PRML by Bishop - chapter 3, Section 3.3.2 - ①

Predictive Distribution - EQUATION 3.59 -

Vectorized Implementation for a batch input

1) Eqⁿ 3.59 states that the variance $\sigma_N^2(x)$ of the predictive distribution is given by

$$\sigma_N^2(x) = \frac{1}{\beta} + \phi(x)^T S_N \phi(x) \quad -(3.59)$$

This is the variance of a single input sample.

Here $\phi(x) \in \mathbb{R}^{M \times 1}$; $S_N \in \mathbb{R}^{M \times M}$

NOTE:- $\phi_0(x) =$ Bias Term $= 1$

& $\phi(x) = [\phi_0(x)\ \phi_1(x) \cdots \phi_{M-1}(x)]^T$

is the set of basis functions.

2) Lets again consider a single input x_1 and define

$\hookrightarrow \phi(x_1) = [\phi_{10}\ \phi_{11} \cdots \phi_{1(M-1)}]^T$ — ⓐ

$$\hookrightarrow S_N = \begin{bmatrix} S_{11} & S_{12} & \cdots & S_{1M} \\ S_{21} & S_{22} & \cdots & S_{2M} \\ \vdots & \vdots & & \vdots \\ S_{M1} & S_{M2} & \cdots & S_{MM} \end{bmatrix} \quad -ⓑ$$

NOTE:- ϕ_{10} = Bias Term = 1.

$\phi_{11} = \phi_1(x_1)$

$\phi_{1(M-1)} = \phi_{M-1}(x_1)$

(next page)

3) The second term is eqn (3.59) is given by

$$\underline{\phi(x_1)^T} \, \underline{\underline{S_n}} \, \underline{\phi(x_1)}$$

$$= \underbrace{\begin{bmatrix} \phi_{10} & \phi_{11} & \cdots & \phi_{1(M-1)} \end{bmatrix}}_{\mathbb{R}^{1\times M}} \underbrace{\begin{bmatrix} S_{11} & S_{12} & \cdots & S_{1M} \\ S_{21} & S_{22} & \cdots & S_{2M} \\ \vdots & \vdots & & \vdots \\ S_{M1} & S_{M2} & \cdots & S_{MM} \end{bmatrix}}_{R^{M\times M}} \underbrace{\begin{bmatrix} \phi_{10} \\ \phi_{11} \\ \vdots \\ \phi_{1(M-1)} \end{bmatrix}}_{R^{M\times 1}} \quad - (d)$$

$$= \underbrace{\begin{bmatrix} S_1 & S_2 & \cdots & S_M \end{bmatrix}}_{\mathbb{R}^{1\times M}} \underbrace{\begin{bmatrix} \phi_{10} \\ \phi_{11} \\ \vdots \\ \phi_{1M-1} \end{bmatrix}}_{\mathbb{R}^{M\times 1}} \quad - (e)$$

where

$$S_1 = S_{11}\phi_{10} + S_{21}\phi_{11} + \cdots + S_{M1}\phi_{1(M-1)}$$

$$S_2 = S_{12}\phi_{10} + S_{22}\phi_{11} + \cdots + S_{M2}\phi_{1(M-1)}$$

$$\vdots$$

$$S_M = S_{1M}\phi_{10} + S_{2M}\phi_{11} + \cdots + S_{MM}\phi_{1(M-1)}$$

(f)

