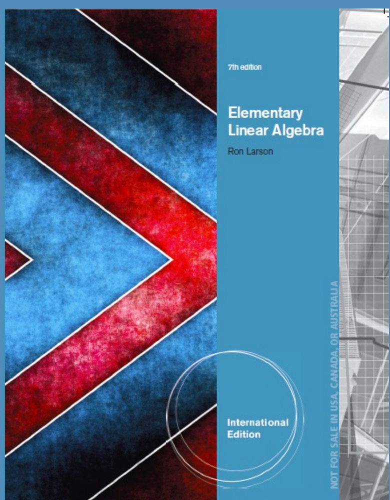
CHAPTER 5
INNER PRODUCT SPACES



5.3 Orthonormal Bases: Gram-Schmidt Process

- Orthogonal:

  A set S of vectors in an inner product space V is called an **orthogonal set** if every pair of vectors in the set is orthogonal.

$$S = \{\mathbf{v}_1, \mathbf{v}_2, \cdots, \mathbf{v}_n\} \subseteq V$$

$$\langle \mathbf{v}_i, \mathbf{v}_j \rangle = 0 \qquad i \neq j$$

- Orthonormal:

  An orthogonal set in which each vector is a unit vector is called **orthonormal**.

$$S = \{\mathbf{v}_1, \mathbf{v}_2, \cdots, \mathbf{v}_n\} \subseteq V$$

$$\langle \mathbf{v}_i, \mathbf{v}_j \rangle = \begin{cases} 1 & i = j \\ 0 & i \neq j \end{cases}$$

- Note:

  If S is a basis, then it is called an orthogonal basis or an orthonormal basis.

- Ex 1: (A nonstandard orthonormal basis for R^3)

Show that the following set is an orthonormal basis.

$$S = \left\{ \overset{\mathbf{v}_1}{\left(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}, 0\right)}, \overset{\mathbf{v}_2}{\left(-\frac{\sqrt{2}}{6}, \frac{\sqrt{2}}{6}, \frac{2\sqrt{2}}{3}\right)}, \overset{\mathbf{v}_3}{\left(\frac{2}{3}, -\frac{2}{3}, \frac{1}{3}\right)} \right\}$$

Sol:

Show that the three vectors are mutually orthogonal.

$$\mathbf{v}_1 \cdot \mathbf{v}_2 = -\tfrac{1}{6} + \tfrac{1}{6} + 0 = 0$$

$$\mathbf{v}_1 \cdot \mathbf{v}_3 = \frac{2}{3\sqrt{2}} - \frac{2}{3\sqrt{2}} + 0 = 0$$

$$\mathbf{v}_2 \cdot \mathbf{v}_3 = -\frac{\sqrt{2}}{9} - \frac{\sqrt{2}}{9} + \frac{2\sqrt{2}}{9} = 0$$

Show that each vector is of length 1.

$$\|\mathbf{v}_1\| = \sqrt{\mathbf{v}_1 \cdot \mathbf{v}_1} = \sqrt{\tfrac{1}{2} + \tfrac{1}{2} + 0} = 1$$

$$\|\mathbf{v}_2\| = \sqrt{\mathbf{v}_2 \cdot \mathbf{v}_2} = \sqrt{\tfrac{2}{36} + \tfrac{2}{36} + \tfrac{8}{9}} = 1$$

$$\|\mathbf{v}_3\| = \sqrt{\mathbf{v}_3 \cdot \mathbf{v}_3} = \sqrt{\tfrac{4}{9} + \tfrac{4}{9} + \tfrac{1}{9}} = 1$$

Thus S is an orthonormal set.

- Ex 2: (An orthonormal basis for $P_3(x)$)

In $P_3(x)$, with the inner product

$$\langle p, q\rangle = a_0b_0 + a_1b_1 + a_2b_2$$

The standard basis $B = \{1, x, x^2\}$ is orthonormal.

Sol:

$$\mathbf{v}_1 = 1 + 0x + 0x^2, \quad \mathbf{v}_2 = 0 + x + 0x^2, \quad \mathbf{v}_3 = 0 + 0x + x^2,$$

Then

$$\langle \mathbf{v}_1, \mathbf{v}_2\rangle = (1)(0) + (0)(1) + (0)(0) = 0,$$
$$\langle \mathbf{v}_1, \mathbf{v}_3\rangle = (1)(0) + (0)(0) + (0)(1) = 0,$$
$$\langle \mathbf{v}_2, \mathbf{v}_3\rangle = (0)(0) + (1)(0) + (0)(1) = 0$$

$$\|\mathbf{v}_1\| = \sqrt{\langle \mathbf{v}_1, \mathbf{v}_1 \rangle} = \sqrt{(1)(1) + (0)(0) + (0)(0)} = 1,$$

$$\|\mathbf{v}_2\| = \sqrt{\langle \mathbf{v}_2, \mathbf{v}_2 \rangle} = \sqrt{(0)(0) + (1)(1) + (0)(0)} = 1,$$

$$\|\mathbf{v}_3\| = \sqrt{\langle \mathbf{v}_3, \mathbf{v}_3 \rangle} = \sqrt{(0)(0) + (0)(0) + (1)(1)} = 1$$

- Thm 5.10: (Orthogonal sets are linearly independent)

  If $S=\{\mathbf{v}_1,\mathbf{v}_2,\cdots,\mathbf{v}_n\}$ is an orthogonal set of nonzero vectors in an inner product space V, then S is linearly independent.

