Chapter 5 : Orthogonality

5.1 scalar product in $\mathbb{R}^n$

① Def :

$u \in \mathbb{R}^n \qquad u = \begin{pmatrix} u_1 \\ \vdots \\ u_n \end{pmatrix}$

$v \in \mathbb{R}^n \qquad v = \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix}$

$u^T v = u_1 v_1 + \cdots u_n v_n \Longleftrightarrow$ sclar product (dot product) of $u, v \in \mathbb{R}^n$

② Length / norm

$\|u\| = \sqrt{u^T u} = \sqrt{u_1^2 + u_2^2 + \cdots u_n^2}$

unit vector : if $\|u\| = 1$

For any $v \in \mathbb{R}^n$, $u = \frac{v}{\|v\|}$ is a unit vector with the same direct as v

distance between u and v : $\|u - v\|$

the angle θ : $\cos\theta = \frac{u^T v}{\|u\| \|v\|}$

$|u^T v| \le \|u\| \|v\|$

$$\boxed{|u_1 v_1 + \cdots u_n v_n| \le \sqrt{u_1^2 + u_2^2 + \cdots u_n^2} \sqrt{v_1^2 + \cdots v_n^2}}$$

Cauchy-Schwarz Inequality

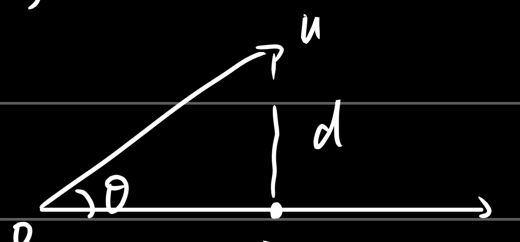
③ orthogonal

$u, v \in \mathbb{R}^n$, u is orthogonal to v if

$u^T v = 0$

④ Scalar projection and vector projection

in $\mathbb{R}^2$



p: vector projection of u onto v

d: Scalar projection of u onto v

For given u and v, <i> Find "θ" . $\cos\theta = \frac{u^Tv}{\|u\|\|v\|}$

<ii> $d = \|u\| \sin\theta$

<iii> $p = \alpha v$ $\quad \|p\| = \|\alpha\|\|v\|$

$\|u\| \cos\theta = \|\alpha\|\|v\|$

$\|\alpha\| = \frac{\|u\|}{\|v\|} \cos\theta = \frac{\cancel{\|u\|}\ u^Tv}{\|v\|\ \cancel{\|u\|}\|v\|}$

$= \frac{u^Tv}{\|v\|^2}$

5.2 orthogonal subspace

① Def

Two subspaces X and Y of $\mathbb{R}^n$ are said to be **orthogonal** if $\mathbf{x}^T\mathbf{y} = 0$ for every $\mathbf{x} \in X$ and every $\mathbf{y} \in Y$. If X and Y are orthogonal, we write $X \perp Y$.

Ex: $V = \{\alpha \begin{pmatrix} 1 \\ 2 \end{pmatrix} \mid \alpha \in R\}$ $\quad \Rightarrow V \perp U$

$U = \{\alpha \begin{pmatrix} -2 \\ 1 \end{pmatrix} \mid \alpha \in R\}$

more general case:

V: base $\{v_1, v_2 \cdots v_r\}$

U: base $\{u_1, u_2 \cdots u_k)$

Then V and U are orthogonal if and only if

$v_i^T u_j = 0 \quad i = 1, 2, \cdots r$

$j = 1, 2 \cdots k$

proof: <i> $\Rightarrow$ V and U are orthogonal

By definition, $v_i \in V$, $u_j \in U$

$v_i^T u_j = 0$

<ii> $\Leftarrow$ If $v_i^T u_j = 0$, for any $v \in V$, $u \in U$

we have $v = x_1 v_1 + \cdots x_r v_r$

$u = y_1 u_1 + \cdots y_k u_k$

Then $v^T u = (x_1 v_1 + x_2 v_2 + \cdots x_r v_r)^T (y_1 u_1 + \cdots y_k u_k)$

$= \sum_i x_i v_i^T \sum_j y_j u_j$

$= \sum_i \sum_j x_i y_i \, v_i^T u_j = 0$

then $v_i^T u_j = 0$

Ex: $N(A^T) = \{x \mid A^T x = 0\}$

let $A = (a_1 \cdots a_n)$, then $N(A^T) = \{x \mid a_j^T x = 0, j = 1, \cdots n\}$ $(x \perp a_j)$

$\Rightarrow N(A^T) \perp \text{Span}\{a_1, a_2 \cdots a_n\}$

Column space $= R(A)$

$x \in N(A^T)$, $y \in R(A)$

Prove <i> $x^T y = 0$

<ii> x and y are linearly independent.

