VIETNAM NATIONAL UNIVERSITY-HCMC
International University

Chapter 4. Inner product space

Applied Linear Algebra

Lecturer: Kha Kim Bao Han, MSc
kkbhan@hcmiu.edu.vn

Contents

Length and Dot Product

- Length:
  The length of a vector $v = (v_1, v_2, ..., v_n)$ in $\mathbb{R}^n$:
  $$||v|| = \sqrt{v_1^2 + v_2^2 + ... + v_n^2}$$
- Notes: The length of a vector is also called its norm.
- Properties of length (norm):
  $||v|| \geq 0$, $||v|| = 0$ if and only if $||v|| = 0$.
  $||v|| = 1 \;\; => v$ is called unit vector.
  $||cv|| = |c|||v||$

Normalizing a vector

Theorem

If v is a nonzero vector in $\mathbb{R}^n$, then the vector $u = \frac{v}{||v||}$ has length 1 and has the same direction as v. This vector u is called the unit vector in the direction of v.

Dot product

- Dot product in $\mathbb{R}^n$: The dot product of $u = (u_1, u_2, ..., u_n)$ and $v = (v_1, v_2, ..., v_n)$ returns a scalar quantity

$$u \cdot v = u_1 v_1 + u_2 v_2 + ... + u_n v_n$$

- Example: Dot product of $u = (1, 2, 0, -3)$ and $v = (3, -2, 4, 2)$ is

$$u \cdot v = 1.3 + 2(-2) + 0.4 + (-3)2 = -7$$

Properties of dot product

Properties of dot product in $\mathbb{R}^n$

- Linearity: $\forall \lambda, \mu \in \mathbb{R}^n$,
  $(\lambda u + \mu v) \cdot w = \lambda u \cdot w + \mu v \cdot w$,
- Symmetry:
  $u \cdot v = v \cdot u$
- Positive-definiteness: $\forall u \in \mathbb{R}^n$
  $u \cdot u \geq 0$, and $u \cdot u = 0$ if and only if $u = 0$.

Orthogonal vectors

- Orthogonal vectors:
  Two vectors u and v in $\mathbb{R}^n$ are orthogonal (perpendicular) if $u \cdot v = 0$
- Note: The vector 0 is said to be orthogonal to every vector
- Example: Determine all vectors in $\mathbb{R}^n$ that are orthogonal to $u = (4, 2)$
- Solution:

$$\mathbf{u} = (4, 2) \quad \text{Let} \quad \mathbf{v} = (v_1, v_2)$$

$$\begin{aligned} \Rightarrow \quad \mathbf{u} \cdot \mathbf{v} &= (4, 2) \cdot (v_1, v_2) \\ &= 4v_1 + 2v_2 \\ &= 0 \end{aligned}$$

$$\Rightarrow \quad v_1 = \frac{-t}{2}, \quad v_2 = t$$

$$\therefore \; \mathbf{v} = \left(\frac{-t}{2}, t \right), \; t \in R$$

Inner product

Definition

Let V be a real vector space. An inner product on V is a function that assigns to each ordered pair of vectors u, v in V real number (u, v) satisfying the following properties:

(a) $(u, u) \geq 0$, $(u, u) = 0$ if and only if $u = 0_V$. ?

(b) $(u, v) = (v, u)$, for any u, v in V.

(c) $(u + v, w) = (u, w) + (v, w)$ for any u, v, w in V.

(d) $(cu, v) = c(u, v)$ for u, v in V and c a real scalar.

$\langle u, v \rangle$

Inner product

Example: Standard inner product or Dot product in $\mathbb{R}^n$

We define the standard inner product, or dot product of each ordered pair of vectors as follow

$$(u, v) = u_1 v_1 + u_2 v_2 + \dots + u_n v_n,$$

where

$$u = \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{bmatrix}, v = \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix}$$

Remark: We can write the standard inner product of u and v in terms of matrix multiplication as

$$(u, v) = u^T v$$

Inner product

Example: Another inner product in $\mathbb{R}^2$

Let

$$u = \begin{bmatrix} u_1 \\ u_2 \end{bmatrix}, v = \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} \in \mathbb{R}^2$$

We define

$$(u, v) = u_1 v_1 - u_2 v_1 - u_1 v_2 + 3 u_2 v_2$$

It is easy to check that this is an inner product in $\mathbb{R}^2$.

Inner product

Example

Let V be the vector space of all continuous real-valued functions on the unit interval $[0, 1]$. For f and g in V, we let

$$(f, g) = \int_0^1 f(t)\, g(t)\, dt$$

This is an inner product in V.

Inner product

Theorem

Let $S = \{u_1, u_2, ..., u_n\}$ be an ordered basis for a finite-dimensional vector space V, and assume that we are given an inner product on V. Let $c_{ij} = (u_i, u_j)$ and $C = [c_{ij}]$. Then

(a) C is a symmetric matrix.

