Machine Learning (ELL409)

Assignment-2

Hand Written Digit Recognition

Submitted on : 10th April, 2016

Santosh Nain (2012EE10476)

Sachin Rinder (2013EE10489)

Ashutosh Kumar (2013EE30550)

A) Logistic Regression :

Cost function:

$$E(w) = \frac{1}{m}\sum_{i=1}^{m}\left[-y^{(i)}\log(h_w(x^{(i)})) - (1-y^{(i)})\log(1-h_w(x^{(i)}))\right] + \frac{\lambda}{2m}\sum_{j=1}^{N} w_j^2.$$

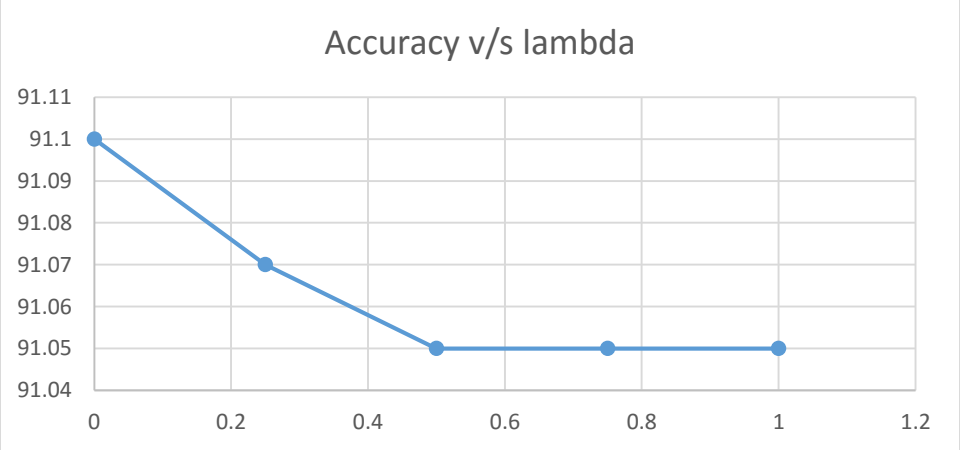
Gradient:

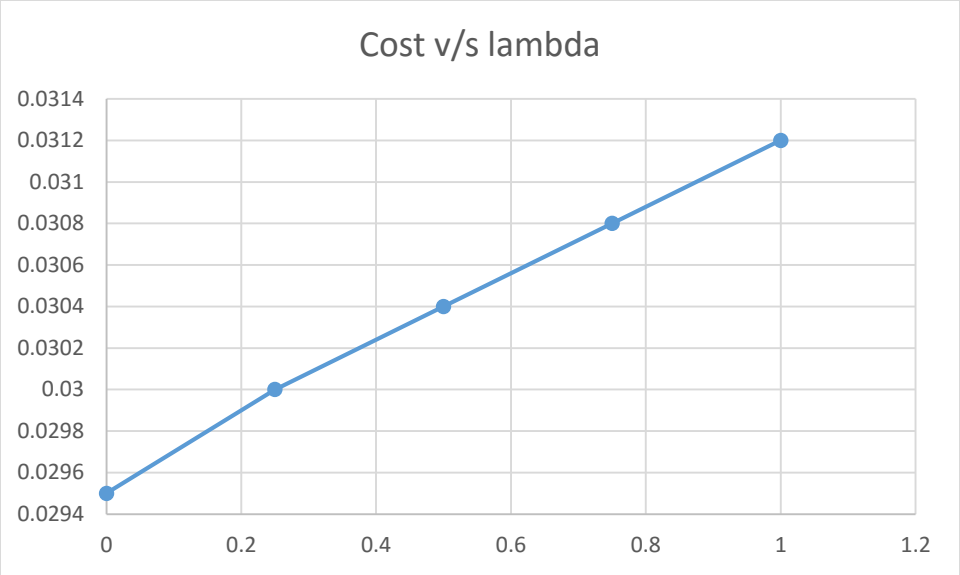
$$J(\theta) = \frac{1}{2m}\left[\sum_{i=1}^{m}(h_\theta(x^{(i)}) - y^{(i)})^2 + \lambda\sum_{j=1}^{n}\theta_j^2\right]$$