Pf:

S is an orthogonal set of nonzero vectors

i.e. $\langle \mathbf{v}_i,\mathbf{v}_j\rangle=0 \quad i\neq j$ and $\langle \mathbf{v}_i,\mathbf{v}_i\rangle>0$

Let $\quad c_1\mathbf{v}_1+c_2\mathbf{v}_2+\cdots+c_n\mathbf{v}_n=0$

$\Rightarrow \quad \langle c_1\mathbf{v}_1+c_2\mathbf{v}_2+\cdots+c_n\mathbf{v}_n,\mathbf{v}_i\rangle=\langle 0,\mathbf{v}_i\rangle=0 \qquad \forall i$

$\Rightarrow \quad c_1\langle \mathbf{v}_1,\mathbf{v}_i\rangle+c_2\langle \mathbf{v}_2,\mathbf{v}_i\rangle+\cdots+c_i\langle \mathbf{v}_i,\mathbf{v}_i\rangle+\cdots+c_n\langle \mathbf{v}_n,\mathbf{v}_i\rangle$

$=c_i\langle \mathbf{v}_i,\mathbf{v}_i\rangle$

$\because \langle \mathbf{v}_i,\mathbf{v}_i\rangle\neq 0 \quad \Rightarrow c_i=0 \quad \forall i \qquad \therefore S$ is linearly independent.

推論

- Corollary to Thm 5.10:

  If V is an inner product space of dimension n, then any orthogonal set of n nonzero vectors is a basis for V.

- Ex 4: (Using orthogonality to test for a basis)

Show that the following set is a basis for R^4.

$$S=\{\overset{\mathbf{v}_1}{(2,3,2,-2)},\overset{\mathbf{v}_2}{(1,0,0,1)},\overset{\mathbf{v}_3}{(-1,0,2,1)},\overset{\mathbf{v}_4}{(-1,2,-1,1)}\}$$

Sol:

$\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3, \mathbf{v}_4$: nonzero vectors

$$\mathbf{v}_1 \cdot \mathbf{v}_2 = 2+0+0-2=0 \qquad \mathbf{v}_2 \cdot \mathbf{v}_3 = -1+0+0+1=0$$

$$\mathbf{v}_1 \cdot \mathbf{v}_3 = -2+0+4-2=0 \qquad \mathbf{v}_2 \cdot \mathbf{v}_4 = -1+0+0+1=0$$

$$\mathbf{v}_1 \cdot \mathbf{v}_4 = -2+6-2-2=0 \qquad \mathbf{v}_3 \cdot \mathbf{v}_4 = 1+0-2+1=0$$

$\Rightarrow S$ is orthogonal.

$\Rightarrow S$ is a basis for R^4 (by Corollary to Theorem 5.10)

- Thm 5.11: (Coordinates relative to an orthonormal basis)

  If $B=\{\mathbf{v}_1, \mathbf{v}_2, \cdots, \mathbf{v}_n\}$ is an orthonormal basis for an inner product space V, then the coordinate representation of a vector $\mathbf{w}$ with respect to B is

$$\mathbf{w}=\langle \mathbf{w}, \mathbf{v}_1\rangle \mathbf{v}_1+\langle \mathbf{w}, \mathbf{v}_2\rangle \mathbf{v}_2+\cdots+\langle \mathbf{w}, \mathbf{v}_n\rangle \mathbf{v}_n$$

Pf:

$B=\{\mathbf{v}_1, \mathbf{v}_2, \cdots, \mathbf{v}_n\}$ is a basis for V

$\mathbf{w}\in V$

$\mathbf{w}=k_1\mathbf{v}_1+k_2\mathbf{v}_2+\cdots+k_n\mathbf{v}_n$ (unique representation)

$\because B=\{\mathbf{v}_1, \mathbf{v}_2, \cdots, \mathbf{v}_n\}$ is orthonormal

$$\Rightarrow \langle \mathbf{v}_i, \mathbf{v}_j\rangle=\begin{cases}1 & i=j\\ 0 & i\neq j\end{cases}$$

$$\begin{aligned}\langle \mathbf{w}, \mathbf{v}_i\rangle &= \langle (k_1\mathbf{v}_1 + k_2\mathbf{v}_2 + \cdots + k_n\mathbf{v}_n), \mathbf{v}_i\rangle \\ &= k_1\langle \mathbf{v}_1, \mathbf{v}_i\rangle + \cdots + k_i\langle \mathbf{v}_i, \mathbf{v}_i\rangle + \cdots + k_n\langle \mathbf{v}_n, \mathbf{v}_i\rangle \\ &= k_i \qquad \forall i\end{aligned}$$

$$\Rightarrow \mathbf{w} = \langle \mathbf{w}, \mathbf{v}_1\rangle \mathbf{v}_1 + \langle \mathbf{w}, \mathbf{v}_2\rangle \mathbf{v}_2 + \cdots + \langle \mathbf{w}, \mathbf{v}_n\rangle \mathbf{v}_n$$

- Note:

If $B = \{\mathbf{v}_1, \mathbf{v}_2, \cdots, \mathbf{v}_n\}$ is an orthonormal basis for V and $\mathbf{w} \in V$,

Then the corresponding coordinate matrix of $\mathbf{w}$ relative to B is

$$[\mathbf{w}]_B = \begin{bmatrix} \langle \mathbf{w}, \mathbf{v}_1\rangle \\ \langle \mathbf{w}, \mathbf{v}_2\rangle \\ \vdots \\ \langle \mathbf{w}, \mathbf{v}_n\rangle \end{bmatrix}$$

