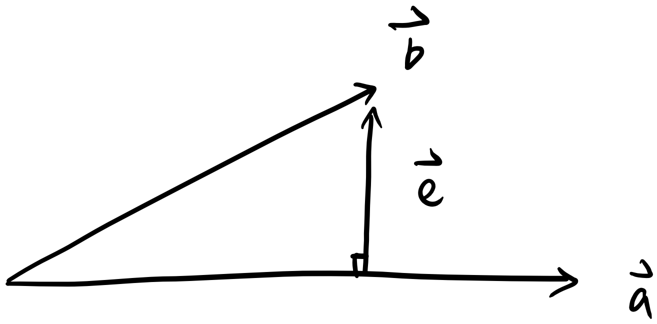
MIT : Gilbert strang



$proj(\vec{b}) = P_b$ project $\vec{b}$ onto $\vec{a}$

$P_b = x\vec{a}$

x ↓ multiple

$\vec{e} = \vec{b} - x\vec{a}$

求 x

$\vec{e} \cdot \vec{a} = 0 \Rightarrow \vec{b} \cdot \vec{a} - x\vec{a} \cdot \vec{a} = 0$

$\Downarrow$

$$x = \frac{a^T b}{a^T a}$$

↓ generalize to higher dimensional case

E.g. 3-dim

P_b project b to the plane $(\vec{a}_1, \vec{a}_2)$

$$\vec{e} = \vec{b} - x_1 \vec{a}_1 - x_2 \vec{a}_2$$

$$\vec{e} \cdot \vec{a}_1 = 0$$

$$\vec{e} \cdot \vec{a}_2 = 0$$

$$\begin{cases} \vec{b} \cdot \vec{a}_1 = x_1 \vec{a}_1 \cdot \vec{a}_1 + x_2 \vec{a}_2 \cdot \vec{a}_1 \\ \vec{b} \cdot \vec{a}_2 = x_1 \vec{a}_1 \cdot \vec{a}_2 + x_2 \vec{a}_2 \cdot \vec{a}_2 \end{cases}$$

$\downarrow$

$$a^T b = \begin{pmatrix} \vec{a}_1 \cdot \vec{a}_1 & \vec{a}_2 \cdot \vec{a}_1 \\ \vec{a}_1 \cdot \vec{a}_2 & \vec{a}_2 \cdot \vec{a}_2 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$$

$$A = \begin{pmatrix} | & | \\ a_1 & a_2 \\ | & | \end{pmatrix}$$

$$a^T a = \begin{pmatrix} - a_1^T - \\ - a_2^T - \end{pmatrix} \begin{pmatrix} | & | \\ a_1 & a_2 \\ | & | \end{pmatrix}$$

a_1, a_2 是三分量矢量 here.

$$X = (a^T a)^{-1} a^T b$$

$A = a$, here

Q:

$p = A\hat{x}$

find $\hat{x}$

$b - A\hat{x}$ is perpendicular to the plane $(\vec{a}_1, \vec{a}_2)$

$$a_1^T (b - A\hat{x}) = 0$$

$$a_2^T (b - A\hat{x}) = 0$$

column vector ↗

$$\begin{pmatrix} a_1^T \\ a_2^T \end{pmatrix} (b - A\hat{x}) = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

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

$e = b - A\hat{x}$ in $N(A^T)$

$e \perp C(A)$ yes!

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

general case

$A_{n \times n}$

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

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

in one-dim

$$P_b = ax = a\frac{a^T b}{a^T a} = P b$$

projection operator on b

to give P_b

$$A(A^TA)^{-1}A^T$$

project onto col (A) space

Q: can we reduce it to

$$AA^T (A^{T^{-1}}) A^T = I \ ?$$

No ! A is not square

Here, col(A) is not the whole space.

