

COMS 474  
Homework 6

Haadi Majeed

Spring 2022

## Problem 1

[8 points]

Recall that for a two-class  $\{0, 1\}$  classification problem with  $p$  features  $\{X_1, \dots, X_p\}$ , logistic regression models have the form

$$P(Y = 1|X_1, \dots, X_p) = \frac{e^{\beta_0 + \beta_1 X_1 + \dots + \beta_p X_p}}{1 + e^{\beta_0 + \beta_1 X_1 + \dots + \beta_p X_p}} \quad P(Y = 0|X_1, \dots, X_p) = \frac{1}{1 + e^{\beta_0 + \beta_1 X_1 + \dots + \beta_p X_p}}$$

For a model with fitted coefficients  $\{\beta_0, \beta_1, \dots, \beta_p\}$ , what is the boundary between  $\hat{Y} = 1$  and  $\hat{Y} = 0$ ? Simply your work as much as possible.

By assigning the equations equal to each other, and since they have identical values in the denominator we can be sitting with  $e^{\beta_0 + \beta_1 X_1 + \dots + \beta_p X_p} = 1$  and we then assign  $X_1, \dots, X_p = 0$  along with  $\beta_0 = 0$ . We can get the following sequence of equations:

$$\begin{aligned} \frac{e^{\beta_0}}{1 + e^{\beta_0}} &= \frac{1}{1 + e^{\beta_0}} \\ \frac{1}{1 + 1} &= \frac{1}{1 + 1} \\ \frac{1}{2} &= \frac{1}{2} \end{aligned}$$

## Problem 2

[15 points (8,2,5)]

In this problem you will identify the max margin classifier for the following data set with three samples.

Sample #	$Y$	$X_1$	$X_2$
1	+1	0	2
2	-1	0	0
3	+1	-2	0

You do not need to draw a scatter plot of these samples, though it may help you to check your answers to the following questions.

### A

What are the set of coefficients  $\{\beta_0, \beta_1, \beta_2\}$  for the line

$$\beta_0 + \beta_1 X_1 + \beta_2 X_2 = 0$$

that passes through the two  $Y = +1$  samples? Normalize the coefficients so that  $\beta_1^2 + \beta_2^2 = 2$  (In lecture we used a different normalization; this is chosen to simplify your work). Since there are two  $Y = +1$  samples, we create a system of equations.

The first sample:

$$\beta_0 + \beta_1(0) + \beta_2(2) = 0$$

$$\beta_0 + 2\beta_2 = 0$$

The second sample:

$$\beta_0 + \beta_1(-2) + \beta_2(0) = 0$$

$$\beta_0 - 2\beta_1 = 0$$

With this, we can determine the following:

$\beta_0$	$\beta_1$	$\beta_2$
2	1	-1

This satisfies the normalization equation resulting with:

$$-1^2 + 1^2 = 2$$

### B

Confirm whether your coefficients  $\{\beta_0, \beta_1, \beta_2\}$  of  $\beta_0 + \beta_1 X_1 + \beta_2 X_2$  evaluates positive or negative for the  $Y = -1$  sample.

If the evaluation is positive, change your formula so that it evaluates as negative for that sample.

The coefficients

$\beta_0$	$\beta_1$	$\beta_2$
2	1	-1

We can enter into the formula:

$$2 + 1 * 0 + -1 * 0 = 2$$

Since it is positive, we can modify all  $\beta$  values while still maintaining validity for part A by doing the following:

$$-1 * \{\beta_0, \beta_1, \beta_2\}$$

## C

Using your results from A and B, guess what the optimal separating line (i.e. the boundary line for the max margin classifier)

$$\tilde{\beta}_0 + \tilde{\beta}_1 X_1 + \tilde{\beta}_2 X_2 = 0$$

is with  $\tilde{\beta}_1^2 + \tilde{\beta}_2^2 = 2$ .

Confirm that using the classifier

$$\hat{Y}(X_1, X_2) = \text{sign}\left(\tilde{\beta}_0 + \tilde{\beta}_1 X_1 + \tilde{\beta}_2 X_2\right)$$

achieves zero classification error and that all training samples have the same magnitude of  $\tilde{\beta}_0 + \tilde{\beta}_1 X_1 + \tilde{\beta}_2 X_2$ .

Using properties from A and B We can fill in the formula

$$\hat{Y}(X_1, X_2) = -\left(\tilde{2} + \tilde{1}X_1 - \tilde{1}X_2\right)$$

When we plug in the values for each of the samples, we get the following:

$$\text{Sample 1: } -(2 + 1 * 0 - 1 * 2 = 0)$$

$$\text{Sample 2: } -(2 + 1 * 0 - 1 * 0 = -2)$$

$$\text{Sample 3: } -(2 + 1 * -2 - 1 * 0 = 0)$$

Thus we need to adjust the value of  $\beta_0$  from 2 to 1 in order to get a more equal distribution.

$$\text{Sample 1: } -(1 + 1 * 0 - 1 * 2 = 1)$$

$$\text{Sample 2: } -(1 + 1 * 0 - 1 * 0 = -1)$$

$$\text{Sample 3: } -(1 + 1 * -2 - 1 * 0 = 1)$$

This modified version now has a correct sign of negative one, and confirms the new classifier with equal magnitudes and zero error.

## Problem 3

[24 points (15, 4, 2, 3)]

You are working on a two-class  $\{A, B\}$  classification problem using two features  $\{X_1, X_2\}$ . You decide to use QDA. You estimate the mean vectors for each class and find that they are the same,

$$\widehat{\mu}_A = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \quad \widehat{\mu}_B = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

When you estimate the covariance matrices, you find they are both diagonal matrices,

$$\widehat{\Sigma}_A = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \quad \widehat{\Sigma}_B = \begin{bmatrix} 4 & 0 \\ 0 & 4 \end{bmatrix}$$

Lastly, the number of samples for Class  $A$  and  $B$  are the same, so  $\widehat{\pi}_A = \frac{1}{2} = \widehat{\pi}_B$ .

### A

What is the equation for the QDA decision boundary for this problem? Simplify it as much as possible.

### B

Sketch a plot by hand showing the boundary. Indicate for what region(s)  $\hat{Y} = A$  and for what region(s)  $\hat{Y} = B$

### C

What would the prediction be for the mean value of both conditional distributions,  $\begin{bmatrix} 0 \\ 0 \end{bmatrix}$ ?

What would the prediction be for the point  $\begin{bmatrix} 3 \\ -3 \end{bmatrix}$ ?

### D

Suppose we had a third class  $C$ , also centered at the origin and with a diagonal covariance matrix,

$$\widehat{\mu}_C = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \quad \widehat{\Sigma}_C = \begin{bmatrix} 9 & 0 \\ 0 & 9 \end{bmatrix}$$

and that all classes had equal priors,  $\pi_A = \pi_B = \pi_C = \frac{1}{3}$ .

Sketch a plot similar to B, of the decision regions for this problem with three classes (e.g. make it clear on the plot what region has  $\hat{Y} = A$ , what region has  $\hat{Y} = B$ , and what region has  $\hat{Y} = C$ ). You do not need to do calculations, but extrapolate from what you just observed for the two class problem.