where λ is the regularization parameter

When alpha (learning rate) = 0.5, m (number of training data) = 4000

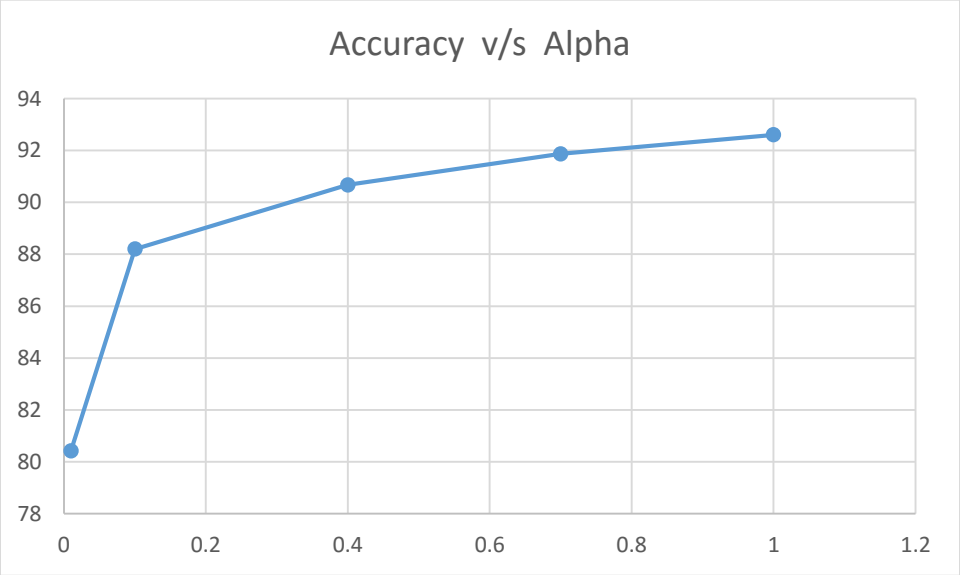
Lambda	Training Accuracy	Cost for Class zero
0	91.1	0.0295
0.25	91.07	0.0300
0.50	91.05	0.0304
0.75	91.05	0.0308
1	91.05	0.0312

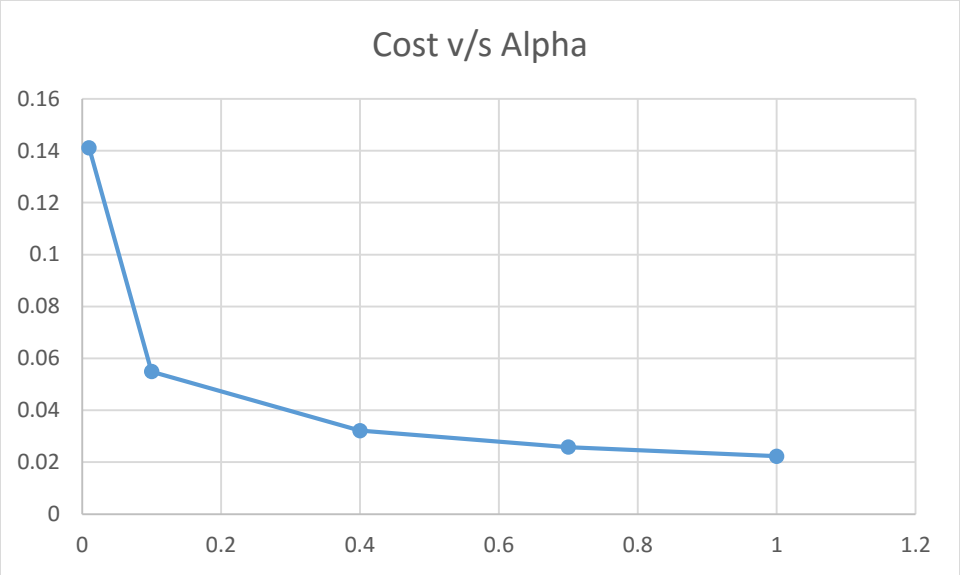




When lambda = 0, m (number of training data) = 4000

Alpha	**Training Accuracy**	**Cost for Class zero**
0.01	80.42	0.1412
0.1	88.2	0.0549
0.4	90.67	0.0322
0.7	91.87	0.0258
1	92.6	0.0223





