MA106-Linear Algebra

Spring 2011

Anant R. Shastri

Department of Mathematics
Indian Institute of Technology, Bombay

February 15, 2011

Inner Product

Parallelogram law, the ability to measure angle between two vectors and in particular, the concept of perpendicularity make the euclidean space quite a special type of a vector space.

Inner Product

Parallelogram law, the ability to measure angle between two vectors and in particular, the concept of perpendicularity make the euclidean space quite a special type of a vector space. Essentially all these are consequences of the dot product.

Inner Product

Parallelogram law, the ability to measure angle between two vectors and in particular, the concept of perpendicularity make the euclidean space quite a special type of a vector space. Essentially all these are consequences of the dot product. Thus, it makes sense to look for operations which share the basic properties of the dot product.

Inner Product

Parallelogram law, the ability to measure angle between two vectors and in particular, the concept of perpendicularity make the euclidean space quite a special type of a vector space. Essentially all these are consequences of the dot product. Thus, it makes sense to look for operations which share the basic properties of the dot product.In this section we shall briefly discuss this.

Inner Product

Definition

Let V be a vector space over $\mathbb{K} = \mathbb{R}$ or $\mathbb{C}$.

Inner Product

Definition

Let V be a vector space over $\mathbb{K} = \mathbb{R}$ or $\mathbb{C}$. (Let $z \mapsto \bar{z}$ denote the complex conjugation if $\mathbb{K} = \mathbb{C}$ or the identity map if $\mathbb{K} = \mathbb{R}$.)

Inner Product

Definition

Let V be a vector space over $\mathbb{K} = \mathbb{R}$ or $\mathbb{C}$. (Let $z \mapsto \bar{z}$ denote the complex conjugation if $\mathbb{K} = \mathbb{C}$ or the identity map if $\mathbb{K} = \mathbb{R}$.)
By an **inner product** on V we mean a binary operation, which associates a scalar say, $\langle \mathbf{u}, \mathbf{v} \rangle$ for each pair of vectors $(\mathbf{u}, \mathbf{v}) \in V \times V$, satisfying the following properties for all $\mathbf{u}, \mathbf{v}, \mathbf{w}$ in V and α, β any scalar.

Inner Product

Definition

Let V be a vector space over $\mathbb{K} = \mathbb{R}$ or $\mathbb{C}$. (Let $z \mapsto \bar{z}$ denote the complex conjugation if $\mathbb{K} = \mathbb{C}$ or the identity map if $\mathbb{K} = \mathbb{R}$.)
By an **inner product** on V we mean a binary operation, which associates a scalar say, $\langle \mathbf{u}, \mathbf{v} \rangle$ for each pair of vectors $(\mathbf{u}, \mathbf{v}) \in V \times V$, satisfying the following properties for all $\mathbf{u}, \mathbf{v}, \mathbf{w}$ in V and α, β any scalar.

(1) $\langle \mathbf{u}, \mathbf{v} \rangle = \overline{\langle \mathbf{v}, \mathbf{u} \rangle}$ (Hermitian property or conjugate symmetry);

Inner Product

Definition

Let V be a vector space over $\mathbb{K} = \mathbb{R}$ or $\mathbb{C}$. (Let $z \mapsto \bar{z}$ denote the complex conjugation if $\mathbb{K} = \mathbb{C}$ or the identity map if $\mathbb{K} = \mathbb{R}$.)
By an **inner product** on V we mean a binary operation, which associates a scalar say, $\langle \mathbf{u}, \mathbf{v} \rangle$ for each pair of vectors $(\mathbf{u}, \mathbf{v}) \in V \times V$, satisfying the following properties for all $\mathbf{u}, \mathbf{v}, \mathbf{w}$ in V and α, β any scalar.

(1) $\langle \mathbf{u}, \mathbf{v} \rangle = \overline{\langle \mathbf{v}, \mathbf{u} \rangle}$ (Hermitian property or conjugate symmetry);
(2) $\langle \mathbf{u}, \alpha\mathbf{v} + \beta\mathbf{w} \rangle = \alpha\langle \mathbf{u}, \mathbf{v} \rangle + \beta\langle \mathbf{u}, \mathbf{w} \rangle$ (sesquilinearity);

Inner Product

Definition

Let V be a vector space over $\mathbb{K} = \mathbb{R}$ or $\mathbb{C}$. (Let $z \mapsto \bar{z}$ denote the complex conjugation if $\mathbb{K} = \mathbb{C}$ or the identity map if $\mathbb{K} = \mathbb{R}$.)
By an **inner product** on V we mean a binary operation, which associates a scalar say, $\langle \mathbf{u}, \mathbf{v} \rangle$ for each pair of vectors $(\mathbf{u}, \mathbf{v}) \in V \times V$, satisfying the following properties for all $\mathbf{u}, \mathbf{v}, \mathbf{w}$ in V and α, β any scalar.

(1) $\langle \mathbf{u}, \mathbf{v} \rangle = \overline{\langle \mathbf{v}, \mathbf{u} \rangle}$ (Hermitian property or conjugate symmetry);
(2) $\langle \mathbf{u}, \alpha\mathbf{v} + \beta\mathbf{w} \rangle = \alpha\langle \mathbf{u}, \mathbf{v} \rangle + \beta\langle \mathbf{u}, \mathbf{w} \rangle$ (sesquilinearity);
(3) $\langle \mathbf{v}, \mathbf{v} \rangle > 0$ if $\mathbf{v} \neq 0$ (positivity).

Inner Product

Definition

Let V be a vector space over $\mathbb{K} = \mathbb{R}$ or $\mathbb{C}$. (Let $z \mapsto \bar{z}$ denote the complex conjugation if $\mathbb{K} = \mathbb{C}$ or the identity map if $\mathbb{K} = \mathbb{R}$.)
By an **inner product** on V we mean a binary operation, which associates a scalar say, $\langle \mathbf{u}, \mathbf{v} \rangle$ for each pair of vectors $(\mathbf{u}, \mathbf{v}) \in V \times V$, satisfying the following properties for all $\mathbf{u}, \mathbf{v}, \mathbf{w}$ in V and α, β any scalar.

(1) $\langle \mathbf{u}, \mathbf{v} \rangle = \overline{\langle \mathbf{v}, \mathbf{u} \rangle}$ (Hermitian property or conjugate symmetry);
(2) $\langle \mathbf{u}, \alpha\mathbf{v} + \beta\mathbf{w} \rangle = \alpha\langle \mathbf{u}, \mathbf{v} \rangle + \beta\langle \mathbf{u}, \mathbf{w} \rangle$ (sesquilinearity);
(3) $\langle \mathbf{v}, \mathbf{v} \rangle > 0$ if $\mathbf{v} \neq 0$ (positivity).

A vector space with an inner product is called an **inner product space**.

Inner Product

Remark

(i) Observe that we have not mentioned whether V is a real vector space or a complex vector space.

Inner Product

Remark

(i) Observe that we have not mentioned whether V is a real vector space or a complex vector space. The above definition includes both the cases.

Inner Product

Remark

(i) Observe that we have not mentioned whether V is a real vector space or a complex vector space. The above definition includes both the cases. The only difference is that if $\mathbb{K} = \mathbb{R}$ then the conjugation is just the identity. Thus for real vector spaces, (1) will becomes 'symmetric property' since for a real number c, we have $\bar{c} = c$.

Inner Product

Remark

(i) Observe that we have not mentioned whether V is a real vector space or a complex vector space. The above definition includes both the cases. The only difference is that if $\mathbb{K} = \mathbb{R}$ then the conjugation is just the identity. Thus for real vector spaces, (1) will becomes 'symmetric property' since for a real number c, we have $\bar{c} = c$.
(ii) Combining (1) and (2) we obtain
(2') $\langle \alpha \mathbf{u} + \beta \mathbf{v}, \mathbf{w} \rangle = \bar{\alpha} \langle \mathbf{u}, \mathbf{w} \rangle + \bar{\beta} \langle \mathbf{v}, \mathbf{w} \rangle$.

Inner Product

Remark

(i) Observe that we have not mentioned whether V is a real vector space or a complex vector space. The above definition includes both the cases. The only difference is that if $\mathbb{K} = \mathbb{R}$ then the conjugation is just the identity. Thus for real vector spaces, (1) will becomes 'symmetric property' since for a real number c, we have $\bar{c} = c$.
(ii) Combining (1) and (2) we obtain
(2') $\langle \alpha\mathbf{u} + \beta\mathbf{v}, \mathbf{w}\rangle = \bar{\alpha}\langle \mathbf{u}, \mathbf{w}\rangle + \bar{\beta}\langle \mathbf{v}, \mathbf{w}\rangle$.
In the special case when $\mathbb{K} = \mathbb{R}$, properties (2) and (2') together are called *bilinearity*

frametitleInner Product: Examples

Example

(i) The dot product in $\mathbb{R}^n$ *is an inner product. This is often called the* standard inner product.

frametitleInner Product: Examples

Example

(i) The dot product in $\mathbb{R}^n$ *is an inner product. This is often called the* standard inner product.
(ii) Let $\mathbf{x} = (x_1, x_2, \ldots, x_n)^t$ *and* $\mathbf{y} = (y_1, y_2, \ldots, y_n)^t \in \mathbb{C}^n$. *Define* $\langle \mathbf{x}, \mathbf{y} \rangle = \sum_{i=1}^{n} \overline{x_i} y_i = \mathbf{x}^* \mathbf{y}$.

frametitleInner Product: Examples

Example

(i) The dot product in $\mathbb{R}^n$ *is an inner product. This is often called the* standard inner product.
(ii) Let $\mathbf{x} = (x_1, x_2, \ldots, x_n)^t$ *and* $\mathbf{y} = (y_1, y_2, \ldots, y_n)^t \in \mathbb{C}^n$. *Define* $\langle \mathbf{x}, \mathbf{y} \rangle = \sum_{i=1}^{n} \overline{x_i} y_i = \mathbf{x}^* \mathbf{y}$. *This is an inner product on the complex vector space* $\mathbb{C}^n$.

frametitleInner Product: Examples

Example

(i) The dot product in $\mathbb{R}^n$ *is an inner product. This is often called the* standard inner product.
(ii) Let $\mathbf{x} = (x_1, x_2, \ldots, x_n)^t$ *and* $\mathbf{y} = (y_1, y_2, \ldots, y_n)^t \in \mathbb{C}^n$. *Define* $\langle \mathbf{x}, \mathbf{y} \rangle = \sum_{i=1}^{n} \overline{x_i} y_i = \mathbf{x}^* \mathbf{y}$. *This is an inner product on the complex vector space* $\mathbb{C}^n$.
(iii) Let $V = C[0, 1]$. *For* $f, g \in V$, *put*

$$\langle f, g \rangle = \int_0^1 \overline{f(t)} g(t)\, dt.$$

It is easy to see that $\langle f, g \rangle$ *is an inner product.*

Norm Associated to an Inner Product

Definition

Let V be an inner product space.

Norm Associated to an Inner Product

Definition

Let V be an inner product space. For any $\mathbf{v} \in V$, the norm of $\mathbf{v}$, denoted by $\|\mathbf{v}\|$, is the positive square root of $\langle \mathbf{v}, \mathbf{v} \rangle$: $\|\mathbf{v}\| = \sqrt{\langle \mathbf{v}, \mathbf{v} \rangle}$.

Norm Associated to an Inner Product

Definition

Let V be an inner product space. For any $\mathbf{v} \in V$, the norm of $\mathbf{v}$, denoted by $\|\mathbf{v}\|$, is the positive square root of $\langle \mathbf{v}, \mathbf{v} \rangle$: $\|\mathbf{v}\| = \sqrt{\langle \mathbf{v}, \mathbf{v} \rangle}$.

For standard inner product in $\mathbb{R}^n$, $\|\mathbf{v}\|$ is the usual length of the vector $\mathbf{v}$.

Norm Associated to an Inner Product

Proposition

Let V be an inner product space. Let $\mathbf{u}, \mathbf{v} \in V$ and c be a scalar.

