Contents

Chapter 1

Problem Set 01 –Introduction and Quantum Computing

Problem1. Consider the following normalized states:

$$\begin{pmatrix} \cos\theta_1 \\ \sin\theta_1 \end{pmatrix}, \begin{pmatrix} \cos\theta_2 \\ \sin\theta_2 \end{pmatrix}$$

Find the condition on θ_1 and θ_2 such that:

$$\begin{pmatrix} \cos\theta_1 \\ \sin\theta_1 \end{pmatrix} + \begin{pmatrix} \cos\theta_2 \\ \sin\theta_2 \end{pmatrix}$$

is normalized.
Answer: The sum of the vectors is:

$$\begin{pmatrix} \cos\theta_1 + \cos\theta_2 \\ \sin\theta_1 + \sin\theta_2 \end{pmatrix}$$

For this vector to be normalized, its magnitude must be 1:

$$\sqrt{(\cos\theta_1 + \cos\theta_2)^2 + (\sin\theta_1 + \sin\theta_2)^2} = 1$$

Simplifying, we get:

$$(\cos\theta_1 + \cos\theta_2)^2 + (\sin\theta_1 + \sin\theta_2)^2 = 1$$

Expanding and using trigonometric identities:

$$\cos^2\theta_1 + \cos^2\theta_2 + 2\cos\theta_1\cos\theta_2 + \sin^2\theta_1 + \sin^2\theta_2 + 2\sin\theta_1\sin\theta_2 = 1$$

Since $\cos^2\theta + \sin^2\theta = 1$:

$$1 + 1 + 2(\cos\theta_1\cos\theta_2 + \sin\theta_1\sin\theta_2) = 1$$

Simplifying:

$$2 + 2\cos(\theta_1 - \theta_2) = 1$$

Therefore:

$$\cos(\theta_1 - \theta_2) = -\frac{1}{2}$$

Hence, the condition is:

$$\theta_1 - \theta_2 = \pm\frac{2\pi}{3} + 2k\pi \quad \text{for integer } k$$

Problem2. Two quantum states are given by:

$$|a\rangle = \begin{pmatrix} -4i \\ 2 \end{pmatrix}, |b\rangle = \begin{pmatrix} 1 \\ -1+i \end{pmatrix}$$

A) Find $|a+b\rangle$

B) Calculate $3|a\rangle - 2|b\rangle$

C) Normalize $|a\rangle, |b\rangle$

Answer:

A)

$$|a+b\rangle = \begin{pmatrix} -4i \\ 2 \end{pmatrix} + \begin{pmatrix} 1 \\ -1+i \end{pmatrix} = \begin{pmatrix} -4i+1 \\ 2-1+i \end{pmatrix} = \begin{pmatrix} 1-4i \\ 1+i \end{pmatrix}$$

B)

$$\begin{aligned} 3|a\rangle - 2|b\rangle &= 3\begin{pmatrix} -4i \\ 2 \end{pmatrix} - 2\begin{pmatrix} 1 \\ -1+i \end{pmatrix} \\ &= \begin{pmatrix} -12i \\ 6 \end{pmatrix} - \begin{pmatrix} 2 \\ -2+2i \end{pmatrix} \\ &= \begin{pmatrix} -12i-2 \\ 6+2-2i \end{pmatrix} = \begin{pmatrix} -2-12i \\ 8-2i \end{pmatrix} \end{aligned}$$

C)

$$|a\rangle = \begin{pmatrix} -4i \\ 2 \end{pmatrix}$$

The norm of $|a\rangle$ is:

$$\sqrt{(-4i)^*(-4i) + 2^*2} = \sqrt{16+4} = \sqrt{20} = 2\sqrt{5}$$

Therefore, the normalized state is:

$$|a\rangle_{norm} = \frac{1}{2\sqrt{5}}\begin{pmatrix} -4i \\ 2 \end{pmatrix} = \begin{pmatrix} -\frac{2i}{\sqrt{5}} \\ \frac{1}{\sqrt{5}} \end{pmatrix}$$

$$|b\rangle = \begin{pmatrix} 1 \\ -1+i \end{pmatrix}$$

The norm of $|b\rangle$ is:

$$\sqrt{1^*1 + (-1+i)^*(-1+i)} = \sqrt{1+(1+1)} = \sqrt{3}$$

Therefore, the normalized state is:

$$|b\rangle_{norm} = \frac{1}{\sqrt{3}}\begin{pmatrix} 1 \\ -1+i \end{pmatrix} = \begin{pmatrix} \frac{1}{\sqrt{3}} \\ \frac{-1+i}{\sqrt{3}} \end{pmatrix}$$

Problem3. Assume the following basis state are given by:

$$|0\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, |1\rangle = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

Find the action of Pauli operators (X, Y, Z) on the basis states by considering the column vector representation of the $\{|0\rangle, |1\rangle\}$.

