Least - Square Approximation

Orthonormality

- $\overset{m}{\text{equations}} > \text{no. of } \overset{n}{\text{variables}} \Rightarrow$ over-determined system

  - $x = A^{-1}b$ can't give any solution, A^{-1} is not invertible as its not a square matrix

- $A = m \times n$ matrix

- solution (approx.) of over-determined system is L.S.A.

$$\underset{Q}{(m \times n)} \; \underset{x}{(n \times 1)} = \underset{b}{(m \times 1)}$$

- $Q = (q_1 \quad q_2 \quad \cdots \quad q_n)$ each $m \times 1$ matrix

Theorem :

Let Q be $m \times n$ matrix. The columns of Q form an orthonormal set iff $Q^T Q = I_n$

* **Orthogonality** :

Let $\bar{x}$ & $\bar{y}$ are 2 vectors (n×1 column matrix) in n-dimensional Euclidean vector space $\mathbb{R}^n$. $\bar{x}, \bar{y} \in \mathbb{R}^n$

Date : ..../......../........

$$\bar{x} \cdot \bar{y} = x^T y = \sum_{i=1}^{n} x_i y_i$$ (scalar / dot product in 3D)

(number)

- inner product with itself = l_2 norm = norm of vector

$$|\bar{x}| = \sqrt{x^T x}$$ = absolute value = magnitude in 3D

- For angle θ between any 2 vectors,

$$x^T y = |\bar{x}| |\bar{y}| \cos\theta$$

$$= xy \cos\theta$$

- $\bar{x}$ & $\bar{y}$ are orthogonal iff $\boxed{x^T y = 0}$ $\Rightarrow \theta = 90°$

(perpendicular to each other)

- a set of n vectors in $\mathbb{R}^n$, $S = \{x_1, x_2, \dots, x_n\}$. S is called orthogonal set if $x_i^T x_j = 0 \;\forall\; i, j = 1, 2, \dots, n$ with $i \neq j$ (any pair)

- $\boxed{x_i^T x_i = 1}$ for each $i = 1, 2, \dots, n$ $\Rightarrow$ vectors x_i has norm unity

(all unit vectors)

Date : ..../........./.........

* Orthonormal set :

The set of vectors in which every pair of vectors are orthogonal & each vector has norm one.

$$S = \{x_i \mid x_i \in \mathbb{R}^n ,\ x_i^T x_j = \delta_{ij} ,\ i,j = 1,2,\dots,n\}$$

↓

kronecker delta

$$\delta_{ij} = \begin{cases} 1 & i = j \text{ (norm-1)} \\ 0 & i \neq j \text{ (orthogonal)} \end{cases}$$

(?) orthonormal $S = \{ \frac{1}{\sqrt{5}} (2,1)^T ,\ \frac{1}{\sqrt{5}} (1,-2)^T \}$. prove it.

⇒ Let $u = \frac{1}{\sqrt{5}} (2,1)^T$, $v = \frac{1}{\sqrt{5}} (1,-2)^T$

$$\bar{u}.\bar{u} = u^T u = \frac{1}{5} \cancel{2} (2,1) \begin{pmatrix} 2 \\ 1 \end{pmatrix} = \frac{1}{5} (2^2 + 1^2) = 1$$

$$\bar{v}.\bar{v} = v^T v = \frac{1}{5} (1^2 + (-2)^2) = 1$$

$$\bar{u}.\bar{v} = u^T v = \frac{1}{5} (2 \ \ 1) \begin{pmatrix} 1 \\ -2 \end{pmatrix} = \frac{1}{5} (2 \times 1 + 1 \times (-2)) = 0$$

∴ S is orthonormal set in $\mathbb{R}^2$ and hence form a basis in 2D vector space.

Date : ....../....../......

Polynomial data fitting

- Normal equation :

$$\underbrace{(A^TA)}_{\substack{n\times n \text{ matrix} \\ \text{(square)}}} x = \underbrace{A^Tb}_{n\times 1 \text{ matrix}} \quad \left([n\times m][m\times 1] = [n\times 1]\right)$$

- If $\det(A^TA) \neq 0$, A^TA is invertible

Theorem :

The matrix A^TA is invertible iff the columns of A are linearly independent, in which case $Ax = b$ has a <u>unique</u> least-squares solution $\boxed{x = (A^TA)^{-1} A^T b}$

$$P_n(x_0) = a_0 + a_1x_0 + \dots + a_nx_0^n = f(x_0)$$

$$P_n(x_1) = a_0 + a_1x_1 + \dots + a_nx_1^n = f(x_1)$$

$$\vdots \qquad\qquad (m > n)$$

$$P_n(x_m) = a_0 + a_1x_m + \dots + a_nx_m^n = f(x_m)$$

$$\Downarrow$$

$$\underbrace{\begin{bmatrix} 1 & x_0 & \dots & x_0^n \\ 1 & x_1 & \dots & x_1^n \\ \vdots & \vdots & & \\ 1 & x_m & \dots & x_m^n \end{bmatrix}}_{(m\times n)} \underbrace{\begin{bmatrix} a_0 \\ a_1 \\ \vdots \\ a_n \end{bmatrix}}_{(n\times 1)} = \underbrace{\begin{bmatrix} f(x_0) \\ f(x_1) \\ \vdots \\ f(x_m) \end{bmatrix}}_{(m\times 1)}$$

Date : ..../........../..........

