Machine Learning - HW1 & 2

Pang Liang
Student No. 201418013229033

April 27, 2015

1 Problem1 - Bayes Classifier

1. Joint Bayes Classifier The joint distribution over all feature is that

$$p(y|x_1, x_2) = \frac{p(x_1, x_2|y) * p(y)}{\sum_y p(x_1, x_2|y)p(y)} \tag{1}$$

So we need estimate these probabilities given the train data. The $p(y, x_1, x_2)$ shows in **??**.

y	$p(x_1 = 0, x_2 = 0\|y)$	$p(x_1 = 0, x_2 = 1\|y)$	$p(x_1 = 1, x_2 = 0\|y)$	$p(x_1 = 1, x_2 = 1\|y)$
0	0.125	0.125	0.375	0.375
1	0.375	0.375	0.000	0.250

Table 1: Values of $p(x_1, x_2|y)$.

$$p(y = 0) = 0.5, p(y = 1) = 0.5 \tag{2}$$

Then we predict the test data below:

$$p(y = 1|x_1 = 0, x_2 = 1) = 0.75 \tag{3}$$

$$p(y = 1|x_1 = 1, x_2 = 0) = 0 \tag{4}$$

$$p(y = 0|x_1 = 1, x_2 = 1) = 0.6 \tag{5}$$

2. Naive Bayes Classifier For Naive Bayes we assume that $p(x_1, x_2|y) = p(x_1|y) * p(x_2|y)$. So

$$p(y|x_1, x_2) = \frac{p(x_1|y) * p(x_2|y) * p(y)}{\sum_y p(x_1|y) * p(x_2|y) * p(y)} \tag{6}$$

Then we only need $p(x_1|y)$ and $p(x_2|y)$ below:

$$p(y = 0) = 0.5, p(y = 1) = 0.5 \tag{7}$$

Then we predict the test data below:

$$\begin{aligned} p(y = 1|x_1 = 0, x_2 = 1) &= \frac{0.75 * 0.625}{0.75 * 0.625 + 0.25 * 0.5} \\ &= 0.789 \end{aligned} \tag{8}$$

$$\begin{aligned} p(y = 1|x_1 = 1, x_2 = 0) &= \frac{0.25 * 0.375}{0.25 * 0.375 + 0.75 * 0.5} \\ &= 0.2 \end{aligned} \tag{9}$$

$$\begin{aligned} p(y = 0|x_1 = 1, x_2 = 1) &= \frac{0.75 * 0.5}{0.75 * 0.5 + 0.25 * 0.625} \\ &= 0.706 \end{aligned} \tag{10}$$

2 Problem2 - Gaussian Bayes Classifiers

See Code **H1P2.m** .

1. Plot the iris data X in 2-D space

y	$p(x_1 = 0\|y)$	$p(x_1 = 1\|y)$
0	0.25	0.75
1	0.75	0.25

Table 2: Values of $p(x_1|y)$.

y	$p(x_2 = 0\|y)$	$p(x_2 = 1\|y)$
0	0.5	0.5
1	0.375	0.625

Table 3: Values of $p(x_2|y)$.

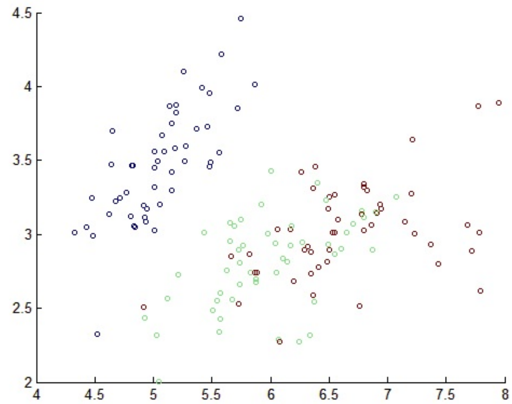


Figure 1: Data distribution

2. Plot each of Gaussian Kernels on top of the data

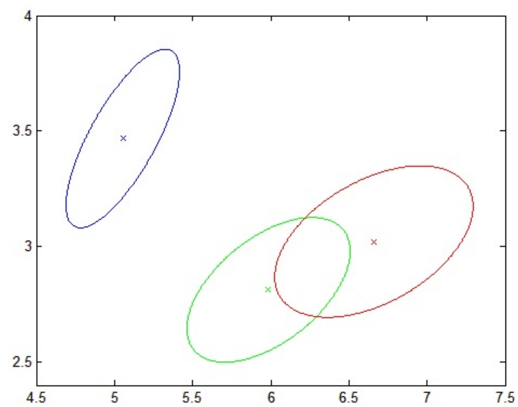


Figure 2: Estimated Gaussian Kernels

3. Using Gaussian Bayes Classify (with free covariances)

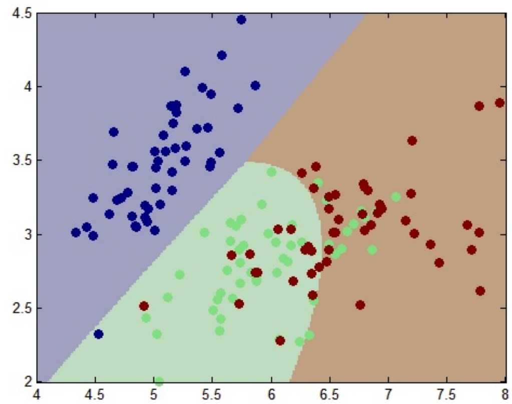


Figure 3: Class Boundary

The shape of the boundary are the parts of a ellipse/circle.