Norm Associated to an Inner Product

Proposition

Let V be an inner product space. Let $\mathbf{u}, \mathbf{v} \in V$ and c be a scalar. Then
(i) $\|c\mathbf{u}\| = |c|\|\mathbf{u}\|$;

Norm Associated to an Inner Product

Proposition

Let V be an inner product space. Let $\mathbf{u}, \mathbf{v} \in V$ and c be a scalar. Then
(i) $\|c\mathbf{u}\| = |c|\|\mathbf{u}\|$;
(ii) $\|\mathbf{u}\| > 0$ for $\mathbf{u} \neq 0$;

Norm Associated to an Inner Product

Proposition

Let V be an inner product space. Let $\mathbf{u}, \mathbf{v} \in V$ and c be a scalar. Then
(i) $\|c\mathbf{u}\| = |c|\|\mathbf{u}\|$;
(ii) $\|\mathbf{u}\| > 0$ for $\mathbf{u} \neq 0$;
(iii) $|\langle \mathbf{u}, \mathbf{v} \rangle| \leq \|\mathbf{u}\|\|\mathbf{v}\|$;

Norm Associated to an Inner Product

Proposition

Let V be an inner product space. Let $\mathbf{u}, \mathbf{v} \in V$ and c be a scalar. Then
(i) $\|c\mathbf{u}\| = |c|\|\mathbf{u}\|$;
(ii) $\|\mathbf{u}\| > 0$ for $\mathbf{u} \neq 0$;
(iii) $|\langle \mathbf{u}, \mathbf{v} \rangle| \leq \|\mathbf{u}\|\|\mathbf{v}\|$;
(iv) $\|\mathbf{u} + \mathbf{v}\| \leq \|\mathbf{u}\| + \|\mathbf{v}\|$, equality holds in this triangle inequality iff one of $\mathbf{u}, \mathbf{v}$ is a nonnegative real multiple of the other.

Norm Associated to an Inner Product

Proposition

Let V be an inner product space. Let $\mathbf{u}, \mathbf{v} \in V$ and c be a scalar. Then
(i) $\|c\mathbf{u}\| = |c|\|\mathbf{u}\|$;
(ii) $\|\mathbf{u}\| > 0$ for $\mathbf{u} \neq 0$;
(iii) $|\langle \mathbf{u}, \mathbf{v} \rangle| \leq \|\mathbf{u}\|\|\mathbf{v}\|$;
(iv) $\|\mathbf{u} + \mathbf{v}\| \leq \|\mathbf{u}\| + \|\mathbf{v}\|$, equality holds in this triangle inequality iff one of $\mathbf{u}, \mathbf{v}$ is a nonnegative real multiple of the other.

Norm Associated to an Inner Product

Proof: (i) and (ii) follow from definitions 1,2.

Norm Associated to an Inner Product

Proof: (i) and (ii) follow from definitions 1,2.
(iii) This is called the **Cauchy-Schwarz inequality**. If either **u** or **v** is zero, it is clear.

Norm Associated to an Inner Product

Proof: (i) and (ii) follow from definitions 1,2.
(iii) This is called the **Cauchy-Schwarz inequality**. If either **u** or **v** is zero, it is clear. So let **u** be nonzero. Consider the vector

$$\mathbf{w} = \mathbf{v} - a\mathbf{u}.$$

where a is a scalar to be specified later.

Norm Associated to an Inner Product

Proof: (i) and (ii) follow from definitions 1,2.
(iii) This is called the **Cauchy-Schwarz inequality**. If either $\mathbf{u}$ or $\mathbf{v}$ is zero, it is clear. So let $\mathbf{u}$ be nonzero. Consider the vector

$$\mathbf{w} = \mathbf{v} - a\mathbf{u}.$$

where a is a scalar to be specified later. Then

$$\begin{aligned} 0 \leq \langle \mathbf{w}, \mathbf{w} \rangle &= \langle \mathbf{v} - a\mathbf{u}, \mathbf{v} - a\mathbf{u} \rangle \\ &= \langle \mathbf{v}, \mathbf{v} \rangle - \langle \mathbf{v}, a\mathbf{u} \rangle - \langle a\mathbf{u}, \mathbf{v} \rangle + \langle a\mathbf{u}, a\mathbf{u} \rangle \\ &= \langle \mathbf{v}, \mathbf{v} \rangle - a\langle \mathbf{v}, \mathbf{u} \rangle - \overline{a}\langle \mathbf{u}, \mathbf{v} \rangle + a\overline{a}\langle \mathbf{u}, \mathbf{u} \rangle \end{aligned}$$

Norm Associated to an Inner Product

Proof: (i) and (ii) follow from definitions 1,2.
(iii) This is called the **Cauchy-Schwarz inequality**. If either $\mathbf{u}$ or $\mathbf{v}$ is zero, it is clear. So let $\mathbf{u}$ be nonzero. Consider the vector

$$\mathbf{w} = \mathbf{v} - a\mathbf{u}.$$

where a is a scalar to be specified later. Then

$$\begin{aligned} 0 \leq \langle \mathbf{w}, \mathbf{w} \rangle &= \langle \mathbf{v} - a\mathbf{u}, \mathbf{v} - a\mathbf{u} \rangle \\ &= \langle \mathbf{v}, \mathbf{v} \rangle - \langle \mathbf{v}, a\mathbf{u} \rangle - \langle a\mathbf{u}, \mathbf{v} \rangle + \langle a\mathbf{u}, a\mathbf{u} \rangle \\ &= \langle \mathbf{v}, \mathbf{v} \rangle - a\langle \mathbf{v}, \mathbf{u} \rangle - \overline{a}\langle \mathbf{u}, \mathbf{v} \rangle + a\overline{a}\langle \mathbf{u}, \mathbf{u} \rangle \end{aligned}$$

Substitute $a = \langle \mathbf{u}, \mathbf{v} \rangle / \langle \mathbf{u}, \mathbf{u} \rangle$ and multiply by $\langle \mathbf{u}, \mathbf{u} \rangle$, we get $|\langle \mathbf{u}, \mathbf{v} \rangle|^2 \leq \langle \mathbf{u}, \mathbf{u} \rangle \langle \mathbf{v}, \mathbf{v} \rangle$. This proves (iii).

Norm Associated to an Inner Product

(iv) Using (iii), we get $\Re(\langle \mathbf{u}, \mathbf{v} \rangle) \leq \|\mathbf{u}\| \, \|\mathbf{v}\|$.

Norm Associated to an Inner Product

(iv) Using (iii), we get $\Re(\langle \mathbf{u}, \mathbf{v} \rangle) \leq \|\mathbf{u}\| \, \|\mathbf{v}\|$.Therefore,

$$\begin{aligned} \|\mathbf{u}+\mathbf{v}\|^2 &= \|\mathbf{u}\|^2 + \langle \mathbf{u}, \mathbf{v} \rangle + \langle \mathbf{v}, \mathbf{u} \rangle + \|\mathbf{v}\|^2 \\ &= \|\mathbf{u}\|^2 + 2\Re\langle \mathbf{u}, \mathbf{v} \rangle + \|\mathbf{v}\|^2 \\ &\leq \|\mathbf{u}\|^2 + 2\|\mathbf{u}\|\|\mathbf{v}\| + \|\mathbf{v}\|^2 \\ &= (\|\mathbf{u}\| + \|\mathbf{v}\|)^2. \end{aligned}$$

Norm Associated to an Inner Product

(iv) Using (iii), we get $\Re(\langle \mathbf{u}, \mathbf{v}\rangle) \leq \|\mathbf{u}\|\,\|\mathbf{v}\|$.Therefore,

$$\begin{aligned}
\|\mathbf{u}+\mathbf{v}\|^2 &= \|\mathbf{u}\|^2 + \langle \mathbf{u}, \mathbf{v}\rangle + \langle \mathbf{v}, \mathbf{u}\rangle + \|\mathbf{v}\|^2 \\
&= \|\mathbf{u}\|^2 + 2\Re\langle \mathbf{u}, \mathbf{v}\rangle + \|\mathbf{v}\|^2 \\
&\leq \|\mathbf{u}\|^2 + 2\|\mathbf{u}\|\|\mathbf{v}\| + \|\mathbf{v}\|^2 \\
&= (\|\mathbf{u}\| + \|\mathbf{v}\|)^2.
\end{aligned}$$

Thus $\|\mathbf{u}+\mathbf{v}\| \leq \|\mathbf{u}\| + \|\mathbf{v}\|$.

Norm Associated to an Inner Product

(iv) Using (iii), we get $\Re(\langle \mathbf{u}, \mathbf{v}\rangle) \leq \|\mathbf{u}\|\,\|\mathbf{v}\|$.Therefore,

$$\begin{aligned}
\|\mathbf{u}+\mathbf{v}\|^2 &= \|\mathbf{u}\|^2 + \langle \mathbf{u}, \mathbf{v}\rangle + \langle \mathbf{v}, \mathbf{u}\rangle + \|\mathbf{v}\|^2 \\
&= \|\mathbf{u}\|^2 + 2\Re\langle \mathbf{u}, \mathbf{v}\rangle + \|\mathbf{v}\|^2 \\
&\leq \|\mathbf{u}\|^2 + 2\|\mathbf{u}\|\|\mathbf{v}\| + \|\mathbf{v}\|^2 \\
&= (\|\mathbf{u}\| + \|\mathbf{v}\|)^2.
\end{aligned}$$

Thus $\|\mathbf{u}+\mathbf{v}\| \leq \|\mathbf{u}\| + \|\mathbf{v}\|$.
Now equality holds in (iv) iff $\Re\langle \mathbf{u}, \mathbf{v}\rangle = \langle \mathbf{u}, \mathbf{v}\rangle$ and equality holds in (iii).

Norm Associated to an Inner Product

(iv) Using (iii), we get $\Re(\langle \mathbf{u}, \mathbf{v}\rangle) \leq \|\mathbf{u}\|\,\|\mathbf{v}\|$.Therefore,

$$\begin{aligned}
\|\mathbf{u}+\mathbf{v}\|^2 &= \|\mathbf{u}\|^2 + \langle \mathbf{u}, \mathbf{v}\rangle + \langle \mathbf{v}, \mathbf{u}\rangle + \|\mathbf{v}\|^2 \\
&= \|\mathbf{u}\|^2 + 2\Re\langle \mathbf{u}, \mathbf{v}\rangle + \|\mathbf{v}\|^2 \\
&\leq \|\mathbf{u}\|^2 + 2\|\mathbf{u}\|\|\mathbf{v}\| + \|\mathbf{v}\|^2 \\
&= (\|\mathbf{u}\| + \|\mathbf{v}\|)^2.
\end{aligned}$$

Thus $\|\mathbf{u}+\mathbf{v}\| \leq \|\mathbf{u}\| + \|\mathbf{v}\|$.
Now equality holds in (iv) iff $\Re\langle \mathbf{u}, \mathbf{v}\rangle = \langle \mathbf{u}, \mathbf{v}\rangle$ and equality holds in (iii). This is the same as saying that $\langle \mathbf{u}, \mathbf{v}\rangle$ is nonnegative real and $\mathbf{w} = 0$. This is the same as saying that $\mathbf{v} = a\mathbf{u}$ and $\langle \mathbf{v}, \mathbf{u}\rangle = a\langle \mathbf{u}, \mathbf{u}\rangle$ is nonnegative real, which, in turn, is the same as saying that $\mathbf{v}$ is a nonnegative real multiple of $\mathbf{u}$. ♠(You may get the other way round if you start with the assumption $\mathbf{v} \neq 0$.

Distance and Angle

Definition

In any inner product space V, the distance between $\mathbf{u}$ and $\mathbf{v}$ is defined as

$$d(\mathbf{u}, \mathbf{v}) = \|\mathbf{u} - \mathbf{v}\|.$$

Distance and Angle

Definition

In any inner product space V, the distance between $\mathbf{u}$ and $\mathbf{v}$ is defined as

$$d(\mathbf{u}, \mathbf{v}) = \|\mathbf{u} - \mathbf{v}\|.$$

Definition

Let V be a real inner product space. The angle between vectors $\mathbf{u}, \mathbf{v}$ is defined to be the angle $\theta, 0 \leq \theta \leq \pi$ so that

$$\cos\theta = \frac{\langle \mathbf{u}, \mathbf{v} \rangle}{\|\mathbf{u}\|\|\mathbf{v}\|}.$$

Distance and Angle

Definition

In any inner product space V, the distance between $\mathbf{u}$ and $\mathbf{v}$ is defined as

$$d(\mathbf{u}, \mathbf{v}) = \|\mathbf{u} - \mathbf{v}\|.$$

Definition

Let V be a real inner product space. The angle between vectors $\mathbf{u}, \mathbf{v}$ is defined to be the angle $\theta, 0 \leq \theta \leq \pi$ so that

$$\cos\theta = \frac{\langle \mathbf{u}, \mathbf{v} \rangle}{\|\mathbf{u}\|\|\mathbf{v}\|}.$$

Proposition

(i) $d(\mathbf{u}, \mathbf{v}) \geq 0$ and $d(\mathbf{u}, \mathbf{v}) = 0$ iff $\mathbf{u} = \mathbf{v}$;
(ii) $d(\mathbf{u}, \mathbf{v}) = d(\mathbf{v}, \mathbf{u})$;
(iii) $d(\mathbf{u}, \mathbf{w}) \leq d(\mathbf{u}, \mathbf{v}) + d(\mathbf{v}, \mathbf{w})$.