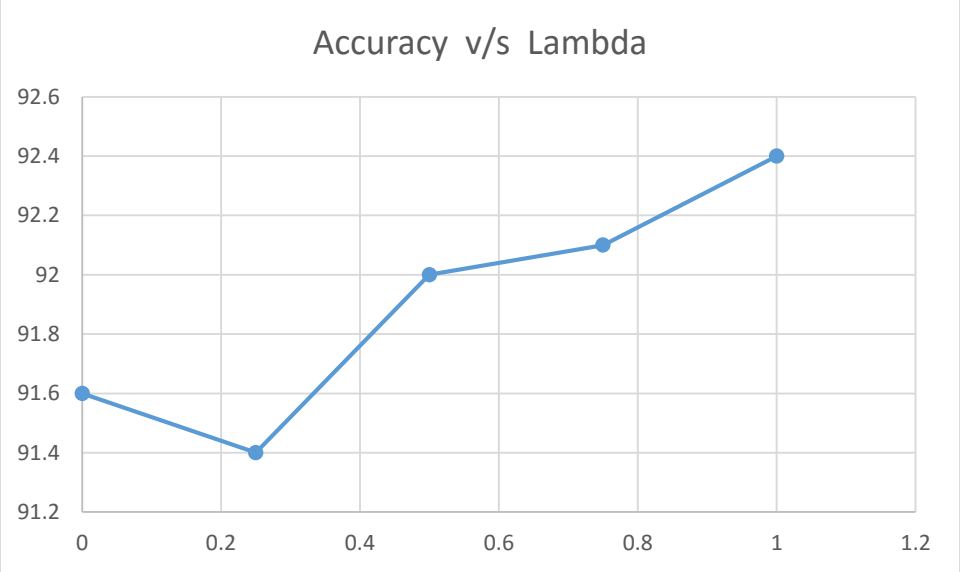
Conclusion:

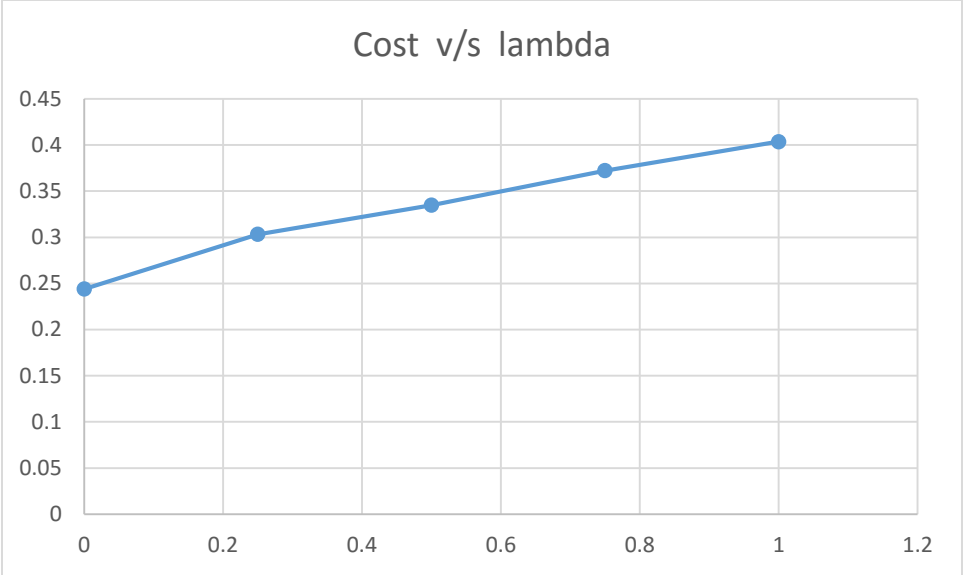
1) Training accuracy decreases when we increase lambda and finally gets saturated at approx. 91.05, considering no of training data set is 4000.
2) Cost increases with the lambda linearly, as expected from the cost equation.
3) Also training Accuracy increases linearly when varying learning rate and finally converges to 92.04.
4) And Cost decreases (seems to be exponentially) as we increase alpha.

