{y}$ is:

$$\delta(\mathbf{x}, \mathbf{y}) = ||\mathbf{x} - \mathbf{y}||$$

Reference: Leon, 2014 [2]

Example: Norms

Suppose $\mathbf{v} = (1, -2, 3)^T$, compute the following norms:

1. $||\mathbf{v}||_1 = 1 + 2 + 3 = 6$
2. $||\mathbf{v}||_2 = \sqrt{1 + 4 + 9} = \sqrt{14}$
3. $||\mathbf{v}||_5 = \left[1 + |-2|^5 + 3^5\right]^{1/5} = 3.0774$
4. $||\mathbf{v}||_\infty = \max\limits_{j=\overline{1,3}}\{1, 2, 3\} = 3$

udemy

Table of Contents

Orthogonal & Orthonormal Sets of Vectors

Definition (An Orthogonal Set of Vectors)

Suppose $V \sim$ Inner Product Space and $S = \{\mathbf{v_1}, \mathbf{v_2}, \ldots, \mathbf{v_j}, \ldots, \mathbf{v_n}\}$ is a collection of vectors: $\mathbf{v_j} \in V$.

$$\langle \mathbf{v_i}, \mathbf{v_j} \rangle = 0 \ \textit{When } i \neq j \longleftrightarrow S \sim \textit{Orthogonal Set of Vectors}$$

udemy

Orthogonal & Orthonormal Sets of Vectors

Definition (An Orthogonal Set of Vectors)

Suppose $V \sim$ Inner Product Space and $S = \{\mathbf{v_1}, \mathbf{v_2}, \ldots, \mathbf{v_j}, \ldots, \mathbf{v_n}\}$ is a collection of vectors: $\mathbf{v_j} \in V$.

$$\langle \mathbf{v_i}, \mathbf{v_j} \rangle = 0 \ \textit{When } i \neq j \longleftrightarrow S \sim \textit{Orthogonal Set of Vectors}$$

Definition (An Orthonormal Set of Vectors)

Given $S = \{\mathbf{v_1}, \mathbf{v_2}, \ldots, \mathbf{v_j}, \ldots, \mathbf{v_n}\}$ is an **Orthogonal** set of vectors on Inner Product Space V, then:

$$||\mathbf{v_j}|| = 1 \longleftrightarrow S \sim \textit{Orthonormal Set of Vectors}$$

Orthogonal & Orthonormal Sets of Vectors

Definition (An Orthogonal Set of Vectors)

Suppose $V \sim$ Inner Product Space and $S = \{\mathbf{v_1}, \mathbf{v_2}, \ldots, \mathbf{v_j}, \ldots, \mathbf{v_n}\}$ is a collection of vectors: $\mathbf{v_j} \in V$.

$$\langle \mathbf{v_i}, \mathbf{v_j} \rangle = 0 \textit{ When } i \neq j \longleftrightarrow S \sim \textit{Orthogonal Set of Vectors}$$

Definition (An Orthonormal Set of Vectors)

Given $S = \{\mathbf{v_1}, \mathbf{v_2}, \ldots, \mathbf{v_j}, \ldots, \mathbf{v_n}\}$ is an **Orthogonal** set of vectors on Inner Product Space V, then:

$$||\mathbf{v_j}|| = 1 \longleftrightarrow S \sim \textit{Orthonormal Set of Vectors}$$

Remark: In other words, an Orthonormal Set of Vectors is an Orthogonal set of unit vectors!

Example: Orthogonal Set in $\mathbb{R}^3$

$$\text{Suppose we are given } S = \left\{ \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 2 \\ 1 \\ -3 \end{bmatrix}, \begin{bmatrix} 4 \\ -5 \\ 1 \end{bmatrix} \right\} \triangleq \{\mathbf{x_1}, \mathbf{x_2}, \mathbf{x_3}\}$$

$$\text{Show that } S \sim \text{Orthogonal Set of Vectors under } \langle \mathbf{x}, \mathbf{y} \rangle = \mathbf{x}^T \mathbf{y}$$

Example: Orthogonal Set in $\mathbb{R}^3$

$$\text{Suppose we are given } S = \left\{ \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 2 \\ 1 \\ -3 \end{bmatrix}, \begin{bmatrix} 4 \\ -5 \\ 1 \end{bmatrix} \right\} \triangleq \{\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3\}$$

$$\text{Show that } S \sim \text{Orthogonal Set of Vectors under } \langle \mathbf{x}, \mathbf{y} \rangle = \mathbf{x}^T \mathbf{y}$$

Proof.

$$\langle \mathbf{x}_1, \mathbf{x}_2 \rangle = 1(2) + 1(1) + 1(-3) = 0$$
$$\langle \mathbf{x}_1, \mathbf{x}_3 \rangle = 1(4) + 1(-5) + 1(1) = 0$$
$$\langle \mathbf{x}_2, \mathbf{x}_3 \rangle = 2(4) + 1(-5) + (-3)1 = 0$$

$\square$

udemy

Orthogonal Sets of Vectors are Linearly Independent

Theorem (Linear Independence of an Orthogonal Collection of Vectors)

Suppose $S = \{\mathbf{v_1}, \mathbf{v_2}, \ldots, \mathbf{v_j}, \ldots, \mathbf{v_n}\} \sim$ *Orthogonal Collection where* $\langle \mathbf{v_i}, \mathbf{v_j} \rangle = 0$ *when* $i \neq j$*, then* S *is Linearly Independent.*

Orthogonal Sets of Vectors are Linearly Independent

Theorem (Linear Independence of an Orthogonal Collection of Vectors)

Suppose $S = \{\mathbf{v_1}, \mathbf{v_2}, \ldots, \mathbf{v_j}, \ldots, \mathbf{v_n}\} \sim$ *Orthogonal Collection where* $\langle \mathbf{v_i}, \mathbf{v_j} \rangle = 0$ *when* $i \neq j$, *then* S *is Linearly Independent.*

Proof.

Suppose $V \sim$ Inner Product Space and let S be an Orthogonal Collection of non-null vectors aforementioned. Consider:

$$\sum_{i=1}^{n} c_i \mathbf{v_i} = 0 \sim \textit{For some } c_i \in \mathbb{R}$$

□

Orthogonal Sets of Vectors are Linearly Independent

Theorem (Linear Independence of an Orthogonal Collection of Vectors)

Suppose $S = \{\mathbf{v_1}, \mathbf{v_2}, \ldots, \mathbf{v_j}, \ldots, \mathbf{v_n}\} \sim$ *Orthogonal Collection where* $\langle \mathbf{v_i}, \mathbf{v_j} \rangle = 0$ *when* $i \neq j$*, then* S *is Linearly Independent.*

Proof.

Suppose $V \sim$ Inner Product Space and let S be an Orthogonal Collection of non-null vectors aforementioned. Consider:

$$\sum_{i=1}^{n} c_i \mathbf{v_i} = 0 \sim \textit{For some } c_i \in \mathbb{R}$$

Taking the inner product with one $\mathbf{v_j} \in S$ gives:

$$c_j * \langle \mathbf{v_j}, \mathbf{v_j} \rangle = 0 \rightarrow c_j = 0$$

$\square$

Orthogonal Sets of Vectors are Linearly Independent

Theorem (Linear Independence of an Orthogonal Collection of Vectors)

Suppose $S = \{\mathbf{v_1}, \mathbf{v_2}, \ldots, \mathbf{v_j}, \ldots, \mathbf{v_n}\} \sim$ *Orthogonal Collection where* $\langle \mathbf{v_i}, \mathbf{v_j} \rangle = 0$ *when* $i \neq j$*, then* S *is Linearly Independent.*

Proof.

Suppose $V \sim$ Inner Product Space and let S be an Orthogonal Collection of non-null vectors aforementioned. Consider:

$$\sum_{i=1}^{n} c_i \mathbf{v_i} = 0 \sim \textit{For some } c_i \in \mathbb{R}$$

Taking the inner product with one $\mathbf{v_j} \in S$ gives:

$$c_j * \langle \mathbf{v_j}, \mathbf{v_j} \rangle = 0 \rightarrow c_j = 0$$

Since $\mathbf{v_j}$ was chosen arbitrary, then $c_j = 0$ for each $j = \overline{1, n} \rightarrow S \sim \textit{Linearly Independent}$

□

Reference: Leon, 2014 [2]

Example: Orthonormal

$$\text{Suppose we are given } S = \left\{ \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 2 \\ 1 \\ -3 \end{bmatrix}, \begin{bmatrix} 4 \\ -5 \\ 1 \end{bmatrix} \right\} \triangleq \{\mathbf{x_1}, \mathbf{x_2}, \mathbf{x_3}\}$$

Determine an Orthonormal Collection of vectors on $\mathbb{R}^3$

Example: Orthonormal

Suppose we are given $S = \left\{ \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 2 \\ 1 \\ -3 \end{bmatrix}, \begin{bmatrix} 4 \\ -5 \\ 1 \end{bmatrix} \right\} \triangleq \{\mathbf{x_1}, \mathbf{x_2}, \mathbf{x_3}\}$

Determine an Orthonormal Collection of vectors on $\mathbb{R}^3$

Proof.

