

Week 2 HW

1. Forward Stepwise Selection

```
library(leaps)
library(mlbench)

data(BostonHousing)
head(BostonHousing)
sum(is.na(BostonHousing))

regfit.fwd = regsubsets(medv ~., data=BostonHousing,
nvmax=ncol(BostonHousing), method='forward')
summary(regfit.fwd)

coef(regfit.fwd, 4)
```

1-1.

(Intercept)	rm	dis	ptratio	lstat
24.4713576	4.2237922	-0.5519263	-0.9736458	-0.6654360

1-2.

```
summary(regfit.fwd)$rsq[4]
[1] 0.6903077
```

2. ESL 7.4

Ex. 7.4 Consider the in-sample prediction error (7.18) and the training error $\overline{\text{err}}$ in the case of squared-error loss:

$$\begin{aligned}\text{Err}_{\text{in}} &= \frac{1}{N} \sum_{i=1}^N E_{Y^0} (Y_i^0 - \hat{f}(x_i))^2 \\ \overline{\text{err}} &= \frac{1}{N} \sum_{i=1}^N (y_i - \hat{f}(x_i))^2.\end{aligned}$$

Add and subtract $f(x_i)$ and $E\hat{f}(x_i)$ in each expression and expand. Hence establish that the average optimism in the training error is

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

as given in (7.21).

$$W = E_y (\text{Err}_{\text{in}} - \overline{\text{err}})$$

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

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

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

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