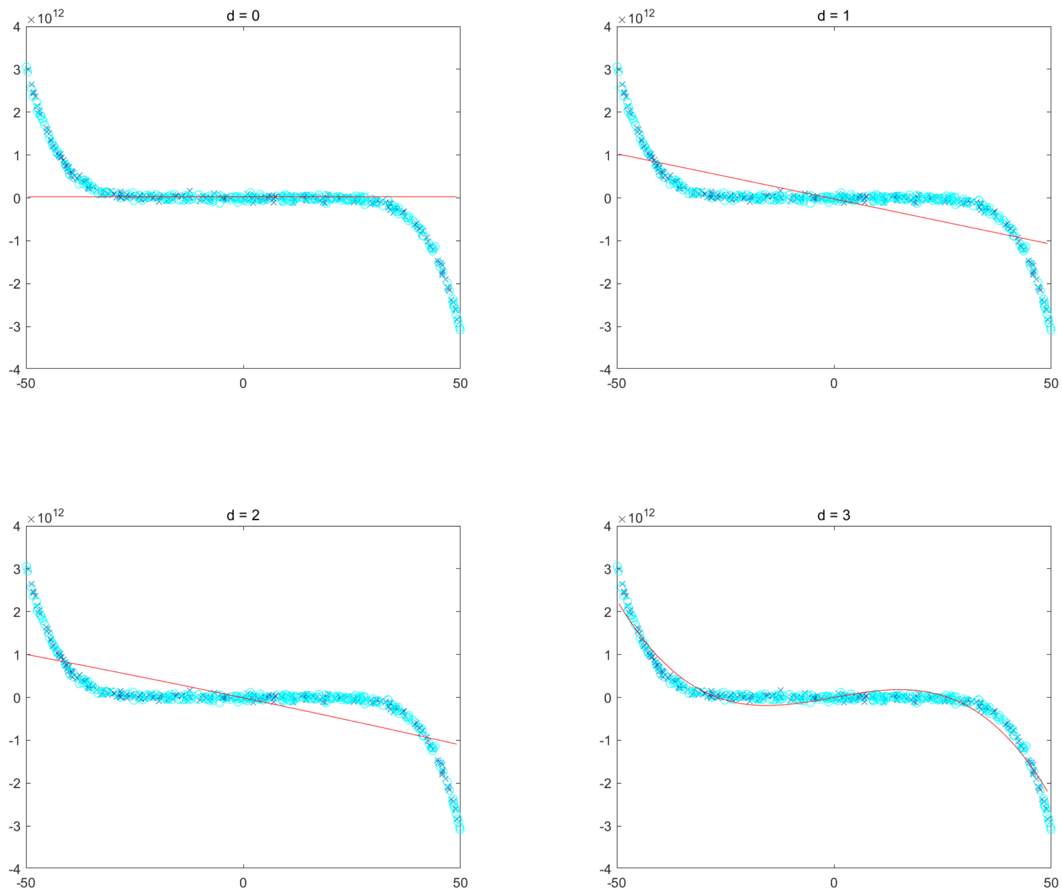
ECE 6143 Homework 1

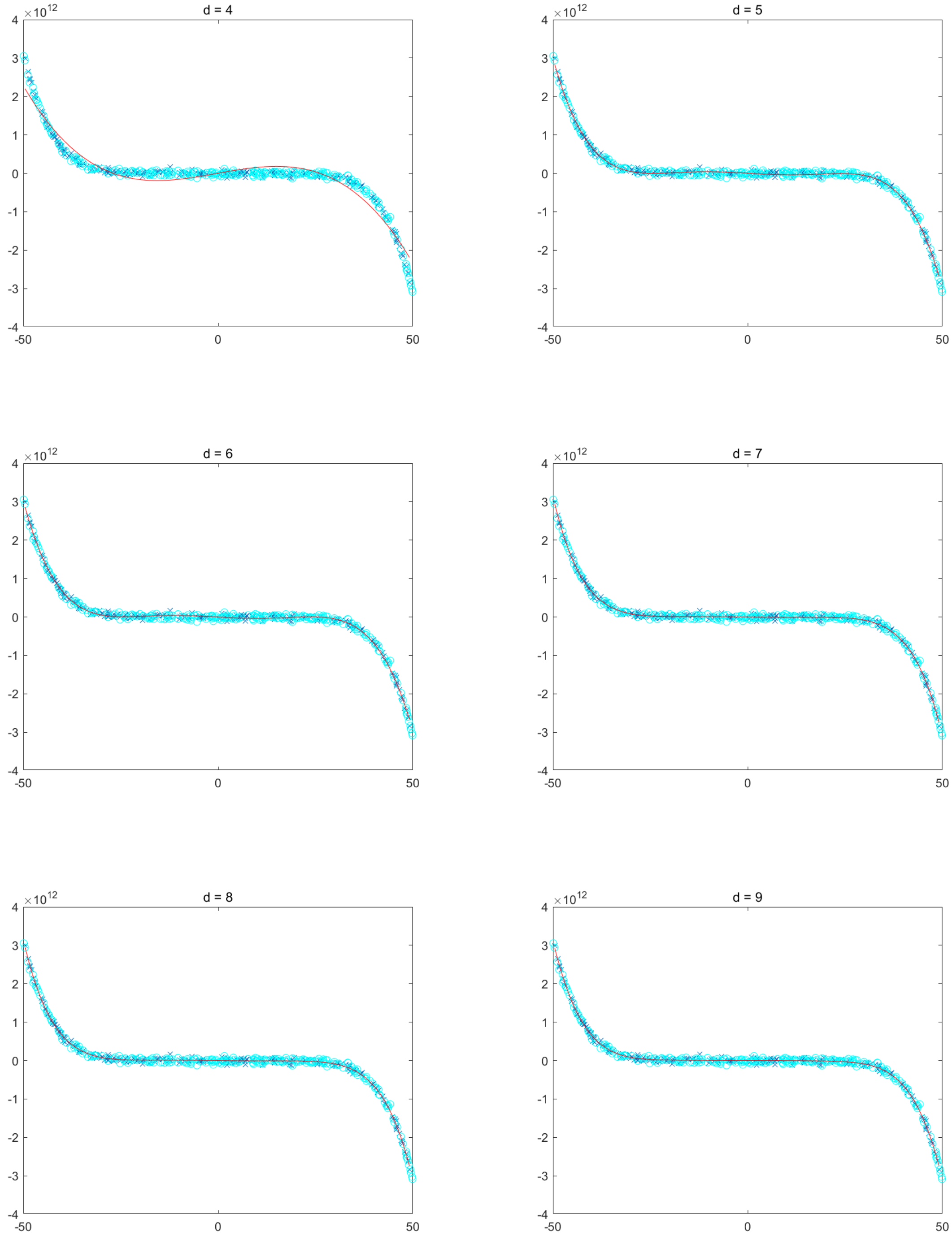
Tang Qianyong

September 14, 2019

1 Problem 1

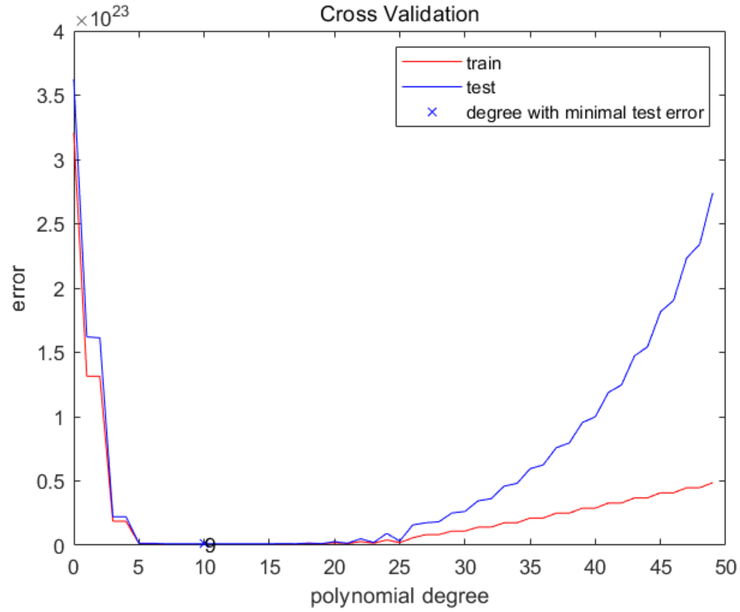
Using empirical measure, I choose degree of polynomial from 0 to 9. Applying two-fold cross validation, the regression plots are shown below. The degree of 9 seems reasonable since it has the smallest test error.