(b) C determines (v, w) for every v and w in V.

Remark: This theorem shows that every inner product on a finite-dimensional vector space V is completely determined, in terms of a given basis, by a certain matrix $C = [c_{ij}]$.

Inner product

Proof outline

(a) $c_{ij} = (u_i, u_j) = (u_j, u_i) = c_{ji}$.

(b) If v and w are in V, then

$$v = a_1 u_1 + a_2 u_2 + ... + a_n u_n, w = b_1 u_1 + b_2 u_2 + ... + b_n u_n$$

$$(v, w) = \sum_{i=1}^{n} \sum_{j=1}^{n} a_i c_{ij} b_j = [v]_S^T \, C \, [w]_S$$

Remark: The equation above showed that $x^T Cx = (x, x) > 0$ for every nonzero x in $\mathbb{R}^n$.

Inner product

Definition: Positive definite matrix

An $n \times n$ symmetric matrix C with the property that $x^T C x > 0$ for every nonzero vector x in $\mathbb{R}^n$ is called positive definite.

Remark: A positive definite matrix C is nonsingular.

Example: Positive definite matrix

The matrix

$$C = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}$$

is positive definite since

$$x^T C x = \begin{bmatrix} x_1 & x_2 \end{bmatrix} \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = x_1^2 + x_2^2 + (x_1 + x_2)^2 > 0$$

Euclidean space and Cauchy-Schwarz Inequality

Definition: Euclidean space

A real vector space that has an inner product defined on it is called an inner product space. If the space is finite dimensional it is called a Euclidean space.

Cauchy-Schwarz Inequality

If u and v are any two vectors in an inner product space V, then

$$|(u, v)| \leqslant \|u\| \, \|v\| \, ,$$

where $\|u\| = \sqrt{(u, u)}$ is the length (norm) of u.

We define the angle between u and v, the angle θ, such that

$$\cos \theta = \frac{(u, v)}{\|u\| \, \|v\|}.$$

Cauchy-Schwarz Inequality

Example: Cauchy-Schwarz Inequality

$$\left|\sum_{i=1}^{n} u_i v_i\right| \leqslant \left(\sum_{i=1}^{n} u_i^2\right)^{1/2} \left(\sum_{i=1}^{n} v_i^2\right)^{1/2}$$

Or,

$$\left|\int_0^1 f(t)\,g(t)\,dt\right| \leqslant \left(\int_0^1 (f(t))^2\,dt\right)^{1/2} \left(\int_0^1 (g(t))^2\,dt\right)^{1/2}$$

Triangle Inequality

If u and v are any vectors in an inner product space V, then $\|u+v\| \leq \|u\| + \|v\|$.

Orthogonal

Definition

Let V be an inner product space. Two vectors u and v in V are orthogonal if $(u, v) = 0$.

Example

Let V be the Euclidean space $\mathbb{R}^4$ with the standard inner product. If

$$u = \begin{bmatrix} -1 \\ 0 \\ 0 \\ 6 \end{bmatrix}, v = \begin{bmatrix} 0 \\ 3 \\ 5 \\ 0 \end{bmatrix}$$

then $(u, v) = 0$, so u and v are orthogonal.

Orthogonal

Definition: Orthogonal set

Let V be an inner product space. A set S of vectors in V is called orthogonal if any two distinct vectors in S are orthogonal. If, in addition, each vector in S is of unit length, then S is called orthonormal.

Example

Let

$$x_1 = \begin{bmatrix} 1 \\ 0 \\ 2 \end{bmatrix}, x_2 = \begin{bmatrix} -2 \\ 0 \\ 1 \end{bmatrix}, x_3 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$$

then $\{x_1, x_2, x_3\}$ is an orthogonal set since $(x_i, x_j) = 0$ for $i \neq j$.

Orthogonal

Remark on orthogonal set

Let x be a nonzero vector in an inner product space, we define

$$u = \frac{x}{\|x\|}$$

then u is a vector of unit length (called a unit vector) in the same direction as x.

Example

Let

$$u_1 = \begin{bmatrix} \frac{1}{\sqrt{5}} \\ 0 \\ \frac{2}{\sqrt{5}} \end{bmatrix}, u_2 = \begin{bmatrix} \frac{-2}{\sqrt{5}} \\ 0 \\ \frac{1}{\sqrt{5}} \end{bmatrix}, u_3 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$$

then $\{u_1, u_2, u_3\}$ is also an orthogonal set and $\|u_i\| = 1, \;\; i = 1, 2, 3$

Orthogonal

Remark on orthogonal set

The natural bases for $\mathbb{R}^n$ are orthonormal sets with respect to the standard inner products on these vector spaces.

