Assignment 1 (CS18BTECH11021 & CS18BTECH11039)

Q 2.1.1 we have $y(x, w) = W^T \times \phi(x)$ {given}.

$D_y = K \times 1$, $D_w = M \times K$, $D_{\phi_x} = M \times 1$ --- (i).

→ Assuming the conditional distribution of Target vector, here y to be an isotropic Gaussian of form:

$$P(y / x, W, \beta) = \mathcal{N}(y / W^T \phi(x), \sigma) \quad \text{--- (ii)}$$

→ Probability distribution of y_i as a Gaussion :-

$$f(y_i) = \frac{1}{\sigma\sqrt{2\pi}} e^{-(x_i - \mu)^2 / 2\sigma^2} \quad \text{---- (iii)}$$

→ given K such y_i & assuming all $\{y_i\}$ are independent we can express joint Probability distribution function as the Probability Distribution function of y_i as Gaussian is we know given by :-

$$f(Y) = \prod_{i=1}^{n} f(y_i) = \prod_{i=1}^{n} \frac{1}{\sigma\sqrt{2\pi}} e^{-\| y_i - W^T \phi(x_i) \|^2 / 2\sigma^2} \quad \text{----- (iv)}$$

$$f(Y) = \left(\frac{1}{\sigma\sqrt{2\pi}}\right)^n e^{\sum_{i=1}^{n} (-\| y_i - W^T \phi(x_i) \|^2 / 2\sigma^2)}$$

To maximize the likelihood function $f(Y)$ we need to maximize the exponent as $\left(\frac{1}{\sigma\sqrt{2\pi}}\right)^n$ is a constant;

so we Maximize $-\sum_{i=1}^{n} \| y_i - W^T \phi(x_i) \|^2$

↳ i.e.

<u>minimize</u> $\sum_{i=1}^{n} \| y_i - W^T \phi(x_i) \|^2$ ... (v)

→ In Matrix Notation of equation (v) combining all dimention it becomes :-

minimize $\| Y - W^T \phi(X) \|^2$ ... (vi)

→ expanding L2-norm for equation (vi) ~~for~~ using scalar product

$$\text{minimize} \rightarrow \left\{ Y \cdot Y - 2 \sum_{i=1}^{n} w_i^T (Y \cdot \phi(x_i)) + \sum_{i,j=1}^{n} w_i^T w_j^T (\phi(x_i) \cdot \phi(x_j)) \right\}.$$

... (vii)

To find Minima for equation (vii), we take the derivative with respect to W_i for $\phi(x_i)$ & setting derivative 0

$$-2\left(Y \cdot \phi(x_i)\right) + 2\sum_{i,j=1}^{n} W_j^T\left(\phi(x_i) \cdot \phi(x_j)\right) = 0$$

$$\Rightarrow \forall\, i,j \in \{1,2,\ldots n\};\; \left(Y \cdot \phi(x_i)\right) = \sum_{i,j=1}^{n} w_j^T\left(\phi(x_i) \cdot \phi(x_j)\right)$$

$$\Rightarrow \quad \phi^T y_i = w_i\, \phi^T \phi$$

$$w_i = (\phi^T \phi)^{-1} \phi^T y_i$$

→ for entire column of output Y, combining these individual y_i, as we have assumed each output y_i is independent of output y_j, $i \neq j$ we have found the Maximum Likelihood estimate for multioutput regression.

$$\underline{W_{MLE} = (\phi^T \phi)^{-1} (\phi^T Y)}.$$

Q 2.1.2 MAP estimate

given α, assume gaussian Prior for w_i

$$\rightarrow P(w_i/\alpha) = \left(\sqrt{\frac{\alpha}{2\pi}}\right)^{M} e^{-\frac{+\alpha}{2}(w_i^T \cdot w_i)}$$

MAP is given as

$$\hookrightarrow P(W|D,\alpha) \propto \left(P(D|W) \times P(W|\alpha)\right)$$

$$\propto \left(P(D|W) \times \prod_{i=1}^{K} P(w_i|\alpha)\right)$$

$$P(W/D,\alpha) = C_1 e^{-\frac{1}{2}\left\{\sum_{n=1}^{N}\sum_{i=1}^{K}\frac{(t_{ni} - w_i^T\phi(x_i))^2}{\sigma_i^2}\right.} \cdot C_2 e^{-\sum_{i=1}^{K}\left(\frac{\alpha}{2} w_i^T w_i\right)}$$

where C_1 & C_2 are constants

$$\log\left(P(W|D,\alpha)\right) = -\frac{1}{2}\sum_{n=1}^{N}\sum_{i=1}^{K}\frac{(t_{ni} - w_i^T\phi(x_i))^2}{\sigma_i^2} - \sum_{i=1}^{K}\frac{\alpha}{2} w_i^T w_i + C$$

where C is $\log(C_1 C_2)$ a constant.

To find w_i's we need to maximize log-MAP, thus taking Partial derivative w.r.t w_i

$$\frac{\partial \log(P(w/D, \alpha))}{\partial w_i} = -\phi(X)\, T_i + \phi(X)\, \phi(X)^T w_i + \alpha w_i = 0$$

$$w_i = (\phi(X)\, \phi(X)^T + \alpha I)^{-1} \phi(X) T_i$$

let $T = [t_1, t_2 \text{--} t_n]^T$

In Matrix form.

$$W = (\phi(X)\, \phi(X)^T + \alpha I)^{-1} \phi(X)\, T$$

Q2.2 :-

$$\phi(X) = \begin{bmatrix} 1 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 1 \end{bmatrix}$$

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

$$\phi(x) \cdot \phi(x)^T = \begin{bmatrix} 1 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 1 \end{bmatrix} \times \begin{bmatrix} 1 & 0 \\ 1 & 0 \\ 1 & 0 \\ 0 & 1 \\ 0 & 1 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 3 & 0 \\ 0 & 3 \end{bmatrix}$$

$$(\phi(x) \cdot \phi(x)^T)^{-1} = \begin{bmatrix} 1/3 & 0 \\ 0 & 1/3 \end{bmatrix}$$

$$\phi(x)\, T = \begin{bmatrix} 1 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 1 \end{bmatrix} \times \begin{bmatrix} -1 & -1 \\ -1 & -2 \\ -2 & -1 \\ 1 & 1 \\ 1 & 2 \\ 2 & 1 \end{bmatrix} = \begin{bmatrix} -4 & -4 \\ 4 & 4 \end{bmatrix}$$

$$W = (\phi(X)\, \phi(x)^T)^{-1} \phi(x)\, T$$

$$W_{MLE} = \begin{bmatrix} 1/3 & 0 \\ 0 & 1/3 \end{bmatrix} \times \begin{bmatrix} -4 & -4 \\ 4 & 4 \end{bmatrix} = \begin{array}{cc} w_1 & w_2 \\ \downarrow & \downarrow \\ \end{array} \begin{bmatrix} -4/3 & -4/3 \\ 4/3 & 4/3 \end{bmatrix}$$

$$w_1 = [-4/3, 4/3] \; ; \quad w_2 = [-4/3, 4/3]$$