Math 221: LINEAR ALGEBRA

Chapter 5. Vector Space $\mathbb{R}^n$

§5-3. Orthogonality

Le Chen[1]

Emory University, 2020 Fall

(last updated on 10/26/2020)

[1] Slides are adapted from those by Karen Seyffarth from University of Calgary.

The Dot Product

The Cauchy Inequality

Orthogonality

Orthogonality and Independence

Fourier Expansion

Dot Product

Dot Product

Definitions

Let $\vec{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$ and $\vec{y} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}$ be vectors in $\mathbb{R}^n$.

Dot Product

Definitions

Let $\vec{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$ and $\vec{y} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}$ be vectors in $\mathbb{R}^n$.

1. The dot product of $\vec{x}$ and $\vec{y}$ is

$$\vec{x} \cdot \vec{y} = x_1 y_1 + x_2 y_2 + \cdots x_n y_n = \vec{x}^T \vec{y}.$$

Dot Product

Definitions

Let $\vec{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$ and $\vec{y} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}$ be vectors in $\mathbb{R}^n$.

1. The dot product of $\vec{x}$ and $\vec{y}$ is

$$\vec{x} \cdot \vec{y} = x_1 y_1 + x_2 y_2 + \cdots x_n y_n = \vec{x}^T \vec{y}.$$

2. The length or norm of $\vec{x}$, denoted $||\vec{x}||$ is

$$||\vec{x}|| = \sqrt{x_1^2 + x_2^2 \cdots + x_n^2} = \sqrt{\vec{x} \cdot \vec{x}} = \sqrt{\vec{x}^T \vec{x}}.$$

Dot Product

Definitions

Let $\vec{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$ and $\vec{y} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}$ be vectors in $\mathbb{R}^n$.

1. The dot product of $\vec{x}$ and $\vec{y}$ is

$$\vec{x} \cdot \vec{y} = x_1 y_1 + x_2 y_2 + \cdots x_n y_n = \vec{x}^T \vec{y}.$$

2. The length or norm of $\vec{x}$, denoted $||\vec{x}||$ is

$$||\vec{x}|| = \sqrt{x_1^2 + x_2^2 \cdots + x_n^2} = \sqrt{\vec{x} \cdot \vec{x}} = \sqrt{\vec{x}^T \vec{x}}.$$

3. $\vec{x}$ is called a unit vector if $||\vec{x}|| = 1$.

Theorem (Properties of length and the dot product)

Let $\vec{x}, \vec{y}, \vec{z} \in \mathbb{R}^n$, and let $a \in \mathbb{R}$. Then

Theorem (Properties of length and the dot product)

Let $\vec{x}, \vec{y}, \vec{z} \in \mathbb{R}^n$, and let $a \in \mathbb{R}$. Then

1. $\vec{x} \cdot \vec{y} = \vec{y} \cdot \vec{x}$ (the dot product is commutative)

Theorem (Properties of length and the dot product)

Let $\vec{x}, \vec{y}, \vec{z} \in \mathbb{R}^n$, and let $a \in \mathbb{R}$. Then

1. $\vec{x} \cdot \vec{y} = \vec{y} \cdot \vec{x}$ (the dot product is commutative)
2. $\vec{x} \cdot (\vec{y} + \vec{z}) = \vec{x} \cdot \vec{y} + \vec{x} \cdot \vec{z}$ (the dot product distributes over addition)

Theorem (Properties of length and the dot product)

Let $\vec{x}, \vec{y}, \vec{z} \in \mathbb{R}^n$, and let $a \in \mathbb{R}$. Then

1. $\vec{x} \cdot \vec{y} = \vec{y} \cdot \vec{x}$ (the dot product is commutative)
2. $\vec{x} \cdot (\vec{y} + \vec{z}) = \vec{x} \cdot \vec{y} + \vec{x} \cdot \vec{z}$ (the dot product distributes over addition)
3. $(a\vec{x}) \cdot \vec{y} = a(\vec{x} \cdot \vec{y}) = \vec{x} \cdot (a\vec{y})$

Theorem (Properties of length and the dot product)

Let $\vec{x}, \vec{y}, \vec{z} \in \mathbb{R}^n$, and let $a \in \mathbb{R}$. Then

1. $\vec{x} \cdot \vec{y} = \vec{y} \cdot \vec{x}$ (the dot product is commutative)
2. $\vec{x} \cdot (\vec{y} + \vec{z}) = \vec{x} \cdot \vec{y} + \vec{x} \cdot \vec{z}$ (the dot product distributes over addition)
3. $(a\vec{x}) \cdot \vec{y} = a(\vec{x} \cdot \vec{y}) = \vec{x} \cdot (a\vec{y})$
4. $||\vec{x}||^2 = \vec{x} \cdot \vec{x}$.

Theorem (Properties of length and the dot product)

Let $\vec{x}, \vec{y}, \vec{z} \in \mathbb{R}^n$, and let $a \in \mathbb{R}$. Then

1. $\vec{x} \cdot \vec{y} = \vec{y} \cdot \vec{x}$ (the dot product is commutative)
2. $\vec{x} \cdot (\vec{y} + \vec{z}) = \vec{x} \cdot \vec{y} + \vec{x} \cdot \vec{z}$ (the dot product distributes over addition)
3. $(a\vec{x}) \cdot \vec{y} = a(\vec{x} \cdot \vec{y}) = \vec{x} \cdot (a\vec{y})$
4. $||\vec{x}||^2 = \vec{x} \cdot \vec{x}$.
5. $||\vec{x}|| \geq 0$ with equality if and only if $\vec{x} = \vec{0}_n$.

Theorem (Properties of length and the dot product)

Let $\vec{x}, \vec{y}, \vec{z} \in \mathbb{R}^n$, and let $a \in \mathbb{R}$. Then

1. $\vec{x} \cdot \vec{y} = \vec{y} \cdot \vec{x}$ (the dot product is commutative)
2. $\vec{x} \cdot (\vec{y} + \vec{z}) = \vec{x} \cdot \vec{y} + \vec{x} \cdot \vec{z}$ (the dot product distributes over addition)
3. $(a\vec{x}) \cdot \vec{y} = a(\vec{x} \cdot \vec{y}) = \vec{x} \cdot (a\vec{y})$
4. $||\vec{x}||^2 = \vec{x} \cdot \vec{x}$.
5. $||\vec{x}|| \geq 0$ with equality if and only if $\vec{x} = \vec{0}_n$.
6. $||a\vec{x}|| = |a| \; ||\vec{x}||$.

Example

Let $\vec{\mathrm{x}}, \vec{\mathrm{y}} \in \mathbb{R}^{\mathrm{n}}$. Then

$$\begin{aligned} ||\vec{\mathrm{x}}+\vec{\mathrm{y}}||^2 &= (\vec{\mathrm{x}}+\vec{\mathrm{y}}) \cdot (\vec{\mathrm{x}}+\vec{\mathrm{y}}) \\ &= \vec{\mathrm{x}} \cdot \vec{\mathrm{x}} + \vec{\mathrm{x}} \cdot \vec{\mathrm{y}} + \vec{\mathrm{y}} \cdot \vec{\mathrm{x}} + \vec{\mathrm{y}} \cdot \vec{\mathrm{y}} \\ &= \vec{\mathrm{x}} \cdot \vec{\mathrm{x}} + 2(\vec{\mathrm{x}} \cdot \vec{\mathrm{y}}) + \vec{\mathrm{y}} \cdot \vec{\mathrm{y}} \\ &= ||\vec{\mathrm{x}}||^2 + 2(\vec{\mathrm{x}} \cdot \vec{\mathrm{y}}) + ||\vec{\mathrm{y}}||^2. \end{aligned}$$

Problem

Let $\{\vec{f}_1, \vec{f}_2, \ldots, \vec{f}_k\} \in \mathbb{R}^n$ and suppose $\mathbb{R}^n = \text{span}\{\vec{f}_1, \vec{f}_2, \ldots, \vec{f}_k\}$. Furthermore, suppose that there exists a vector $\vec{x} \in \mathbb{R}^n$ for which $\vec{x} \cdot \vec{f}_j = 0$ for all j, $1 \leq j \leq k$. Show that $\vec{x} = \vec{0}_n$.

Problem

Let $\{\vec{f}_1, \vec{f}_2, \ldots, \vec{f}_k\} \in \mathbb{R}^n$ and suppose $\mathbb{R}^n = \text{span}\{\vec{f}_1, \vec{f}_2, \ldots, \vec{f}_k\}$. Furthermore, suppose that there exists a vector $\vec{x} \in \mathbb{R}^n$ for which $\vec{x} \cdot \vec{f}_j = 0$ for all j, $1 \leq j \leq k$. Show that $\vec{x} = \vec{0}_n$.

