Homework 3: Model Order Selection

Introduction to Machine Learning

Shangwen Yan | N17091204 | sy2160

1.

(a)

(i) only when $\beta_2 = 0$, the model is linear

(ii) no under-modeling

(iii) $\beta_0 = 1, \beta_1 = 2, \beta_2 = 0$

(b)

(i) only when $b_1 = 0, b_0 \neq 0$, the model is linear

(ii) no under-modeling

(iii) $a_0 = 3, a_1 = 3, b_0 = 2, b_1 = 3$

(c)

(i) only when $c_1 = 0, c_2 = 0$, the model is linear

(ii) there is under-modeling

(iii)

2.

(a)

$$A = [1, \boldsymbol{x}]$$

$$\hat{\boldsymbol{\beta}} = (A^T A)^{-1} A^T \boldsymbol{y}$$

(b)

$$B = [1, \boldsymbol{x}, \boldsymbol{x^2}]$$

$$\hat{\boldsymbol{\beta}} = (A^T A)^{-1} A^T B \boldsymbol{\beta_0}$$

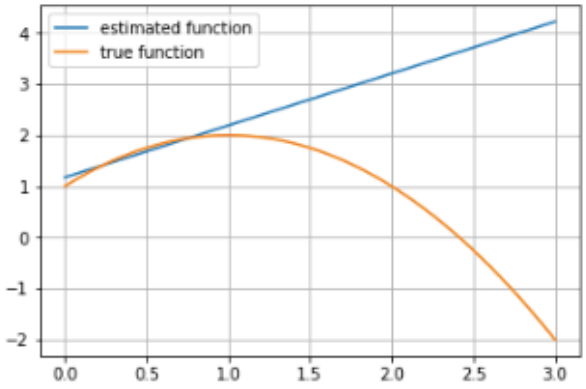
(c)

```
import numpy as np
import matplotlib
import matplotlib.pyplot as plt
from sklearn import datasets, linear_model, preprocessing
%matplotlib inline
import numpy.polynomial.polynomial as poly

beta = np.array([1,2,-1])
x = np.random.uniform(0,1,10)
y = poly.polyval(x,beta)
beta_hat = poly.polyfit(x,y,1)

x_plt = np.linspace(0,3,100)
y_plt = poly.polyval(x_plt,beta)
y_hat = poly.polyval(x_plt,beta_hat)
plt.plot(x_plt,y_hat,label="estimated function")
plt.plot(x_plt,y_plt,label ="true function")
plt.grid()
plt.legend()
```

```
<matplotlib.legend.Legend at 0x11cf6dac8>
```



(d) x = 3

3.

(a) let cancer volume be x_1 , age be x_2 and cancer type be x_3

$$Model_1 : \hat{y} = \beta_0 + \beta_1 x_1$$

$$Model_2 : \hat{y} = \beta_0 + \beta_1 x_1 + \beta_2 x_2$$

$$Model_3 : \hat{y} = \beta_0 + \beta_1 u x_1 + \beta_2 u x_1 + \beta_3 x_2, u \in \{0, 1\}$$

(b) 2 ; 3 ; 4 ; Model3 is the most complex one

(c)

$$A_1 = \begin{bmatrix} 1 & 0.7 \\ 1 & 1.3 \\ 1 & 1.6 \\ \vdots & \vdots \end{bmatrix}$$

$$A_2 = \begin{bmatrix} 1 & 0.7 & 55 \\ 1 & 1.3 & 65 \\ 1 & 1.6 & 75 \\ \vdots & \vdots & \vdots \end{bmatrix}$$

$$A_3 = \begin{bmatrix} 1 & 0.7 & 0 & 55 \\ 1 & 0 & 1.3 & 65 \\ 1 & 0 & 1.6 & 75 \\ \vdots & \vdots & \vdots & \vdots \end{bmatrix}$$

(d) select Model 2