SC4001 Exam 2019-2020 Sem 1

1. (a) State whether each of the following statements is "TRUE" or "FALSE". Each subquestion carries one mark.

(7 marks)

Answer

F (i) Gradient Descent (GD) learning always converges to the global minimum of the cost function.

F (ii) Stochastic Gradient Descent (SGD) learning generally finds better weights and biases than GD learning.

F (iii) A linear neuron can be trained to learn nonlinear mapping.

T (iv) Discrete perceptron learns a hyperplane as the decision boundary in the feature space.

F (v) Discrete perceptron learning is derived using GD learning.

T (vi) Weights of neurons are initialized to smaller values in order to operate in the linear region of sigmoid activation function.

F (vii) As the batch size is increased, the time to update weights drops in mini-batch GD learning.

1. (b) The softmax layer shown in Q1(b) has three neurons, receives 2-dimensional inputs $(x_1, x_2) \in \boldsymbol{R}^2$ and produces an output class label $y \in \{1, 2, 3\}$. Weights and biases to the neurons are given in the figure Q1b.

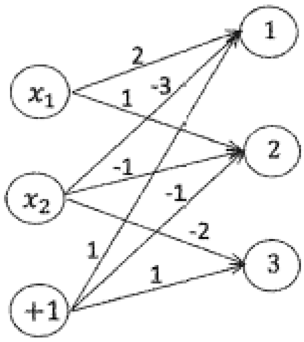


Figure Q1b

1. (b) (i) Write the weight vectors and biases of the neurons.

(2 marks)

Answer

$$W = \begin{pmatrix} 2 & 1 & 0 \\ -3 & 1 & -2 \end{pmatrix}$$

$$b = \begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix}$$

1. (b) (ii) Find the decision boundaries separating each pair of the classes.

(6 marks)

Answer

Neuron of class A, $u_1 = 2x_1 - 3x_2 + 1$

Neuron of class B, $u_2 = x_1 - x_2 - 1$

Neuron of class C $u_3 = -2x_2 + 1$

Decision boundaries:

Between class 1 and class 2 $u_1 = u_2$:

$$u_1 = u_2$$
$$2x_1 - 3x_2 + 1 = x_1 - x_2 - 1$$
$$2x_2 = x_1 + 2$$
$$x_2 = 0.5x_1 + 1$$

Between class 2 and class 3 $u_2 = u_3$:

$$u_2 = u_3$$
$$x_1 - x_2 - 1 = -2x_2 + 1$$
$$x_2 = -x_1 + 2$$

Between class 3 and class 1 $u_3 = u_1$:

$$u_3 = u_1$$
$$-2x_2 + 1 = 2x_1 - 3x_2 + 1$$
$$x_2 = 2x_1$$

1. (b) (iii) Plot the decision boundaries separating the three classes, clearly indicating the regions belonging to each class.

(6 marks)

Answer

Between class 1 and class 2 $u_1 = u_2$:

$x_2 = 0.5x_1 + 1$

plot

$when\ x_2 = 0, x_1 = -2$
$when\ x_1 = 0, x_2 = 1$

Between class 2 and class 3 $u_2 = u_3$:

$x_2 = 2x_1$

plot

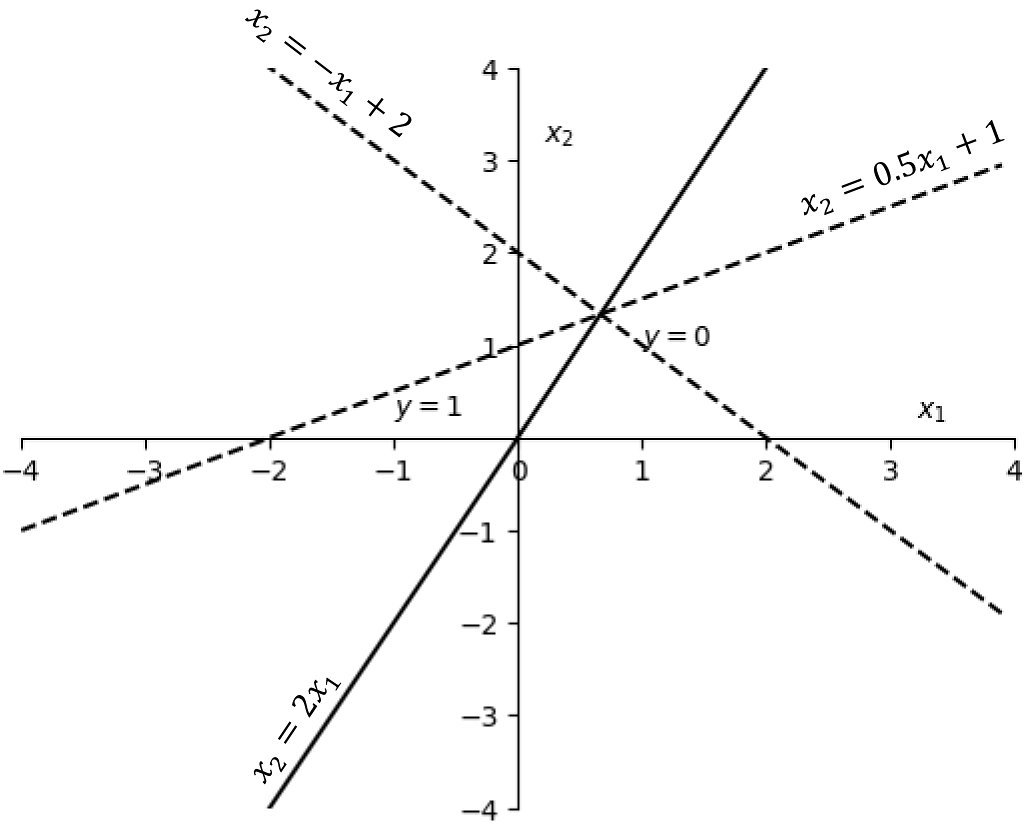
$when\ x_2 = 0, x_1 = 0$
$when\ x_1 = 1, x_2 = 2$

Between class 3 and class 1 $u_3 = u_1$:

$x_2 = -x_1 + 2$

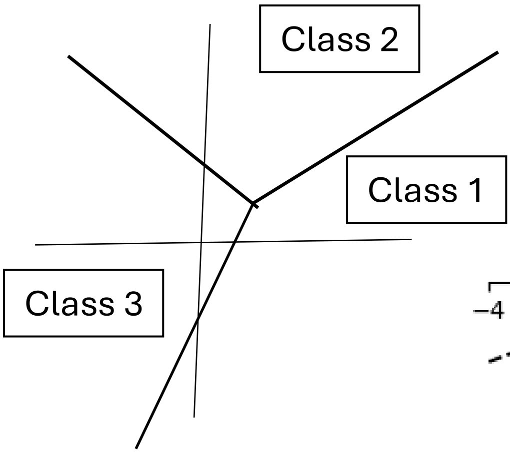
plot

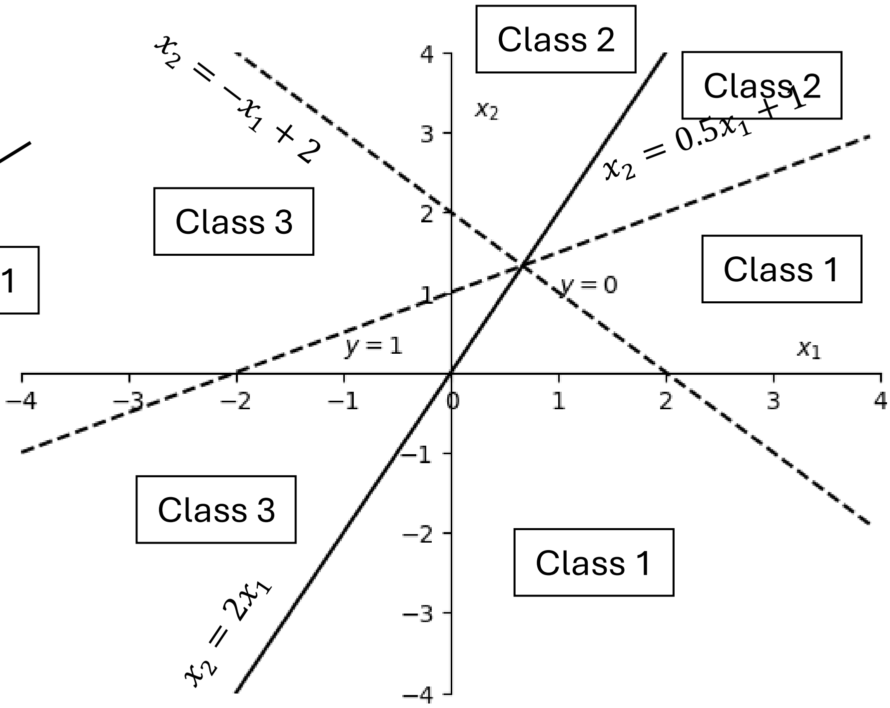
$when\ x_2 = 0, x_1 = 2$
$when\ x_1 = 0, x_2 = 2$