Distance and Angle

Definition

In any inner product space V, the distance between $\mathbf{u}$ and $\mathbf{v}$ is defined as

$$d(\mathbf{u}, \mathbf{v}) = \|\mathbf{u} - \mathbf{v}\|.$$

Definition

Let V be a real inner product space. The angle between vectors $\mathbf{u}, \mathbf{v}$ is defined to be the angle $\theta, 0 \leq \theta \leq \pi$ so that

$$\cos\theta = \frac{\langle \mathbf{u}, \mathbf{v} \rangle}{\|\mathbf{u}\|\|\mathbf{v}\|}.$$

Proposition

(i) $d(\mathbf{u}, \mathbf{v}) \geq 0$ and $d(\mathbf{u}, \mathbf{v}) = 0$ iff $\mathbf{u} = \mathbf{v}$;
(ii) $d(\mathbf{u}, \mathbf{v}) = d(\mathbf{v}, \mathbf{u})$;
(iii) $d(\mathbf{u}, \mathbf{w}) \leq d(\mathbf{u}, \mathbf{v}) + d(\mathbf{v}, \mathbf{w})$.
(iv) $d(\alpha\mathbf{u}, \alpha\mathbf{v}) = |\alpha| d(\mathbf{u}, \mathbf{v})$.

Gram-Schmidt

The standard basis $E = \{e_1, e_2, \ldots, e_n\}$ of $\mathbb{R}^n$ has two properties :

Gram-Schmidt

The standard basis $E = \{e_1, e_2, \ldots, e_n\}$ of $\mathbb{R}^n$ has two properties :
(i) each e_i has length one;

Gram-Schmidt

The standard basis $E = \{e_1, e_2, \ldots, e_n\}$ of $\mathbb{R}^n$ has two properties :
(i) each e_i has length one;
(ii) any two elements of E are mutually perpendicular.

Gram-Schmidt

The standard basis $E = \{e_1, e_2, \ldots, e_n\}$ of $\mathbb{R}^n$ has two properties :
(i) each e_i has length one;
(ii) any two elements of E are mutually perpendicular.
When vectors are expressed in terms of the standard basis, calculations become easy due to the above properties.

Gram-Schmidt

The standard basis $E = \{e_1, e_2, \ldots, e_n\}$ of $\mathbb{R}^n$ has two properties :
(i) each e_i has length one;
(ii) any two elements of E are mutually perpendicular.
When vectors are expressed in terms of the standard basis, calculations become easy due to the above properties. Now, let V be any inner product space of dimension d.

Gram-Schmidt

The standard basis $E = \{e_1, e_2, \ldots, e_n\}$ of $\mathbb{R}^n$ has two properties :
(i) each e_i has length one;
(ii) any two elements of E are mutually perpendicular.
When vectors are expressed in terms of the standard basis, calculations become easy due to the above properties. Now, let V be any inner product space of dimension d. We wish to construct an analogue of E for V.

Gram-Schmidt

Definition

Two vectors $\mathbf{u}, \mathbf{v}$ of an inner product space V are called *orthogonal to each other* or *perpendicular to each other* if $\langle \mathbf{u}, \mathbf{v} \rangle = 0$.

Gram-Schmidt

Definition

Two vectors $\mathbf{u}, \mathbf{v}$ of an inner product space V are called *orthogonal to each other* or *perpendicular to each other* if $\langle \mathbf{u}, \mathbf{v} \rangle = 0$. We express this symbolically by $\mathbf{u} \perp \mathbf{v}$.

Gram-Schmidt

Definition

Two vectors $\mathbf{u}, \mathbf{v}$ of an inner product space V are called *orthogonal to each other* or *perpendicular to each other* if $\langle \mathbf{u}, \mathbf{v} \rangle = 0$. We express this symbolically by $\mathbf{u} \perp \mathbf{v}$. A set S consisting of nonzero elements of V is called *orthogonal* if $\langle \mathbf{u}, \mathbf{v} \rangle = 0$ for every pair of elements of S.

Gram-Schmidt

Definition

Two vectors $\mathbf{u}, \mathbf{v}$ of an inner product space V are called *orthogonal to each other* or *perpendicular to each other* if $\langle \mathbf{u}, \mathbf{v} \rangle = 0$. We express this symbolically by $\mathbf{u} \perp \mathbf{v}$. A set S consisting of nonzero elements of V is called *orthogonal* if $\langle \mathbf{u}, \mathbf{v} \rangle = 0$ for every pair of elements of S. Further if every element of V is of unit norm then S is called an *orthonormal* set.

Gram-Schmidt

Definition

Two vectors $\mathbf{u}, \mathbf{v}$ of an inner product space V are called *orthogonal to each other* or *perpendicular to each other* if $\langle \mathbf{u}, \mathbf{v} \rangle = 0$. We express this symbolically by $\mathbf{u} \perp \mathbf{v}$. A set S consisting of nonzero elements of V is called *orthogonal* if $\langle \mathbf{u}, \mathbf{v} \rangle = 0$ for every pair of elements of S. Further if every element of V is of unit norm then S is called an *orthonormal* set. Further, if S is a basis for V then it is called an *orthonormal basis*.

Gram-Schmidt

Definition

Two vectors $\mathbf{u}, \mathbf{v}$ of an inner product space V are called *orthogonal to each other* or *perpendicular to each other* if $\langle \mathbf{u}, \mathbf{v} \rangle = 0$. We express this symbolically by $\mathbf{u} \perp \mathbf{v}$. A set S consisting of nonzero elements of V is called *orthogonal* if $\langle \mathbf{u}, \mathbf{v} \rangle = 0$ for every pair of elements of S. Further if every element of V is of unit norm then S is called an *orthonormal* set. Further, if S is a basis for V then it is called an *orthonormal basis*.

One of the biggest advantage of an orthogonal set is that Pythagorus theorem and its general form are valid, which makes many computations easy.

Gram-Schmidt

Theorem

Pythagorus Theorem: *Given* $\mathbf{w}, \mathbf{u} \in V$ *such that* $\mathbf{w} \perp \mathbf{u}$ *we have*

$$\|\mathbf{w} + \mathbf{u}\|^2 = \|\mathbf{w}\|^2 + \|\mathbf{u}\|^2. \tag{1}$$

Gram-Schmidt

Theorem

Pythagorus Theorem: *Given* $\mathbf{w}, \mathbf{u} \in V$ *such that* $\mathbf{w} \perp \mathbf{u}$ *we have*

$$\|\mathbf{w}+\mathbf{u}\|^2 = \|\mathbf{w}\|^2 + \|\mathbf{u}\|^2. \tag{1}$$

Proof: We have,

$$\|\mathbf{w}+\mathbf{u}\|^2 = \langle \mathbf{w}+\mathbf{u}, \mathbf{w}+\mathbf{u}\rangle = \langle \mathbf{w}, \mathbf{w}\rangle + \langle \mathbf{u}, \mathbf{w}\rangle + \langle \mathbf{w}, \mathbf{u}\rangle + \langle \mathbf{u}, \mathbf{u}\rangle = \|\mathbf{w}\|^2 + \|\mathbf{u}\|^2$$

since $\langle \mathbf{u}, \mathbf{w}\rangle = \langle \mathbf{w}, \mathbf{u}\rangle = 0$. ♠

Gram-Schmidt

Theorem

Pythagorus Theorem: *Given* $\mathbf{w}, \mathbf{u} \in V$ *such that* $\mathbf{w} \perp \mathbf{u}$ *we have*

$$\|\mathbf{w} + \mathbf{u}\|^2 = \|\mathbf{w}\|^2 + \|\mathbf{u}\|^2. \tag{1}$$

Proof: We have,

$$\|\mathbf{w} + \mathbf{u}\|^2 = \langle \mathbf{w} + \mathbf{u}, \mathbf{w} + \mathbf{u} \rangle = \langle \mathbf{w}, \mathbf{w} \rangle + \langle \mathbf{u}, \mathbf{w} \rangle + \langle \mathbf{w}, \mathbf{u} \rangle + \langle \mathbf{u}, \mathbf{u} \rangle = \|\mathbf{w}\|^2 + \|\mathbf{u}\|^2$$

since $\langle \mathbf{u}, \mathbf{w} \rangle = \langle \mathbf{w}, \mathbf{u} \rangle = 0$. ♠

Proposition

Let $S = \{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_n\}$ be an orthogonal set. Then S is linearly independent.

Gram-Schmidt

Theorem

Pythagorus Theorem: *Given* $\mathbf{w}, \mathbf{u} \in V$ *such that* $\mathbf{w} \perp \mathbf{u}$ *we have*

$$\|\mathbf{w}+\mathbf{u}\|^2 = \|\mathbf{w}\|^2 + \|\mathbf{u}\|^2. \tag{1}$$

Proof: We have,

$$\|\mathbf{w}+\mathbf{u}\|^2 = \langle \mathbf{w}+\mathbf{u}, \mathbf{w}+\mathbf{u}\rangle = \langle \mathbf{w}, \mathbf{w}\rangle + \langle \mathbf{u}, \mathbf{w}\rangle + \langle \mathbf{w}, \mathbf{u}\rangle + \langle \mathbf{u}, \mathbf{u}\rangle = \|\mathbf{w}\|^2 + \|\mathbf{u}\|^2$$

since $\langle \mathbf{u}, \mathbf{w}\rangle = \langle \mathbf{w}, \mathbf{u}\rangle = 0$. ♠

Proposition

Let $S = \{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_n\}$ be an orthogonal set. Then S is linearly independent.

Proof: Suppose $c_1, c_2, \ldots, c_n$ are scalars with

$$c_1\mathbf{u}_1 + c_2\mathbf{u}_2 + \ldots + c_n\mathbf{u}_n = 0.$$

Gram-Schmidt

Theorem

Pythagorus Theorem: *Given* $\mathbf{w}, \mathbf{u} \in V$ *such that* $\mathbf{w} \perp \mathbf{u}$ *we have*

$$\|\mathbf{w}+\mathbf{u}\|^2 = \|\mathbf{w}\|^2 + \|\mathbf{u}\|^2. \tag{1}$$

Proof: We have,

$$\|\mathbf{w}+\mathbf{u}\|^2 = \langle \mathbf{w}+\mathbf{u}, \mathbf{w}+\mathbf{u}\rangle = \langle \mathbf{w}, \mathbf{w}\rangle + \langle \mathbf{u}, \mathbf{w}\rangle + \langle \mathbf{w}, \mathbf{u}\rangle + \langle \mathbf{u}, \mathbf{u}\rangle = \|\mathbf{w}\|^2 + \|\mathbf{u}\|^2$$

since $\langle \mathbf{u}, \mathbf{w}\rangle = \langle \mathbf{w}, \mathbf{u}\rangle = 0$. ♠

Proposition

Let $S = \{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_n\}$ be an orthogonal set. Then S is linearly independent.

Proof: Suppose $c_1, c_2, \ldots, c_n$ are scalars with

$$c_1\mathbf{u}_1 + c_2\mathbf{u}_2 + \ldots + c_n\mathbf{u}_n = 0.$$

Take inner product with $\mathbf{u}_i$ on both sides to get

$$c_i\langle \mathbf{u}_i, \mathbf{u}_i\rangle = 0.$$

Gram-Schmidt

Theorem

Pythagorus Theorem: *Given* $\mathbf{w}, \mathbf{u} \in V$ *such that* $\mathbf{w} \perp \mathbf{u}$ *we have*

$$\|\mathbf{w}+\mathbf{u}\|^2 = \|\mathbf{w}\|^2 + \|\mathbf{u}\|^2. \tag{1}$$

Proof: We have,

$$\|\mathbf{w}+\mathbf{u}\|^2 = \langle \mathbf{w}+\mathbf{u}, \mathbf{w}+\mathbf{u}\rangle = \langle \mathbf{w}, \mathbf{w}\rangle + \langle \mathbf{u}, \mathbf{w}\rangle + \langle \mathbf{w}, \mathbf{u}\rangle + \langle \mathbf{u}, \mathbf{u}\rangle = \|\mathbf{w}\|^2 + \|\mathbf{u}\|^2$$

since $\langle \mathbf{u}, \mathbf{w}\rangle = \langle \mathbf{w}, \mathbf{u}\rangle = 0$. ♠

Proposition

Let $S = \{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_n\}$ be an orthogonal set. Then S is linearly independent.

