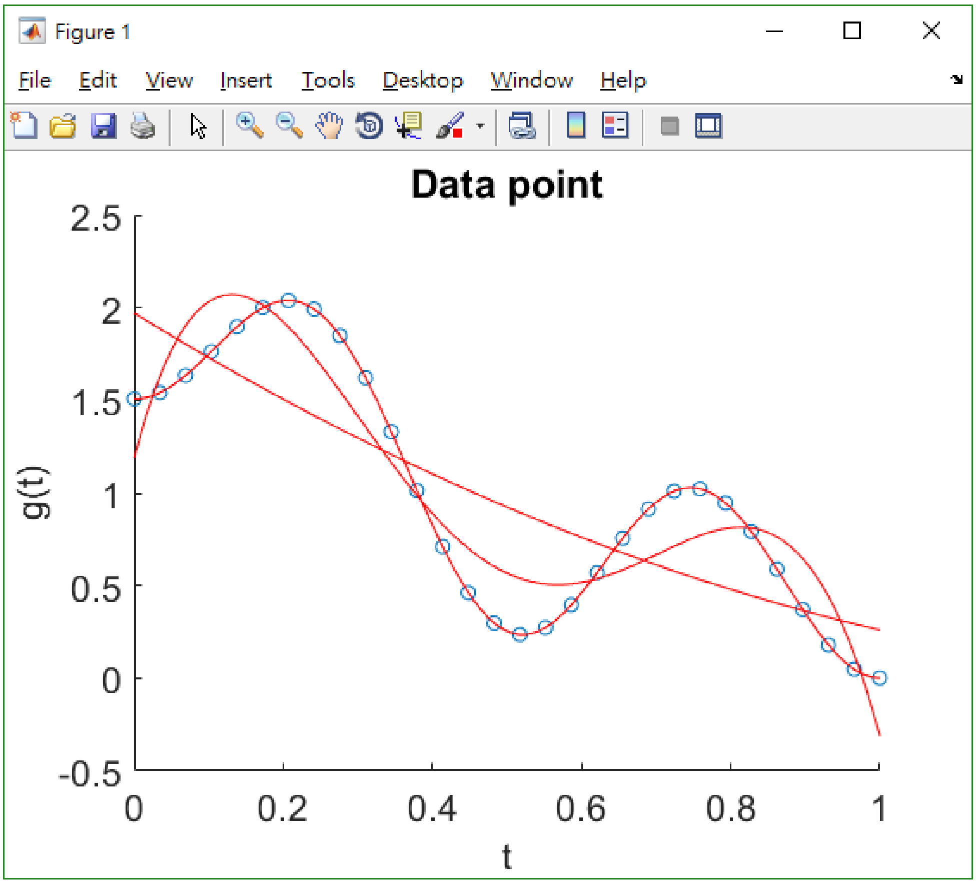
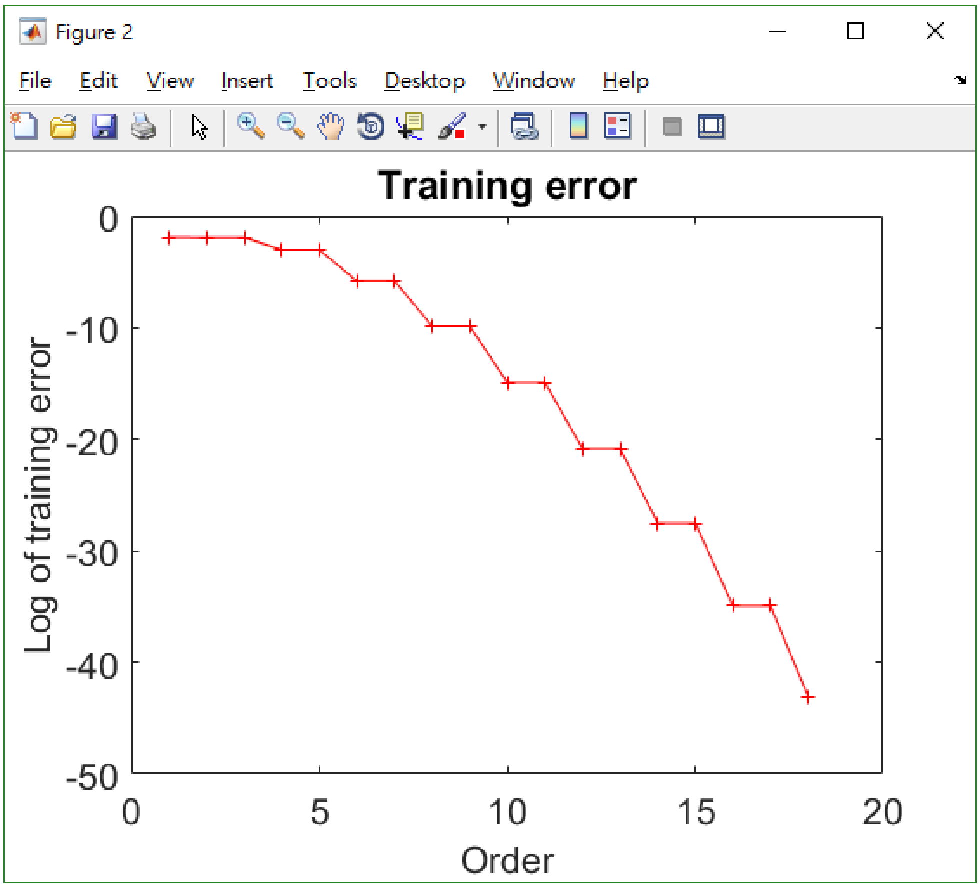
R05631027_楊皓文_機器學習_作業#3

