COL 774: Machine Learning Assignment 1

Saurabh Godse [2018MCS2019]

Due date: February 12, 2019, 11:59pm IST

1 Gradient Descent

- Learning Rate: 0.01.
- Value of θ_0 and θ_1 are [0.9963085 0.00133978] respectively.
- The Stopping Criteria is the difference in the cost function of 10^{-9}.

LearningRate	TotalIterations	Overshoots	Converges
0.1	89	No	Yes
0.5	16	No	Yes
0.9	6	No	Yes
1.3	10	Yes	Yes
1.7	29	Yes	Yes
2.1	Nan	Yes	No
2.5	Nan	Yes	No

Table 1: Learning Rates and Observation

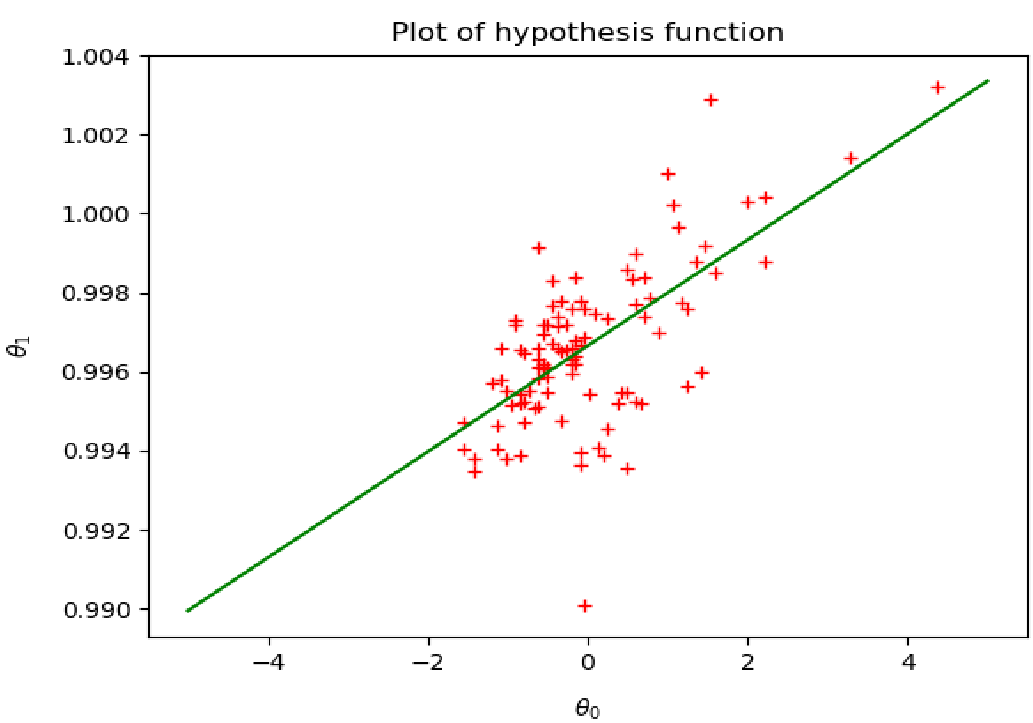

Figure 1: Gradient Descent

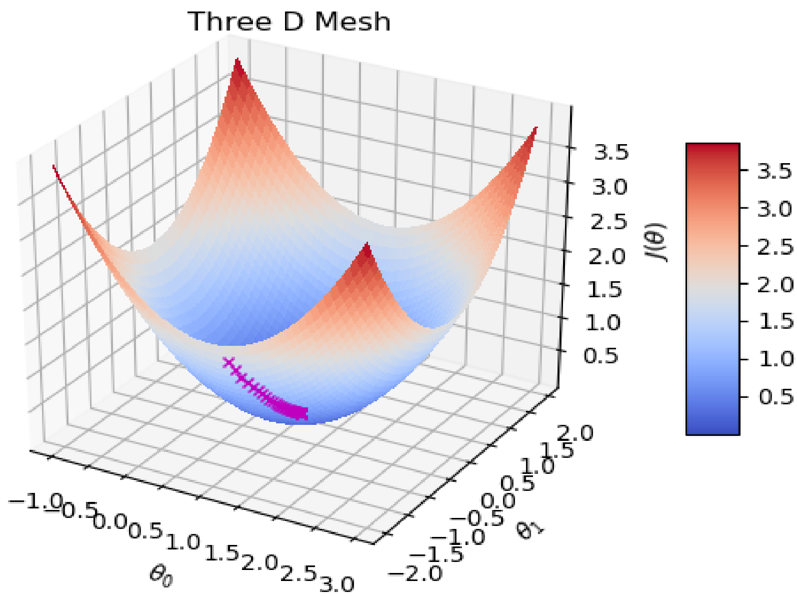

Figure 2: 3d mesh plot of Costfunction

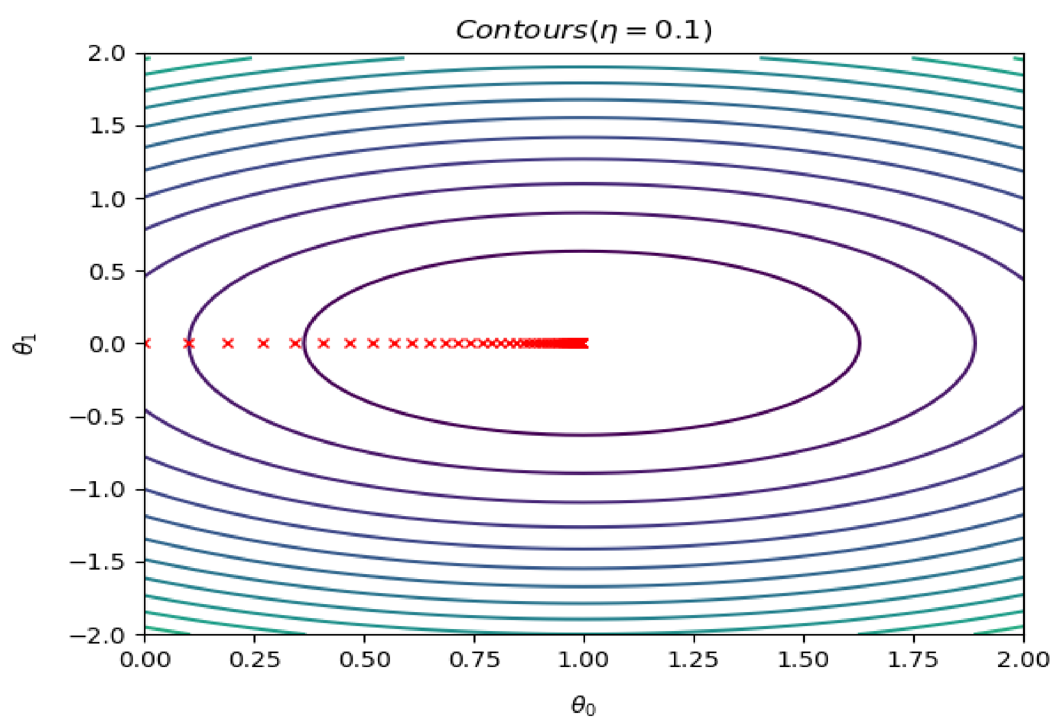

Figure 3: Contour at eta = 0.1

