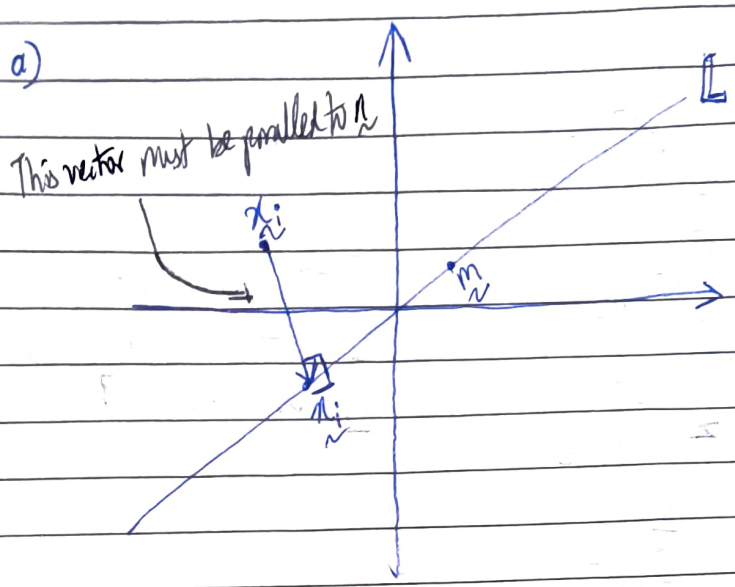
Subhas
Kumuth

Assignment 3

Quick Work
Page No.: 1
Date: 25/4/25
M T W T F S S

25945

3) a)



Because $x_i - \bar{x}_i$ should be parallel to n, we can say that:

$$x_i - \bar{x}_i = ((x_i - m)^T n) n$$

Solving this equation, we get:

$$\bar{x}_i = x_i - ((x_i - m)^T n) n$$

b) We have to minimise the sum of the squares of Euclidean distances between each ~~data point~~ x_i & $\bar{x}_i$. Let us call this sum S.

$$\therefore S = \sum_{i=1}^{N} \| x_i - \bar{x}_i \|_2^2$$

Substituting the expression from part a), we get

$$S = \sum_{i=1}^{N} \| ((\underset{\sim}{x}_i - \underset{\sim}{m})^T \underset{\sim}{n}) \underset{\sim}{n} \|_2^2$$

$$S = \sum_{i=1}^{N} \| [(\underset{\sim}{x}_i - \underset{\sim}{m})^T \underset{\sim}{n}]^2 \cdot \| \underset{\sim}{n} \|_2^2 \quad \left[\text{Because } (\underset{\sim}{x}_i - \underset{\sim}{m})^T \underset{\sim}{n} \text{ is a scaler}\right]$$

$$= \sum_{i=1}^{N} [(\underset{\sim}{x}_i - \underset{\sim}{m})^T \underset{\sim}{n}]^2 \quad \left[\text{Because } \underset{\sim}{n} \text{ is a unit vector}\right]$$

Because we know that the optimal $\underset{\sim}{n}$ is $\underset{\sim}{m}^*$, we can replace $\underset{\sim}{m}$ with $\underset{\sim}{m}^*$.

$$\therefore S = \sum_{i=1}^{N} [(\underset{\sim}{x}_i - \underset{\sim}{m}^*)^T \underset{\sim}{n}]^2$$

c) Let $\underset{\sim}{q}$ be the unit direction vector of L. $\underset{\sim}{q}$ will be orthogonal to $\underset{\sim}{n}$. Hence, $\underset{\sim}{q}$ and $\underset{\sim}{n}$ will form a basis for $\mathbb{R}^2$.

The projection matrix onto the line's direction is $\underset{\sim}{q}\underset{\sim}{q}^T$. Similarly, the projection matrix onto $\underset{\sim}{n}$'s direction is $\underset{\sim}{n}\underset{\sim}{n}^T$.