Solution. <i> let $y = \alpha_1 a_1 + \alpha_2 a_2 + \cdots \alpha_n a_n = A\alpha$

to prove $x^T y = 0$, we can prove $y^T x = 0$

i.e. $\alpha^T A^T x = 0$, since $x \in N(A^T)$, then $A^T x = 0$

so $y^T x = 0$, $x y^T = 0$

<ii> Suppose $C_1 x + C_2 y = 0$

<1> $x^T (C_1 x + C_2 y) = 0$ $\quad C_1 x^T x + C_2 x^T y = 0$

$\Rightarrow C_1 x^T x = 0$ $\quad C_1 = 0$, so we also have $C_2 = 0$

so $C_1 x + C_2 y = 0$ has only zero sol.

$\Rightarrow x, y$ are linearly independent.

② orthogonal complement

Def.

Let Y be a subspace of $\mathbb{R}^n$. The set of all vectors in $\mathbb{R}^n$ that are orthogonal to every vector in Y will be denoted $Y^\perp$. Thus,

$$Y^\perp = \{\mathbf{x} \in \mathbb{R}^n \mid \mathbf{x}^T\mathbf{y} = 0 \quad \text{for every} \quad \mathbf{y} \in Y\}$$

The set $Y^\perp$ is called the **orthogonal complement** of Y.

Ex: $S \subset \mathbb{R}^n$, then $S^\perp$ is a subspace of $\mathbb{R}^n$

<i> nonempty $0^T \cdot x = 0$. $0 \in S^\perp$

<ii> $u, v \in S^\perp$, $u^T x = v^T x = 0$ for all $x \in S$

$\Rightarrow (u^T + v^T) x = (u+v)^T x = 0$

which shows $u+v \in S^\perp$

<iii> $u \in S^\perp$, $u^T x = 0$, $\alpha u^T x = 0$

so $S^\perp$ is a subspace of $\mathbb{R}^n$

Ex: $S = \text{span}\{\underbrace{v_1 \cdots v_k}_{}\} \subset \mathbb{R}^n$ $R(A)$ $A = (v_1 \cdots v_k)$

Find a basis of $S^\perp$

sol. $S^\perp = \{x \mid x^T y = 0,\ x \in S\}$

$v_j^T y = 0 \quad j = 1, 2 \cdots k$

$(v_1 \cdots v_k)^T \cdot y = 0$

$S^\perp = \{y \in \mathbb{R}^n \mid A^T \cdot y = 0\}$

$S^\perp = N(A^T)$

Ex: $S = \text{span}\{a_1, a_2, a_3\} \subset \mathbb{R}^4$

$a_1 = \begin{Bmatrix} 1 \\ 0 \\ 1 \\ 0 \end{Bmatrix}$ $a_2 = \begin{Bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{Bmatrix}$ $a_3 = \begin{Bmatrix} 2 \\ 1 \\ 0 \\ 0 \end{Bmatrix}$

Find a basis for $S^\perp$

sol. let $A = \begin{pmatrix} 1 & 0 & 2 \\ 0 & 1 & 5 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$ $A^T = \begin{pmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 2 & 5 & 0 & 0 \end{pmatrix}$

③ Four Fundamental Subspace

$A_{m\times n} \Rightarrow$ subspace

$R(A) = \{ y \mid y = Ax \mid x \in \mathbb{R}^n \}$

$N(A) = \{ x \in \mathbb{R}^n \mid Ax = 0 \}$

$R(A^T)$ = row space

$N(A^T)$

Theorem 5.2.1 Fundamental Subspaces Theorem

If A is an $m \times n$ matrix, then $N(A) = R(A^T)^\perp$ and $N(A^T) = R(A)^\perp$.