e.g. (n-1) vectors q_i $(i=1,2,\dots,n-1)$ with n components (n-dim)

proj b to the whole space : $P = I$

$$P = A(A^TA)^{-1}A^T$$

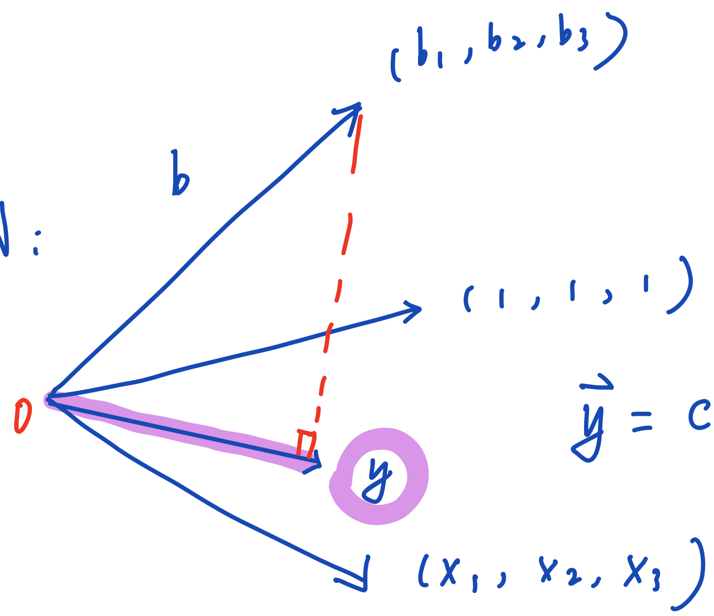
$$P = P^T$$

$$P^2 = P$$

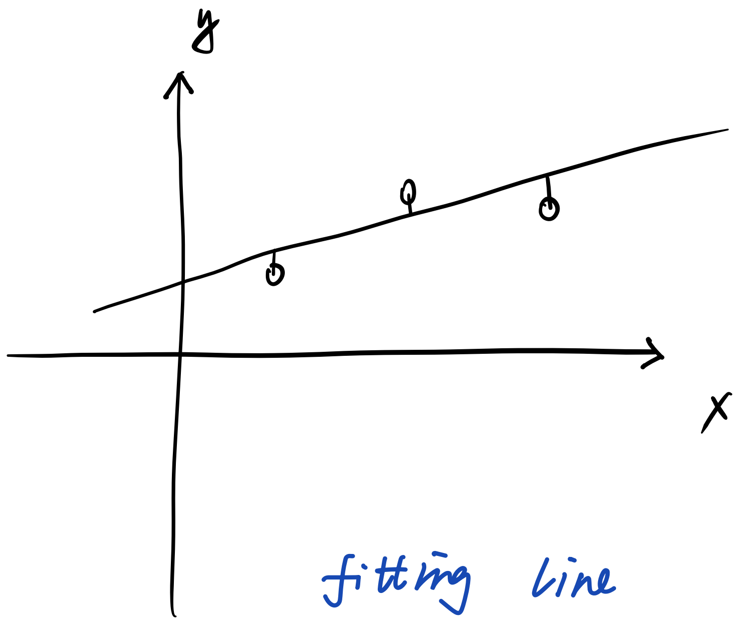
proj once, into col(A)

proj again, still there

几何直观图：



$\vec{y} = C\vec{x} + D$



$y = Cx + D$

已知 (x_1, b_1), (x_2, b_2), (x_3, b_3)

$\min |\vec{b} - \vec{y}|$

is equivalent to the projection of b

onto y

Solve C, D

$$C\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} + D\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix}$$

$$\begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} = \begin{pmatrix} x_1 & 1 \\ x_2 & 1 \\ x_3 & 1 \end{pmatrix}\begin{pmatrix} C \\ D \end{pmatrix} = A(A^TA)^{-1}A^Tb$$

$$\begin{pmatrix} C \\ D \end{pmatrix} = \left[\begin{pmatrix} x_1 & x_2 & x_3 \\ 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} x_1 & 1 \\ x_2 & 1 \\ x_3 & 1 \end{pmatrix} \right]^{-1}$$

$$\begin{pmatrix} x_1 & x_2 & x_3 \\ 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix}$$

最终的距离即投影出的 error

$$\vec{e} = \vec{b} - \vec{y} = \vec{b} - C\vec{x} - D\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$$

