Written Exercises

Adam Wolfson (amw337) & John Waldo (jw922)

October 5, 2021

1 Exercise 1: Ridge Regression

To begin, we can denote X_{aug}, Y_{aug} as the following:

$$X_{aug} = \begin{pmatrix} X \\ \sqrt{\lambda} I_{pxp} \end{pmatrix}, Y_{aug} = \begin{pmatrix} y \\ 0_{px1} \end{pmatrix}$$

OLS tells us that $\hat{\beta} = (X^T X)^{-1} X^T y$
Using this, we can see that for X_{aug}, Y_{aug} we see $\hat{\beta} = (X_{aug}^T X_{aug})^{-1} X_{aug}^T Y_{aug}$
Looking first at $X_{aug}^T X_{aug}$, this is:

$$\begin{pmatrix} X^T & \sqrt{\lambda} I \end{pmatrix} * \begin{pmatrix} X \\ \lambda I \end{pmatrix}$$

which yields

$$\begin{pmatrix} X^T X + \lambda I \end{pmatrix}$$

Now we look at $X_{aug}^T Y_{aug}$

$$\begin{pmatrix} X^T & \sqrt{\lambda} I \end{pmatrix} \begin{pmatrix} y \\ 0 \end{pmatrix} = X^T y$$

Combining these we see that $\hat{\beta} = (X^T X + \lambda I)^{-1} X^T y$ which is the Ridge Regression solution to the minimization, solving for β. We therefore can see that OLS on X_{aug}, Y_{aug} has the same solution as Ridge Regression on X, Y with parameter λ

2 Exercise 2: Naive Bayes Classifiers

2.1 Part A: Optimal Bayes

For the Bayes Optimal classifier, we want (Note: Fever = F, No Headache = NH):

$$P(Pneumonia|F, NH) = \frac{P(F, NH|Pneumonia)P(Pneumonia)}{P(F, NH)}$$

$$P(Flu|F, NH) = \frac{P(F, NH|Flu)P(Flu)}{P(F, NH)}$$

$$P(Healthy|F, NH) = \frac{P(F, NH|Healthy)P(Healthy)}{P(F, NH)}$$

We gather the following probabilities from the data to be able to compute this:

$$P(Pneumonia) = \frac{1}{10}, P(Flu) = \frac{2}{10}, P(Healthy) = \frac{7}{10}, P(Fever, NoHeadache) = \frac{9}{100}$$

$$P(F, NH|Pneumonia) = \frac{0}{10}, P(F, NH|Flu) = \frac{6}{20}, P(F, NH|Healthy) = \frac{3}{70}$$

We then get:

$$P(Pneumonia|F, NH) = \frac{\frac{0}{10}\frac{1}{10}}{\frac{9}{100}} = 0$$

$$P(Flu|F, NH) = \frac{\frac{6}{20}\frac{2}{10}}{\frac{9}{100}} = \frac{2}{3}$$

$$P(Healthy|F, NH) = \frac{\frac{3}{70}\frac{7}{10}}{\frac{9}{100}} = \frac{1}{3}$$

2.2 Part B: Naive Bayes

For Naive Bayes, we incorporate our naive assumption and calculate the probabilities as:

$$P(Pneumonia|F, NH) = \frac{P(F|Pneumonia)P(NH|Pneumonia)P(Pneumonia)}{P(F, NH)}$$

$$P(Flu|F,NH) = \frac{P(F|Flu)P(NH|Flu)P(Flu)}{P(F,NH)}$$

$$P(Healthy|F,NH) = \frac{P(F|Healthy)P(NH|Healthy)P(Healthy)}{P(F,NH)}$$

P(Pneumonia), P(Flu), and P(Healthy) will be the same as in the previous section. We next turn towards gathering the other probabilities.

$$P(F|Pneumonia) = \frac{1}{2}, P(NH|Pneumonia) = \frac{1}{10}$$

$$P(F|Flu) = \frac{3}{4}, P(NH|Flu) = \frac{8}{20}$$

$$P(F|Healthy) = \frac{5}{70}, P(NH|Healthy) = \frac{61}{70}$$

Finally, we have to figure out our denominator $P(F,NH)$ for each class k (Pneumonia, Flu, Healthy):

$$P(F,NH) = \sum_k P(C_k)P(F|C_k)P(NH|C_k)$$

As this is the summation of all 3 of our numerators, we will hold off on this for now and first compute the numerators for each class.

$$P(Pneumonia|F,NH) = \frac{\frac{1}{2}\frac{1}{10}\frac{1}{10}}{P(F,NH)} = \frac{\frac{1}{200}}{P(F,NH)}$$

$$P(Flu|F,NH) = \frac{\frac{3}{4}\frac{8}{20}\frac{2}{10}}{P(F,NH)} = \frac{\frac{3}{50}}{P(F,NH)}$$

$$P(Healthy|F,NH) = \frac{\frac{5}{70}\frac{61}{70}\frac{7}{10}}{P(F,NH)} = \frac{\frac{61}{1400}}{P(F,NH)}$$

Summing up the numerators here we get $P(F,NH) = \frac{1}{200} + \frac{3}{50} + \frac{61}{1400} = \frac{19}{175}$
Finally, we arrive at:

$$P(Pneumonia|F,NH) = \frac{\frac{1}{200}}{\frac{19}{175}} = \frac{7}{152}$$

$$P(Flu|F,NH) = \frac{\frac{3}{50}}{\frac{19}{175}} = \frac{21}{38}$$

$$P(Healthy|F,NH) = \frac{\frac{61}{1400}}{\frac{19}{175}} = \frac{61}{152}$$

3 Exercise 3: Naive Bayes with Data for Nominal Attributes

In this exercise, we want to predict whether our subject will buy or not buy the computer. We can compute this by comparing the probabilities $P(Buy|X)$ and $P(Don'tBuy|X)$ where $X = \{\leq 30, medium, yes, fair\}$

Using naive bayes, we can write the probabilities as (Note: Buy = YES, Don't Buy = NO:

$$P(YES|X) = \frac{P(\leq 30|YES)P(medium|YES)P(yes|YES)P(fair|YES)P(YES)}{P(\leq 30, medium, yes, fair)}$$

$$P(NO|X) = \frac{P(\leq 30|NO)P(medium|NO)P(yes|NO)P(fair|NO)P(NO)}{P(\leq 30, medium, yes, fair)}$$

Because the denominators are the same and we are comparing the two, we can ignore it and focus purely on the numerator, where we have

$$P(YES) = \frac{9}{14}, P(NO) = \frac{5}{14}$$

$$P(\leq 30|YES) = \frac{2}{9}, P(medium|YES) = \frac{4}{9}, P(yes|YES) = \frac{6}{9}, P(fair|YES) = \frac{6}{9}$$

$$P(\leq 30|NO) = \frac{3}{5}, P(medium|NO) = \frac{2}{5}, P(yes|NO) = \frac{1}{5}, P(fair|NO) = \frac{2}{5}$$

Inserting these numbers into our previous equations gives the numerator for a classification of YES as 0.028 and for a classification of no it is 0.007. As we remember that the denominators are the same, since the probability of YES for a purchase is higher than the probability of NO, we predict that this individual will buy the computer.