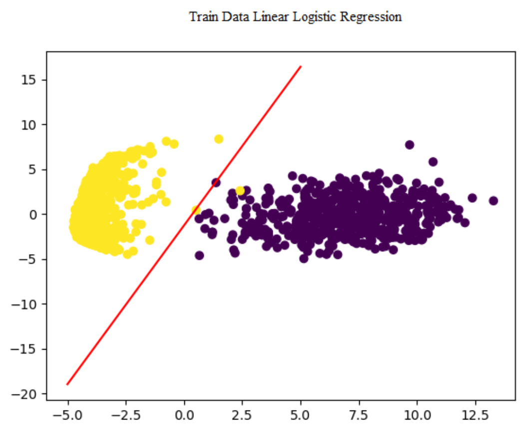
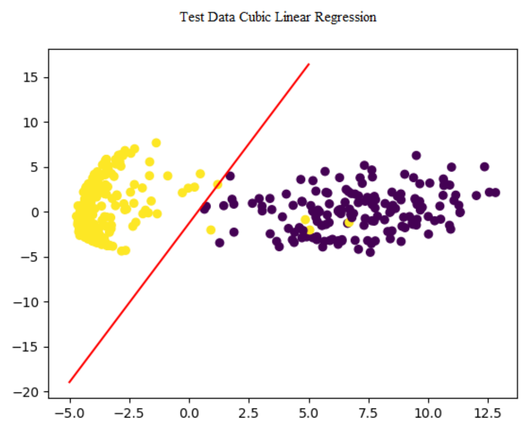
Homework 7

Karan Sarkar
sarkak2@rpi.edu

October 23, 2019

Problem 1. We chose to use logistic regression with gradient descent.





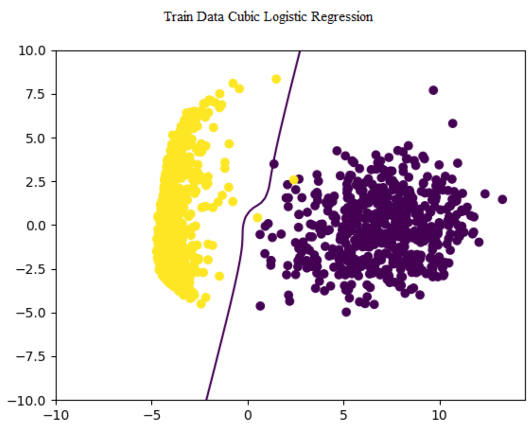
(a)

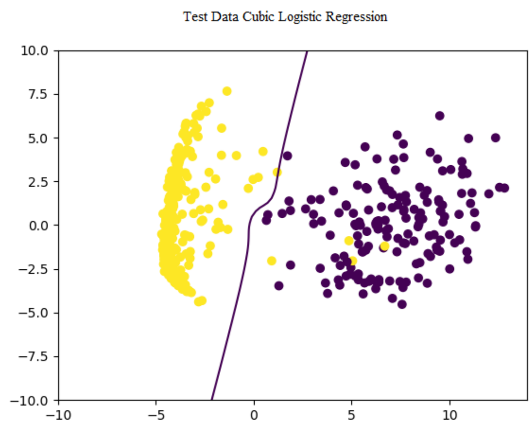
(b) We computed $E_{\text{in}} = 0.0012$ and $E_{\text{test}} = 0.012$.

(c) We compute two bounds on E_{out}.

$$\begin{aligned}
E_{\text{out}} &\leq E_{\text{in}} + \sqrt{\frac{8}{N} \ln \frac{4 m_{\mathcal{H}}(N)}{\delta}} \\
&\leq 0.0012 + \sqrt{\frac{8}{1561} \ln \frac{4 \cdot 1561^3}{0.05}} \\
&\leq 0.369 \\
E_{\text{out}} &\leq E_{\text{test}} + \sqrt{\frac{1}{2N} \ln \frac{2}{\delta}} \\
&\leq 0.012 + \sqrt{\frac{1}{2 * 1561} \ln \frac{2}{0.05}} \\
&\leq 0.0463
\end{aligned}$$

The bound from the test error is much better.