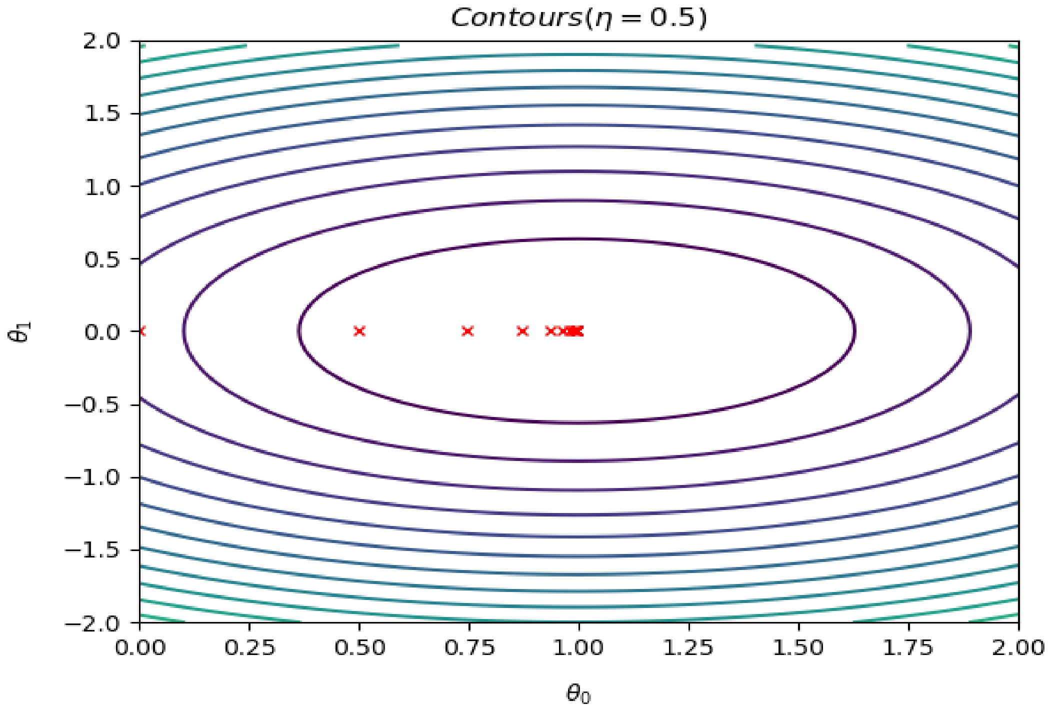

Figure 4: Contour at eta = 0.5

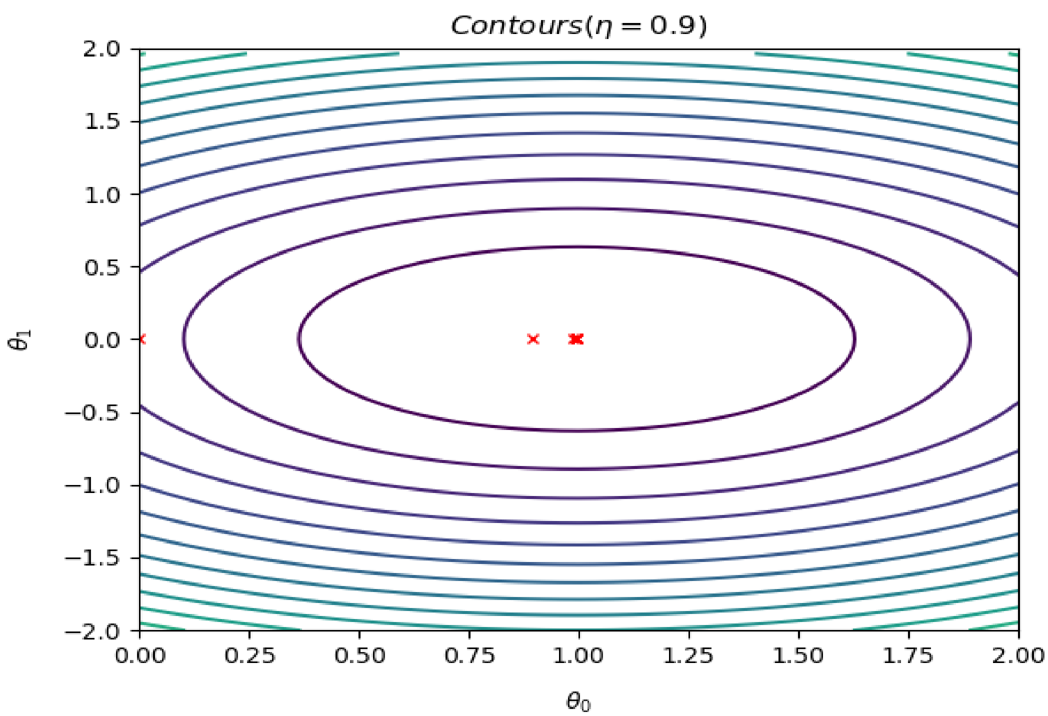

Figure 5: Contour at eta = 0.9

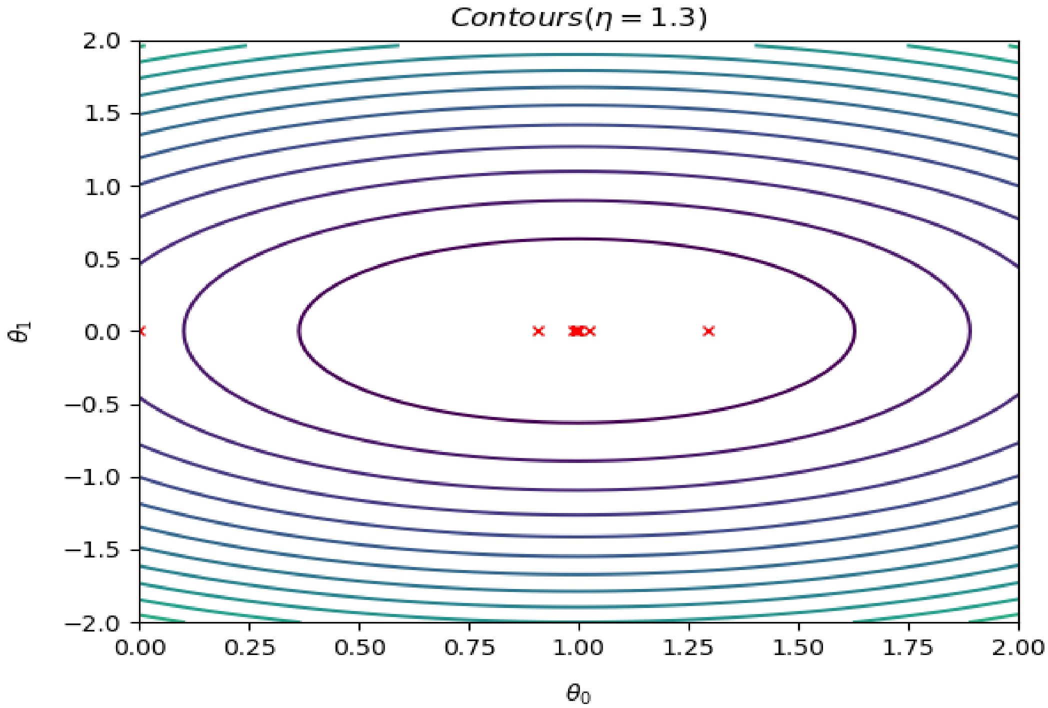

Figure 6: Contour at eta = 1.3

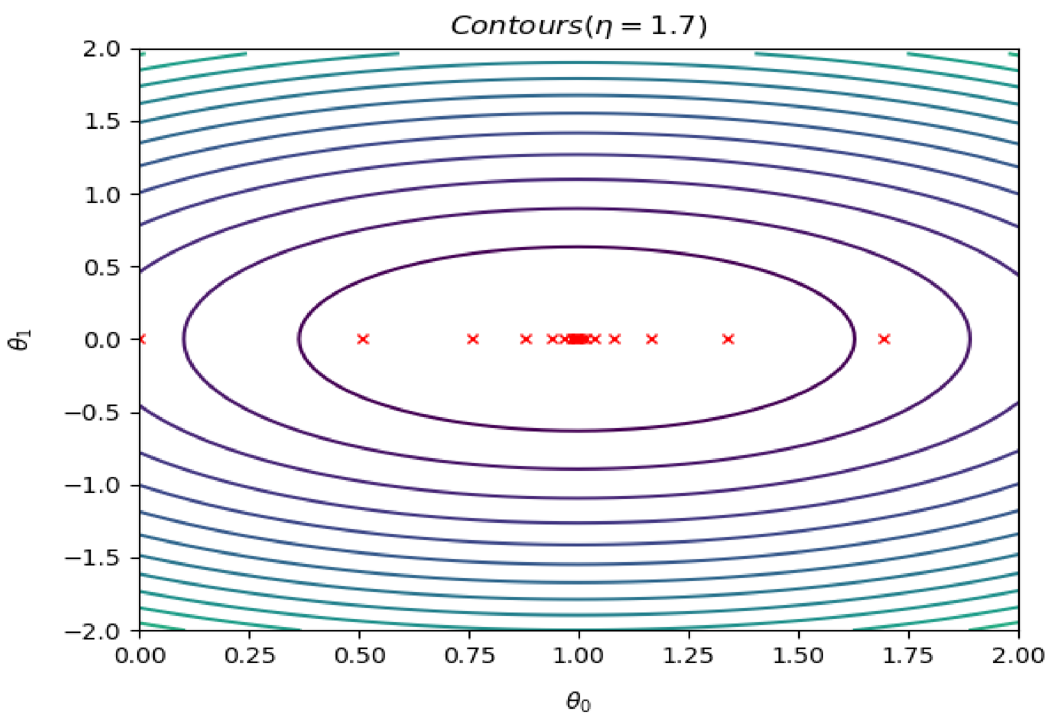

Figure 7: Contour at eta = 1.7

2 Locally Weighted Linear Regression

$$J(\theta) = \frac{1}{2m}(X\theta - Y)^T W(X\theta - Y)$$

$$On\ solving\ for\ \frac{dJ(\theta)}{d\theta} = 0,\ we\ get\ \theta = (X^T W X)^{-1} X^T W Y$$

- Best Value of tau should be 0.3 because it generalizes the data well.
- Larger the value of tau it underfits the data or produces a linear fit to the data.
- Smaller value of tau will lead to overfitting issues.