Q：如果已知 4 个点，$\{(x_1, b_1), (x_2, b_2), (x_3, b_3), (x_4, b_4)\}$

对上面的表达式作出怎样的修改？

A：
$$\begin{pmatrix} x_1 & 1 \\ x_2 & 1 \\ x_3 & 1 \\ x_4 & 1 \end{pmatrix}$$

四维空间，两个 basis，再好的拟合也有限，应该如何处理？

Q： _ dim to _ dim is 4-point case

E.g. a : column vector (n-component)

$$a a^T = A_{n \times n} \quad \text{projection matrix}$$

$$\begin{pmatrix} | \\ a \\ | \end{pmatrix} \begin{pmatrix} — & a^T & — \end{pmatrix}$$

rank 1, A still in a

col(A) is a multiple of a

e.g.

$$\begin{pmatrix} 1 \\ 2 \\ 4 \\ 5 \end{pmatrix} \begin{pmatrix} 1 & 2 & 4 & 5 \end{pmatrix}$$

$1 \quad 2 \quad 4 \quad 5$ — multiple of a

$$= \begin{pmatrix} 1 & 2 & 4 & 5 \\ 2 & 4 & 8 & 10 \\ 4 & 8 & 16 & 20 \\ 5 & 10 & 20 & 25 \end{pmatrix}$$

$$\begin{pmatrix} | \\ a \\ | \end{pmatrix} \underbrace{\begin{pmatrix} — a^T — \end{pmatrix} X}_{c \text{ number}}$$

↓ c number

↓ multiple of a

projection

$$T: \quad R^2 \to R$$

$$T(v+w) = T(v) + T(w)$$

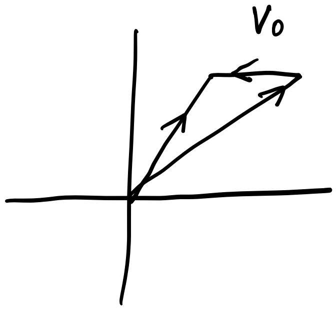
$$T(cv) = cT(v)$$

linear transformation

$$T(cv + dw) = cT(v) + dT(w)$$

Eg. 2

shift whole plane by v_0



T does to v

$\downarrow$

$$T(v) = v_0 + v$$

T not a linear trans

non - example

Another perspective :

e.g. $W = I$, prove the equivalence of the variation and equation

$$\min \| y - \hat{y} \|$$

$$\min \underline{\sum_i (y_i - (\beta_0 + \beta_1 x_i))^2} = f$$

$$\frac{\partial f}{\partial \beta_0} = -2 \sum_i [y_i - (\beta_0 + \beta_1 x_i)] = 0$$

$$\frac{\partial f}{\partial \beta_1} = -2 \sum_i x_i [y_i - (\beta_0 + \beta_1 x_i)] = 0$$

$$\sum_i y_i = \sum_i (\beta_0 + \beta_1 x_i)$$

$$\sum_i x_i y_i = \sum_i x_i (\beta_0 + \beta_1 x_i)$$

$$\Downarrow$$

$$A^T y = A^T A \begin{pmatrix} \beta_0 \\ \beta_1 \end{pmatrix}$$

举例：已知四个点 $(x_1, y_1), \cdots (x_4, y_4)$

$$A = \begin{pmatrix} 1 & x_1 \\ 1 & x_2 \\ 1 & x_3 \\ 1 & x_4 \end{pmatrix}$$

$$\begin{pmatrix} 1 & 1 & 1 & 1 \\ x_1 & x_2 & x_3 & x_4 \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \end{pmatrix} = \begin{pmatrix} \sum_i y_i \\ \sum_i x_i y_i \end{pmatrix}$$

$$\begin{pmatrix} 1 & 1 & 1 & 1 \\ x_1 & x_2 & x_3 & x_4 \end{pmatrix} \begin{pmatrix} 1 & x_1 \\ 1 & x_2 \\ 1 & x_3 \\ 1 & x_4 \end{pmatrix} \begin{pmatrix} \beta_0 \\ \beta_1 \end{pmatrix} =$$

$$\begin{pmatrix} 4 & \sum_i x_i \\ \sum_i x_i & \sum_i x_i^2 \end{pmatrix} \begin{pmatrix} \beta_0 \\ \beta_1 \end{pmatrix} = \begin{pmatrix} 4\beta_0 + \sum_i x_i \beta_1 \\ \sum_i x_i \beta_0 + \sum_i x_i^2 \beta_1 \end{pmatrix}$$

$$\begin{pmatrix} \beta_0 \\ \beta_1 \end{pmatrix} = (A^T A)^{-1} A^T y$$