Linear algebra for AI & ML

$x_i^T \in \mathbb{R}^p$, $y_i \in \mathbb{R}$

$f: \mathbb{R}^p \to \mathbb{R}$

$\tilde{f} : \mathbb{R}^p \to \mathbb{R}$

$f_1, f_2, \dots, f_m : \mathbb{R}^p \to \mathbb{R}$

$y_i = f(x_i)$ for $i = 1, 2, \dots, N$

$\{ (\underbrace{x_i}_{p\text{- vector}}, \overset{\text{labels}}{y_i}) \}_{i=1}^{N}$: data set

obj: Approximate f by $\tilde{f}$

where $\tilde{f} = \alpha_1 f_1 + \alpha_2 f_2 + \dots + \alpha_m f_m$

and $f_1, f_2, \dots, f_m$ are known fns (basis fns)

$$T \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_m \end{pmatrix} = \alpha_1 f_1 + \alpha_2 f_2 + \dots + \alpha_m f_m$$

$$\alpha_1 f_1(x_i) + \alpha_2 f_2(x_i) + \cdots + \alpha_m f_m(x_i) = \hat{y}_i \quad \forall i = 1, 2, \ldots, N$$

$$\hat{y} = \begin{bmatrix} \hat{y}_1 \\ \vdots \\ \hat{y}_N \end{bmatrix} \quad \& \quad y = \begin{bmatrix} y_1 \\ \vdots \\ y_N \end{bmatrix}, \quad r = \begin{pmatrix} \hat{y}_1 - y_1 \\ \vdots \\ \hat{y}_N - y_N \end{pmatrix}$$

$$\min_{\alpha_1, \ldots, \alpha_m} \|r\|_2^2 = \min_{\alpha \in \mathbb{R}^m} \|y - A\alpha\|_2^2$$

where $A_i = [f_1(x_i) \; f_2(x_i) \ldots f_m(x_i)] \in \mathbb{R}^{1 \times m}$

i^{th} row of A & $\alpha = \begin{bmatrix} \alpha_1 \\ \vdots \\ \alpha_m \end{bmatrix}$

$A \in \mathbb{R}^{N \times m}$, $\alpha \in \mathbb{R}^m$, $y \in \mathbb{R}^N$

$$\boxed{\hat{\alpha} = (A^T A)^{-1} A^T y}$$

(x_i, y_i) $\quad x_i \in \mathbb{R}, \quad y_i \in \mathbb{R}, \quad i = 1, 2, \ldots, N$

i) $y_i = f(x_i)$

$$\tilde{f}(x_i) = \alpha_1 f_1(x_i) + \alpha_2 f_2(x_i)$$

$$f_1(x) = 1$$

$$f_2(x) = x$$

$$\tilde{f}(x) = \alpha_1 + \alpha_2 x$$

ii) $$\tilde{f}(x_j) = \sum_{i=1}^{m} \alpha_i f_i(x_j)$$

$$f_1(x) = 1, \quad f_2(x) = x, \quad f_3(x) = x^2, \quad \ldots, \quad f_m(x) = x^{m-1}$$

$$\tilde{f} = \alpha_1 + \alpha_2 x + \cdots + \alpha_m x^{m-1}$$

$$A = \begin{bmatrix} 1 & x_1 & x_1^2 & \cdots & x_1^{m-1} \\ 1 & x_2 & x_2^2 & & x_2^{m-1} \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & x_N & x_N^2 & \cdots & x_N^{m-1} \end{bmatrix} \in \mathbb{R}^{N \times m}$$

Vandermonde matrix.

3) $f_1(x) = 1$

$$\tilde{f}(x) = \alpha_1$$

$$\min_{\alpha_1} \sum_{i=1}^{N} \frac{1}{N} \left(y_i - \alpha_1 \right)^2$$

$$A = \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix}_{N \times 1} \qquad \alpha = [\alpha_1]_{1 \times 1} \qquad y = \begin{bmatrix} y_1 \\ \vdots \\ y_N \end{bmatrix}$$

$$\alpha = (A^T A)^{-1} A^T y = \frac{y_1 + y_2 + \cdots + y_N}{N}$$

4) $\{y_1, y_2, y_3, y_4, \ldots, y_{100}\}$ ← time series data.

$$y_4 = \alpha_1 y_1 + \alpha_2 y_2 + \alpha_3 y_3$$

$$y_5 = \alpha_1 y_2 + \alpha_2 y_3 + \alpha_3 y_{\cancel{5}4}$$

$$\underbrace{\begin{bmatrix} y_1 & y_2 & y_3 \\ y_2 & y_3 & y_4 \\ y_3 & y_4 & y_5 \\ y_4 & y_5 & y_6 \\ y_5 & y_6 & y_7 \\ & \vdots & \\ y_{97} & y_{98} & y_{99} \end{bmatrix}}_{A} \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \end{bmatrix} = \begin{bmatrix} y_4 \\ y_5 \\ y_6 \\ y_7 \\ y_8 \\ \vdots \\ y_{100} \end{bmatrix}$$

Iterative method for least squares.

$$A = QR$$

$$Ax = b$$

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

$$x = (A^TA)^{-1}A^Tb$$

$$= R^{-1}Q^T$$

$$\min \|Ax - b\|_2^2$$

Iterations:

$$x^{(0)} = 0$$

$k = 1, 2, \ldots$

$$\mu = \frac{1}{\|A\|_2^2}$$

$$x^{(k+1)} = x^{(k)} - \mu A^T(Ax^{(k)} - b)$$

$\mu > 0$ is a positive constant.

$$x^{(k+1)} = (I - \mu A^TA)x^{(k)} + \mu A^Tb$$

i) Remember right inverse of a matrix A?

$A \in \mathbb{R}^{10 \times 3}$

$\nexists \; X \in \mathbb{R}^{3 \times 10}$

s.t. $AX = I_{10}$

$$A x_i = e_i$$

$$A \left[x_1 \; \cdots \; x_{10} \right] = \left[e_1 \;\; e_2 \; \cdots \; e_{10} \right]$$

$$\min \left[\|A x_1 - e_1\|_2^2 + \|A x_2 - e_2\|_2^2 + \cdots + \|A x_{10} - e_{10}\|_2^2 \right]$$

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

$$\hat{X}_{LS} = (A^T A)^{-1} A^T \left[e_1 \;\; e_2 \;\; \cdots \;\; e_{10} \right]$$