Proof: Suppose $c_1, c_2, \ldots, c_n$ are scalars with

$$c_1\mathbf{u}_1 + c_2\mathbf{u}_2 + \ldots + c_n\mathbf{u}_n = 0.$$

Take inner product with $\mathbf{u}_i$ on both sides to get

$$c_i\langle \mathbf{u}_i, \mathbf{u}_i\rangle = 0.$$

Since $\mathbf{u}_i \neq 0$, we get $c_i = 0$. Thus S is linearly independent. ♠

Gram-Schmidt

Definition

If $\mathbf{u}$ and $\mathbf{v}$ are vectors in an inner product space V and $\mathbf{v} \neq 0$ then the vector

$$\langle \mathbf{v}, \mathbf{u} \rangle \frac{\mathbf{v}}{\|\mathbf{v}\|^2}$$

is called the **orthogonal projection** of $\mathbf{u}$ along $\mathbf{v}$.

Gram-Schmidt

Definition

If $\mathbf{u}$ and $\mathbf{v}$ are vectors in an inner product space V and $\mathbf{v} \neq 0$ then the vector

$$\langle \mathbf{v}, \mathbf{u} \rangle \frac{\mathbf{v}}{\|\mathbf{v}\|^2}$$

is called the **orthogonal projection** of $\mathbf{u}$ along $\mathbf{v}$.

Theorem

(Gram-Schmidt Orthonormalization Process): *Let V be any inner product space.*

Gram-Schmidt

Definition

If $\mathbf{u}$ and $\mathbf{v}$ are vectors in an inner product space V and $\mathbf{v} \neq 0$ then the vector

$$\langle \mathbf{v}, \mathbf{u} \rangle \frac{\mathbf{v}}{\|\mathbf{v}\|^2}$$

is called the **orthogonal projection** of $\mathbf{u}$ along $\mathbf{v}$.

Theorem

(Gram-Schmidt Orthonormalization Process): *Let V be any inner product space. Given a linearly independent subset $\{\mathbf{u}_1, \ldots, \mathbf{u}_n\}$, there exists an orthonormal set of vectors $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ such that $L(\{\mathbf{u}_1, \ldots, \mathbf{u}_k\}) = L(\{\mathbf{v}_1, \ldots, \mathbf{v}_k\})$ for all $1 \leq k \leq n$.*

Gram-Schmidt

Definition

If $\mathbf{u}$ and $\mathbf{v}$ are vectors in an inner product space V and $\mathbf{v} \neq 0$ then the vector

$$\langle \mathbf{v}, \mathbf{u} \rangle \frac{\mathbf{v}}{\|\mathbf{v}\|^2}$$

is called the **orthogonal projection** of $\mathbf{u}$ along $\mathbf{v}$.

Theorem

(Gram-Schmidt Orthonormalization Process): *Let V be any inner product space. Given a linearly independent subset $\{\mathbf{u}_1, \ldots, \mathbf{u}_n\}$, there exists an orthonormal set of vectors $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ such that $L(\{\mathbf{u}_1, \ldots, \mathbf{u}_k\}) = L(\{\mathbf{v}_1, \ldots, \mathbf{v}_k\})$ for all $1 \leq k \leq n$. In particular, if V is finite dimensional, then it has an orthonormal basis.*

Gram-Schmidt

Proof: The construction of the orthonormal set is algorithmic and of course, inductive.

Gram-Schmidt

Proof: The construction of the orthonormal set is algorithmic and of course, inductive. First we construct an intermediate orthogonal set $\{\mathbf{w}_1, \ldots, \mathbf{w}_n\}$ with the same property for each k. Then we simply 'normalize' each element, viz, take

$$\mathbf{v}_i = \frac{\mathbf{w}_i}{\|\mathbf{w}_i\|}.$$

Gram-Schmidt

Proof: The construction of the orthonormal set is algorithmic and of course, inductive. First we construct an intermediate orthogonal set $\{\mathbf{w}_1, \ldots, \mathbf{w}_n\}$ with the same property for each k. Then we simply 'normalize' each element, viz, take

$$\mathbf{v}_i = \frac{\mathbf{w}_i}{\|\mathbf{w}_i\|}.$$

That will give us an orthonormal set as required.

Gram-Schmidt

Proof: The construction of the orthonormal set is algorithmic and of course, inductive. First we construct an intermediate orthogonal set $\{\mathbf{w}_1, \ldots, \mathbf{w}_n\}$ with the same property for each k. Then we simply 'normalize' each element, viz, take

$$\mathbf{v}_i = \frac{\mathbf{w}_i}{\|\mathbf{w}_i\|}.$$

That will give us an orthonormal set as required. The last part of the theorem follows if we begin with $\{\mathbf{u}_1, \ldots, \mathbf{u}_m\}$ as any basis for V.

Gram-Schmidt

Proof: The construction of the orthonormal set is algorithmic and of course, inductive. First we construct an intermediate orthogonal set $\{\mathbf{w}_1, \ldots, \mathbf{w}_n\}$ with the same property for each k. Then we simply 'normalize' each element, viz, take

$$\mathbf{v}_i = \frac{\mathbf{w}_i}{\|\mathbf{w}_i\|}.$$

That will give us an orthonormal set as required. The last part of the theorem follows if we begin with $\{\mathbf{u}_1, \ldots, \mathbf{u}_m\}$ as any basis for V.
Take $\mathbf{w}_1 := \mathbf{u}_1$.

Gram-Schmidt

Proof: The construction of the orthonormal set is algorithmic and of course, inductive. First we construct an intermediate orthogonal set $\{\mathbf{w}_1, \ldots, \mathbf{w}_n\}$ with the same property for each k. Then we simply 'normalize' each element, viz, take

$$\mathbf{v}_i = \frac{\mathbf{w}_i}{\|\mathbf{w}_i\|}.$$

That will give us an orthonormal set as required. The last part of the theorem follows if we begin with $\{\mathbf{u}_1, \ldots, \mathbf{u}_m\}$ as any basis for V.
Take $\mathbf{w}_1 := \mathbf{u}_1$. So, the construction is over for $k = 1$.

Gram-Schmidt

Proof: The construction of the orthonormal set is algorithmic and of course, inductive. First we construct an intermediate orthogonal set $\{\mathbf{w}_1, \ldots, \mathbf{w}_n\}$ with the same property for each k. Then we simply 'normalize' each element, viz, take

$$\mathbf{v}_i = \frac{\mathbf{w}_i}{\|\mathbf{w}_i\|}.$$

That will give us an orthonormal set as required. The last part of the theorem follows if we begin with $\{\mathbf{u}_1, \ldots, \mathbf{u}_m\}$ as any basis for V.
Take $\mathbf{w}_1 := \mathbf{u}_1$. So, the construction is over for $k = 1$. To construct $\mathbf{w}_2$, subtract from $\mathbf{u}_2$, its orthogonal projection along $\mathbf{w}_1$.

Gram-Schmidt

Proof: The construction of the orthonormal set is algorithmic and of course, inductive. First we construct an intermediate orthogonal set $\{\mathbf{w}_1, \ldots, \mathbf{w}_n\}$ with the same property for each k. Then we simply 'normalize' each element, viz, take

$$\mathbf{v}_i = \frac{\mathbf{w}_i}{\|\mathbf{w}_i\|}.$$

That will give us an orthonormal set as required. The last part of the theorem follows if we begin with $\{\mathbf{u}_1, \ldots, \mathbf{u}_m\}$ as any basis for V.
Take $\mathbf{w}_1 := \mathbf{u}_1$. So, the construction is over for $k = 1$. To construct $\mathbf{w}_2$, subtract from $\mathbf{u}_2$, its orthogonal projection along $\mathbf{w}_1$. Thus,

$$\mathbf{w}_2 = \mathbf{u}_2 - \langle \mathbf{w}_1, \mathbf{u}_2 \rangle \frac{\mathbf{w}_1}{\|\mathbf{w}_1\|^2}.$$

Gram-Schmidt

Check that $\langle \mathbf{w}_2, \mathbf{w}_1 \rangle = 0$.

Gram-Schmidt

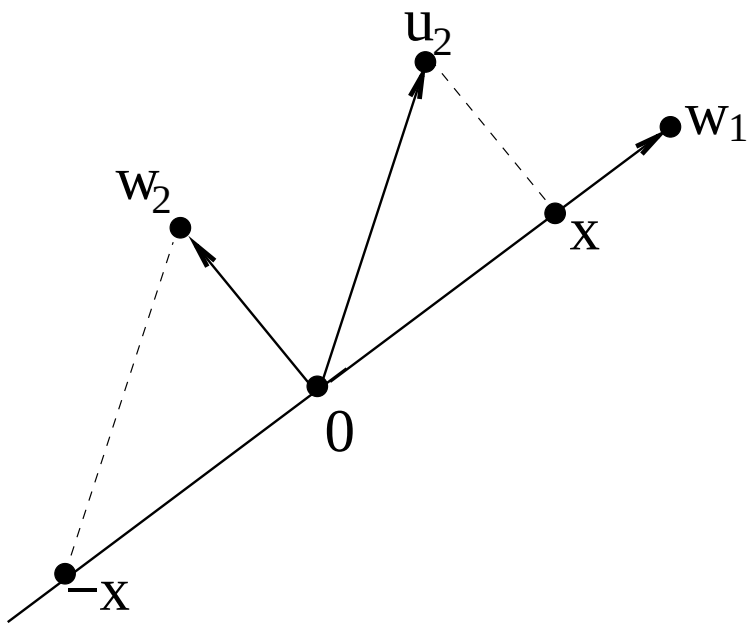
Check that $\langle \mathbf{w}_2, \mathbf{w}_1 \rangle = 0$. Thus $\{\mathbf{w}_1, \mathbf{w}_2\}$ is an orthogonal set.

Gram-Schmidt

Check that $\langle \mathbf{w}_2, \mathbf{w}_1 \rangle = 0$. Thus $\{\mathbf{w}_1, \mathbf{w}_2\}$ is an orthogonal set.
Check also that $L(\{\mathbf{u}_1, \mathbf{u}_2\}) = L(\{\mathbf{w}_1, \mathbf{w}_2\})$.

Gram-Schmidt

Check that $\langle \mathbf{w}_2, \mathbf{w}_1 \rangle = 0$. Thus $\{\mathbf{w}_1, \mathbf{w}_2\}$ is an orthogonal set.
Check also that $L(\{\mathbf{u}_1, \mathbf{u}_2\}) = L(\{\mathbf{w}_1, \mathbf{w}_2\})$.



Gram-Schmidt: Continued

Now suppose we have constructed $\mathbf{w}_i$ for $i \leq k < n$ as required.

Gram-Schmidt: Continued

Now suppose we have constructed $\mathbf{w}_i$ for $i \leq k < n$ as required. Put

$$\mathbf{w}_{k+1} = \mathbf{u}_{k+1} - \sum_{j=1}^{k} \langle \mathbf{w}_j, \mathbf{u}_{k+1} \rangle \frac{\mathbf{w}_j}{\|\mathbf{w}_j\|^2}.$$

Gram-Schmidt: Continued

Now suppose we have constructed $\mathbf{w}_i$ for $i \leq k < n$ as required. Put

$$\mathbf{w}_{k+1} = \mathbf{u}_{k+1} - \sum_{j=1}^{k} \langle \mathbf{w}_j, \mathbf{u}_{k+1} \rangle \frac{\mathbf{w}_j}{\|\mathbf{w}_j\|^2}.$$

Now check that $\mathbf{w}_{k+1}$ is orthogonal to $\mathbf{w}_1, \ldots, \mathbf{w}_k$ and $L(\{\mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_{k+1}\}) = L(\{\mathbf{u}_1, \ldots, \mathbf{u}_{k+1}\})$.