Proof.

Write $\vec{x} = t_1\vec{f}_1 + t_2\vec{f}_2 + \cdots + t_k\vec{f}_k$ for some $t_1, t_2, \ldots, t_k \in \mathbb{R}$ (this is possible because $\vec{f}_1, \vec{f}_2, \ldots, \vec{f}_k$ span $\mathbb{R}^n$, is this representation unique?).

Problem

Let $\{\vec{f}_1, \vec{f}_2, \ldots, \vec{f}_k\} \in \mathbb{R}^n$ and suppose $\mathbb{R}^n = \mathrm{span}\{\vec{f}_1, \vec{f}_2, \ldots, \vec{f}_k\}$. Furthermore, suppose that there exists a vector $\vec{x} \in \mathbb{R}^n$ for which $\vec{x} \cdot \vec{f}_j = 0$ for all j, $1 \leq j \leq k$. Show that $\vec{x} = \vec{0}_n$.

Proof.

Write $\vec{x} = t_1\vec{f}_1 + t_2\vec{f}_2 + \cdots + t_k\vec{f}_k$ for some $t_1, t_2, \ldots, t_k \in \mathbb{R}$ (this is possible because $\vec{f}_1, \vec{f}_2, \ldots, \vec{f}_k$ span $\mathbb{R}^n$, is this representation unique?). Then

$$||\vec{x}||^2 = \vec{x} \cdot \vec{x}$$

Problem

Let $\{\vec{f}_1, \vec{f}_2, \ldots, \vec{f}_k\} \in \mathbb{R}^n$ and suppose $\mathbb{R}^n = \text{span}\{\vec{f}_1, \vec{f}_2, \ldots, \vec{f}_k\}$. Furthermore, suppose that there exists a vector $\vec{x} \in \mathbb{R}^n$ for which $\vec{x} \cdot \vec{f}_j = 0$ for all j, $1 \leq j \leq k$. Show that $\vec{x} = \vec{0}_n$.

Proof.

Write $\vec{x} = t_1\vec{f}_1 + t_2\vec{f}_2 + \cdots + t_k\vec{f}_k$ for some $t_1, t_2, \ldots, t_k \in \mathbb{R}$ (this is possible because $\vec{f}_1, \vec{f}_2, \ldots, \vec{f}_k$ span $\mathbb{R}^n$, is this representation unique?). Then

$$\begin{aligned} ||\vec{x}||^2 &= \vec{x} \cdot \vec{x} \\ &= \vec{x} \cdot (t_1\vec{f}_1 + t_2\vec{f}_2 + \cdots + t_k\vec{f}_k) \end{aligned}$$

Problem

Let $\{\vec{f}_1, \vec{f}_2, \ldots, \vec{f}_k\} \in \mathbb{R}^n$ and suppose $\mathbb{R}^n = \mathrm{span}\{\vec{f}_1, \vec{f}_2, \ldots, \vec{f}_k\}$. Furthermore, suppose that there exists a vector $\vec{x} \in \mathbb{R}^n$ for which $\vec{x} \cdot \vec{f}_j = 0$ for all j, $1 \leq j \leq k$. Show that $\vec{x} = \vec{0}_n$.

Proof.

Write $\vec{x} = t_1\vec{f}_1 + t_2\vec{f}_2 + \cdots + t_k\vec{f}_k$ for some $t_1, t_2, \ldots, t_k \in \mathbb{R}$ (this is possible because $\vec{f}_1, \vec{f}_2, \ldots, \vec{f}_k$ span $\mathbb{R}^n$, is this representation unique?). Then

$$\begin{aligned} ||\vec{x}||^2 &= \vec{x} \cdot \vec{x} \\ &= \vec{x} \cdot (t_1\vec{f}_1 + t_2\vec{f}_2 + \cdots + t_k\vec{f}_k) \\ &= \vec{x} \cdot (t_1\vec{f}_1) + \vec{x} \cdot (t_2\vec{f}_2) + \cdots + \vec{x} \cdot (t_k\vec{f}_k) \end{aligned}$$

Problem

Let $\{\vec{f}_1, \vec{f}_2, \ldots, \vec{f}_k\} \in \mathbb{R}^n$ and suppose $\mathbb{R}^n = \text{span}\{\vec{f}_1, \vec{f}_2, \ldots, \vec{f}_k\}$. Furthermore, suppose that there exists a vector $\vec{x} \in \mathbb{R}^n$ for which $\vec{x} \cdot \vec{f}_j = 0$ for all j, $1 \leq j \leq k$. Show that $\vec{x} = \vec{0}_n$.

Proof.

Write $\vec{x} = t_1\vec{f}_1 + t_2\vec{f}_2 + \cdots + t_k\vec{f}_k$ for some $t_1, t_2, \ldots, t_k \in \mathbb{R}$ (this is possible because $\vec{f}_1, \vec{f}_2, \ldots, \vec{f}_k$ span $\mathbb{R}^n$, is this representation unique?). Then

$$\begin{aligned}
||\vec{x}||^2 &= \vec{x} \cdot \vec{x} \\
&= \vec{x} \cdot (t_1\vec{f}_1 + t_2\vec{f}_2 + \cdots + t_k\vec{f}_k) \\
&= \vec{x} \cdot (t_1\vec{f}_1) + \vec{x} \cdot (t_2\vec{f}_2) + \cdots + \vec{x} \cdot (t_k\vec{f}_k) \\
&= t_1(\vec{x} \cdot \vec{f}_1) + t_2(\vec{x} \cdot \vec{f}_2) + \cdots + t_k(\vec{x} \cdot \vec{f}_k)
\end{aligned}$$

Problem

Let $\{\vec{f}_1, \vec{f}_2, \ldots, \vec{f}_k\} \in \mathbb{R}^n$ and suppose $\mathbb{R}^n = \text{span}\{\vec{f}_1, \vec{f}_2, \ldots, \vec{f}_k\}$. Furthermore, suppose that there exists a vector $\vec{x} \in \mathbb{R}^n$ for which $\vec{x} \cdot \vec{f}_j = 0$ for all j, $1 \leq j \leq k$. Show that $\vec{x} = \vec{0}_n$.

Proof.

Write $\vec{x} = t_1\vec{f}_1 + t_2\vec{f}_2 + \cdots + t_k\vec{f}_k$ for some $t_1, t_2, \ldots, t_k \in \mathbb{R}$ (this is possible because $\vec{f}_1, \vec{f}_2, \ldots, \vec{f}_k$ span $\mathbb{R}^n$, is this representation unique?). Then

$$\begin{aligned}
||\vec{x}||^2 &= \vec{x} \cdot \vec{x} \\
&= \vec{x} \cdot (t_1\vec{f}_1 + t_2\vec{f}_2 + \cdots + t_k\vec{f}_k) \\
&= \vec{x} \cdot (t_1\vec{f}_1) + \vec{x} \cdot (t_2\vec{f}_2) + \cdots + \vec{x} \cdot (t_k\vec{f}_k) \\
&= t_1(\vec{x} \cdot \vec{f}_1) + t_2(\vec{x} \cdot \vec{f}_2) + \cdots + t_k(\vec{x} \cdot \vec{f}_k) \\
&= t_1(0) + t_2(0) + \cdots + t_k(0) = 0.
\end{aligned}$$

Problem

Let $\{\vec{f}_1, \vec{f}_2, \ldots, \vec{f}_k\} \in \mathbb{R}^n$ and suppose $\mathbb{R}^n = \text{span}\{\vec{f}_1, \vec{f}_2, \ldots, \vec{f}_k\}$. Furthermore, suppose that there exists a vector $\vec{x} \in \mathbb{R}^n$ for which $\vec{x} \cdot \vec{f}_j = 0$ for all j, $1 \leq j \leq k$. Show that $\vec{x} = \vec{0}_n$.

Proof.