Problem 1 (Overfitting)

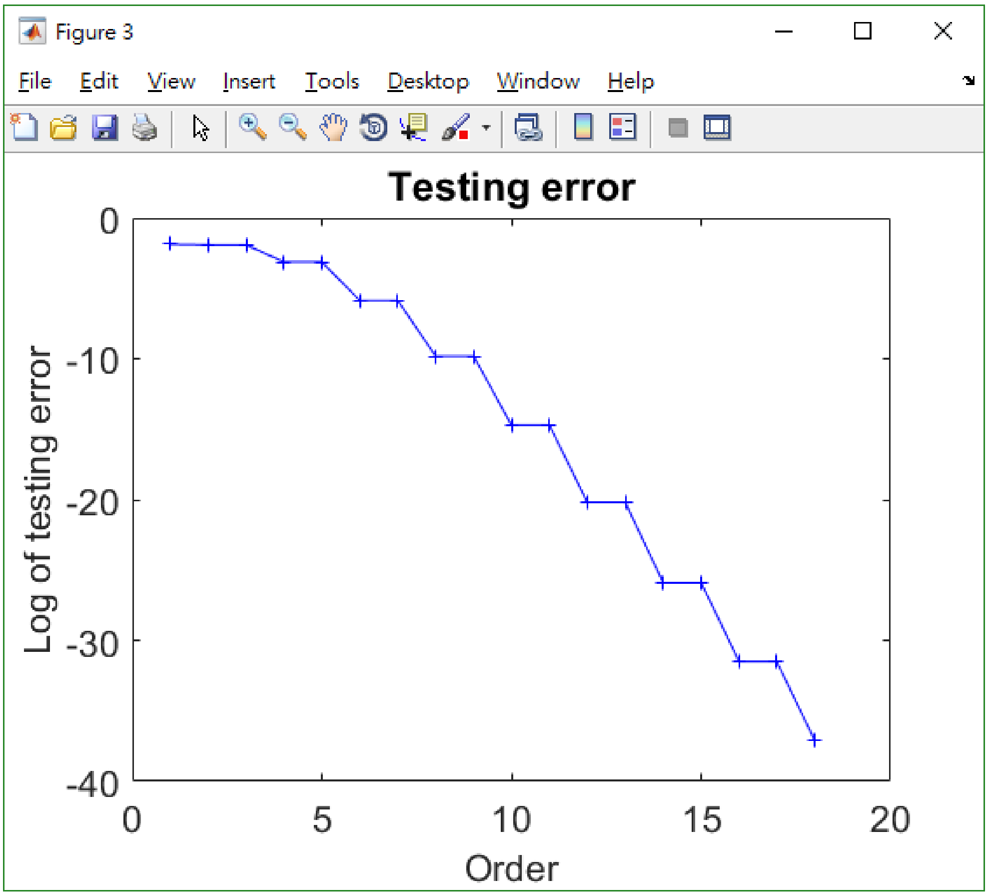
(a) Plot the 5 curves superimposed over a plot of the data points



