# Linear Algebra Assignment

## Case Study 2

Given :- $A = \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix}^T$ $v_i \in \mathbb{R}^m$ $\Rightarrow A = \begin{bmatrix} v_{11} & v_{12} & \cdots & v_{1n} \\ v_{21} & v_{22} & \cdots & v_{2n} \\ \vdots \\ v_{m1} & v_{m2} & \cdots & v_{mn} \end{bmatrix}$

Aim $\rightarrow$ To minimize $L = \| \sum_{i=1}^{n} x_i v_i - w \|_2$

(a) To prove :- Problem can be solved directly if A is a diagonal matrix.

If $A = [v_1 \ v_2 \ \ldots \ v_n]$ and $x \in \mathbb{R}^n$ $\Rightarrow x = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$

$\Rightarrow L = \|Ax - w\|_2$

When A is a diagonal matrix :-

$$A = \begin{bmatrix} d_1 & 0 & 0 & \cdots & 0 \\ 0 & d_2 & \cdots & \cdots & 0 \\ \vdots & & \ddots & & \\ 0 & & & & d_n \end{bmatrix}$$

$A \in \mathbb{R}^{m \times n}$

☆ Each diagonal entry $d_i$ corresponds to the scaling of the coefficient $x_i$, and all non-diagonal entries are zero.

$$\Rightarrow Ax = \begin{bmatrix} d_1 x_1 \\ d_2 x_2 \\ \vdots \\ d_m x_m \end{bmatrix}$$

$$\Rightarrow L = \sqrt{\sum_{i=1}^{m} (d_i x_i - w_i)^2}$$

To minimize L $\Rightarrow \frac{dL_i}{dx_i} = 0 \Rightarrow \frac{d\left[(d_i x_i - w_i)^2\right]}{dx_i} = 0$

$$\Rightarrow 2(d_i x_i - w_i)(d_i) = 0$$

$$\Rightarrow \boxed{x_i = \frac{w_i}{d_i}} \quad (\text{if } d_i \neq 0)$$

$$\Rightarrow x = \begin{bmatrix} \frac{w_1}{d_1} \\ \frac{w_2}{d_2} \\ \vdots \\ \frac{w_m}{d_m} \end{bmatrix}$$

$\Rightarrow$ solution is direct as it involves only division when A is diagonal.

★ Special cases :-

① if $d_i = 0$ and $w_i \neq 0 \Rightarrow$ no solution exists.

② if $d_i = 0$ and $w_i = 0 \Rightarrow \infty$ solutions.

Hence, For a diagonal matrix A, the problem reduces to minimizing independent quadratic terms and sol is given by :- $x_i = \frac{w_i}{d_i}$

(b) Geometric interpretation of SVD (singular value decomposition) :-

$$A_{m \times n} = U \Sigma V^T$$

where → U is an $m \times n$ orthogonal matrix

→ $\Sigma$ is an $m \times n$ diagonal matrix containing singular values of A.

→ V is an $n \times n$ orthogonal matrix ($VV^T = I$)

☆ Geometrically :-

① Transformation of a vector :-

→ Any vector $x \in R^n$ is first rotated by $V^T$.

→ Then it is stretched along the axes defined by the singular values in $\Sigma$.

→ Finally, it is rotated again by U.

☆ Meaning of singular values of a matrix :-

① Geometric interpretation :-

→ Singular values represent how much a matrix A stretches or compresses a vector when transforming it.

→ If x is a vector, applying A to x scales it by a factor of $\sigma_i$ along the corresponding singular vector.

Objective → minimize $\|Ax - w\|$ → least squares eq$^n$ — ①

Assumption :- $A = U \Sigma V^T$

$U$ → orthogonal matrix $(m \times m)$

$\Sigma$ → diagonal matrix $(m \times n)$

$V^T$ → orthogonal matrix $(n \times n)$

Using SVD, we write eq ① as :-

$\Rightarrow Ax - w = U \Sigma V^T x - w$

$\Rightarrow U^T (Ax - w) = U^T (U \Sigma V^T x - w)$ {multiplying it with $U^T$}

$\Rightarrow \Sigma V^T x - U^T w = U^T (Ax - w)$ — ② → minimize this now.

when A is diagonal (special case) :-

$A = D = \text{diag}(\sigma_1, \sigma_2, \ldots, \sigma_n)$

which means :- $U = I_m$ (Identity matrix)

$V = I_n$

$\Sigma = D$ (already diagonal)

So, SVD of A $\Rightarrow A = I_m \Sigma I_n^T = \Sigma = D$

Putting this in eq ② :-

$\Rightarrow \Sigma V^T x - U^T w = Dx - w$

$\Rightarrow \min \|Dx - w\|$

→ Since D is diagonal, solving this problem is much easier because each eq$^n$ is transformed into independent scalar eq$^n$:-

$$\sigma_i x_i - w_i = 0$$

$$\Rightarrow \sigma_i x_i = w_i$$

$$\Rightarrow \boxed{x_i = \frac{w_i}{\sigma_i}} \quad \text{if } \sigma_i \neq 0$$

(C) given n data points $(x_1, y_1), (x_2, y_2) \ldots (x_n, y_n) \in R^2$

Aim → to find best fitting linear eq$^n$ (this is linear regression)

$$\Rightarrow y = ax + b$$

So, we want to minimize L:-

$$L = \sum_{i=1}^{n} (ax_i + b - y_i)^2 \quad \text{— (1)}$$

Converting this to a least squares problem:-

$$A = \begin{bmatrix} x_1 & 1 \\ x_2 & 1 \\ \vdots & \vdots \\ x_n & 1 \end{bmatrix} \in R^{n \times 2}$$

$$x = \begin{bmatrix} a \\ b \end{bmatrix}$$

$$y = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}$$

$$\Rightarrow L = \|Ax - y\|^2$$

$\Rightarrow \quad Ax = y$

$\Rightarrow \quad A^T A x = A^T y$

$\Rightarrow \quad \boxed{x = (A^T A)^{-1} A^T y}$ → This yields the best-fit values of a and b, making this a special case of general least squares formulation.

(d) Polynomial Regression :-

$$y = ax^2 + bx + c$$

We want to minimize :- $L = \sum_{i=1}^{n} (ax_i^2 + bx_i + c - y_i)^2$ — ①

Converting this to a least squares problem :-

$$A = \begin{bmatrix} x_1^2 & x_1 & 1 \\ x_2^2 & x_2 & 1 \\ \vdots & \vdots & \vdots \\ x_n^2 & x_n & 1 \end{bmatrix} \in R^{n \times 3}$$

$$x = \begin{bmatrix} a \\ b \\ c \end{bmatrix}$$ → unknowns we have to find.

$$y = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}$$

eq ① becomes :- $L = \|Ax - y\|^2$

$\Rightarrow \quad A^T A x = A^T y$

$\Rightarrow \quad \boxed{x = (A^T A)^{-1} A^T y}$