Inner Product and Orthogonality

CE282: Linear Algebra

Computer Engineering Department

Sharif University of Technology

Hamid R. Rabiee

Maryam Ramezani

Definition

Suppose V and W are vector spaces over the same field $\mathbb{F}$. Then a function $f: V \times W \rightarrow \mathbb{F}$ is called a **bilinear form** if it satisfies the following properties:

a) It is linear in its first argument:

i. $f(\mathbf{v_1} + \mathbf{v_2}, \mathbf{w}) = f(\mathbf{v_1}, \mathbf{w}) + f(\mathbf{v_2}, \mathbf{w})$ and

ii. $f(c\mathbf{v_1}, \mathbf{w}) = cf(\mathbf{v_1}, \mathbf{w})$ for all $c \in \mathbb{F}, \mathbf{v_1}, \mathbf{v_2} \in V$, and $\mathbf{w} \in W$.

b) It is linear in its second argument:

i. $f(\mathbf{v}, \mathbf{w_1} + \mathbf{w_2}) = f(\mathbf{v}, \mathbf{w_1}) + f(\mathbf{v}, \mathbf{w_2})$ and

ii. $f(\mathbf{v}, c\mathbf{w_1}) = cf(\mathbf{v}, \mathbf{w_1})$ for all $c \in \mathbb{F}, \mathbf{v} \in V$, and $\mathbf{w_1}, \mathbf{w_2} \in W$.

Note

Let V be a vector space over a field $\mathbb{F}$. Then the **dual** of V, denoted by V^*, is the vector space consisting of all linear forms on V.

Example

Let V be a vector space over a field $\mathbb{F}$. Show that the function $g: V^* \times V \to \mathbb{F}$ defined by

$$g(f, \mathbf{v}) = f(\mathbf{v}) \text{ for all } f \in V^*, \mathbf{v} \in V$$

is a bilinear form.

Review: Inner products over real field

- An inner product on V is a function $\langle , \rangle : V \times V \rightarrow \mathbb{R}$ such that

1. $\langle v, v\rangle \geq 0$ for all $v \in V$.
2. $\langle v, v\rangle = 0$ if and only if $v = 0$.
3. $\langle u + v, w\rangle = \langle u, w\rangle + \langle v, w\rangle$ for all u, $v, w \in V$.
4. $\langle cw, u\rangle = c\langle w, u\rangle$ for all $u, w \in V$ and $c \in \mathbb{R}$.
5. $\langle w, v\rangle = \langle v, w\rangle$.

General Inner products

Definition

Suppose that $\mathbb{F} = \mathbb{R}$ or $\mathbb{F} = \mathbb{C}$ and that V is a vector space over $\mathbb{F}$. Then an **inner product** on V is a function $\langle\cdot,\cdot\rangle : V \times V \rightarrow \mathbb{F}$ such that the following three properties hold for all $c \in \mathbb{F}$ and all $\mathbf{v}, \mathbf{w}, \mathbf{x} \in V$:

a) $\langle \mathbf{v}, \mathbf{w} \rangle = \overline{\langle \mathbf{w}, \mathbf{v} \rangle}$ (conjugate symmetry)

b) $\langle \mathbf{v} + c\mathbf{x}, \mathbf{w} \rangle = \langle \mathbf{v}, \mathbf{w} \rangle + c\langle \mathbf{x}, \mathbf{w} \rangle$ (linearity)

c) $\langle \mathbf{v}, \mathbf{v} \rangle \geq 0$, with equality if and only if $\mathbf{v} = \mathbf{0}$. (pos. definiteness)

General Inner products

Note

- $F = \mathbb{R}$ bilinear forms
- $F = \mathbb{C}$ sesquilinear forms—they are linear in their first argument, but only conjugate linear in their second argument

$$\langle v, w + cx \rangle = \overline{(\langle w + cx, v \rangle)} = \overline{(\langle w, v \rangle)} + \overline{(c\langle x, v \rangle\ )} = \langle v, w \rangle + \bar{c}\langle v, x \rangle$$

Complex Dot Product

Example

Show that the function $\langle \cdot,\cdot \rangle : \mathbb{C}^n \times \mathbb{C}^n \to \mathbb{C}$ defined by

$$\langle \mathbf{v}, \mathbf{w} \rangle = \mathbf{v}^* \mathbf{w} = \sum_{\boldsymbol{i=1}}^{\boldsymbol{n}} \overline{v}_i w_i \quad \text{for all} \quad \mathbf{v}, \mathbf{w} \in \mathbb{C}^n$$

is an inner product on $\mathbb{C}^n$.