(b) Log of training error



(c) Log of testing error

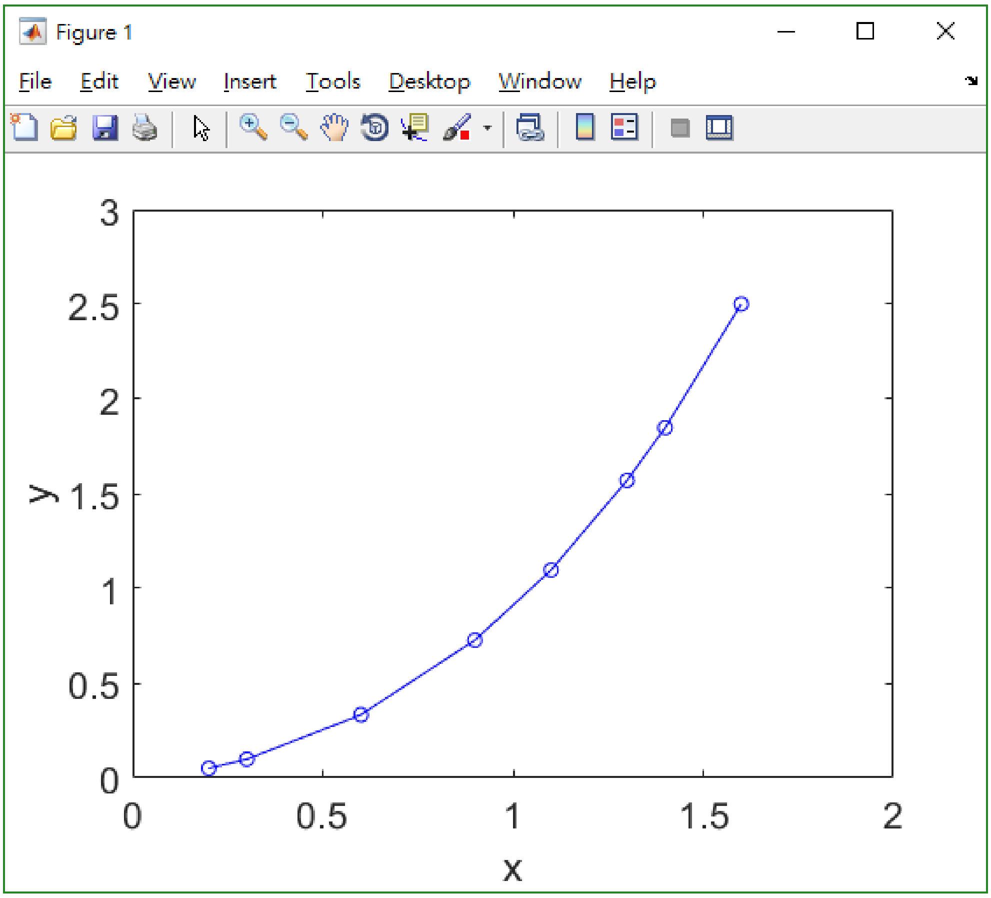


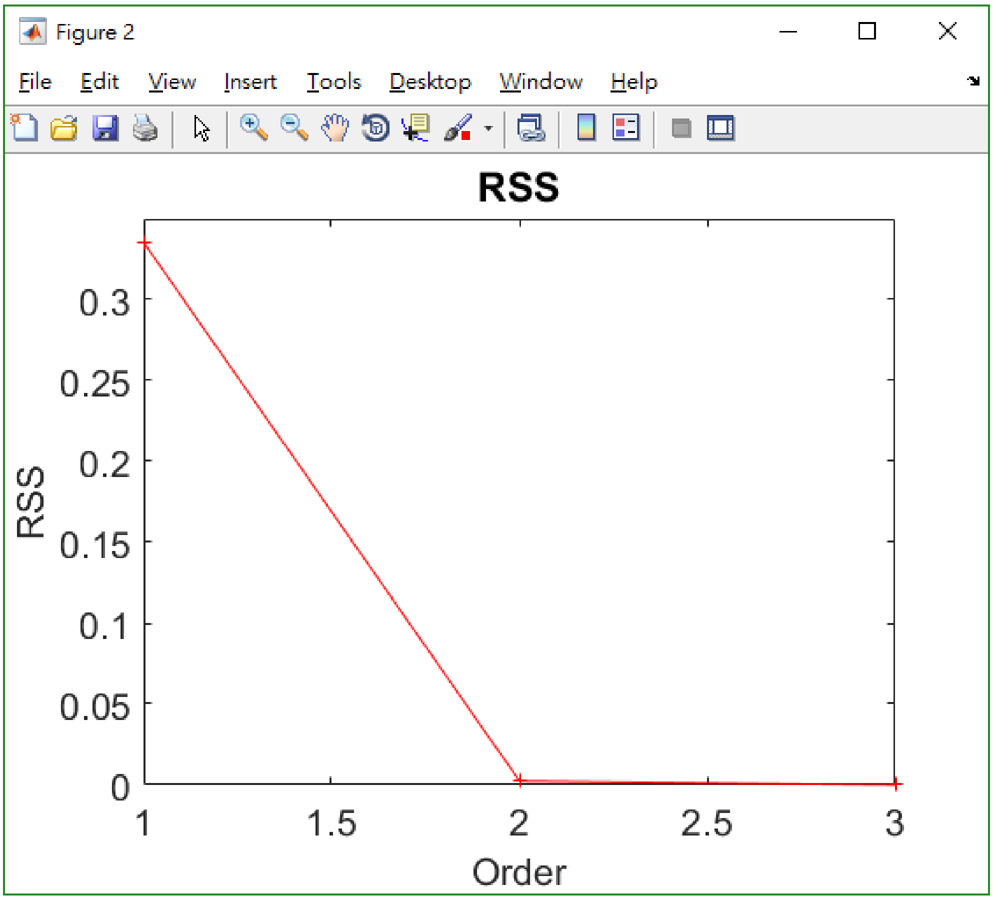
(d) Compare the training error with the test error