Theorem on a finite orthogonal set

Let $S = \{u_1, u_2, ..., u_n\}$ be a finite orthogonal set of nonzero vectors in an inner product space V. Then S is linearly independent.

Proof:

Suppose that

$$a_1 u_1 + a_2 u_2 + ... + a_n u_n = 0$$

We want to show that $a_i = 0, \; i = 1, 2, .., n$.

Orthogonal

Proof (Cont.):

Taking the inner product of both sides with u_i, we have

$$(a_1u_1 + a_2u_2 + ... + a_nu_n, u_i) = 0$$

$$(a_1u_1, u_i) + (a_2u_2, u_i) + ... + (a_iu_i, u_i) + ... + (a_nu_n, u_i) = 0$$

This implies

$$a_1(u_1, u_i) + a_2(u_2, u_i) + ... + a_i(u_i, u_i) + ... + a_n(u_n, u_i) = 0$$

Since $(u_j, u_i) = 0$ if $j \neq i$, we obtain

$$a_i(u_i, u_i) = a_i\|u_i\|^2 = 0$$

Thus, $a_i = 0$.

Orthogonal set

Example

Let V be the vector space of all continuous real-valued functions on $[-\pi, \pi]$. For f and g in V, we let

$$(f, g) = \int_{-\pi}^{\pi} f(t)g(t)\,dt$$

which is an inner product on V.

Consider the following functions in V

$$S = \{1, \cos t, \sin t, \cos 2t, \sin 2t, ..., \cos nt, \sin nt, ...\}$$

Orthogonal set

Example

The relationships

$$\int_{-\pi}^{\pi} \cos nt\,dt = \int_{-\pi}^{\pi} \sin nt\,dt = \int_{-\pi}^{\pi} \sin nt \cos nt\,dt = 0$$

$$\int_{-\pi}^{\pi} \cos mt \cos nt\,dt = \int_{-\pi}^{\pi} \sin mt \sin nt\,dt = 0, \text{ if } m \neq n$$

demonstrate that $(f, g) = 0$ whenever f and g are distinct functions in S. Hence every finite subset of functions of S is an orthogonal set.

Gram-Schmidt Process

In this section, we prove that for every Euclidean space V we can obtain a basis S for V such that S is an orthonormal set!

Such a basis is called an orthonormal basis, and the method we use to obtain it is called the Gram-Schmidt process.

Gram-Schmidt Process

In this section, we prove that for every Euclidean space V we can obtain a basis S for V such that S is an orthonormal set!

Such a basis is called an orthonormal basis, and the method we use to obtain it is called the Gram-Schmidt process.

Theorem

Let $S = \{u_1, u_2, ..., u_n\}$ be an orthonormal basis for a Euclidean space V and let v be any vector in V. Then

$$v = c_1 u_1 + c_2 u_2 + ... + c_n u_n$$

where $c_i = (v, u_i)$.

Gram-Schmidt Process

Example

Let $S = \{u_1, u_2, u_3\}$ be an orthonormal basis for $\mathbb{R}^3$, where

$$u_1 = \begin{bmatrix} \frac{2}{3} \\ -\frac{2}{3} \\ \frac{1}{3} \end{bmatrix}, u_2 = \begin{bmatrix} \frac{2}{3} \\ \frac{1}{3} \\ -\frac{2}{3} \end{bmatrix}, u_3 = \begin{bmatrix} \frac{1}{3} \\ \frac{2}{3} \\ \frac{2}{3} \end{bmatrix}$$

Write the vector $v = \begin{bmatrix} 3 \\ 4 \\ 5 \end{bmatrix}$ as a linear combination of the vectors in S.

Gram-Schmidt Process

Example

Let $S = \{u_1, u_2, u_3\}$ be an orthonormal basis for $\mathbb{R}^3$, where

$$u_1 = \begin{bmatrix} \frac{2}{3} \\ -\frac{2}{3} \\ \frac{1}{3} \end{bmatrix}, u_2 = \begin{bmatrix} \frac{2}{3} \\ \frac{1}{3} \\ -\frac{2}{3} \end{bmatrix}, u_3 = \begin{bmatrix} \frac{1}{3} \\ \frac{2}{3} \\ \frac{2}{3} \end{bmatrix}$$

Write the vector $v = \begin{bmatrix} 3 \\ 4 \\ 5 \end{bmatrix}$ as a linear combination of the vectors in S.

Solution

We have $v = c_1 u_1 + c_2 u_2 + c_3 u_3$, where

$c_1 = (v, u_1) = 1, c_2 = (v, u_2) = 0, c_3 = (v, u_3) = 7$. Hence $v = u_1 + 7u_3$.

Q: How to find an orthonormal basis?

