CSci 5523: Machine Learning

Homework 1

Name: Anuradha Uduwage (Anu) *Student ID:* ***0427

Question 1

We calculate mean for each class, using equation(4.21) in the Bishop's book, we use equation 4.24 in Bishops book to calculate within class variance Then we find the projection using 4.26 in the book.

(i) diagFisher.m file implements the first section, with Identity Matrix this method use GetTenFold.m to get the 10 folds. GetTenFolds.m returns a cell with ten folds separated as data and labels. It return the test error for each fold and overall test error as an average of all 10 folds.

(ii) Fisher.m file implements the first section, with S_B and S_W as eigeon vectors, this method use GetTenFold.m to get the 10 folds. GetTenFolds.m returns a cell with ten folds separated as data and labels. It return the test error for each fold and overall test error as an average of all 10 folds.

Question 2

(a) (i) Logistic Regression, The Gradient Ascent methodology has been used.
We wish to maximise the likelihood of the observed data. To do this, we can make use of gradient information of the likelihood, and then ascend the likelihood.
After the derevatives we see, that gradient ascent $w^n ew = w + \eta\delta_w L$

Question 3

(a)

$$TheBayesclassieristhefollowingmapping:$$

$$f(a) = \begin{cases} +1, & ifP(1|X) \geq 0 \\ -1, & otherwise \end{cases} \tag{1}$$

Expected value of the lost function we have,

$$L(f) = E_X Y[l(f(X), Y)] = E_X Y 1_f(X) \neq Y = P_X Y(f(X) \neq Y)$$

Let g(x) be any classier. We will show that

$$P(g(X) = Y|X = x) \geq P(f(x)) \neq Y|X = x).$$

for any g, we can construct

$$\begin{aligned} P(g(X) \neq Y|X = x) =& 1 - P(Y = g(X)|X = x) \\ =& 1 - [P(Y = 1, g(X) = 1|X = x) + P(Y = 0, g(X) = 0|X = x)] \\ =& 1 - E[1_{\{Y=1\}} 1_{\{g(X)=1\}}|X = x] + E[1_{\{Y=0\}} 1_{\{g(X)=0\}}|X = x] \\ =& 1 - 1_{\{g(X)=1\}} E[1_{\{Y=1\}}|X = x] + 1_{\{g(x)=0\}} E[1_{\{Y=0\}}|X = x] \\ =& 1 - 1_{\{g(x)=1\}} P(Y = 1|X = x) + 1_{\{g(x)=0\}} P(Y = 0|X = x) \\ =& 1 - 1_{\{g(x)=1\}} \eta(x) + 1_{\{g(x)=0\}} (1\eta(x)) \end{aligned}$$

now we look at the difference

$$
\begin{aligned}
P(g(x) \neq Y|X=x) - P(f(x) \neq Y|X=x) \\
&= \eta(x)[1_{\{f^*(x)=1\}} - 1_{\{g(x)=1\}} + (1-\eta(x))[1_{\{f(x)=0\}} - 1_{\{g(x)=0\}}] \\
&= \eta(x)[1_{\{f(x)=1\}} - 1_{\{g(x)=1\}} - (1-\eta(x))[1_{\{f(x)=1\}} 1_{\{g(x)=1\}}] \\
&= (2\eta(x)1)(1_{\{f(x)=1\}} - 1_{\{g(x)=1\}}
\end{aligned}
$$

where $\eta(x) = PY|X(Y = 1|X = x)$. we see that second equation follows such that, $1_{\{g(x)=0\}} = 1 - 1\{g(x) = 1\}$ based on equation (1), For x such that $\eta(x)1/2$ we have

$$2\eta(x) - 1 \geq 0 and$$

$$1_{\{f(x)=1\}} - 1_{\{g(x)=1\}} \geq 0$$

and for x such that $\eta(x) < 1/2$, we have

$$2\eta(x) - 1 < 0 and$$

$$1_{\{f(x)=1\}} - 1_{\{g(x)=1\}} \leq 0$$

which implies that,

$$(2\eta(x) - 1)(1_{\{f(x)=1\}} - 1_{\{g(x)=1\}}) \geq 0$$

and we can rewrite that as

$$P(g(X) = Y|X = x) \geq P(f(x)) \neq Y|X = x)$$