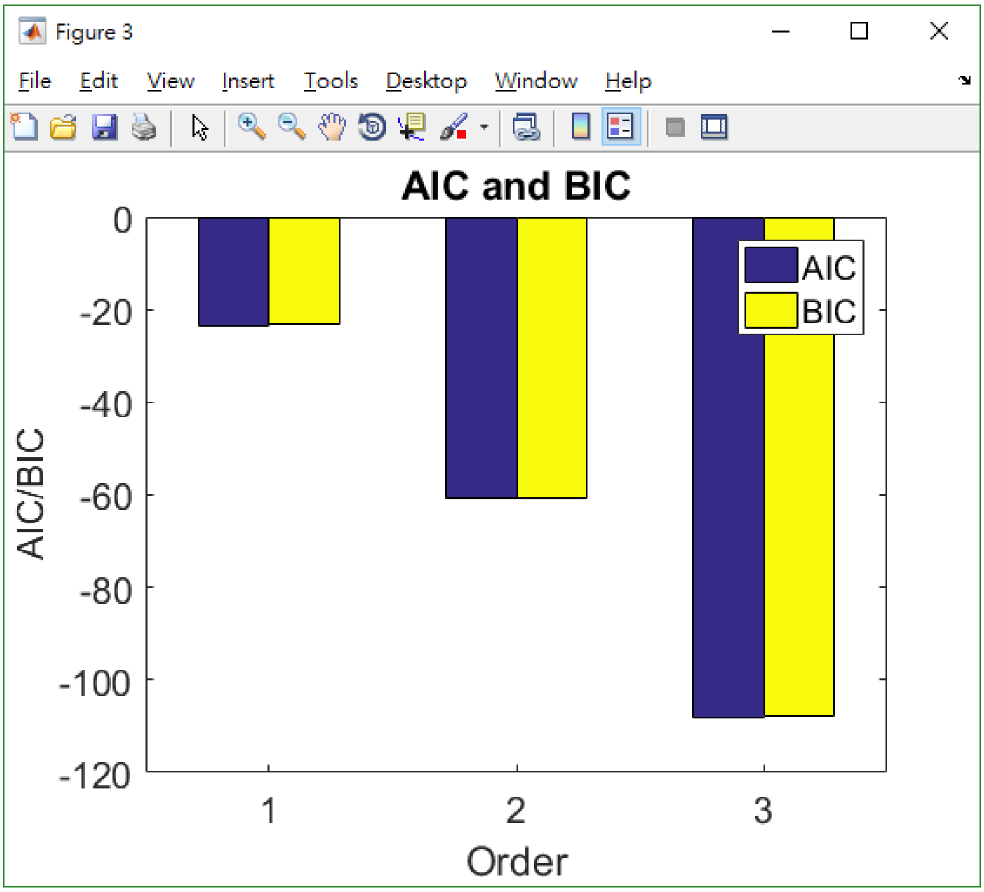
結果討論 ：

由上圖可以看出，當 order 數越高時，log of training error 下降的幅度越快，也就是 training accuracy 上升的比 testing accuracy 快，造成高 order 時，會有 overfitting 的現象。