Since $S \sim$ Orthogonal Collection of Vectors, we only need to divide each $\mathbf{x_j}$ by its norm:

$$\mathbf{u_1} = \frac{\mathbf{x_1}}{||\mathbf{x_1}||} = \frac{1}{\sqrt{3}} * \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$$

$$\mathbf{u_2} = \frac{\mathbf{x_2}}{||\mathbf{x_2}} = \frac{1}{\sqrt{14}} * \begin{bmatrix} 2 \\ 1 \\ -3 \end{bmatrix}$$

$$\mathbf{u_3} = \frac{\mathbf{x_3}}{||\mathbf{x_3}||} = \frac{1}{\sqrt{42}} * \begin{bmatrix} 4 \\ -5 \\ 1 \end{bmatrix}$$

Then: $S_u = \{\mathbf{u_1}, \mathbf{u_2}, \mathbf{u_3}\} \sim$ Orthonormal set of Vectors.

□

Orthonormal Set on $C[\pi, \pi]$

Theorem (Trigonometric Orthogonal Functions wrt $\langle \cdot \rangle$)

Given the Function Space: $C[-\pi, \pi]$ *with Inner Product:*

$$\langle f, g \rangle = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) * g(x) dx$$

The following will hold:

udemy

Orthonormal Set on $C[\pi, \pi]$

Theorem (Trigonometric Orthogonal Functions wrt $\langle \cdot \rangle$)

Given the Function Space: $C[-\pi, \pi]$ with Inner Product:

$$\langle f, g \rangle = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) * g(x) dx$$

The following will hold:

$$1 : \langle 1, Cos(mx) \rangle = 0 \ \forall \ m \in \mathbb{N}$$

Orthonormal Set on $C[\pi, \pi]$

Theorem (Trigonometric Orthogonal Functions wrt $\langle \cdot \rangle$)

Given the Function Space: $C[-\pi, \pi]$ *with Inner Product:*

$$\langle f, g \rangle = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) * g(x) dx$$

The following will hold:

$\mathit{1}$: $\langle 1, Cos(mx) \rangle = 0 \; \forall \; m \in \mathbb{N}$

$\mathit{2}$: $\langle 1, Sin(nx) \rangle = 0 \; \forall \; n \in \mathbb{N}$

udemy

Orthonormal Set on $C[\pi, \pi]$

Theorem (Trigonometric Orthogonal Functions wrt $\langle \cdot \rangle$)

Given the Function Space: $C[-\pi, \pi]$ *with Inner Product:*

$$\langle f, g \rangle = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) * g(x) dx$$

The following will hold:

1 : $\langle 1, Cos(mx) \rangle = 0 \ \forall \ m \in \mathbb{N}$

2 : $\langle 1, Sin(nx) \rangle = 0 \ \forall \ n \in \mathbb{N}$

3 : $\langle Cos(mx), Sin(nx) \rangle = 0 \ \forall m \neq n \ \& \ m, n \in \mathbb{N}$

udemy

Orthonormal Set on $C[\pi, \pi]$

Theorem (Trigonometric Orthogonal Functions wrt $\langle \cdot \rangle$)

Given the Function Space: $C[-\pi, \pi]$ *with Inner Product:*

$$\langle f, g \rangle = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) * g(x) dx$$

The following will hold:

1 : $\langle 1, Cos(mx) \rangle = 0 \; \forall \; m \in \mathbb{N}$

2 : $\langle 1, Sin(nx) \rangle = 0 \; \forall \; n \in \mathbb{N}$

3 : $\langle Cos(mx), Sin(nx) \rangle = 0 \; \forall m \neq n \; \& \; m, n \in \mathbb{N}$

4 : $\langle Cos(mx), Cos(nx) \rangle = 0 \; \forall m \neq n \; \& \; m, n \in \mathbb{N}$

Orthonormal Set on $C[\pi, \pi]$

Theorem (Trigonometric Orthogonal Functions wrt $\langle \cdot \rangle$)

Given the Function Space: $C[-\pi, \pi]$ *with Inner Product:*

$$\langle f, g \rangle = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) * g(x) dx$$

The following will hold:

1 : $\langle 1, Cos(mx) \rangle = 0 \ \forall \ m \in \mathbb{N}$

2 : $\langle 1, Sin(nx) \rangle = 0 \ \forall \ n \in \mathbb{N}$

3 : $\langle Cos(mx), Sin(nx) \rangle = 0 \ \forall m \neq n \ \& \ m, n \in \mathbb{N}$

4 : $\langle Cos(mx), Cos(nx) \rangle = 0 \ \forall m \neq n \ \& \ m, n \in \mathbb{N}$

5 : $\langle Sin(mx), Sin(nx) \rangle = 0 \ \forall m \neq n \ \& \ m, n \in \mathbb{N}$

Orthonormal Set on $C[\pi, \pi]$

Theorem (Trigonometric Orthogonal Functions wrt $\langle \cdot \rangle$)

Given the Function Space: $C[-\pi, \pi]$ *with Inner Product:*

$$\langle f, g \rangle = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) * g(x) dx$$

The following will hold:

1 : $\langle 1, Cos(mx) \rangle = 0 \ \forall \ m \in \mathbb{N}$

2 : $\langle 1, Sin(nx) \rangle = 0 \ \forall \ n \in \mathbb{N}$

3 : $\langle Cos(mx), Sin(nx) \rangle = 0 \ \forall m \neq n \ \& \ m, n \in \mathbb{N}$

4 : $\langle Cos(mx), Cos(nx) \rangle = 0 \ \forall m \neq n \ \& \ m, n \in \mathbb{N}$

5 : $\langle Sin(mx), Sin(nx) \rangle = 0 \ \forall m \neq n \ \& \ m, n \in \mathbb{N}$

6 : $||1|| = 1 = ||Sin(nx)|| = ||Cos(nx)|| \ \forall n \in \mathbb{N}$

udemy

Orthonormal Set on $C[-\pi, \pi]$

To prove conditions: 3 through 5 we will need the following two lemmas:

Given the trigonometric functions $Cos(mx), Sin(nx)$ with $m, n \in \mathbb{N}$ the following holds:

Orthonormal Set on $C[-\pi, \pi]$

To prove conditions: 3 through 5 we will need the following two lemmas:
Given the trigonometric functions $Cos(mx), Sin(nx)$ with $m, n \in \mathbb{N}$ the following holds:

Lemma (Product to Sum Trigonometric Functions)

$$1: \quad Sin[(m \pm n)x] = Sin(mx) * Cos(nx) \pm Sin(nx) * Cos(nx)$$

udemy

Orthonormal Set on $C[-\pi, \pi]$

To prove conditions: 3 through 5 we will need the following two lemmas:
Given the trigonometric functions $Cos(mx), Sin(nx)$ with $m, n \in \mathbb{N}$ the following holds:

Lemma (Product to Sum Trigonometric Functions)

$$1: \quad Sin[(m \pm n)x] = Sin(mx) * Cos(nx) \pm Sin(nx) * Cos(nx)$$

$$2: \quad Cos[(m \mp n)x] = Cos(mx) * Cos(nx) \pm Sin(mx) * Sin(nx)$$

udemy

Orthonormal Set on $C[-\pi, \pi]$

To prove conditions: (3) through (5) we will need the following two lemmas:
Given the trigonometric functions $Cos(mx), Sin(nx)$ with $m, n \in \mathbb{N}$ the following holds:

Lemma (Product to Sum Trigonometric Functions)

$$(1): \quad Sin[(m \pm n)x] = Sin(mx) * Cos(nx) \pm Sin(nx) * Cos(nx)$$

$$(2): \quad Cos[(m \mp n)x] = Cos(mx) * Cos(nx) \pm Sin(mx) * Sin(nx)$$

Lemma (Sum to Product of Trigonometric Functions)

$$(A): \quad Sin(mx) * Cos(nx) = \frac{Sin[(m+n)x] - Sin[(m-n)x]}{2}$$

Orthonormal Set on $C[-\pi, \pi]$

To prove conditions: (3) through (5) we will need the following two lemmas:
Given the trigonometric functions $Cos(mx), Sin(nx)$ with $m, n \in \mathbb{N}$ the following holds:

Lemma (Product to Sum Trigonometric Functions)

$$(1): \quad Sin[(m \pm n)x] = Sin(mx) * Cos(nx) \pm Sin(nx) * Cos(nx)$$

$$(2): \quad Cos[(m \mp n)x] = Cos(mx) * Cos(nx) \pm Sin(mx) * Sin(nx)$$

Lemma (Sum to Product of Trigonometric Functions)

$$(A): \quad Sin(mx) * Cos(nx) = \frac{Sin[(m+n)x] - Sin[(m-n)x]}{2}$$

$$(B): \quad Cos(mx) * Cos(nx) = \frac{Cos[(m+n)x] + Cos[(m-n)x]}{2}$$

Reference: Leon, 2014 [2]

Orthonormal Set on $C[-\pi, \pi]$

To prove conditions: 3 through 5 we will need the following two lemmas:
Given the trigonometric functions $Cos(mx), Sin(nx)$ with $m, n \in \mathbb{N}$ the following holds:

Lemma (Product to Sum Trigonometric Functions)

$$1: \quad Sin[(m \pm n)x] = Sin(mx) * Cos(nx) \pm Sin(nx) * Cos(nx)$$

$$2: \quad Cos[(m \mp n)x] = Cos(mx) * Cos(nx) \pm Sin(mx) * Sin(nx)$$

Lemma (Sum to Product of Trigonometric Functions)

$$A: \quad Sin(mx) * Cos(nx) = \frac{Sin[(m+n)x] - Sin[(m-n)x]}{2}$$

$$B: \quad Cos(mx) * Cos(nx) = \frac{Cos[(m+n)x] + Cos[(m-n)x]}{2}$$

$$C: \quad Sin(mx) * Sin(nx) = \frac{Cos[(m-n)x] - Cos[(m+n)x]}{2}$$

Reference: Leon, 2014 [2]

Orthonormal Basis

Theorem (Coefficient of the Orthogonal Basis)