1. (b) (iii) cont

Answer





$$u_1 = 2x_1 - 3x_2 + 1$$
$$u_2 = x_1 - x_2 - 1$$
$$u_3 = -2x_2 + 1$$

Test point (x, y):

The u with highest value is the region

Test point (3, 3):

$$u_1 = 2(3) - 3(3) + 1 = -2$$
$$u_2 = (3) - (3) - 1 = -1$$
$$u_3 = -2(3) + 1 = -5$$

Region: Class 2

Test point (1, 3):

$$u_1 = 2(1) - 3(3) + 1 = -6$$
$$u_2 = (1) - (3) - 1 = -3$$
$$u_3 = -2(3) + 1 = -5$$

Region: Class 2

Test point (2, 1):

$$u_1 = 2(2) - 3(1) + 1 = 2$$
$$u_2 = (2) - (1) - 1 = 0$$
$$u_3 = -2(1) + 1 = -1$$

Region: Class 1

Test point (2, -1):

$$u_1 = 2(2) - 3(-1) + 1 = 8$$
$$u_2 = (2) - (-1) - 1 = 2$$
$$u_3 = -2(-1) + 1 = 3$$

Region: Class 1

Test point (-2, 2):

$$u_1 = 2(-2) - 3(2) + 1 = -9$$
$$u_2 = (-2) - (2) - 1 = -5$$
$$u_3 = -2(2) + 1 = -3$$

Region: Class 3

Test point (-2, -2):

$$u_1 = 2(-2) - 3(-2) + 1 = 3$$
$$u_2 = (-2) - (-2) - 1 = -1$$
$$u_3 = -2(-2) + 1 = 5$$

Region: Class 3

1. (b) (iv) Find the input that yields equal class probabilities.

(4 marks)

Answer

Take any 2 lines that intersect:

Decision boundary separating Class 1 and Class 2:

$x_2 = 0.5x_1 + 1$ ---------------------------- (1)

Decision boundary separating Class 1 and Class 3:

$x_2 = 2x_1$ ---------------------------- (2)

Substitute (2) into (1),

$$x_2 = 0.5x_1 + 1$$

$$2x_1 = \frac{1}{2}x_1 + 1$$

$$\frac{3}{2}x_1 = 1$$

$$x_1 = \frac{2}{3}$$ ---------------------------- (3)

Substitute (3) into (1),

$$x_2 = 0.5x_1 + 1$$

$$x_2 = \frac{1}{2}\left(\frac{2}{3}\right) + 1$$

$$x_2 = \frac{2}{6} + 1$$

$$x_2 = \frac{4}{3}$$

Input that yields equal class probabilities = (x_1, x_2) = $\left(\frac{2}{3}, \frac{4}{3}\right)$

2. The three-layer feedforward neural network shown in Figure Q2 receives 2-dimensional inputs $(x_1, x_2) \in \boldsymbol{R}^2$ and produces an output $y \in \{0, 1\}$. The hidden-layer consists of three perceptrons and the output layer consists of a logistic regression neuron. The weights and biases of the networks are initialized as indicated in the figure.

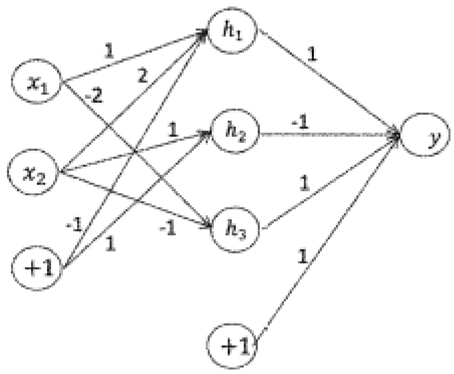


Figure Q2

The network is trained to produce a desired output $d_1 = 1$ and $d_2 = 0$ for inputs $x_1 = \begin{pmatrix}0\\1\end{pmatrix}$ and $x_2 = \begin{pmatrix}1\\0\end{pmatrix}$, respectively, by gradient decent learning. The learning factor $\alpha = 0.4$.

Perform one iteration of batch gradient descent learning. Give your answers up to two decimal places.

2. (a) Write initial weight matrices $\boldsymbol{W}$ and bias vector $\boldsymbol{b}$ of the hidden layer, and initial weight matrix $\boldsymbol{V}$ and bias vector $\boldsymbol{c}$ of the output layers.

(2 marks)

Answer

$$W = \begin{pmatrix}1 & 0 & -2\\2 & 1 & -1\end{pmatrix}, \quad b = \begin{pmatrix}-1\\1\\0\end{pmatrix}, \quad V = \begin{pmatrix}1\\-1\\1\end{pmatrix}, \quad c = (1)$$

2. (b) Write data matrix $\boldsymbol{X}$ and target matrix $\boldsymbol{D}$.

(1 marks)

Answer

$$x_1 = \begin{pmatrix}0\\1\end{pmatrix} \qquad x_2 = \begin{pmatrix}1\\0\end{pmatrix}, \qquad d_1 = 1 \qquad d_2 = 0$$

$$X = x^T = \begin{pmatrix}0 & 1\\1 & 0\end{pmatrix}, \qquad D = \begin{pmatrix}1\\0\end{pmatrix}$$

2. (c) Find synaptic input matrix $\boldsymbol{Z}$ and output matrix $\boldsymbol{H}$ of the hidden layer, and synaptic input matrix $\boldsymbol{U}$ and output matrix $\boldsymbol{Y}$ of the output layer.

(6 marks)

Answer

$$W = \begin{pmatrix} 1 & 0 & -2 \\ 2 & 1 & -1 \end{pmatrix}, \quad b = \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix}, \quad V = \begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix}, \quad c = (1) \quad X = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad D = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$$

$$B = b^T = \begin{pmatrix} -1 & 1 & 0 \\ -1 & 1 & 0 \end{pmatrix}, \qquad C = c^T = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

Synaptic input to hidden-layer,

$$\boldsymbol{Z} = \boldsymbol{XW} + B = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} 1 & 0 & -2 \\ 2 & 1 & -1 \end{pmatrix} + \begin{pmatrix} -1 & 1 & 0 \\ -1 & 1 & 0 \end{pmatrix}$$

$$= \begin{pmatrix} (0)(1)+(1)(2) & (0)(0)+(1)(1) & (0)(-2)+(1)(-1) \\ (1)(1)+(0)(2) & (1)(0)+(0)(1) & (1)(-2)+(0)(-1) \end{pmatrix} + \begin{pmatrix} -1 & 1 & 0 \\ -1 & 1 & 0 \end{pmatrix}$$

$$= \begin{pmatrix} 2 & 1 & -1 \\ 1 & 0 & -2 \end{pmatrix} + \begin{pmatrix} -1 & 1 & 0 \\ -1 & 1 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 2 & -1 \\ 0 & 1 & -2 \end{pmatrix}$$

