

linear regression: $\hat{f}(x) = x^T \hat{\beta}$
 We have seen earlier that for a new x^0 , we have

$$\begin{aligned} \text{Var}(\hat{f}(x^0)) &= \text{Var}(x^0)^T \hat{\beta}) = (x^0)^T \underbrace{\text{Var}(\hat{\beta})}_{= \sigma_\varepsilon^2 (X^T X)^{-1}} x^0 \\ &= \sigma_\varepsilon^2 (x^0)^T (X^T X)^{-1} x^0 \end{aligned}$$

$$h(x^0) = X (X^T X)^{-1} x^0$$

We get

$$\begin{aligned} \|h(x^0)\|^2 &= h(x^0)^T h(x^0) = (x^0)^T (X^T X)^{-1} \underbrace{X^T X}_{= I} (X^T X)^{-1} x^0 \\ &= (x^0)^T (X^T X)^{-1} x^0 \end{aligned}$$

$$\rightarrow \text{Var}(\hat{f}(x^0)) = \sigma_\varepsilon^2 \|h(x^0)\|^2$$

It can be shown that $\frac{1}{N} \sum_{i=1}^N \|h(x_i)\|^2 = \frac{p}{N}$

We get

$$\begin{aligned} \frac{1}{N} \sum_{i=1}^N \text{Err}(x_i) &= \frac{1}{N} \sum_{i=1}^N \left(\sigma_\varepsilon^2 + \left(f(x_i) - E(\hat{f}(x_i)) \right)^2 + \|h(x_i)\|^2 \sigma_\varepsilon^2 \right) \\ &= \underbrace{\frac{1}{N} \sum_{i=1}^N \sigma_\varepsilon^2}_{= \sigma_\varepsilon^2} + \frac{1}{N} \sum_{i=1}^N \left(f(x_i) - E(\hat{f}(x_i)) \right)^2 + \sigma_\varepsilon^2 \cdot \frac{1}{N} \sum_{i=1}^N \|h(x_i)\|^2 \\ &= \sigma_\varepsilon^2 + \frac{1}{N} \sum_{i=1}^N \left(f(x_i) - E(\hat{f}(x_i)) \right)^2 + \frac{p}{N} \sigma_\varepsilon^2 \end{aligned}$$