Suppose $S = \{\mathbf{u_1}, \mathbf{u_2}, \ldots, \mathbf{u_j}, \ldots, \mathbf{u_n}\}$ *is an Orthonormal Basis on Inner Product Space* V*, then for each* $\mathbf{v} \in V$*:*

$$\mathbf{v} = \sum_{i=1}^{n} c_i \mathbf{u_i} \longrightarrow c_j = \langle \mathbf{v}, \mathbf{u_j} \rangle$$

Orthonormal Basis

Theorem (Coefficient of the Orthogonal Basis)

Suppose $S = \{\mathbf{u_1}, \mathbf{u_2}, \ldots, \mathbf{u_j}, \ldots, \mathbf{u_n}\}$ *is an Orthonormal Basis on Inner Product Space* V*, then for each* $\mathbf{v} \in V$*:*

$$\mathbf{v} = \sum_{i=1}^{n} c_i \mathbf{u_i} \longrightarrow c_j = \langle \mathbf{v}, \mathbf{u_j} \rangle$$

Proof.

$$\langle \mathbf{v}, \mathbf{u_j} \rangle = \left\langle \sum_{i=1}^{n} c_i \mathbf{u_i}, \mathbf{u_j} \right\rangle = \sum_{i=1}^{n} c_i \langle \mathbf{u_i}, \mathbf{u_j} \rangle = c_j \langle \mathbf{u_j}, \mathbf{u_j} \rangle = c_j$$

$\square$

Orthonormal Basis

Theorem (Coefficient of the Orthogonal Basis)

Suppose $S = \{\mathbf{u_1}, \mathbf{u_2}, \dots, \mathbf{u_j}, \dots, \mathbf{u_n}\}$ *is an Orthonormal Basis on Inner Product Space* V*, then for each* $\mathbf{v} \in V$*:*

$$\mathbf{v} = \sum_{i=1}^{n} c_i \mathbf{u_i} \longrightarrow c_j = \langle \mathbf{v}, \mathbf{u_j} \rangle$$

Proof.

$$\langle \mathbf{v}, \mathbf{u_j} \rangle = \left\langle \sum_{i=1}^{n} c_i \mathbf{u_i}, \mathbf{u_j} \right\rangle = \sum_{i=1}^{n} c_i \langle \mathbf{u_i}, \mathbf{u_j} \rangle = c_j \langle \mathbf{u_j}, \mathbf{u_j} \rangle = c_j$$

$\square$

Corollary

Suppose $S = \{\mathbf{u_1}, \mathbf{u_2}, \dots, \mathbf{u_j}, \dots, \mathbf{u_n}\}$ *is an* **Orthonormal Basis** *on Inner Product Space* V*, then for each* $\mathbf{v}, \mathbf{w} \in V$*:*

$$\mathbf{v} = \sum_{j=1}^{n} \alpha_j \mathbf{u_j} \;\&\; \mathbf{w} = \sum_{j=1}^{n} \beta_j \mathbf{u_j} \longrightarrow \langle \mathbf{v}, \mathbf{w} \rangle = \sum_{j=1}^{n} \alpha_j * \beta_j$$

Parseval's Theorem

Theorem (Parseval's Theorem for an Orthonormal Basis)

Suppose $S = \{\mathbf{u_1}, \mathbf{u_2}, \ldots, \mathbf{u_j}, \ldots, \mathbf{u_n}\} \sim$ *Orthonormal Basis on Inner Product Space* V*, then:*

$$\mathbf{v} = \sum_{j=1}^{n} \alpha_j \mathbf{u_j} \longrightarrow ||\mathbf{v}||^2 = \sum_{j=1}^{n} {\alpha_j}^2$$

udemy

Example: Parseval's Formula

Consider the Orthonormal Basis on $\mathbb{R}^2$: $S = \left\{ \begin{bmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{bmatrix}, \begin{bmatrix} \frac{1}{\sqrt{2}} \\ \frac{-1}{\sqrt{2}} \end{bmatrix} \right\}$

Determine the magnitude of any vector in $\mathbb{R}^2$ under the L_2 norm.

udemy

Example: Parseval's Formula

Consider the Orthonormal Basis on $\mathbb{R}^2$: $S = \left\{ \begin{bmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{bmatrix}, \begin{bmatrix} \frac{1}{\sqrt{2}} \\ \frac{-1}{\sqrt{2}} \end{bmatrix} \right\}$

Determine the magnitude of any vector in $\mathbb{R}^2$ under the L_2 norm.

Proof.

Since $S \sim$ **Orthonormal Basis** on $\mathbb{R}^2$ it follows that $\mathbf{x} = c_1\mathbf{u_1} + c_2\mathbf{u_2}$ for some $c_1, c_2 \in \mathbb{R}$.

□

udemy

Example: Parseval's Formula

Consider the Orthonormal Basis on $\mathbb{R}^2$: $S = \left\{ \begin{bmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{bmatrix}, \begin{bmatrix} \frac{1}{\sqrt{2}} \\ \frac{-1}{\sqrt{2}} \end{bmatrix} \right\}$

Determine the magnitude of any vector in $\mathbb{R}^2$ under the L_2 norm.

Proof.

Since $S \sim$ **Orthonormal Basis** on $\mathbb{R}^2$ it follows that $\mathbf{x} = c_1\mathbf{u_1} + c_2\mathbf{u_2}$ for some $c_1, c_2 \in \mathbb{R}$.

By previous theory:

$$c_1 = \langle \mathbf{x}, \mathbf{u_1} \rangle = \frac{x_1 + x_2}{\sqrt{2}} \ \& \ c_2 = \langle \mathbf{x}, \mathbf{u_2} \rangle = \frac{x_1 - x_2}{\sqrt{2}}$$

□

udemy

Example: Parseval's Formula

Consider the Orthonormal Basis on $\mathbb{R}^2$: $S = \left\{ \begin{bmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{bmatrix}, \begin{bmatrix} \frac{1}{\sqrt{2}} \\ \frac{-1}{\sqrt{2}} \end{bmatrix} \right\}$

Determine the magnitude of any vector in $\mathbb{R}^2$ under the L_2 norm.

Proof.

Since $S \sim$ **Orthonormal Basis** on $\mathbb{R}^2$ it follows that $\mathbf{x} = c_1\mathbf{u_1} + c_2\mathbf{u_2}$ for some $c_1, c_2 \in \mathbb{R}$.

By previous theory:

$$c_1 = \langle \mathbf{x}, \mathbf{u_1} \rangle = \frac{x_1 + x_2}{\sqrt{2}} \ \& \ c_2 = \langle \mathbf{x}, \mathbf{u_2} \rangle = \frac{x_1 - x_2}{\sqrt{2}}$$

By Parseval's Formula:

$$||\mathbf{x}||_2^2 = \langle \mathbf{x}, \mathbf{x} \rangle = c_1^2 + c_2^2 = x_1^2 + x_2^2$$

$\square$

udemy

Example: Parseval's Formula

The set: $S = \{1/\sqrt{2}, Cos(2x)\} \sim$ an Orthonormal Set in $C[-\pi, \pi]$ with Inner Product: $\langle \cdot \rangle = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) * g(x) dx$.

Find the value of $\int_{-\pi}^{\pi} Sin^4(x) dx$ without computing Anti-derivatives.

udemy

Example: Parseval's Formula

The set: $S = \{1/\sqrt{2}, Cos(2x)\} \sim$ an Orthonormal Set in $C[-\pi, \pi]$ with Inner Product: $\langle \cdot \rangle = \frac{1}{\pi}\int_{-\pi}^{\pi} f(x) * g(x)dx$.

Find the value of $\int_{-\pi}^{\pi} Sin^4(x)dx$ without computing Anti-derivatives.

Proof.

$$Cos(2x) = Cos^2(x) - Sin^2(x) = (1 - Sin^2(x)) - Sin^2(x)$$

$$\rightarrow Sin^2(x) = \frac{1 - Cos(2x)}{2} = \left(\frac{1}{\sqrt{2}}\right)\frac{1}{\sqrt{2}} + \left(-\frac{1}{2}\right)Cos(2x)$$

□

Example: Parseval's Formula

The set: $S = \{1/\sqrt{2}, Cos(2x)\}$ ~ an Orthonormal Set in $C[-\pi, \pi]$ with Inner Product: $\langle \cdot \rangle = \frac{1}{\pi}\int_{-\pi}^{\pi} f(x) * g(x)dx$.

Find the value of $\int_{-\pi}^{\pi} Sin^4(x)dx$ without computing Anti-derivatives.