Write $\vec{x} = t_1\vec{f}_1 + t_2\vec{f}_2 + \cdots + t_k\vec{f}_k$ for some $t_1, t_2, \ldots, t_k \in \mathbb{R}$ (this is possible because $\vec{f}_1, \vec{f}_2, \ldots, \vec{f}_k$ span $\mathbb{R}^n$, is this representation unique?). Then

$$\begin{aligned}
||\vec{x}||^2 &= \vec{x} \cdot \vec{x} \\
&= \vec{x} \cdot (t_1\vec{f}_1 + t_2\vec{f}_2 + \cdots + t_k\vec{f}_k) \\
&= \vec{x} \cdot (t_1\vec{f}_1) + \vec{x} \cdot (t_2\vec{f}_2) + \cdots + \vec{x} \cdot (t_k\vec{f}_k) \\
&= t_1(\vec{x} \cdot \vec{f}_1) + t_2(\vec{x} \cdot \vec{f}_2) + \cdots + t_k(\vec{x} \cdot \vec{f}_k) \\
&= t_1(0) + t_2(0) + \cdots + t_k(0) = 0.
\end{aligned}$$

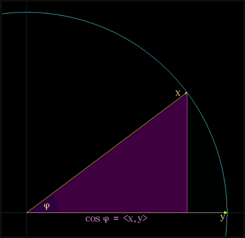
Since $||\vec{x}||^2 = 0$, $||\vec{x}|| = 0$. By the previous theorem, $||\vec{x}|| = 0$ if and only if $\vec{x} = \vec{0}_n$. Therefore, $\vec{x} = \vec{0}_n$. ■

Cauchy-Schwartz Inequality

Cauchy-Schwartz Inequality

Theorem (Cauchy-Schwartz Inequality)

If $\vec{x}, \vec{y} \in \mathbb{R}^n$, then $|\vec{x} \cdot \vec{y}| \leq ||\vec{x}||\ ||\vec{y}||$ with equality if and only if $\{\vec{x}, \vec{y}\}$ is linearly dependent.



$$\left| \frac{\vec{x}}{||\vec{x}||} \cdot \frac{\vec{y}}{||\vec{y}||} \right| \leq 1$$

$$\{\vec{x}, \vec{y}\} \text{ is linearly dependent} \quad \Leftrightarrow \quad \vec{x} = t\vec{y}, \quad \text{for some } t \in \mathbb{R}.$$

Proof.

Let $\vec{x}, \vec{y} \in \mathbb{R}^n$ and $t \in \mathbb{R}$. Then

$$\begin{aligned} 0 \leq ||t\vec{x} + \vec{y}||^2 &= (t\vec{x} + \vec{y}) \cdot (t\vec{x} + \vec{y}) \\ &= t^2\vec{x} \cdot \vec{x} + 2t\vec{x} \cdot \vec{y} + \vec{y} \cdot \vec{y} \\ &= t^2||\vec{x}||^2 + 2t(\vec{x} \cdot \vec{y}) + ||\vec{y}||^2. \end{aligned}$$

The quadratic $t^2||\vec{x}||^2 + 2t(\vec{x} \cdot \vec{y}) + ||\vec{y}||^2$ in t is always nonnegative, so it does not have distinct real roots. Thus, if we use the quadratic formula to solve for t, the discriminant must be non-positive, i.e.,

$$\Delta = (2\vec{x} \cdot \vec{y})^2 - 4||\vec{x}||^2||\vec{y}||^2 \leq 0$$

Therefore, $(2\vec{x} \cdot \vec{y})^2 \leq 4||\vec{x}||^2||\vec{y}||^2$. Since both sides of the inequality are nonnegative, we can take (positive) square roots of both sides:

$$|2\vec{x} \cdot \vec{y}| \leq 2||\vec{x}||\ ||\vec{y}||$$

Therefore, $|\vec{x} \cdot \vec{y}| \leq ||\vec{x}||\ ||\vec{y}||$. What remains is to show that $|\vec{x} \cdot \vec{y}| = ||\vec{x}||\ ||\vec{y}||$ if and only if $\{\vec{x}, \vec{y}\}$ is linearly dependent. ■

Proof. (continued)

First suppose that $\{\vec{x}, \vec{y}\}$ is dependent. Then by symmetry (of $\vec{x}$ and $\vec{y}$), $\vec{x} = k\vec{y}$ for some $k \in \mathbb{R}$. Hence

$$|\vec{x}\cdot\vec{y}| = |(k\vec{y})\cdot\vec{y}| = |k|\, |\vec{y}\cdot\vec{y}| = |k|\, ||\vec{y}||^2, \quad \text{and} \quad ||\vec{x}||\, ||\vec{y}|| = ||k\vec{y}||\, ||\vec{y}|| = |k|\, ||\vec{y}||^2,$$

so $|\vec{x} \cdot \vec{y}| = ||\vec{x}||\, ||\vec{y}||$.

Conversely, suppose $\{\vec{x}, \vec{y}\}$ is independent; then $t\vec{x} + \vec{y} \neq \vec{0}_n$ for all $t \in \mathbb{R}$, so $||t\vec{x} + \vec{y}||^2 > 0$ for all $t \in \mathbb{R}$. Thus the quadratic

$$t^2||\vec{x}||^2 + 2t(\vec{x} \cdot \vec{y}) + ||\vec{y}||^2 > 0$$

so has no real roots. It follows that the the discriminant is negative, i.e.,

$$(2\vec{x} \cdot \vec{y})^2 - 4||\vec{x}||^2||\vec{y}||^2 < 0.$$

Therefore, $(2\vec{x} \cdot \vec{y})^2 < 4||\vec{x}||^2||\vec{y}||^2$; taking square roots of both sides (they are both nonnegative) and dividing by two gives us

$$|\vec{x} \cdot \vec{y}| < ||\vec{x}||\, ||\vec{y}||,$$

showing that equality is impossible. ■

Corollary (Triangle Inequality I)

If $\vec{x}, \vec{y} \in \mathbb{R}^n$, then $||\vec{x} + \vec{y}|| \leq ||\vec{x}|| + ||\vec{y}||$.

Corollary (Triangle Inequality I)

If $\vec{x}, \vec{y} \in \mathbb{R}^n$, then $||\vec{x} + \vec{y}|| \leq ||\vec{x}|| + ||\vec{y}||$.

Proof.

$$||\vec{x} + \vec{y}||^2 = (\vec{x} + \vec{y}) \cdot (\vec{x} + \vec{y})$$

Corollary (Triangle Inequality I)

If $\vec{x}, \vec{y} \in \mathbb{R}^n$, then $||\vec{x}+\vec{y}|| \leq ||\vec{x}|| + ||\vec{y}||$.

Proof.

$$\begin{aligned} ||\vec{x}+\vec{y}||^2 &= (\vec{x}+\vec{y}) \cdot (\vec{x}+\vec{y}) \\ &= \vec{x}\cdot\vec{x} + 2\vec{x}\cdot\vec{y} + \vec{y}\cdot\vec{y} \end{aligned}$$

Corollary (Triangle Inequality I)

If $\vec{x}, \vec{y} \in \mathbb{R}^n$, then $||\vec{x}+\vec{y}|| \leq ||\vec{x}|| + ||\vec{y}||$.

Proof.

$$\begin{aligned} ||\vec{x}+\vec{y}||^2 &= (\vec{x}+\vec{y}) \cdot (\vec{x}+\vec{y}) \\ &= \vec{x} \cdot \vec{x} + 2\vec{x} \cdot \vec{y} + \vec{y} \cdot \vec{y} \\ &= ||\vec{x}||^2 + 2\vec{x} \cdot \vec{y} + ||\vec{y}||^2 \end{aligned}$$

Corollary (Triangle Inequality I)

If $\vec{x}, \vec{y} \in \mathbb{R}^n$, then $||\vec{x} + \vec{y}|| \leq ||\vec{x}|| + ||\vec{y}||$.

Proof.

$$\begin{aligned}
||\vec{x} + \vec{y}||^2 &= (\vec{x} + \vec{y}) \cdot (\vec{x} + \vec{y}) \\
&= \vec{x} \cdot \vec{x} + 2\vec{x} \cdot \vec{y} + \vec{y} \cdot \vec{y} \\
&= ||\vec{x}||^2 + 2\vec{x} \cdot \vec{y} + ||\vec{y}||^2 \\
&\leq ||\vec{x}||^2 + 2||\vec{x}|| \; ||\vec{y}|| + ||\vec{y}||^2 \text{ by the Cauchy Inequality}
\end{aligned}$$

Corollary (Triangle Inequality I)

If $\vec{x}, \vec{y} \in \mathbb{R}^n$, then $||\vec{x}+\vec{y}|| \leq ||\vec{x}|| + ||\vec{y}||$.

Proof.

$$
\begin{aligned}
||\vec{x}+\vec{y}||^2 &= (\vec{x}+\vec{y}) \cdot (\vec{x}+\vec{y}) \\
&= \vec{x} \cdot \vec{x} + 2\vec{x} \cdot \vec{y} + \vec{y} \cdot \vec{y} \\
&= ||\vec{x}||^2 + 2\vec{x} \cdot \vec{y} + ||\vec{y}||^2 \\
&\leq ||\vec{x}||^2 + 2||\vec{x}|| \; ||\vec{y}|| + ||\vec{y}||^2 \text{ by the Cauchy Inequality} \\
&= (||\vec{x}|| + ||\vec{y}||)^2.
\end{aligned}
$$

Corollary (Triangle Inequality I)

If $\vec{x}, \vec{y} \in \mathbb{R}^n$, then $||\vec{x}+\vec{y}|| \leq ||\vec{x}|| + ||\vec{y}||$.