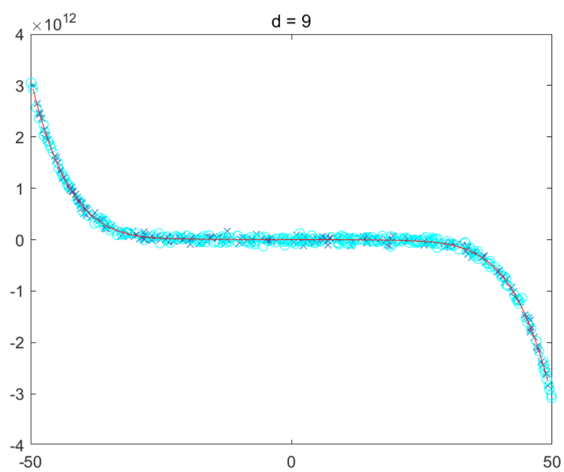


θ with different degrees

degree	θ
0	25245155745
1	[-21109992128;-25715357481]
2	[-17099436;-21133647921;-11154246673]
3	[-25711478;-15745480;18100077353;-863129976]
4	[5745;-25707895;-27942691;18087103266;2055959051]
5	[-16307;5934;19001008;-13640720;-5607084728;3283491171]
6	[17.02;-16267;-51231;18868934;33350838;-5526035561;-2055956173]
7	[-3.5804;9.4457;-1948.269;-26398;2694694;16695013;-1069440979;-2501583916]
8	[-0.00155;-3.5804;14.338;-1954.278;-27575.096;2713941;11080115;-1082317635;-6085965]
9	[0.00077;-0.00143;-7.775;13.0734;5708.342;-24298;-2610388;8814595;2182;52410]

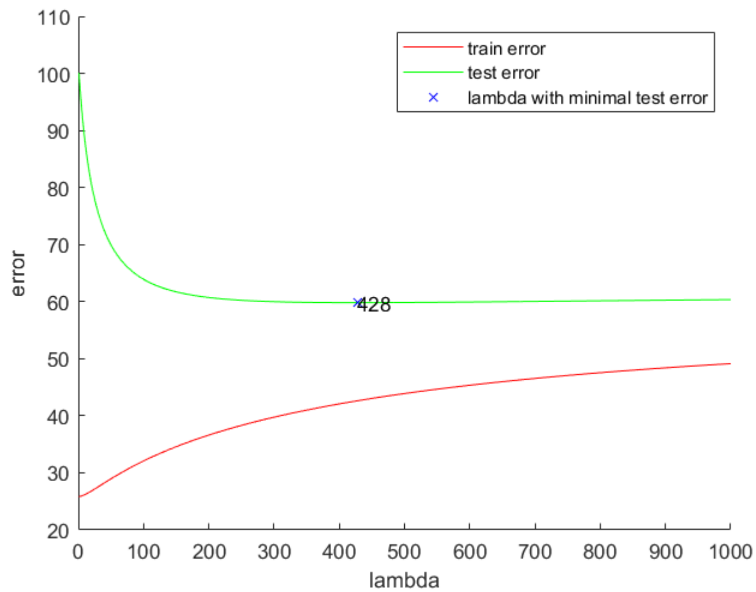


Picture above shows the cross validation plot. The test error is minimized at degree equals to 9.
The outcome may change based on how the data is split into training set and testing set. The best plot of $f(x;\theta)$ overlaid on data is shown below:



2 Problem 2

Using two-fold cross validation, we get the following plot:



As shown in the figure above, the best value for λ is 428, where the test error reaches minimum. The best value for lambda will change based on how the data is split.
It can be noticed that as λ increases, training error will increase while testing error will decrease.

3 Problem 3

3.1

$$g(z) = (1+\exp(-z))^{-1}$$

$$g(-z) = (1+(\exp(z))^{-1}$$

$$1-g(z) = 1-\frac{1}{1+\exp(-z)} = \frac{1+\exp(-z)-1}{1+\exp(-z)} = \frac{\exp(-z)}{1+\exp(-z)}$$