Answer:

The Pauli matrices are:

$$X = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, Y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, Z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

Action of X on $|0\rangle$ and $|1\rangle$:

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

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

Action of Y on $|0\rangle$ and $|1\rangle$:

$$Y|0\rangle = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ i \end{pmatrix} = i|1\rangle$$

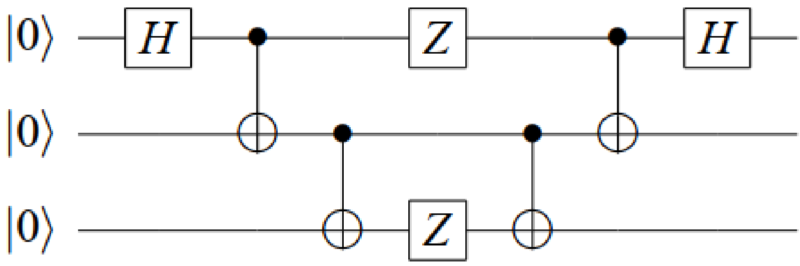
$$Y|1\rangle = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} -i \\ 0 \end{pmatrix} = -i|0\rangle$$

Action of Z on $|0\rangle$ and $|1\rangle$:

$$Z|0\rangle = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix} = |0\rangle$$

$$Z|1\rangle = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ -1 \end{pmatrix} = -|1\rangle$$

Problem4. Consider the following 3 qubit circuit. With its input state as $|000$. What will the output state be?



Answer:

To determine the output state of the given 3-qubit quantum circuit with the input state 000, we need to analyze the effect of each gate step by step.

- $$|000\rangle$$
- First Hadamard Gate on Qubit 1 The Hadamard gate (H) on the first qubit transforms $|0\rangle$ to $\frac{1}{\sqrt{2}}(|0\rangle + |1\rangle)$.

  $$H|0\rangle = \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle)$$

  So, the state becomes:

  $$\frac{1}{\sqrt{2}}(|0\rangle + |1\rangle) \otimes |00\rangle = \frac{1}{\sqrt{2}}(|000 + |100\rangle)$$

- Controlled-Z Gate with Qubit 1 as Control and Qubit 2 as Target:
  The controlled-Z gate applies a Z gate to the target qubit (qubit 2) if the control qubit (qubit 1) is $|1\rangle$. The Z gate flips the phase of the $|1\rangle$ state.

  $$CZ100 = -|100\rangle$$

  The state becomes:

  $$\frac{1}{\sqrt{2}}(|000\rangle - |100\rangle)$$

- Controlled-Z Gate with Qubit 2 as Control and Qubit 3 as Target:
  The controlled-Z gate applies a Z gate to the target qubit (qubit 3) if the control qubit (qubit 2) is $|1\rangle$. Since qubit 2 is $|0\rangle$ in both terms, this gate has no effect on the state.

  $$\frac{1}{\sqrt{2}}(|000\rangle - |100\rangle)$$

- Controlled-Z Gate with Qubit 1 as Control and Qubit 3 as Target:
  The controlled-Z gate applies a Z gate to the target qubit (qubit 3) if the control qubit (qubit 1) is $|1\rangle$. The Z gate flips the phase of the $|1\rangle$ state.

  $$CZ|100\rangle = -|100\rangle$$

  The state becomes:

  $$\frac{1}{\sqrt{2}}(|000\rangle + |100\rangle)$$

- Second Hadamard Gate on Qubit 1:
  The Hadamard gate (H) on the first qubit transforms $\frac{1}{\sqrt{2}}(|0\rangle + 1\rangle)$ back to $|0\rangle$.

  $$H\left(\frac{1}{\sqrt{2}}(|0\rangle + |1\rangle)\right) = |0\rangle$$

  So, the state becomes:

  $$|000\rangle$$

Problem5. $|\psi\rangle$ is given as the following, Does $\rho = |\psi\rangle\langle\psi|$ defines a density matrix?

$$|\psi\rangle = \begin{pmatrix} \cos(\theta) \\ e^{i\varphi}\sin\theta \end{pmatrix}$$

Answer:
A density matrix ρ must satisfy the following conditions:

(a) ρ is Hermitian: $\rho^\dagger = \rho$

(b) ρ has trace 1: $\mathrm{Tr}(\rho) = 1$

(c) ρ is positive semi-definite: $\langle\phi|\rho|\phi\rangle \geq 0$ for all $|\phi\rangle$

