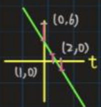
8

4月28日 星期三 上午10:16

Least square solution

- $A\underline{x} \neq \underline{b}$ $\quad \underline{b} - A\underline{x} = \underline{e}$ The length of $\underline{e}$ or $\|\underline{e}\|^2$ 越小越好

  ex: Find the closest line to the points

  $b = C + Dt$

  $$\begin{cases} C + D\cdot 0 = 6 \\ C + D\cdot 1 = 0 \\ C + D\cdot 2 = 0 \end{cases} \Rightarrow A\underline{x} = \underline{b},\ A = \begin{bmatrix} 1 & 0 \\ 1 & 1 \\ 1 & 2 \end{bmatrix},\ \underline{x} = \begin{bmatrix} C \\ D \end{bmatrix},\ \underline{b} = \begin{bmatrix} 6 \\ 0 \\ 0 \end{bmatrix} \Rightarrow A\underline{x} = \underline{b} \text{ 無解}$$

  The best fit is $\hat{\underline{x}}$ such that $\underline{p} = A\hat{\underline{x}}$ is the projection of $\underline{b}$ onto $C(A)$.

  $$A^T(\underline{b} - A\hat{\underline{x}}) = \underline{0} \Rightarrow A^TA\hat{\underline{x}} = A^T\underline{b} \Rightarrow \begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 2 \end{bmatrix}\begin{bmatrix} 1 & 0 \\ 1 & 1 \\ 1 & 2 \end{bmatrix}\begin{bmatrix} C \\ D \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 2 \end{bmatrix}\begin{bmatrix} 6 \\ 0 \\ 0 \end{bmatrix} \Rightarrow \begin{bmatrix} 3 & 3 \\ 3 & 5 \end{bmatrix}\begin{bmatrix} C \\ D \end{bmatrix} = \begin{bmatrix} 6 \\ 0 \end{bmatrix} \Rightarrow \hat{\underline{x}} = \begin{bmatrix} C \\ D \end{bmatrix} = \begin{bmatrix} 5 \\ -3 \end{bmatrix}$$

  Best line : $b = 5 - 3t$

- By linear algebra, $\underline{b} = \underset{C(A)}{\underset{\uparrow}{\underline{p}}} + \underset{N(A^T)}{\underset{\uparrow}{\underline{e}}}$ We know that $A\underline{x} = \underline{b} = \underline{p} + \underline{e}$ 無解 and $A\hat{\underline{x}} = \underline{p}$ 有解

  $$\|A\underline{x} - \underline{b}\|^2 = \|A\underline{x} - \underline{p} - \underline{e}\|^2 = (A\underline{x} - \underline{p} - \underline{e})^T(A\underline{x} - \underline{p} - \underline{e}) = \|A\underline{x} - \underline{p}\|^2 + \|\underline{e}\|^2$$

  $\left\langle (A\underline{x} - \underline{p})^T\underline{e} = 0 \text{ since } A\underline{x} - \underline{p} \in C(A) \text{ and } \underline{e} \in N(A^T) \right\rangle \Rightarrow \|A\underline{x} - \underline{b}\|^2$ is minimized when $x = \hat{x}$ $(A\hat{\underline{x}} = \underline{p})$

  i.e. $\min_{\underline{x}} \|A\underline{x} - \underline{b}\|^2 = \|A\hat{\underline{x}} - \underline{b}\|^2 = \|A\hat{\underline{x}} - \underline{p}\|^2 + \|\underline{e}\|^2 = \|\underline{e}\|^2$ where $A^TA\hat{\underline{x}} = A^T\underline{b}$ and $\underline{p} = A\hat{\underline{x}}$

- By calculus, $E = \|A\underline{x} - \underline{b}\|^2 = (C + D\cdot 0 - 6)^2 + (C + D\cdot 1)^2 + (C + D\cdot 2)^2$

  $\frac{\partial E}{\partial C} = 2(C + D\cdot 0 - 6) + 2(C + D\cdot 1) + 2(C + D\cdot 2) = 0$

  $\frac{\partial E}{\partial D} = 2(C + D\cdot 1) + 4(C + D\cdot 2) = 0$

  $\begin{cases} 6C + 6D = 12 \\ 6C + 10D = 0 \end{cases} \Rightarrow D = -3,\ C = 5 \Rightarrow b = 5 - 3t$

  The partial derivatives of $\|A\underline{x} - \underline{b}\|^2$ are 0 when $A^TA\hat{\underline{x}} = A^T\underline{b}$, $\min_{\underline{x}} \|A\underline{x} - \underline{b}\|^2 = \|\underline{e}\|^2$

