CS536 Project #002

LaTeX Mini Project 3

Sachin Srivastava
Rutgers Univeristy
sachin.srivastava@rutgers.edu

1. Problem 1

The python scripts is attached. The bias and variance estimate are as follows:

2. Problem 2

Since, θ is 0.5 for all class. We have $p(x_1) = p(x_1|y = c)$ for all c and x_1 and y are independent random variables. As consequences we get
i) $p(y|x_1) = p(y)$
ii) $p(y|x_1, x_2) = p(y|x_2)$
iii)

$$p(x_2) = \sum_{1}^{3} p(y = c)p(x_2|y = c) = \frac{1}{\sqrt{2\pi}}(\frac{\exp^{\frac{-(x_2+1)^2}{2}}}{2} + \frac{\exp^{\frac{-(x_2)^2}{2}}}{4} + \frac{\exp^{\frac{-(x_2-1)^2}{2}}}{4})$$

Thus, $p(y|x_2) = \frac{p(x_2|y)p(y)}{p(x_2)}$, with $p(x_2|y) = \frac{1}{2\pi} \exp^{\frac{(x_2-\mu_y)^2}{2}}$
(μ_y depends on what values y takes) and $p(x_2)$ as above.

3. Problem 3

Since $log\frac{P(y=1|x)}{P(y=1|x)} = 0$, then $logp(x|y = 1) + logP(y = 1) = log(p(x|y = 0) + logP(y = 0)$ and since the conditional densities $p(x|y)$ are Gaussians the formula is equivalent to:

$$(x - \mu_1)^T \sum_{1}^{-1} (x - \mu_1) + log|\sum_{1}| - 2logP(y = 1)$$

$$= (x - \mu_0)^T \sum_{0}^{-1} (x - \mu_0) + log|\sum_{0}| - 2logP(y = 0)$$

and since $\sum_1 = k\sum_0$, the above formula is a quadratic formula with the main term being $(k - 1)x^T \sum_0^{-1}$ which means that the decision boundary is an ellipse.

4. Problem 4

For a fixed x_i, y_i are i.i.d random variables with y_i $N(w_1x_i + w_0, \sigma^2)$. So the probability distribution of y_1, y_2, ... is defined by:

$$f(y_1, ..., y_n|w_1, w_0) = \pi_{i=1}^{n} f(y_i|w_1, w_0)$$

$$= \pi_{i=1}^{n} \frac{1}{(2\sigma^2)^{\frac{n}{2}}} \exp^{\frac{-(y_i - w_1 x_i - w_0)^2}{2\sigma^2}}$$

$$= \frac{1}{(2\sigma^2)^{\frac{n}{2}}} \exp^{\frac{1}{(2\sigma^2)^{\frac{n}{2}}} \sum_{i=1}^{n} (y_i - w_1 x_i - w_0)^2}$$

To get the MLE estimates of w_1 and w_0 we will set $\frac{\partial f}{\partial w_1} = 0$ and $\frac{\partial f}{\partial w_0} = 0$ which leads to the equations:

$$\sum_{i=1}^{n} x_i (y_i - w_1 x_i - w_0) = 0$$

$$\sum_{i=1}^{n} (y_i - w_1 x_i - w_0) = 0$$

Solving the second equation for w_0 yields $w_0 = y - w_1 x$ and replacing w_0 in the first equation yields:

$$\sum_{i=1}^{n} (x_i - \bar{x} + \bar{x})(y_i - w_1 x_i - \bar{y} + w_1 \bar{x}) = 0$$

which in turn leads to

$$w_1 = \frac{\sum_{i=1}^{n} (x_i - \bar{x})(y_i - \bar{y})}{\sum_{i=1}^{n} (x_i - \bar{x})}$$

5. Problem 5

i) When the number of learning observation is small the estimate $\mu \sum^*$ become inexact and result in the classification error of observation vectors which do not participate in the design of the classification rule.

ii) In the two class case we will have $p(y = 1|x, \theta) = \sigma(\beta_1 - \beta_0)^T x + (\gamma_1 - \gamma_0)$
In this case, The decision boundary will get shifted depending on the priors.
iii) It will not be a problem in that case as the co-variance matrix has the correct estimate.