Output of the hidden layer,

$$\boldsymbol{H} = g(\boldsymbol{z}) = \frac{1}{1+e^{-z}} = \begin{pmatrix} 0.73 & 0.88 & 0.27 \\ 0.50 & 0.73 & 0.12 \end{pmatrix}$$

Synaptic input to output-layer

$$\boldsymbol{U} = \boldsymbol{HV} + \boldsymbol{C} = \begin{pmatrix} 0.73 & 0.88 & 0.27 \\ 0.50 & 0.73 & 0.12 \end{pmatrix}\begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix} + \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

$$= \begin{pmatrix} (1)(0.73)+(-1)(0.88)+(1)(0.27) \\ (1)(0.50)+(-1)(0.73)+(1)(0.12) \end{pmatrix} + \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 0.12 \\ -0.11 \end{pmatrix} + \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 1.12 \\ 0.89 \end{pmatrix}$$

Output

$$Y = sigmoid(u) = \frac{1}{1+\ e^{-u}} = \begin{pmatrix} 0.75 \\ 0.71 \end{pmatrix}$$

<u>Summary</u>

synaptic input $\quad Z = \begin{pmatrix} 1 & 2 & -1 \\ 0 & 1 & -2 \end{pmatrix}$

output $\quad H = \begin{pmatrix} 0.73 & 0.88 & 0.27 \\ 0.50 & 0.73 & 0.12 \end{pmatrix}$

synaptic input $\quad U = \begin{pmatrix} 1.12 \\ 0.89 \end{pmatrix}$

output $\quad Y = \begin{pmatrix} 0.75 \\ 0.71 \end{pmatrix}$

2. (d) Find the cross-entropy cost J and classification error of the inputs.

(3 marks)

Answer

output $Y = \begin{pmatrix} 0.75 \\ 0.71 \end{pmatrix}, Y_1 = 0.75, Y_2 = 0.71$ $\qquad D = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$

$$J = \begin{cases} -\log\big(f(u)\big) & , \quad if\ d = 1 \\ -\log\big(1 - f(u)\big), & \quad if\ d = 0 \end{cases}$$

cross-entropy

$$J = \sum_{p=1}^{P} d_p \log\left(f(u_p)\right) - (1 - d_p)\log\left(1 - f(u_p)\right)$$
$$= -\log(0.75) - \log(1 - 0.71)$$
$$= 1.526$$

$$Y = 1(f(\boldsymbol{u}) > 0.5) = \left(\begin{pmatrix} 0.75 \\ 0.71 \end{pmatrix} > 0.5\right) = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

Classification error $\quad \sum_{p=1}^{2} 1\left(d_p \neq y_p\right) = 1$

<u>Summary</u>

cross-entropy $\quad J = 1.526$

Classification error = 1

2. (e) Find gradients $\nabla_{\boldsymbol{U}} J$ and $\nabla_{\boldsymbol{Z}} J$ of the cost J with respect to $\boldsymbol{U}$ and $\boldsymbol{Z}$, respectively.

(7 marks)

Answer

Backpropagation for FFN:

$$D = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \qquad Y = \begin{pmatrix} 0.75 \\ 0.71 \end{pmatrix} \qquad H = \begin{pmatrix} 0.73 & 0.88 & 0.27 \\ 0.50 & 0.73 & 0.12 \end{pmatrix} \quad V = \begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix}$$

Gradient $\nabla_{\boldsymbol{U}} J$,

$$\nabla_{\boldsymbol{U}} J = -(D - Y) = -\left(\begin{pmatrix} 1 \\ 0 \end{pmatrix} - \begin{pmatrix} 0.75 \\ 0.71 \end{pmatrix}\right) = -\begin{pmatrix} 0.25 \\ -0.71 \end{pmatrix} = \begin{pmatrix} -0.25 \\ 0.71 \end{pmatrix}$$

Gradient $\nabla_{\boldsymbol{Z}} J$,

$$\begin{aligned} g'(\boldsymbol{Z}) = \boldsymbol{H} \cdot (\boldsymbol{1} - \boldsymbol{H}) &= \begin{pmatrix} 0.73 & 0.88 & 0.27 \\ 0.50 & 0.73 & 0.12 \end{pmatrix} \cdot \left(\begin{pmatrix} 1 \\ 1 \end{pmatrix} - \begin{pmatrix} 0.73 & 0.88 & 0.27 \\ 0.50 & 0.73 & 0.12 \end{pmatrix}\right) \\ &= \begin{pmatrix} 0.73(1-0.73) & 0.88(1-0.88) & 0.27(1-0.27) \\ 0.50(1-0.50) & 0.73(1-0.73) & 0.12(1-0.12 \end{pmatrix} \\ &= \begin{pmatrix} 0.20 & 0.11 & 0.20 \\ 0.25 & 0.20 & 0.11 \end{pmatrix} \end{aligned}$$

$$\begin{aligned} \nabla_Z J = \nabla_u J V^T \cdot g'(\boldsymbol{Z}) &= \begin{pmatrix} -0.25 \\ 0.71 \end{pmatrix} (1 \quad -1 \quad 1) \cdot \begin{pmatrix} 0.20 & 0.11 & 0.20 \\ 0.25 & 0.20 & 0.11 \end{pmatrix} \\ &= \begin{pmatrix} (-0.25)(1) & (-0.25)(-1) & (-0.25)(1) \\ (\ 0.71)(1) & (\ 0.71)(-1) & (\ 0.71)(1) \end{pmatrix} \cdot \begin{pmatrix} 0.20 & 0.11 & 0.20 \\ 0.25 & 0.20 & 0.11 \end{pmatrix} \\ &= \begin{pmatrix} -0.25 & 0.25 & -0.25 \\ 0.71 & -0.71 & 0.71 \end{pmatrix} \cdot \begin{pmatrix} 0.20 & 0.11 & 0.20 \\ 0.25 & 0.20 & 0.11 \end{pmatrix} \\ &= \begin{pmatrix} -0.05 & 0.03 & -0.05 \\ 0.18 & -0.14 & 0.08 \end{pmatrix} \end{aligned}$$

<u>Summary</u>

Gradient $\nabla_{\boldsymbol{u}} J = \begin{pmatrix} -0.25 \\ 0.71 \end{pmatrix}$

Gradient $\nabla_Z J = \begin{pmatrix} -0.05 & 0.03 & -0.05 \\ 0.18 & -0.14 & 0.08 \end{pmatrix}$

2. (f) Find gradients $\nabla_{V}J$, $\nabla_{c}J$, $\nabla_{W}J$, and $\nabla_{b}J$ of the cost J with respect to $\boldsymbol{V}, \boldsymbol{c}, \boldsymbol{W}$, and $\boldsymbol{b}$, respectively.

(4 marks)

Answer

$$\nabla_{u}J = \begin{pmatrix} -0.25 \\ 0.71 \end{pmatrix} \qquad \boldsymbol{H} = \begin{pmatrix} 0.73 & 0.88 & 0.27 \\ 0.50 & 0.73 & 0.12 \end{pmatrix}$$

$$\nabla_{Z}J = \begin{pmatrix} -0.05 & 0.03 & -0.05 \\ 0.18 & -0.14 & 0.08 \end{pmatrix}$$

Output layer:

$$\nabla_{V}J = H^{T}\nabla_{U}J = \begin{pmatrix} 0.73 & 0.50 \\ 0.88 & 0.73 \\ 0.27 & 0.12 \end{pmatrix} \begin{pmatrix} -0.25 \\ 0.71 \end{pmatrix} = \begin{pmatrix} (0.73)(-0.25) + (0.50)(0.71) \\ (0.88)(-0.25) + (0.73)(0.71) \\ (0.27)(-0.25) + (0.12)(0.71) \end{pmatrix} = \begin{pmatrix} 0.17 \\ 0.30 \\ 0.02 \end{pmatrix}$$

$$\nabla_{c}J = (\nabla_{U}J)^{T} 1_{p} = (-0.25 \quad 0.71) \begin{pmatrix} 1 \\ 1 \end{pmatrix} = (-0.25)(1) + (0.71)(1) = 0.46$$