4) Expanding (e) using (f), we get 3

$\phi(x_1)^T S_n \phi(x_1)$

$$= [s_1 \phi_{10} + s_2 \phi_{11} + \dots + s_M \phi_{1(M-1)}] \quad \text{— (g)}$$

$$= [s_{11} \phi_{10} \phi_{10} + s_{21} \phi_{11} \phi_{10} + \dots + s_{M1} \phi_{1(M-1)} \phi_{10}$$

$$+ s_{12} \phi_{10} \phi_{11} + s_{22} \phi_{11} \phi_{11} + \dots + s_{M2} \phi_{1(M-1)} \phi_{11}$$

$$+ \vdots$$

$$+ s_{1M} \phi_{10} \phi_{1(M-1)} + s_{2M} \phi_{11} \phi_{1(M-1)} + \dots + s_{MM} \phi_{1(M-1)} \phi_{1(M-1)}] \rightarrow (h)$$

$$= \sum_{i=1}^{M} \sum_{j=1}^{M} [s_{ij} \phi_{1(i-1)} \phi_{1(j-1)}] \rightarrow (i)$$

5) Lets now consider a batch input X with N samples $[x_1 \; x_2 \; \dots \; x_N]$. The design matrix is given by

$$\Phi = \begin{bmatrix} \phi_0(x_1) & \phi_1(x_1) & \dots & \phi_{M-1}(x_1) \\ \phi_0(x_2) & \phi_1(x_2) & \dots & \phi_{M-1}(x_2) \\ \vdots & \vdots & & \vdots \\ \phi_0(x_N) & \phi_1(x_N) & \dots & \phi_{M-1}(x_N) \end{bmatrix} \in \mathbb{R}^{N \times M} \rightarrow (j)$$

$$= \begin{bmatrix} \phi_{10} & \phi_{11} & \dots & \phi_{1(M-1)} \\ \phi_{20} & \phi_{21} & \dots & \phi_{2(M-1)} \\ \vdots & \vdots & & \vdots \\ \phi_{N0} & \phi_{N1} & \dots & \phi_{N(M-1)} \end{bmatrix} \in \mathbb{R}^{N \times M} \quad \text{— (k)}$$

6) we would like to obtain (4)

$$\sigma_N^2(\underline{X}) = \frac{1}{\beta} + \begin{bmatrix} \sigma_N^2(x_1) \\ \sigma_N^2(x_2) \\ \vdots \\ \sigma_N^2(x_N) \end{bmatrix} \longrightarrow (L)$$

where each $\sigma_N^2(x_i)$ is given by eq^n (i) above.

7) Consider $\underline{\underline{\Phi}} \cdot \underline{\underline{S_n}}$

$$= \underbrace{\begin{bmatrix} \phi_{10} & \phi_{11} & \cdots & \phi_{1M-1} \\ \phi_{20} & \phi_{21} & \cdots & \phi_{2M-1} \\ \vdots & \vdots & & \vdots \\ \phi_{N0} & \phi_{N1} & \cdots & \phi_{NM-1} \end{bmatrix}}_{\mathbb{R}^{N\times M}} \underbrace{\begin{bmatrix} S_{11} & S_{12} & \cdots & S_{1M} \\ S_{21} & S_{22} & \cdots & S_{2M} \\ \vdots & \vdots & & \vdots \\ S_{M1} & S_{M2} & \cdots & S_{MM} \end{bmatrix}}_{\mathbb{R}^{M\times M}} \to (M)$$

$$= \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1M} \\ a_{21} & a_{22} & \cdots & a_{2M} \\ \vdots & \vdots & & \vdots \\ a_{N1} & a_{N2} & \cdots & a_{NM} \end{bmatrix} \in \mathbb{R}^{N\times M} \longrightarrow (N)$$

where

$$a_{11} = [S_{11}\phi_{10} + S_{21}\phi_{11} + \cdots + S_{M1}\phi_{1M-1}]$$

$$a_{12} = [S_{12}\phi_{10} + S_{22}\phi_{11} + \cdots + S_{M2}\phi_{1M-1}]$$

$$\vdots$$

$$a_{1M} = [S_{1M}\phi_{10} + S_{2M}\phi_{11} + \cdots + S_{MM}\phi_{1M-1}]$$

(next page). $\to$ (O)

$$a_{21} = [S_{11}\phi_{20} + S_{21}\phi_{21} + \cdots + S_{M1}\phi_{2M-1}]$$ (5)

