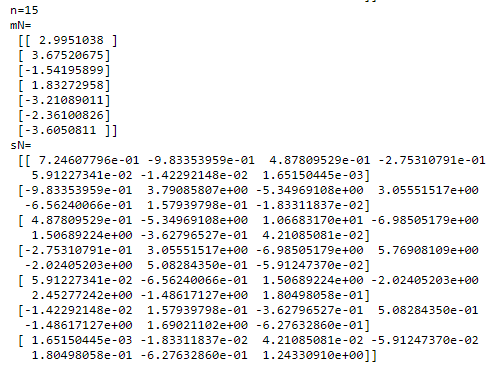
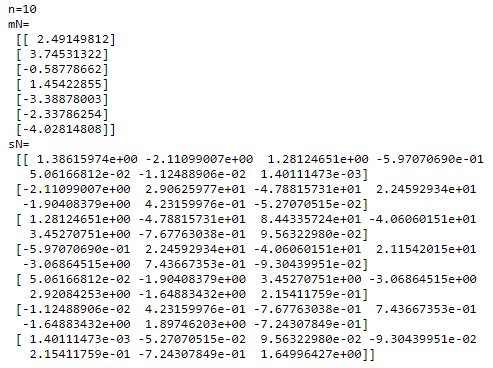
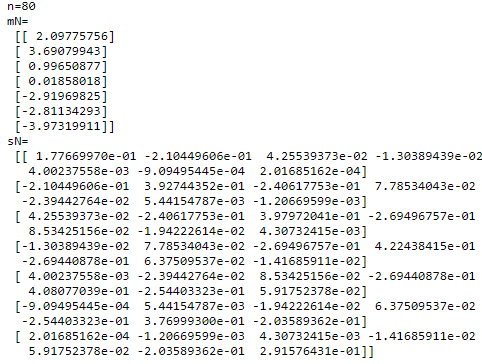
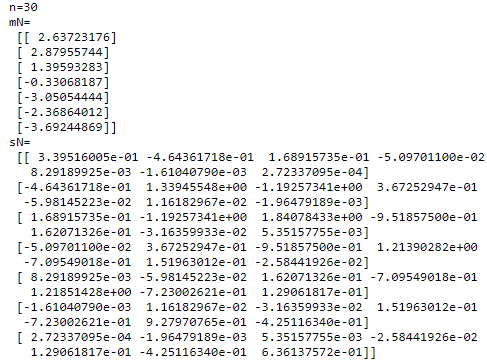
ML LAB2 學號:0412538 姓名:黃宣學

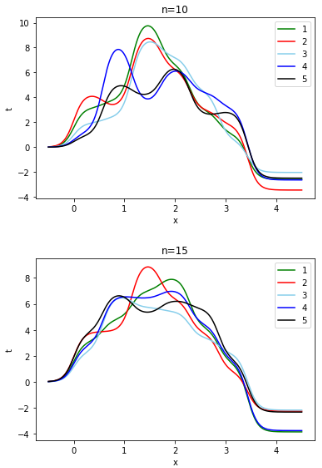
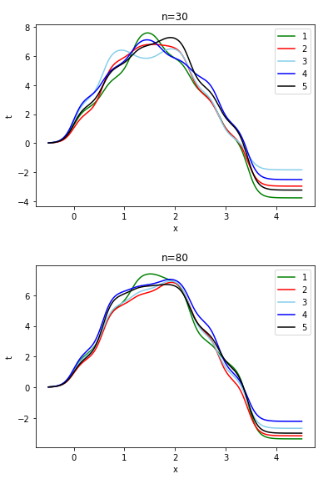
1. Bayesian Linear Regression

1. compute the mean vector mN and the covariance matrix SN



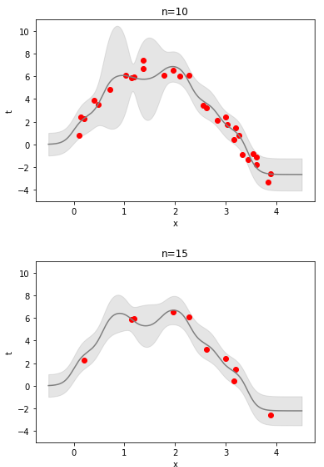
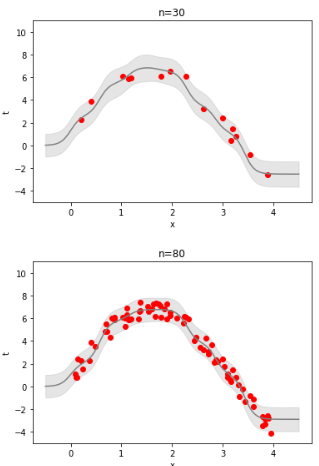