(?) Fit a L.S. straight line to the data : $f(-3) = f(0) = 0$ and $f(6) = 2$

$\Rightarrow$ straight line, so $n = 1$

$\therefore P_1(x) = a_0 + a_1 x$

3 nodes, $x_0 = -3$, $x_1 = 0$, $x_2 = 6$

$\therefore m = 2$

$$\therefore P_1(x_0) = a_0 + a_1 x_0 = f(x_0)$$

$$P_1(x_1) = a_0 + a_1 x_1 = f(x_1)$$

$$P_1(x_2) = a_0 + a_1 x_2 = f(x_2)$$

$$\Downarrow$$

$$\underbrace{\begin{bmatrix} 1 & x_0 \\ 1 & x_1 \\ 1 & x_2 \end{bmatrix}}_{A} \underbrace{\begin{bmatrix} a_0 \\ a_1 \end{bmatrix}}_{x} = \underbrace{\begin{bmatrix} f(x_0) \\ f(x_1) \\ f(x_2) \end{bmatrix}}_{b}$$

$$\Rightarrow \begin{bmatrix} 1 & -3 \\ 1 & 0 \\ 1 & 6 \end{bmatrix} \begin{bmatrix} a_0 \\ a_1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 2 \end{bmatrix}$$

$$\Rightarrow \begin{bmatrix} 1 & 1 & 1 \\ -3 & 0 & 6 \end{bmatrix} \begin{bmatrix} 1 & -3 \\ 1 & 0 \\ 1 & 6 \end{bmatrix} \begin{bmatrix} a_0 \\ a_1 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 \\ -3 & 0 & 6 \end{bmatrix} \begin{bmatrix} 0 \\ 0 \\ 2 \end{bmatrix}$$

Date : ..../........../..........

$$\Rightarrow \begin{bmatrix} 3 & 3 \\ 3 & 45 \end{bmatrix} \begin{bmatrix} a_0 \\ a_1 \end{bmatrix} = \begin{bmatrix} 2 \\ 12 \end{bmatrix}$$

$A^T A$ $\qquad\qquad$ $A^T b$

$$\det(A^T A) = 135 - 9 = 126 \neq 0$$

$$\therefore x = (A^T A)^{-1} A^T b$$

$$= \frac{1}{126} \begin{bmatrix} 45 & -3 \\ -3 & 3 \end{bmatrix} \begin{bmatrix} 2 \\ 12 \end{bmatrix}$$

$$\Rightarrow \begin{bmatrix} a_0 \\ a_1 \end{bmatrix} = \begin{bmatrix} 3/7 \\ 5/21 \end{bmatrix}$$

$$\therefore P_1(x) = \underbrace{3/7}_{\text{intecept}} + \underbrace{5/21}_{\text{slope}}\, x$$

Date : ..../........../..........

QR Decomposition

Theorem :

Any ~~~~ real ~~~~ $m \times n$ matrix A, with $m \geq n$, can be written in the form $A = QR$, where Q is an $m \times n$ matrix with orthonormal columns & R is an upper triangular $n \times n$ matrix.

$\therefore A_{m \times n} = Q_{m \times n} R_{n \times n}$ where $Q = \{q_1\ q_1\ q_2 \ldots q_n\}$

* Gram-Schmidt orthogonalization process :

Mathematical process to obtain an orthonormal basis or set from a set of linearly independent vectors.

(m×1 column matrix)

$$A = \begin{array}{c} \begin{array}{cccc} v_1 & v_2 & & U_n \end{array} \\ \begin{bmatrix} u_{11} & U_{12} & \cdots & U_{1n} \\ U_{21} & U_{22} & \cdots & U_{1n} \\ \vdots & \vdots & & \vdots \\ U_{m1} & U_{m2} & \cdots & U_{mn} \end{bmatrix} \end{array}$$

; U_i for $i = 1, 2, \ldots, n$

- linearly independent

$$U_i = (U_{1i}\quad U_{2i}\quad \cdots\quad U_{mi})^T$$

$=$ ~~•~~ $m \times 1$ column matrix

Date : ..../......./.......

- Let $\{q_1, q_2, \dots, q_n\}$ be the set of orthonormal vectors constructed ~~form~~ from set $A = \{u_1, u_2, \dots, u_n\}$ which is a set of n vectors in $\mathbb{R}^m$.