Gram-Schmidt: Continued

Now suppose we have constructed $\mathbf{w}_i$ for $i \leq k < n$ as required. Put

$$\mathbf{w}_{k+1} = \mathbf{u}_{k+1} - \sum_{j=1}^{k} \langle \mathbf{w}_j, \mathbf{u}_{k+1} \rangle \frac{\mathbf{w}_j}{\|\mathbf{w}_j\|^2}.$$

Now check that $\mathbf{w}_{k+1}$ is orthogonal to $\mathbf{w}_1, \ldots, \mathbf{w}_k$ and
$L(\{\mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_{k+1}\}) = L(\{\mathbf{u}_1, \ldots, \mathbf{u}_{k+1}\})$.
$[L(\{\mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_{k+1}\}) = L(\{\mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_k\} \cup \{\mathbf{w}_{k+1}\}) =$
$L(\{\mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_k\} \cup \{\mathbf{u}_{k+1}\}) = L(\{\mathbf{u}_1, \ldots, \mathbf{u}_k, \mathbf{u}_{k+1}\})$.]
By induction, this completes the proof. ♠

Gram-Schmidt: An example

Example

Let $V = P_3[-1,-1]$ *denote the real vector space of polynomials of degree at most 3 defined on* $[-1,1]$, *together with the zero polynomial.*

Gram-Schmidt: An example

Example

Let $V = P_3[-1,-1]$ *denote the real vector space of polynomials of degree at most 3 defined on* $[-1,1]$, *together with the zero polynomial.* V *is an inner product space under the inner product*

$$\langle f, g \rangle = \int_{-1}^{1} f(t)g(t)\,dt.$$

Gram-Schmidt: An example

Example

Let $V = P_3[-1,-1]$ *denote the real vector space of polynomials of degree at most 3 defined on* $[-1,1]$, *together with the zero polynomial.* V *is an inner product space under the inner product*

$$\langle f, g\rangle = \int_{-1}^{1} f(t)g(t)\,dt.$$

To find an orthonormal basis, we begin with the basis $\{1, x, x^2, x^3\}$.

Gram-Schmidt: An example

Set $\mathbf{w}_1 = 1$. Then

$$\mathbf{w}_2 \quad = \quad x - \langle x, 1\rangle \frac{1}{\|1\|^2} = x - \frac{1}{2}\int_{-1}^{1} t dt = x,$$

Gram-Schmidt: An example

Set $\mathbf{w}_1 = 1$. Then

$$\begin{aligned}
\mathbf{w}_2 &= x - \langle x, 1\rangle \frac{1}{\|1\|^2} = x - \frac{1}{2}\int_{-1}^{1} t dt = x, \\
\mathbf{w}_3 &= x^2 - \langle x^2, 1\rangle \frac{1}{2} - \langle x^2, x\rangle \frac{x}{(2/3)} \\
&= x^2 - \frac{1}{2}\int_{-1}^{1} t^2 dt - \frac{3}{2}x\int_{-1}^{1} t^3 dt \\
&= x^2 - \frac{1}{3},
\end{aligned}$$

Gram-Schmidt: An example

Set $\mathbf{w}_1 = 1$. Then

$$\begin{aligned}
\mathbf{w}_2 &= x - \langle x, 1\rangle \frac{1}{\|1\|^2} = x - \frac{1}{2}\int_{-1}^{1} t dt = x, \\
\mathbf{w}_3 &= x^2 - \langle x^2, 1\rangle \frac{1}{2} - \langle x^2, x\rangle \frac{x}{(2/3)} \\
&= x^2 - \frac{1}{2}\int_{-1}^{1} t^2 dt - \frac{3}{2}x\int_{-1}^{1} t^3 dt \\
&= x^2 - \frac{1}{3}, \\
\mathbf{w}_4 &= x^3 - \langle x^3, 1\rangle \frac{1}{2} - \langle x^3, x\rangle \frac{x}{(2/3)} - \langle x^3, x^2 - \frac{1}{3}\rangle \frac{x^2 - \frac{1}{3}}{(2/5)} \\
&= x^3 - \frac{3}{5}x.
\end{aligned}$$

Gram-Schmidt

Thus $\{1, x, x^2 - \frac{1}{3}, x^3 - \frac{3}{5}x\}$ is an orthogonal basis.

Gram-Schmidt

Thus $\{1, x, x^2 - \frac{1}{3}, x^3 - \frac{3}{5}x\}$ is an orthogonal basis. We divide these by respective norms to get an orthonormal basis.

$$\left\{ \frac{1}{\sqrt{2}},\ x\sqrt{\frac{3}{2}},\ (x^2 - \frac{1}{3})\frac{3\sqrt{5}}{2\sqrt{2}},\ (x^3 - \frac{3}{5}x)\frac{5\sqrt{7}}{2\sqrt{2}} \right\}.$$

Gram-Schmidt

Thus $\{1, x, x^2 - \frac{1}{3}, x^3 - \frac{3}{5}x\}$ is an orthogonal basis. We divide these by respective norms to get an orthonormal basis.

$$\left\{ \frac{1}{\sqrt{2}},\ x\sqrt{\frac{3}{2}},\ (x^2 - \frac{1}{3})\frac{3\sqrt{5}}{2\sqrt{2}},\ (x^3 - \frac{3}{5}x)\frac{5\sqrt{7}}{2\sqrt{2}} \right\}.$$

You will meet these polynomials while studying differential equations.

Orthogonal Complement

Definition

Let V be an inner product space. Given $\mathbf{v}, \mathbf{w} \in V$ we say $\mathbf{v}$ is perpendicular to $\mathbf{W}$ and write $\mathbf{v} \perp \mathbf{w}$ if $\langle \mathbf{v}, \mathbf{w} \rangle = 0$.

Orthogonal Complement

Definition

Let V be an inner product space. Given $\mathbf{v}, \mathbf{w} \in V$ we say $\mathbf{v}$ is perpendicular to $\mathbf{W}$ and write $\mathbf{v} \perp \mathbf{w}$ if $\langle \mathbf{v}, \mathbf{w} \rangle = 0$. Given a subspace W of V, we define

$$W^{\perp} := \{\mathbf{v} \in V \ : \ \mathbf{v} \perp \mathbf{w}, \ \forall \ \mathbf{w} \in W\}$$

and call it the *orthogonal complement* of W.

Orthogonal Complement

Definition

Let V be an inner product space. Given $\mathbf{v}, \mathbf{w} \in V$ we say $\mathbf{v}$ is perpendicular to $\mathbf{W}$ and write $\mathbf{v} \perp \mathbf{w}$ if $\langle \mathbf{v}, \mathbf{w} \rangle = 0$. Given a subspace W of V, we define

$$W^{\perp} := \{\mathbf{v} \in V \ : \ \mathbf{v} \perp \mathbf{w}, \ \forall \ \mathbf{w} \in W\}$$

and call it the *orthogonal complement* of W.

Orthogonal Complement

Proposition

Let W be a subspace of an inner product space. Then

Orthogonal Complement

Proposition

Let W be a subspace of an inner product space. Then
(i) $W^{\perp}$ is a subspace of V;

Orthogonal Complement

Proposition

Let W be a subspace of an inner product space. Then
(i) $W^{\perp}$ is a subspace of V;
(ii) $W \cap W^{\perp} = (0)$;

Orthogonal Complement

Proposition

Let W be a subspace of an inner product space. Then
(i) $W^{\perp}$ is a subspace of V;
(ii) $W \cap W^{\perp} = (0)$;
(iii) $W \subset (W^{\perp})^{\perp}$;

Orthogonal Complement

Proposition

Let W be a subspace of an inner product space. Then
(i) $W^{\perp}$ is a subspace of V;
(ii) $W \cap W^{\perp} = (0)$;
(iii) $W \subset (W^{\perp})^{\perp}$;

Proof: All the proofs are easy.

Orthogonal Complement

Proposition

Let W be a subspace of an inner product space. Then
(i) $W^{\perp}$ is a subspace of V;
(ii) $W \cap W^{\perp} = (0)$;
(iii) $W \subset (W^{\perp})^{\perp}$;

Proof: All the proofs are easy.

Orthogonal Complement

Theorem

Let W be a finite dimensional subspace of an inner product space.

Orthogonal Complement

Theorem

Let W be a finite dimensional subspace of an inner product space. Then every element $\mathbf{v} \in V$ is expressible as a unique sum

$$\mathbf{v} = \mathbf{w} + \mathbf{w}^{\perp}$$

Orthogonal Complement

Theorem

Let W be a finite dimensional subspace of an inner product space. Then every element $\mathbf{v} \in V$ is expressible as a unique sum

$$\mathbf{v} = \mathbf{w} + \mathbf{w}^{\perp}$$

where $\mathbf{w} \in W$ and $\mathbf{w}^{\perp} \in W^{\perp}$.

Orthogonal Complement

Theorem

Let W be a finite dimensional subspace of an inner product space. Then every element $\mathbf{v} \in V$ is expressible as a unique sum

$$\mathbf{v} = \mathbf{w} + \mathbf{w}^{\perp}$$

where $\mathbf{w} \in W$ and $\mathbf{w}^{\perp} \in W^{\perp}$. Further, $\mathbf{w}$ is characterised by the property that it is the near most point in W to $\mathbf{v}$.

Orthogonal Complement

Theorem

Let W be a finite dimensional subspace of an inner product space. Then every element $\mathbf{v} \in V$ is expressible as a unique sum

$$\mathbf{v} = \mathbf{w} + \mathbf{w}^{\perp}$$

where $\mathbf{w} \in W$ and $\mathbf{w}^{\perp} \in W^{\perp}$. Further, $\mathbf{w}$ is characterised by the property that it is the near most point in W to $\mathbf{v}$.

Proof: Fix an orthonormal basis $\{\mathbf{w}_1, \ldots, \mathbf{w}_k\}$ for W.

Orthogonal Complement

Theorem

Let W be a finite dimensional subspace of an inner product space. Then every element $\mathbf{v} \in V$ is expressible as a unique sum

$$\mathbf{v} = \mathbf{w} + \mathbf{w}^{\perp}$$

where $\mathbf{w} \in W$ and $\mathbf{w}^{\perp} \in W^{\perp}$. Further, $\mathbf{w}$ is characterised by the property that it is the near most point in W to $\mathbf{v}$.

Proof: Fix an orthonormal basis $\{\mathbf{w}_1, \ldots, \mathbf{w}_k\}$ for W. Define $\mathbf{w} = \sum_{i=1}^{k} \langle \mathbf{w}_i, \mathbf{v} \rangle \mathbf{w}_i$.

Orthogonal Complement

Theorem

Let W be a finite dimensional subspace of an inner product space. Then every element $\mathbf{v} \in V$ is expressible as a unique sum

$$\mathbf{v} = \mathbf{w} + \mathbf{w}^{\perp}$$

where $\mathbf{w} \in W$ and $\mathbf{w}^{\perp} \in W^{\perp}$. Further, $\mathbf{w}$ is characterised by the property that it is the near most point in W to $\mathbf{v}$.

Proof: Fix an orthonormal basis $\{\mathbf{w}_1, \ldots, \mathbf{w}_k\}$ for W. Define $\mathbf{w} = \sum_{i=1}^{k} \langle \mathbf{w}_i, \mathbf{v} \rangle \mathbf{w}_i$. Take

$$\mathbf{w}^{\perp} = \mathbf{v} - \mathbf{w} \tag{2}$$

and verify that $\mathbf{w}^{\perp} \in W^{\perp}$.

Orthogonal Complement

Theorem

Let W be a finite dimensional subspace of an inner product space. Then every element $\mathbf{v} \in V$ is expressible as a unique sum

$$\mathbf{v} = \mathbf{w} + \mathbf{w}^{\perp}$$

where $\mathbf{w} \in W$ and $\mathbf{w}^{\perp} \in W^{\perp}$. Further, $\mathbf{w}$ is characterised by the property that it is the near most point in W to $\mathbf{v}$.

Proof: Fix an orthonormal basis $\{\mathbf{w}_1, \ldots, \mathbf{w}_k\}$ for W. Define $\mathbf{w} = \sum_{i=1}^{k} \langle \mathbf{w}_i, \mathbf{v} \rangle \mathbf{w}_i$. Take

$$\mathbf{w}^{\perp} = \mathbf{v} - \mathbf{w} \tag{2}$$

and verify that $\mathbf{w}^{\perp} \in W^{\perp}$.
This proves the existence of such a decomposition.

Orthogonal Complement

Theorem

Let W be a finite dimensional subspace of an inner product space. Then every element $\mathbf{v} \in V$ is expressible as a unique sum

$$\mathbf{v} = \mathbf{w} + \mathbf{w}^{\perp}$$

where $\mathbf{w} \in W$ and $\mathbf{w}^{\perp} \in W^{\perp}$. Further, $\mathbf{w}$ is characterised by the property that it is the near most point in W to $\mathbf{v}$.

Proof: Fix an orthonormal basis $\{\mathbf{w}_1, \ldots, \mathbf{w}_k\}$ for W. Define $\mathbf{w} = \sum_{i=1}^{k} \langle \mathbf{w}_i, \mathbf{v} \rangle \mathbf{w}_i$. Take

$$\mathbf{w}^{\perp} = \mathbf{v} - \mathbf{w} \tag{2}$$

and verify that $\mathbf{w}^{\perp} \in W^{\perp}$.
This proves the existence of such a decomposition.
Now suppose, we have $\mathbf{v} = \mathbf{u} + \mathbf{u}^{\perp}$ is another similar expression.