Proof.

$$Cos(2x) = Cos^2(x) - Sin^2(x) = (1 - Sin^2(x)) - Sin^2(x)$$

$$\rightarrow Sin^2(x) = \frac{1 - Cos(2x)}{2} = \left(\frac{1}{\sqrt{2}}\right)\frac{1}{\sqrt{2}} + \left(-\frac{1}{2}\right)Cos(2x)$$

$$\int_{-\pi}^{\pi} Sin^4(x)dx = \pi * \left[\frac{1}{\pi}\int_{-\pi}^{\pi} Sin^2(x) * Sin^2(x)dx\right]$$

$$= \pi * ||Sin^2(x)||^2 = \pi\left[\left(\frac{1}{\sqrt{2}}\right)^2 + \left(\frac{-1}{2}\right)^2\right] = \frac{3\pi}{4}$$

$\square$

udemy

The Orthogonal Matrix

Definition (Orthogonal Matrix)

Given $Q \in \mathbb{R}^{nxn}$ is an **Orthogonal Matrix** provided:

$$Q^T Q = I_n \longleftrightarrow \mathbf{q_i}^T \mathbf{q_j} = \delta_{ij} = \begin{cases} 1 \ni: i = j \\ 0 \ni: i \neq j \end{cases}$$

udemy

The Orthogonal Matrix

Definition (Orthogonal Matrix)

Given $Q \in \mathbb{R}^{nxn}$ is an **Orthogonal Matrix** provided:

$$Q^T Q = I_n \longleftrightarrow \mathbf{q_i}^T \mathbf{q_j} = \delta_{ij} = \begin{cases} 1 \ni: i = j \\ 0 \ni: i \neq j \end{cases}$$

Remark: Equivalently: $Q \sim$ Orthogonal Square Matrix $\longleftrightarrow$ the columns of Q form an Orthonormal Basis on $\mathbb{R}^n$

udemy

The Orthogonal Matrix

Definition (Orthogonal Matrix)

Given $Q \in \mathbb{R}^{nxn}$ is an **Orthogonal Matrix** provided:

$$Q^T Q = I_n \longleftrightarrow \mathbf{q_i}^T \mathbf{q_j} = \delta_{ij} = \begin{cases} 1 \ni: i = j \\ 0 \ni: i \neq j \end{cases}$$

Remark: Equivalently: $Q \sim$ Orthogonal Square Matrix $\longleftrightarrow$ the columns of Q form an Orthonormal Basis on $\mathbb{R}^n$

Lemma (Properties of the Orthogonal Matrix)

Let $Q \in \mathbb{R}^{nxn}$ be an **Orthogonal Matrix** *then:*

1 : The columns of Q form an Orthonormal basis on $\mathbb{R}^n$

The Orthogonal Matrix

Definition (Orthogonal Matrix)

Given $Q \in \mathbb{R}^{nxn}$ is an **Orthogonal Matrix** provided:

$$Q^T Q = I_n \longleftrightarrow \mathbf{q_i}^T \mathbf{q_j} = \delta_{ij} = \begin{cases} 1 \ni: i = j \\ 0 \ni: i \neq j \end{cases}$$

Remark: Equivalently: $Q \sim$ Orthogonal Square Matrix $\longleftrightarrow$ the columns of Q form an Orthonormal Basis on $\mathbb{R}^n$

Lemma (Properties of the Orthogonal Matrix)

Let $Q \in \mathbb{R}^{nxn}$ be an **Orthogonal Matrix** *then:*

1 : *The columns of Q form an Orthonormal basis on $\mathbb{R}^n$*

2 : $Q^T Q = I_n \longleftrightarrow Q^{-1} = Q^T$

The Orthogonal Matrix

Definition (Orthogonal Matrix)

Given $Q \in \mathbb{R}^{nxn}$ is an **Orthogonal Matrix** provided:

$$Q^T Q = I_n \longleftrightarrow \mathbf{q_i}^T \mathbf{q_j} = \delta_{ij} = \begin{cases} 1 \ni: i = j \\ 0 \ni: i \neq j \end{cases}$$

Remark: Equivalently: $Q \sim$ Orthogonal Square Matrix $\longleftrightarrow$ the columns of Q form an Orthonormal Basis on $\mathbb{R}^n$

Lemma (Properties of the Orthogonal Matrix)

Let $Q \in \mathbb{R}^{nxn}$ be an **Orthogonal Matrix** *then:*

1 : *The columns of Q form an Orthonormal basis on $\mathbb{R}^n$*

2 : $Q^T Q = I_n \longleftrightarrow Q^{-1} = Q^T$

3 : $\langle Q\mathbf{x}, Q\mathbf{y} \rangle = \langle \mathbf{x}, \mathbf{y} \rangle$ *for each* $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$

The Orthogonal Matrix

Definition (Orthogonal Matrix)

Given $Q \in \mathbb{R}^{nxn}$ is an **Orthogonal Matrix** provided:

$$Q^T Q = I_n \longleftrightarrow \mathbf{q_i}^T \mathbf{q_j} = \delta_{ij} = \begin{cases} 1 \ni: i = j \\ 0 \ni: i \neq j \end{cases}$$

Remark: Equivalently: $Q \sim$ Orthogonal Square Matrix $\longleftrightarrow$ the columns of Q form an Orthonormal Basis on $\mathbb{R}^n$

Lemma (Properties of the Orthogonal Matrix)

Let $Q \in \mathbb{R}^{nxn}$ be an **Orthogonal Matrix** *then:*

1 : *The columns of Q form an Orthonormal basis on $\mathbb{R}^n$*

2 : $Q^T Q = I_n \longleftrightarrow Q^{-1} = Q^T$

3 : $\langle Q\mathbf{x}, Q\mathbf{y} \rangle = \langle \mathbf{x}, \mathbf{y} \rangle$ *for each* $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$

4 : $||Q\mathbf{x}|| = ||\mathbf{x}||$ *for each* $\mathbf{x} \in \mathbb{R}^n$

Reference: Leon, 2014 [2]

Example: Orthogonal Matrix in $\mathbb{R}^{2x2}$

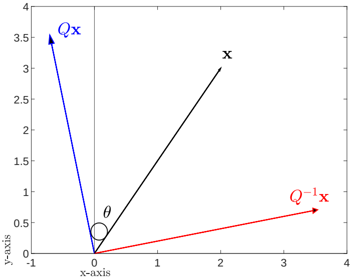
Lemma (Matrix Representation of CCW Rotation & Orthogonality)

Given the Linear Space $\mathbb{R}^2$ and the Linear Operator that rotates a $\mathbf{x} \in \mathbb{R}^2$ by an angle of $\theta \in \mathbb{R}$, then

$$Q = \begin{bmatrix} Cos(\theta) & -Sin(\theta) \\ Sin(\theta) & Cos(\theta) \end{bmatrix}$$

$$Q^T = \begin{bmatrix} Cos(\theta) & Sin(\theta) \\ -Sin(\theta) & Cos(\theta) \end{bmatrix}$$

Is an **Orthogonal Matrix**.



udemy

Reference: Leon, 2014 [2]

Example: Orthogonal Matrix in $\mathbb{R}^{2x2}$

Lemma (Matrix Representation of CCW Rotation & Orthogonality)

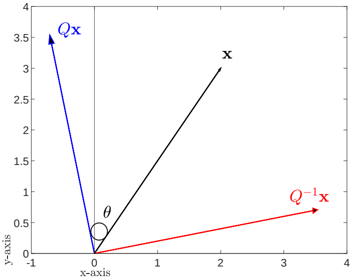
Given the Linear Space $\mathbb{R}^2$ and the Linear Operator that rotates a $\mathbf{x} \in \mathbb{R}^2$ by an angle of $\theta \in \mathbb{R}$, then

$$Q = \begin{bmatrix} Cos(\theta) & -Sin(\theta) \\ Sin(\theta) & Cos(\theta) \end{bmatrix}$$

$$Q^T = \begin{bmatrix} Cos(\theta) & Sin(\theta) \\ -Sin(\theta) & Cos(\theta) \end{bmatrix}$$

Is an **Orthogonal Matrix**.

Remark: *Q represents the CCW rotation by an angle of θ and Q^T represents the CW rotation by an angle of θ.*



udemy

The Permutation Matrix

Definition (Permutation Matrix)

The Permutation Matrix is obtained by re-ordering the columns of I_n by a collection of indices: $\mathcal{K} = \{k_1, k_2, \ldots, k_n\}$.

$$P = [\mathbf{e_{k_1}}, \mathbf{e_{k_2}}, \ldots, \mathbf{e_{k_n}}] \in \mathbb{R}^{nxn}$$

The Permutation Matrix

Definition (Permutation Matrix)

The Permutation Matrix is obtained by re-ordering the columns of I_n by a collection of indices: $\mathcal{K} = \{k_1, k_2, \ldots, k_n\}$.

$$P = [\mathbf{e_{k_1}}, \mathbf{e_{k_2}}, \ldots, \mathbf{e_{k_n}}] \in \mathbb{R}^{nxn}$$

Remarks: 1: Post-multiplication of $A \in \mathbb{R}^{mxn}$ by $P \in \mathbb{R}^{nxn}$ re-orders the columns of A:

$$AP = [\mathbf{a_{k_1}}, \mathbf{a_{k_2}}, \ldots, \mathbf{a_{k_n}}]$$

The Permutation Matrix

Definition (Permutation Matrix)

The Permutation Matrix is obtained by re-ordering the columns of I_n by a collection of indices: $\mathcal{K} = \{k_1, k_2, \ldots, k_n\}$.

$$P = [\mathbf{e_{k_1}}, \mathbf{e_{k_2}}, \ldots, \mathbf{e_{k_n}}] \in \mathbb{R}^{nxn}$$

Remarks: 1: Post-multiplication of $A \in \mathbb{R}^{mxn}$ by $P \in \mathbb{R}^{nxn}$ re-orders the columns of A:

$$AP = [\mathbf{a_{k_1}}, \mathbf{a_{k_2}}, \ldots, \mathbf{a_{k_n}}]$$

2: Likewise, Pre-multiplication of $A \in \mathbb{R}^{nxr}$ by $P \in \mathbb{R}^{nxn}$ re-orders the rows of A:

$$PA = (\mathbf{a}'_{\mathbf{k_1}}, \mathbf{a}'_{\mathbf{k_2}}, \ldots, \mathbf{a}'_{\mathbf{k_r}})$$

The Permutation Matrix

Definition (Permutation Matrix)

The Permutation Matrix is obtained by re-ordering the columns of I_n by a collection of indices: $\mathcal{K} = \{k_1, k_2, \ldots, k_n\}$.

$$P = [\mathbf{e_{k_1}}, \mathbf{e_{k_2}}, \ldots, \mathbf{e_{k_n}}] \in \mathbb{R}^{nxn}$$

Remarks: (1): Post-multiplication of $A \in \mathbb{R}^{mxn}$ by $P \in \mathbb{R}^{nxn}$ re-orders the columns of A:

$$AP = [\mathbf{a_{k_1}}, \mathbf{a_{k_2}}, \ldots, \mathbf{a_{k_n}}]$$

(2): Likewise, Pre-multiplication of $A \in \mathbb{R}^{nxr}$ by $P \in \mathbb{R}^{nxn}$ re-orders the rows of A:

$$PA = (\mathbf{a}'_{\mathbf{k_1}}, \mathbf{a}'_{\mathbf{k_2}}, \ldots, \mathbf{a}'_{\mathbf{k_r}})$$

(3): Every Permutation Matrix, P, is an Orthogonal Matrix: $P^T P = I_n$

Example: Permutation Matrices

Consider:

$$A = \begin{bmatrix} 8 & 1 & 6 \\ 3 & 5 & 7 \\ 4 & 9 & 2 \end{bmatrix} \quad P = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix}$$

Since $P = [\mathbf{e_3}, \mathbf{e_1}, \mathbf{e_2}]$, then $P \sim$ **Permutation Matrix**.