Hidden layer:

$$X = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

$$\nabla_{W}J = X^{T}(\nabla_{Z}J) = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} -0.05 & 0.03 & -0.05 \\ 0.18 & -0.14 & 0.08 \end{pmatrix}$$

$$= \begin{pmatrix} (0)(-0.05) + (1)(0.18) & (0)(0.03) + (1)(-0.14) & (0)(-0.05) + (1)(0.08) \\ (1)(-0.05) + (0)(0.18) & (1)(0.03) + (0)(-0.14) & (1)(-0.05) + (0)(0.08) \end{pmatrix}$$

$$= \begin{pmatrix} 0.18 & -0.14 & 0.08 \\ -0.05 & 0.03 & -0.05 \end{pmatrix}$$

$$\nabla_{b}J = (\nabla_{Z}J)^{T} 1_{p} = \begin{pmatrix} -0.05 & 0.18 \\ 0.03 & -0.14 \\ -0.05 & 0.08 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} (-0.05)(1) + (\ \ 0.18)(1) \\ (\ \ 0.03)(1) + (-0.14)(1) \\ (-0.05)(1) + (\ \ 0.08)(1) \end{pmatrix} = \begin{pmatrix} 0.13 \\ -0.11 \\ 0.03 \end{pmatrix}$$

Summary

$$\nabla_{V}J = \begin{pmatrix} 0.17 \\ 0.30 \\ 0.02 \end{pmatrix}$$

$$\nabla_{c}J = 0.46$$

$$\nabla_{W}J = \begin{pmatrix} 0.18 & -0.14 & 0.08 \\ -0.05 & 0.03 & -0.05 \end{pmatrix}$$

$$\nabla_{b}J = \begin{pmatrix} 0.13 \\ -0.11 \\ 0.03 \end{pmatrix}$$

2. (g) Find the updated values of $\boldsymbol{V}, \boldsymbol{c}, \boldsymbol{W}$, and $\boldsymbol{b}$.

(2 marks)

Answer

$$W = \begin{pmatrix} 1 & 0 & -2 \\ 2 & 1 & -1 \end{pmatrix}, \quad b = \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix}, \quad V = \begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix}, \quad c = (1) \qquad \alpha = 0.4$$

$$\nabla_V J = \begin{pmatrix} 0.17 \\ 0.30 \\ 0.02 \end{pmatrix}$$

$$\nabla_c J = 0.46$$

$$\nabla_W J = \begin{pmatrix} 0.18 & -0.14 & 0.08 \\ -0.05 & 0.03 & -0.05 \end{pmatrix}$$

$$\nabla_b J = \begin{pmatrix} 0.13 \\ -0.11 \\ 0.03 \end{pmatrix}$$

$$\boldsymbol{V} = \boldsymbol{V} - \alpha \nabla_V J = \begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix} - 0.4 \begin{pmatrix} 0.17 \\ 0.30 \\ 0.02 \end{pmatrix} = \begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix} - \begin{pmatrix} 0.07 \\ 0.12 \\ 0.01 \end{pmatrix} = \begin{pmatrix} 0.93 \\ -1.12 \\ 0.99 \end{pmatrix}$$

$$\boldsymbol{c} = \boldsymbol{c} - \alpha \nabla_c J = (1) - 0.4(0.46) = (1) - (0.18) = 0.82$$

$$\boldsymbol{W} = W - \alpha \nabla_W J = \begin{pmatrix} 1 & 0 & -2 \\ 2 & 1 & -1 \end{pmatrix} - 0.4 \begin{pmatrix} 0.18 & -0.14 & 0.08 \\ -0.05 & 0.03 & -0.05 \end{pmatrix}$$

$$= \begin{pmatrix} 1 & 0 & -2 \\ 2 & 1 & -1 \end{pmatrix} - \begin{pmatrix} 0.07 & -0.06 & 0.03 \\ -0.02 & 0.01 & -0.02 \end{pmatrix} = \begin{pmatrix} 0.93 & 0.06 & -2.03 \\ 2.02 & 0.99 & -0.98 \end{pmatrix}$$

$$\boldsymbol{b} = \boldsymbol{b} - \alpha \nabla_b J = \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix} - 0.4 \begin{pmatrix} 0.13 \\ -0.11 \\ 0.03 \end{pmatrix} = \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix} - \begin{pmatrix} 0.05 \\ -0.04 \\ 0.01 \end{pmatrix} = \begin{pmatrix} -1.05 \\ 1.04 \\ 0.01 \end{pmatrix}$$

Summary

Updated $V = \begin{pmatrix} 0.93 \\ -1.12 \\ 0.99 \end{pmatrix}$

Updated $\boldsymbol{c} = 0.82$

Updated $W = \begin{pmatrix} 0.93 & 0.06 & -2.03 \\ 2.02 & 0.99 & -0.98 \end{pmatrix}$

Updated $\boldsymbol{b} = \begin{pmatrix} -1.05 \\ 1.04 \\ 0.01 \end{pmatrix}$

3. (a) (The following input I is applied to the input layer of a convolutional neural network (CNN).

$$\boldsymbol{I} = \begin{pmatrix} -0.5 & -0.3 & 0.4 & -0.2 & 1.5 & 1.0 & -0.1 & 1.0 \\ 1.0 & 0.1 & 0.5 & 0.2 & 1.2 & -1.0 & 0.1 & 0.3 \\ -0.5 & -0.2 & 1.5 & 0.7 & 0.9 & 0.6 & -0.5 & -0.4 \\ 1.6 & -0.4 & 0.4 & -1.0 & -0.1 & 0.5 & 1.1 & 0.0 \end{pmatrix}$$

The input is processed by a convolution layer of neurons having weights $w = \begin{pmatrix} 1.0 & -0.2 \\ 0.2 & 0.5 \end{pmatrix}$ and rectified linear unit (ReLU) activation functions, and thereafter an average pooling layer having a pooling window of 2 × 2 size. The convolution is performed with stride = 2 and with 'VALID' padding. The bias connected to the convolution layer is 0.1.

Find the feature maps at the first convolution layer and pooling layer.

(10 marks)

Answer

- For 'VALID' padding, padding = 0
- stride = 2
- bias = 0.1

convolution size:

$$\text{row} = \frac{R - F + 2P}{S} + 1 = \frac{4 - 2 + 2(0)}{2} + 1 = 2$$

$$\text{col} = \frac{C - F + 2P}{S} + 1 = \frac{8 - 2 + 2(0)}{2} + 1 = 4$$

Synaptic inputs to the feature map with filter w: $u = Conv(\boldsymbol{I}, w) + b$

$$u(1,1) = (-0.5)(1.0) + (-0.3)(-0.2) + (1.0)(0.2) + (0.1)(0.5) + 0.1$$
$$= -0.5 + 0.06 + 0.2 + 0.05 + 0.1 = -0.09$$

$$u(1,2) = (0.4)(1.0) + (-0.2)(-0.2) + (0.5)(0.2) + (0.2)(0.5) + 0.1$$
$$= 0.4 + 0.04 + 0.1 + 0.1 + 0.1 = 0.74$$

$$u(1,3) = (1.5)(1.0) + (1.0)(-0.2) + (1.2)(0.2) + (-1.0)(0.5) + 0.1$$
$$= 1.5 - 0.2 + 0.24 - 0.5 + 0.1 = 1.14$$

$$u(1,4) = (-0.1)(1.0) + (1.0)(-0.2) + (0.1)(0.2) + (0.3)(0.5) + 0.1$$
$$= -0.1 - 0.2 + 0.02 + 0.15 + 0.1 = -0.03$$

$$u(2,1) = (-0.5)(1.0) + (-0.2)(-0.2) + (1.6)(0.2) + (-0.4)(0.5) + 0.1$$
$$= -0.5 + 0.04 + 0.32 - 0.2 + 0.1 = -0.24$$

$$u(2,2) = (1.5)(1.0) + (0.7)(-0.2) + (0.4)(0.2) + (-1.0)(0.5) + 0.1$$
$$= 1.5 - 0.14 + 0.08 - 0.5 + 0.1 = 1.04$$

$$u(2,3) = (0.9)(1.0) + (0.6)(-0.2) + (-0.1)(0.2) + (0.5)(0.5) + 0.1$$
$$= 0.9 - 0.12 - 0.02 + 0.25 + 0.1 = 1.11$$

$$u(2,4) = (-0.5)(1.0) + (-0.4)(-0.2) + (1.1)(0.2) + (0.0)(0.5) + 0.1$$
$$= -0.5 + 0.08 + 0.22 + 0 + 0.1 = -0.1$$