Proof.

$$
\begin{aligned}
||\vec{x}+\vec{y}||^2 &= (\vec{x}+\vec{y})\cdot(\vec{x}+\vec{y}) \\
&= \vec{x}\cdot\vec{x} + 2\vec{x}\cdot\vec{y} + \vec{y}\cdot\vec{y} \\
&= ||\vec{x}||^2 + 2\vec{x}\cdot\vec{y} + ||\vec{y}||^2 \\
&\leq ||\vec{x}||^2 + 2||\vec{x}||\ ||\vec{y}|| + ||\vec{y}||^2 \text{ by the Cauchy Inequality} \\
&= (||\vec{x}|| + ||\vec{y}||)^2.
\end{aligned}
$$

Since both sides of the inequality are nonnegative, we take (positive) square roots of both sides:

$$||\vec{x}+\vec{y}|| \leq ||\vec{x}|| + ||\vec{y}||.$$

■

Definition

If $\vec{x}, \vec{y} \in \mathbb{R}^n$, then the distance between $\vec{x}$ and $\vec{y}$ is defined as

$$d(\vec{x}, \vec{y}) = ||\vec{x} - \vec{y}||.$$

Definition

If $\vec{x}, \vec{y} \in \mathbb{R}^n$, then the distance between $\vec{x}$ and $\vec{y}$ is defined as

$$d(\vec{x}, \vec{y}) = ||\vec{x} - \vec{y}||.$$

Theorem (Properties of the distance function)

Let $\vec{x}, \vec{y}, \vec{z} \in \mathbb{R}^n$. Then

Definition

If $\vec{x}, \vec{y} \in \mathbb{R}^n$, then the distance between $\vec{x}$ and $\vec{y}$ is defined as

$$d(\vec{x}, \vec{y}) = ||\vec{x} - \vec{y}||.$$

Theorem (Properties of the distance function)

Let $\vec{x}, \vec{y}, \vec{z} \in \mathbb{R}^n$. Then

1. $d(\vec{x}, \vec{y}) \geq 0$.

Definition

If $\vec{x}, \vec{y} \in \mathbb{R}^n$, then the distance between $\vec{x}$ and $\vec{y}$ is defined as

$$d(\vec{x}, \vec{y}) = ||\vec{x} - \vec{y}||.$$

Theorem (Properties of the distance function)

Let $\vec{x}, \vec{y}, \vec{z} \in \mathbb{R}^n$. Then

1. $d(\vec{x}, \vec{y}) \geq 0$.
2. $d(\vec{x}, \vec{y}) = 0$ if and only if $\vec{x} = \vec{y}$.

Definition

If $\vec{x}, \vec{y} \in \mathbb{R}^n$, then the distance between $\vec{x}$ and $\vec{y}$ is defined as

$$d(\vec{x}, \vec{y}) = ||\vec{x} - \vec{y}||.$$

Theorem (Properties of the distance function)

Let $\vec{x}, \vec{y}, \vec{z} \in \mathbb{R}^n$. Then

1. $d(\vec{x}, \vec{y}) \geq 0$.
2. $d(\vec{x}, \vec{y}) = 0$ if and only if $\vec{x} = \vec{y}$.
3. $d(\vec{x}, \vec{y}) = d(\vec{y}, \vec{x})$.

Definition

If $\vec{x}, \vec{y} \in \mathbb{R}^n$, then the distance between $\vec{x}$ and $\vec{y}$ is defined as

$$d(\vec{x}, \vec{y}) = ||\vec{x} - \vec{y}||.$$

Theorem (Properties of the distance function)

Let $\vec{x}, \vec{y}, \vec{z} \in \mathbb{R}^n$. Then

1. $d(\vec{x}, \vec{y}) \geq 0$.
2. $d(\vec{x}, \vec{y}) = 0$ if and only if $\vec{x} = \vec{y}$.
3. $d(\vec{x}, \vec{y}) = d(\vec{y}, \vec{x})$.
4. $d(\vec{x}, \vec{z}) \leq d(\vec{x}, \vec{y}) + d(\vec{y}, \vec{z})$ (Triangle Inequality II).

Definition

If $\vec{x}, \vec{y} \in \mathbb{R}^n$, then the distance between $\vec{x}$ and $\vec{y}$ is defined as

$$d(\vec{x}, \vec{y}) = ||\vec{x} - \vec{y}||.$$

Theorem (Properties of the distance function)

Let $\vec{x}, \vec{y}, \vec{z} \in \mathbb{R}^n$. Then

1. $d(\vec{x}, \vec{y}) \geq 0$.
2. $d(\vec{x}, \vec{y}) = 0$ if and only if $\vec{x} = \vec{y}$.
3. $d(\vec{x}, \vec{y}) = d(\vec{y}, \vec{x})$.
4. $d(\vec{x}, \vec{z}) \leq d(\vec{x}, \vec{y}) + d(\vec{y}, \vec{z})$ (Triangle Inequality II).

Proof. (Proof of the Triangle Inequality II)

$$d(\vec{x}, \vec{z}) = ||\vec{x} - \vec{z}|| \quad = \quad ||(\vec{x} - \vec{y}) + (\vec{y} - \vec{z})||$$

Definition

If $\vec{x}, \vec{y} \in \mathbb{R}^n$, then the distance between $\vec{x}$ and $\vec{y}$ is defined as

$$d(\vec{x}, \vec{y}) = ||\vec{x} - \vec{y}||.$$

Theorem (Properties of the distance function)

Let $\vec{x}, \vec{y}, \vec{z} \in \mathbb{R}^n$. Then

1. $d(\vec{x}, \vec{y}) \geq 0$.
2. $d(\vec{x}, \vec{y}) = 0$ if and only if $\vec{x} = \vec{y}$.
3. $d(\vec{x}, \vec{y}) = d(\vec{y}, \vec{x})$.
4. $d(\vec{x}, \vec{z}) \leq d(\vec{x}, \vec{y}) + d(\vec{y}, \vec{z})$ (Triangle Inequality II).

Proof. (Proof of the Triangle Inequality II)

$$\begin{aligned} d(\vec{x}, \vec{z}) = ||\vec{x} - \vec{z}|| &= ||(\vec{x} - \vec{y}) + (\vec{y} - \vec{z})|| \\ &\leq ||\vec{x} - \vec{y}|| + ||\vec{y} - \vec{z}|| \text{ by Triangle Inequality I} \end{aligned}$$

Definition

If $\vec{x}, \vec{y} \in \mathbb{R}^n$, then the distance between $\vec{x}$ and $\vec{y}$ is defined as

$$d(\vec{x}, \vec{y}) = ||\vec{x} - \vec{y}||.$$

Theorem (Properties of the distance function)

Let $\vec{x}, \vec{y}, \vec{z} \in \mathbb{R}^n$. Then

1. $d(\vec{x}, \vec{y}) \geq 0$.
2. $d(\vec{x}, \vec{y}) = 0$ if and only if $\vec{x} = \vec{y}$.
3. $d(\vec{x}, \vec{y}) = d(\vec{y}, \vec{x})$.
4. $d(\vec{x}, \vec{z}) \leq d(\vec{x}, \vec{y}) + d(\vec{y}, \vec{z})$ (Triangle Inequality II).

Proof. (Proof of the Triangle Inequality II)

$$\begin{aligned} d(\vec{x}, \vec{z}) = ||\vec{x} - \vec{z}|| &= ||(\vec{x} - \vec{y}) + (\vec{y} - \vec{z})|| \\ &\leq ||\vec{x} - \vec{y}|| + ||\vec{y} - \vec{z}|| \text{ by Triangle Inequality I} \\ &= d(\vec{x}, \vec{y}) + d(\vec{y}, \vec{z}). \end{aligned}$$

∎

Orthogonality

Orthogonality

Definitions

- Let $\vec{x}, \vec{y} \in \mathbb{R}^n$. We say that two vectors $\vec{x}$ and $\vec{y}$ are orthogonal if $\vec{x} \cdot \vec{y} = 0$.

