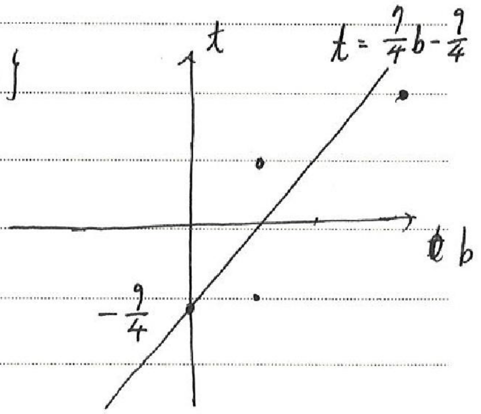
$$\text{Min } E^2 = \|Ax - b\|^2 = \sum_{i=1}^{m} (C + Dt_i - b_i)^2$$

$$\Rightarrow A^TA\hat{x} = A^Tb. \Rightarrow \begin{bmatrix} m & \Sigma t_i \\ \Sigma t_i & \Sigma t_i^2 \end{bmatrix} \begin{bmatrix} \hat{C} \\ \hat{D} \end{bmatrix} = \begin{bmatrix} \Sigma b_i \\ \Sigma b_i t_i \end{bmatrix}$$

ex) 3 samples. $(b, t) = \{(1,-1), (1,1), (3,2)\}$



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

$$\Rightarrow \begin{bmatrix} 3 & 2 \\ 2 & 6 \end{bmatrix} \begin{bmatrix} \hat{C} \\ \hat{D} \end{bmatrix} = \begin{bmatrix} 5 \\ 6 \end{bmatrix}$$

$$\begin{bmatrix} \hat{C} \\ \hat{D} \end{bmatrix} = \begin{bmatrix} 3 & 2 \\ 2 & 6 \end{bmatrix}^{-1} \begin{bmatrix} 5 \\ 6 \end{bmatrix} = \begin{bmatrix} \frac{9}{7} \\ \frac{4}{7} \end{bmatrix}$$

$\frac{9}{7} + \frac{4}{7}t = b$ → best line.

$$e = b - Pb = \begin{bmatrix} 1 \\ 1 \\ 3 \end{bmatrix} - \underbrace{A(A^TA)^{-1}A^T}_{Pb} \begin{bmatrix} 5 \\ 6 \end{bmatrix} = \begin{bmatrix} \frac{2}{7} \\ -\frac{6}{7} \\ \frac{4}{7} \end{bmatrix}$$

$$e \perp C(A) \rightarrow \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}^T \begin{bmatrix} \frac{2}{7} \\ -\frac{6}{7} \\ \frac{4}{7} \end{bmatrix} = 0, \quad [-1, 1, 2] \begin{bmatrix} \frac{2}{7} \\ -\frac{6}{7} \\ \frac{4}{7} \end{bmatrix} = 0.$$

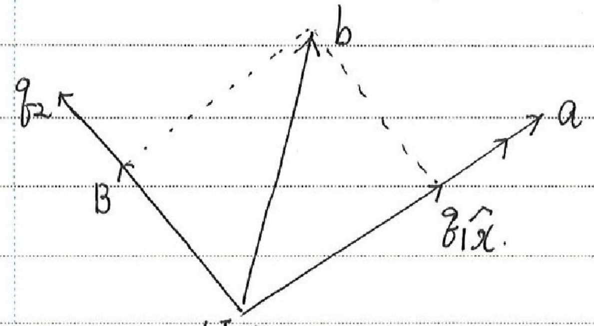
⊙ Weighted Least Squares.

- there are weights (probability) of observed samples.

  $x_i \rightarrow W_i$

⊙ Gram - Schmidt Orthogonalization.

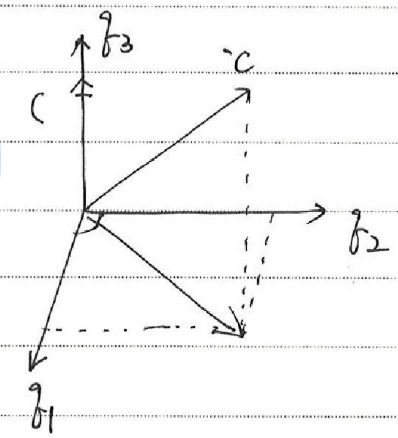
→ find the orthonormal basis given independent vectors.



$$q_1 = \frac{a}{\|a\|}$$

$$B = b - \frac{b^T a}{a^T a} a = b - (q_1^T b) q_1 = (q_2^T b) q_2$$

$$b = (q_1^T b) q_1 + (q_2^T b) q_2$$



$$C = c - \left[(q_1^T c) q_1 + (q_2^T c) q_2 \right]$$
$$= (q_3^T c) q_3$$

$$c = \sum_{i=1}^{3} (q_i^T c) q_i$$

⇒ For given independent vectors, $\{a_1, \cdots, a_n\}$

$$\begin{cases} A_j = a_j - \sum_{i=1}^{j-1} (q_i^T a_j) q_i = (q_j^T a_j) q_j \\ a_j = \sum_{i=1}^{j} (q_i^T a_j) q_i \end{cases}$$

$$q_j = \frac{A_j}{\|A_j\|}$$ → normalization.