Example

Let $a < b$ be real numbers and let $C[a, b]$ be the vector space of continuous functions on the real interval $[a, b]$.
Show that the function $\langle\cdot,\cdot\rangle : C[a, b] \times C[a, b] \to \mathbb{R}$ defined by

$$\langle f, g\rangle = \int_a^b f(x)g(x)\,dx \quad \text{for all } f, g \in C[a, b]$$

is and inner product on $C[a, b]$.

Example

Find $\langle p, q \rangle, \|p\|, \|p - q\|$ which $p(x) = 3 - x + 2x^2$ and $q(x) = 4x + x^2$.

Inner product and norm

Theorem

Take any inner product $\langle\cdot,\cdot\rangle$ and define $f(x) = \sqrt{\langle x, x\rangle}$. Then f is a norm.

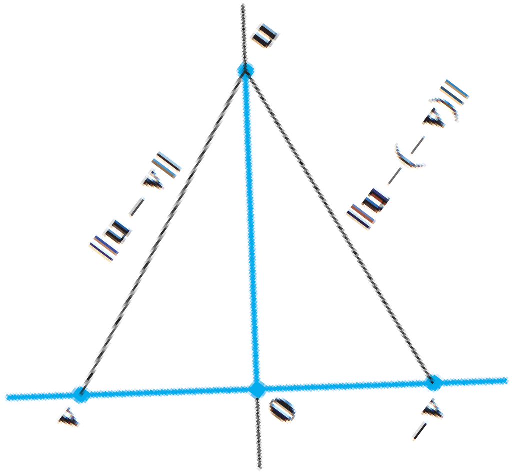
Proof

Note

Every inner product gives rise to a norm, but not every norm comes from an inner product. (Think about norm 2 and norm max)

Orthogonal vectors

- Geometry



- Algebra

Two vectors **u** and **v** in $\mathbb{R}^n$ are **orthogonal** (to each other) if $\mathbf{u} \cdot \mathbf{v} = \mathbf{0}$.

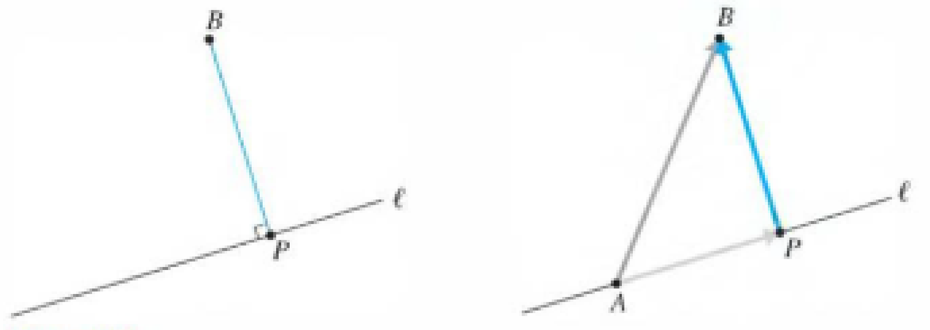
Suppose V is an inner product space.
Two vectors $\mathbf{v}, \mathbf{w} \in V$ are called **orthogonal** if $\langle \mathbf{v}, \mathbf{w} \rangle = 0$.

The Pythagorean Theorem

Two vectors **u** and **v** are orthogonal if and only if $\|\mathbf{u} + \mathbf{v}\|^2 = \|\mathbf{u}\|^2 + \|\mathbf{v}\|^2$

Projection

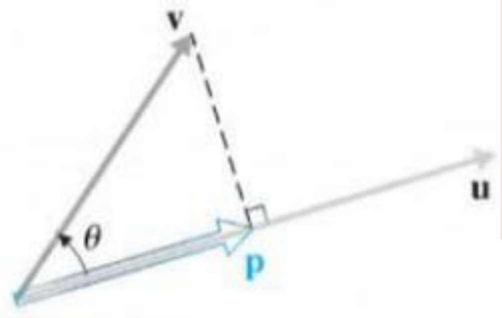
- Finding the distance from a point B to line l = Finding the length of line segment BP
- AP: projection of AB onto the line l



Definition

If $\mathbf{u}$ and $\mathbf{v}$ are vectors in $\mathbb{R}^n$ and $\mathbf{u} \neq \mathbf{0}$, then the **projection of v onto u** is the vector $proj_{\mathbf{u}}(\mathbf{v})$ defined by

$$proj_{\mathbf{u}}(\mathbf{v}) = \left(\frac{\mathbf{u}\cdot\mathbf{v}}{\mathbf{u}\cdot\mathbf{u}}\right)\mathbf{u}$$



The projection of **v** onto **u**

Orthogonal Sets

- A set of vectors $\{a_1, \dots, a_k\}$ in R^n is orthogonal set if each pair of distinct vectors is orthogonal (mutually orthogonal vectors).

