COM S 474/574 Spring 2022
Introduction to Machine Learning

Homework 6

1 Directions:

- **Due: Thursday March 31, 2022 at 9pm.** Late submissions will be accepted for 24 hours with a 15% penalty. (the enforcement is strict, beginning at 9:01pm, except for extreme situations; having a poor wifi connection or minor computer problems is not sufficient for the penalty to be waived.)
- Upload the homework to Canvas as a single pdf file.
- If the graders cannot easily read your submission (writing is illegible, image is too dark, or if the contrast is too low) then you might receive a zero or only partial credit.
- Any non-administrative questions must be asked in office hours or (if a brief response is sufficient) Piazza.

2 Problems

Problem 1. [8 points] Recall that for a two-class $\{0,1\}$ classification problem with p features $\{X_1,\ldots,X_p\}$, logistic regression models have the form

$$P(Y=1|X_1,\ldots,X_p)=\frac{e^{\beta_0+\beta_1X_1+\cdots+\beta_pX_p}}{1+e^{\beta_0+\beta_1X_1+\cdots+\beta_pX_p}} \qquad P(Y=0|X_1,\ldots,X_p)=\frac{1}{1+e^{\beta_0+\beta_1X_1+\cdots+\beta_pX_p}}$$

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

(The assignment continues on the next page)

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 <u>not</u> 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 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).

(b). Confirm whether for your coefficients $\{\beta_0, \beta_1, \beta_2\}$ if $\beta_0 + \beta_1 X_1 + \beta_2 X_2$ evaluates as 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.

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

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

is, with $\widetilde{\beta}^2 + \widetilde{\beta}^2 = 2$.

Confirm that using the classifier

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

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

(The assignment continues on the next page)

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,

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

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

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

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) $\widehat{Y} = A$ and for what region(s) $\widehat{Y} = B$

(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,

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

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 (eg make it clear on the plot what region has $\widehat{Y} = A$, what region has $\widehat{Y} = B$, and what region has $\widehat{Y} = C$). You do not need to do calculations, but extrapolate from what you just observed for the two class problem.