practice midterm 2

Go Ito

May 13, 2019

Question 1.

Given the following information for Neural Network, obtain the output $\hat{y}$.

$$
\begin{aligned}
X &= \begin{bmatrix} 5.1 & 4.9 \\ 4.9 & 3.0 \end{bmatrix} \\
W^{(1)} &= \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \\
B^{(1)} &= \begin{bmatrix} 1.0 & -4.0 \end{bmatrix} \\
W^{(2)} &= \begin{bmatrix} 1 \\ -1 \end{bmatrix} \\
B^{(2)} &= \begin{bmatrix} 0.5 \end{bmatrix}
\end{aligned}
$$

Use ReLu as the activation function (check 6-3).

Question 2.

Now you are given the test data

$$
y = \begin{bmatrix} 6.0 \\ 6.0 \end{bmatrix}
$$

First, write down the 4 partial derivatives of loss function. That is,

$$
\frac{\partial J}{\partial W^{(2)}}, \frac{\partial J}{\partial B^{(2)}}, \frac{\partial J}{\partial W^{(1)}}, \frac{\partial J}{\partial B^{(2)}}
$$

Then, use the test data y, data from Question 1., AND new activation function: Sigmoid, obtain

$$
\frac{\partial J}{\partial W^{(1)}}
$$

numerically.

Question 3.

Let $x \sim exp(\lambda)$. Its pdf is given as

$$
f(x) = \lambda e^{-\lambda x}
$$

Obtain the Maxlimum Likelihood Estimate of the parameter λ.

Also obtain the fisher information.

Question 4.

Let

$$X \sim N(\begin{bmatrix}1\\2\end{bmatrix}, \begin{bmatrix}1/2 & 0\\0 & 1\end{bmatrix})$$

Obtain the Fisher Information Matrix.

Question 5.

We have three movie theaters: A, B and C. When we watch Detective Pikachu movie, a pack of Pokemon cards that contains 2 cards will be gifted. We are trying to classify which movie theater that a pack of Pokemon cards come from via the portion of the Pikachu card.

From each theater, we got : 8 packs from A, 5 from B and 5 from C. Again, each pack contains 2 cards.

We opened up the packs, and it turns out that the portion of Pikachu cards are as follows: 50% for pack A, 20% for B, and 80% for C.

Now we have 3 packs obtained from one of the movie theaters, and it contained 3 Pikachu cards!

Calculate the "posterior" probability; that is, the probability that these 3 packs come from each theaters.

Question 6.

We have 2 Bivariate Normals:

$$X_1 \sim N(\begin{bmatrix}1\\2\end{bmatrix}, \begin{bmatrix}1/2 & 0\\0 & 1\end{bmatrix})$$
$$X_2 \sim N(\begin{bmatrix}-4\\-2\end{bmatrix}, \begin{bmatrix}1 & 0\\0 & 2\end{bmatrix})$$

Given new data point $X_{new} = (0, 0)$, calculate the the probability that this point belong to each Bivariate normals (assume the number of points that come from each Bivariate normals are the same).

Question 7.

- What is Naive Bayes Classifier? Pros and cons?
- How is Hessian useful in general? How is it useful in terms of log-likelihood?
- Recall the marvel example from the class (6-1). We have 50% red marvels from factory B. However, when the test case (10/20 are red) were concerned, the Bayes Classifier suggested that this test case bag is more likely from factory A, where only 40% of the marvels are red. Why did this happen?
- What are the pros and cons of Neural Network?
- What are the four main components of Baye's classifier? Explain each.
- What is the score? How is it used in Fisher Information?
- What does it mean when the value of a Hessian is overall high? When would this happen?