Theorem

If $S = \{a_1, \dots, a_k\}$ is an orthogonal set of nonzero vectors in R^n, then S is linearly independent and is a basis for the subspace spanned by S.

Proof

If k = n, then prove that S is a basis for R^n

Orthonormal vectors

Definition

A basis B of an inner product space V is called an **orthonormal basis** of V if

a) $\langle \mathbf{v}, \mathbf{w} \rangle = 0$ for all $\mathbf{v} \neq \boldsymbol{w} \in B$, and (mutual orthogonality)

b) $\|\mathbf{v}\| = 1$ for all $\mathbf{v} \in B$. (normalization)

- set of n-vectors $a_1, \dots, a_k$ are *(mutually) orthogonal* if $a_i \perp a_j$ for $i \neq j$
- they are *normalized* if $\|a_i\| = 1$ for $i = 1, \dots, k$
- they are *orthonormal* if both hold
- can be expressed using inner products as

$$a_i^T a_j = \begin{cases} 1 & i = j \\ 0 & i \neq j \end{cases}$$

Orthonormal vectors

Independence-dimension inequality

If the n-vectors $a_1, \ldots, a_k$ are linearly independent, then $k \leq n$.

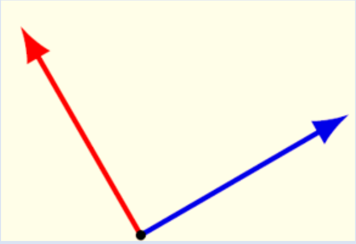
- orthonormal sets of vectors are linearly independent
- by independence-dimension inequality, must have $k \leq n$
- when $k = n, a_1, \ldots, a_n$ are an *orthonormal basis*

Example

- Standard unit n-vectors $e_1, \ldots, e_n$
- The 3-vectors

$$\begin{bmatrix} 0 \\ 0 \\ -1 \end{bmatrix}, \qquad \frac{1}{\sqrt{2}}\begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \qquad \frac{1}{\sqrt{2}}\begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}$$

- The 2-vectors shown below



- The standard basis in $P^p\ [a, b]$ (be the set of real-valued polynomials of degree at most p.)

Linear combinations of orthonormal vectors

- A simple way to check if an n-vector y is a linear combination of the orthonormal vectors $a_1, \dots, a_k$, if and only if:

$$y = (a_1^T y)a_1 + \ \dots + (a_k^T y)a_k$$

- For orthogonal vectors $a_1, \dots, a_k$:

$$y = c_1 a_1 + \cdots + c_k a_k$$

$$c_j = \frac{y.a_j}{a_j.a_j}$$

Linear combinations of orthonormal vectors

Example

Write x as a linear combination of a_1, a_2, a_3?

$$x = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}, \; a_1 = \begin{bmatrix} 0 \\ 0 \\ -1 \end{bmatrix}, a_2 = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, a_3 = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}$$

Orthogonal Complements

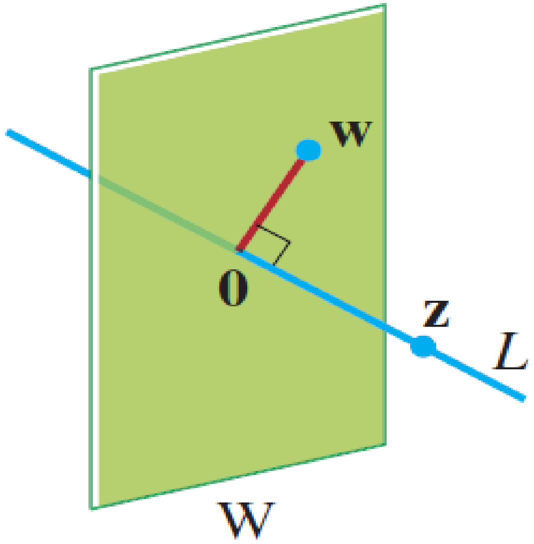
Definition

- If a vector z is orthogonal to every vector in a subspace W of R^n, then z is said to be orthogonal to W.
- **The set of all vectors z that are orthogonal** to W is called the orthogonal complement of W and is denoted by $w^{\perp}$

Example

W be a plane through the origin in $\mathbb{R}^3$.