Orthogonality

Definitions

- Let $\vec{x}, \vec{y} \in \mathbb{R}^n$. We say that two vectors $\vec{x}$ and $\vec{y}$ are orthogonal if $\vec{x} \cdot \vec{y} = 0$.
- More generally, $X = \{\vec{x}_1, \vec{x}_2, \ldots, \vec{x}_k\} \subseteq \mathbb{R}^n$ is an orthogonal set if each $\vec{x}_i$ is nonzero, and every pair of distinct vectors of X is orthogonal, i.e., $\vec{x}_i \cdot \vec{x}_j = 0$ for all $i \neq j$, $1 \leq i, j \leq k$.

Orthogonality

Definitions

- Let $\vec{x}, \vec{y} \in \mathbb{R}^n$. We say that two vectors $\vec{x}$ and $\vec{y}$ are orthogonal if $\vec{x} \cdot \vec{y} = 0$.
- More generally, $X = \{\vec{x}_1, \vec{x}_2, \ldots, \vec{x}_k\} \subseteq \mathbb{R}^n$ is an orthogonal set if each $\vec{x}_i$ is nonzero, and every pair of distinct vectors of X is orthogonal, i.e., $\vec{x}_i \cdot \vec{x}_j = 0$ for all $i \neq j$, $1 \leq i, j \leq k$.
- A set $X = \{\vec{x}_1, \vec{x}_2, \ldots, \vec{x}_k\} \subseteq \mathbb{R}^n$ is an orthonormal set if X is an orthogonal set of unit vectors, i.e., $||\vec{x}_i|| = 1$ for all i, $1 \leq i \leq k$.

Examples

1. The standard basis $\{\vec{e}_1, \cdots, \vec{e}_n\}$ of $\mathbb{R}^n$ is an orthonormal set (and hence an orthogonal set).

Examples

1. The standard basis $\{\vec{e}_1, \cdots, \vec{e}_n\}$ of $\mathbb{R}^n$ is an orthonormal set (and hence an orthogonal set).

2. $$\left\{ \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ -1 \\ -1 \end{bmatrix}, \begin{bmatrix} 1 \\ -1 \\ 1 \\ -1 \end{bmatrix} \right\}$$

   is an orthogonal (but not orthonormal) subset of $\mathbb{R}^4$.

Examples

1. The standard basis $\{\vec{e}_1, \cdots, \vec{e}_n\}$ of $\mathbb{R}^n$ is an orthonormal set (and hence an orthogonal set).

2. $$\left\{ \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ -1 \\ -1 \end{bmatrix}, \begin{bmatrix} 1 \\ -1 \\ 1 \\ -1 \end{bmatrix} \right\}$$

   is an orthogonal (but not orthonormal) subset of $\mathbb{R}^4$.

3. If $\{\vec{x}_1, \vec{x}_2, \ldots, \vec{x}_k\}$ is an orthogonal subset of $\mathbb{R}^n$ and $p \neq 0$, then $\{p\vec{x}_1, p\vec{x}_2, \ldots, p\vec{x}_k\}$ is an orthogonal subset of $\mathbb{R}^n$.

Examples

1. The standard basis $\{\vec{e}_1, \cdots, \vec{e}_n\}$ of $\mathbb{R}^n$ is an orthonormal set (and hence an orthogonal set).

2. $$\left\{ \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ -1 \\ -1 \end{bmatrix}, \begin{bmatrix} 1 \\ -1 \\ 1 \\ -1 \end{bmatrix} \right\}$$
   is an orthogonal (but not orthonormal) subset of $\mathbb{R}^4$.

3. If $\{\vec{x}_1, \vec{x}_2, \ldots, \vec{x}_k\}$ is an orthogonal subset of $\mathbb{R}^n$ and $p \neq 0$, then $\{p\vec{x}_1, p\vec{x}_2, \ldots, p\vec{x}_k\}$ is an orthogonal subset of $\mathbb{R}^n$.

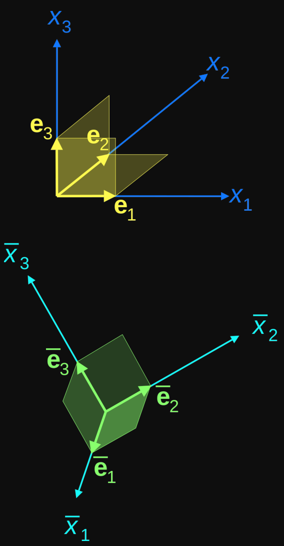
4. $$\left\{ \frac{1}{2}\begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}, \frac{1}{2}\begin{bmatrix} 1 \\ 1 \\ -1 \\ -1 \end{bmatrix}, \frac{1}{2}\begin{bmatrix} 1 \\ -1 \\ 1 \\ -1 \end{bmatrix} \right\}$$
   is an orthonormal subset of $\mathbb{R}^4$.

Definition

Normalizing an orthogonal set is the process of turning an orthogonal (but not orthonormal) set into an orthonormal set. If $\{\vec{x}_1, \vec{x}_2, \ldots, \vec{x}_k\}$ is an orthogonal subset of $\mathbb{R}^n$, then

$$\left\{ \frac{1}{||\vec{x}_1||}\vec{x}_1, \frac{1}{||\vec{x}_2||}\vec{x}_2, \ldots, \frac{1}{||\vec{x}_k||}\vec{x}_k \right\}$$

is an orthonormal set.



x_3
x_2
x_1
$\mathbf{e}_3$
$\mathbf{e}_2$
$\mathbf{e}_1$
$\bar{x}_3$
$\bar{x}_2$
$\bar{x}_1$
$\bar{\mathbf{e}}_3$
$\bar{\mathbf{e}}_2$
$\bar{\mathbf{e}}_1$

Problem

Verify that

$$\left\{ \begin{bmatrix} 1 \\ -1 \\ 2 \end{bmatrix}, \begin{bmatrix} 0 \\ 2 \\ 1 \end{bmatrix}, \begin{bmatrix} 5 \\ 1 \\ -2 \end{bmatrix} \right\}$$

is an orthogonal set, and normalize this set.

Solution

$$\begin{aligned}
\begin{bmatrix} 1 \\ -1 \\ 2 \end{bmatrix} \cdot \begin{bmatrix} 0 \\ 2 \\ 1 \end{bmatrix} &= 0 - 2 + 2 = 0, \\
\begin{bmatrix} 0 \\ 2 \\ 1 \end{bmatrix} \cdot \begin{bmatrix} 5 \\ 1 \\ -2 \end{bmatrix} &= 0 + 2 - 2 = 0, \\
\begin{bmatrix} 1 \\ -1 \\ 2 \end{bmatrix} \cdot \begin{bmatrix} 5 \\ 1 \\ -2 \end{bmatrix} &= 5 - 1 - 4 = 0,
\end{aligned}$$

proving that the set is orthogonal. Normalizing gives us the orthonormal set

$$\left\{ \frac{1}{\sqrt{6}} \begin{bmatrix} 1 \\ -1 \\ 2 \end{bmatrix}, \frac{1}{\sqrt{5}} \begin{bmatrix} 0 \\ 2 \\ 1 \end{bmatrix}, \frac{1}{\sqrt{30}} \begin{bmatrix} 5 \\ 1 \\ -2 \end{bmatrix} \right\}.$$

■

Theorem (Pythagoras' Theorem)

If $\{\vec{x}_1, \vec{x}_2, \ldots, \vec{x}_k\} \subseteq \mathbb{R}^n$ is orthogonal, then

$$||\vec{x}_1 + \vec{x}_2 + \cdots + \vec{x}_k||^2 = ||\vec{x}_1||^2 + ||\vec{x}_2||^2 + \cdots + ||\vec{x}_k||^2.$$

Theorem (Pythagoras' Theorem)

If $\{\vec{x}_1, \vec{x}_2, \ldots, \vec{x}_k\} \subseteq \mathbb{R}^n$ is orthogonal, then

$$||\vec{x}_1 + \vec{x}_2 + \cdots + \vec{x}_k||^2 = ||\vec{x}_1||^2 + ||\vec{x}_2||^2 + \cdots + ||\vec{x}_k||^2.$$

Proof.

Start with

$$||\vec{x}_1 + \vec{x}_2 + \cdots + \vec{x}_k||^2 \quad = \quad (\vec{x}_1 + \vec{x}_2 + \cdots + \vec{x}_k) \cdot (\vec{x}_1 + \vec{x}_2 + \cdots + \vec{x}_k)$$

Theorem (Pythagoras' Theorem)

If $\{\vec{x}_1, \vec{x}_2, \ldots, \vec{x}_k\} \subseteq \mathbb{R}^n$ is orthogonal, then

$$||\vec{x}_1 + \vec{x}_2 + \cdots + \vec{x}_k||^2 = ||\vec{x}_1||^2 + ||\vec{x}_2||^2 + \cdots + ||\vec{x}_k||^2.$$

Proof.