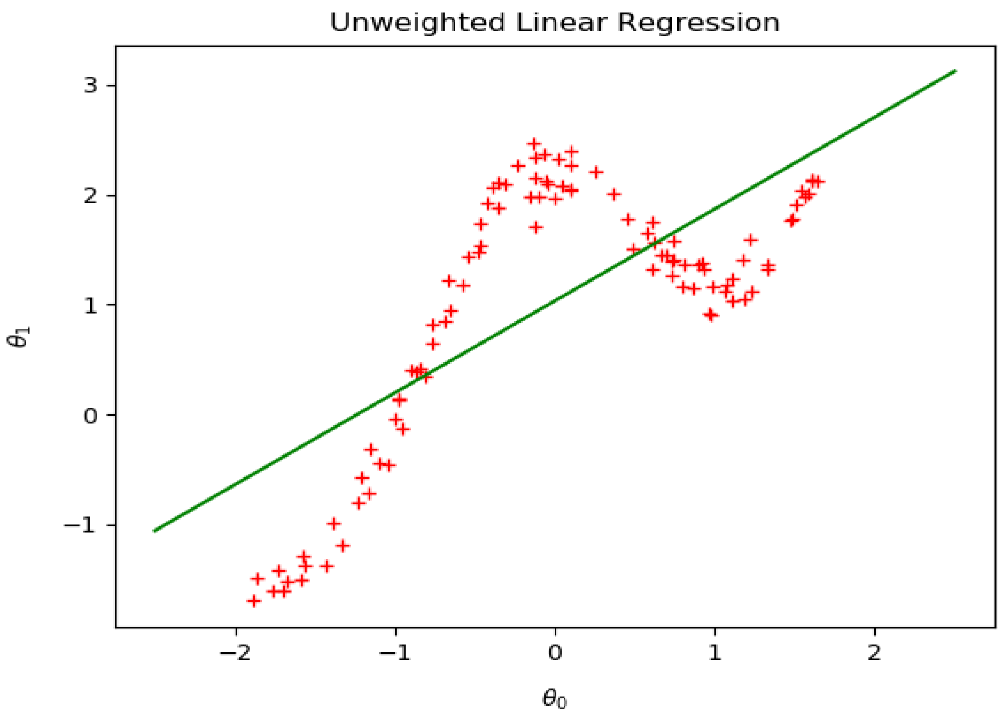

Figure 8: Analytical Solution

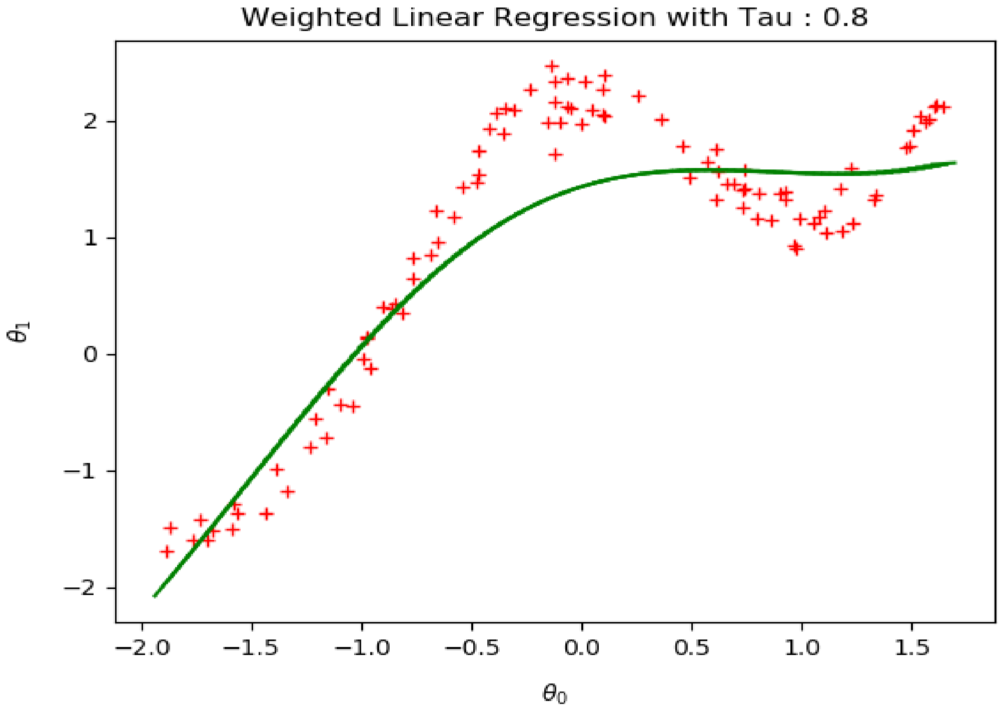

Figure 9: Locally Weighted Solution at tau = 0.8

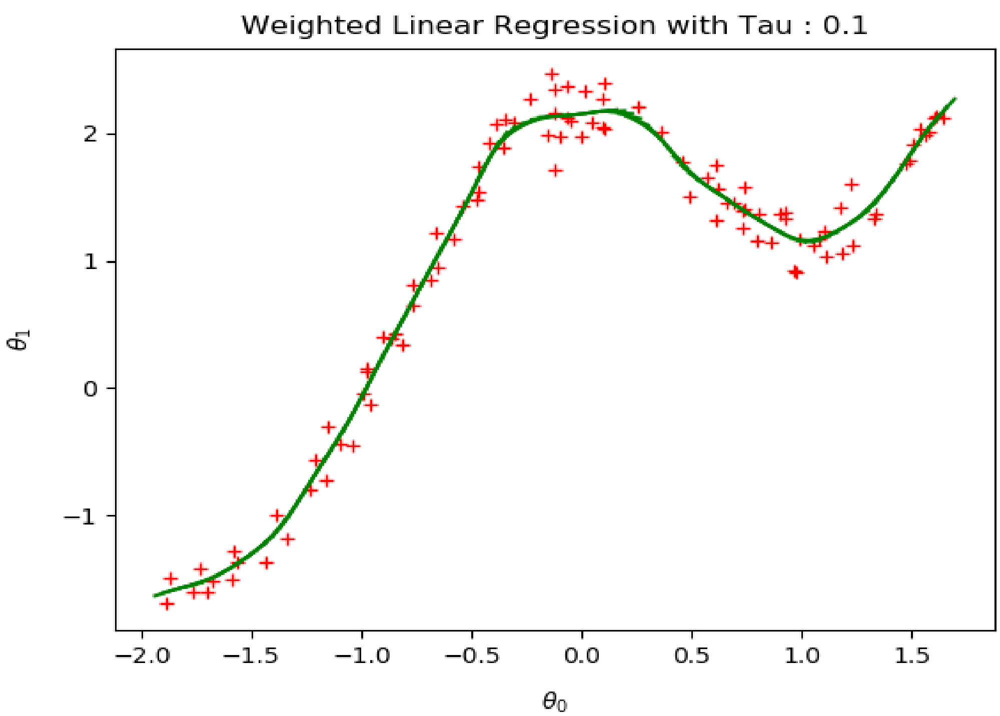

Figure 10: Locally Weighted Solution at tau = 0.1

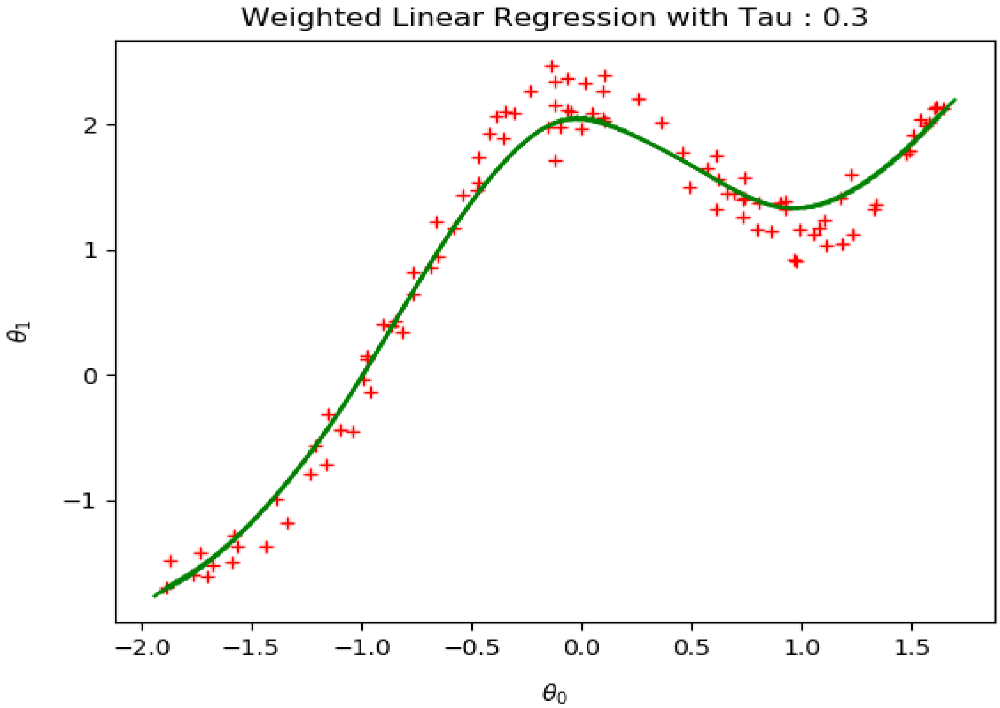

Figure 11: Locally Weighted Solution at tau = 0.3

