

We have seen earlier that for a new $\underline{x}^*$

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

$$\text{Let } h(\underline{x}^*) = \underline{X} (\underline{X}^T \underline{X})^{-1} \underline{x}^*$$

We have:

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

$$\implies \text{Var}(\hat{f}(\underline{x}^*)) = \sigma_\varepsilon^2 \|h(\underline{x}^*)\|^2$$

We get

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

It can be shown that $\sum_{i=1}^N \|h(\underline{x}_i)\|^2 = p$ - We get

$$\text{Err}_in = \frac{1}{N} \sum_{i=1}^N \text{Err}(\underline{x}_i) = \sigma_\varepsilon^2 + \frac{1}{N} \sum_{i=1}^N \left(f(\underline{x}_i) - \mathbb{E}(\hat{f}(\underline{x}_i)) \right)^2 + \frac{p}{N} \sigma_\varepsilon^2$$