proof: $N(A^T) = R(A)^\perp$ $A = (a_1 \cdots a_n)$

(i) For $u \in N(A^T)$, $A^T u = 0$

$\Downarrow$

$\begin{pmatrix} a_1^T \\ \vdots \\ a_n^T \end{pmatrix} u = 0$

$a_j^T u = 0$, $j = 1 \cdots n$

$u \perp \text{Span}\{a_1 \cdots a_n\}$

$u \perp R(A) \Rightarrow u \subset R(A)^\perp \Rightarrow N(A^T) \subset R(A)^\perp$

(ii) For $v \in R(A)^\perp$,

$v^T y = 0$ for all $y \in R(A)$

for $y = a_j$

$v^T a_j = 0$, $a_j^T v = 0$, $\Rightarrow v \in N(A^T) \Rightarrow R(A)^\perp \subset N(A^T)$

So $N(A^T) = R(A)^\perp$

☆ Theorem. S is a subspace of $\mathbb{R}^n$, then

(i) $\dim(S) + \dim(S^\perp) = n$

(ii) If $\{x_1 \cdots x_r\}$ is a base for S

$\{x_{r+1} \cdots x_n\}$ is a base for $S^\perp$

(iii) then $\{x_1 \cdots x_n\}$ is a base for $\mathbb{R}^n$

proof: (i) If $S = \{0\}$, $S^\perp = \mathbb{R}$

If $S = \mathbb{R}$, $S^\perp = \{0\}$

For general case, we assume

$\{x_1 \cdots x_r\}$ is a base for S, $0 < r < n$

Let $X^T = (x_1, x_2 \cdots x_r)$, $x_j \in \mathbb{R}^n$

$S = \text{span}\{x_1\ x_2 \cdots x_r\} = R(X^T)$

By theorem 5.2.1

$$R(X^T)^\perp = N(X)$$

by theorem 3.6.5

$$\underset{\substack{\| \\ \dim(R(X^T)) \\ \| \\ \dim(S)}}{\text{Rank}(X)} + \dim N(X) = n$$

$\dim(S) + \dim(S^\perp) = n$

(ii) we only need to prove $\{x_1 \cdots x_n\}$ are linearly ind

$$\underbrace{\alpha_1 x_1 + \cdots \alpha_r x_r}_{u \in S} + \underbrace{\alpha_{r+1} x_{r+1} + \cdots \alpha_n x_n}_{v \in S^\perp} = 0, \quad \alpha_j \in \mathbb{R}$$

$\Rightarrow$ $u + v = 0$ $\qquad$ $v^T(u+v) = 0$

$u^T(u+v) = 0$ $\qquad$ $v^T u + v^T v = 0$

$u^T u = 0$ $\qquad$ $v^T v = 0$

$u = 0$ $\qquad$ $v = 0$

$\Rightarrow$ $u = v = 0$

Another proof: $u+v=0$, $S \ni u = -v \in S^\perp$

So $u, v \in S \cap S^\perp = \{0\}$

So $u = v = \{0\}$

Since $u = v = 0 \Rightarrow a_j = 0 \quad j = 1 \cdots n$

$\{x_1 \cdots x_n\}$ are linearly independent.

④ Direct Sum

Definition If U and V are subspaces of a vector space W and each $\mathbf{w} \in W$ can be written uniquely as a sum $\mathbf{u} + \mathbf{v}$, where $\mathbf{u} \in U$ and $\mathbf{v} \in V$, then we say that W is a **direct sum** of U and V, and we write $W = U \oplus V$.

Theorem 5.2.3 *If S is a subspace of $\mathbb{R}^n$, then*

$$\mathbb{R}^n = S \oplus S^\perp$$

proof. $\dim(S) = r \quad 0 < r < n$

By theorem 5.2.2

$\{x_1 \cdots x_r\}$ is a basis for S

$\{x_{r+1} \cdots x_n\}$ is a basis for $S^\perp$

then $\{x_1 \cdots x_n\}$ is a basis for $\mathbb{R}^n$

For any $w \in \mathbb{R}^n$

$$w = \underbrace{c_1x_1 + \cdots c_rx_r}_{u} + \underbrace{c_{r+1}x_{r+1} + \cdots c_nx_n}_{v}$$