Orthogonal Complement

Theorem

Let W be a finite dimensional subspace of an inner product space. Then every element $\mathbf{v} \in V$ is expressible as a unique sum

$$\mathbf{v} = \mathbf{w} + \mathbf{w}^{\perp}$$

where $\mathbf{w} \in W$ and $\mathbf{w}^{\perp} \in W^{\perp}$. Further, $\mathbf{w}$ is characterised by the property that it is the near most point in W to $\mathbf{v}$.

Proof: Fix an orthonormal basis $\{\mathbf{w}_1, \ldots, \mathbf{w}_k\}$ for W. Define $\mathbf{w} = \sum_{i=1}^{k} \langle \mathbf{w}_i, \mathbf{v} \rangle \mathbf{w}_i$. Take

$$\mathbf{w}^{\perp} = \mathbf{v} - \mathbf{w} \tag{2}$$

and verify that $\mathbf{w}^{\perp} \in W^{\perp}$.
This proves the existence of such a decomposition.
Now suppose, we have $\mathbf{v} = \mathbf{u} + \mathbf{u}^{\perp}$ is another similar expression. Then it follows that $\mathbf{w} - \mathbf{u} = \mathbf{u}^{\perp} - \mathbf{w}^{\perp} \in W \cap W^{\perp} = (0)$.

Orthogonal Complement

Theorem

Let W be a finite dimensional subspace of an inner product space. Then every element $\mathbf{v} \in V$ is expressible as a unique sum

$$\mathbf{v} = \mathbf{w} + \mathbf{w}^{\perp}$$

where $\mathbf{w} \in W$ and $\mathbf{w}^{\perp} \in W^{\perp}$. Further, $\mathbf{w}$ is characterised by the property that it is the near most point in W to $\mathbf{v}$.

Proof: Fix an orthonormal basis $\{\mathbf{w}_1, \ldots, \mathbf{w}_k\}$ for W. Define $\mathbf{w} = \sum_{i=1}^{k} \langle \mathbf{w}_i, \mathbf{v} \rangle \mathbf{w}_i$. Take

$$\mathbf{w}^{\perp} = \mathbf{v} - \mathbf{w} \tag{2}$$

and verify that $\mathbf{w}^{\perp} \in W^{\perp}$.
This proves the existence of such a decomposition.
Now suppose, we have $\mathbf{v} = \mathbf{u} + \mathbf{u}^{\perp}$ is another similar expression. Then it follows that $\mathbf{w} - \mathbf{u} = \mathbf{u}^{\perp} - \mathbf{w}^{\perp} \in W \cap W^{\perp} = (0)$. Therefore $\mathbf{w} = \mathbf{u}$ and $\mathbf{w}^{\perp} = \mathbf{u}^{\perp}$.

Orthogonal Complement

Theorem

Let W be a finite dimensional subspace of an inner product space. Then every element $\mathbf{v} \in V$ is expressible as a unique sum

$$\mathbf{v} = \mathbf{w} + \mathbf{w}^{\perp}$$

where $\mathbf{w} \in W$ and $\mathbf{w}^{\perp} \in W^{\perp}$. Further, $\mathbf{w}$ is characterised by the property that it is the near most point in W to $\mathbf{v}$.

Proof: Fix an orthonormal basis $\{\mathbf{w}_1, \ldots, \mathbf{w}_k\}$ for W. Define $\mathbf{w} = \sum_{i=1}^{k} \langle \mathbf{w}_i, \mathbf{v} \rangle \mathbf{w}_i$. Take

$$\mathbf{w}^{\perp} = \mathbf{v} - \mathbf{w} \tag{2}$$

and verify that $\mathbf{w}^{\perp} \in W^{\perp}$.
This proves the existence of such a decomposition.
Now suppose, we have $\mathbf{v} = \mathbf{u} + \mathbf{u}^{\perp}$ is another similar expression. Then it follows that $\mathbf{w} - \mathbf{u} = \mathbf{u}^{\perp} - \mathbf{w}^{\perp} \in W \cap W^{\perp} = (0)$. Therefore $\mathbf{w} = \mathbf{u}$ and $\mathbf{w}^{\perp} = \mathbf{u}^{\perp}$. This proves the uniqueness.

Orthogonal Complement

Finally, let $\mathbf{u} \in W$ be any vector.

Orthogonal Complement

Finally, let $\mathbf{u} \in W$ be any vector. Then $\mathbf{v} - \mathbf{u} = (\mathbf{v} - \mathbf{w}) + (\mathbf{w} - \mathbf{u})$. Since $\mathbf{w} - \mathbf{u} \in W$ and $\mathbf{v} - \mathbf{w} \in W^{\perp}$, we can apply Pythagorus theorem

Orthogonal Complement

Finally, let $\mathbf{u} \in W$ be any vector. Then $\mathbf{v} - \mathbf{u} = (\mathbf{v} - \mathbf{w}) + (\mathbf{w} - \mathbf{u})$. Since $\mathbf{w} - \mathbf{u} \in W$ and $\mathbf{v} - \mathbf{w} \in W^{\perp}$, we can apply Pythagorus theorem to conclude that $\|\mathbf{v} - \mathbf{u}\|^2 \geq \|\mathbf{v} - \mathbf{w}\|^2$ and equality holds iff $\mathbf{u} = \mathbf{w}$.

Orthogonal Complement

Finally, let $\mathbf{u} \in W$ be any vector. Then $\mathbf{v} - \mathbf{u} = (\mathbf{v} - \mathbf{w}) + (\mathbf{w} - \mathbf{u})$. Since $\mathbf{w} - \mathbf{u} \in W$ and $\mathbf{v} - \mathbf{w} \in W^{\perp}$, we can apply Pythagorus theorem to conclude that $\|\mathbf{v} - \mathbf{u}\|^2 \geq \|\mathbf{v} - \mathbf{w}\|^2$ and equality holds iff $\mathbf{u} = \mathbf{w}$. This proves that $\mathbf{w}$ is the (only) nearest point to $\mathbf{v}$ in W. ♠

Orthogonal Projection

Remark

(i) The element **w** so obtained is called the *orthogonal projection* of **v** on W.

Orthogonal Projection

Remark

(i) The element $\mathbf{w}$ so obtained is called the *orthogonal projection* of $\mathbf{v}$ on W. Using the expression (2)

$$P_W(\mathbf{v}) = \sum_{i=1}^{k} \langle \mathbf{w}_i, \mathbf{v} \rangle \mathbf{w}_i$$

as obtained in the provious theorem, one can easily check that the assignment $\mathbf{v} \mapsto \mathbf{w}$ itself defines a linear map $P_W : V \longrightarrow W$. Also observe that P_W is identity on W, i.e., $P_W(\mathbf{w}) = \mathbf{w}$ for every $\mathbf{w} \in W$.

Orthogonal Projection

Remark

(i) The element $\mathbf{w}$ so obtained is called the *orthogonal projection* of $\mathbf{v}$ on W. Using the expression (2)

$$P_W(\mathbf{v}) = \sum_{i=1}^{k} \langle \mathbf{w}_i, \mathbf{v} \rangle \mathbf{w}_i$$

as obtained in the provious theorem, one can easily check that the assignment $\mathbf{v} \mapsto \mathbf{w}$ itself defines a linear map $P_W : V \longrightarrow W$. Also observe that P_W is identity on W, i.e., $P_W(\mathbf{w}) = \mathbf{w}$ for every $\mathbf{w} \in W$. Hence $P_W^2 := P_W \circ P_W = P_W$.

Orthogonal Projection

Remark

(i) The element $\mathbf{w}$ so obtained is called the *orthogonal projection* of $\mathbf{v}$ on W. Using the expression (2)

$$P_W(\mathbf{v}) = \sum_{i=1}^{k} \langle \mathbf{w}_i, \mathbf{v} \rangle \mathbf{w}_i$$

as obtained in the provious theorem, one can easily check that the assignment $\mathbf{v} \mapsto \mathbf{w}$ itself defines a linear map $P_W : V \longrightarrow W$. Also observe that P_W is identity on W, i.e., $P_W(\mathbf{w}) = \mathbf{w}$ for every $\mathbf{w} \in W$. Hence $P_W^2 := P_W \circ P_W = P_W$.
(ii) Observe that if V itself is finite dimensional, then it follows that the above proposition is applicable to all subspaces in V.

Orthogonal Projection

Remark

(i) The element $\mathbf{w}$ so obtained is called the *orthogonal projection* of $\mathbf{v}$ on W. Using the expression (2)

$$P_W(\mathbf{v}) = \sum_{i=1}^{k} \langle \mathbf{w}_i, \mathbf{v} \rangle \mathbf{w}_i$$

as obtained in the provious theorem, one can easily check that the assignment $\mathbf{v} \mapsto \mathbf{w}$ itself defines a linear map $P_W : V \longrightarrow W$. Also observe that P_W is identity on W, i.e., $P_W(\mathbf{w}) = \mathbf{w}$ for every $\mathbf{w} \in W$. Hence $P_W^2 := P_W \circ P_W = P_W$.
(ii) Observe that if V itself is finite dimensional, then it follows that the above proposition is applicable to all subspaces in V. In this case it easily follows that $dim\, W + dim\, W^{\perp} = dim\, V$ (see exercise at the end of section Linear Dependence.)

Orthogonal Projection

Example

As an application of orthogonal projection, let us consider the following problem.

Orthogonal Projection

Example

As an application of orthogonal projection, let us consider the following problem. Suppose we are given a system of linear equations and a likely candidate for the solution.

Orthogonal Projection

Example

As an application of orthogonal projection, let us consider the following problem. Suppose we are given a system of linear equations and a likely candidate for the solution. If this candidate is actually a solution well and good.

Orthogonal Projection

Example

As an application of orthogonal projection, let us consider the following problem. Suppose we are given a system of linear equations and a likely candidate for the solution. If this candidate is actually a solution well and good. Otherwise we would like to get a solution which is nearmost to the candidate.

Orthogonal Projection

Example

As an application of orthogonal projection, let us consider the following problem. Suppose we are given a system of linear equations and a likely candidate for the solution. If this candidate is actually a solution well and good. Otherwise we would like to get a solution which is nearmost to the candidate. This problem is a typical example of finding 'best approximations'.

Orthogonal Projection

Example

As an application of orthogonal projection, let us consider the following problem. Suppose we are given a system of linear equations and a likely candidate for the solution. If this candidate is actually a solution well and good. Otherwise we would like to get a solution which is nearmost to the candidate. This problem is a typical example of finding 'best approximations'. (It can also be discussed using Lagrange multipliers.)
Thus if the system is a homgeneous set of m linear equations involving n variables

Orthogonal Projection

Example

As an application of orthogonal projection, let us consider the following problem. Suppose we are given a system of linear equations and a likely candidate for the solution. If this candidate is actually a solution well and good. Otherwise we would like to get a solution which is nearmost to the candidate. This problem is a typical example of finding 'best approximations'. (It can also be discussed using Lagrange multipliers.)

Thus if the system is a homgeneous set of m linear equations involving n variables we get a linear transformation $f : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ *associated to it and the space of solutions is nothing but the subspace* $W = \mathcal{N}(f)$*, the null space.*

Orthogonal Projection

Example

As an application of orthogonal projection, let us consider the following problem. Suppose we are given a system of linear equations and a likely candidate for the solution. If this candidate is actually a solution well and good. Otherwise we would like to get a solution which is nearmost to the candidate. This problem is a typical example of finding 'best approximations'. (It can also be discussed using Lagrange multipliers.)
Thus if the system is a homgeneous set of m linear equations involving n variables we get a linear transformation $f : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ *associated to it and the space of solutions is nothing but the subspace* $W = \mathcal{N}(f)$*, the null space. Given a point* $\mathbf{v} \in \mathbb{R}^n$ *the nearmost point in* W *to* $\mathbf{v}$ *is nothing but the orthogonal projection of* $\mathbf{v}$ *on* W*.*

Orthogonal Projection

Example

As an application of orthogonal projection, let us consider the following problem. Suppose we are given a system of linear equations and a likely candidate for the solution. If this candidate is actually a solution well and good. Otherwise we would like to get a solution which is nearmost to the candidate. This problem is a typical example of finding 'best approximations'. (It can also be discussed using Lagrange multipliers.)
Thus if the system is a homgeneous set of m linear equations involving n variables we get a linear transformation $f : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ associated to it and the space of solutions is nothing but the subspace $W = \mathcal{N}(f)$, the null space. Given a point $\mathbf{v} \in \mathbb{R}^n$ the nearmost point in W to $\mathbf{v}$ is nothing but the orthogonal projection of $\mathbf{v}$ on W. In case the system is inhomogeneous, we still consider the linear map f defined by the homogeneous part and put $W = \mathcal{N}(f)$. Then we must first of all have one correct solution $\mathbf{p}$ of the inhomogeneous system. We then take $\mathbf{v} - \mathbf{p}$, take its projection on W as above and then add back $\mathbf{p}$ to it to get the nearmost solution to $\mathbf{v}$.