Start with

$$\begin{aligned}
||\vec{x}_1 + \vec{x}_2 + \cdots + \vec{x}_k||^2 &= (\vec{x}_1 + \vec{x}_2 + \cdots + \vec{x}_k) \cdot (\vec{x}_1 + \vec{x}_2 + \cdots + \vec{x}_k) \\
&= \ \ (\vec{x}_1 \cdot \vec{x}_1 + \vec{x}_1 \cdot \vec{x}_2 + \cdots + \vec{x}_1 \cdot \vec{x}_k) \\
&\quad +(\vec{x}_2 \cdot \vec{x}_1 + \vec{x}_2 \cdot \vec{x}_2 + \cdots + \vec{x}_2 \cdot \vec{x}_k) \\
&\qquad \vdots \qquad\qquad\qquad \vdots \qquad\qquad\qquad \vdots \\
&\quad +(\vec{x}_k \cdot \vec{x}_1 + \vec{x}_k \cdot \vec{x}_2 + \cdots + \vec{x}_k \cdot \vec{x}_k) \\
&= \vec{x}_1 \cdot \vec{x}_1 + \vec{x}_2 \cdot \vec{x}_2 + \cdots + \vec{x}_k \cdot \vec{x}_k
\end{aligned}$$

Theorem (Pythagoras' Theorem)

If $\{\vec{x}_1, \vec{x}_2, \ldots, \vec{x}_k\} \subseteq \mathbb{R}^n$ is orthogonal, then

$$||\vec{x}_1 + \vec{x}_2 + \cdots + \vec{x}_k||^2 = ||\vec{x}_1||^2 + ||\vec{x}_2||^2 + \cdots + ||\vec{x}_k||^2.$$

Proof.

Start with

$$\begin{aligned}
||\vec{x}_1 + \vec{x}_2 + \cdots + \vec{x}_k||^2 &= (\vec{x}_1 + \vec{x}_2 + \cdots + \vec{x}_k) \cdot (\vec{x}_1 + \vec{x}_2 + \cdots + \vec{x}_k) \\
&= \quad (\vec{x}_1 \cdot \vec{x}_1 + \vec{x}_1 \cdot \vec{x}_2 + \cdots + \vec{x}_1 \cdot \vec{x}_k) \\
&\quad +(\vec{x}_2 \cdot \vec{x}_1 + \vec{x}_2 \cdot \vec{x}_2 + \cdots + \vec{x}_2 \cdot \vec{x}_k) \\
&\qquad \vdots \qquad\qquad\qquad \vdots \qquad\qquad\qquad \vdots \\
&\quad +(\vec{x}_k \cdot \vec{x}_1 + \vec{x}_k \cdot \vec{x}_2 + \cdots + \vec{x}_k \cdot \vec{x}_k) \\
&= \vec{x}_1 \cdot \vec{x}_1 + \vec{x}_2 \cdot \vec{x}_2 + \cdots + \vec{x}_k \cdot \vec{x}_k \\
&= ||\vec{x}_1||^2 + ||\vec{x}_2||^2 + \cdots + ||\vec{x}_k||^2.
\end{aligned}$$

Theorem (Pythagoras' Theorem)

If $\{\vec{x}_1, \vec{x}_2, \ldots, \vec{x}_k\} \subseteq \mathbb{R}^n$ is orthogonal, then

$$||\vec{x}_1 + \vec{x}_2 + \cdots + \vec{x}_k||^2 = ||\vec{x}_1||^2 + ||\vec{x}_2||^2 + \cdots + ||\vec{x}_k||^2.$$

Proof.

Start with

$$\begin{aligned}
||\vec{x}_1 + \vec{x}_2 + \cdots + \vec{x}_k||^2 &= (\vec{x}_1 + \vec{x}_2 + \cdots + \vec{x}_k) \cdot (\vec{x}_1 + \vec{x}_2 + \cdots + \vec{x}_k) \\
&= \quad (\vec{x}_1 \cdot \vec{x}_1 + \vec{x}_1 \cdot \vec{x}_2 + \cdots + \vec{x}_1 \cdot \vec{x}_k) \\
&\quad +(\vec{x}_2 \cdot \vec{x}_1 + \vec{x}_2 \cdot \vec{x}_2 + \cdots + \vec{x}_2 \cdot \vec{x}_k) \\
&\qquad \vdots \qquad\qquad\qquad \vdots \qquad\qquad\qquad \vdots \\
&\quad +(\vec{x}_k \cdot \vec{x}_1 + \vec{x}_k \cdot \vec{x}_2 + \cdots + \vec{x}_k \cdot \vec{x}_k) \\
&= \vec{x}_1 \cdot \vec{x}_1 + \vec{x}_2 \cdot \vec{x}_2 + \cdots + \vec{x}_k \cdot \vec{x}_k \\
&= ||\vec{x}_1||^2 + ||\vec{x}_2||^2 + \cdots + ||\vec{x}_k||^2.
\end{aligned}$$

The second last equality follows from the fact that the set is orthogonal, so for all i and j, $i \neq j$ and $1 \leq i, j \leq k$, $\vec{x}_i \cdot \vec{x}_j = 0$. Thus, the only nonzero terms are the ones of the form $\vec{x}_i \cdot \vec{x}_i$, $1 \leq i \leq k$. ■

Orthogonality and Independence

Orthogonality and Independence

Theorem

If $S = \{\vec{x}_1, \vec{x}_2, \ldots, \vec{x}_k\} \subseteq \mathbb{R}^n$ is an orthogonal set, then S is independent.

Orthogonality and Independence

Theorem

If $S = \{\vec{x}_1, \vec{x}_2, \ldots, \vec{x}_k\} \subseteq \mathbb{R}^n$ is an orthogonal set, then S is independent.

Proof.

Form the linear equation: $t_1\vec{x}_1 + t_2\vec{x}_2 + \cdots + t_k\vec{x}_k = \vec{0}$. We need to check whether there is only trivial solution.

Orthogonality and Independence

Theorem

If $S = \{\vec{x}_1, \vec{x}_2, \ldots, \vec{x}_k\} \subseteq \mathbb{R}^n$ is an orthogonal set, then S is independent.

Proof.

Form the linear equation: $t_1\vec{x}_1 + t_2\vec{x}_2 + \cdots + t_k\vec{x}_k = \vec{0}$. We need to check whether there is only trivial solution. Notice that for all i, $1 \leq i \leq k$,

$$0 = (t_1\vec{x}_1 + t_2\vec{x}_2 + \cdots + t_k\vec{x}_k) \cdot \vec{x}_i = t_i\vec{x}_i \cdot \vec{x}_i = t_i||\vec{x}_i||^2,$$

since $t_j\vec{x}_j \cdot \vec{x}_i = 0$ for all j, $1 \leq j \leq k$ where $j \neq i$.

Theorem

If $S = \{\vec{x}_1, \vec{x}_2, \ldots, \vec{x}_k\} \subseteq \mathbb{R}^n$ is an orthogonal set, then S is independent.

Proof.

Form the linear equation: $t_1\vec{x}_1 + t_2\vec{x}_2 + \cdots + t_k\vec{x}_k = \vec{0}$. We need to check whether there is only trivial solution. Notice that for all i, $1 \leq i \leq k$,

$$0 = (t_1\vec{x}_1 + t_2\vec{x}_2 + \cdots + t_k\vec{x}_k) \cdot \vec{x}_i = t_i\vec{x}_i \cdot \vec{x}_i = t_i||\vec{x}_i||^2,$$

since $t_j\vec{x}_j \cdot \vec{x}_i = 0$ for all j, $1 \leq j \leq k$ where $j \neq i$. Since $\vec{x}_i \neq \vec{0}_n$ and $t_i||\vec{x}_i||^2 = 0$, it follows that $t_i = 0$ for all i, $1 \leq i \leq k$. Therefore, S is linearly independent. ■

Example

Given an arbitrary vector

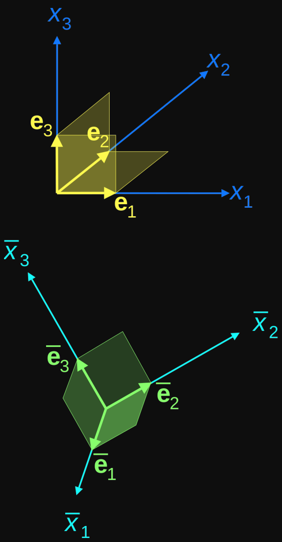
$$\vec{x} = \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix} \in \mathbb{R}^n,$$

it is trivial to express $\vec{x}$ as a linear combination of the standard basis vectors of $\mathbb{R}^n$, $\{\vec{e}_1, \vec{e}_2, \ldots, \vec{e}_n\}$:

$$\vec{x} = a_1\vec{e}_1 + a_2\vec{e}_2 + \cdots + a_n\vec{e}_n.$$

Problem

Given any orthogonal basis B of $\mathbb{R}^n$ (so not necessarily the standard basis), and an arbitrary vector $\vec{x} \in \mathbb{R}^n$, how do we express $\vec{x}$ as a linear combination of the vectors in B?