B) Neural Networks

When alpha (learning rate) = 1, m (number of training data) = 4000
(After 1000 iterations)

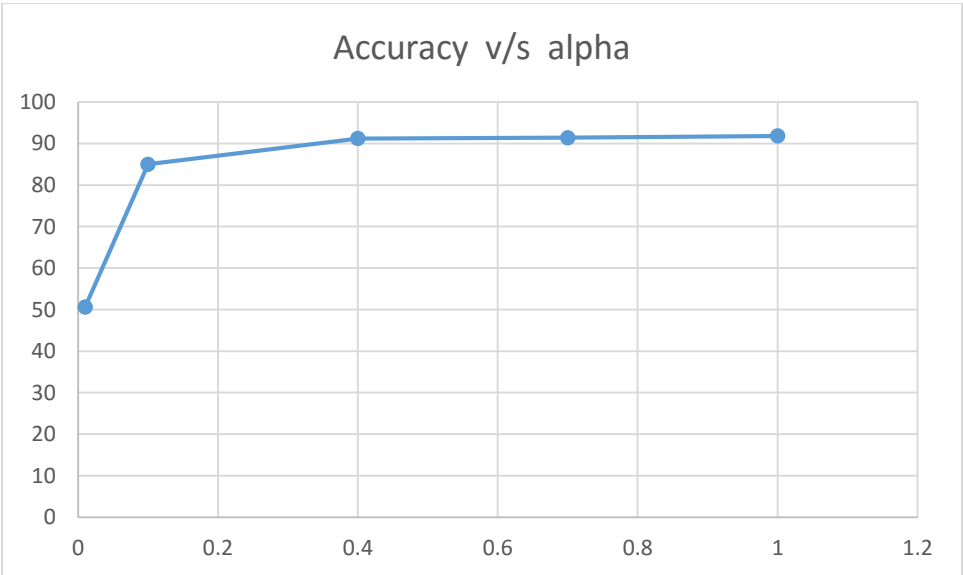
Lambda	Training Accuracy	Cost for Class zero
0	91.6	0.2440
0.25	91.4	0.3033
0.50	92	0.3348
0.75	92.1	0.3722
1	92.4	0.4036

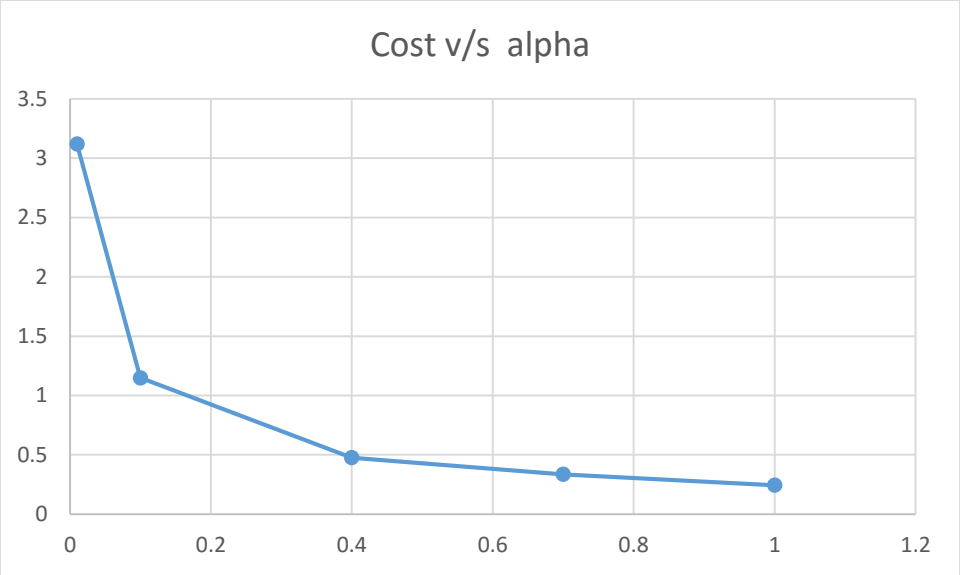




When lambda = 0, m (number of training data) = 4000

Alpha	**Training Accuracy**	**Cost for Class zero**
0.01	50.6	3.1204
0.1	85	1.1488
0.4	91.2	0.4756
0.7	91.4	0.3354
1	91.8	0.2434





CONCLUSIONS :

1) We can see here that initially when we increase lambda, accuracy has a dip but later it increases.
2) Cost varies linearly with lambda.
3) Accuracy seems to be getting saturated when we increase alpha.
4) And Cost decreases (seems to be exponentially) as we increase alpha.