Compute PA and AP:

udemy

Example: Permutation Matrices

Consider:

$$A = \begin{bmatrix} 8 & 1 & 6 \\ 3 & 5 & 7 \\ 4 & 9 & 2 \end{bmatrix} \quad P = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix}$$

Since $P = [\mathbf{e_3}, \mathbf{e_1}, \mathbf{e_2}]$, then $P \sim$ **Permutation Matrix**.

Compute PA and AP:

Computation.

$$PA = (\mathbf{a}_2', \mathbf{a}_3', \mathbf{a}_1') = \begin{bmatrix} 3 & 5 & 7 \\ 4 & 9 & 2 \\ 8 & 1 & 6 \end{bmatrix} \quad AP = [\mathbf{a_3}, \mathbf{a_1}, \mathbf{a_2}] = \begin{bmatrix} 6 & 8 & 1 \\ 7 & 3 & 5 \\ 2 & 4 & 9 \end{bmatrix}$$

$\square$

Orthonormal Sets and Least Squares

Lemma (Orthonormal Sets for Least Squares Problems)

Suppose one solves $A\mathbf{x} = \mathbf{b}$ where $A \in \mathbb{R}^{mxn}$, $\mathbf{x} \in \mathbb{R}^n$, and $\mathbf{b} \in \mathbb{R}^m$ where $m > n$ and the system is inconsistent.

If Rank$(A) = n$ and the columns of A form an Orthonormal Set of vectors, then the Unique Least Squares Solution is:

$$A^T A = I_n \longrightarrow \hat{\mathbf{x}} = A^T \mathbf{b}$$

udemy

Orthonormal Basis

Theorem (The Orthonormal Basis)

Consider $S \subset V$ *where* $V \sim$ **Inner Product Space** *and consider* $\mathbf{x} \in V$.
Take $\mathcal{B} = \{\mathbf{u_1}, \mathbf{u_2}, \ldots, \mathbf{u_j}, \ldots, \mathbf{u_n}\}$ *as an Orthonormal Basis on* S*, then:*

$$\mathbf{p} = \sum_{j=1}^{n} c_j \mathbf{u_j} \longrightarrow c_j = \langle \mathbf{x}, \mathbf{u_j} \rangle \text{ for each } j = 1, 2, \ldots, n$$

udemy

Orthonormal Basis

Theorem (The Orthonormal Basis)

Consider $S \subset V$ *where* $V \sim$ **Inner Product Space** *and consider* $\mathbf{x} \in V$.
Take $\mathcal{B} = \{\mathbf{u_1}, \mathbf{u_2}, \ldots, \mathbf{u_j}, \ldots, \mathbf{u_n}\}$ *as an Orthonormal Basis on* S, *then:*

$$\mathbf{p} = \sum_{j=1}^{n} c_j \mathbf{u_j} \longrightarrow c_j = \langle \mathbf{x}, \mathbf{u_j} \rangle \text{ for each } j = 1, 2, \ldots, n$$

Remark: *Additionally, one has* $\mathbf{p} - \mathbf{x} \perp S$

udemy

Orthonormal Basis

Theorem (The Orthonormal Basis)

Consider $S \subset V$ *where* $V \sim$ **Inner Product Space** *and consider* $\mathbf{x} \in V$*.*
Take $\mathcal{B} = \{\mathbf{u_1}, \mathbf{u_2}, \dots, \mathbf{u_j}, \dots, \mathbf{u_n}\}$ *as an Orthonormal Basis on* S*, then:*

$$\mathbf{p} = \sum_{j=1}^{n} c_j \mathbf{u_j} \longrightarrow c_j = \langle \mathbf{x}, \mathbf{u_j} \rangle \text{ for each } j = 1, 2, \dots, n$$

Remark*: Additionally, one has* $\mathbf{p} - \mathbf{x} \perp S$

Theorem (Projection Vector & Closest Element)

Take $S \subset V$ *where* $V \sim$ *Inner Product and take* $\mathbf{x} \in V$*, then the projection of* V *onto* S *(*$\mathbf{p}$*) is the closest vector in* S *to* $\mathbf{x}$*:*

$$||\mathbf{x} - \mathbf{y}|| > ||\mathbf{x} - \mathbf{p}|| \text{ for each } y \neq p \text{ and } \mathbf{y}, \mathbf{p} \in S$$

Orthonormal Basis

Theorem (The Orthonormal Basis)

Consider $S \subset V$ *where* $V \sim$ **Inner Product Space** *and consider* $\mathbf{x} \in V$*.*
Take $\mathcal{B} = \{\mathbf{u_1}, \mathbf{u_2}, \ldots, \mathbf{u_j}, \ldots, \mathbf{u_n}\}$ *as an Orthonormal Basis on* S*, then:*

$$\mathbf{p} = \sum_{j=1}^{n} c_j \mathbf{u_j} \longrightarrow c_j = \langle \mathbf{x}, \mathbf{u_j} \rangle \text{ for each } j = 1, 2, \ldots, n$$

Remark: *Additionally, one has* $\mathbf{p} - \mathbf{x} \perp S$

Theorem (Projection Vector & Closest Element)

Take $S \subset V$ *where* $V \sim$ *Inner Product and take* $\mathbf{x} \in V$*, then the projection of* V *onto* S *(*$\mathbf{p}$*) is the closest vector in* S *to* $\mathbf{x}$*:*

$$||\mathbf{x} - \mathbf{y}|| > ||\mathbf{x} - \mathbf{p}|| \text{ for each } y \neq p \text{ and } \mathbf{y}, \mathbf{p} \in S$$

Remark: *Taking* $V = \mathbb{R}^m$*,* $p = A\hat{x}$*,* $\mathbf{x} = \mathbf{b}$*, and* $S = R(A)$ *gives the formulation of the Normal Equations.*

Projection and Orthonormal Basis

Corollary

Let $S \subset \mathbb{R}^n$ and $\mathbf{b} \in \mathbb{R}^m$.

Suppose $\mathbb{B} = \{\mathbf{u_1}, \mathbf{u_2}, \ldots, \mathbf{u_j}, \ldots, \mathbf{u_n}\}$ is an Orthonormal Basis on S.

The Projection of $\mathbf{b}$ onto S is:

$$\mathbf{p} = UU^T\mathbf{b} \ \ni: \ U = [\mathbf{u_1}, \mathbf{u_2}, \ldots, \mathbf{u_n}]$$

udemy

Projection and Orthonormal Basis

Corollary

Let $S \subset \mathbb{R}^n$ and $\mathbf{b} \in \mathbb{R}^m$.
Suppose $\mathbb{B} = \{\mathbf{u_1}, \mathbf{u_2}, \ldots, \mathbf{u_j}, \ldots, \mathbf{u_n}\}$ is an Orthonormal Basis on S.
The Projection of $\mathbf{b}$ onto S is:

$$\mathbf{p} = UU^T\mathbf{b} \ \ni: \ U = [\mathbf{u_1}, \mathbf{u_2}, \ldots, \mathbf{u_n}]$$

Remark: *$Q = UU^T \longleftrightarrow Q\mathbf{b} = UU^T\mathbf{b}$ is also known as the Projection Matrix (Unique) of $\mathbf{b}$ onto S.*

udemy

Table of Contents

Construction of an Orthonormal Basis

Corollary

Suppose $\mathcal{B}_x = \{\mathbf{x_1}, \mathbf{x_2}, \ldots, \mathbf{x_n}\}$ *is an* **Ordered Basis** *on Inner Product Space* V *and* $\mathcal{B}_u = \{\mathbf{u_1}, \mathbf{u_2}, \ldots, \mathbf{u_n}\}$ *is an* **Orthonormal Basis** *on* V*, then:*

$$Span(\mathcal{B}_x) = V = Span(\mathcal{B}_u)$$

udemy

Intuition behind Construction of the Orthonormal Basis

Suppose $S = \{\mathbf{x_1}, \mathbf{x_2}, \ldots, \mathbf{x_n}\}$ is a Basis on Inner Product Space V.
First, take the Unit Vector: $\mathbf{u_1} = \frac{\mathbf{x_1}}{||\mathbf{x_1}||}$.

