

20835281

BI Qiuyu

2.1.

(a). expect flexible method to be better.

p is small, flexible method will fit the sample data better.
 n is large, sample distribution is very close to real dist.
generally, flexible method will be better.

(b). expect flexible method to be worse.

p is large. flexible method will fit the noise well.
 n is small. sample dist is greatly different from real dist.
generally, flexible method will be worse.

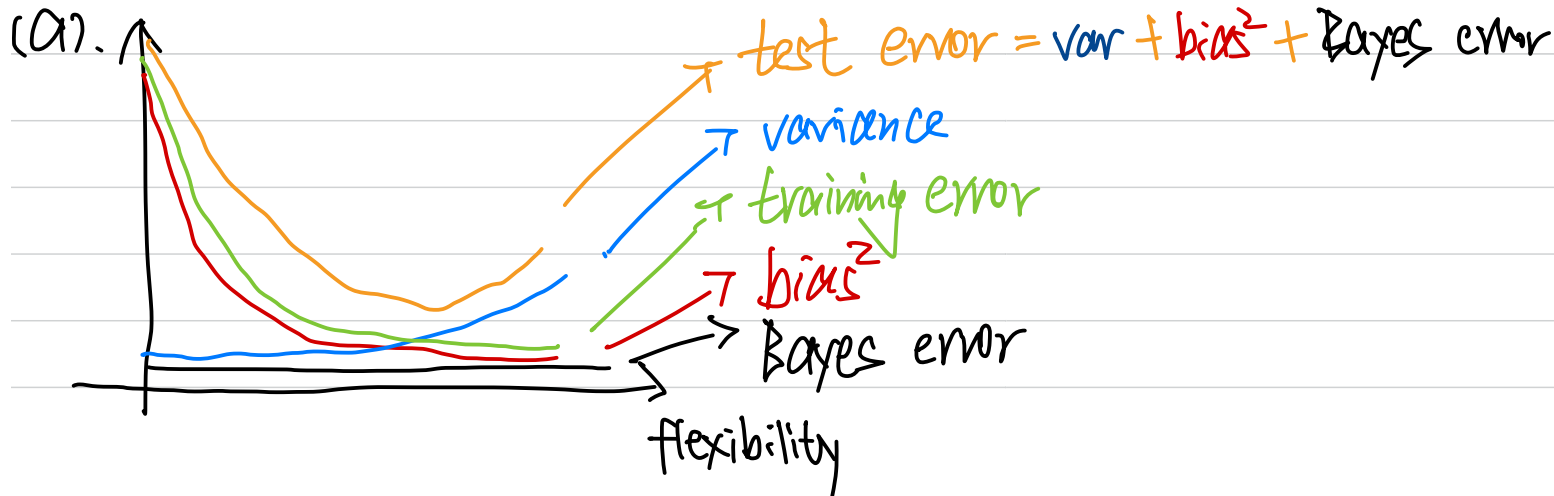
(c). expect flexible method to be better.

Predictors and response is highly non-linear, which means underlying model is more complex. A flexible method will have less bias and fit it better.

(d). expect flexible method to be worse.

G^2 is large, flexible method will fit the noise well.

2.3.



(b) **Bayes error**: It's the irreducible error between real data and underlying model, so it's constant.

Bias $= (E\hat{f}(x) - f(x))^2$, with flexibility increasing, $E\hat{f}(x)$ will be more and more close to the underlying function $f(x)$.

Var $= E(\hat{f}(x) - E\hat{f}(x))^2$, with flexibility increasing, $\hat{f}(x)$ will fit more and more to sample data, so variance of $\hat{f}(x)$ will increase slowly when fit the model and then fastly when fit the noise.

test error $= \text{Var} + \text{Bias}^2 + \text{Bayes error}$, according to above three curves, test error will decrease at the beginning and then increase when model fit more to the noise.

training error, the objective in training period is minimizing training error. With flexibility increasing, model will fit more and more to sample data, at the same time, training error is calculated using sample data, so it will decreasing continuously.

3.1

Null hypothesis are that :

There is no relationship between sales and Intercept.

There is no relationship between sales and TV

There is no relationship between sales and radio

There is no relationship between sales and newspaper.

Conclusions:

The relationship exists between sales and Intercept, TV, radio

There is no relationship between sales and newspaper

3.1

$$\hat{\beta} = \left(\sum_{i=1}^n x_i y_i \right) / \left(\sum_{i=1}^n x_i^2 \right)$$

$$\begin{aligned} \hat{y}_i &= x_i \cdot \hat{\beta} = x_i \cdot \left(\sum_{j=1}^n x_j y_j \right) / \left(\sum_{j=1}^n x_j^2 \right) \\ &= \sum_{j=1}^n \left(x_i x_j / \left(\sum_{j=1}^n x_j^2 \right) \right) \cdot y_j \end{aligned}$$

$$\text{So. } \alpha_i' = \frac{x_{i\cdot} \cdot x_i'}{\sum_{j=1}^n x_j^2}$$