Fourier Expansion

Fourier Expansion

Theorem (Fourier Expansion)

Let $\{\vec{f}_1, \vec{f}_2, \ldots, \vec{f}_m\}$ be an orthogonal basis of a subspace U of $\mathbb{R}^n$. Then for any $\vec{x} \in U$,

$$\vec{x} = \left(\frac{\vec{x} \cdot \vec{f}_1}{||\vec{f}_1||^2}\right)\vec{f}_1 + \left(\frac{\vec{x} \cdot \vec{f}_2}{||\vec{f}_2||^2}\right)\vec{f}_2 + \cdots + \left(\frac{\vec{x} \cdot \vec{f}_m}{||\vec{f}_m||^2}\right)\vec{f}_m.$$

This expression is called the Fourier expansion of $\vec{x}$, and

$$\frac{\vec{x} \cdot \vec{f}_j}{||\vec{f}_j||^2}, \quad j = 1, 2, \ldots, m$$

are called the Fourier coefficients.

Example

Let $\vec{f}_1 = \begin{bmatrix} 1 \\ -1 \\ 2 \end{bmatrix}$, $\vec{f}_2 = \begin{bmatrix} 0 \\ 2 \\ 1 \end{bmatrix}$, and $\vec{f}_3 = \begin{bmatrix} 5 \\ 1 \\ -2 \end{bmatrix}$, and let $\vec{x} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$.

We have seen that $B = \{\vec{f}_1, \vec{f}_2, \vec{f}_3\}$ is an orthogonal subset of $\mathbb{R}^3$.

Example

Let $\vec{f}_1 = \begin{bmatrix} 1 \\ -1 \\ 2 \end{bmatrix}$, $\vec{f}_2 = \begin{bmatrix} 0 \\ 2 \\ 1 \end{bmatrix}$, and $\vec{f}_3 = \begin{bmatrix} 5 \\ 1 \\ -2 \end{bmatrix}$, and let $\vec{x} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$.

We have seen that $B = \{\vec{f}_1, \vec{f}_2, \vec{f}_3\}$ is an orthogonal subset of $\mathbb{R}^3$.

It follows that B is an orthogonal basis of $\mathbb{R}^3$. (Why?)

Example

Let $\vec{f}_1 = \begin{bmatrix} 1 \\ -1 \\ 2 \end{bmatrix}, \vec{f}_2 = \begin{bmatrix} 0 \\ 2 \\ 1 \end{bmatrix}$, and $\vec{f}_3 = \begin{bmatrix} 5 \\ 1 \\ -2 \end{bmatrix}$, and let $\vec{x} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$.

We have seen that $B = \{\vec{f}_1, \vec{f}_2, \vec{f}_3\}$ is an orthogonal subset of $\mathbb{R}^3$.

It follows that B is an orthogonal basis of $\mathbb{R}^3$. (Why?)

To express $\vec{x}$ as a linear combination of the vectors of B, apply the Fourier Expansion Theorem. Assume $\vec{x} = t_1\vec{f}_1 + t_2\vec{f}_2 + t_3\vec{f}_3$.

Example

Let $\vec{f}_1 = \begin{bmatrix} 1 \\ -1 \\ 2 \end{bmatrix}$, $\vec{f}_2 = \begin{bmatrix} 0 \\ 2 \\ 1 \end{bmatrix}$, and $\vec{f}_3 = \begin{bmatrix} 5 \\ 1 \\ -2 \end{bmatrix}$, and let $\vec{x} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$.

We have seen that $B = \{\vec{f}_1, \vec{f}_2, \vec{f}_3\}$ is an orthogonal subset of $\mathbb{R}^3$.

It follows that B is an orthogonal basis of $\mathbb{R}^3$. (Why?)

To express $\vec{x}$ as a linear combination of the vectors of B, apply the Fourier Expansion Theorem. Assume $\vec{x} = t_1\vec{f}_1 + t_2\vec{f}_2 + t_3\vec{f}_3$. Then

$$t_1 = \frac{\vec{x} \cdot \vec{f}_1}{||\vec{f}_1||^2} = \frac{2}{6}, \quad t_2 = \frac{\vec{x} \cdot \vec{f}_2}{||\vec{f}_2||^2} = \frac{3}{5}, \quad \text{and} \quad t_3 = \frac{\vec{x} \cdot \vec{f}_3}{||\vec{f}_3||^2} = \frac{4}{30}.$$

Example

Let $\vec{f}_1 = \begin{bmatrix} 1 \\ -1 \\ 2 \end{bmatrix}, \vec{f}_2 = \begin{bmatrix} 0 \\ 2 \\ 1 \end{bmatrix}$, and $\vec{f}_3 = \begin{bmatrix} 5 \\ 1 \\ -2 \end{bmatrix}$, and let $\vec{x} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$.

We have seen that $B = \{\vec{f}_1, \vec{f}_2, \vec{f}_3\}$ is an orthogonal subset of $\mathbb{R}^3$.

It follows that B is an orthogonal basis of $\mathbb{R}^3$. (Why?)

To express $\vec{x}$ as a linear combination of the vectors of B, apply the Fourier Expansion Theorem. Assume $\vec{x} = t_1\vec{f}_1 + t_2\vec{f}_2 + t_3\vec{f}_3$. Then

$$t_1 = \frac{\vec{x} \cdot \vec{f}_1}{||\vec{f}_1||^2} = \frac{2}{6}, \quad t_2 = \frac{\vec{x} \cdot \vec{f}_2}{||\vec{f}_2||^2} = \frac{3}{5}, \quad \text{and} \quad t_3 = \frac{\vec{x} \cdot \vec{f}_3}{||\vec{f}_3||^2} = \frac{4}{30}.$$

Therefore,

$$\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} = \frac{1}{3}\begin{bmatrix} 1 \\ -1 \\ 2 \end{bmatrix} + \frac{3}{5}\begin{bmatrix} 0 \\ 2 \\ 1 \end{bmatrix} + \frac{2}{15}\begin{bmatrix} 5 \\ 1 \\ -2 \end{bmatrix}.$$

Proof. (Fourier Expansion)

Let $\vec{x} \in U$. Since $\{\vec{f}_1, \vec{f}_2, \ldots, \vec{f}_m\}$ is a basis of U, $\vec{x} = t_1\vec{f}_1 + t_2\vec{f}_2 + \cdots + t_m\vec{f}_m$ for some $t_1, t_2, \ldots, t_m \in \mathbb{R}$.

Proof. (Fourier Expansion)

Let $\vec{x} \in U$. Since $\{\vec{f}_1, \vec{f}_2, \ldots, \vec{f}_m\}$ is a basis of U, $\vec{x} = t_1\vec{f}_1 + t_2\vec{f}_2 + \cdots + t_m\vec{f}_m$ for some $t_1, t_2, \ldots, t_m \in \mathbb{R}$. Notice that for any i, $1 \leq i \leq m$,

$$\vec{x} \cdot \vec{f}_i = (t_1\vec{f}_1 + t_2\vec{f}_2 + \cdots + t_m\vec{f}_m) \cdot \vec{f}_i$$

Proof. (Fourier Expansion)

Let $\vec{x} \in U$. Since $\{\vec{f}_1, \vec{f}_2, \dots, \vec{f}_m\}$ is a basis of U, $\vec{x} = t_1\vec{f}_1 + t_2\vec{f}_2 + \cdots + t_m\vec{f}_m$ for some $t_1, t_2, \dots, t_m \in \mathbb{R}$. Notice that for any i, $1 \leq i \leq m$,

$$\begin{aligned}
\vec{x} \cdot \vec{f}_i &= (t_1\vec{f}_1 + t_2\vec{f}_2 + \cdots + t_m\vec{f}_m) \cdot \vec{f}_i \\
&= t_i\vec{f}_i \cdot \vec{f}_i \quad \text{since } \{\vec{f}_1, \vec{f}_2, \dots, \vec{f}_m\} \text{ is orthogonal} \\
&= t_i||\vec{f}_i||^2.
\end{aligned}$$