udemy

Intuition behind Construction of the Orthonormal Basis

Suppose $S = \{\mathbf{x_1}, \mathbf{x_2}, \ldots, \mathbf{x_n}\}$ is a Basis on Inner Product Space V.
First, take the Unit Vector: $\mathbf{u_1} = \frac{\mathbf{x_1}}{||\mathbf{x_1}||}$.
Then, take the Projection of $\mathbf{x_2}$ onto $\mathbf{u_1}$ as:

$$\mathbf{p_1} = \frac{\langle \mathbf{x_2}, \mathbf{u_1} \rangle}{\langle \mathbf{u_1}, \mathbf{u_1} \rangle} \mathbf{u_1}$$

udemy

Intuition behind Construction of the Orthonormal Basis

Suppose $S = \{\mathbf{x_1}, \mathbf{x_2}, \ldots, \mathbf{x_n}\}$ is a Basis on Inner Product Space V.
First, take the Unit Vector: $\mathbf{u_1} = \frac{\mathbf{x_1}}{||\mathbf{x_1}||}$.
Then, take the Projection of $\mathbf{x_2}$ onto $\mathbf{u_1}$ as:

$$\mathbf{p_1} = \frac{\langle \mathbf{x_2}, \mathbf{u_1} \rangle}{\langle \mathbf{u_1}, \mathbf{u_1} \rangle} \mathbf{u_1}$$

Because $\mathbf{x_1}, \mathbf{x_2}$ are Linearly Independent, take:

$$\mathbf{u_2} = \frac{\mathbf{x_2} - \mathbf{p_1}}{||\mathbf{x_2} - \mathbf{p_1}||}$$

Intuition behind Construction of the Orthonormal Basis

Suppose $S = \{\mathbf{x_1}, \mathbf{x_2}, \ldots, \mathbf{x_n}\}$ is a Basis on Inner Product Space V.
First, take the Unit Vector: $\mathbf{u_1} = \frac{\mathbf{x_1}}{||\mathbf{x_1}||}$.
Then, take the Projection of $\mathbf{x_2}$ onto $\mathbf{u_1}$ as:

$$\mathbf{p_1} = \frac{\langle \mathbf{x_2}, \mathbf{u_1} \rangle}{\langle \mathbf{u_1}, \mathbf{u_1} \rangle} \mathbf{u_1}$$

Because $\mathbf{x_1}, \mathbf{x_2}$ are Linearly Independent, take:

$$\mathbf{u_2} = \frac{\mathbf{x_2} - \mathbf{p_1}}{||\mathbf{x_2} - \mathbf{p_1}||}$$

Take the projection of $Span(\mathbf{x_1}, \mathbf{x_2})$ onto $\mathbf{u_2}$:

$$\mathbf{p_2} = \frac{\langle \mathbf{x_1}, \mathbf{u_1} \rangle}{\langle \mathbf{u_1}, \mathbf{u_1} \rangle} \mathbf{u_1} + \frac{\langle \mathbf{x_2}, \mathbf{u_2} \rangle}{\langle \mathbf{u_2}, \mathbf{u_2} \rangle} \mathbf{u_2} \longrightarrow \mathbf{u_3} = \frac{\mathbf{x_3} - \mathbf{p_2}}{||\mathbf{x_3} - \mathbf{p_2}||}$$

Iterate the process until the $Span(\mathbf{x_1}, \mathbf{x_2}, \ldots, \mathbf{x_n})$ is reached (requires $n - 1$ iterations).

Construction of an Orthonormal Basis

Theorem (Gram-Schmidt Orthogonalization Process)

Consider $\mathcal{B}_x = \{\mathbf{x_1}, \mathbf{x_2}, \ldots, \mathbf{x_n}\}$ *as a Basis on Inner Product Space* V.
Take: $\mathbf{u_1} = \frac{\mathbf{x_1}}{||\mathbf{x_1}||}$ *and define:*

$$\mathbf{u}_{k+1} = \frac{\mathbf{x}_{k+1} - \mathbf{p}_k}{||\mathbf{x}_{k+1} - \mathbf{p}_k||} \quad \ni: \ \mathbf{p}_k = \sum_{j=1}^{k} \frac{\langle \mathbf{x}_{k+1}, \mathbf{u}_j \rangle}{\langle \mathbf{u_j}, \mathbf{u_j} \rangle} \mathbf{u_j}$$

For each $k = 1, 2, \ldots, (n-1)$

Then $\mathcal{B}_u = \{\mathbf{u_1}, \mathbf{u_2}, \ldots, \mathbf{u_n}\}$ *is an Orthonormal Basis on* V.

udemy

Gram-Schmidt Orthogonalization Process

By Induction.

Note that $Span(\mathbf{x_1}) = Span(\mathbf{u_1})$. Assume for some $k \in \mathbb{N}$ where $1 \leq k \leq n-1$ such that: $\mathcal{B}_u = \{\mathbf{u_1}, \mathbf{u_2}, \ldots, \mathbf{u_k}\}$ is an Orthonormal Set such that:

$$Span(\mathcal{B}_x) = Span(\mathbf{u_1}, \mathbf{u_2}, \ldots, \mathbf{u_k})$$

$\square$

udemy

Gram-Schmidt Orthogonalization Process

By Induction.

Note that $Span(\mathbf{x_1}) = Span(\mathbf{u_1})$. Assume for some $k \in \mathbb{N}$ where $1 \leq k \leq n-1$ such that: $\mathcal{B}_u = \{\mathbf{u_1}, \mathbf{u_2}, \dots, \mathbf{u_k}\}$ is an Orthonormal Set such that:

$$Span(\mathcal{B}_x) = Span(\mathbf{u_1}, \mathbf{u_2}, \dots, \mathbf{u_k})$$

$$\mathbf{p_k} \in Span(\mathbf{u_1}, \mathbf{u_2}, \dots, \mathbf{u_k}) \longrightarrow \mathbf{p_k} \in Span(\mathcal{B}_x)$$

$\square$

udemy

Gram-Schmidt Orthogonalization Process

By Induction.

Note that $Span(\mathbf{x_1}) = Span(\mathbf{u_1})$. Assume for some $k \in \mathbb{N}$ where $1 \leq k \leq n-1$ such that: $\mathcal{B}_u = \{\mathbf{u_1}, \mathbf{u_2}, \ldots, \mathbf{u_k}\}$ is an Orthonormal Set such that:

$$Span(\mathcal{B}_x) = Span(\mathbf{u_1}, \mathbf{u_2}, \ldots, \mathbf{u_k})$$

$$\mathbf{p_k} \in Span(\mathbf{u_1}, \mathbf{u_2}, \ldots, \mathbf{u_k}) \longrightarrow \mathbf{p_k} \in Span(\mathcal{B}_x)$$

$$\mathbf{x_{k+1}} - \mathbf{p_k} = \mathbf{x_{k+1}} - \sum_{j=1}^{k} c_j \mathbf{u_j} \ \textit{for some}\ c_j \in \mathbb{R}$$

$$Since:\ \mathcal{B}_x \sim Linearly\ Independent \longrightarrow x_{k+1} - p_k \neq 0$$

$\square$

Gram-Schmidt Orthogonalization Process

By Induction.

Note that $Span(\mathbf{x_1}) = Span(\mathbf{u_1})$. Assume for some $k \in \mathbb{N}$ where $1 \le k \le n-1$ such that: $\mathcal{B}_u = \{\mathbf{u_1}, \mathbf{u_2}, \dots, \mathbf{u_k}\}$ is an Orthonormal Set such that:

$$Span(\mathcal{B}_x) = Span(\mathbf{u_1}, \mathbf{u_2}, \dots, \mathbf{u_k})$$

$$\mathbf{p_k} \in Span(\mathbf{u_1}, \mathbf{u_2}, \dots, \mathbf{u_k}) \longrightarrow \mathbf{p_k} \in Span(\mathcal{B}_x)$$

$$\mathbf{x_{k+1}} - \mathbf{p_k} = \mathbf{x_{k+1}} - \sum_{j=1}^{k} c_j \mathbf{u_j} \ \textit{for some } c_j \in \mathbb{R}$$

$$\textit{Since}: \ \mathcal{B}_x \sim \textit{Linearly Independent} \longrightarrow x_{k+1} - p_k \neq 0$$

$$\textit{Since}: \ \mathbf{x_{k+1}} - \mathbf{p_k} \in Span(\mathcal{B}_u)^{\perp} \longrightarrow \mathbf{x_{k+1}} - \mathbf{p_k} \perp \mathbf{u_j} \ \ni: \ j = 1, 2, \dots, k$$

$\square$

udemy

Gram-Schmidt Orthogonalization Process

By Induction.