2. generate ﬁve curve samples from the parameter posterior distribution

心得:從上面四張圖可以發現當訓練的data愈多時 , predict的 curve分佈的愈集中

3. plot the predictive distribution of target value t and show the mean curve and the region of variance with one standard deviation on either side of the mean curve.

心得:從上面4張圖可以發現在training data愈集中的地方 uncertainty愈低

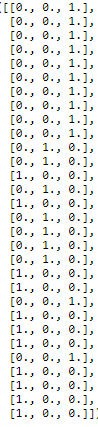
2. Logistic Regression

1. 因為我跑完第一個iteration後，w變得有點大所以要跑第二個iteraion時，exp() 會overflow 算 出inf 所以我只跑了1個iteration

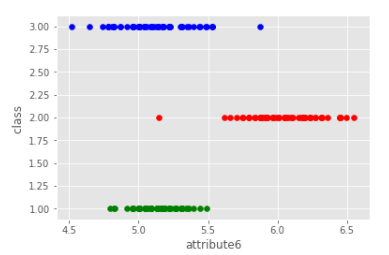
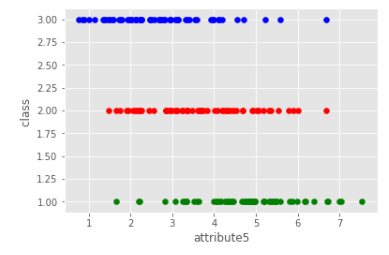
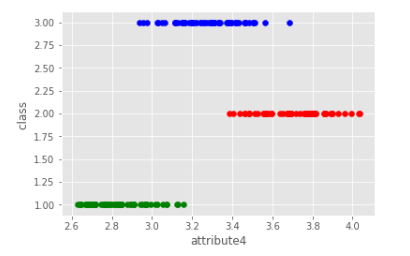
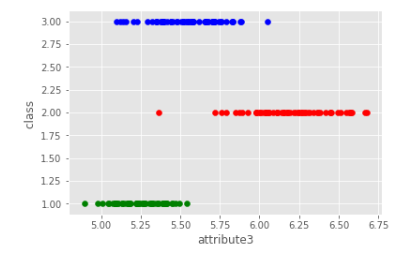
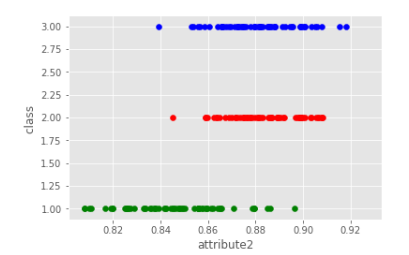
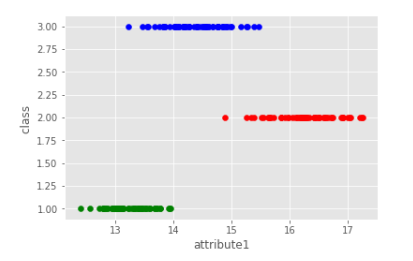
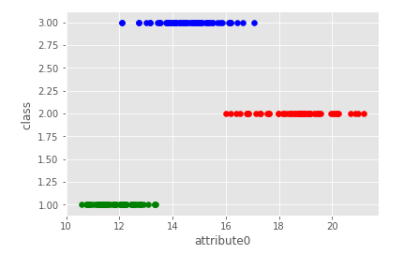
結果為:

(有給王煜閔助教看過了 ,我的code應該沒什麼問題只是python的function真d母湯 , 他說我這一題可以不用畫curve ,寫結果就好)

2.