$w = u + v \quad u \in S,\ v \in S^\perp$ (existence)

If $w = x + y \quad x \in S,\ y \in S^\perp$

we have $u + v = x + y$

$S \ni u - x = y - v \in S^\perp$

So $u-x,\ y-v \in S \cap S^{\perp}$

So $u-x = y-v = 0$, then $u=x,\ y=v$

we prove the uniqueness

Theorem 5.2.4 *If S is a subspace of $\mathbb{R}^n$, then $(S^{\perp})^{\perp} = S$.*

proof: For any $x \in S$,

$x \perp y$, for any $y \in S^{\perp} \Rightarrow x \in (S^{\perp})^{\perp}$

$S \subset (S^{\perp})^{\perp}$

for any $z \in (S^{\perp})^{\perp}$, $z \perp y$ for any $y \in S^{\perp}$

By theorem 5.2.3

$z = u + v,\ u \in S,\ v \in S^{\perp}$

$0 = v^T z = v^T(u+v) = v^T u + v^T v = v^T v$

$\Rightarrow v = 0$

$\Rightarrow z = u \in S$

so $(S^{\perp})^{\perp} \subset S$

so $(S^{\perp})^{\perp} = S$

Remarks

(1) $N(A)^{\perp} = R(A^T)$, $N(A^T)^{\perp} = R(A)$

(2) If $\{v_1 \cdots v_r\}$ is a base for $R(A^T)$

then $\{Av_1 \cdots Av_r\}$ is a base for $R(A)$

(Hint: $\{Av_1 \cdots Av_r\}$ is linearly independent)

(3) $N(A) = N(A^T A)$

$N(A^T) = N(AA^T)$

5.3 Least square problems

① Def $Ax = b$

residual: $r(x) = b - Ax$

$$\min_{x \in \mathbb{R}^n} \underline{\|r(x)\| = \sqrt{r^T r} = \|b - Ax\|} = r(\hat{x})$$

least $\Updownarrow$ square $\qquad$ $\hat{x}$: least square solution.

$$\|r(\hat{x})\| \le \|r(x)\| \text{ for any } x \in \mathbb{R}^n$$

Ex: $A = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix}$ $\quad b = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$

$$r(x) = b - Ax = \begin{bmatrix} 1 - x_1 \\ 2 - x_2 \\ 3 \end{bmatrix}$$

$$\|r(x)\| = \sqrt{r^T r} = \sqrt{(1-x_1)^2 + (2-x_2)^2 + 3^2}$$

Least square sol $\hat{x} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$ $\quad \|r(x)\| = 3$

$\min \|b - y\|,\ y \in R(A) = \min \|b - Ax\|,\ x \in \mathbb{R}^n$

Since $R(A)$ is a vector space.

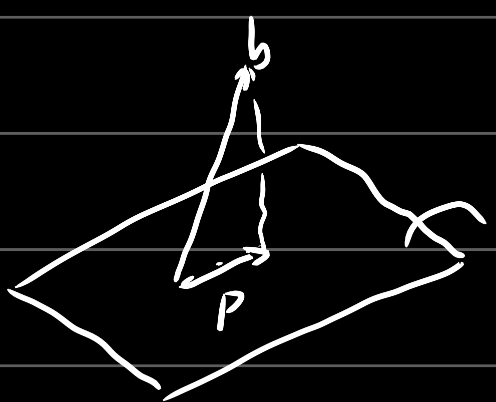
we assume $\{a_1, a_2 \cdots a_n\}$

is a base of $R(A)$

For any $y \in R(A)$

$$y = x_1 a_1 + x_2 a_2 + \cdots + x_n a_n$$

$$y = Ax, \quad A = (a_1 \cdots a_n) \quad x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$$