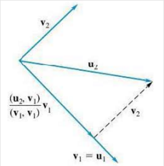
Gram-Schmidt Process

Theorem: Gram-Schmidt Process

Let V be an inner product space and $W \neq \{0\}$ an $m-$dimensional subspace of V.
Then there exists an orthonormal basis $T = \{w_1, w_2, ..., w_m\}$ for W.

Proof:
The proof is constructive. We first find an orthogonal basis $T^* = \{v_1, v_2, ..., v_m\}$ for W. Let $S = \{u_1, u_2, ..., u_m\}$ be any basis for W.

- Let $v_1 = u_1$
- Let $v_2 = u_2 - \frac{(u_2,v_1)}{(v_1,v_1)} v_1$. Since $(v_2, v_1) = (u_2, v_1) - \frac{(u_2,v_1)}{(v_1,v_1)} (v_1, v_1) = 0$, we thus have an orthogonal subset $\{v_1, v_2\}$ of W.



Gram-Schmidt Process

Proof (Cont.):

- Let $v_3 = u_3 - \frac{(u_3,v_1)}{(v_1,v_1)}v_1 - \frac{(u_3,v_2)}{(v_2,v_2)}v_2$. We have $(v_3, v_1) = (v_3, v_2) = 0$. We thus have an orthogonal subset $\{v_1, v_2, v_3\}$ of W.

- In general, we let $v_k = u_k - \sum_{i=1}^{k-1} \frac{(u_k,v_i)}{(v_i,v_i)}v_i$, for $k = 2, 3, ..., m$ to obtain an orthogonal basis $T^* = \{v_1, v_2, ..., v_m\}$ for W.

- Finally, let $w_i = \frac{v_i}{\|v_i\|}$ to get an orthonormal basis then $T = \{w_1, w_2, ..., w_m\}$ for W.

Gram-Schmidt Process

Remark: The projection operator is defined by

$$proj_u(v) = \frac{(v, u)}{(u, u)} u$$

So one can write v_k as

$$v_k = u_k - \sum_{i=1}^{k-1} proj_{v_i}(u_k)$$

Gram-Schmidt Process

Example: Gram-Schmidt Process

Let $S = \{u_1, u_2\}$ where

$$u_1 = \begin{bmatrix} 2 \\ -1 \end{bmatrix}, u_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

Transform S to an orthogonal basis T.

Gram-Schmidt Process

Example: Gram-Schmidt Process

Let $S = \{u_1, u_2\}$ where

$$u_1 = \begin{bmatrix} 2 \\ -1 \end{bmatrix}, u_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

Transform S to an orthogonal basis T.

Solution:

- By Gram-Schmidt Process, let $v_1 = u_1 = \begin{bmatrix} 2 \\ -1 \end{bmatrix}$
- Let

$$v_2 = u_2 - \frac{(u_2, v_1)}{(v_1, v_1)} v_1 = \begin{bmatrix} 0 \\ 1 \end{bmatrix} - \frac{-1}{5} \begin{bmatrix} 2 \\ -1 \end{bmatrix} = \begin{bmatrix} 2/5 \\ 4/5 \end{bmatrix}$$

So $T = \{v_1, v_2\}$ is an orthogonal process.

Gram-Schmidt Process

Example: Gram-Schmidt Process

Let W be the subspace of the Euclidean space $\mathbb{R}^4$ with the standard inner product with basis $S = \{u_1, u_2, u_3\}$ where

$$u_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 0 \end{bmatrix}, u_2 = \begin{bmatrix} -1 \\ 0 \\ -1 \\ 1 \end{bmatrix}, u_3 = \begin{bmatrix} -1 \\ 0 \\ 0 \\ -1 \end{bmatrix}$$

Transform S to an orthonormal basis $T = \{w_1, w_2, w_3\}$.

Gram-Schmidt Process

Example: Gram-Schmidt Process

Let W be the subspace of the Euclidean space $\mathbb{R}^4$ with the standard inner product with basis $S = \{u_1, u_2, u_3\}$ where

$$u_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 0 \end{bmatrix}, u_2 = \begin{bmatrix} -1 \\ 0 \\ -1 \\ 1 \end{bmatrix}, u_3 = \begin{bmatrix} -1 \\ 0 \\ 0 \\ -1 \end{bmatrix}$$

Transform S to an orthonormal basis $T = \{w_1, w_2, w_3\}$.

Solution:

- By Gram-Schmidt Process, let $v_1 = u_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 0 \end{bmatrix}$

Gram-Schmidt Process

Solution (Cont.):

- Let $v_2 = u_2 - \frac{(u_2,v_1)}{(v_1,v_1)} v_1 = \begin{bmatrix} -1 \\ 0 \\ -1 \\ 1 \end{bmatrix} - \left(\frac{-2}{3}\right) \begin{bmatrix} 1 \\ 1 \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} -1/3 \\ 2/3 \\ -1/3 \\ 1 \end{bmatrix}$
- Similarly, $v_3 = u_3 - \frac{(u_3,v_1)}{(v_1,v_1)} v_1 - \frac{(u_3,v_2)}{(v_2,v_2)} v_2 = \begin{bmatrix} -4/5 \\ 3/5 \\ 1/5 \\ -3/5 \end{bmatrix}$.

