

Ex. 7.7

$$GCV(\hat{f}) = \frac{1}{N} \sum_{i=1}^N \left[ \frac{y_i - \hat{f}(x_i)}{1 - \frac{\text{trace}(S)}{N}} \right]^2$$

$$= \bar{err} \cdot \left( \frac{1}{1 - \frac{\text{trace}(S)}{N}} \right)^2 \quad \left( \begin{array}{l} \text{by definition} \\ \text{of } \bar{err} \text{ (7.4)} \end{array} \right)$$

$$\approx \bar{err} \cdot \left( 1 + 2 \cdot \frac{\text{trace}(S)}{N} \right)$$

$$= \bar{err} + 2 \frac{d}{N} \bar{err} \quad (*) \quad \left( \begin{array}{l} \text{by definition of} \\ \text{effective number of} \\ \text{Parameters (7.32)} \end{array} \right)$$

Notice that AIC is defined as

$$AIC = \bar{err} + 2 \cdot \frac{d}{N} \cdot \hat{\sigma}_\varepsilon^2$$

Which is the same as (\*) if  $\bar{err}$  is used to estimate the noise Variance  $\hat{\sigma}_\varepsilon^2$ , as required.