4. Using Gaussian Bayes Classify (with equal covariances)

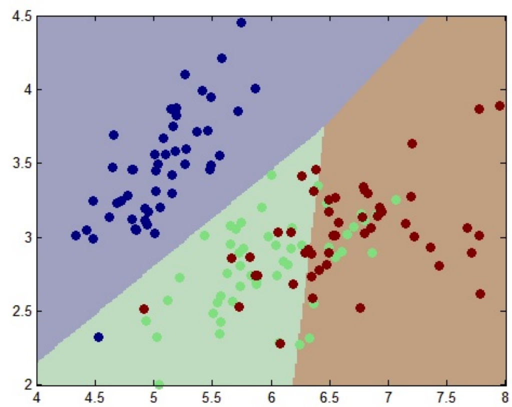
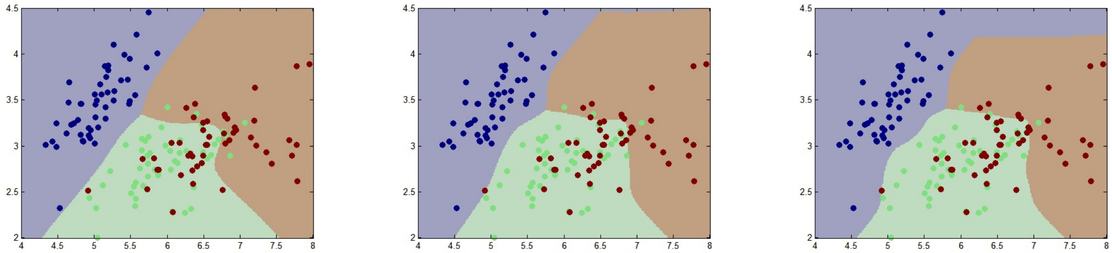


Figure 4: Class Boundary

The shape of the boundary are the straight lines.

5. Poly Classify ($p = 2, 3, 4$)



(a) $p = 2$ (b) $p = 3$ (c) $p = 4$

Figure 5: Different Poly Level, Draw Class Boundary

With the increasing of p, the boundary of each classes become more complex. And it's become easier to overfit on training set.

3 Problem3 - SVM

See Code **H1P3.m** .

1. The optimal hyperplane is $x_1 + x_2 - 3 = 0$.

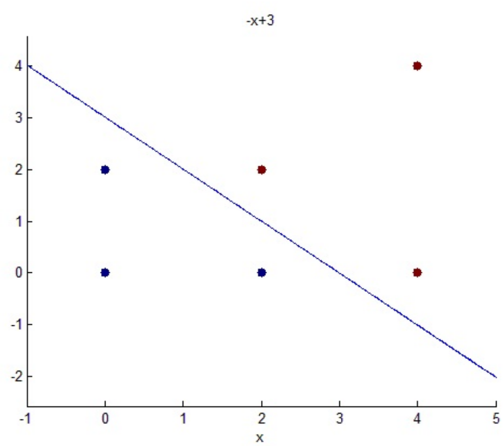


Figure 6: Class Boundary

2. The support vectors are: $(2, 2)$,$(4, 0)$,$(2, 0)$,$(0, 2)$.

4 Problem4 - Run SVMs

1. See Code **H1P4.m** .

2. Use linear SVM get Accuracy of 89.3%.
   Use default SVM get Accuracy of 92.4%.

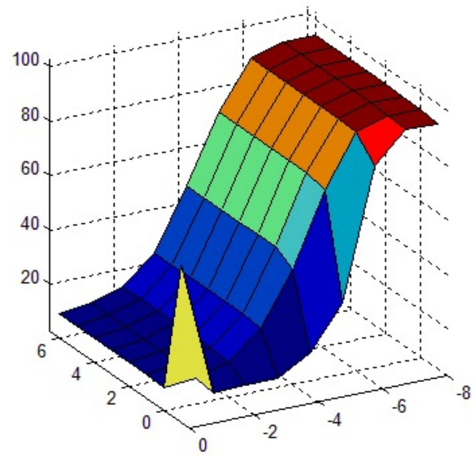


Figure 7: Grid Search.

In my experiments, to set $C = 2^3 = 8$ and $gamma = 2^{-7}$ is the best parameter. Cross Validation accuracy is 94.45%, and Test accuracy is 95%.