  Thus,

$$S = \left\{ \begin{bmatrix} 1 \\ 1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} -1 \\ 2 \\ -1 \\ 3 \end{bmatrix}, \begin{bmatrix} -4 \\ 3 \\ 1 \\ -3 \end{bmatrix} \right\}$$

  is an orthogonal basis.

Gram-Schmidt Process

Solution (Cont.): Hence

$$T = \left\{ \begin{bmatrix} 1/\sqrt{3} \\ 1/\sqrt{3} \\ 1/\sqrt{3} \\ 0 \end{bmatrix}, \begin{bmatrix} -1/\sqrt{15} \\ 2/\sqrt{15} \\ -1/\sqrt{15} \\ 3/\sqrt{15} \end{bmatrix}, \begin{bmatrix} -4/\sqrt{35} \\ 3/\sqrt{35} \\ 1/\sqrt{35} \\ -3/\sqrt{35} \end{bmatrix} \right\}$$

is an orthonormal basis for W.

Alternative form of the Gram Schmidt orthonormalization process:
$S = \{u_1, u_2, ..., u_n\}$ is a basis for an inner product space V
$T^* = \{v_1, v_2, ..., v_n\}$ is orthogonal basis for V
$T = \{w_1, w_2, ..., w_n\}$ is orthonormal basis for V

$w_1 = \frac{v_1}{||v_1||}$ where $v_1 = u_1$
$w_2 = \frac{v_2}{||v_2||}$ where $v_2 = u_2 - (u_2, w_1)w_1$
$w_3 = \frac{v_3}{||v_3||}$ where $v_3 = u_3 - (u_3, w_1)w_1 - (u_3, w_2)w_2$

.

.

.

$w_n = \frac{v_n}{||v_n||}$ where $v_n = u_n - \sum_{i=1}^{n-1}(u_n, w_i)w_i$

Orthogonal Complements

Orthogonal Complements of V

a) A vector v in V is said to be orthogonal to S, if v is orthogonal to every vector in S, i.e.,
$(v, w) = 0, \forall w \in S$.

b) The set of all vectors in V that are orthogonal to S is called the orthogonal complement of S
$S^{\perp} = \{v \in V | (v, w) = 0, \forall w \in S\}$

- Note:
  1)$(\{0\})^{\perp} = V$
  2)$V^{\perp} = \{0\}$

Note

Given S to be a subspace of V,
1) $S^{\perp}$ is a subspace of V
2) $S \cap S^{\perp} = \{0\}$

- Example:
  If $V = R^2, S =$x-axis. Then:
  1) $S^{\perp} =$y-axis is a subspace of R^2
  2) $S \cap S^{\perp} = \{(0,0)\}$

Direct sum

Definition

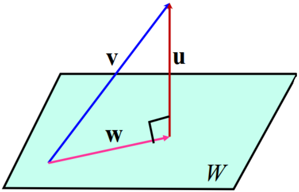
Let S_1 and S_2 be two subspaces of V. If each vector $x \in V$ can be uniquely written as a sum of a vector v_1 from S_1 and a vector v_2 from S_2, i.e., $x = v_1 + v_2$, then V is the direct sum of S_1 and S_2 , and we can write

$$V = S_1 \oplus S_2$$

Theorem

Let W be a finite-dimensional subspace of an inner product space V. Then $V = W \oplus W^{\perp}$ and $(W^{\perp})^{\perp} = W$.

Projections and Least Squares



- If W is a finite-dimensional subspace of an inner product space V, then $\forall v \in V, \exists w \in W, u \in W^{\perp} : v = w + u$
- w is called orthogonal projection of v on W, denoted by: $w = \text{proj}_W v$
- $\{w_1, w_2, ..., w_m\}$:orthonormal basis for W: $w = (v, w_1)w_1 + (v, w_2)w_2 + ... + (v, w_m)w_m$

$$w = \text{proj}_W v = \sum_{i=1}^{m} \frac{(v, w_i)}{(w_i, w_i)} w_i$$

Example 5: Projection onto a subspace

$w_1 = (0,3,1),\ w_2 = (2,0,0),\ v = (1,1,3)$

Find the projection of v onto the subspace $W = span(\{w_1, w_2\})$

Solution:

$\{\mathbf{w}_1, \mathbf{w}_2\}$: an orthogonal basis for W

$$\{\mathbf{u}_1, \mathbf{u}_2\} = \left\{\frac{\mathbf{w}_1}{\|\mathbf{w}_1\|}, \frac{\mathbf{w}_2}{\|\mathbf{w}_2\|}\right\} = \left\{(0, \frac{3}{\sqrt{10}}, \frac{1}{\sqrt{10}}), (1,0,0)\right\}$$: an orthonormal basis for W

$$\begin{aligned}\text{proj}_W \mathbf{v} &= \langle \mathbf{v},\ \mathbf{u}_1 \rangle \mathbf{u}_1 + \langle \mathbf{v},\ \mathbf{u}_2 \rangle \mathbf{u}_2 \\ &= \frac{6}{\sqrt{10}}(0, \frac{3}{\sqrt{10}}, \frac{1}{\sqrt{10}}) + (1,0,0) = (1, \frac{9}{5}, \frac{3}{3})\end{aligned}$$

Orthogonal projection and distance

Theorem

Let S be a subspace of an inner product space V, and $v \in V$. Then for all $u \in S, u \neq \text{proj}_S v$

$||v - \text{proj}_S v|| < ||v - u||$

or $||v - \text{proj}_S v|| = min||v - u||$

Proof:

$$\mathbf{v}-\mathbf{u}=(\mathbf{v}-\text{proj}_S\mathbf{v})+(\text{proj}_S\mathbf{v}-\mathbf{u})$$

$$\because \quad \text{proj}_S\mathbf{v}-\mathbf{u}\in S \text{ and } \mathbf{v}-\text{proj}_S\mathbf{v}\in S^{\perp} \Rightarrow \mathbf{v}-\text{proj}_S\mathbf{v} \perp \text{proj}_S\mathbf{v}-\mathbf{u}$$

$$\Rightarrow \langle \mathbf{v}-\text{proj}_S\mathbf{v}, \text{proj}_S\mathbf{v}-\mathbf{u}\rangle = 0$$

Thus the Pythagoream Theorem can be applied:

$$\|\mathbf{v}-\mathbf{u}\|^2 = \|\mathbf{v}-\text{proj}_S\mathbf{v}\|^2 + \|\text{proj}_S\mathbf{v}-\mathbf{u}\|^2.$$

Since $\mathbf{u}\neq \text{proj}_S\mathbf{v}$, the second term on the right hand side is positive,

and we can have $\|\mathbf{v}-\text{proj}_S\mathbf{v}\| < \|\mathbf{v}-\mathbf{u}\|$

Fundamental subspaces

Theorem

Fundamental subspaces of a matrix, including CS(A), CS(A^T), NS(A), and NS(A^T)

If A is an $m \times n$ matrix, then

(1) $CS(A) \perp NS(A^T)$ (or expressed as $CS(A)^{\perp} = NS(A^T)$)

Pf: Consider any $\mathbf{v} \in CS(A)$ and any $\mathbf{u} \in NS(A^T)$, and the goal is to prove $\mathbf{v} \cdot \mathbf{u} = 0$

$$\because \mathbf{u} \in NS(A^T) \quad \therefore A^T\mathbf{u} = \begin{bmatrix} (A^{(1)})^T \\ \vdots \\ (A^{(n)})^T \end{bmatrix} \mathbf{u} = \begin{bmatrix} (A^{(1)})^T\mathbf{u} \\ \vdots \\ (A^{(n)})^T\mathbf{u} \end{bmatrix} = \begin{bmatrix} 0 \\ \vdots \\ 0 \end{bmatrix} = \mathbf{0}$$

$$\Rightarrow \mathbf{v} \cdot \mathbf{u} = (c_1 A^{(1)} + \cdots + c_n A^{(n)}) \cdot \mathbf{u} = (c_1 A^{(1)} \cdot \mathbf{u} + \cdots + c_n A^{(n)} \cdot \mathbf{u}) = (0 + \cdots + 0) = 0$$

(Proved by setting $B = A^T$ and B satisfies the first property)

(2) $CS(A) \oplus NS(A^T) = R^m$ (because $CS(A) \oplus CS(A)^{\perp} = R^m$)

(3) $CS(A^T) \perp NS(A)$ (or expressed as $CS(A^T)^{\perp} = NS(A)$)

(4) $CS(A^T) \oplus NS(A) = R^n$ (because $CS(A^T) \oplus CS(A^T)^{\perp} = R^n$)

Example

- Find the four fundamental subspaces of the matrix

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

- Solution

$CS(A) = \text{span}\left(\{(1,0,0,0),\ (0,1,0,0)\}\right)$ is a subspace of R^4

$CS(A^T) = RS(A) = \text{span}\left(\{(1,2,0),\ (0,0,1)\}\right)$ is a subspace of R^3

$NS(A) = \text{span}(\{(-2,1,0)\})$ is a subspace of R^3

(The nullspace of A is a solution space of the homogeneous system $A\mathbf{x} = \mathbf{0}$, i.e., you need to solve $A\mathbf{x} = \mathbf{0}$ to derive (–2, 1, 0))

$$[A^T \mid \mathbf{0}] = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 2 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \end{bmatrix} \xrightarrow{\text{G.-J. E.}} \begin{array}{c} \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} \\ \begin{matrix} & & s & t & \end{matrix} \end{array}$$

$NS(A^T) = \text{span}\left(\{(0,0,1,0),\ (0,0,0,1)\}\right)$ is a subspace of R^4

- Example 2

$W = \text{span}(\{\mathbf{w}_1, \mathbf{w}_2\})$

Let W is a subspace of R^4 and $\mathbf{w}_1 = (1, 2, 1, 0)$, $\mathbf{w}_2 = (0, 0, 0, 1)$.