Because Projection matrix of $\underset{\sim}{q}$ + Projection matrix of $\underset{\sim}{n}$ = Identity matrix,

$$\underset{\sim}{q}\underset{\sim}{q}^T + \underset{\sim}{n}\underset{\sim}{n}^T = \underset{\sim}{I}$$

Let us get down to the minimisation problem: we want to minimise S, where:

$$S = \sum_{i=0}^{N} \left[(\underset{\sim}{x}_i - \underset{\sim}{m}^*)^T \underset{\sim}{n} \right]^2$$

To simplify this expression, let us have $\underset{\sim}{y}_i = \underset{\sim}{x}_i - \underset{\sim}{m}^*$. Then:

$$S = \sum_{i=0}^{N} \left(\underset{\sim}{y}_i^T \underset{\sim}{n} \right)^2$$

We can see that this summation is the squared Frobenius norm of $\underset{\sim\sim}{X} \underset{\sim\sim}{n} \underset{\sim}{n}^T$ where:

$$\underset{\sim}{X} = \begin{bmatrix} - \ \underset{\sim}{y}_1^T \ - \\ \vdots \\ - \ \underset{\sim}{y}_n^T \ - \end{bmatrix}$$

Hence, S can be further simplified: $S = \| X n n^T \|_F^2$

Because we know that $\underset{\sim\sim}{n} \underset{\sim\sim}{n}^T = \underset{\sim}{I} - \underset{\sim}{q} \underset{\sim}{q}^T$,

$$S = \| \underset{\sim}{X} (\underset{\sim}{I} - \underset{\sim}{q} \underset{\sim}{q}^T) \|_F^2$$

$$= \| \underset{\sim}{X} \|_F^2 - \| \underset{\sim}{X} \underset{\sim}{q} \|_F^2$$

Because $\| \underset{\sim}{X} \|_F^2$ is a constant we have to maximise $\| \underset{\sim}{X} \underset{\sim}{q} \|_F^2$ so that S is minimised. For this we shall have to find a $\underset{\sim}{q}$ s.t. $\| \underset{\sim}{X} \underset{\sim}{q} \|_F^2$ is maximised.

d) Because we know that the maximum value of $\|\underset{\sim}{X}\underset{\sim}{q}\|_2^2$ occurs when $\underset{\sim}{q}$ is the eigenvector corresponding to the largest eigenvalue of $\underset{\sim}{X}$.

Furthermore, the maximum value of $\|\underset{\sim}{X}\underset{\sim}{q}\|_2^2$ will be equal to this largest eigenvalue.

If we perform an SVD on $\underset{\sim}{X}$ we know that:

→ The largest eigenvalue will be σ_1^2 (square of largest singular value)

→ $\underset{\sim}{q} = \underset{\sim}{v}_1$, where $\underset{\sim}{v}_1$ is the largest principal component

Hence, the problem is now reduced to finding the best 1-rank approximation for $\underset{\sim}{X}$. We know, from the Eckhart-Young theorem, that the best way to do this is to perform a low rank approximation.

The one-rank approximation of $\underset{\sim}{X} = \underset{\sim}{u}_1 \sigma_1 \underset{\sim}{v}_1^T$

Because σ_1 is a scalar, $\underset{\sim}{X} = \sigma_1 \underset{\sim}{u}_1 \underset{\sim}{v}_1^T$

We can also see that the unit vector of the line, $\underset{\sim}{q} = \underset{\sim}{v}_1$.

Because $\underset{\sim}{V}^T$ is an orthogonal matrix, $\underset{\sim}{v}_1$ & $\underset{\sim}{v}_2$ are orthonormal to each other. Also, because we know that $\underset{\sim}{q}_k = \underset{\sim}{v}_1$, $\underset{\sim}{v}_k$ must be equal to $\underset{\sim}{n}$.

In other words, $\underset{\sim}{V}^T = \begin{bmatrix} | & | \\ \underset{\sim}{q} & \underset{\sim}{n} \\ | & | \end{bmatrix}$

e) To get the equations of the best fit ~~line~~ line, we need a point on the line. We also shall require a unit normal vector to the line

→ We have $\underset{\sim}{m}^*$ as the point on the line.
→ We have $\underset{\sim}{n}$ is the unit normal vector. We may use $\underset{\sim}{v}_2^T$ also.

Hence, the equation for the best-fit line is:

$$\underset{\sim}{L} = \underset{\sim}{n}(\underset{\sim}{x}_i - \underset{\sim}{m}^*)$$