Note that $Span(\mathbf{x_1}) = Span(\mathbf{u_1})$. Assume for some $k \in \mathbb{N}$ where $1 \leq k \leq n-1$ such that: $\mathcal{B}_u = \{\mathbf{u_1}, \mathbf{u_2}, \ldots, \mathbf{u_k}\}$ is an Orthonormal Set such that:

$$Span(\mathcal{B}_x) = Span(\mathbf{u_1}, \mathbf{u_2}, \ldots, \mathbf{u_k})$$

$$\mathbf{p_k} \in Span(\mathbf{u_1}, \mathbf{u_2}, \ldots, \mathbf{u_k}) \longrightarrow \mathbf{p_k} \in Span(\mathcal{B}_x)$$

$$\mathbf{x_{k+1}} - \mathbf{p_k} = \mathbf{x_{k+1}} - \sum_{j=1}^{k} c_j \mathbf{u_j} \ \textit{for some } c_j \in \mathbb{R}$$

$$\textit{Since}: \ \mathcal{B}_x \sim \textit{Linearly Independent} \longrightarrow x_{k+1} - p_k \neq 0$$

$$\textit{Since}: \ \mathbf{x_{k+1}} - \mathbf{p_k} \in Span(\mathcal{B}_u)^{\perp} \longrightarrow \mathbf{x_{k+1}} - \mathbf{p_k} \perp \mathbf{u_j} \ \ni: \ j = 1, 2, \ldots, k$$

$$\textit{The Projection of } \mathbf{x_{k+1}} \textit{ onto } Span(\mathcal{B}_u): \mathbf{p_k} = \sum_{j=1}^{k} c_j \mathbf{u_j} \ni: c_j = \langle \mathbf{x_{k+1}}, \mathbf{u_j} \rangle$$

$\square$

Example: Gram-Schmidt Orthogonalization

Construct an Orthonormal Basis on the Column Space of A.

$$A = \begin{bmatrix} 1 & -1 & 4 \\ 1 & 4 & -2 \\ 1 & 4 & -2 \\ 1 & -1 & 0 \end{bmatrix}$$

Example: Gram-Schmidt Orthogonalization

Construct an Orthonormal Basis on the Column Space of A.

$$A = \begin{bmatrix} 1 & -1 & 4 \\ 1 & 4 & -2 \\ 1 & 4 & -2 \\ 1 & -1 & 0 \end{bmatrix}$$

$$1 : r_{11} = ||\mathbf{a_1}|| = 2$$

$$\mathbf{q_1} = \frac{\mathbf{a_1}}{||\mathbf{a_1}||} = (1/2, 1/2, 1/2, 1/2)^T$$

$$r_{12} = \langle \mathbf{a_2}, \mathbf{q_1} \rangle = 3$$

$$\mathbf{p_1} = \langle \mathbf{a_2}, \mathbf{q_1} \rangle \mathbf{q_1} = r_{12}\mathbf{q_1} = (3/2, 3/2, 3/2, 3/2)^T$$

udemy

Example: Gram-Schmidt Orthogonalization

Construct an Orthonormal Basis on the Column Space of A.

$$A = \begin{bmatrix} 1 & -1 & 4 \\ 1 & 4 & -2 \\ 1 & 4 & -2 \\ 1 & -1 & 0 \end{bmatrix}$$

$$1: r_{11} = ||\mathbf{a_1}|| = 2$$

$$\mathbf{q_1} = \frac{\mathbf{a_1}}{||\mathbf{a_1}||} = (1/2, 1/2, 1/2, 1/2)^T$$

$$r_{12} = \langle \mathbf{a_2}, \mathbf{q_1} \rangle = 3$$

$$\mathbf{p_1} = \langle \mathbf{a_2}, \mathbf{q_1} \rangle \mathbf{q_1} = r_{12}\mathbf{q_1} = (3/2, 3/2, 3/2, 3/2)^T$$

$$2: r_{22} = ||\mathbf{a_2} - \mathbf{p_1}|| = 5$$

$$\mathbf{q_2} = \frac{\mathbf{a_2} - \mathbf{p_1}}{||\mathbf{a_2} - \mathbf{p_1}||} = (-1/2, 1/2, 1/2, -1/2)^T$$

$$r_{13} = \langle \mathbf{a_3}, \mathbf{q_1} \rangle = 2 \quad r_{23} = \langle \mathbf{a_3}, \mathbf{q_2} \rangle = -2$$

$$\mathbf{p_2} = r_{13}\mathbf{q_1} + r_{23}\mathbf{q_2} = (2, 0, 0, 2)^T$$

udemy

Example: Gram-Schmidt Orthogonalization

Construct an Orthonormal Basis on the Column Space of A.

$$A = \begin{bmatrix} 1 & -1 & 4 \\ 1 & 4 & -2 \\ 1 & 4 & -2 \\ 1 & -1 & 0 \end{bmatrix}$$

$$\text{1}: r_{11} = ||\mathbf{a_1}|| = 2$$

$$\mathbf{q_1} = \frac{\mathbf{a_1}}{||\mathbf{a_1}||} = (1/2, 1/2, 1/2, 1/2)^T$$

$$r_{12} = \langle \mathbf{a_2}, \mathbf{q_1} \rangle = 3$$

$$\mathbf{p_1} = \langle \mathbf{a_2}, \mathbf{q_1} \rangle \mathbf{q_1} = r_{12}\mathbf{q_1} = (3/2, 3/2, 3/2, 3/2)^T$$

$$\text{2}: r_{22} = ||\mathbf{a_2} - \mathbf{p_1}|| = 5$$

$$\mathbf{q_2} = \frac{\mathbf{a_2} - \mathbf{p_1}}{||\mathbf{a_2} - \mathbf{p_1}||} = (-1/2, 1/2, 1/2, -1/2)^T$$

$$r_{13} = \langle \mathbf{a_3}, \mathbf{q_1} \rangle = 2 \quad r_{23} = \langle \mathbf{a_3}, \mathbf{q_2} \rangle = -2$$

$$\mathbf{p_2} = r_{13}\mathbf{q_1} + r_{23}\mathbf{q_2} = (2, 0, 0, 2)^T$$

$$\text{3}: r_{33} = ||\mathbf{a_3} - p_2|| = 4$$

$$\mathbf{q_3} = \frac{\mathbf{a_3} - \mathbf{p_2}}{||\mathbf{a_3} - \mathbf{p_2}||} = (1/2, -1/2, 1/2, -1/2)^T$$

udemy

Reference: Leon, 2014 [2]

Example: Gram-Schmidt Orthogonalization Continued

One can show

$$\langle \mathbf{q_i}, \mathbf{q_j} \rangle = \delta_{ij} = \begin{cases} 1 & \ni: i = j \\ 0 & \ni: i \neq j \end{cases}$$

Thus, taking: $Q = [\mathbf{q_1}, \mathbf{q_2}, \mathbf{q_3}]$ will give an Orthonormal Basis on $C(A)$.
Putting $r_{ij} \in \mathbb{R}$ into the Matrix R will give the Upper Triangular Residual Matrix where:

$$A = QR$$

udemy

Example: Gram-Schmidt Orthogonalization Continued

One can show

$$\langle \mathbf{q_i}, \mathbf{q_j} \rangle = \delta_{ij} = \begin{cases} 1 & \ni: i = j \\ 0 & \ni: i \neq j \end{cases}$$

Thus, taking: $Q = [\mathbf{q_1}, \mathbf{q_2}, \mathbf{q_3}]$ will give an Orthonormal Basis on $C(A)$.
Putting $r_{ij} \in \mathbb{R}$ into the Matrix R will give the Upper Triangular Residual Matrix where:

$$A = QR$$

Remark: Provided the columns of A do not form an Orthonormal set of vectors and $Rank(A) = n$, then the factorization can be used to solve Least Squares.

udemy

Gram-Schmidt QR Factorization for Least Squares

Theorem (QR Factorization of Full Column Rank $A \in \mathbb{R}^{mxn}$)

Suppose $A \in \mathbb{R}^{mxn}$ *where* $Rank(A) = n$, *then* A *can be factored in terms of* QR *where* $\{\mathbf{q_1}, \mathbf{q_2}, \ldots, \mathbf{q_n}\} \sim$ *Orthonormal Basis on* $C(A)$.

Suppose one solves the system $A\mathbf{x} = \mathbf{b}$ *and the system is* **Inconsistent**, *yet* A *has full column Rank, then* $A = QR$ *where* Q *has Orthonormal columns and one may alternatively solve:*

$$QR\mathbf{x} = \mathbf{b} \longrightarrow R\mathbf{x} = Q^T\mathbf{b} \longleftrightarrow R\hat{x} = \hat{b} \text{ Where : } Q^T\mathbf{b} = \hat{b}$$

Gram-Schmidt QR Factorization for Least Squares

Theorem (QR Factorization of Full Column Rank $A \in \mathbb{R}^{mxn}$)

Suppose $A \in \mathbb{R}^{mxn}$ where $Rank(A) = n$, then A can be factored in terms of QR where $\{\mathbf{q_1}, \mathbf{q_2}, \ldots, \mathbf{q_n}\} \sim$ Orthonormal Basis on $C(A)$.