R(A) $\quad$ P: projection of b onto R(A)

$$\min \|b - y\| \quad y \in V_n = \text{span}\{v_1 \cdots v_n\}$$

$y = (v_1 \cdots v_n)x \quad , \quad y = Ax$

② Theorem

Theorem 5.3.1 *Let S be a subspace of $\mathbb{R}^m$. For each $\mathbf{b} \in \mathbb{R}^m$, there is a unique element $\mathbf{p}$ of S that is closest to $\mathbf{b}$; that is,*

$$\|\mathbf{b} - \mathbf{y}\| > \|\mathbf{b} - \mathbf{p}\|$$

for any $\mathbf{y} \neq \mathbf{p}$ in S. Furthermore, a given vector $\mathbf{p}$ in S will be closest to a given vector $\mathbf{b} \in \mathbb{R}^m$ if and only if $\mathbf{b} - \mathbf{p} \in S^\perp$.

proof: $S \subset \mathbb{R}^n$

$\mathbb{R}^n = S \oplus S^\perp$

by theorem, for any $b \in \mathbb{R}^n$

$b = p + z$ uniquely for $p \in S$, $z \in S^\perp$

For any $y \in S$

$b - y = \underbrace{b-p}_{z \in S^\perp = N(A^T)} + \underbrace{p-y}_{\in S = R(A)}$

$\|b-y\|^2 = \|b-p\|^2 + \|p-y\|^2$

$\|b-p\|^2 = \|b-y\|^2 - \|p-y\|^2$ for any $y \in S$

$\leq \|b-y\|^2$

p is the solution

when $S = R(A)$

$p \in R(A)$

$p = A\hat{x}$

$b - p \in R(A)^\perp = N(A^T)$

$A^T(b-p) = 0$

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

Theorem 5.3.2 *If A is an $m \times n$ matrix of rank n, the normal equations*

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

have a unique solution

$$\hat{\mathbf{x}} = (A^TA)^{-1}A^T\mathbf{b}$$

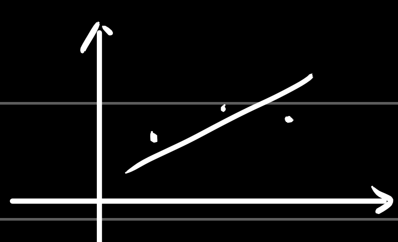
and $\hat{\mathbf{x}}$ is the unique least squares solution of the system $A\mathbf{x} = \mathbf{b}$.

$P = A\hat{x} = \boxed{A(A^TA)^{-1}A^T}b$

projection matrix P

③ Remarks

x	x_1	$\cdots$	x_n
y	y_1	$\cdots$	y_n



$$y_j = a x_j + b$$

$$\underset{A^T}{\begin{bmatrix} 1 & x_1 \\ \vdots & \vdots \\ 1 & x_n \end{bmatrix}} \underset{x}{\begin{bmatrix} b \\ a \end{bmatrix}} = \underset{b}{\begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix}}$$

$$A^T A x = A^T b$$

$$A^T A = \begin{bmatrix} n & \sum_{j=1}^{n} x_j \\ \sum_{j=1}^{n} x_j & \sum_{j=1}^{n} x_j^2 \end{bmatrix}$$

$$\min_{x \in \mathbb{R}^n} \|b - Ax\| = \|b - A\hat{x}\|$$

$$\min_{y \in V_n} \|b - y\| = \|b - p\|$$

$y \in V_n \quad V_n = R(A)$

$$y = a_0 x + b_0$$

$$b = \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix}$$

$$\|b - y\|^2 = \sum_{j=1}^{n} (a_0 x_j + b_0 - y_j)^2$$

5.4 Inner product Spaces

① Ex $C[a,b]$: {real value continous function on $[a,b]$}

Inner product $\langle f, g\rangle = \int_a^b f(x)\, g(x)\, dx$ for $f, g \in [a,b]$

$\langle f, g \rangle = 0$, then $f \perp g$

Ex $\mathbb{R}^{m\times n} = \{A_{m\times n}$ matrices$\}$

Inner product : $A, B \in \mathbb{R}^{m\times n}$

$$\langle A, B\rangle = \sum_{i,j=1}^{m,n} a_{ij} b_{ij}$$

W: vector space

Def: Inner product

$x, y \in W \rightarrow$ real value

satisfying

<i> $(x,y) = (y,x)$

(ii) $(x,x) \geq 0$ $(x,x)=0$ if and only if $x=0$

<iii> $(\alpha x+\beta y, z) = \alpha(x,z) + \beta(y,z)$

$C^0(a,b) = \{$ Continous function on $[a,b]\}$

$u, v \in C^0$, Def: $(u,v) = \int_a^b u(x)v(x)\,dx$

$\boxed{\text{Inner product space}}$

5.5 Orthonormal Set

①

☆

for orthonormal set of vectors

$$v_i^T v_j = \delta_{ij} = \begin{cases} 1 & i=j \\ 0 & i\neq j \end{cases}$$

Orthogonal + unit = orthonormal

Theorem

Theorem 5.5.1 *If $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ is an orthogonal set of nonzero vectors in an inner product space V, then $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$ are linearly independent.*