$L = W^T$ and $W = L^T$



Orthogonal Complements

Theorem

1) A vector x is in W^T if and only if x is orthogonal to every vector in a set that spans W.
2) W^T is a subspace of $\mathbb{R}^n$.

Proof

Important

We emphasize that W_1 and W_2 can be orthogonal without being complements.

$W_1 = span((1,0,0))$ and $W_2 = span((0,1,0))$.

The Orthogonal Decomposition Theorem

Let W be a subspace of $\mathbb{R}^n$. Then each $\mathbf{y}$ in $\mathbb{R}^n$ can be written **uniquely** in the form:

$$\mathbf{y} = \hat{\mathbf{y}} + \mathbf{z} \tag{1}$$

$\text{proj}_W\, \mathbf{y}$

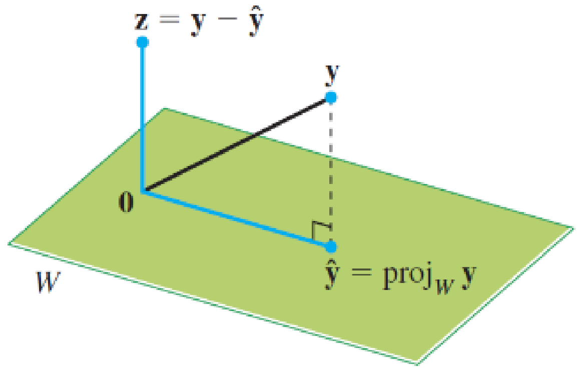
where $\hat{\mathbf{y}}$ is in W and $\mathbf{z}$ is in W^T. In fact, if $\{\mathbf{u}_1, \ldots, \mathbf{u}_p\}$ is any orthogonal basis of W, then

$$\hat{\mathbf{y}} = \frac{\mathbf{y}\cdot\mathbf{u}_1}{\mathbf{u}_1.\mathbf{u}_1}\mathbf{u}_1 + \cdots + \frac{\mathbf{y}\cdot\mathbf{u}_p}{\mathbf{u}_p\cdot\mathbf{u}_p}\mathbf{u}_p \tag{2}$$

and $z = \mathbf{y} - \hat{\mathbf{y}}$

Important

The uniqueness of the decomposition (1) shows that the orthogonal projection $\hat{\mathbf{y}}$ depends only on W and not on the particular basis used in (2).



The orthogonal projection of $\mathbf{y}$ onto W.

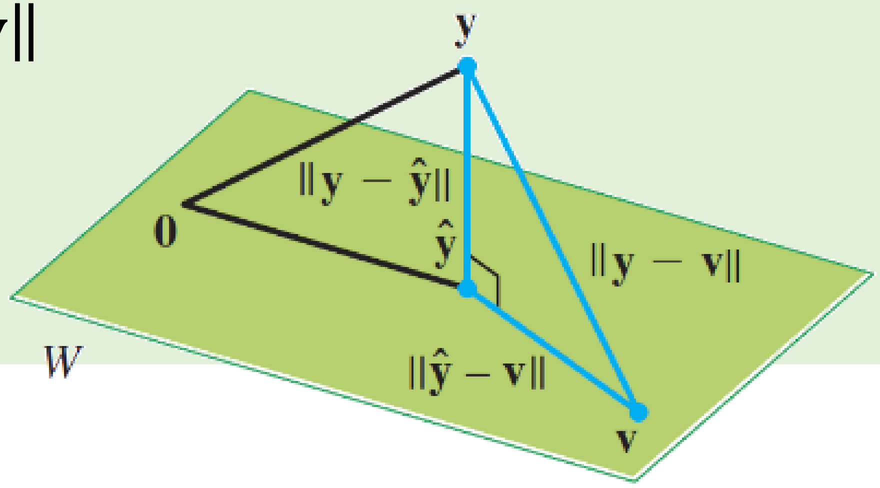
The Best Approximation Theorem

Let W be a subspace of $\mathbb{R}^n$. let $\mathbf{y}$ be any vector in $\mathbb{R}^n$, and let $\hat{\mathbf{y}}$ be the orthogonal projection of $\mathbf{y}$ onto W. Then $\hat{\mathbf{y}}$ is the closest point in W to $\mathbf{y}$, in the sense that