(a) Find a basis for W

(b) Find a basis for the orthogonal complement of W

- Solution

$$A = \begin{array}{cc} \begin{bmatrix} 1 & 0 \\ 2 & 0 \\ 1 & 0 \\ 0 & 1 \end{bmatrix} \\ \mathbf{w}_1 \quad \mathbf{w}_2 \end{array} \xrightarrow{\text{G.-J. E.}} \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \\ 0 & 0 \end{bmatrix} \quad \text{(reduced row-echelon form)}$$

(a) $W = CS(A)$, and since G.-J. E. will not affect the dependency among columns, we can conlude that $\{(1,2,1,0),(0,0,0,1)\}$ are linearly independent and could be a basis of W

(b) $W^{\perp} = CS(A)^{\perp} = NS(A^{T})$ (The nullspace of A^T is a solution space of the homogeneous system $A^T\mathbf{x} = \mathbf{0}$)

$$\because A^T = \begin{bmatrix} 1 & 2 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \qquad \therefore \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} -2s-t \\ s \\ t \\ 0 \end{bmatrix} = s\begin{bmatrix} -2 \\ 1 \\ 0 \\ 0 \end{bmatrix} + t\begin{bmatrix} -1 \\ 0 \\ 1 \\ 0 \end{bmatrix}$$

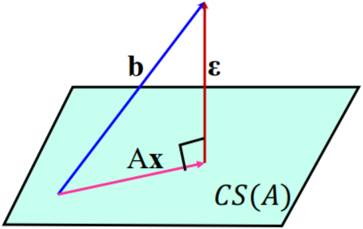
$\Rightarrow \{(-2,1,0,0)\ \ (-1,0,1,0)\}$ is a basis for $W^{\perp}$

Least squares

- Least squares problem:

$$\underset{m\times n}{A}\ \underset{n\times 1}{\mathbf{x}} = \underset{m\times 1}{\mathbf{b}}$$

Note: $A\mathbf{x} = \mathbf{b}$ is consistent if and only if $\mathbf{b} \in CS(A)$



(1) $\mathbf{b} \in CS(A)$, the system is consistent, we can use the Gaussian elimination to get exact solution $\mathbf{x}$

(2) $\mathbf{b} \notin CS(A)$, the system is inconsistent, only the "best possible" solution of the system can be found, i.e., to find a solution of $\mathbf{x}$ for which the error $D=||\boldsymbol{\varepsilon}||$ is minimum, where $\boldsymbol{\varepsilon} = \mathbf{b} - A\mathbf{x}$

$A \in M_{m\times n}$

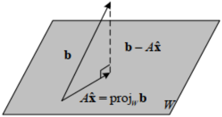
$\mathbf{x} \in R^n$

$A\mathbf{x} \in CS(A)$

$A\mathbf{x}$ can expressed as $x_1A^{(1)}+x_2A^{(2)}+\ldots+x_nA^{(n)}$

That is, find $\hat{x}_1A^{(1)}+\hat{x}_2A^{(2)}+\ldots+\hat{x}_nA^{(n)}$, which is closest to $\mathbf{b}$

Define $W = CS(A)$, and the problem to find $\hat{\mathbf{x}}$ such that $A\hat{\mathbf{x}}$ is closest to $\mathbf{b}$ is equivalent to find the vector in $CS(A)$ closest to $\mathbf{b}$, that is $\text{proj}_W\mathbf{b}$



Thus $A\hat{\mathbf{x}} = \text{proj}_W\mathbf{b}$ (To find the best solution $\hat{\mathbf{x}}$ which should satisfy this equation)

$\Rightarrow (\mathbf{b} - \text{proj}_W\mathbf{b}) = (\mathbf{b} - A\hat{\mathbf{x}}) \perp W \Rightarrow (\mathbf{b} - A\hat{\mathbf{x}}) \perp CS(A)$

$\Rightarrow \mathbf{b} - A\hat{\mathbf{x}} \in CS(A)^\perp = NS(A^T)$ (The nullspace of A^T is a solution space of the homogeneous system $A^T\mathbf{x}=\mathbf{0}$)

