Stats 415 Assignment 6

19342932 Shu Zhou.

1. $y_i = x_i^T \beta^* + \varepsilon$

$AIC(M) = -2\log L(M) + 2d$

So, we first calculate the maximum likelihood function

From the p.d.f: $f_i(x) = \frac{1}{\sqrt{2\pi}\,\sigma} e^{-\frac{1}{2\sigma^2}(x - \beta_1 x_{i2} - \cdots \beta_p x_{ip})^2)}$

Hence, the likelihood function is

$$\prod_{i=1}^{n} f_i Y_{(i)} = \left(\frac{1}{\sqrt{2\pi}\,\sigma}\right)^n e^{(-\frac{1}{2\sigma^2}|Y - X\beta|^2)} = L(M)$$

Then $AIC(M) = -2\log L(M) + 2d$

$$= -2\log\left[\left(\frac{1}{\sqrt{2\pi}\,\sigma}\right)^n e^{-\frac{1}{2\sigma^2}|Y - X\beta|^2}\right] + 2d$$

$$= 2d + n \times \left[(-2) \times -\tfrac{1}{2}\log 2\pi + -2 \times -\log\sigma\right] + \frac{1}{-2\sigma^2} \times (-2) \times \underset{\overset{\shortparallel}{|Y - X\beta|^2}}{SSE} + 2d$$

$$= n\left[\log 2\pi + 2\log\sigma\right] + \frac{SSE}{\sigma^2} + 2d.$$

$$= n\left[\log 2\pi + 2\log\sigma\right] + \frac{SSE}{\sigma^2} + 2p \text{ (\# of predictors)}$$

2. $\sum_{i=1}^{n}(y_i - \beta_0 - \sum_{j=1}^{p}\beta_j x_{ij})$ subject to $\sum_{j=1}^{p}|B_j| \le s$.

a. The number of included variable will steadily increases. Since we are restricting B_j in a smaller region. So the model becomes more flexible and more variables will be included

b. The training error will steadily decreases, according to the explanation in (a), the increase in the flexibility of a model will cause an decrease in training error.

c. The test error will decrease initially and then eventually starts increasing. So at first, our model is underfitting, so if we restrict the B_j coefficients, the test error will decrease. However, this restriction would eventually cause overfitting and increase the test error.

d. The variance of $\hat\beta$ will steadily increase. The model becomes more and more flexible, so the variance would steadily increas

e. The squared bias of $\hat\beta$ will steadily decrease. Since we fit our model closer to their least-square estimate, our model became more and more unbiased.