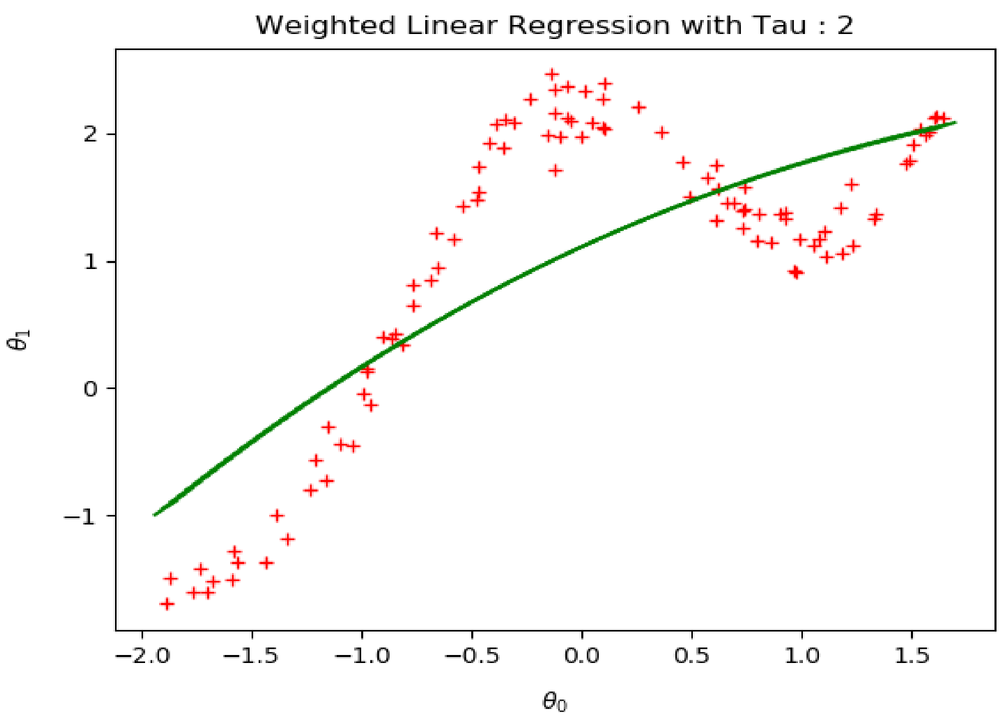

Figure 12: Locally Weighted Solution at tau = 2

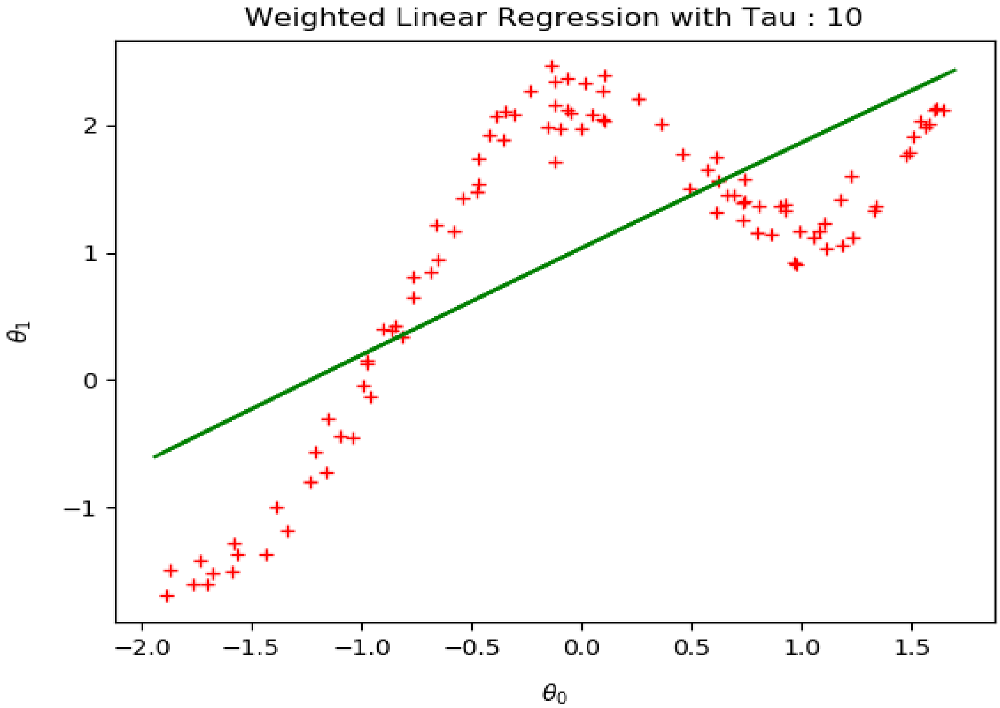

Figure 13: Locally Weighted Solution at tau = 10

3 Logistic Regression

$$Hessian = -X^T(g(\theta^T X)(1 - g(\theta^T X))X$$

$$Gradient = X^T Y - X^T g(\theta^T X))$$

On Applying Newton's Method we get