3. 我用藍紅綠分別代表3個class ,並且為了方便觀察設給不同class不同的y值

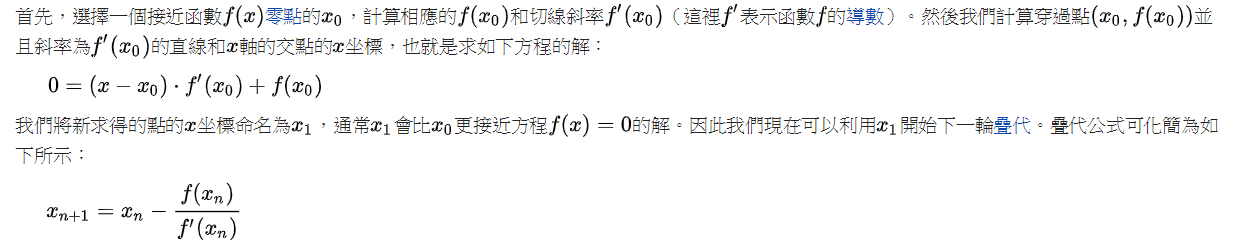


4.

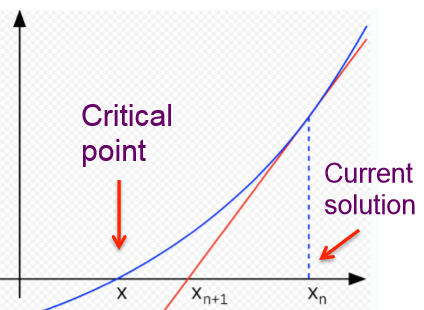
我們的目標是找到E(w)最小的點 , 且因為式子中沒有quadratic form 所以無法使用令一次微分為0的方式找到這個點 , 因此我們用牛頓法以iterative的方式來求解 , 且因為error function是concave的形式 , 所以我們可以保證用牛頓法可以算出最佳解

原理:

(wiki截圖)

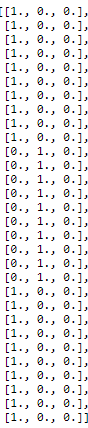


圖例

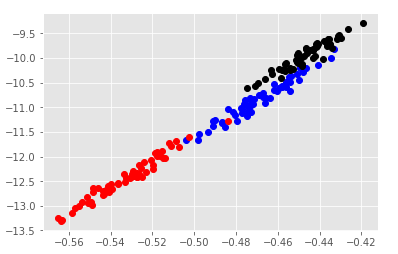


5. 由第三小題可以發現attribute 0、1、3、4在分佈上比較沒有重疊 , 較適合用來當作區別class的 性質=>選擇0和1

6. (只有跑一個iteration , 原因同第一小題)



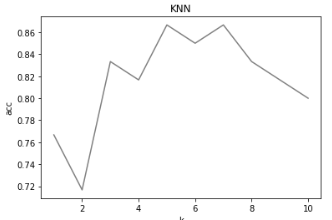
7.



從上圖來看，從7D投影到2D後，三個類別還算分的蠻清楚的，只有極少數的重疊

3 Nonparametric Methods (Bonus Question 30%)

1. K-Nearest-Neighbor Classiﬁer



(1)從上圖來看,可以發現KNN對於K值的選擇敏感度很高,在太大or太小or偶數時表現較不佳

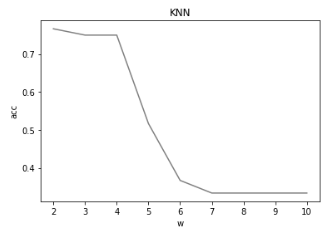
而且k為偶數時可能會發生兩個類別一樣多的情況，所以應該要用奇數以避免這種狀況發生

另外k太小時對於噪音會太過敏感，適當調大k值可以防止因為噪音而誤判

(2)雖然這個lab只用了很基本的majority-voting的KNN並且也得到了不錯的accuracy >86% ，但是這是因為我們的樣本分布很平均，各類別的數目一樣多，如果在樣本分布很不平均的狀況下(每個類別>>其他類別)，可能會發生新的樣本容易被歸類為這個主導的類別，相對應的改進版本為根據距離加上權重(EX:1/d)

(3)KNN有個很好的優點就是只有1個參數要調，所以只要把每個範圍的k全部跑一下就弄完了，也不需要複雜的數學，不過他的缺點就是時間複雜度(每次都要求出和全部training data的距離)和空間複雜度(存全部training data)都蠻高的

2. Alternative solution by ﬁxing the distance and determining the K from training data



從上圖來看，可以發現acc會隨者w下降，而且結果比原本的版本差