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

$\Rightarrow A^TA\hat{\mathbf{x}} = A^T\mathbf{b}$ (the $n\times n$ linear system of normal equations associated with $A\mathbf{x} = \mathbf{b}$)

Theorem

A: m x n matrix, B: n x m: matrix

a) $Ax \cdot y = x \cdot A^T y$

b) $x \cdot By = B^T x \cdot y$

Proof

$$\mathbf{u} = \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{bmatrix}, \quad \mathbf{v} = \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix} \qquad \mathbf{u} \cdot \mathbf{v} = \mathbf{u}^T \mathbf{v} = u_1 v_1 + u_2 v_2 + \cdots + u_n v_n$$

(a) $Ax \bullet y = (Ax)^T y = (x^T A^T) y = x^T (A^T y) = x \bullet A^T y$

(b) $A = B^T$

Theorem

A: m x n matrix

a) $NS(A^TA) = NS(A)$

b) $rank(A^TA) = rank(A)$

Proof (a) $u \in NS(A) \Leftrightarrow Au = 0 \Rightarrow A^T Au = A^T 0 = 0$

$\Rightarrow u \in NS(A^T A)$

$v \in NS(A^T A) \Rightarrow A^T Av = 0 \quad \Rightarrow A^T Av \bullet v = 0 \bullet v = 0$

$Av \bullet Av = 0 \Rightarrow Av = 0 \quad \Rightarrow v \in NS(A)$

(b) $rk(A) + \dim(NS(A)) = n = rk(A^T A) + \dim(NS(A^T A))$

By (a), we get the result

Corollary

A : mxn matrix, $m \geq n$, and $rank(A) = n$, then A^TA is invertible.

Normal Equation

- Theorem 4.16: A least squares solution to the sytem $A\mathbf{x}=\mathbf{b}$ is an exact solution of the **normal equation**

$$\boxed{A^T A\mathbf{x}' = A^T\mathbf{b}}$$

If the columns of A are linearly independent, then $A^T A$ is invertible, so the above equation has a unique solution

$$\boxed{\mathbf{x}' = (A^T A)^{-1} A^T\mathbf{b}}$$

- Example: Solving the normal equations

Find the least squares solution of the following system

$$A\mathbf{x} = \mathbf{b}$$

$$\begin{bmatrix} 1 & 1 \\ 1 & 2 \\ 1 & 3 \end{bmatrix} \begin{bmatrix} c_0 \\ c_1 \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \\ 3 \end{bmatrix}$$

and find the orthogonal projection of $\mathbf{b}$ onto the column space of A

- Solution:

$$A^T A = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & 3 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 1 & 2 \\ 1 & 3 \end{bmatrix} = \begin{bmatrix} 3 & 6 \\ 6 & 14 \end{bmatrix}$$

$$A^T \mathbf{b} = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & 3 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \\ 3 \end{bmatrix} = \begin{bmatrix} 4 \\ 11 \end{bmatrix}$$

the corresponding normal system

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

$$\begin{bmatrix} 3 & 6 \\ 6 & 14 \end{bmatrix} \begin{bmatrix} \hat{c}_0 \\ \hat{c}_1 \end{bmatrix} = \begin{bmatrix} 4 \\ 11 \end{bmatrix}$$

the least squares solution of $A\mathbf{x} = \mathbf{b}$

$$\hat{\mathbf{x}} = \begin{bmatrix} \hat{c}_0 \\ \hat{c}_1 \end{bmatrix} = \begin{bmatrix} -\frac{5}{3} \\ \frac{3}{2} \end{bmatrix}$$

the orthogonal projection of $\mathbf{b}$ onto the column space of A

$$\operatorname{proj}_{CS(A)}\mathbf{b} = A\hat{\mathbf{x}} = \begin{bmatrix} 1 & 1 \\ 1 & 2 \\ 1 & 3 \end{bmatrix} \begin{bmatrix} \frac{-5}{3} \\ \frac{3}{2} \end{bmatrix} = \begin{bmatrix} \frac{-1}{6} \\ \frac{8}{6} \\ \frac{17}{6} \end{bmatrix}$$

※ Find an orthogonal basis for $CS(A)$ by performing the Gram-Schmidt process, and then calculate $\operatorname{proj}_{CS(A)}\mathbf{b}$ directly, you will derive the same result

Exercises

1. Find perpendicular projection of v into the subspace spanned by the given vectors c_1, c_2

$$v = (-2, 1, 0),\ c_1 = (2, 1, 1), \qquad c_2 = (3, 2, -1)$$

2. Find the least squares solution of the following system Ax=b, where

$$A = \begin{bmatrix} -3 & 2 \\ 2 & -2 \\ 4 & 5 \end{bmatrix}, \quad b = \begin{bmatrix} 4 \\ -2 \\ 3 \end{bmatrix}$$