Suppose one solves the system $A\mathbf{x} = \mathbf{b}$ and the system is **Inconsistent**, *yet A has full column Rank, then $A = QR$ where Q has Orthonormal columns and one may alternatively solve:*

$$QR\mathbf{x} = \mathbf{b} \longrightarrow R\mathbf{x} = Q^T\mathbf{b} \longleftrightarrow R\hat{x} = \hat{b} \text{ Where : } Q^T\mathbf{b} = \hat{b}$$

Where R is an Upper Triangular Matrix with $(i,j)_{th}$ element:

$$r_{ij} = \begin{cases} ||\mathbf{a_1}|| \ni: i = 1 = j \\ \mathbf{q_i}^T\mathbf{a_j} \ni: i < j \\ ||\mathbf{a_j} - \mathbf{p_{j-1}}|| \ni: i = j = 2, 3, \ldots, n \end{cases}$$

Gram-Schmidt QR Factorization for Least Squares

Theorem (QR Factorization of Full Column Rank $A \in \mathbb{R}^{mxn}$)

Suppose $A \in \mathbb{R}^{mxn}$ *where* $Rank(A) = n$*, then* A *can be factored in terms of* QR *where* $\{\mathbf{q_1}, \mathbf{q_2}, \ldots, \mathbf{q_n}\} \sim$ *Orthonormal Basis on* $C(A)$*.*
Suppose one solves the system $A\mathbf{x} = \mathbf{b}$ *and the system is* **Inconsistent***, yet* A *has full column Rank, then* $A = QR$ *where* Q *has Orthonormal columns and one may alternatively solve:*

$$QR\mathbf{x} = \mathbf{b} \longrightarrow R\mathbf{x} = Q^T\mathbf{b} \longleftrightarrow R\hat{x} = \hat{b} \text{ Where : } \ Q^T\mathbf{b} = \hat{b}$$

Where R *is an Upper Triangular Matrix with* $(i,j)_{th}$ *element:*

$$r_{ij} = \begin{cases} ||\mathbf{a_1}|| \ni: i = 1 = j \\ \mathbf{q_i}^T\mathbf{a_j} \ni: i < j \\ ||\mathbf{a_j} - \mathbf{p_{j-1}}|| \ni: i = j = 2, 3, \ldots, n \end{cases}$$

Remark*: If* $A \in \mathbb{R}^{nxn} \longrightarrow Det(A) = Det(R) = \prod_{j=1}^{n} r_{jj}$

Example: QR for Least Squares

Find the Least Squares solution to the following system:

$$
\begin{aligned}
x_1 - 2x_2 - x_1 &= -1 \\
2x_1 + 0x_2 + x_3 &= 1 \\
2x_1 - 4x_2 + 2x_3 &= 1 \\
4x_1 + 0x_2 + 0x_3 &= -2
\end{aligned}
$$

.

udemy

Example: QR for Least Squares

Find the Least Squares solution to the following system:

$$\begin{aligned} x_1 - 2x_2 - x_1 &= -1 \\ 2x_1 + 0x_2 + x_3 &= 1 \\ 2x_1 - 4x_2 + 2x_3 &= 1 \\ 4x_1 + 0x_2 + 0x_3 &= -2 \end{aligned}$$

Expressing the system in its Augmented form gives:

$$[A|\mathbf{b}] = \left[\begin{array}{ccc|c} 1 & -2 & -1 & -1 \\ 2 & 0 & 1 & 1 \\ 2 & -4 & 2 & 1 \\ 4 & 0 & 0 & -2 \end{array}\right] \sim RREF \sim [U|\mathbf{b_u}] = \left[\begin{array}{ccc|c} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{array}\right]$$

.

udemy

Example: QR for Least Squares

Find the Least Squares solution to the following system:

$$\begin{aligned} x_1 - 2x_2 - x_1 &= -1 \\ 2x_1 + 0x_2 + x_3 &= 1 \\ 2x_1 - 4x_2 + 2x_3 &= 1 \\ 4x_1 + 0x_2 + 0x_3 &= -2 \end{aligned}$$

Expressing the system in its Augmented form gives:

$$[A|\mathbf{b}] = \left[\begin{array}{ccc|c} 1 & -2 & -1 & -1 \\ 2 & 0 & 1 & 1 \\ 2 & -4 & 2 & 1 \\ 4 & 0 & 0 & -2 \end{array}\right] \sim RREF \sim [U|\mathbf{b_u}] = \left[\begin{array}{ccc|c} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{array}\right]$$

The last row of RREF implies: $0x_1 + 0x_2 + 0x_3 = 1 \longrightarrow [A|\mathbf{b}] \sim$ *Inconsistent*.

udemy

Example: QR for Least Squares

Applying the QR factorization method:

$$\text{(1)} : r_{11} = ||\mathbf{a_1}|| = \sqrt{1 + 2^4 + 2^4 + 4^2} = 5$$

$$\mathbf{q_1} = \mathbf{a_1}/r_{11} = (1/5, 2/5, 2/5, 4/5)^T$$

udemy

Example: QR for Least Squares

Applying the QR factorization method:

$$\textcircled{1}: r_{11} = ||\mathbf{a_1}|| = \sqrt{1+2^4+2^4+4^2} = 5$$

$$\mathbf{q_1} = \mathbf{a_1}/r_{11} = (1/5, 2/5, 2/5, 4/5)^T$$

$$\textcircled{2}: r_{12} = \mathbf{q_1}^T\mathbf{a_2} = -2$$

$$p_1 = r_{12}\mathbf{q_1} = (-2/5, -4/5, -4/5, -8/5)^T$$

$$\mathbf{a_2} - \mathbf{p_1} = (-8/5, 4/5, -16/5, 8/5)^T$$

$$r_{22} = ||\mathbf{a_2} - \mathbf{p_1}|| = 4$$

$$\mathbf{q_2} = \frac{\mathbf{a_2} - \mathbf{p_1}}{r_{22}} = (-2/5, 1/5, -4/5, 2/5)^T$$

Example: QR for Least Squares

Applying the QR factorization method:

$$\textcircled{1}: r_{11} = ||\mathbf{a_1}|| = \sqrt{1 + 2^4 + 2^4 + 4^2} = 5$$

$$\mathbf{q_1} = \mathbf{a_1}/r_{11} = (1/5, 2/5, 2/5, 4/5)^T$$

$$\textcircled{2}: r_{12} = \mathbf{q_1}^T \mathbf{a_2} = -2$$

$$p_1 = r_{12}\mathbf{q_1} = (-2/5, -4/5, -4/5, -8/5)^T$$

$$\mathbf{a_2} - \mathbf{p_1} = (-8/5, 4/5, -16/5, 8/5)^T$$

$$r_{22} = ||\mathbf{a_2} - \mathbf{p_1}|| = 4$$

$$\mathbf{q_2} = \frac{\mathbf{a_2} - \mathbf{p_1}}{r_{22}} = (-2/5, 1/5, -4/5, 2/5)^T$$

$$\textcircled{3}: r_{13} = \mathbf{q_1}^T \mathbf{a_3} = 1 \quad r_{23} = \mathbf{q_2}^T \mathbf{a_3} = -1$$

$$\mathbf{p_2} = r_{13}\mathbf{q_1} + r_{23}\mathbf{q_2} = \mathbf{q_1} - \mathbf{q_2} = (3/5, 1/5, 6/5, 2/5)^T$$

$$\mathbf{a_3} - \mathbf{p_2} = (-8/5, 4/5, 4/5, -2/5)^T$$

$$r_{33} = ||\mathbf{a_3} - \mathbf{p_2}|| = 2$$

$$\mathbf{q_3} = \frac{\mathbf{a_3} - \mathbf{p_2}}{r_{33}} = (-4/5, 2/5, 2/5, -1/5)^T$$

Example: QR for Least Squares

Collecting r_{ij} to form an Upper Triangular Matrix and $\mathbf{q_j}$ to form an Orthogonal Matrix:

$$A = QR = \begin{bmatrix} 1 & -2 & -1 \\ 2 & 0 & 1 \\ 2 & -4 & 2 \\ 4 & 0 & 0 \end{bmatrix} = \begin{bmatrix} 1/5 & -2/5 & -4/5 \\ 2/5 & 1/5 & 2/5 \\ 2/5 & -4/5 & 2/5 \\ 4/5 & 2/5 & -1/5 \end{bmatrix} \begin{bmatrix} 5 & -2 & 1 \\ 0 & 4 & -1 \\ 0 & 0 & 2 \end{bmatrix}$$

Example: QR for Least Squares

Collecting r_{ij} to form an Upper Triangular Matrix and $\mathbf{q_j}$ to form an Orthogonal Matrix:

$$A = QR = \begin{bmatrix} 1 & -2 & -1 \\ 2 & 0 & 1 \\ 2 & -4 & 2 \\ 4 & 0 & 0 \end{bmatrix} = \begin{bmatrix} 1/5 & -2/5 & -4/5 \\ 2/5 & 1/5 & 2/5 \\ 2/5 & -4/5 & 2/5 \\ 4/5 & 2/5 & -1/5 \end{bmatrix} \begin{bmatrix} 5 & -2 & 1 \\ 0 & 4 & -1 \\ 0 & 0 & 2 \end{bmatrix}$$

One only needs solve: $R\hat{\mathbf{x}} = Q^T\mathbf{b}$

$$\begin{aligned} 5x_1 - 2x_2 + x_3 &= -1 \\ 4x_2 - x_3 &= -1 \\ 2x_3 &= 2 \end{aligned}$$

udemy

Example: QR for Least Squares

Collecting r_{ij} to form an Upper Triangular Matrix and $\mathbf{q_j}$ to form an Orthogonal Matrix:

$$A = QR = \begin{bmatrix} 1 & -2 & -1 \\ 2 & 0 & 1 \\ 2 & -4 & 2 \\ 4 & 0 & 0 \end{bmatrix} = \begin{bmatrix} 1/5 & -2/5 & -4/5 \\ 2/5 & 1/5 & 2/5 \\ 2/5 & -4/5 & 2/5 \\ 4/5 & 2/5 & -1/5 \end{bmatrix} \begin{bmatrix} 5 & -2 & 1 \\ 0 & 4 & -1 \\ 0 & 0 & 2 \end{bmatrix}$$

One only needs solve: $R\hat{\mathbf{x}} = Q^T\mathbf{b}$

$$\begin{aligned} 5x_1 - 2x_2 + x_3 &= -1 \\ 4x_2 - x_3 &= -1 \\ 2x_3 &= 2 \end{aligned}$$

After back substitution, one obtains:

$$\hat{\mathbf{x}} = \begin{bmatrix} -2/5 \\ 0 \\ 1 \end{bmatrix}$$

udemy

References I

[1] David Harville. *Matrix Algebra From a Statistician's Perspective*. New York: Springer-Verlag, 1997.

[2] Leon Stephen. *Linear Algebra with Applications (9th Edition) (Featured Titles for Linear Algebra*. London, England: Pearson, 2014.

udemy