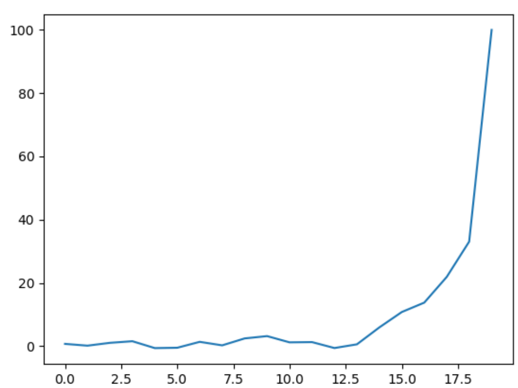
(d) We computed $E_{\text{in}} = 0.0013$ and $E_{\text{test}} = 0.009$. We compute two bounds on E_{out}.

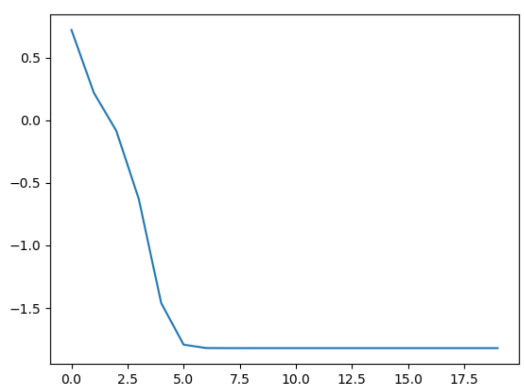
$$
\begin{aligned}
E_{\text{out}} &\leq E_{\text{in}} + \sqrt{\frac{8}{N} \ln \frac{4 m_{\mathcal{H}}(N)}{\delta}} \\
&\leq 0.0012 + \sqrt{\frac{8}{1561} \ln \frac{4 \cdot 1561^9}{0.05}} \\
&\leq 0.603 \\
E_{\text{out}} &\leq E_{\text{test}} + \sqrt{\frac{1}{2N} \ln \frac{2}{\delta}} \\
&\leq 0.009 + \sqrt{\frac{1}{2 * 1561} \ln \frac{2}{0.05}} \\
&\leq 0.0433
\end{aligned}
$$

The bound from the test error is much better.

(e) I would use the linear model bound without the transform because it gives a tighter bound on the out of sample error.

Problem 2.





Here are the two plots for learning rate at 0.1 and 0.01 respectively. For 0.1 it diverges to infinity. For 0.01 it converges. This is a table of calculated 'minimums'.

x	y	η	$x_{\min}$	$y_{\min}$	$f_{\min}$
0.1	0.1	0.1	-3.13e59	-9.82e117	1.92e236
0.1	0.1	0.01	0.245	0.237	-1.82
1	1	0.1	-0.931	0.094	1.36
1	1	0.01	0.689	1.18	1.58
-0.5	-0.5	0.1	-1.48	-0.677	2.96
-0.5	-0.5	0.01	-0.733	-0.244	-1.33
-1	-1	0.1	-3.23e54	1.04e108	2.19e216
-1	-1	0.01	-1.232	0.732	0.615

Problem 3.16.

(a) The expected cost when we accept someone is $0 \cdot P[y = 1|x] + c_a P[y = -1|x] = c_a(1 - g(x))$. The expected cost when we reject someone is $c_r \cdot P[y = 1|x] + 0 \cdot P[y = -1|x] = c_r g(x)$.

(b) We decide to accept someone when the cost to accept is less than the cost to reject. Thus, the threshold occurs at $c_a(1 - g(x)) = c_r g(x)$ when the costs are equal. This translates to $g(x) = \frac{c_a}{c_a+c_r} = \kappa$.

(c) For the supermarket, we have $\kappa = \frac{1}{11}$. This is logical as we want to avoid false negatives thus we only reject when $g(x)$ is very small. For the CIA, we have $\kappa = \frac{1000}{1001}$. This is logical as we want to avoid false positives thus we accept when $g(x)$ is very high.