Name: Aman Kumar
Hawk ID: A20538809
Assignment 2
CS 584

Exercise 7.4.

Solution :-

start with Err in. Let's denote $\hat{y_i} = \hat{f}(x_i)$

$$Y_i^\circ - \hat{f}(x_i) = Y_i^\circ - f(x_i) + f(x_i) - E\hat{y_i} + E\hat{y_i} - \hat{y_i}$$

thus,

$$Err_{in} = \frac{1}{N} \sum_{i=1}^{N} E_{Y^\circ} \left(Y_i^\circ - f(x_i) + f(x_i) - E\hat{y_i} + E\hat{y_i} - \hat{y_i}\right)^2$$

$$= \frac{1}{N} \sum_{i=1}^{N} A_i + B_i + C_i + D_i + E_i + F_i$$

where, $A_i = E_{Y^\circ}(Y_i^\circ - f(x_i))^2$

$$B_i = (f(x_i) - E\hat{y_i})^2$$
$$C_i = (E\hat{y_i} - \hat{y_i})^2$$
$$D_i = 2E_{Y^\circ}(Y_i^\circ - f(x_i))(f(x_i) - E\hat{y_i})$$
$$E_i = 2E_{Y^\circ}(Y_i^\circ - f(x_i))(E\hat{y_i} - \hat{y_i})$$
$$F_i = 2(f(x_i) - E\hat{y_i})(E\hat{y_i} - \hat{y_i}).$$

for $\overline{Err}$, we have,

$$y_i - \hat{f}(x_i) = y_i - f(x_i) + f(x_i) - E\hat{y_i} + E\hat{y_i} - \hat{y_i}$$

and,

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

$$= \frac{1}{N} \sum_{i=1}^{N} G_i + B_i + C_i + H_i + J_i + F_i$$

where, new terms,

$$G_i = (y_i - f(x_i))^2$$

$$H_i = 2(y_i - f(x_i))(f(x_i) - E\hat{y_i})$$

$$J_i = 2(y_i - f(x_i))(E\hat{y_i} - \hat{y_i}).$$

$$\therefore \; E_y(op) = E_y(Err_{in} - \overline{err})$$

$$= \frac{1}{N} \sum_{i=1}^{N} E_y[(A_i - G_i) + (D_i - H_i) + (E_i - J_i)]$$

as for A_i & G_i, the expectation over y captures unpredictable error, so, $E_y(A_i - G_i) = 0$
and thus, $E_y D_i = E_y H_i = E_y E_i = 0$

$$E_y(op) = -\frac{2}{N} \sum_{i=1}^{N} J_i$$

$$= -\frac{2}{N} \sum_{i=1}^{N} E_y (y_i - f(x_i))(E\hat{y_i} - \hat{y_i}).$$

$$= \frac{2}{N} \sum_{i=1}^{N} [E_y(y_i \hat{y_i}) - E_y \cdot y_i \; E_y \cdot \hat{y_i}]$$

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

proved

Exercise 7.5

Solution given, $\hat{y} = Sy$.

$$\sum_{i=1}^{N} \text{Cov}(\hat{y_i}, y_i) = \text{trace}(\text{Cov}(\hat{y}, y))$$

$$= \text{trace}(\text{Cov}(Sy, y)).$$

$$= \text{trace}(S \,\text{Cov}(y, y)).$$

$$= \text{trace}(S \,\text{Var}(y))$$

$$= \text{trace}(S) \cdot \sigma_\epsilon^2$$ proved

For the questions related to the programming is provided below:-

1, Do you get the same optimal parameters with each method?

Answer:- From the results, it has been observed that while K-fold-cross-validation values of 5 and 10 are nearly identical for each of the harmonic values, bias and variance rise as the harmonic value increases. The RMSE, AIC and bootstrap values start to fall as the harmonic value increases.

~~Furthermore, the lowest value of AIC is observed at means 3~~.

2, i, AIC gave higher bias but lower variance.

ii, K-fold-cross validation at

a, K = 5 :- gave higher variance and low bias

b, K=10 :- " lower " " high "

iii, Bootstrapping :- gave lower bias and higher variance.