Orthogonal Projection

Problem (Method of Least Squares) Given a linear system of m equations in n variables $A\mathbf{x} = b$, and an n-vector $\mathbf{v}$, find the solution $\mathbf{y}$ such that $\|\mathbf{y} - \mathbf{v}\|$ is minimal.

Answer: Step 1. Apply GJEM to find a basis for the null-space W of A and a particular solution $\mathbf{p}$ of the system.

Step 2. Apply Gram-Schmidt process to the basis of W to get an orthonomal basis $\{\mathbf{v}_1, \ldots, \mathbf{v}_k\}$ for W.

Orthogonal Projection

Problem (Method of Least Squares) Given a linear system of m equations in n variables $A\mathbf{x} = b$, and an n-vector $\mathbf{v}$, find the solution $\mathbf{y}$ such that $\|\mathbf{y} - \mathbf{v}\|$ is minimal.

Answer: Step 1. Apply GJEM to find a basis for the null-space W of A and a particular solution $\mathbf{p}$ of the system.

Step 2. Apply Gram-Schmidt process to the basis of W to get an orthonomal basis $\{\mathbf{v}_1, \ldots, \mathbf{v}_k\}$ for W.

Step 3. Let $P_W(\mathbf{v} - \mathbf{p}) = \sum_{j=1}^{k} \langle (\mathbf{v} - \mathbf{p}), \mathbf{v}_j \rangle \mathbf{v}_j$ denote the orthogonal projection of $\mathbf{v} - \mathbf{p}$ on the space W.

Orthogonal Projection

Problem (Method of Least Squares) Given a linear system of m equations in n variables $A\mathbf{x} = b$, and an n-vector $\mathbf{v}$, find the solution $\mathbf{y}$ such that $\|\mathbf{y} - \mathbf{v}\|$ is minimal.

Answer: Step 1. Apply GJEM to find a basis for the null-space W of A and a particular solution $\mathbf{p}$ of the system.

Step 2. Apply Gram-Schmidt process to the basis of W to get an orthonomal basis $\{\mathbf{v}_1, \ldots, \mathbf{v}_k\}$ for W.

Step 3. Let $P_W(\mathbf{v} - \mathbf{p}) = \sum_{j=1}^{k} \langle(\mathbf{v} - \mathbf{p}), \mathbf{v}_j\rangle \mathbf{v}_j$ denote the orthogonal projection of $\mathbf{v} - \mathbf{p}$ on the space W. Take $\mathbf{y} = P_w(\mathbf{v} - \mathbf{p}) + \mathbf{p}$. (Verify that this is the required solution.)

Least-square problem

Example

Consider a system of three linear equations in 5 variables $x_1, \ldots, x_5$ given by

$$\begin{array}{rrrrcl} & 2x_3 & -2x_4 & -13x_5 & = & 2 \\ 2x_2 & -8x_3 & +14x_4 & -13x_5 & = & 2 \\ x_2 & +3x_3 & & +17x_5 & = & 8 \end{array}$$

(a) Write down a basis for the space of solution of the associated homogeneous system.
(b) Carry out Gram-Schmidt on this above basis.
(c) Given the vector $\mathbf{v} = (1, 5, 1, 0, 1)^t$, find the solution to the system which is nearmost to $\mathbf{v}$.
Solution: *The matrix form of the system is*

$$\left[\begin{array}{ccccc|c} 0 & 0 & 2 & -2 & -13 & 2 \\ 0 & 2 & -8 & 14 & -13 & 2 \\ 0 & 1 & 3 & 0 & 16 & 8 \end{array}\right]$$

Least-square problem; Example continued:

Applying GJEM, we see that the Gauss-Jordan form of this system is

$$\left[\begin{array}{ccccc|c} 0 & 1 & 0 & 3 & -1/2 & 5 \\ 0 & 0 & 1 & -1 & 3/2 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right]$$

Least-square problem; Example continued:

Applying GJEM, we see that the Gauss-Jordan form of this system is

$$\left[\begin{array}{ccccc|c} 0 & 1 & 0 & 3 & -1/2 & 5 \\ 0 & 0 & 1 & -1 & 3/2 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right]$$

Therefore the general solution is of the form

$$\begin{bmatrix} x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 5 & 0 & -3 & 1/2 \\ 1 & 0 & 1 & -1/2 \end{bmatrix} \begin{bmatrix} 1 \\ x_1 \\ x_4 \\ x_5 \end{bmatrix}.$$

Least-square problem; Example continued:

Applying GJEM, we see that the Gauss-Jordan form of this system is

$$\left[\begin{array}{ccccc|c} 0 & 1 & 0 & 3 & -1/2 & 5 \\ 0 & 0 & 1 & -1 & 3/2 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right]$$

Therefore the general solution is of the form

$$\begin{bmatrix} x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 5 & 0 & -3 & 1/2 \\ 1 & 0 & 1 & -1/2 \end{bmatrix} \begin{bmatrix} 1 \\ x_1 \\ x_4 \\ x_5 \end{bmatrix}.$$

A particular solution of the system is obtained by taking $x_1 = x_4 = x_5 = 0$ and hence $x_2 = 5, x_3 = 1$, i.e., $\mathbf{p} = (0, 5, 1, 0, 0)$.

Least-square problem; Example continued:

Applying GJEM, we see that the Gauss-Jordan form of this system is

$$\left[\begin{array}{ccccc|c} 0 & 1 & 0 & 3 & -1/2 & 5 \\ 0 & 0 & 1 & -1 & 3/2 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right]$$

Therefore the general solution is of the form

$$\begin{bmatrix} x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 5 & 0 & -3 & 1/2 \\ 1 & 0 & 1 & -1/2 \end{bmatrix} \begin{bmatrix} 1 \\ x_1 \\ x_4 \\ x_5 \end{bmatrix}.$$

A particular solution of the system is obtained by taking $x_1 = x_4 = x_5 = 0$ and hence $x_2 = 5, x_3 = 1$, i.e., $\mathbf{p} = (0, 5, 1, 0, 0)$.
Also the space W of solutions of the homogeneous system is given by

$$\begin{bmatrix} x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 5 & 0 & -3 & 1/2 \\ 1 & 0 & 1 & -3/2 \end{bmatrix} \begin{bmatrix} 0 \\ x_1 \\ x_4 \\ x_5 \end{bmatrix}.$$

Least-square problem

Therefore a basis for this space W (which is the null-space of the un-augemented matrix) is

$$\{(1,0,0,0,0)^t, (0,-3,1,1,0)^t, (0,-1/2,-3/2,0,1)^t\}$$

(b) We note that the given vectors are already orthogonal to each other. Hence Gram-Schmidt process just dives eac of them by their norm. This yields:

$$\{(1,0,0,0,0), \frac{1}{1}\sqrt{11}(0,-3,1,1,0), \sqrt{\frac{2}{7}}(0,-1/2,-3/2,0,1)\}$$

Least-square problem

Therefore a basis for this space W (which is the null-space of the un-augemented matrix) is

$$\{(1,0,0,0,0)^t, (0,-3,1,1,0)^t, (0,-1/2,-3/2,0,1)^t\}$$

(b) We note that the given vectors are already orthogonal to each other. Hence Gram-Schmidt process just dives eac of them by their norm. This yields:

$$\{(1,0,0,0,0), \frac{1}{1}\sqrt{11}(0,-3,1,1,0), \sqrt{\frac{2}{7}}(0,-1/2,-3/2,0,1)\}$$

(c) Now we consider the orthogonal projection of $\mathbf{v}-\mathbf{p} = (1,0,0,0,1)^t$ on the space W which is equal to

Least-square problem

Therefore a basis for this space W (which is the null-space of the un-augemented matrix) is

$$\{(1,0,0,0,0)^t, (0,-3,1,1,0)^t, (0,-1/2,-3/2,0,1)^t\}$$

(b) We note that the given vectors are already orthogonal to each other. Hence Gram-Schmidt process just dives eac of them by their norm. This yields:

$$\{(1,0,0,0,0), \frac{1}{1}\sqrt{11}(0,-3,1,1,0), \sqrt{\frac{2}{7}}(0,-1/2,-3/2,0,1)\}$$

(c) Now we consider the orthogonal projection of $\mathbf{v}-\mathbf{p} = (1,0,0,0,1)^t$ on the space W which is equal to
$\mathbf{w} = p_W(\mathbf{v}-\mathbf{p}) = (1,0,0,0,0)^t + \frac{2}{7}(0,-1/2,-3/2,0,1)^t$
$= (1,-1/7,-3/7,0,2/7)^t$

Least-square problem

Therefore a basis for this space W (which is the null-space of the un-augemented matrix) is

$$\{(1,0,0,0,0)^t, (0,-3,1,1,0)^t, (0,-1/2,-3/2,0,1)^t\}$$

(b) We note that the given vectors are already orthogonal to each other. Hence Gram-Schmidt process just dives eac of them by their norm. This yields:

$$\{(1,0,0,0,0), \frac{1}{1}\sqrt{11}(0,-3,1,1,0), \sqrt{\frac{2}{7}}(0,-1/2,-3/2,0,1)\}$$

(c) Now we consider the orthogonal projection of $\mathbf{v}-\mathbf{p} = (1,0,0,0,1)^t$ on the space W which is equal to
$\mathbf{w} = p_W(\mathbf{v}-\mathbf{p}) = (1,0,0,0,0)^t + \frac{2}{7}(0,-1/2,-3/2,0,1)^t$
$= (1,-1/7,-3/7,0,2/7)^t$
Now the solution to the original system which is nearmost to the given vector $\mathbf{v}$ is obtained by adding $\mathbf{p}$ to $\mathbf{w}$, viz. to
Answer: $(1,34/7,4/7,0,2/7)^t$.

Lagrange Interpolation formula

Example

As another example of an application of orthonormal bases we shall present Lagrange interpolation formula:

Lagrange Interpolation formula

Example

As another example of an application of orthonormal bases we shall present Lagrange interpolation formula:
Consider the $n+1$*-dimensional vector space* $\mathcal{P}_n[t]$ *of all polynomials of degree* $\neq n$ *in one variable* t *over the field* $\mathbb{K} = \mathbb{R}$ *or* $\mathbb{C}$.

Lagrange Interpolation formula

Example

As another example of an application of orthonormal bases we shall present Lagrange interpolation formula:
Consider the $n+1$*-dimensional vector space* $\mathcal{P}_n[t]$ *of all polynomials of degree* $\neq n$ *in one variable* t *over the field* $\mathbb{K} = \mathbb{R}$ *or* $\mathbb{C}$.
Fix any $n+1$ *distinct points* $x_0, x_1, \ldots, x_n \in \mathbb{K}$.

Lagrange Interpolation formula

Example

As another example of an application of orthonormal bases we shall present Lagrange interpolation formula:
Consider the $n+1$*-dimensional vector space* $\mathcal{P}_n[t]$ *of all polynomials of degree* $\neq n$ *in one variable* t *over the field* $\mathbb{K} = \mathbb{R}$ *or* $\mathbb{C}$.
Fix any $n+1$ *distinct points* $x_0, x_1, \ldots, x_n \in \mathbb{K}$.
Define

$$\langle f, g \rangle = \sum_{i=0}^{n} \overline{f(x_i)} g(x_i), \quad f, g \in \mathcal{P}_n[t].$$

Lagrange Interpolation formula

Example

As another example of an application of orthonormal bases we shall present Lagrange interpolation formula:
Consider the $n+1$*-dimensional vector space* $\mathcal{P}_n[t]$ *of all polynomials of degree* $\neq n$ *in one variable* t *over the field* $\mathbb{K} = \mathbb{R}$ *or* $\mathbb{C}$.
Fix any $n+1$ *distinct points* $x_0, x_1, \ldots, x_n \in \mathbb{K}$.
Define

$$\langle f, g \rangle = \sum_{i=0}^{n} \overline{f(x_i)} g(x_i), \quad f, g \in \mathcal{P}_n[t].$$

It can be easily verified that this defines an inner product on $\mathcal{P}_n[t]$. *The important thing to note here is that a a polynomial of degree n has more than n roots iff it is identically zero.*

Lagrange Interpolation formula

Example

(continued) For each $0 \leq k \leq n$, *consider the* **Lagrange Polynomials**

$$L_k(t) = \frac{\prod_{i \neq k}(t - x_i)}{\prod_{i \neq k}(x_k - x_i)}$$

each of degree n.