- Ex 5: (Representing vectors relative to an orthonormal basis)

  Find the coordinates of $\mathbf{w} = (5, -5, 2)$ relative to the following orthonormal basis for R^3.

$$B = \{(\tfrac{3}{5}, \tfrac{4}{5}, 0), (-\tfrac{4}{5}, \tfrac{3}{5}, 0), (0, 0, 1)\}$$

Sol:

$$\langle \mathbf{w}, \mathbf{v}_1 \rangle = \mathbf{w} \cdot \mathbf{v}_1 = (5, -5, 2) \cdot (\tfrac{3}{5}, \tfrac{4}{5}, 0) = -1$$

$$\langle \mathbf{w}, \mathbf{v}_2 \rangle = \mathbf{w} \cdot \mathbf{v}_2 = (5, -5, 2) \cdot (-\tfrac{4}{5}, \tfrac{3}{5}, 0) = -7$$

$$\langle \mathbf{w}, \mathbf{v}_3 \rangle = \mathbf{w} \cdot \mathbf{v}_3 = (5, -5, 2) \cdot (0, 0, 1) = 2$$

$$\Rightarrow [\mathbf{w}]_B = \begin{bmatrix} -1 \\ -7 \\ 2 \end{bmatrix}$$

- Gram-Schmidt orthonormalization process:

$B=\{\mathbf{u}_1, \mathbf{u}_2, \cdots, \mathbf{u}_n\}$ is a basis for an inner product space V

Let $\mathbf{v}_1=\mathbf{u}_1$

$\mathbf{w}_1 = span(\{\mathbf{v}_1\})$

$$\mathbf{v}_2=\mathbf{u}_2-\text{proj}_{\mathbf{W}_1}\mathbf{u}_2=\mathbf{u}_2-\frac{\langle\mathbf{u}_2, \mathbf{v}_1\rangle}{\langle\mathbf{v}_1, \mathbf{v}_1\rangle}\mathbf{v}_1$$

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

$$\mathbf{v}_3=\mathbf{u}_3-\text{proj}_{\mathbf{W}_2}\mathbf{u}_3=\mathbf{u}_3-\frac{\langle\mathbf{u}_3, \mathbf{v}_1\rangle}{\langle\mathbf{v}_1, \mathbf{v}_1\rangle}\mathbf{v}_1-\frac{\langle\mathbf{u}_3, \mathbf{v}_2\rangle}{\langle\mathbf{v}_2, \mathbf{v}_2\rangle}\mathbf{v}_2$$

$$\vdots$$

$$\mathbf{v}_n=\mathbf{u}_n-\text{proj}_{\mathbf{W}_{n-1}}\mathbf{u}_n=\mathbf{u}_n-\sum_{i=1}^{n-1}\frac{\langle\mathbf{v}_n, \mathbf{v}_i\rangle}{\langle\mathbf{v}_i, \mathbf{v}_i\rangle}\mathbf{v}_i$$

$\Rightarrow B'=\{\mathbf{v}_1, \mathbf{v}_2, \cdots, \mathbf{v}_n\}$ is an orthogonal basis.

$\Rightarrow B''=\left\{\frac{\mathbf{v}_1}{\|\mathbf{v}_1\|}, \frac{\mathbf{v}_2}{\|\mathbf{v}_2\|}, \cdots, \frac{\mathbf{v}_n}{\|\mathbf{v}_n\|}\right\}$ is an orthonormal basis.

- Ex 7: (Applying the Gram-Schmidt orthonormalization process)

Apply the Gram-Schmidt process to the following basis.

$$B = \{\underset{\mathbf{u}_1}{(1,1,0)},\ \underset{\mathbf{u}_2}{(1,2,0)},\ \underset{\mathbf{u}_3}{(0,1,2)}\}$$

Sol: $\mathbf{v}_1 = \mathbf{u}_1 = (1,1,0)$

$$\mathbf{v}_2 = \mathbf{u}_2 - \frac{\mathbf{u}_2 \cdot \mathbf{v}_1}{\mathbf{v}_1 \cdot \mathbf{v}_1}\mathbf{v}_1 = (1,2,0) - \frac{3}{2}(1,1,0) = (-\frac{1}{2}, \frac{1}{2}, 0)$$

$$\mathbf{v}_3 = \mathbf{u}_3 - \frac{\mathbf{u}_3 \cdot \mathbf{v}_1}{\mathbf{v}_1 \cdot \mathbf{v}_1}\mathbf{v}_1 - \frac{\mathbf{u}_3 \cdot \mathbf{v}_2}{\mathbf{v}_2 \cdot \mathbf{v}_2}\mathbf{v}_2$$

$$= (0,1,2) - \frac{1}{2}(1,1,0) - \frac{1/2}{1/2}(-\frac{1}{2}, \frac{1}{2}, 0) = (0,0,2)$$

Orthogonal basis

$$\Rightarrow B' = \{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\} = \{(1, 1, 0), (\frac{-1}{2}, \frac{1}{2}, 0), (0, 0, 2)\}$$

Orthonormal basis

$$\Rightarrow B'' = \{\frac{\mathbf{v}_1}{\|\mathbf{v}_1\|}, \frac{\mathbf{v}_2}{\|\mathbf{v}_2\|}, \frac{\mathbf{v}_3}{\|\mathbf{v}_3\|}\} = \{(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}, 0), (\frac{-1}{\sqrt{2}}, \frac{1}{\sqrt{2}}, 0), (0, 0, 1)\}$$