3. (a) cont

Answer

$$u = \begin{pmatrix} -0.09 & 0.74 & 1.14 & -0.03 \\ -0.24 & 1.04 & 1.11 & -0.10 \end{pmatrix}$$

Feature maps at the convolution layer:

$$y = relu(u) = \max\{0, u\} = \left(\begin{array}{cc|cc} 0 & 0.74 & 1.14 & 0 \\ 0 & 1.04 & 1.11 & 0 \end{array}\right)$$

Pooling 2×2 and strides = 2

$$ave(1,1) = \frac{0 + 0.74 + 0 + 1.04}{4} = \frac{1.78}{4} = 0.45$$

$$ave(1,2) = \frac{1.14 + 0 + 1.11 + 0}{4} = \frac{2.25}{4} = 0.56$$

pooling size:

$$\text{row} = \frac{R - F}{S} + 1 = \frac{2 - 2}{2} + 1 = 1$$

$$\text{col} = \frac{C - F}{S} + 1 = \frac{4 - 2}{2} + 1 = 2$$

Feature maps at the pooling layer: $P_{ave} = (0.45 \quad 0.56)$

<u>Summary</u>

Feature maps at the first convolution layer: $u = \begin{pmatrix} 0 & 0.74 & 1.14 & 0 \\ 0 & 1.04 & 1.11 & 0 \end{pmatrix}$

Feature maps at the pooling layer:

$$P_{ave} = \begin{pmatrix} \frac{0 + 0.74 + 0 + 1.04}{4} & \frac{1.14 + 0 + 1.11 + 0}{4} \end{pmatrix} = (0.45 \quad 0.56)$$

3. (b) A trained CNN is found to contain some dead neurons in its hidden layers. Neurons in these hidden layers use ReLU activation functions. Explain this phenomenon and provide a way to avoid dead neurons during training.

(2 marks)

Answer

- Dead neurons are neurons that always output zero, regardless of the input.
- This happens when the ReLU activation receives only negative inputs.
- To avoid dead neurons during training:
  - Replace ReLU with Leaky ReLU, which allows a small, non-zero gradient when the input is negative, helping neurons continue learning.

$$\text{Leaky ReLU}(x) = \begin{cases} x, & if\ x > 0 \\ 0.01x, & if\ x \leq 0 \end{cases}$$

3. (c) A convolutional neural network has five consecutive 3×3 convolutional layers with stride 1 and no pooling. How large is the receptive field (in terms of number of pixels) of a neuron in the fifth non-image layer of this network?

(3 marks)

Answer

- n = 5
- k = 3

- Initial receptive field: 1 pixel (the neuron itself)

$$\mathrm{RF} = 1 + n * (k - 1) = 1 + 5 * (3 - 1) = 1 + (5 * 2) = 11$$

- Receptive field size = 11×11

3. (d) An autoencoder with a single hidden-layer is used to reconstruct 4- dimensional input patterns $\boldsymbol{x} \in \boldsymbol{R}^4$. The hidden-layer has three neurons with activation functions $\phi(u) = \frac{e^u - e^{-u}}{e^u + e^{-u}}$ and the output layer has four $\sigma(u) = \frac{1}{1+e^{-u}}$. The weight matrix $\boldsymbol{W}$ connected to the hidden layer, the bias vector $\boldsymbol{b}$ of the hidden-layer, and bias vector $\boldsymbol{c}$ of the output layer are given by

$$W = \begin{pmatrix} -2.0 & 1.2 & -0.8 \\ 0.5 & 1.5 & 2.2 \\ -2.2 & 3.2 & -1.2 \\ 2.0 & -1.0 & -3.6 \end{pmatrix},\ b = \begin{pmatrix} 0.2 \\ -0.5 \\ 0.0 \end{pmatrix},\ \text{and}\ c = \begin{pmatrix} 0.5 \\ 1.0 \\ -0.6 \\ 2.0 \end{pmatrix}$$

For input patterns $\boldsymbol{x}_1 = \begin{pmatrix} 1 \\ 1 \\ 0 \\ 0 \end{pmatrix}$ and $\boldsymbol{x}_2 = \begin{pmatrix} 0 \\ 1 \\ 1 \\ 1 \end{pmatrix}$, find

3. (d) (i) the hidden layer activations.

(3 marks)

Answer

$$W = \begin{pmatrix} -2.0 & 1.2 & -0.8 \\ 0.5 & 1.5 & 2.2 \\ -2.2 & 3.2 & -1.2 \\ 2.0 & -1.0 & -3.6 \end{pmatrix}, \; b = \begin{pmatrix} 0.2 \\ -0.5 \\ 0.0 \end{pmatrix}, \; c = \begin{pmatrix} 0.5 \\ 1.0 \\ -0.6 \\ 2.0 \end{pmatrix} \quad \boldsymbol{x}_1 = \begin{pmatrix} 1 \\ 1 \\ 0 \\ 0 \end{pmatrix} \quad \boldsymbol{x}_2 = \begin{pmatrix} 0 \\ 1 \\ 1 \\ 1 \end{pmatrix}$$

$$b = \begin{pmatrix} 0.2 \\ -0.5 \\ 0.0 \end{pmatrix} \qquad c = \begin{pmatrix} 0.5 \\ 1.0 \\ -0.6 \\ 2.0 \end{pmatrix}$$

$$\phi(u) = \tanh(u) \quad = \frac{e^u - e^{-u}}{e^u + e^{-u}}$$

$$\sigma(u) = \text{sigmoid}(u) = \frac{1}{1 + e^{-u}}$$

For input $\boldsymbol{x}_1 = \begin{pmatrix} 1 \\ 1 \\ 0 \\ 0 \end{pmatrix}$,

$$u_1 = W^T x_1 + b = \begin{pmatrix} -2.0 & 0.5 & -2.2 & 2.0 \\ 1.2 & 1.5 & 3.2 & -1.0 \\ -0.8 & 2.2 & -1.2 & -3.6 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \\ 0 \\ 0 \end{pmatrix} + \begin{pmatrix} 0.2 \\ -0.5 \\ 0.0 \end{pmatrix}$$

$$= \begin{pmatrix} (-2.0)(1) + (0.5)(1) + (-2.2)(0) + (\ \ 2.0)(0) \\ (\ \ 1.2)(1) + (1.5)(1) + (\ \ 3.2)(0) + (-1.0)(0) \\ (-0.8)(1) + (2.2)(1) + (-1.2)(0) + (-3.6)(0) \end{pmatrix} + \begin{pmatrix} 0.2 \\ -0.5 \\ 0.0 \end{pmatrix}$$

$$= \begin{pmatrix} -1.5 \\ 2.7 \\ 1.4 \end{pmatrix} + \begin{pmatrix} 0.2 \\ -0.5 \\ 0.0 \end{pmatrix} = \begin{pmatrix} -1.3 \\ 2.2 \\ 1.4 \end{pmatrix}$$

$$h_1 = \phi(u_1) = \tanh(u) = \frac{e^u - e^{-u}}{e^u + e^{-u}} = \begin{pmatrix} -0.86 \\ 0.98 \\ 0.89 \end{pmatrix}$$