Proof. (Fourier Expansion)

Let $\vec{x} \in U$. Since $\{\vec{f}_1, \vec{f}_2, \ldots, \vec{f}_m\}$ is a basis of U, $\vec{x} = t_1\vec{f}_1 + t_2\vec{f}_2 + \cdots + t_m\vec{f}_m$ for some $t_1, t_2, \ldots, t_m \in \mathbb{R}$. Notice that for any i, $1 \leq i \leq m$,

$$\begin{aligned} \vec{x} \cdot \vec{f}_i &= (t_1\vec{f}_1 + t_2\vec{f}_2 + \cdots + t_m\vec{f}_m) \cdot \vec{f}_i \\ &= t_i\vec{f}_i \cdot \vec{f}_i \quad \text{since } \{\vec{f}_1, \vec{f}_2, \ldots, \vec{f}_m\} \text{ is orthogonal} \\ &= t_i||\vec{f}_i||^2. \end{aligned}$$

Since $\vec{f}_i$ is nonzero, we obtain

$$t_i = \frac{\vec{x} \cdot \vec{f}_i}{||\vec{f}_i||^2}.$$

The result now follows. ■

Proof. (Fourier Expansion)

Let $\vec{x} \in U$. Since $\{\vec{f}_1, \vec{f}_2, \ldots, \vec{f}_m\}$ is a basis of U, $\vec{x} = t_1\vec{f}_1 + t_2\vec{f}_2 + \cdots + t_m\vec{f}_m$ for some $t_1, t_2, \ldots, t_m \in \mathbb{R}$. Notice that for any i, $1 \leq i \leq m$,

$$\begin{aligned} \vec{x} \cdot \vec{f}_i &= (t_1\vec{f}_1 + t_2\vec{f}_2 + \cdots + t_m\vec{f}_m) \cdot \vec{f}_i \\ &= t_i\vec{f}_i \cdot \vec{f}_i \quad \text{since } \{\vec{f}_1, \vec{f}_2, \ldots, \vec{f}_m\} \text{ is orthogonal} \\ &= t_i||\vec{f}_i||^2. \end{aligned}$$

Since $\vec{f}_i$ is nonzero, we obtain

$$t_i = \frac{\vec{x} \cdot \vec{f}_i}{||\vec{f}_i||^2}.$$

The result now follows. ■

Remark

If $\{\vec{f}_1, \vec{f}_2, \ldots, \vec{f}_m\}$ is an orthonormal basis, then the Fourier coefficients are simply $t_j = \vec{x} \cdot \vec{f}_j$, $j = 1, 2, \ldots, m$.

Problem

$$\text{Let}\quad \vec{f}_1 = \begin{bmatrix} 1 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \quad \vec{f}_2 = \begin{bmatrix} 1 \\ -1 \\ 0 \\ 0 \end{bmatrix}, \quad \vec{f}_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \\ 1 \end{bmatrix}, \quad \vec{f}_4 = \begin{bmatrix} 0 \\ 0 \\ 1 \\ -1 \end{bmatrix}.$$

Show that $B = \{\vec{f}_1, \vec{f}_2, \vec{f}_3, \vec{f}_4\}$ is an orthogonal basis of $\mathbb{R}^4$, and express $\vec{x} = \begin{bmatrix} a & b & c & d \end{bmatrix}^T$ as a linear combination of $\vec{f}_1, \vec{f}_2, \vec{f}_3$ and $\vec{f}_4$.

Problem

Let $$\vec{f}_1 = \begin{bmatrix} 1 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \quad \vec{f}_2 = \begin{bmatrix} 1 \\ -1 \\ 0 \\ 0 \end{bmatrix}, \quad \vec{f}_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \\ 1 \end{bmatrix}, \quad \vec{f}_4 = \begin{bmatrix} 0 \\ 0 \\ 1 \\ -1 \end{bmatrix}.$$

Show that $B = \{\vec{f}_1, \vec{f}_2, \vec{f}_3, \vec{f}_4\}$ is an orthogonal basis of $\mathbb{R}^4$, and express $\vec{x} = \begin{bmatrix} a & b & c & d \end{bmatrix}^T$ as a linear combination of $\vec{f}_1, \vec{f}_2, \vec{f}_3$ and $\vec{f}_4$.

Solution

Computing $\vec{f}_i \cdot \vec{f}_j$ for $1 \leq i < j \leq 4$ gives us

$$\begin{aligned} \vec{f}_1 \cdot \vec{f}_2 &= 0, & \vec{f}_1 \cdot \vec{f}_3 &= 0, & \vec{f}_1 \cdot \vec{f}_4 &= 0, \\ \vec{f}_2 \cdot \vec{f}_3 &= 0, & \vec{f}_2 \cdot \vec{f}_4 &= 0, & \vec{f}_3 \cdot \vec{f}_4 &= 0. \end{aligned}$$

Hence, B is an orthogonal set.

Problem

$$\text{Let}\quad \vec{f}_1 = \begin{bmatrix} 1 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \quad \vec{f}_2 = \begin{bmatrix} 1 \\ -1 \\ 0 \\ 0 \end{bmatrix}, \quad \vec{f}_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \\ 1 \end{bmatrix}, \quad \vec{f}_4 = \begin{bmatrix} 0 \\ 0 \\ 1 \\ -1 \end{bmatrix}.$$

Show that $B = \{\vec{f}_1, \vec{f}_2, \vec{f}_3, \vec{f}_4\}$ is an orthogonal basis of $\mathbb{R}^4$, and express $\vec{x} = \begin{bmatrix} a & b & c & d \end{bmatrix}^T$ as a linear combination of $\vec{f}_1, \vec{f}_2, \vec{f}_3$ and $\vec{f}_4$.

Solution

Computing $\vec{f}_i \cdot \vec{f}_j$ for $1 \leq i < j \leq 4$ gives us

$$\begin{array}{lllllllll} \vec{f}_1 \cdot \vec{f}_2 & = & 0, & \vec{f}_1 \cdot \vec{f}_3 & = & 0, & \vec{f}_1 \cdot \vec{f}_4 & = & 0, \\ \vec{f}_2 \cdot \vec{f}_3 & = & 0, & \vec{f}_2 \cdot \vec{f}_4 & = & 0, & \vec{f}_3 \cdot \vec{f}_4 & = & 0. \end{array}$$

Hence, B is an orthogonal set. It follows that B is independent, and since $|B| = 4 = \dim(\mathbb{R}^4)$, B also spans $\mathbb{R}^4$. Therefore, B is an orthogonal basis of $\mathbb{R}^4$.

Problem

$$\text{Let} \quad \vec{f}_1 = \begin{bmatrix} 1 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \quad \vec{f}_2 = \begin{bmatrix} 1 \\ -1 \\ 0 \\ 0 \end{bmatrix}, \quad \vec{f}_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \\ 1 \end{bmatrix}, \quad \vec{f}_4 = \begin{bmatrix} 0 \\ 0 \\ 1 \\ -1 \end{bmatrix}.$$

Show that $B = \{\vec{f}_1, \vec{f}_2, \vec{f}_3, \vec{f}_4\}$ is an orthogonal basis of $\mathbb{R}^4$, and express $\vec{x} = \begin{bmatrix} a & b & c & d \end{bmatrix}^T$ as a linear combination of $\vec{f}_1, \vec{f}_2, \vec{f}_3$ and $\vec{f}_4$.

Solution

Computing $\vec{f}_i \cdot \vec{f}_j$ for $1 \leq i < j \leq 4$ gives us

$$\begin{aligned} \vec{f}_1 \cdot \vec{f}_2 &= 0, & \vec{f}_1 \cdot \vec{f}_3 &= 0, & \vec{f}_1 \cdot \vec{f}_4 &= 0, \\ \vec{f}_2 \cdot \vec{f}_3 &= 0, & \vec{f}_2 \cdot \vec{f}_4 &= 0, & \vec{f}_3 \cdot \vec{f}_4 &= 0. \end{aligned}$$

Hence, B is an orthogonal set. It follows that B is independent, and since $|B| = 4 = \dim(\mathbb{R}^4)$, B also spans $\mathbb{R}^4$. Therefore, B is an orthogonal basis of $\mathbb{R}^4$. By the Fourier Expansion Theorem,

$$\vec{x} = \left(\frac{a+b}{2}\right)\vec{f}_1 + \left(\frac{a-b}{2}\right)\vec{f}_2 + \left(\frac{c+d}{2}\right)\vec{f}_3 + \left(\frac{c-d}{2}\right)\vec{f}_4.$$

■