MCA/MSc. (CS) Semester-II Examination 2022-23
CS322: Deep Learning

Time: 1 hour **20 marks**

1. Find the SVD of A, $U\sum V^T$, where A $=\begin{bmatrix} 3 & 2 & 2 \\ 2 & 3 & -2 \end{bmatrix}$. Explain each step clearly. [11 marks]

2. Let's assume we have a two layer neural network, as defined below:

$$z_1 = W_1 x^{(i)} + b_1, \qquad a_1 = ReLU(z_1), \qquad z_2 = W_2 a_1 + b_2, \qquad \hat{y}^{(i)} = \sigma(z_2)$$
$$L^{(i)} = y^{(i)} * \log\left(\hat{y}^{(i)}\right) + (1 - y^{(i)}) * \log\left(1 - \hat{y}^{(i)}\right), \qquad J = \frac{-1}{m}\sum_{i=1}^{m} L^{(i)}$$

Note that $x^{(i)}$ represents a single input example, and is of shape $D_x \times 1$. Further $y^{(i)}$ is a single output label and is a scalar. There are m examples in our dataset. We will use D_{a_1} nodes in our hidden layer; that is, z_1's shape is $D_{a_1} \times 1$.

MCA/MSc. (CS) Semester-II Examination 2022-23
CS322: Deep Learning

Time: 1 hour **20 marks**

1. Find the SVD of A, $U\sum V^T$, where A $=\begin{bmatrix} 3 & 2 & 2 \\ 2 & 3 & -2 \end{bmatrix}$. Explain each step clearly. [11 marks]

2. Let's assume we have a two layer neural network, as defined below:

$$z_1 = W_1 x^{(i)} + b_1, \qquad a_1 = ReLU(z_1), \qquad z_2 = W_2 a_1 + b_2, \qquad \hat{y}^{(i)} = \sigma(z_2)$$
$$L^{(i)} = y^{(i)} * \log\left(\hat{y}^{(i)}\right) + (1 - y^{(i)}) * \log\left(1 - \hat{y}^{(i)}\right), \qquad J = \frac{-1}{m}\sum_{i=1}^{m} L^{(i)}$$

Note that $x^{(i)}$ represents a single input example, and is of shape $D_x \times 1$. Further $y^{(i)}$ is a single output label and is a scalar. There are m examples in our dataset. We will use D_{a_1} nodes in our hidden layer; that is, z_1's shape is $D_{a_1} \times 1$.

MCA/MSc. (CS) Semester-II Examination 2022-23
CS322: Deep Learning

Time: 1 hour **20 marks**

1. Find the SVD of A, $U\sum V^T$, where A $=\begin{bmatrix} 3 & 2 & 2 \\ 2 & 3 & -2 \end{bmatrix}$. Explain each step clearly. [11 marks]

2. Let's assume we have a two layer neural network, as defined below:

$$z_1 = W_1 x^{(i)} + b_1, \qquad a_1 = ReLU(z_1), \qquad z_2 = W_2 a_1 + b_2, \qquad \hat{y}^{(i)} = \sigma(z_2)$$
$$L^{(i)} = y^{(i)} * \log\left(\hat{y}^{(i)}\right) + (1 - y^{(i)}) * \log\left(1 - \hat{y}^{(i)}\right), \qquad J = \frac{-1}{m}\sum_{i=1}^{m} L^{(i)}$$

Note that $x^{(i)}$ represents a single input example, and is of shape $D_x \times 1$. Further $y^{(i)}$ is a single output label and is a scalar. There are m examples in our dataset. We will use D_{a_1} nodes in our hidden layer; that is, z_1's shape is $D_{a_1} \times 1$.

MCA/MSc. (CS) Semester-II Examination 2022-23
CS322: Deep Learning

Time: 1 hour **20 marks**

1. Find the SVD of A, $U\sum V^T$, where A $=\begin{bmatrix} 3 & 2 & 2 \\ 2 & 3 & -2 \end{bmatrix}$. Explain each step clearly. [11 marks]

2. Let's assume we have a two layer neural network, as defined below:

$$z_1 = W_1 x^{(i)} + b_1, \qquad a_1 = ReLU(z_1), \qquad z_2 = W_2 a_1 + b_2, \qquad \hat{y}^{(i)} = \sigma(z_2)$$
$$L^{(i)} = y^{(i)} * \log\left(\hat{y}^{(i)}\right) + (1 - y^{(i)}) * \log\left(1 - \hat{y}^{(i)}\right), \qquad J = \frac{-1}{m}\sum_{i=1}^{m} L^{(i)}$$

Note that $x^{(i)}$ represents a single input example, and is of shape $D_x \times 1$. Further $y^{(i)}$ is a single output label and is a scalar. There are m examples in our dataset. We will use D_{a_1} nodes in our hidden layer; that is, z_1's shape is $D_{a_1} \times 1$.