- Ex 10: (Alternative form of Gram-Schmidt orthonormalization process)

  Find an orthonormal basis for the solution space of the homogeneous system of linear equations.

$$\begin{aligned} x_1 + x_2 \qquad\quad + 7x_4 &= 0 \\ 2x_1 + x_2 + 2x_3 + 6x_4 &= 0 \end{aligned}$$

Sol:

$$\begin{bmatrix} 1 & 1 & 0 & 7 & 0 \\ 2 & 1 & 2 & 6 & 0 \end{bmatrix} \xrightarrow{G.J.E} \begin{bmatrix} 1 & 0 & 2 & -1 & 0 \\ 0 & 1 & -2 & 8 & 0 \end{bmatrix}$$

$$\Rightarrow \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} -2s+t \\ 2s-8t \\ s \\ t \end{bmatrix} = s\begin{bmatrix} -2 \\ 2 \\ 1 \\ 0 \end{bmatrix} + t\begin{bmatrix} 1 \\ -8 \\ 0 \\ 1 \end{bmatrix}$$

Thus one basis for the solution space is

$$B=\{\boldsymbol{u}_1, \boldsymbol{u}_2\}=\{(-2, 2, 1, 0), (1, -8, 0, 1)\}$$

$$\boldsymbol{v}_1=\boldsymbol{u}_1=(-2, 2, 1, 0)$$

$$\boldsymbol{v}_2=\boldsymbol{u}_2-\frac{\langle \boldsymbol{u}_2, \boldsymbol{v}_1\rangle}{\langle \boldsymbol{v}_1, \boldsymbol{v}_1\rangle}\boldsymbol{v}_1=(1, -8, 0, 1)-\frac{-18}{9}(-2, 2, 1, 0)$$

$$=(-3, -4, 2, 1)$$

$$\Rightarrow B'=\{(-2,2,1,0)(-3,-4,2,1)\}$$ (orthogonal basis)

$$\Rightarrow B''=\left\{\left(\frac{-2}{3},\frac{2}{3},\frac{1}{3},0\right),\left(\frac{-3}{\sqrt{30}},\frac{-4}{\sqrt{30}},\frac{2}{\sqrt{30}},\frac{1}{\sqrt{30}}\right)\right\}$$

(orthonormal basis)

Key Learning in Section 5.3

- Show that a set of vectors is orthogonal and forms an orthonormal basis, and represent a vector relative to an orthonormal basis.
- Apply the Gram-Schmidt orthonormalization process.

Keywords in Section 5.3

- orthogonal set: 正交集合
- orthonormal set: 單範正交集合
- orthogonal basis: 正交基底
- orthonormal basis: 單範正交基底
- linear independent: 線性獨立
- Gram-Schmidt Process: Gram-Schmidt過程

Review exercise

39. Apply the Gram-Schmidt orthonormalization process to transform the given basis for R^n into an orthonormal basis. Use the Euclidean inner product for R^n and use the vectors in the order in which they are given.

$$B = \{(0,3,4), (1,0,0), (1,1,0)\}$$

- Orthogonal subspaces:

The subspaces W_1 and W_2 of an inner product space V are orthogonal if $\langle \boldsymbol{v}_1, \boldsymbol{v}_2 \rangle = 0$ for all $\boldsymbol{v}_1$ in W_1 and all $\boldsymbol{v}_2$ in W_2.

- Ex 2: (Orthogonal subspaces)

The subspaces

$$W_1 = \text{span}\left(\begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}\right) \text{ and } W_2 = \text{span}\left(\begin{bmatrix} -1 \\ 1 \\ 1 \end{bmatrix}\right)$$

are orthogonal because $\langle \boldsymbol{v}_1, \boldsymbol{v}_2 \rangle = 0$ for any vector in W_1 and any vector in W_2 is zero.

- Orthogonal complement of W:

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

(a) A vector $\mathbf{u}$ in V is said to **orthogonal to W**,

if $\mathbf{u}$ is orthogonal to every vector in W.

(b) The set of all vectors in V that are orthogonal to every

vector in W is called the **orthogonal complement of W**.

$$W^{\perp} = \{\mathbf{v} \in V \mid \langle \mathbf{v}, \mathbf{w} \rangle = 0, \forall\, \mathbf{w} \in W\}$$

$W^{\perp}$ (read " W perp")

- Notes:

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

- Notes:

W is a subspace of V

(1) $W^{\perp}$ is a subspace of V

(2) $W \cap W^{\perp} = \{\mathbf{0}\}$

(3) $(W^{\perp})^{\perp} = W$

- Ex:

If $V = R^2, \quad W = x-\text{axis}$

Then (1) $W^{\perp} = y\text{-axis}$ is a subspace of R^2

(2) $W \cap W^{\perp} = \{(0,0)\}$

(3) $(W^{\perp})^{\perp} = W$

- Direct sum:

Let W_1 and W_2 be two subspaces of R^n. If each vector $\mathbf{x} \in R^n$ can be uniquely written as a sum of a vector $\mathbf{w}_1$ from W_1 and a vector $\mathbf{w}_2$ from W_2, $\mathbf{x} = \mathbf{w}_1 + \mathbf{w}_2$, then R^n is the direct sum of W_1 and W_2, and you can write .

$$R^n = W_1 \oplus W_2$$

- Thm 5.13: (Properties of orthogonal subspaces)

Let W be a subspace of R^n. Then the following properties are true.

