9. training data set $(\vec{x}_1, y_1)\ (\vec{x}_2, y_2), \cdots, (\vec{x}_N, y_N)$

virtual examples $(\tilde{x}_1, y_1)\ (\tilde{x}_2, y_2), \cdots, (\tilde{x}_N, y_N)$

each $\tilde{x}_n = \vec{x}_n + \vec{\varepsilon}$ where $\vec{\varepsilon} \sim N(\vec{0}_{d+1}, \sigma^2 I_{d+1})$ $\vec{\varepsilon} \in \mathbb{R}^{d+1}$ let $\vec{\varepsilon} = [\varepsilon_0\ \varepsilon_1 \cdots \varepsilon_d]^T$

i.e. $E[\varepsilon_i] = 0 \quad \forall i = 0 \sim d$ $\quad (\varepsilon_i \sim N(0,1) \quad \forall i = 0 \sim d)$

$Var(\varepsilon_i) = 1 \quad \forall i = 0 \sim d$

$$X_h = [\vec{x}_1 \cdots \vec{x}_N\ \tilde{x}_1 \cdots \tilde{x}_N]^T \rightarrow X_h^T = [\vec{x}_1 \cdots \vec{x}_N\ \tilde{x}_1 \cdots \tilde{x}_N]$$

$$= \begin{bmatrix} \vec{x}_1^T \\ \vdots \\ \vec{x}_N^T \\ \tilde{x}_1^T \\ \vdots \\ \tilde{x}_N^T \end{bmatrix}$$

$$X^T X = [\vec{x}_1^T \cdots \vec{x}_N^T] \begin{bmatrix} \vec{x}_1 \\ \vdots \\ \vec{x}_N \end{bmatrix} = \boxed{\sum_{n=1}^{N} \vec{x}_n \vec{x}_n^T}$$

$$\boxed{E[X_h^T X_h] = \alpha X^T X + \beta \sigma^2 I_{d+1} \qquad \vec{\varepsilon} \sim N(\vec{0}_{d+1}, \sigma^2 I_{d+1})}$$

$$E[X_h^T X_h] = E\left[[\vec{x}_1 \cdots \vec{x}_N\ \tilde{x}_1 \cdots \tilde{x}_N] \begin{bmatrix} \vec{x}_1^T \\ \vdots \\ \vec{x}_N^T \\ \tilde{x}_1^T \\ \vdots \\ \tilde{x}_N^T \end{bmatrix} \right]$$

$$= E\left[\vec{x}_1 \vec{x}_1^T + \cdots + \vec{x}_N \vec{x}_N^T + \tilde{x}_1 \tilde{x}_1^T + \cdots + \tilde{x}_N \tilde{x}_N^T \right]$$

$$= E\left(\sum_{n=1}^{N} \left[\vec{x}_n \vec{x}_n^T + (\vec{x}_n + \vec{\varepsilon})(\vec{x}_n + \vec{\varepsilon})^T \right] \right)$$

$$(\vec{x}_n + \vec{\varepsilon})(\vec{x}_n + \vec{\varepsilon})^T \qquad \vec{x}_n \in \mathbb{R}^{d+1}\ \vec{\varepsilon} \in \mathbb{R}^{d+1}$$

$$= (\vec{x}_n + \vec{\varepsilon})(\vec{x}_n^T + \vec{\varepsilon}^T)$$

$$= \vec{x}_n \vec{x}_n^T + \vec{x}_n \vec{\varepsilon}^T + \vec{\varepsilon} \vec{x}_n^T + \vec{\varepsilon} \vec{\varepsilon}^T$$

$$= \sum_{n=1}^{N} E\left[2\vec{x}_n \vec{x}_n^T + \vec{x}_n \vec{\varepsilon}^T + \vec{\varepsilon} \vec{x}_n^T + \vec{\varepsilon} \vec{\varepsilon}^T \right]$$

$$= \sum_{n=1}^{N} \left[E[2\vec{x}_n \vec{x}_n^T] + E[\vec{x}_n \vec{\varepsilon}^T] + E[\vec{\varepsilon} \vec{x}_n^T] + E[\vec{\varepsilon} \vec{\varepsilon}^T] \right]$$

$$= \sum_{n=1}^{N} \left[2\vec{x}_n \vec{x}_n^T + \vec{x}_n E[\vec{\varepsilon}^T] + E[\vec{\varepsilon}] \vec{x}_n^T + E[\vec{\varepsilon} \vec{\varepsilon}^T] \right]$$

$$\vec{\varepsilon} \vec{\varepsilon}^T = \begin{bmatrix} \varepsilon_0 \\ \varepsilon_1 \\ \vdots \\ \varepsilon_d \end{bmatrix} [\varepsilon_0\ \varepsilon_1 \cdots \varepsilon_d] = \begin{bmatrix} \varepsilon_0^2 & \varepsilon_0\varepsilon_1 & \cdots & \varepsilon_0\varepsilon_d \\ \varepsilon_1\varepsilon_0 & \varepsilon_1^2 & & \vdots \\ \vdots & & \ddots & \\ \varepsilon_d\varepsilon_0 & \cdots & & \varepsilon_d^2 \end{bmatrix}$$

$$= 2\boxed{\sum_{n=1}^{N} \vec{x}_n \vec{x}_n^T} + \sum_{n=1}^{N} \vec{x}_n [0 \cdots 0]_{1,d+1} + \sum_{n=1}^{N} \begin{bmatrix} 0 \\ \vdots \\ 0 \end{bmatrix}_{d+1,1} \vec{x}_n^T + \sum_{n=1}^{N} \sigma^2 I_{d+1}$$

$$E[\vec{\varepsilon} \vec{\varepsilon}^T] = E[(\vec{\varepsilon} - E[\vec{\varepsilon}])(\vec{\varepsilon} - E[\vec{\varepsilon}])^T]$$

$$= Var(\vec{\varepsilon})$$

= covariance matrix

$$\begin{bmatrix} \sigma^2 & & 0 \\ & \sigma^2 & \\ 0 & & \ddots\ \sigma^2 \end{bmatrix}_{(d+1)\times(d+1)}$$

$$= \sigma^2 I_{d+1}$$

$$= 2X^T X + N\sigma^2 I_{d+1}$$

Thus, $\alpha = 2$, $\beta = N$ □