Based on the given information, answer the following questions:

(a) What are the shapes of W_1 , b_1 , W_2 , b_2? If we were vectorizing this network across multiple examples, what would the shapes of the weights/biases be instead? If we were vectorizing across multiple examples, what would the shapes of X and Y be instead? [2 marks]

(b) What is $\partial J/\partial \hat{y}^{(i)}$? Refer to this result as $\partial_1^{(i)}$. Using this result, what is $\partial J/\partial \hat{y}$? [1 marks]

(c) What is $\partial \hat{y}^{(i)}/\partial z_2$? Refer to this result as $\partial_2^{(i)}$. [1 marks]

(d) What is $\partial z_2/\partial a_1$? Refer to this result as $\partial_3^{(i)}$. [1 marks]

(e) What is $\partial a_1/\partial z_1$? Refer to this result as $\partial_4^{(i)}$. [1 marks]

(f) What is $\partial z_1/\partial W_1$? Refer to this result as $\partial_5^{(i)}$. [1 marks]

(g) What is $\partial J/\partial W_1$? [2 marks]

Based on the given information, answer the following questions:

(a) What are the shapes of W_1 , b_1 , W_2 , b_2? If we were vectorizing this network across multiple examples, what would the shapes of the weights/biases be instead? If we were vectorizing across multiple examples, what would the shapes of X and Y be instead? [2 marks]

(b) What is $\partial J/\partial \hat{y}^{(i)}$? Refer to this result as $\partial_1^{(i)}$. Using this result, what is $\partial J/\partial \hat{y}$? [1 marks]

(c) What is $\partial \hat{y}^{(i)}/\partial z_2$? Refer to this result as $\partial_2^{(i)}$. [1 marks]

(d) What is $\partial z_2/\partial a_1$? Refer to this result as $\partial_3^{(i)}$. [1 marks]

(e) What is $\partial a_1/\partial z_1$? Refer to this result as $\partial_4^{(i)}$. [1 marks]

(f) What is $\partial z_1/\partial W_1$? Refer to this result as $\partial_5^{(i)}$. [1 marks]

(g) What is $\partial J/\partial W_1$? [2 marks]

Based on the given information, answer the following questions:

(a) What are the shapes of W_1 , b_1 , W_2 , b_2? If we were vectorizing this network across multiple examples, what would the shapes of the weights/biases be instead? If we were vectorizing across multiple examples, what would the shapes of X and Y be instead? [2 marks]

(b) What is $\partial J/\partial \hat{y}^{(i)}$? Refer to this result as $\partial_1^{(i)}$. Using this result, what is $\partial J/\partial \hat{y}$? [1 marks]

(c) What is $\partial \hat{y}^{(i)}/\partial z_2$? Refer to this result as $\partial_2^{(i)}$. [1 marks]

(d) What is $\partial z_2/\partial a_1$? Refer to this result as $\partial_3^{(i)}$. [1 marks]

(e) What is $\partial a_1/\partial z_1$? Refer to this result as $\partial_4^{(i)}$. [1 marks]

(f) What is $\partial z_1/\partial W_1$? Refer to this result as $\partial_5^{(i)}$. [1 marks]

(g) What is $\partial J/\partial W_1$? [2 marks]

Based on the given information, answer the following questions:

(a) What are the shapes of W_1 , b_1 , W_2 , b_2? If we were vectorizing this network across multiple examples, what would the shapes of the weights/biases be instead? If we were vectorizing across multiple examples, what would the shapes of X and Y be instead? [2 marks]

(b) What is $\partial J/\partial \hat{y}^{(i)}$? Refer to this result as $\partial_1^{(i)}$. Using this result, what is $\partial J/\partial \hat{y}$? [1 marks]

(c) What is $\partial \hat{y}^{(i)}/\partial z_2$? Refer to this result as $\partial_2^{(i)}$. [1 marks]

(d) What is $\partial z_2/\partial a_1$? Refer to this result as $\partial_3^{(i)}$. [1 marks]

(e) What is $\partial a_1/\partial z_1$? Refer to this result as $\partial_4^{(i)}$. [1 marks]

(f) What is $\partial z_1/\partial W_1$? Refer to this result as $\partial_5^{(i)}$. [1 marks]

(g) What is $\partial J/\partial W_1$? [2 marks]