(1) $\dim(W) + \dim(W^{\perp}) = n$

(2) $R^n = W \oplus W^{\perp}$

(3) $(W^{\perp})^{\perp} = W$

Exercise 71 of Section 5.3

- Let A be a general $m \times n$ matrix.

(a) Explain why $CS(A^T)$ is the same as the row space of A.

(b) Prove that $NS(A) \subset CS(A^T)^{\perp}$

(c) Prove that $NS(A) = CS(A^T)^{\perp}$

(d) Prove that $NS(A^T) = CS(A)^{\perp}$

Proof of Exercise 71

(a) The row space of A is the column space of A^T.

(b) Let $x \in NS(A) \rightarrow Ax = 0 \rightarrow x$ is orthogonal to all the rows of A.$\rightarrow x$ is orthogonal to all the columns of A^T.$\rightarrow x \in CS(A^T)^{\perp}$

(c) Let $x \in CS(A^T)^{\perp} \rightarrow x$ is orthogonal to all the columns of A^T $\rightarrow x$ is orthogonal to all the rows of A. $\rightarrow Ax = 0 \rightarrow x \in NS(A)$. Combining this with part (b) $\rightarrow NS(A) = CS(A^T)^{\perp}$.

(d) Substitute A^T for A in part (c).

Proof of Theorem 5.13 (1)

- If $S = R^n$ or $S = \{0\}$, then Property 1 is trivial. So, let $\{v_1, v_2, \cdots, v_t\}$ be a basis for S, $0 < t < n$.

- Let A be the $n \times t$ matrix whose columns are the basis vectors v_i.

- Then $S = CS(A)$ (the column space of A), which implies that $S^{\perp} = NS(A^T)$, where A^T is a $t \times n$ matrix of rank t.

- Because the dimension of $NS(A^T)$ is $n - t$, we have shown that $\dim(S) + \dim(S^{\perp}) = t + (n - t) = n$.

rank(A^T)+nullity(A^T)=n

Proof of Theorem 5.13 (2)

- If $S = R^n$ or $S = \{0\}$, then Property 2 is trivial. So, let $\{v_1, v_2, \cdots, v_t\}$ be a basis for S, $0 < t < n$ and let $\{v_{t+1}, v_{t+2}, \cdots, v_n\}$ be a basis for $S^\perp$.

- It can be shown that $\{v_1, v_2, \cdots, v_t, v_{t+1}, v_{t+2}, \cdots, v_n\}$ is linearly independent and forms a basis for R^n.

- If you write $v = c_1 v_1 + \cdots + c_t v_t$ and $w = c_{t+1} v_{t+1} + \cdots + c_n v_n$, then you have expressed an arbitrary vector x as a vector from S and a vector from $S^\perp$, $x = v + w$.

Proof of Theorem 5.13 (2)

- To show the uniqueness of this representation, assume $x = v + w = v' + w'$, where v' is in S and w' is in $S^{\perp}$.

- This implies $v' - v = w' - w$.

- So, the two vectors $v' - v$ and $w' - w$ are in both S and $S^{\perp}$.

- Because $S \cap S^{\perp} = \{0\}$, we have $v' = v$ and $w' = w$.

Proof of Theorem 5.13 (3)

- Let $v \in S$. Then $v \cdot u = 0$ for all $u \in S^{\perp}$, which implies that $v \in (S^{\perp})^{\perp}$. $\therefore S \subseteq (S^{\perp})^{\perp}$

- On the other hand, if $v \in (S^{\perp})^{\perp}$, then, because $R^n = S \oplus S^{\perp}$, you can write v as the unique sum of the vector from S and a vector from $S^{\perp}$, $v = s + w, s \in S, w \in S^{\perp}$.

- Because w is in $S^{\perp}$, it is orthogonal to every vector in S, and in particular to v.

- So, $0 = w \cdot v = w \cdot (s + w) = w \cdot s + w \cdot w = w \cdot w$. This implies that $w = 0$ and $v = s + w = s \in S$.

$\therefore (S^{\perp})^{\perp} \subseteq S$

- Thm 5.14: (Projection onto a subspace)

If $\{\boldsymbol{u}_1, \boldsymbol{u}_2, \cdots, \boldsymbol{u}_t\}$ is an orthonormal basis for the subspace W of V, and $\mathbf{v} \in V$, then

$$\text{proj}_W \mathbf{v} = \langle \mathbf{v}, \mathbf{u}_1 \rangle \mathbf{u}_1 + \langle \mathbf{v}, \mathbf{u}_2 \rangle \mathbf{u}_2 + \cdots + \langle \mathbf{v}, \mathbf{u}_t \rangle \mathbf{u}_t$$

Pf:

$\because \text{proj}_W \boldsymbol{v} \in$ W and $\{\boldsymbol{u}_1, \boldsymbol{u}_2, \cdots, \boldsymbol{u}_t\}$ is an orthonormal basis for W

$$\Rightarrow \text{proj}_W \boldsymbol{v} = \langle \text{proj}_W \boldsymbol{v}, \boldsymbol{u}_1 \rangle \boldsymbol{u}_1 + \cdots + \langle \text{proj}_W \boldsymbol{v}, \boldsymbol{u}_t \rangle \boldsymbol{u}_t$$

$$= \langle \boldsymbol{v} - \text{proj}_{W^\perp} \boldsymbol{v}, \boldsymbol{u}_1 \rangle \boldsymbol{u}_1 + \cdots + \langle \boldsymbol{v} - \text{proj}_{W^\perp} \boldsymbol{v}, \boldsymbol{u}_t \rangle \boldsymbol{u}_t$$