$$\theta^{t+1} = \theta^t - \frac{Gradient}{Hessian}$$

- Value of $\theta_0, \theta_1, \theta_2$ are [0.22329537 1.96261552 -1.9648612] respectively.

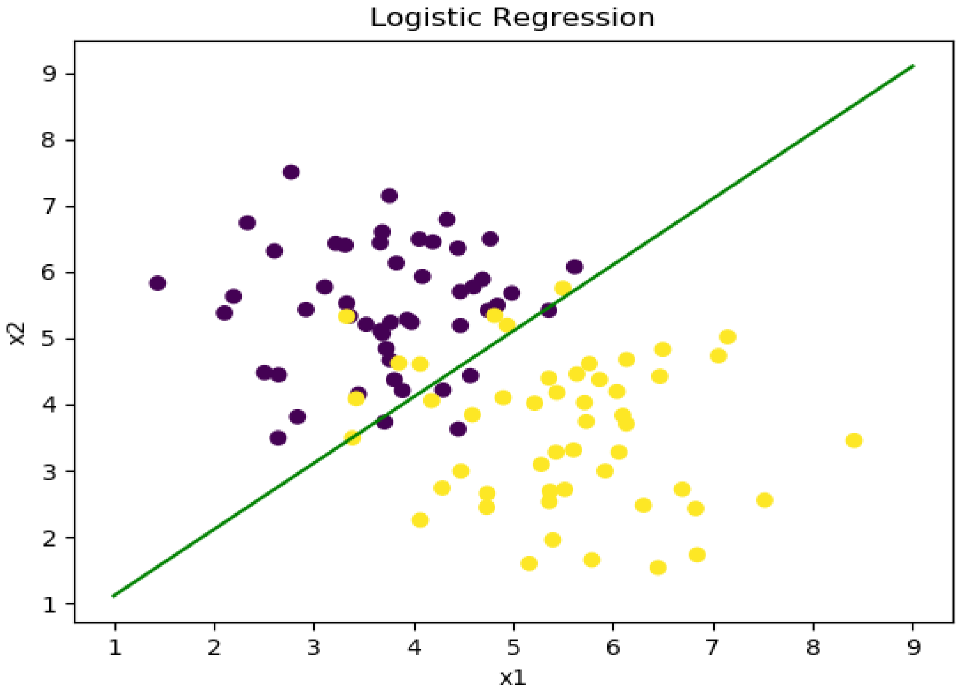

Figure 14: Logistic Regression with Newton's Method

4 Gaussian Discriminant Analysis

4.1 In case of Same Covariance

- Covariance

Σ

28748.2	-2674.8
-2674.8	112325.