Lagrange Interpolation formula

Example

(continued) For each $0 \leq k \leq n$, *consider the* **Lagrange Polynomials**

$$L_k(t) = \frac{\prod_{i \neq k}(t - x_i)}{\prod_{i \neq k}(x_k - x_i)}$$

each of degree n.
Then check that
(i) $L_k(x_j) = 0$ *for all* $j \neq k$ *and* $L_k(x_k) = 1$.

Lagrange Interpolation formula

Example

(continued) For each $0 \leq k \leq n$, *consider the* **Lagrange Polynomials**

$$L_k(t) = \frac{\prod_{i \neq k}(t - x_i)}{\prod_{i \neq k}(x_k - x_i)}$$

each of degree n.
Then check that
(i) $L_k(x_j) = 0$ *for all* $j \neq k$ *and* $L_k(x_k) = 1$.
(ii) $\langle L_p, L_q \rangle = \delta_{pq}$, *i.e., the* $\{L_0, L_1, \ldots, L_n\}$ *is an orthonormal set.*

Lagrange Interpolation formula

Example

(continued) For each $0 \leq k \leq n$, *consider the* **Lagrange Polynomials**

$$L_k(t) = \frac{\prod_{i \neq k}(t - x_i)}{\prod_{i \neq k}(x_k - x_i)}$$

each of degree n.
Then check that
(i) $L_k(x_j) = 0$ *for all* $j \neq k$ *and* $L_k(x_k) = 1$.
(ii) $\langle L_p, L_q \rangle = \delta_{pq}$, *i.e., the* $\{L_0, L_1, \ldots, L_n\}$ *is an orthonormal set.*
In particular, these are linearly independent.

Lagrange Interpolation formula

Example

(continued) For each $0 \leq k \leq n$, *consider the* **Lagrange Polynomials**

$$L_k(t) = \frac{\prod_{i \neq k}(t - x_i)}{\prod_{i \neq k}(x_k - x_i)}$$

each of degree n.
Then check that
(i) $L_k(x_j) = 0$ *for all* $j \neq k$ *and* $L_k(x_k) = 1$.
(ii) $\langle L_p, L_q \rangle = \delta_{pq}$, *i.e., the* $\{L_0, L_1, \ldots, L_n\}$ *is an orthonormal set.*
In particular, these are linearly independent. Since the dimension of $\mathcal{P}_n[t]$ *is equal to* $n + 1$, *it follows that they form an orthonormal basis.*

Lagrange Interpolation formula

Example

(continued) For each $0 \leq k \leq n$, *consider the* **Lagrange Polynomials**

$$L_k(t) = \frac{\prod_{i \neq k}(t - x_i)}{\prod_{i \neq k}(x_k - x_i)}$$

each of degree n.
Then check that
(i) $L_k(x_j) = 0$ *for all* $j \neq k$ *and* $L_k(x_k) = 1$.
(ii) $\langle L_p, L_q \rangle = \delta_{pq}$, *i.e., the* $\{L_0, L_1, \ldots, L_n\}$ *is an orthonormal set.*
In particular, these are linearly independent. Since the dimension of $\mathcal{P}_n[t]$ *is equal to* $n+1$, *it follows that they form an orthonormal basis.*
(iii) Given any polynomial $P \in P_n[t]$,

$$\langle P, L_k \rangle = P(x_k).$$

Lagrange Interpolation formula

Example

(continued)
Therefore it follows that each P is a unique linear combination of the Lagrange Polynomials

$$P(t) = \sum_i P(x_i) L_i(t).$$

Lagrange Interpolation formula

Example

(continued)
Therefore it follows that each P is a unique linear combination of the Lagrange Polynomials

$$P(t) = \sum_i P(x_i)L_i(t).$$

In general, for any arbitrary function $f : \mathbb{K} \to \mathbb{K}$ the formula

$$P_f(t) = \sum_i P(x_i)L_i(t)$$

is known as the Lagrange interpolation formula which gives a polynomial function which has the same value as the given function at the specific points $x_0, \ldots, x_n$. This is very useful in numerical analysis.

Eigenvalues and Eigenvectors: Introduction

The simplest of matrices are the diagonal ones.

Eigenvalues and Eigenvectors: Introduction

The simplest of matrices are the diagonal ones. Thus a linear map will be also easy to handle if its associated matrix is a diagonal matrix.

Eigenvalues and Eigenvectors: Introduction

The simplest of matrices are the diagonal ones. Thus a linear map will be also easy to handle if its associated matrix is a diagonal matrix.
Then again we have seen that the matrix associated depends upon the choice of the bases to some extent.

Eigenvalues and Eigenvectors: Introduction

The simplest of matrices are the diagonal ones. Thus a linear map will be also easy to handle if its associated matrix is a diagonal matrix.
Then again we have seen that the matrix associated depends upon the choice of the bases to some extent. This naturally leads us to the problem of investigating the existence and construction of a suitable basis with respect to which the matrix associated to a given linear transformation is diagonal.

Problem of Diagonalizability

Definition

An $n \times n$ matrix A is called diagonalizable if there exists an invertible $n \times n$ matrix M such that $M^{-1}AM$ is a diagonal matrix.

Problem of Diagonalizability

Definition

An $n \times n$ matrix A is called diagonalizable if there exists an invertible $n \times n$ matrix M such that $M^{-1}AM$ is a diagonal matrix. In other words, a square matrix A is diagonalizable if it is similar to a diagonal matrix.

Remark

For $1 \leq i \leq n$, put $\mathbf{v}_i = M\mathbf{e}_i$, the i^{th} column vector of the matrix M. Then we know M is invertible iff $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ forms a basis for $\mathbb{K}^n$. Put $M^{-1}AM = D = diag(\lambda_1, \ldots, \lambda_n)$ where $\lambda_i \in \mathbb{K}$. We see that

$$M^{-1}AM = D \iff AM = MD \iff AM\mathbf{e}_i = MD\mathbf{e}_i = \lambda_i M\mathbf{e}_i, 1 \leq i \leq n.$$

$$\iff A\mathbf{v}_i = \lambda_i \mathbf{v}_i, 1 \leq i \leq n.$$

Therefore we have following complete answer:

Diagonalization: A Reformulation

Proposition

An $n \times n$ matrix A is diagonalizable iff there exists a basis $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ of $\mathbb{K}^n$ such that $A\mathbf{v}_i = \lambda_i \mathbf{v}_i$ for some $\lambda_i \in \mathbb{K}$.

Definition

A linear map $f : V \longrightarrow V$ is called diagonalizable if the matrix associated to f with respect to some basis is diagonal.

Diagonalization: A Reformulation

Proposition

An $n \times n$ matrix A is diagonalizable iff there exists a basis $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ of $\mathbb{K}^n$ such that $A\mathbf{v}_i = \lambda_i \mathbf{v}_i$ for some $\lambda_i \in \mathbb{K}$.

Definition

A linear map $f : V \longrightarrow V$ is called diagonalizable if the matrix associated to f with respect to some basis is diagonal.

Remark

Let $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ be a basis. The matrix M_f of a linear transformation f w.r.t. this basis is diagonal iff $f(\mathbf{v}_i) = \lambda_i \mathbf{v}_i,\ 1 \leq i \leq n$ for some scalars λ_i.

Diagonalization: A Reformulation

Proposition

An $n \times n$ matrix A is diagonalizable iff there exists a basis $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ of $\mathbb{K}^n$ such that $A\mathbf{v}_i = \lambda_i \mathbf{v}_i$ for some $\lambda_i \in \mathbb{K}$.

Definition

A linear map $f : V \longrightarrow V$ is called diagonalizable if the matrix associated to f with respect to some basis is diagonal.

Remark

Let $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ be a basis. The matrix M_f of a linear transformation f w.r.t. this basis is diagonal iff $f(\mathbf{v}_i) = \lambda_i \mathbf{v}_i,\ 1 \leq i \leq n$ for some scalars λ_i. Thus diagonalization question can be put in the following way:

Diagonalization: A Reformulation

Proposition

An $n \times n$ matrix A is diagonalizable iff there exists a basis $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ of $\mathbb{K}^n$ such that $A\mathbf{v}_i = \lambda_i \mathbf{v}_i$ for some $\lambda_i \in \mathbb{K}$.

Definition

A linear map $f : V \longrightarrow V$ is called diagonalizable if the matrix associated to f with respect to some basis is diagonal.

Remark

Let $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ be a basis. The matrix M_f of a linear transformation f w.r.t. this basis is diagonal iff $f(\mathbf{v}_i) = \lambda_i \mathbf{v}_i,\ 1 \leq i \leq n$ for some scalars λ_i. Thus diagonalization question can be put in the following way:
Given a linear map $f : V \to V$, does there exist a basis for V consisting of vectors $\mathbf{v}$ such that $f(\mathbf{v}) = \lambda_{\mathbf{v}} \mathbf{v}$?

Diagonalizability: Another Reformulation

Definition

Given a linear map $f : V \longrightarrow V$ we say $\mathbf{v} \in V$ is an eigenvector for f if $\mathbf{v} \neq 0$ and $f(\mathbf{v}) = \lambda\mathbf{v}$ for some $\lambda \in \mathbb{K}$.

Diagonalizability: Another Reformulation

Definition

Given a linear map $f : V \longrightarrow V$ we say $\mathbf{v} \in V$ is an eigenvector for f if $\mathbf{v} \neq 0$ and $f(\mathbf{v}) = \lambda\mathbf{v}$ for some $\lambda \in \mathbb{K}$.
In that case λ is called an eigenvalue of f. For a square matrix A we say λ is an eigenvalue if there exists a nonzero column vector $\mathbf{v}$ such that $A\mathbf{v} = \lambda\mathbf{v}$.

Diagonalizability: Another Reformulation

Definition

Given a linear map $f : V \longrightarrow V$ we say $\mathbf{v} \in V$ is an eigenvector for f if $\mathbf{v} \neq 0$ and $f(\mathbf{v}) = \lambda\mathbf{v}$ for some $\lambda \in \mathbb{K}$.
In that case λ is called an eigenvalue of f. For a square matrix A we say λ is an eigenvalue if there exists a nonzero column vector $\mathbf{v}$ such that $A\mathbf{v} = \lambda\mathbf{v}$. Of course $\mathbf{v}$ is then called the eigenvector of A corresponding to λ.

Proposition

A linear map $f : V \to V$ is diagonalizable iff V has a basis $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ consisting of eigenvectors of f.

Eigenvectors and Eigenvalues

Remark

(i) It is easy to see that eigenvalues and eigenvectors of a linear transformation are the same as those of the associated matrix.

Eigenvectors and Eigenvalues

Remark

(i) It is easy to see that eigenvalues and eigenvectors of a linear transformation are the same as those of the associated matrix.
(ii) Even if a linear map is not diagonalizable, the existence of eigenvectors and eigenvalues itself throws some light on the nature of the linear map.

Eigenvectors and Eigenvalues

Remark

(i) It is easy to see that eigenvalues and eigenvectors of a linear transformation are the same as those of the associated matrix.
(ii) Even if a linear map is not diagonalizable, the existence of eigenvectors and eigenvalues itself throws some light on the nature of the linear map. Thus the study of eigenvalues becomes extremely important.

Eigenvectors and Eigenvalues

Remark

(i) It is easy to see that eigenvalues and eigenvectors of a linear transformation are the same as those of the associated matrix.
(ii) Even if a linear map is not diagonalizable, the existence of eigenvectors and eigenvalues itself throws some light on the nature of the linear map. Thus the study of eigenvalues becomes extremely important. They arise naturally in the study of differential equations, extremum problem etc.

Eigenvectors and Eigenvalues

Remark

(i) It is easy to see that eigenvalues and eigenvectors of a linear transformation are the same as those of the associated matrix.
(ii) Even if a linear map is not diagonalizable, the existence of eigenvectors and eigenvalues itself throws some light on the nature of the linear map. Thus the study of eigenvalues becomes extremely important. They arise naturally in the study of differential equations, extremum problem etc. Here we shall use them to address the problem of diagonalization

Eigenvectors and Eigenvalues

Remark

(i) It is easy to see that eigenvalues and eigenvectors of a linear transformation are the same as those of the associated matrix.
(ii) Even if a linear map is not diagonalizable, the existence of eigenvectors and eigenvalues itself throws some light on the nature of the linear map. Thus the study of eigenvalues becomes extremely important. They arise naturally in the study of differential equations, extremum problem etc. Here we shall use them to address the problem of diagonalization and then see some geometric applications of diagonalization itself.