$$a_{22} = [S_{12}\phi_{20} + S_{22}\phi_{21} + \cdots + S_{M2}\phi_{2M-1}] \Rightarrow (P)$$

$$\vdots$$

$$a_{2M} = [S_{1M}\phi_{20} + S_{2M}\phi_{21} + \cdots + S_{MM}\phi_{2M-1}]$$

$$\vdots$$

$$a_{N1} = [S_{11}\phi_{N0} + S_{21}\phi_{N1} + \cdots + S_{M1}\phi_{NM-1}]$$

$$a_{N2} = [S_{12}\phi_{N0} + S_{22}\phi_{N1} + \cdots + S_{M2}\phi_{NM-1}]$$

$$\vdots$$

$$a_{NM} = [S_{1M}\phi_{N0} + S_{2M}\phi_{N1} + \cdots + S_{MM}\phi_{NM-1}] \rightarrow (q)$$

8) Now, consider $\left(\underline{\underline{\Phi}} \cdot \underline{\underline{S_n}}\right) * \underline{\underline{\Phi}}$ → element-wise multiplication

Let $\left(\underline{\underline{\Phi}} \cdot \underline{\underline{S_n}}\right) * \underline{\underline{\Phi}}$

$$= \begin{bmatrix} b_{11} & b_{12} & \cdots & b_{1M} \\ b_{21} & b_{22} & \cdots & b_{2M} \\ \vdots & \vdots & \ddots & \vdots \\ b_{N1} & b_{N2} & \cdots & b_{NM} \end{bmatrix} \in \mathbb{R}^{N \times M} \longrightarrow (R)$$

Then, we have

$b_{11} = a_{11}\phi_{10}$

$b_{12} = a_{12}\phi_{11}$

$\vdots$

$b_{1M} = a_{1M}\phi_{1(M-1)}$

$b_{21} = a_{21}\phi_{20}$

$b_{22} = a_{22}\phi_{21}$

$\vdots$

$b_{2M} = a_{2M}\phi_{2(M-1)}$

$b_{N1} = a_{1M}\phi_{N0}$

$b_{N2} = a_{N2}\phi_{N1}$

$\vdots$

$b_{NM} = a_{NM}\phi_{N(M-1)}$

9) Next, we compute the sum of elements of the above matrix along the column axis [axis = 1]

$\hookrightarrow b_{11} + b_{12} + \cdots + b_{1M}$

$$= [S_{11}\phi_{10}\phi_{10} + S_{21}\phi_{11}\phi_{10} + \cdots + S_{M1}\phi_{1(M-1)}\phi_{10}$$
$$+ S_{12}\phi_{10}\phi_{11} + S_{22}\phi_{11}\phi_{11} + \cdots + S_{M2}\phi_{1(M-1)}\phi_{11}$$
$$+ \vdots$$
$$+ S_{1M}\phi_{10}\phi_{1(M-1)} + S_{2M}\phi_{11}\phi_{1(M-1)} + \cdots + S_{MM}\phi_{1(M-1)}\phi_{1(M-1)}]$$

$\hookrightarrow$ (T)

$$= \underline{\phi(x_1)^T \underline{\underline{S_n}} \phi(x_1)} \Rightarrow \text{(U)}$$

$\hookrightarrow [= \sigma_N^2(x_1)]$

$\hookrightarrow$ |||ry $b_{22} + b_{22} + \cdots + b_{2M}$

$$= \underline{\phi(x_2)^T \underline{\underline{S_n}} \phi(x_2)} \text{ — (V)}$$

10) $\therefore$ $\sigma_n^2(\underline{X}) = \text{np.sum}[(\underline{\Phi} \cdot \underline{\underline{S_n}}) * \underline{\Phi},\ \text{axis}=1] + \frac{1}{\beta}$

$\hookrightarrow$ Eq^n (W)