Gram - Schmidth process yields that,

$$p_k = u_k - \sum_{i=1}^{k-1} (u_k^T q_i)\, q_i \quad \&$$

$$q_k = \frac{p_k}{|p_k|}$$

where $k = 1, 2, \dots, n$ & norm $|p_k| = \sqrt{(p_k^T \cdot p_k)}$

- each p_k is constructed from u_k by subtracting projections of u_k on each of previous q_i for $i<k$

(?) $u_1 = (3, 6, 0)^T$, $u_2 = (1, 2, 2)^T$. Find orthonormal vectors.

$\Rightarrow$

$$p_1 = u_1 = \begin{bmatrix} 3 \\ 6 \\ 0 \end{bmatrix}, \quad |p_1| = \sqrt{p_1^T p}$$

$k = 1$

$$= \sqrt{3^2 + 6^2 + 0} = \sqrt{45}$$

$i = 0$

$$\therefore q_1 = \frac{1}{\sqrt{45}} \begin{bmatrix} 3 \\ 6 \\ 0 \end{bmatrix}$$

Date : ..../......../........

$k = 2$

$i = 1$

$$p_2 = u_2 - (u_2^T q_1) q_1$$

$$= \begin{bmatrix} 1 \\ 2 \\ 2 \end{bmatrix} - \left(\begin{bmatrix} 1 & 2 & 2 \end{bmatrix} \begin{bmatrix} 3 \\ 6 \\ 0 \end{bmatrix} \frac{1}{\sqrt{45}} \right) \frac{1}{\sqrt{45}} \begin{bmatrix} 3 \\ 6 \\ 0 \end{bmatrix}$$

$$= \begin{bmatrix} 1 \\ 2 \\ 2 \end{bmatrix} - \frac{15}{\sqrt{45}} \cdot \frac{1}{\sqrt{45}} \begin{bmatrix} 3 \\ 6 \\ 0 \end{bmatrix}$$

$$= \begin{bmatrix} 1 \\ 2 \\ 2 \end{bmatrix} - \frac{1}{3} \begin{bmatrix} 3 \\ 6 \\ 0 \end{bmatrix}$$

$$= \begin{bmatrix} 1 \\ 2 \\ 2 \end{bmatrix} - \begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 2 \end{bmatrix} \quad ; \quad |p_2| = \sqrt{2^2} = 2$$

$$\therefore \quad q_2 = \frac{1}{2} \begin{bmatrix} 0 \\ 0 \\ 2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

Date : ...... / ...... / ......

$$u_k = |P_k| q_k + \sum_{i=1}^{k-1} (u_k^T q_i) q_i$$

$$\Rightarrow u_k \cdot q_k = |P_k| \underbrace{q_k \cdot q_k}_{1} + \sum_{i=1}^{k-1} (u_k^T q_i) \underbrace{q_i \cdot q_k}_{0}$$

$$= |P_k|$$

$$\therefore u_k = (u_k \cdot q_k) q_k + \sum_{i=1}^{k-1} (u_k^T q_i) q_i$$

$$= \sum_{i=1}^{k} (u_k^T q_i) q_i \quad ; k = 1, 2, \ldots, m$$

$$u_1 = (q_1)(u_1^T q_1)$$

$$u_2 = (q_1 \quad q_2) \begin{pmatrix} u_2^T q_1 \\ u_2^T q_2 \end{pmatrix}$$

$$u_3 = (q_1 \quad q_2 \quad q_3) \begin{pmatrix} u_3^T q_1 \\ u_3^T q_2 \\ u_3^T q_3 \end{pmatrix}$$

Up to 3rd term,

$$(u_1 \quad u_2 \quad u_3) = (q_1 \quad q_2 \quad q_3) \begin{pmatrix} u_1^T q_1 & u_2^T q_1 & u_3^T q_1 \\ 0 & u_2^T q_2 & u_3^T q_2 \\ 0 & 0 & u_3^T q_3 \end{pmatrix}$$

Date : ....../....../......

Up to nth term,

$$\left(u_1 \mid u_2 \mid \cdots \mid u_n\right) = \left(q_1 \mid q_2 \mid \cdots \mid q_n\right) \begin{pmatrix} u_1^T q_1 & u_2^T q_1 & \cdots & u_n^T q_1 \\ 0 & u_2^T q_2 & \cdots & u_n^T q_2 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & u_n^T q_n \end{pmatrix}$$

$A_{m\times n}$ $\qquad$ $Q_{m\times n}$ $\qquad$ $R_{n\times n}$

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

$$\Rightarrow (QR)^T (QR) x = (QR)^T b$$

$$\Rightarrow R^T (Q^T Q) R x = R^T Q^T b$$

$$\Rightarrow Rx = Q^T b \qquad \boxed{Q^T Q = 1}$$

$$\Rightarrow x = R^{-1} Q^T b$$

Date : ..../......./.......

