ELS Ex 7.4 p258

Squared error loss.

Training error: $\overline{err} = \frac{1}{N}\sum_{i=1}^{N}(y_i - \hat{f}(x_i))^2$

In sample prediction error $\quad Err_{in} = \frac{1}{N}\sum_{i=1}^{N} E_{Y^0}\left[(Y_i^0 - \hat{f}(x_i))^2\right]$

observations in training set, x's kept fixed. E only over new Y^0's

TASK: establish that the average optimism is

$$E_y(op) = \omega = \frac{2}{N}\sum_{i=1}^{N} Cov(\hat{y}_i, y_i) \qquad (7.21)$$

$$E_y(op) = E_y(Err_{in} - \overline{err}) \qquad \text{"definition"}$$

expected value over the responses in the training set

$$= E_y\left[\frac{1}{N}\sum_{i=1}^{N} E_{Y^0}\left[(Y_i^0 - \hat{f}(x_i))^2\right] - \frac{1}{N}\sum_{i=1}^{N}(y_i - \hat{f}(x_i))^2\right]$$

$$= E_y\left[\frac{1}{N}\sum_{i=1}^{N}\left(E_{Y_0}\left[(Y_i^0)^2 - 2Y_i^0\cdot\hat{f}(x_i) + \hat{f}(x_i)^2\right] - \left(y_i^2 - 2y_i\hat{f}(x_i) + \hat{f}(x_i)^2\right)\right)\right]$$

$$= E_y\left[\frac{1}{N}\sum_{i=1}^{N}\left(E_{Y_0}[(Y_i^0)^2] - 2\hat{f}(x_i)E_{Y_0}(Y_i^0) + \cancel{\hat{f}(x_i)^2} - y_i^2 + 2y_i\hat{f}(x_i) - \cancel{\hat{f}(x_i)^2}\right)\right]$$

not a function of Y_i^0

$E_{Y_0}(\hat{f}(x_i)) = \hat{f}(x_i)$ and $E_{Y_0}(\hat{f}(x_i)^2) = \hat{f}(x_i)^2$

$$= \frac{1}{N}\sum_{i=1}^{N}\left(\underbrace{E_y \underbrace{E_{Y_0}(Y_i^{0\,2})}_{E_y(y_i^2)}}_{\cancel{E_y(y_i^2)}} - 2\,E_y\left(\hat{f}(x_i)\cdot \underset{Y_0}{E}(Y_i^0)\right) - \cancel{E_y(y_i^2)} + 2E_y\left(y_i\cdot \hat{f}(x_i)\right)\right)$$

same distribution

Also $E_{Y_0}(Y_i^0) = E_y(y_i)$ and

$E_y(E_{Y_0}(Y_i^0)) = E_y(y_i)$

$$= \frac{1}{N}\sum_{i=1}^{N}\left(-2\;E_y(\hat{f}(x_i))\cdot E_y(y_i) + 2\,E_y(y_i\cdot \hat{f}(x_i)) \right)$$

$$\mathrm{Cov}(y_i,\hat{y}_i) = \mathrm{Cov}(y_i, \hat{f}(x_i)) = E_y\left(y_i\cdot\hat{f}(x_i)\right) - E_y(y_i)\cdot E_y(\hat{f}(x_i))$$

$\mathrm{Cov}(V,W) = E(V\cdot W) - E(V)\cdot E(W)$

def. of coverence

$$= \frac{2}{N}\sum_{i=1}^{N}\mathrm{Cov}(y_i, \hat{f}(x_i))$$

$$= \frac{2}{N}\sum_{i=1}^{N}\mathrm{Cov}(y_i, \hat{y}_i)$$