proof: Consider $C_1v_1 + C_2v_2 + \cdots + C_nV_n = 0$.

$v_i^T(C_1v_1 + C_2v_2 + \cdots C_nV_n) = 0$

$C_1v_1^Tv_1 = 0$, $v_1 = 0$

Similary: $C_j = 0$

② orthogonal / orthonormal basis

if $\{v_1 \cdots v_k\}$ is a set of vectors and forms a basis for $V \subset \mathbb{R}^n$, it's called Orthogonal basis of V

$V \subset \mathbb{R}^n$

V: basis $\{u_1, u_2 \cdots u_k\}$ → linearly independent

Existence $\{u_1, u_2 \cdots u_k\}$ is a basis for V_k

$V_k = \text{span}\{u_1, u_2 \cdots u_k\}$

$V_{k-1} = \text{span}\{u_1, u_2 \cdots u_{k-1}\}$

Let p be vector projection of u_k onto u_{k-1}

$v = u - p_k$

then you go to work on V_{k-1} (math induction)

There exists

$\{v_1, v_2 - v_{k-1}\}$ being an orthogonal base for V_{k-1}

So $\{v_1 \cdots v_k\}$ is an orthogonal basis for V_k

$\left\{\frac{v_1}{\|v_1\|} \cdots \frac{v_k}{\|v_k\|}\right\}$ is an orthonormal basis for V_k

Ex: $P_2 = \{\text{polynomials of degree } 2\}$ $\{1, x, x^2\}$ is a basis.

Inner product, $u, v \in P_3$, $\langle u, v\rangle = \int_{-1}^{1} u(x)v(x)\,dx$

orthogonal basis:

$P_0(x) = 1$

$P_1(x) = x$

$P_2(x) = \frac{1}{2}(3x^2 - 1)$

$P_3 = \text{span}\{P_0, P_1, P_2\}$

② $\int_{-1}^{1} P_i(x) P_j(x)\,dx = 0 \quad i \neq j$

Theorem 5.5.2 *Let $\{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_n\}$ be an orthonormal basis for an inner product space V. If $\mathbf{v} = \sum_{i=1}^{n} c_i \mathbf{u}_i$, then $c_i = \langle \mathbf{v}, \mathbf{u}_i \rangle$.*

let $u \in \text{span}\{u_1, u_2 \cdots u_n\}$

$u = \alpha_1 u_1 + \alpha_2 u_2 + \cdots \alpha_n u_n$

$u_1^T u = \alpha_1 u_1^T u_1 \quad \Rightarrow \quad \alpha_j = \dfrac{u_j^T u}{u_j^T u_j}$

if $\{u_1, u_2 \cdots u_n\}$ are orthonormal

$u_j^T u_j = 1$

$\alpha_j = u_j^T u$

Corollary 5.5.3 *Let $\{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_n\}$ be an orthonormal basis for an inner product space V. If $\mathbf{u} = \sum_{i=1}^{n} a_i \mathbf{u}_i$ and $\mathbf{v} = \sum_{i=1}^{n} b_i \mathbf{u}_i$, then*

$$\langle \mathbf{u}, \mathbf{v} \rangle = \sum_{i=1}^{n} a_i b_i$$

for $v = \beta_1 v_1 + \beta_2 v_2 + \cdots \beta_n v_n$

$u^T v = (\alpha_1 v_1 + \cdots \alpha_n v_n)^T (\beta_1 v_1 + \cdots \beta_n v_n)$

$= \alpha_1 \beta_1 u_1^T u_1 + \cdots \alpha_n \beta_n u_n^T u_n$

orthonormal $= \alpha_1 \beta_1 + \cdots \alpha_n \beta_n$

③ orthogonal matrix

Definition An $n \times n$ matrix Q is said to be an **orthogonal matrix** if the column vectors of Q form an orthonormal set in $\mathbb{R}^n$. ☆

Properties of Orthogonal Matrices

If Q is an $n \times n$ orthogonal matrix, then

※ square matrix

(a) the column vectors of Q form an orthonormal basis for $\mathbb{R}^n$.