- When $A\underline{x} = \underline{b}$ 無解, the least square solution $\hat{\underline{x}}$ satisfies $A^TA\hat{\underline{x}} = A^T\underline{b}$

Curve Fitting

- Fitting a Straight Line : Fit height b_i at t_i by $C+Dt$

  solution: $\begin{matrix} C+Dt_1=b_1 \\ \vdots \\ C+Dt_m=b_m \end{matrix} \Rightarrow \begin{bmatrix} 1 & t_1 \\ \vdots & \vdots \\ 1 & t_m \end{bmatrix}\begin{bmatrix} C \\ D \end{bmatrix} = \begin{bmatrix} b_1 \\ \vdots \\ b_m \end{bmatrix}$ ($A\underline{x}=\underline{b}$ 無解)

  best fit: $A^TA = \begin{bmatrix} 1 & \cdots & 1 \\ t_1 & \cdots & t_m \end{bmatrix}\begin{bmatrix} 1 & t_1 \\ \vdots & \vdots \\ 1 & t_m \end{bmatrix} = \begin{bmatrix} m & \sum_i t_i \\ \sum_i t_i & \sum_i t_i^2 \end{bmatrix}$ $\quad A^T\underline{b} = \begin{bmatrix} 1 & \cdots & 1 \\ t_1 & \cdots & t_m \end{bmatrix}\begin{bmatrix} b_1 \\ \vdots \\ b_m \end{bmatrix} = \begin{bmatrix} \sum_i b_i \\ \sum_i t_i b_i \end{bmatrix}$

  Solve $\begin{bmatrix} m & \sum_i t_i \\ \sum_i t_i & \sum_i t_i^2 \end{bmatrix}\begin{bmatrix} C \\ D \end{bmatrix} = \begin{bmatrix} \sum_i b_i \\ \sum_i t_i b_i \end{bmatrix}$

  The best $\hat{\underline{x}} = (C, D)$ minimizes $\|A\underline{x}-\underline{b}\|^2 = \|\underline{e}\|^2 = \sum_i^m (C+Dt_i-b_i)^2$

- Fitting a Parabola : Fit heights b_i at t_i by $C+Dt+Et^2$

  solution: $\begin{matrix} C+Dt_1+Et_1^2=b_1 \\ \vdots \\ C+Dt_m+Et_m^2=b_m \end{matrix} \Rightarrow \begin{bmatrix} 1 & t_1 & t_1^2 \\ & \vdots & \\ 1 & t_m & t_m^2 \end{bmatrix}\begin{bmatrix} C \\ D \\ E \end{bmatrix} = \begin{bmatrix} b_1 \\ \vdots \\ b_m \end{bmatrix}$ ($A\underline{x}=\underline{b}$ 無解) $\Rightarrow$ Solve $A^TA\hat{\underline{x}} = A^T\underline{b}$ for $\hat{\underline{x}}$

Orthogonal Bases and Gram-Schmidt

Def: The vectors $\underline{q}_1 \cdots \underline{q}_n$ are orthogonal if $\underline{q}_i^T\underline{q}_j = 0$ whenever $i \neq j$

Claim: If nonzero vector $\underline{q}_1 \cdots \underline{q}_n$ are orthogonal, then they are L.I.

pf: Consider $x_1\underline{q}_1 + \cdots + x_n\underline{q}_n = \underline{0}$, then $x_1\underline{q}_1^T\underline{q}_1 + \cdots + x_n\underline{q}_1^T\underline{q}_n = \underline{q}_1^T\underline{0} = 0 \Rightarrow x_1\|\underline{q}_1\|^2 = 0$

$\Rightarrow x_1 = 0$, similarly, $x_2 = \cdots = x_n = 0 \Rightarrow \underline{q}_1 \cdots \underline{q}_n$ are L.I.

Def: The vectors $\underline{q}_1 \cdots \underline{q}_n$ are orthonormal if $\underline{q}_i^T\underline{q}_j = \delta_{ij}$

- A matrix Q with orthonormal columns satisfies $Q^TQ = \underbrace{\begin{bmatrix} \underline{q}_1^T \\ \vdots \\ \underline{q}_n^T \end{bmatrix}}_{n\times m}\underbrace{[\underline{q}_1 \cdots \underline{q}_n]}_{m\times n} = \underbrace{I}_{n\times n}$

  When Q is 方陣, $Q^TQ = I \Rightarrow Q^T = Q^{-1}$. In this case, Q is called an orthogonal matrix.

  ex:

  Rotation: $Q = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$ $\quad$ Reflection: $Q = I - 2\underline{u}\,\underline{u}^T$, $\underline{u}$: any unit vector