$$= \frac{1}{1+\exp(z)} = g(-z)$$

proved

3.2

if $g(z) = (1+\exp(-z))^{-1}$, show that $g^{-1}(y) = \ln(y/(1-y))$

assume $g^{-1}(y) = \ln \frac{y}{1-y}$

then $g^{-1}(g(z)) = \ln \dfrac{g(z)}{1-g(z)} = \ln \dfrac{g(z)}{g(-z)}$

$$= \ln \frac{1+\exp(z)}{1+\exp(-z)} = \ln(1+e^z) - \ln(e^z+1) + \ln(e^z)$$

$$= z$$

so proved

4 Problem 4

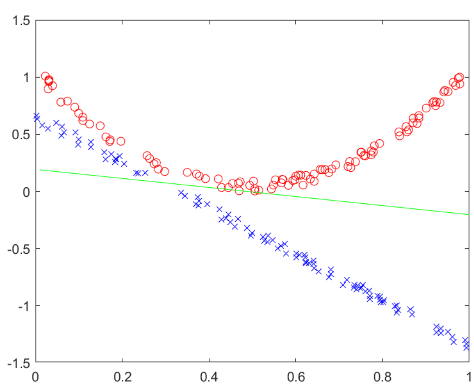
denote $f_i = f(x_i, \theta) = \frac{1}{1+e^{-\theta^T x}}$
$R = \frac{1}{N}\sum_{i=1}^{N}(y_i - 1)\log(1-f_i) - y_i \log(f_i)$

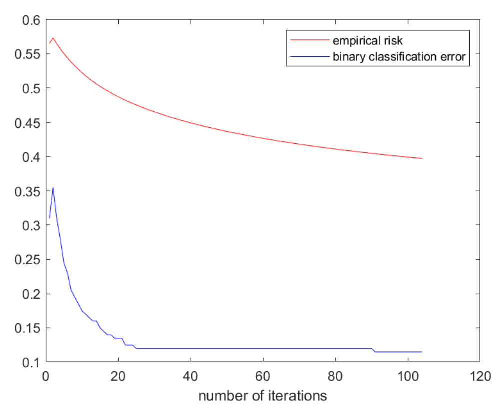
$$\nabla_\theta R = \nabla_\theta \left(\frac{1}{N}\sum_{i=1}^{N} (y_i-1)\log(1-f_i) - y_i \log(f_i) \right)$$

$$= \frac{1}{N}\sum_{i=1}^{N}\left[(y_i-1)\frac{d}{d\theta}\log\left(1-(1+e^{-\theta^T x_i})^{-1}\right) - y_i \frac{d}{d\theta}\log\left(1+e^{-\theta^T x_i}\right)^{-1}\right.$$

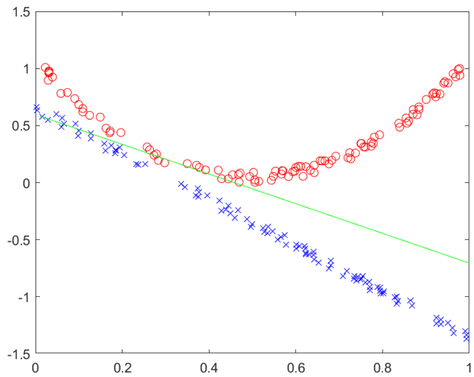
$$= \frac{1}{N}\sum_{i=1}^{N} (y_i-1)\left(-x_i + \frac{x_i e^{-\theta^T x_i}}{1+e^{-\theta^T x_i}}\right) - y_i \left(\frac{x_i e^{-\theta^T x_i}}{1+e^{-\theta^T x_i}}\right)$$

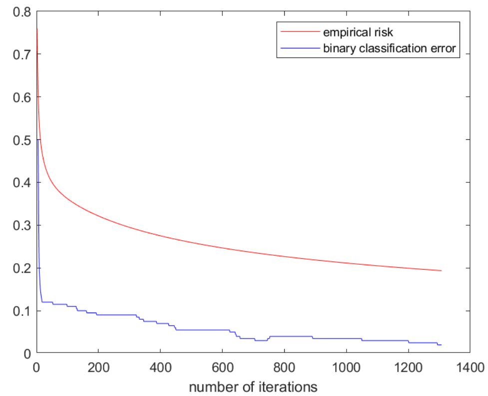
For $\epsilon = 0.01$ and $\eta = 0.2$, $\theta = [0.9642; 2.4353; -0.4690]$





For $\epsilon = 0.005$ and $\eta = 0.5$, $\theta = [9.2952; 7.1489; -4.2585]$





For $\epsilon = 0.0001$ and $\eta = 0.2$, $\theta = [17.4863; 11.2519; -7.5186]$