$$\|\mathbf{y} - \hat{\mathbf{y}}\| < \|\mathbf{y} - \mathbf{v}\|$$

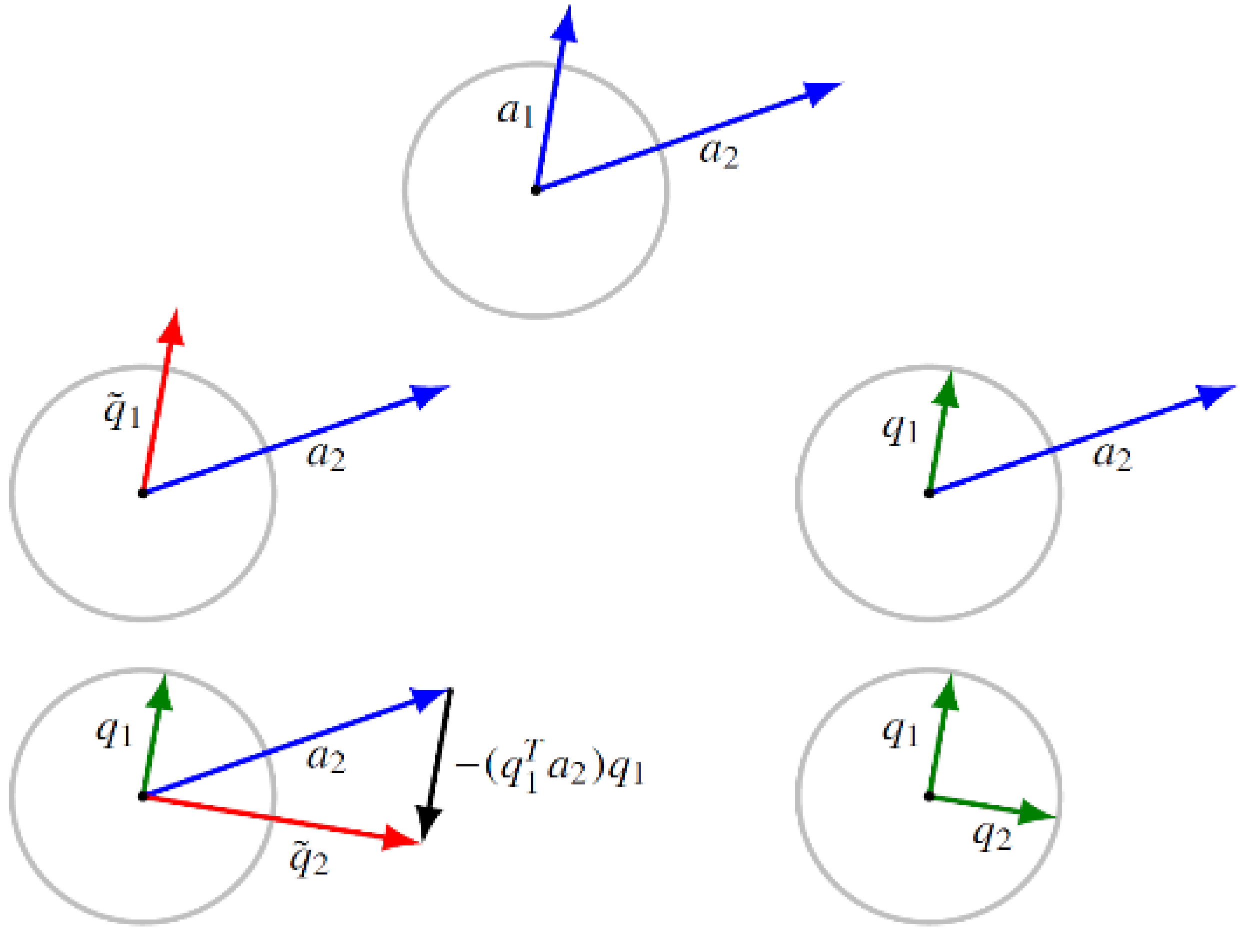
for all $\mathbf{v}$ in W distinct from $\hat{\mathbf{y}}$.

Proof



The orthogonal projection of $\mathbf{y}$ onto W is the closest point in W to $\mathbf{y}$.