$V = W \oplus W^\perp$

$$(\because \text{proj}_W \boldsymbol{v} = \boldsymbol{v} - \text{proj}_{W^\perp} \boldsymbol{v})$$

$$= \langle \boldsymbol{v}, \boldsymbol{u}_1 \rangle \boldsymbol{u}_1 + \cdots + \langle \boldsymbol{v}, \boldsymbol{u}_t \rangle \boldsymbol{u}_t \quad (\because \langle \text{proj}_{W^\perp} \boldsymbol{v}, \boldsymbol{u}_i \rangle = 0, \forall\, i)$$

- Ex 5: (Projection onto a subspace)

$$\mathbf{w}_1 = (0, 3, 1),\ \mathbf{w}_2 = (2, 0, 0),\ \mathbf{v} = (1, 1, 3)$$

Find the projection of the vector **v** onto the subspace W.

Sol:

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

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

$$\Rightarrow \{\mathbf{u}_1, \mathbf{u}_2\} = \left\{\frac{\mathbf{w}_1}{\|\mathbf{w}_1\|}, \frac{\mathbf{w}_2}{\|\mathbf{w}_2\|}\right\} = \left\{(0, \frac{3}{\sqrt{10}}, \frac{1}{\sqrt{10}}), (1,0,0)\right\} \quad :$$

an orthonormal basis for W

$$\text{proj}_W \boldsymbol{v} = \langle \boldsymbol{v}, \boldsymbol{u}_1 \rangle \boldsymbol{u}_1 + \langle \boldsymbol{v}, \boldsymbol{u}_2 \rangle \boldsymbol{u}_2$$

$$= \frac{6}{\sqrt{10}}(0, \frac{3}{\sqrt{10}}, \frac{1}{\sqrt{10}}) + (1,0,0) = (1, \frac{9}{5}, \frac{3}{5})$$

- Find by the other method:

$$A=[\mathbf{w}_1,\mathbf{w}_2],\quad \mathbf{b}=\mathbf{v}$$

$$Ax=\mathbf{b}$$

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

$$\Rightarrow \operatorname{proj}_{cs(A)}\mathbf{b}=Ax=A(A^{\mathrm{T}}A)^{-1}A^{\mathrm{T}}\mathbf{b}$$

(will later be shown by Least Square Problem)

- Thm 5.15: (Orthogonal projection and distance)

Let W be a subspace of an inner product space V, and $\mathbf{v} \in V$

Then for all $\mathbf{w} \in W$, $\mathbf{w} \neq \text{proj}_W \mathbf{v}$

$$\|\mathbf{v} - \text{proj}_W \mathbf{v}\| < \|\mathbf{v} - \mathbf{w}\|$$

or $\|\boldsymbol{v} - \text{proj}_W \boldsymbol{v}\| = \min_{\boldsymbol{w} \in \mathrm{W}} \|\boldsymbol{v} - \boldsymbol{w}\|$

($\text{proj}_W \mathbf{v}$ is the best approximation to **v** from W)

- Pf:

$$\mathbf{v}-\mathbf{w}=(\mathbf{v}-\text{proj}_W\mathbf{v})+(\text{proj}_W\mathbf{v}-\mathbf{w})$$

$$(\mathbf{v}-\text{proj}_W\mathbf{v})\perp(\text{proj}_W\mathbf{v}-\mathbf{w})$$

By the Pythagorean theorem

$$\Rightarrow\|\mathbf{v}-\mathbf{w}\|^2=\|\mathbf{v}-\text{proj}_W\mathbf{v}\|^2+\|\text{proj}_W\mathbf{v}-\mathbf{w}\|^2$$

$$\mathbf{w}\neq\text{proj}_W\mathbf{v}\Rightarrow\|\text{proj}_W\mathbf{v}-\mathbf{w}\|>0$$

$$\Rightarrow\|\mathbf{v}-\mathbf{w}\|^2>\|\mathbf{v}-\text{proj}_W\mathbf{v}\|^2$$

$$\Rightarrow\|\mathbf{v}-\text{proj}_W\mathbf{v}\|<\|\mathbf{v}-\mathbf{w}\|$$

- Notes:

(1) Among all the scalar multiples of a vector **u**, the orthogonal projection of **v** onto **u** is the one that is closest to **v**. (p.244 Thm 5.9)

(2) Among all the vectors in the subspace W, the vector $\text{proj}_W \mathbf{v}$ is the closest vector to **v**.

- Thm 5.16: (Fundamental subspaces of a matrix)

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

(1) $(CS(A))^{\perp} = NS(A^{\mathrm{T}})$

$(NS(A^{\mathrm{T}}))^{\perp} = CS(A)$

(2) $(CS(A^{\mathrm{T}}))^{\perp} = NS(A)$

$(NS(A))^{\perp} = CS(A^{\mathrm{T}})$

(3) $CS(A) \oplus NS(A^{T}) = R^{m}$ $\qquad CS(A) \oplus (\overset{C}{\cancel{N}}S(A))^{\perp} = R^{m}$

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

Proof of Theorem 5.16

- To prove property 1, let $v \in CS(A)$ and $u \in NS(A^T)$. Because the column space of A is equal to the row space of A^T, you can see that $A^T u = 0$ implies $u \cdot v = 0$.

- Property 2 follows from applying Property 1 to A^T.

- To prove Property 3, observe that $CS(A)^{\perp} = NS(A^T)$ and $R^m = CS(A) \oplus CS(A)^{\perp}$. So, $R^m = CS(A) \oplus NS(A^T)$.

