CS229 Fall 2017
Problem Set #1: Supervised Learning

Author: LFhase   rimemosa@163.com

Logistic regression

(a) Given that

$$J(\theta) = \frac{1}{m}\sum_{i=1}^{m} log(1 + e^{-y^{(i)}\theta^T x^{(i)}})$$

we can get

$$\frac{\partial J(\theta)}{\partial \theta_i} = \frac{1}{m}\sum_{k=1}^{m} \frac{-y^{(k)}x_i^{(k)}}{1 + e^{y^{(k)}\theta^T x^{(k)}}}$$

then

$$\frac{\partial J(\theta)}{\partial \theta_i \partial \theta_j} = \frac{1}{m}\sum_{k=1}^{m} \frac{x_i^{(k)}x_j^{(k)}e^{y^{(k)}\theta^T x^{(k)}}}{(1 + e^{y^{(k)}\theta^T x^{(k)}})^2}$$

which is H_{ij}
so

$$\begin{aligned} Z^T H Z &= \sum_{i=1}^{n}\sum_{j=1}^{n} z_i H_{ij} z_j \\ &= \frac{1}{m}\sum_{k=1}^{m} \frac{\sum_{i=1}^{n}\sum_{j=1}^{n} z_i x_i^{(k)} x_j^{(k)} z_j e^{y^{(k)}\theta^T x^{(k)}}}{(1 + e^{y^{(k)}\theta^T x^{(k)}})^2} \end{aligned}$$

known that

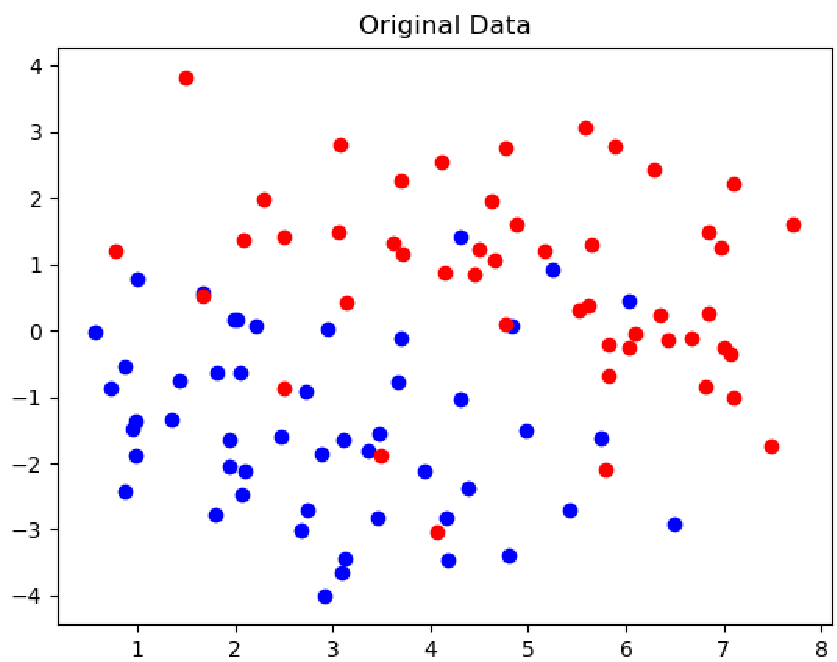
$$\sum_{i=1}^{n}\sum_{j=1}^{n} z_i x_i^{(k)} x_j^{(k)} z_j = (X^T Z)^2 \geq 0$$

$$\frac{e^{y^{(k)}\theta^T x^{(k)}}}{(1 + e^{y^{(k)}\theta^T x^{(k)}})^2} > 0$$

we can easily get

$$Z^T H Z \geq 0$$

(b) Firstly, we plot the original data



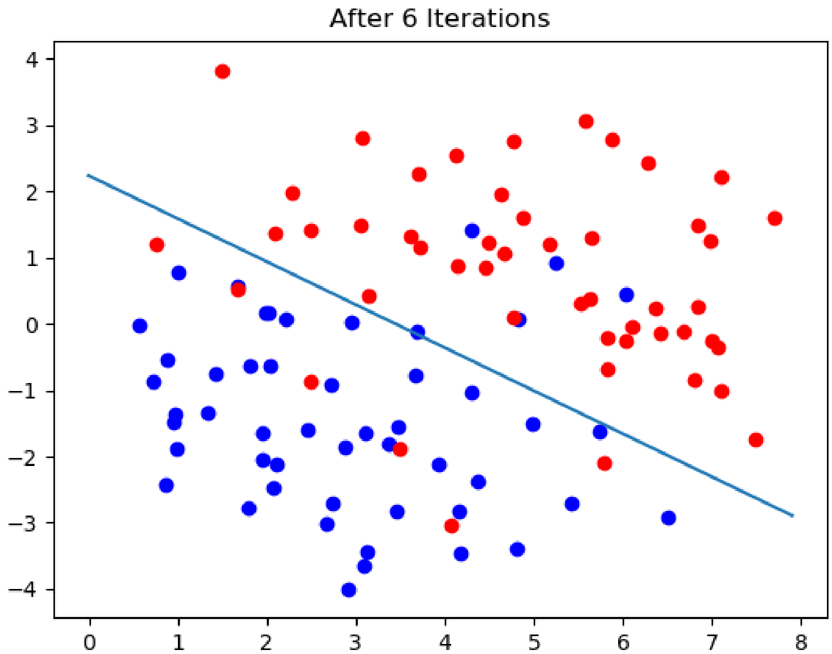
To implement Newton's Method, we calculate the value of

$$\frac{1}{m}\sum_{k=1}^{m}\frac{-y^{(k)}x_i^{(k)}}{1+e^{y^{(k)}\theta^T x^{(k)}}}$$

and **Hessian**
then using the **update rule**

$$\theta := \theta - H^{-1}\bigtriangledown_\theta\, l(\theta)$$

(c) Through 6 iterations, we finally get the result



and the θ is [-2.6205116, 0.76037154, 1.17194674].