(b) $Q^TQ = I$

(c) $Q^T = Q^{-1}$

(d) $\langle Q\mathbf{x}, Q\mathbf{y}\rangle = \langle \mathbf{x}, \mathbf{y}\rangle$

(e) $\|Q\mathbf{x}\|_2 = \|\mathbf{x}\|_2$

$\det(Q) = \pm 1$

Let $\{u_1, u_2 \cdots u_k\} \subset \mathbb{R}^n$ be a orthonormal basis

$P = (u_1, u_2 \cdots u_k)_{n\times k}$ $n > k$

P^TP = a $k\times k$ identity matrix.

④ orthogonal basis and least square

Theorem 5.5.6 *If the column vectors of A form an orthonormal set of vectors in $\mathbb{R}^m$, then $A^TA = I$ and the solution to the least squares problem is*

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

least square problems: $\min \|b - y\|$, $y \in V_k$

if $V_k = \text{span}\{a_1 \cdots a_k\}$, $\{a_1 \cdots a_k\}$ is a orthonormal basis

let $A = (a_1, a_2 \cdots a_k)$

Since $A^TA = I_{k\times k}$ then $A^TA\hat{x} = A^Tb$ let sol. be p

Since $p \in \text{Span}\{a_1 \cdots a_k\}$

$p = \alpha_1 a_1 + \cdots \alpha_k a_k$

$\alpha_j = p^T\alpha_j$

Theorem 5.5.7 *Let S be a subspace of an inner product space V and let $\mathbf{x} \in V$. Let $\{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_n\}$ be an orthonormal basis for S. If*

$$\mathbf{p} = \sum_{i=1}^{n} c_i\mathbf{u}_i \tag{3}$$

where

$$c_i = \langle \mathbf{x}, \mathbf{u}_i\rangle \qquad \textit{for each } i \tag{4}$$

then $\mathbf{p} - \mathbf{x} \in S^{\perp}$ (see Figure 5.5.2).

then $\mathbf{p} - \mathbf{x} \in S^{\perp}$ (see Figure 5.5.2).

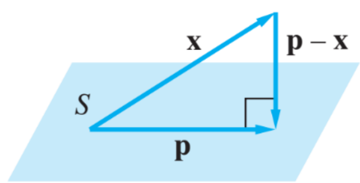


Figure 5.5.2.

proof: $a_j^T(b-p) = a_j^T b - a_j^T p = a_j^T b - a_j^T \sum a_i a_i$

$= a_j^T b - a_j^T a_j \alpha_j$

$= \alpha_j^T b - \alpha_j = 0$

Corollary 5.5.9 *Let S be a nonzero subspace of $\mathbb{R}^m$ and let $\mathbf{b} \in \mathbb{R}^m$. If $\{\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_k\}$ is an orthonormal basis for S and $U = (\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_k)$, then the projection $\mathbf{p}$ of $\mathbf{b}$ onto S is given by*

$$\mathbf{p} = UU^T\mathbf{b}$$

5.6 The Gram-Schmidt Orthogonalization Process

Ex: $V_1 = \begin{Bmatrix} 4 \\ 1 \\ 2 \\ 3 \end{Bmatrix}$ $V_2 = \begin{Bmatrix} -2 \\ 3 \\ 1 \\ 4 \end{Bmatrix}$

Find a orthonormal basis for span $\{V_1, V_2\}$

Let $u_1 = \frac{V_1}{\|V_1\|}$, Find $u_2 = \alpha_1 u_1 + \alpha_2 V_2$

such that $u_1^T u_2 = 0$ and $\|u_2\| = 1$

$u_1^T u_2 = u_1^T(\alpha_1 u_1 + \alpha_2 V_2) = 0$.