(?) Least - square straight line fitting to the data : $f(-3) = f(0) = 0$; $f(6) = 2$

$\Rightarrow$

$n = 1$ (straight line)

$m = \cancel{3}\,2$ (3 data point)

$P_1 = a_0 + a_1 x$

$$\begin{bmatrix} 1 & x_0 \\ 1 & x_1 \\ 1 & x_2 \end{bmatrix} \begin{bmatrix} a_0 \\ a_1 \end{bmatrix} = \begin{bmatrix} f(x_0) \\ f(x_1) \\ f(x_2) \end{bmatrix}$$

$$= \underbrace{\begin{bmatrix} 1 & -3 \\ 1 & 0 \\ 1 & 6 \end{bmatrix}}_{A} \underbrace{\begin{bmatrix} a_0 \\ a_1 \end{bmatrix}}_{x} = \underbrace{\begin{bmatrix} 0 \\ 0 \\ 2 \end{bmatrix}}_{b}$$

linearly independent columns of A,

$$u_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \quad u_2 = \begin{bmatrix} -3 \\ 0 \\ 6 \end{bmatrix}$$

Using Gram - Schmidt,

$$P_1 = u_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \qquad \therefore |P_1| = \sqrt{P_1^T P_1}$$

$$= \sqrt{1^2 + 1^2 + 1^2} = \sqrt{3}$$

$$\therefore q_1 = \frac{1}{\sqrt{3}} \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$$

Date : ______/______/______

$P_2 = u_2 - (u_2^T q_1) q_1$

$$= \begin{bmatrix} -3 \\ 0 \\ 6 \end{bmatrix} - \left(\begin{bmatrix} -3 & 0 & 6 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \frac{1}{\sqrt{3}} \right) \frac{1}{\sqrt{3}} \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$$

$$= \begin{bmatrix} -3 \\ 0 \\ 6 \end{bmatrix} - \frac{3}{\sqrt{3}} \cdot \frac{1}{\sqrt{3}} \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$$

$$= \begin{bmatrix} -4 \\ -1 \\ 5 \end{bmatrix}$$

$\therefore |P_2| = \sqrt{P_2^T P} = \sqrt{4^2 + 1^2 + 5^2} = \sqrt{42}$

$$\therefore q_2 = \frac{1}{\sqrt{42}} \begin{bmatrix} -4 \\ -1 \\ 5 \end{bmatrix}$$

$$\therefore Q = \begin{bmatrix} 1/\sqrt{3} & -4/\sqrt{42} \\ 1/\sqrt{3} & -1/\sqrt{42} \\ 1/\sqrt{3} & 5/\sqrt{42} \end{bmatrix}$$

$$R = \begin{bmatrix} u_1^T q_1 & u_2^T q_1 \\ 0 & u_2^T q_2 \end{bmatrix} = \begin{bmatrix} \sqrt{3} & \sqrt{3} \\ 0 & \sqrt{42} \end{bmatrix}$$

Date :........../........../..........

$\therefore$ $Rx = Q^T b$

$$\Rightarrow \begin{bmatrix} \sqrt{3} & \sqrt{3} \\ 0 & \sqrt{42} \end{bmatrix} \begin{bmatrix} a_0 \\ a_1 \end{bmatrix} = \begin{bmatrix} 1/\sqrt{3} & 1/\sqrt{3} & 1/\sqrt{3} \\ -4/\sqrt{42} & -1/\sqrt{42} & 5/\sqrt{42} \end{bmatrix} \begin{bmatrix} 0 \\ 0 \\ 2 \end{bmatrix}$$

$$\Rightarrow \begin{bmatrix} \sqrt{3}a_0 + \sqrt{3}a_1 \\ \sqrt{42}a_1 \end{bmatrix} = \begin{bmatrix} 2/\sqrt{3} \\ 10/\sqrt{42} \end{bmatrix}$$

$$\therefore a_1 = 10/\sqrt{42} \times \frac{1}{\sqrt{42}} = \frac{10}{42} = \frac{5}{21}$$

$$\therefore \sqrt{3}a_0 = -\sqrt{3} \cdot \frac{5}{21} + 2/\sqrt{3}$$

$$\Rightarrow a_0 = \cancel{-\frac{5\sqrt{3}}{21}} \cancel{\times \frac{1}{\sqrt{3}}} \frac{3\sqrt{3}}{7} \times \frac{1}{\sqrt{3}} = \frac{3}{7}$$

$$\therefore \begin{bmatrix} a_0 \\ a_1 \end{bmatrix} = \begin{bmatrix} 3/7 \\ 5/21 \end{bmatrix}$$