3. (d) (i) cont

Answer

For input $\boldsymbol{x}_2 = \begin{pmatrix} 0 \\ 1 \\ 1 \\ 1 \end{pmatrix}$,

$$u_2 = W^T x_2 + b = \begin{pmatrix} -2.0 & 0.5 & -2.2 & 2.0 \\ 1.2 & 1.5 & 3.2 & -1.0 \\ -0.8 & 2.2 & -1.2 & -3.6 \end{pmatrix} \begin{pmatrix} 0 \\ 1 \\ 1 \\ 1 \end{pmatrix} + \begin{pmatrix} 0.2 \\ -0.5 \\ 0.0 \end{pmatrix}$$

$$= \begin{pmatrix} (-2.0)(0) + (0.5)(1) + (-2.2)(1) + (\ \ 2.0)(1) \\ (\ \ 1.2)(0) + (1.5)(1) + (\ \ 3.2)(1) + (-1.0)(1) \\ (-0.8)(0) + (2.2)(1) + (-1.2)(1) + (-3.6)(1) \end{pmatrix} + \begin{pmatrix} 0.2 \\ -0.5 \\ 0.0 \end{pmatrix}$$

$$= \begin{pmatrix} 0.3 \\ 3.7 \\ -2.6 \end{pmatrix} + \begin{pmatrix} 0.2 \\ -0.5 \\ 0.0 \end{pmatrix} = \begin{pmatrix} 0.5 \\ 3.2 \\ -2.6 \end{pmatrix}$$

$$h_2 = \phi(u_2) = \tanh(u) = \frac{e^u - e^{-u}}{e^u + e^{-u}} = \begin{pmatrix} 0.46 \\ 1.00 \\ -0.99 \end{pmatrix}$$

<u>Summary</u>

hidden activation for $\boldsymbol{x}_1$:

$$h_1 = \begin{pmatrix} -0.86 \\ 0.98 \\ 0.89 \end{pmatrix}$$

hidden activation for $\boldsymbol{x}_2$:

$$h_2 = \begin{pmatrix} 0.46 \\ 1.00 \\ -0.99 \end{pmatrix}$$

3. (d) (ii) the reconstruction errors.

(4 marks)

Answer

$$W = \begin{pmatrix} -2.0 & 1.2 & -0.8 \\ 0.5 & 1.5 & 2.2 \\ -2.2 & 3.2 & -1.2 \\ 2.0 & -1.0 & -3.6 \end{pmatrix}, \; b = \begin{pmatrix} 0.2 \\ -0.5 \\ 0.0 \end{pmatrix}, \; c = \begin{pmatrix} 0.5 \\ 1.0 \\ -0.6 \\ 2.0 \end{pmatrix} \quad \boldsymbol{x}_1 = \begin{pmatrix} 1 \\ 1 \\ 0 \\ 0 \end{pmatrix} \quad \boldsymbol{x}_2 = \begin{pmatrix} 0 \\ 1 \\ 1 \\ 1 \end{pmatrix}$$

$$h_1 = \begin{pmatrix} -0.86 \\ 0.98 \\ 0.89 \end{pmatrix} \quad h_2 = \begin{pmatrix} 0.46 \\ 1.00 \\ -0.99 \end{pmatrix}$$

For $\boldsymbol{x}_1 = \begin{pmatrix} 1 \\ 1 \\ 0 \\ 0 \end{pmatrix}$ P = 4

$$u_1 = Wh_1 + c = \begin{pmatrix} -2.0 & 1.2 & -0.8 \\ 0.5 & 1.5 & 2.2 \\ -2.2 & 3.2 & -1.2 \\ 2.0 & -1.0 & -3.6 \end{pmatrix} \begin{pmatrix} -0.86 \\ 0.98 \\ 0.89 \end{pmatrix} + \begin{pmatrix} 0.5 \\ 1.0 \\ -0.6 \\ 2.0 \end{pmatrix}$$

$$= \begin{pmatrix} (-2.0)(-0.86) + (\ 1.2)(0.98) + (-0.8)(0.89) \\ (\ 0.5)(-0.86) + (\ 1.5)(0.98) + (\ 2.2)(0.89) \\ (-2.2)(-0.86) + (\ 3.2)(0.98) + (-1.2)(0.89) \\ (\ 2.0)(-0.86) + (-1.0)(0.98) + (-3.6)(0.89) \end{pmatrix} + \begin{pmatrix} 0.5 \\ 1.0 \\ -0.6 \\ 2.0 \end{pmatrix}$$

$$= \begin{pmatrix} 2.18 \\ 3.00 \\ 3.96 \\ -5.90 \end{pmatrix} + \begin{pmatrix} 0.5 \\ 1.0 \\ -0.6 \\ 2.0 \end{pmatrix} = \begin{pmatrix} 2.68 \\ 4.00 \\ 3.36 \\ -3.90 \end{pmatrix}$$

$$y_1 = \sigma(u_1) = \mathrm{sigmoid}(u) = \frac{1}{1 + e^{-u}} = \begin{pmatrix} 0.94 \\ 0.98 \\ 0.97 \\ 0.02 \end{pmatrix}$$

Reconstruction error (MSE):

$$J_{mse1} = \frac{1}{P} \sum_{p=1}^{P} \left\| \boldsymbol{y}_p - \boldsymbol{x}_p \right\|^2$$

$$= \frac{1}{4} [\|0.94 - 1\|^2 + \|0.98 - 1\|^2 + \|0.97 - 0\|^2 + \|0.02 - 0\|^2]$$

$$= \frac{1}{4} [(0.06)^2 + (0.02)^2 + (0.97)^2 + (0.02)^2] = \frac{1}{4}(0.95) = 0.24$$

3. (d) (ii) cont

Answer

$$W = \begin{pmatrix} -2.0 & 1.2 & -0.8 \\ 0.5 & 1.5 & 2.2 \\ -2.2 & 3.2 & -1.2 \\ 2.0 & -1.0 & -3.6 \end{pmatrix},\ b = \begin{pmatrix} 0.2 \\ -0.5 \\ 0.0 \end{pmatrix},\ c = \begin{pmatrix} 0.5 \\ 1.0 \\ -0.6 \\ 2.0 \end{pmatrix} \quad \boldsymbol{x}_1 = \begin{pmatrix} 1 \\ 1 \\ 0 \\ 0 \end{pmatrix} \quad \boldsymbol{x}_2 = \begin{pmatrix} 0 \\ 1 \\ 1 \\ 1 \end{pmatrix}$$

$$h_1 = \begin{pmatrix} -0.86 \\ 0.98 \\ 0.89 \end{pmatrix} \quad h_2 = \begin{pmatrix} 0.46 \\ 1.00 \\ -0.99 \end{pmatrix}$$

For $\boldsymbol{x}_2 = \begin{pmatrix} 0 \\ 1 \\ 1 \\ 1 \end{pmatrix}$ P = 4

$$u_2 = Wh_2 + c = \begin{pmatrix} -2.0 & 1.2 & -0.8 \\ 0.5 & 1.5 & 2.2 \\ -2.2 & 3.2 & -1.2 \\ 2.0 & -1.0 & -3.6 \end{pmatrix} \begin{pmatrix} 0.46 \\ 1.00 \\ -0.99 \end{pmatrix} + \begin{pmatrix} 0.5 \\ 1.0 \\ -0.6 \\ 2.0 \end{pmatrix}$$

$$= \begin{pmatrix} (-2.0)(0.46) + (\ 1.2)(1) + (-0.8)(-0.99) \\ (\ 0.5)(0.46) + (\ 1.5)(1) + (\ 2.2)(-0.99) \\ (-2.2)(0.46) + (\ 3.2)(1) + (-1.2)(-0.99) \\ (\ 2.0)(0.46) + (-1.0)(1) + (-3.6)(-0.99) \end{pmatrix} + \begin{pmatrix} 0.5 \\ 1.0 \\ -0.6 \\ 2.0 \end{pmatrix}$$

$$= \begin{pmatrix} 1.07 \\ -0.45 \\ 3.38 \\ 3.48 \end{pmatrix} + \begin{pmatrix} 0.5 \\ 1.0 \\ -0.6 \\ 2.0 \end{pmatrix} = \begin{pmatrix} 1.57 \\ 0.55 \\ 2.78 \\ 5.48 \end{pmatrix}$$

$$y_1 = \sigma(u_1) = \mathrm{sigmoid}(u) = \frac{1}{1 + e^{-u}} = \begin{pmatrix} 0.83 \\ 0.63 \\ 0.94 \\ 1.00 \end{pmatrix}$$

Reconstruction error (MSE):

$$J_{mse2} = \frac{1}{P} \sum_{p=1}^{P} \left\| \boldsymbol{y}_p - \boldsymbol{x}_p \right\|^2$$

$$= \frac{1}{4} [\|0.83 - 0\|^2 + \|0.63 - 1\|^2 + \|0.94 - 1\|^2 + \|1 - 1\|^2]$$

$$= \frac{1}{4} [(0.83)^2 + (0.37)^2 + (0.06)^2 + (0)^2] = \frac{1}{4}(0.83) = 0.21$$

Summary

Reconstruction error (MSE):

$J_{mse1} = 0.24$

$J_{mse2} = 0.21$

3. (e) Undercomplete autoencoders are capable of learning hidden structures of data. Give the reasons.