$= \alpha_1 u_1^T u_1 + \alpha_2 u_1^T v_2 = \alpha_1 + \alpha_2 u_1^T v_2 = 0$

$u_2 = -\alpha_2 u_1^T v_2 \cdot u_1 + \alpha_2 V_2$

$= \alpha_2 (V_2 - (u_1^T V_2) u_1)$

② Grass - Schmidt orthogonalization.

Given $\{v_1, v_2 \cdots V_k\}$ is a basis

(i) let $u_1 = \frac{v_1}{\|v_1\|}$

(ii) By theorem 5.5.7

$v_2 - p_1 \perp u_1$ (or span $\{u_1\}$)

$\Rightarrow u_2 = \frac{v_2 - p}{\|v_2 - p\|}$

(iii) $u_3 = v_3 - p_2$

p_j is the projection vector of v_{j+1} onto span $\{v_1 \cdots v_j\}$

Theorem 5.6.1 The Gram–Schmidt Process

Let $\{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n\}$ be a basis for the inner product space V. Let

$$\mathbf{u}_1 = \left(\frac{1}{\|\mathbf{x}_1\|}\right)\mathbf{x}_1$$

and define $\mathbf{u}_2, \ldots, \mathbf{u}_n$ recursively by

$$\mathbf{u}_{k+1} = \frac{1}{\|\mathbf{x}_{k+1} - \mathbf{p}_k\|}(\mathbf{x}_{k+1} - \mathbf{p}_k) \quad \text{for } k = 1, \ldots, n-1$$

where

$$\mathbf{p}_k = \langle \mathbf{x}_{k+1}, \mathbf{u}_1 \rangle \mathbf{u}_1 + \langle \mathbf{x}_{k+1}, \mathbf{u}_2 \rangle \mathbf{u}_2 + \cdots + \langle \mathbf{x}_{k+1}, \mathbf{u}_k \rangle \mathbf{u}_k$$

is the projection of $\mathbf{x}_{k+1}$ onto $\text{Span}(\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_k)$. *Then the set*

$$\{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_n\}$$

is an orthonormal basis for V.

Theorem 5.6.2 Gram–Schmidt QR Factorization

If A is an $m \times n$ matrix of rank n, then A can be factored into a product QR, where Q is an $m \times n$ matrix with orthonormal column vectors and R is an upper triangular $n \times n$ matrix whose diagonal entries are all positive. [Note: R must be nonsingular since $\det(R) > 0$.*]*

Given $\{a_1, a_2 \cdots a_n\}$ are Linearly ind.

Find an orthonormal basis $\{q_1, q_2 \cdots q_n\}$ as follows

(i) $q_1 = \frac{a_1}{\|a_1\|}$ (ii) $q_2 = \frac{a_2 - p_1}{\|a_2 - p_1\|}$ p_1 is the projection of a_2 onto $S_1 = \text{span}\{a_1\}$

$a_2 = p_1 + q_2$ $\quad \|a_2 - p_1\| q_2 + a_1^T q_1 q_2 = a_2$

(iii) $q_3 = a_3 - p_2$ $\quad p_2$: projection of a_3 onto $S_2 = \text{span}\{a_1, a_2\}$

$\|a_3 - p_2\| q_3 + a_3^T q_1 q_1 + a_3^T q_2 q_2 = a_3$

$r_{13}q_1 + r_{23}q_2 + r_{33}q_3 = a_3$

Generally : $q_k = \frac{a_k - p_{k-1}}{\|a_k - p_{k-1}\|} - r_{kk}$

$$r_{kk}q_k + \underbrace{r_{1k}q_1 + r_{2k}q_2 + \cdots r_{k-1}q_{k-1}}_{p_{k-1}} = a_k$$

$$(a_1\ a_2 \cdots a_n) = (q_1 \cdots q_n)\begin{pmatrix} r_{11} & r_{12} & \cdots & r_{1n} \\ 0 & r_{22} & \ddots & \vdots \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & & r_{nn} \end{pmatrix}$$

$$A = QR$$

③ QR-factorization

$A = QR$, for least square problems

$A^TA\hat{x} = A^Tb \quad (QR)^TQR\hat{x} = (QR)^Tb$

$\hat{x} = R^{-1}Q^Tb$

$R\hat{x} = Q^Tb$

④ modified Gram-Schmidt orthogonalization