- Find orthonormal basis for span$\{a_1, a_2, \dots, a_k\}$
- Geometry:



Gram–Schmidt (orthogonalization) algorithm

- Find orthonormal basis for span$\{a_1, a_2, \dots, a_k\}$
- Algebra:

$$1)\ q1 = \frac{a_1}{\|a_1\|}$$

$$2)\ \widetilde{q_2} = a_2 - (q_1^T a_2)q_1 \rightarrow q_2 = \frac{\widetilde{q_2}}{\|\widetilde{q_2}\|}$$

$$3)\ \widetilde{q_3} = a_3 - (q_1^T a_3)q_1 - (q_2^T a_3)q_2 \rightarrow q_3 = \frac{\widetilde{q_3}}{\|\widetilde{q_3}\|}$$

.

.

.

$$k)\ \widetilde{q_k} = a_k - (q_1^T a_k)q_1 - \cdots - (q_{k-1}^T a_k)q_{k-1} \rightarrow q_k = \frac{\widetilde{q_k}}{\|\widetilde{q_k}\|}$$

Gram–Schmidt (orthogonalization) algorithm

- Why $\{q_1, q_2, \dots, q_k\}$ is a orthonormal basis for span$\{a_1, a_2, \dots, a_k\}$?
  - $\{q_1, q_2, \dots, q_k\}$ are normalized.
  - $\{q_1, q_2, \dots, q_k\}$ is a orthogonal set
  - a_i is a linear combination of $\{q_1, q_2, \dots, q_i\}$

$$\Downarrow$$

$$\text{span}\{q_1, q_2, \dots, q_k\} = \text{span}\{a_1, a_2, \dots, a_k\}$$

- q_i is a linear combination of $\{a_1, a_2, \dots, a_i\}$

Gram–Schmidt (orthogonalization) algorithm

- given n-vectors $a_1, \dots, a_k$

  for $i = 1, \dots, k$

  1. Orthogonalization: $\tilde{q}_i = a_i - (q_1^T a_i)q_1 - \cdots - (q_{i-1}^T a_i)q_{i-1}$
  2. Test for linear dependence: if $\tilde{q}_i = 0$, quit
  3. Normalization: $q_i = \dfrac{\tilde{q}_i}{\|\tilde{q}_i\|}$

Note

- If G-S does not stop early (in step 2), $a_1, \dots, a_k$ are linearly independent.
- If G-S stops early in iteration $i = j$, then a_j is a linear combination of $a_1, \dots, a_{j-1}$ (so $a_1, \dots, a_k$ are linearly dependent)

$$a_j = (q_1^T a_j)q_1 + \cdots + \left(q_{j-1}^T a_j\right)q_{j-1}$$

Complexity of Gram–Schmidt algorithm

- given n-vectors $a_1, \dots, a_k$

  for $i = 1, \dots, k$

  1. Orthogonalization: $\tilde{q}_i = a_i - (q_1^T a_i)q_1 - \cdots - (q_{i-1}^T a_i)q_{i-1}$
  2. Test for linear dependence: if $\tilde{q}_i = 0$, quit
  3. Normalization: $q_i = \frac{\tilde{q}_i}{\|\tilde{q}_i\|}$

Theorem

Suppose B $= \{a_1, a_2, \dots, a_n\}$ is a basis of an inner product space A. Then C $= \{q_1, q_2, \dots, q_n\}$ is an orthonormal basis of $span\{a_1, a_2, \dots, a_n\}$.

$$q_1 = \frac{a_1}{||a_1||} \qquad \mathrm{q_k} = \frac{a_k - \sum_{i=1}^{k-1} < q_i, a_k > q_i}{\left\| a_k - \sum_{i=1}^{k-1} < q_i, a_k > q_i \right\|} \quad \text{for } 2 \le k \le n$$

Proof

We prove this result by induction on k.

Example

Find an orthonormal basis for $P^2[-1,1]$ with respect to the inner product

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

Conclusion

Existence of Orthonormal Bases

- Every finite-dimensional inner product space has an orthonormal basis.
- Since finite-dimensional inner product spaces (by definition) have a basis consisting of finitely many vectors, and the Gram-Schmidt process tells us how to convert that basis into an orthonormal basis, we now know that every finite-dimensional inner product space has an orthonormal basis.

Reference

- Chapter 1: Advanced Linear and Matrix Algebra, Nathaniel Johnston
- Chapter 6: Linear Algebra David Cherney
- Linear Algebra and Optimization for Machine Learning
- Introduction to Applied Linear Algebra Vectors, Matrices, and Least Squares