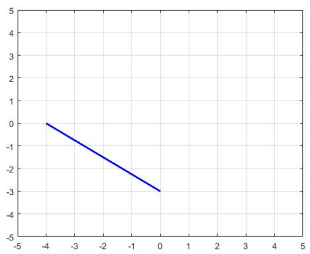
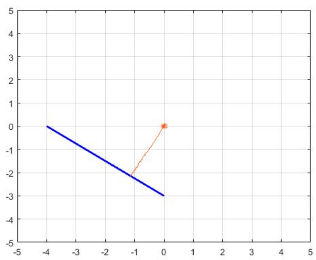
COMP6229(2017/18): Machine Learning Lab 2
Name: Ivanova Iustina
Email: ii1n17@soton.ac.uk

1. There is a line given by the equation 3x + 4y + 12 = 0, where formula of this line is a1x + b1y + c1 = 0. Plot in Matlab:



Img1. Graphic of function: 3x+4y+12=0



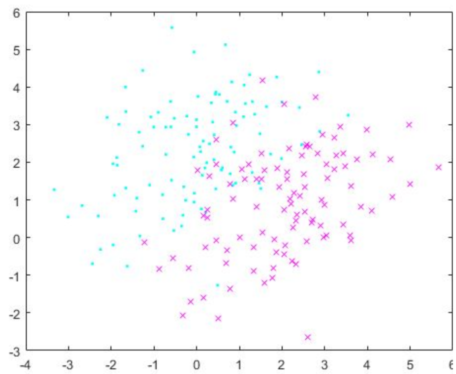
Img2. The distance between line and the origin

The perpendicular distance between line and origin = −c/(a^2 + b^2)^(½) = -12/(3*3+4*4)^(½) = -12/5 = -2,4. From the Img2 the distance is close to this value.

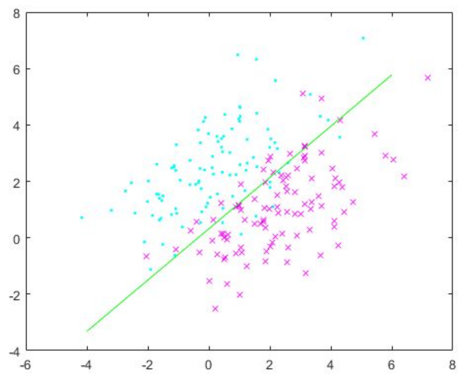
2. There are 2 classes w1, w2 generated by Gaussian distribution with different mean value $m1 = [\frac{0}{2}]$ and $m2 = [\frac{1.5}{0}]$ and identical covariance matrices. Plot in Matlab:

```
N = 100;
C = [2 1; 1 2]
A = chol(C);
X = randn(100,2);
m1 = [0 2];
m2 = [1.5 0];
X1 = X + kron(ones(N,1), m1); !shifting the mean
X2 = X1*A;
Y1 = X + kron(ones(N,1), m2);
Y2 = Y1*A;
C1 = C^(-1);
w = 2*C1*(m2-m1)';
b = m1*C1*m1' - m2*C1*m2';
a1 = w(1);
b1 = w(2);
c1 = b;
x1 = -4;
y1 = (-c1-a1*x1)/b1;
x2 = 6;
y2 = (-c1-a1*x2)/b1;

plot(X2(:,1),X2(:,2),'c.',Y2(:,1),Y2(:,2),'mx');
hold on;
plot([x1 x2],[y1 y2],'g'); ! plot of the
```



Img3. Plot of 2 classes with means m1,m2



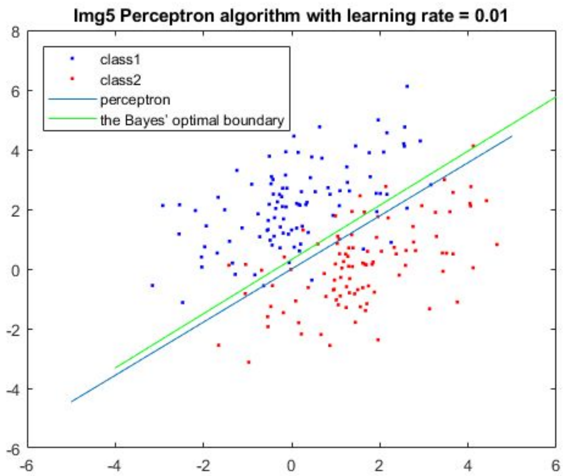
Img4. Bayes' optimal class boundary

3. Bayes' optimal class boundary:

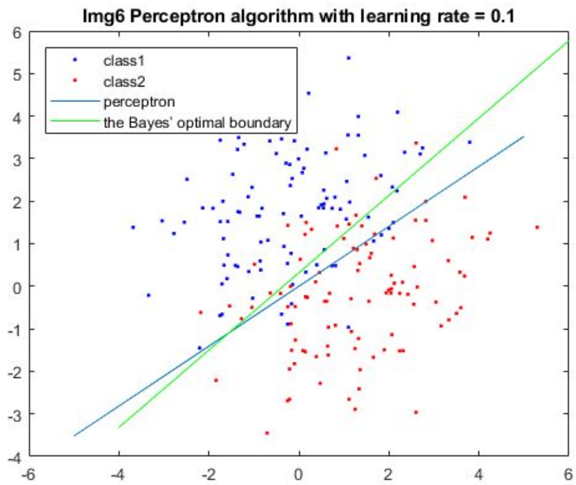
W = 2*C1*(m2-m1)';
b = m1*C1*m1' - m2*C1*m2'; ! as log(P[w1]/P[w2]) = log(1/1)=0
a1x + b1y + c1 = w^t x + b = 0 => a1=w[1]; a2=w[2];

4. Perceptron Learning:

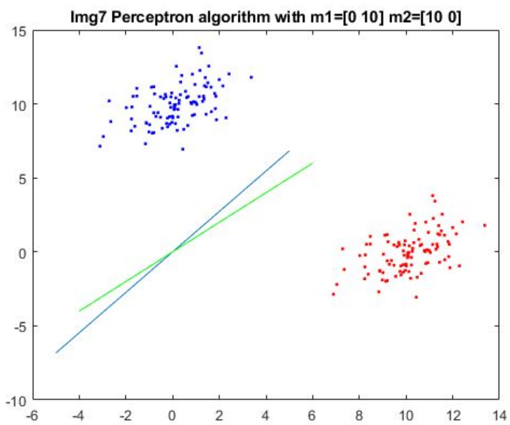
PercentageError = 11

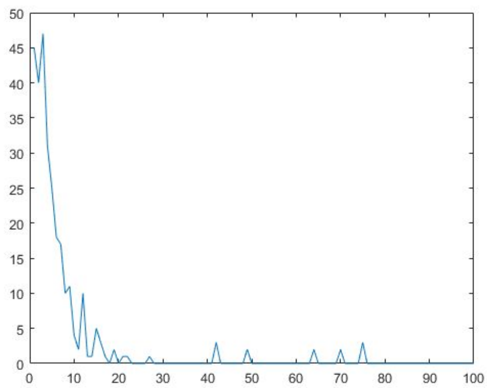


PercentageError = 15



When we change learning rate we can
get similar error, but the difference is the larger learning rates appear to help the model locate regions of general, large-scale optima, while smaller rates help the model focus on one particular local optimum. Usually perceptron start to search with a big learning rate, gradually decreasing it during the process of finding weights.





Img8. Evolution of percentage error (eta= 0.01)
Final error = 0

If the separation of the data that we want to classify is more evident(such as one of the data shown on Img7), it would be easier to determine which samples belong to one class or another.