- A similar argument applied to $CS(A^T)$ proves Property 4.

- Ex 6: (Fundamental subspaces)

Find the four fundamental subspaces of the matrix.

$$A=\begin{bmatrix}1 & 2 & 0\\ 0 & 0 & 1\\ 0 & 0 & 0\\ 0 & 0 & 0\end{bmatrix}$$ (reduced row-echelon form)

Sol:

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

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

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

$$A^{\mathrm{T}} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 2 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix} \sim R = \begin{array}{c} \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \\ \begin{array}{cccc} & & s & t \end{array} \end{array}$$

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

- Check:

$$(CS(A))^{\perp} = NS(A^{\mathrm{T}})$$

$$(CS(A^{\mathrm{T}}))^{\perp} = NS(A)$$

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

$$CS(A^{T}) \oplus NS(A) = R^3$$

- Ex 3 & Ex 4:

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

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.

Sol:

$$A = \begin{array}{c} \begin{bmatrix} 1 & 0 \\ 2 & 0 \\ 1 & 0 \\ 0 & 1 \end{bmatrix} \\ \mathbf{w}_1 \ \ \mathbf{w}_2 \end{array} \sim R = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \\ 0 & 0 \end{bmatrix} \quad \text{(reduced row-echelon form)}$$

$(a)\quad W = CS(A)$

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

$(b)\quad W^{\perp} = (CS(A))^{\perp} = NS(A^{\mathrm{T}})$

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

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

- Notes:

  (1) $\dim(W) + \dim(W^{\perp}) = \dim(R^4)$

  (2) $W \oplus W^{\perp} = R^4$

- Least squares problem:

$$\underset{m\times n}{A}\ \underset{n\times 1}{\mathbf{x}} = \underset{m\times 1}{\mathbf{b}} \qquad \text{(A system of linear equations)}$$

(1) When the system is consistent, we can use the Gaussian elimination with back-substitution to solve for $\mathbf{x}$

(2) When the system is inconsistent, how to find the "best possible" solution of the system. That is, the value of $\mathbf{x}$ for which the difference between $A\mathbf{x}$ and $\mathbf{b}$ is small.

- Least squares solution:

  Given a system $A\mathbf{x} = \mathbf{b}$ of m linear equations in n unknowns, the least squares problem is to find a vector $\mathbf{x}$ in R^n that minimizes $\|A\mathbf{x} - \mathbf{b}\|$ with respect to the Euclidean inner product on R^n. Such a vector is called a least squares solution of $A\mathbf{x} = \mathbf{b}$.

- Notes:

  The least square problem is to find a vector $\hat{\boldsymbol{x}}$ in R^n such that $A\hat{\boldsymbol{x}} = \operatorname{proj}_{CS(A)}\boldsymbol{b}$ in the column space of A (i.e., $A\hat{\boldsymbol{x}} \in CS(A)$) is as close as possible to $\boldsymbol{b}$. That is,

$$\left\|\boldsymbol{b} - \operatorname{proj}_{CS(A)}\boldsymbol{b}\right\| = \left\|\boldsymbol{b} - A\hat{\boldsymbol{x}}\right\| = \min_{\boldsymbol{x}\in R^n}\left\|\boldsymbol{b} - A\boldsymbol{x}\right\|$$

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

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

$$A\boldsymbol{x} \in CS(A) \quad (CS(A) \text{ is a subspace of } R^m)$$

$$\because \mathrm{b} \notin CS(A) \quad (A\boldsymbol{x} = \mathrm{b} \text{ is an inconsistent system})$$

Let $A\hat{\boldsymbol{x}} = proj_{CS(A)}\boldsymbol{b}$

$$\Rightarrow (\boldsymbol{b} - A\hat{\boldsymbol{x}}) \perp CS(A)$$

$$\Rightarrow \boldsymbol{b} - A\hat{\boldsymbol{x}} \in (CS(A))^{\perp} = NS(A^{\mathrm{T}})$$

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

i.e. $A^{\mathrm{T}}A\hat{\boldsymbol{x}} = A^{\mathrm{T}}\boldsymbol{b}$ (the normal system associated with $A\mathbf{x} = \mathbf{b}$)

- Note: ($A\mathbf{x} = \mathbf{b}$ is an inconsistent system)

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

is equal to he problem of finding an exact solution of the

associated normal system $A^{\mathrm{T}} A\hat{\mathbf{x}} = A^{\mathrm{T}}\mathbf{b}$.

- Ex 7: (Solving the normal equations)

Find the least squares solution of the following system

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

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

(this system is inconsistent)

and find the orthogonal projection of **b** on the column space of *A*.

- Sol:

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

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

the associated normal system

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

$$\begin{bmatrix} 3 & 6 \\ 6 & 14 \end{bmatrix} \begin{bmatrix} c_0 \\ c_1 \end{bmatrix} = \begin{bmatrix} 4 \\ 11 \end{bmatrix}$$

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

$$\hat{\boldsymbol{x}} = \begin{bmatrix} -\frac{5}{3} \\ \frac{3}{2} \end{bmatrix}$$

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

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

Key Learning in Section 5.4

- Define the least squares problem.
- Find the orthogonal complement of a subspace and the projection of a vector onto a subspace.
- Find the four fundamental subspaces of a matrix.
- Solve a least squares problem.
- Use least squares for mathematical modeling.