Problem 2 (AIC and BIC)







結果討論 ：

若單純由上圖 AIC 和 BIC 的圖來看，似乎無法直接看出要選擇哪一個 order 的 model 最好，因為 order 的數量只有到 3，不過若是從 RSS 的曲線圖，可以看出 order=2 之後，RSS 的數值就變化不大了。

Problem 3 (Lagrange Multiplier)

(a)

r05631027 楊皓文 ML H.W. #3

(a) $\nabla f(x,y,z) = \langle 2(x+y+z), 2(x+y+z), 2(x+y+z) \rangle$

$\nabla g(x,y,z) = \langle 2x, 4y, 6z \rangle$

By Lagrange's Theorem, $\nabla f = \lambda \nabla g$

$\langle 2(x+y+z), 2(x+y+z), 2(x+y+z) \rangle = \lambda \langle 2x, 4y, 6z \rangle$

$$\begin{cases} \cancel{2}(x+y+z) = \cancel{2}\lambda x & \cdots ① \\ \cancel{2}(x+y+z) = \overset{2}{\cancel{4}}\lambda y & \cdots ② \\ \cancel{2}(x+y+z) = \overset{3}{\cancel{6}}\lambda z & \cdots ③ \\ x^2 + 2y^2 + 3z^2 = 1 & \cdots ④ \end{cases}$$

① − ② : $\lambda(x-2y) = 0 \Rightarrow \cancel{\lambda = 0}$ or $x = 2y$

② − ③ : $\lambda(2y-3z) = 0 \Rightarrow \cancel{\lambda = 0}$ or $2y = 3z$

④ : $(2y)^2 + 2y^2 + 3(\frac{2}{3}y)^2 = 1$

$\Rightarrow \frac{22}{3}y^2 = 1 \Rightarrow y = \pm\sqrt{\frac{3}{22}} \Rightarrow x = \pm 2\sqrt{\frac{3}{22}}, z = \pm\frac{2}{3}\sqrt{\frac{3}{22}}$

$f(\mp 2\sqrt{\frac{3}{22}}, \pm\sqrt{\frac{3}{22}}, \pm\frac{2}{3}\sqrt{\frac{3}{22}}) = (\mp\frac{1}{3}\sqrt{\frac{3}{22}})^2 = \frac{1}{66}$ #

(b)

(b) $\nabla f(x,y,z) = \langle y, x, 2z \rangle$

$\nabla g(x,y,z) = \langle 2x, 2y, 2z \rangle$

By Lagrange's Theorem, $\nabla f = \lambda \nabla g$

$\langle y, x, 2z \rangle = \lambda \langle 2x, 2y, 2z \rangle$

$$\begin{cases} y = 2\lambda x & \cdots ① \\ x = 2\lambda y & \cdots ② \\ 2z = 2\lambda z & \cdots ③ \\ x^2 + y^2 + z^2 = 1 & \cdots ④ \end{cases}$$

v culture

(i) ①·②: $\lambda \neq 0$, $x = y = 0$

③: $(\lambda - 1)z = 0 \Rightarrow z \neq 0$ or $\lambda = 1$

④: $z = \pm 1$

$f(0, 0, \pm 1) = 1$ (max)

(ii) ①+②: $\cancel{(x+y)} = 2\lambda \cancel{(x+y)} \Rightarrow \lambda = \frac{1}{2} \Rightarrow x = y$

③: $z = 0$

④: $2x^2 = 1 \Rightarrow x = \pm\sqrt{\frac{1}{2}}$, $y = \pm\sqrt{\frac{1}{2}}$

$f\left(\pm\sqrt{\frac{1}{2}}, \mp\sqrt{\frac{1}{2}}, 0\right) = -\frac{1}{2}$