Kashi Vishwanath.B

20012715

1) Given $\mu = \begin{pmatrix} \mu_a \\ \mu_b \\ \mu_c \end{pmatrix}$ $\Sigma = \begin{pmatrix} \Sigma_{aa} & \Sigma_{ab} & \Sigma_{ac} \\ \Sigma_{ba} & \Sigma_{bb} & \Sigma_{bc} \\ \Sigma_{ca} & \Sigma_{cb} & \Sigma_{cc} \end{pmatrix}$

• joint distribution of (X_a, X_b) can be expressed as

$$\begin{pmatrix} x_a \\ x_b \end{pmatrix} \sim N \left[\begin{pmatrix} \mu_a \\ \mu_b \\ \cancel{\mu} \end{pmatrix} \cdot \begin{pmatrix} \Sigma_{aa} & \Sigma_{bb} \\ \Sigma_{ba} & \Sigma_{bb} \end{pmatrix} \right]$$

Gaussian with vectors mean $(\mu_a, \mu_b)'$

Covariance matrix $\begin{bmatrix} \Sigma_{aa} & \Sigma_{ab} \\ \Sigma_{ba} & \Sigma_{bb} \end{bmatrix}$

We can partition the components into 2 categories x_a and x_b

We can derive conditional distribution

$$X_a / X_b \sim N [\mu', \Sigma']$$

Here $\mu' = \mu_a + \Sigma_{ab}\Sigma_{bb}^{-1}(x_b - \mu_b)$

and $\Sigma' = \Sigma_{aa} - \Sigma_{ab}\Sigma_{bb}^{-1}\Sigma_{ba}$

Since x_c has been marginalized out no need to consider it.

$P(X_a/X_b) \cap P(N[\mu', \Sigma'])$

2) Given $J(w) = f(w^T \phi(x_1), \dots w^T \phi(x_N))$
$+ g(w^T w)$

$g(\cdot)$ is a monotonically increasing

with $$w = \sum_{n=1}^{N} \alpha_n \phi(x_n) + w_\perp$$

$$J(w) = f(w^T \phi(x_1), \dots, w^T \phi(x_N))$$

$$\Rightarrow f(T) + g(w + (z_1 \phi(x)) + z_2 \phi(x_2) + \dots z_n \phi(x_n) + w_\perp$$

$$T = w^T \sum_{n=1}^{N} z_n \phi(x_n) + w_\perp$$

where $z_n = 1$

$$J(w) = f(w^T \phi(x_1) \dots w^T \phi(x_n)) + g(w^T (z_1 \phi(x_1) + \dots z_n \phi(x_n))^T + w_\perp)$$

This signifies the linearity of the basis function

$\phi(x_n)$ for $n = 1, 2, \ldots N$

Each of these $J(w)$ takes a corresponding $\phi(x_n)$ form in a linear way.

Hence proved, that the value of w that minimizes $J(w)$ takes the form of a linear combination of the basis functions $\phi(x_n)$ for ~~=~~ $n = 1 \ldots N$