- Mean of Alaska

μ_0

98.38	429.66

- Mean of Canada

μ_1

137.46	366.62

4.2 In case of different Covariance

E0 : [[255.3956 -184.3308] [-184.3308 1371.1044]] E1 : [[319.5684 130.8348] [130.8348 875.3956]]

- Covariance of Alaska

Σ_0

255.3956	-184.3308
-184.3308	1371.1044

- Covariance of Canada

Σ_1

319.5684	130.8348
130.8348	875.3956

- Mean of Alaska

μ_0

98.38	429.66

- Mean of Canada

μ_1

137.46	366.62

4.3 Equation of the boundary

$$P(y=1|x;\theta) = P(y=0|x;\theta)$$

$$P(x|y=1;\theta)P(y=1;\phi) = P(x|y=0;\theta)P(y=0;\phi)$$

Solving the above equation yields

$$\frac{1}{2}[(X-\mu_1)^T\Sigma_1^{-1}(X-\mu_1) - (X-\mu_0)^T\Sigma_0^{-1}(X-\mu_0)] = log(\frac{\phi}{1-\phi}) - log(\frac{|\Sigma_1|^{\frac{-1}{2}}}{|\Sigma_0|^{\frac{-1}{2}}})$$

$\Sigma_0 = \Sigma_1 = \Sigma$ Same Covariance

We obtain a linear boundary whose equation is given by

$$X^T\Sigma^{-1}\mu_0 - X^T\Sigma^{-1}\mu_1 + \mu_1^T\Sigma^{-1}\mu_1 - \mu_0^T\Sigma^{-1}\mu_0 - log(\frac{\phi}{1-\phi}) = 0$$

$\Sigma_0 \neq \Sigma_1$ Different Covariance

We obtain a Quadratic boundary whose equation is given by

$$X^T\Sigma_0^{-1}X - X^T\Sigma_1^{-1}X + X^T\Sigma_0^{-1}\mu_0 - X^T\Sigma_1^{-1}\mu_1 + \mu_1^T\Sigma_1^{-1}\mu_1 - \mu_0^T\Sigma_0^{-1}\mu_0 - log(\frac{\phi}{1-\phi}) + log(\frac{|\Sigma_1|^{\frac{-1}{2}}}{|\Sigma_0|^{\frac{-1}{2}}}) = 0$$

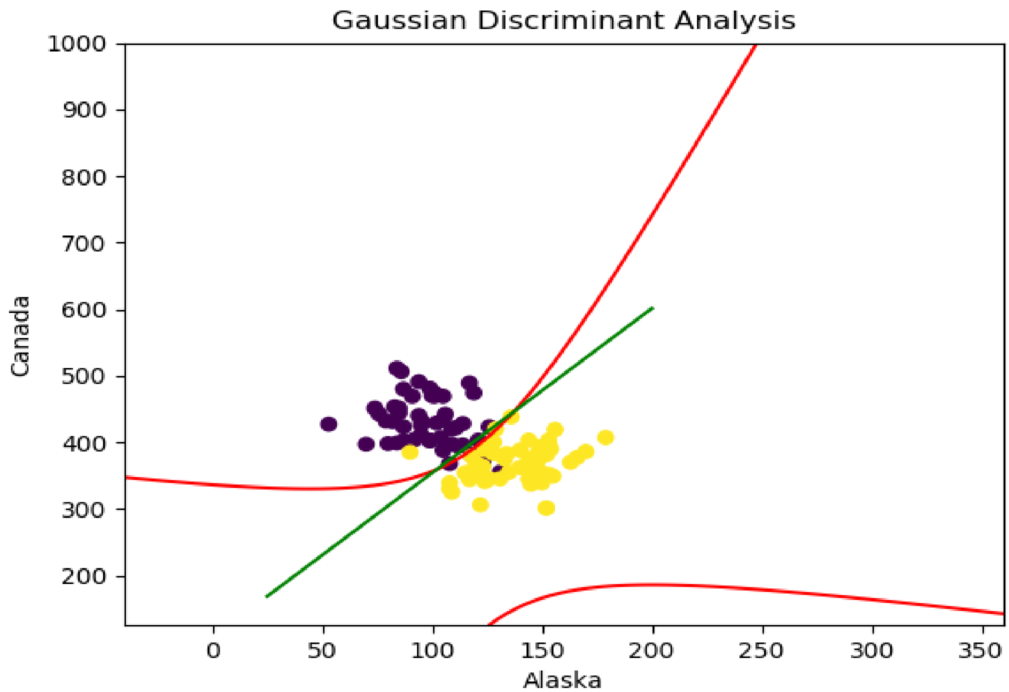

Figure 15: Gaussain Discriminant Analysis Containing Both Linear and Quadratic Boundary