(3 marks)

Answer

- In undercomplete autoencoders, the hidden-layer has a lower dimension than the input layer.

- By learning to approximate an n-dimensional inputs with M ($< n$) number of hidden units, we obtain *a lower dimensional representation* of the input signals.

- The network reconstructs the input signals from the reduced-dimensional hidden representation.

- Learning an undercomplete representation forces the autoencoder to capture the most salient features.

- By limiting the number of hidden neurons, hidden structures of input data can be inferred from autoencoders.

- For example, correlations among input variables, learning principal components of data, etc.

4. (a) Consider an Elman-type recurrent neural network (RNN) that receives 2-dimensional input patterns $\boldsymbol{x} \in \boldsymbol{R}^2$ and has one hidden layer, which is initialized to zeros. The RNN has three neurons in the hidden layer and one neuron in the output layer. The hidden layer neurons have tanh activation functions and the output layer neurons use logistic activation functions.

The weight matrices $\boldsymbol{U}$ connecting the input to the hidden layer, $\boldsymbol{W}$ connecting the previous hidden state to the next hidden state, and $\boldsymbol{V}$ connecting the hidden output to the output layer are given by

$$\boldsymbol{U} = \begin{pmatrix} -1.0 & 1.5 & 2.0 \\ 0.2 & 1.0 & -1.0 \end{pmatrix}, \boldsymbol{W} = \begin{pmatrix} -1.0 & -0.6 & -0.2 \\ 1.5 & 1.7 & 1.5 \\ -0.9 & 2.0 & 0.3 \end{pmatrix} \text{ and } \boldsymbol{V} = \begin{pmatrix} 0.3 \\ 1.0 \\ 0.4 \end{pmatrix}$$

All bias connections to neurons are set to 0.2.

Find the output of the network for a sequence $(\boldsymbol{x}(1), \boldsymbol{x}(2), \boldsymbol{x}(3))$ of input patterns.

$$\boldsymbol{x}(1) = \begin{pmatrix} 0.5 \\ 1.5 \end{pmatrix}, x(2) = \begin{pmatrix} 1.0 \\ -1.0 \end{pmatrix} \text{ and } \boldsymbol{x}(3) = \begin{pmatrix} 2.5 \\ -2.0 \end{pmatrix}$$

(15 marks)

Answer

$$\boldsymbol{U} = \begin{pmatrix} -1.0 & 1.5 & 2.0 \\ 0.2 & 1.0 & -1.0 \end{pmatrix}, \boldsymbol{W} = \begin{pmatrix} -1.0 & -0.6 & -0.2 \\ 1.5 & 1.7 & 1.5 \\ -0.9 & 2.0 & 0.3 \end{pmatrix} \text{ and } \boldsymbol{V} = \begin{pmatrix} 0.3 \\ 1.0 \\ 0.4 \end{pmatrix}$$

$$\boldsymbol{x}(1) = \begin{pmatrix} 0.5 \\ 1.5 \end{pmatrix}, x(2) = \begin{pmatrix} 1.0 \\ -1.0 \end{pmatrix} \text{ and } \boldsymbol{x}(3) = \begin{pmatrix} 2.5 \\ -2.0 \end{pmatrix}$$

$$b = \begin{pmatrix} 0.2 \\ 0.2 \\ 0.2 \end{pmatrix}, \quad h = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \qquad c = b = 0.2$$

$$\phi(u) = \tanh(u) \qquad = \frac{e^u - e^{-u}}{e^u + e^{-u}}$$

$$\sigma(u) = \text{sigmoid}(u) = \frac{1}{1 + e^{-u}}$$

4. (a) cont

Answer

$$\phi(u) = \tanh(u) = \frac{e^{u} - e^{-u}}{e^{u} + e^{-u}}$$

$$\sigma(u) = \text{sigmoid}(u) = \frac{1}{1 + e^{-u}}$$

At $t = 1$, $\boldsymbol{x}(1) = \begin{pmatrix} 0.5 \\ 1.5 \end{pmatrix}$,

$$\boldsymbol{h}(t) = \phi(\boldsymbol{U}^T \boldsymbol{x}(t) + \boldsymbol{W}^T \boldsymbol{h}(t-1) + \boldsymbol{b})$$

$$\boldsymbol{h}(1) = tanh(\boldsymbol{U}^T \boldsymbol{x}(1) + \boldsymbol{W}^T \boldsymbol{h}(0) + \boldsymbol{b})$$

$$= \tanh\left(\begin{pmatrix} -1.0 & 0.2 \\ 1.5 & 1.0 \\ 2.0 & -1.0 \end{pmatrix}\begin{pmatrix} 0.5 \\ 1.5 \end{pmatrix} + \begin{pmatrix} -1.0 & 1.5 & -0.9 \\ -0.6 & 1.7 & 2.0 \\ -0.2 & 1.5 & 0.3 \end{pmatrix}\begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} + \begin{pmatrix} 0.2 \\ 0.2 \\ 0.2 \end{pmatrix}\right)$$

$$= \tanh\left(\begin{pmatrix} (-1.0)(0.5) + (\ \ 0.2)(1.5) \\ (\ \ 1.5)(0.5) + (\ \ 1.0)(1.5) \\ (\ \ 2.0)(0.5) + (-1.0)(1.5) \end{pmatrix} + \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} + \begin{pmatrix} 0.2 \\ 0.2 \\ 0.2 \end{pmatrix}\right)$$

$$= \tanh\left(\begin{pmatrix} -0.20 \\ 2.25 \\ -0.50 \end{pmatrix} + \begin{pmatrix} 0.2 \\ 0.2 \\ 0.2 \end{pmatrix}\right) = \tanh\begin{pmatrix} 0 \\ 2.45 \\ -0.30 \end{pmatrix} = \begin{pmatrix} 0 \\ 0.9852 \\ -0.2913 \end{pmatrix}$$

$$\boldsymbol{y}(t) = \sigma(\boldsymbol{V}^T \boldsymbol{h}(t) + c)$$

$$\boldsymbol{y}(1) = sigmoid(\boldsymbol{V}^T \boldsymbol{h}(1) + c)$$

$$= \text{sigmoid}\left((0.3 \quad 1.0 \quad 0.4)\begin{pmatrix} 0 \\ 0.9852 \\ -0.2913 \end{pmatrix} + 0.2\right)$$

$$= \text{sigmoid}((0.3)(0) + (1)(0.9852) + (0.4)(-0.2913) + 0.2)$$

$$= \text{sigmoid}(1.0687) = 0.7443$$

4. (a) cont

Answer

At $t = 2$, $x(2) = \begin{pmatrix} 1.0 \\ -1.0 \end{pmatrix}$,

$$\phi(u) = \tanh(u) \quad = \frac{e^u - e^{-u}}{e^u + e^{-u}}$$

$$\sigma(u) = \text{sigmoid}(u) = \frac{1}{1 + e^{-u}}$$

$$\boldsymbol{h}(t) = \quad \phi(\boldsymbol{U}^T\boldsymbol{x}(t) + \boldsymbol{W}^T\boldsymbol{h}(t-1) + \boldsymbol{b})$$