Keywords in Section 5.4

- orthogonal to W: 正交於W
- orthogonal complement: 正交補集
- direct sum: 直和
- projection onto a subspace: 在子空間的投影
- fundamental subspaces: 基本子空間
- least squares problem: 最小平方問題
- normal equations: 一般方程式

Review exercises

59. Find the orthogonal complement $S^{\perp}$ of the subspace S of R^3 spanned by the two column vectors of the matrix

$$A = \begin{bmatrix} 1 & 2 \\ 2 & 1 \\ 0 & -1 \end{bmatrix}.$$

60. Find the projection of the vector $v = [1 \quad 0 \quad -2]^T$ onto the subspace

$$S = span\left\{ \begin{bmatrix} 0 \\ -1 \\ 1 \end{bmatrix}, \quad \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix} \right\}.$$

5.1 Linear Algebra Applied

- Electric/Magnetic Flux

Electrical engineers can use the dot product to calculate electric or magnetic *flux*, which is a measure of the strength of the electric or magnetic field penetrating a surface. Consider an arbitrarily shaped surface with an element of area dA, normal (perpendicular) vector $d\mathbf{A}$, electric field vector $\mathbf{E}$ and magnetic field vector $\mathbf{B}$. The electric flux Φ_e can be found using the surface integral $\Phi_e = \int \mathbf{E}\cdot d\mathbf{A}$ and the magnetic flux can be found using the surface integral $\Phi_e = \int \mathbf{B}\cdot d\mathbf{A}$. It is interesting to note that for a given closed surface that surrounds an electrical charge, the net electric flux is proportional to the charge, but the net magnetic flux is zero. This is because electric fields initiate at positive charges and terminate at negative charges, but magnetic fields form closed loops, so they do not initiate or terminate at any point. This means that the magnetic field entering a closed surface must equal the magnetic field leaving the closed surface.

- Work

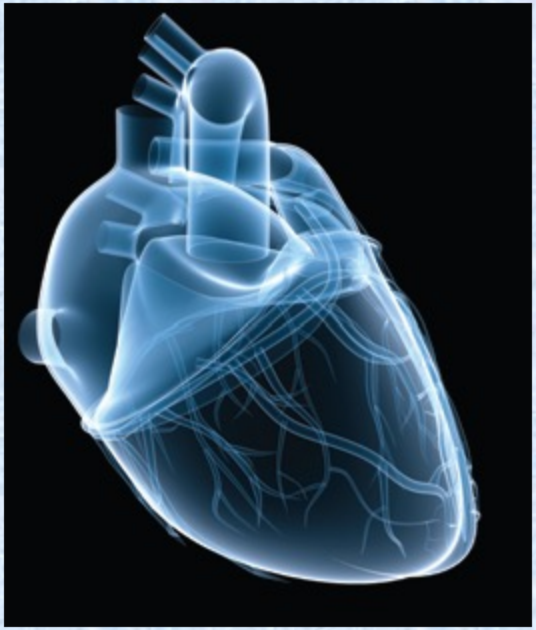
The concept of work is important to scientists and engineers for determining the energy needed to perform various jobs. If a constant force **F** acts at an angle θ with the line of motion of an object to move the object from point A to point B (see figure below), then the work done by the force is given by

$$W = (\cos\theta)\|\mathbf{F}\|\|\overrightarrow{AB}\|$$

$$= \mathbf{F} \cdot \overrightarrow{AB}$$

where $\overrightarrow{AB}$ represents the directed line segment from A to B. The quantity $(\cos\theta)\|\mathbf{F}\|$ is the length of the orthogonal projection of **F** onto $\overrightarrow{AB}$ Orthogonal projections are discussed on the next page.

- Heart Rhythm Analysis



Time-frequency analysis of irregular physiological signals, such as beat-to-beat cardiac rhythm variations (also known as heart rate variability or HRV), can be difficult. This is because the structure of a signal can include multiple periodic, nonperiodic, and pseudo-periodic components. Researchers have proposed and validated a simplified HRV analysis method called orthonormal-basis partitioning and time-frequency representation (OPTR). This method can detect both abrupt and slow changes in the HRV signal's structure, divide a nonstationary HRV signal into segments that are "less nonstationary," and determine patterns in the HRV. The researchers found that although it had poor time resolution with signals that changed gradually, the OPTR method accurately represented multicomponent and abrupt changes in both real-life and simulated HRV signals.

5.4 Linear Algebra Applied

- Revenues

The least squares problem has a wide variety of real-life applications. To illustrate, in Examples 9 and 10 and Exercises 37, 38, and 39, you will use least squares analysis to solve problems involving such diverse subject matter as world population, astronomy, doctoral degrees awarded, revenues for General Dynamics Corporation, and galloping speeds of animals. In each of these applications, you will be given a set of data and you will be asked to come up with mathematical model(s) for the data. For instance, in Exercise 38, you are given the annual revenues from 2005 through 2010 for General Dynamics Corporation, and you are asked to find the least squares regression quadratic and cubic polynomials for the data. With these models, you are asked to predict the revenue for the year 2015, and you are asked to decide which of the models appears to be more accurate for predicting future revenues.

- Torque

In physics, the cross product can be used to measure torque—the moment **M** of a force **F** about a point A as shown in the figure below. When the point of application of the force is B, the moment of **F** about A is given by

$$\mathbf{M} = \overrightarrow{AB} \times \mathbf{F}$$

where $\overrightarrow{AB}$ represents the vector whose initial point is A and whose terminal point is B. The magnitude of the moment **M** measures the tendency of $\overrightarrow{AB}$ to rotate counterclockwise about an axis directed along the vector **M**.