$$\boldsymbol{h}(2) = tanh(\boldsymbol{U}^T\boldsymbol{x}(2) + \boldsymbol{W}^T\boldsymbol{h}(1) + \boldsymbol{b})$$

$$= \tanh\left(\begin{pmatrix} -1.0 & 0.2 \\ 1.5 & 1.0 \\ 2.0 & -1.0 \end{pmatrix}\begin{pmatrix} 1.0 \\ -1.0 \end{pmatrix} + \begin{pmatrix} -1.0 & 1.5 & -0.9 \\ -0.6 & 1.7 & 2.0 \\ -0.2 & 1.5 & 0.3 \end{pmatrix}\begin{pmatrix} 0 \\ 0.9852 \\ -0.2913 \end{pmatrix} + \begin{pmatrix} 0.2 \\ 0.2 \\ 0.2 \end{pmatrix}\right)$$

$$= \tanh\left(\begin{pmatrix} (-1.0)(1) + (\ \ 0.2)(-1) \\ (\ \ 1.5)(1) + (\ \ 1.0)(-1) \\ (\ \ 2.0)(1) + (-1.0)(-1) \end{pmatrix} + \begin{pmatrix} (-1.0)(0) + (1.5)(0.9852) + (-0.9)(-0.2913) \\ (-0.6)(0) + (1.7)(0.9852) + (\ \ 2.0)(-0.2913) \\ (-0.2)(0) + (1.5)(0.9852) + (\ \ 0.3)(-0.2913) \end{pmatrix} + \begin{pmatrix} 0.2 \\ 0.2 \\ 0.2 \end{pmatrix}\right)$$

$$= \tanh\left(\begin{pmatrix} -1.20 \\ 0.50 \\ 3.00 \end{pmatrix} + \begin{pmatrix} 1.7400 \\ 1.0922 \\ 1.3904 \end{pmatrix} + \begin{pmatrix} 0.2 \\ 0.2 \\ 0.2 \end{pmatrix}\right) = \tanh\begin{pmatrix} 0.7400 \\ 1.7922 \\ 4.5904 \end{pmatrix} = \begin{pmatrix} 0.6291 \\ 0.9460 \\ 0.9998 \end{pmatrix}$$

$$\boldsymbol{y}(t) = \quad \sigma(\boldsymbol{V}^T\boldsymbol{h}(t) + c)$$

$$\boldsymbol{y}(2) = sigmoid(\boldsymbol{V}^T\boldsymbol{h}(2) + c)$$

$$= \text{sigmoid}\left((0.3 \quad 1.0 \quad 0.4)\begin{pmatrix} 0.6291 \\ 0.9460 \\ 0.9998 \end{pmatrix} + 0.2\right)$$

$$= \text{sigmoid}((0.3)(0.6291) + (1)(0.9460) + (0.4)(0.9998) + 0.2)$$

$$= \text{sigmoid}(1.7347) = 0.8500$$

4. (a) cont

Answer

At $t = 3$, $\boldsymbol{x}(3) = \begin{pmatrix} 2.5 \\ -2.0 \end{pmatrix}$,

$$\phi(u) = \tanh(u) = \frac{e^u - e^{-u}}{e^u + e^{-u}}$$

$$\sigma(u) = \text{sigmoid}(u) = \frac{1}{1 + e^{-u}}$$

$$\boldsymbol{h}(t) = \phi(\boldsymbol{U}^T\boldsymbol{x}(t) + \boldsymbol{W}^T\boldsymbol{h}(t-1) + \boldsymbol{b})$$

$$\boldsymbol{h}(3) = tanh(\boldsymbol{U}^T\boldsymbol{x}(3) + \boldsymbol{W}^T\boldsymbol{h}(2) + \boldsymbol{b})$$

$$= \tanh\left(\begin{pmatrix} -1.0 & 0.2 \\ 1.5 & 1.0 \\ 2.0 & -1.0 \end{pmatrix}\begin{pmatrix} 2.5 \\ -2.0 \end{pmatrix} + \begin{pmatrix} -1.0 & 1.5 & -0.9 \\ -0.6 & 1.7 & 2.0 \\ -0.2 & 1.5 & 0.3 \end{pmatrix}\begin{pmatrix} 0.6291 \\ 0.9460 \\ 0.9998 \end{pmatrix} + \begin{pmatrix} 0.2 \\ 0.2 \\ 0.2 \end{pmatrix}\right)$$

$$= \tanh\left(\begin{pmatrix} (-1.0)(2.5) + (\ 0.2)(-2) \\ (\ 1.5)(2.5) + (\ 1.0)(-2) \\ (\ 2.0)(2.5) + (-1.0)(-2) \end{pmatrix} + \begin{pmatrix} (-1.0)(0.6291) + (1.5)(0.946) + (-0.9)(0.9998) \\ (-0.6)(0.6291) + (1.7)(0.946) + (\ 2.0)(0.9998) \\ (-0.2)(0.6291) + (1.5)(0.946) + (\ 0.3)(0.9998) \end{pmatrix} + \begin{pmatrix} 0.2 \\ 0.2 \\ 0.2 \end{pmatrix}\right)$$

$$= \tanh\left(\begin{pmatrix} -2.90 \\ 1.75 \\ 7.00 \end{pmatrix} + \begin{pmatrix} -0.1099 \\ 3.2303 \\ 1.5931 \end{pmatrix} + \begin{pmatrix} 0.2 \\ 0.2 \\ 0.2 \end{pmatrix}\right) = \tanh\begin{pmatrix} -2.8099 \\ 5.1803 \\ 8.7931 \end{pmatrix} = \begin{pmatrix} -0.9927 \\ 0.9999 \\ 1.0000 \end{pmatrix}$$

$$\boldsymbol{y}(t) = \sigma(\boldsymbol{V}^T\boldsymbol{h}(t) + c)$$

$$\boldsymbol{y}(3) = sigmoid(\boldsymbol{V}^T\boldsymbol{h}(3) + c)$$

$$= \text{sigmoid}\left((0.3 \quad 1.0 \quad 0.4)\begin{pmatrix} -0.9927 \\ 0.9999 \\ 1.0000 \end{pmatrix} + 0.2\right)$$

$$= \text{sigmoid}((0.3)(-0.9927) + (1)(0.9999) + (0.4)(1) + 0.2)$$

$$= \text{sigmoid}(1.3021) = 0.7861$$

<u>Summary</u>

- Output of the network:
  - y(1) = 0.7443
  - y(2) = 0.8500
  - y(3) = 0.7861

4. (b) The word "adversarial" in the acronym for GAN suggests a two-player game.

(i) What are the two players, and what are their respective goals?

(5 marks)

Answer

- The two players are:
  1. Discriminator
  2. Generator

- Their goals:
  1. Discriminator
     - Goal is to distinguish between real data (from the actual dataset) and fake data (produced by the generator).
     - Objective is to correctly identify and classify the real data and the generated data.
  2. Generator
     - Goal is to generate fake data (e.g., images, text) that is indistinguishable from real data.
     - Objective is to fool the Discriminator into classifying its fake outputs as real.

(ii) Describe the phenomenon of mode collapse in GAN and suggest a method to overcome this issue.

(5 marks)

Answer

- Phenomenon of mode collapse in GAN:
  - Generated samples lack variation, even when input noise is different.
  - The generator becomes overly specialized to a few patterns, reducing its ability to generalize to unseen data or edge cases.
  - Mode collapse often correlates with oscillations or divergence in GAN training, where the generator and discriminator fail to reach equilibrium.

- Method to overcome:
  - Mini-batch discrimination allows the discriminator to look at multiple samples in a batch simultaneously instead of evaluating each sample independently.
  - This helps the discriminator detect lack of diversity in the generator's outputs, which is a key symptom of mode collapse.