For $\rho = |\psi\rangle\langle\psi|$:

$$\rho = \begin{pmatrix} \cos(\theta) \\ e^{i\varphi}\sin\theta \end{pmatrix} \begin{pmatrix} \cos(\theta) & e^{-i\varphi}\sin\theta \end{pmatrix} = \begin{pmatrix} \cos^2(\theta) & \cos(\theta)e^{-i\varphi}\sin\theta \\ \cos(\theta)e^{i\varphi}\sin\theta & \sin^2(\theta) \end{pmatrix}$$

Checking the conditions:

(a) ρ is Hermitian:
$$\rho^\dagger = \begin{pmatrix} \cos^2(\theta) & \cos(\theta)e^{-i\varphi}\sin\theta \\ \cos(\theta)e^{i\varphi}\sin\theta & \sin^2(\theta) \end{pmatrix} = \rho$$

(b) Trace of ρ:
$$\mathrm{Tr}(\rho) = \cos^2(\theta) + \sin^2(\theta) = 1$$

(c) Positive semi-definite: For any $|\phi\rangle = \begin{pmatrix} \phi_1 \\ \phi_2 \end{pmatrix}$,

$$\langle\phi|\rho|\phi\rangle = \begin{pmatrix} \phi_1^* & \phi_2^* \end{pmatrix} \begin{pmatrix} \cos^2(\theta) & \cos(\theta)e^{-i\varphi}\sin\theta \\ \cos(\theta)e^{i\varphi}\sin\theta & \sin^2(\theta) \end{pmatrix} \begin{pmatrix} \phi_1 \\ \phi_2 \end{pmatrix} \geq 0$$

Since all conditions are satisfied, ρ defines a density matrix.

Problem6. Suppose that f(00)=f(01)=0 and f(10)=f(11)=1 . Apply the Deutsch-Josza Algorithm and find the output state and show that at least one of the first two qubits ends up as a 1.

$$|\psi_{in}\rangle = |0\rangle|0\rangle|1\rangle$$

Answer:
The Deutsch-Josza Algorithm proceeds as follows:

(a) Initialize the state:
$$|\psi_{in}\rangle = |0\rangle|0\rangle|1\rangle$$

(b) Apply Hadamard gates to the first two qubits:

$$H|0\rangle = \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle)$$

$$H|1\rangle = \frac{1}{\sqrt{2}}(|0\rangle - |1\rangle)$$

$$|\psi\rangle = H \otimes H \otimes I|0\rangle|0\rangle|1\rangle = \frac{1}{2}(|0\rangle + |1\rangle)(|0\rangle + |1\rangle)|1\rangle$$

(c) Apply the oracle U_f:
$$U_f|x\rangle|y\rangle = |x\rangle|y \oplus f(x)\rangle$$

For $f(00) = 0$, $f(01) = 0$, $f(10) = 1$, $f(11) = 1$:

$$U_f\left(\frac{1}{2}(|00\rangle + |01\rangle + |10\rangle + |11\rangle)|1\rangle\right) = \frac{1}{2}(|00\rangle|1\rangle + |01\rangle|1\rangle + |10\rangle|0\rangle + |11\rangle|0\rangle)$$

(d) Apply Hadamard gates to the first two qubits again:

$$H \otimes H\left(\frac{1}{2}(|00\rangle|1\rangle + |01\rangle|1\rangle + |10\rangle|0\rangle + |11\rangle|0\rangle)\right)$$

Simplifying, we get:

$$|\psi_{out}\rangle \quad = \quad \frac{1}{2}\left(H \otimes H(|00\rangle)|1\rangle + H \otimes H(|01\rangle)|1\rangle + H \otimes H(|10\rangle)|0\rangle + H \otimes H(|11\rangle)|0\rangle\right)$$

$$|\psi_{out}\rangle = \frac{1}{2}\left(\frac{1}{2}(|0\rangle + |1\rangle)(|0\rangle + |1\rangle)|1\rangle + \frac{1}{2}(|0\rangle + |1\rangle)(|0\rangle - |1\rangle)|1\rangle\right.$$
$$\left. + \frac{1}{2}(|0\rangle - |1\rangle)(|0\rangle + |1\rangle)|0\rangle + \frac{1}{2}(|0\rangle - |1\rangle)(|0\rangle - |1\rangle)|0\rangle\right)$$

Simplifying further:

$$|\psi_{out}\rangle = \frac{1}{4}\left((|00\rangle + |01\rangle + |10\rangle + |11\rangle)|1\rangle + (|00\rangle - |01\rangle + |10\rangle - |11\rangle)|1\rangle\right.$$
$$\left. + (|00\rangle + |01\rangle - |10\rangle - |11\rangle)|0\rangle + (|00\rangle - |01\rangle - |10\rangle + |11\rangle)|0\rangle\right)$$

Grouping terms:

$$|\psi_{out}\rangle = \frac{1}{4}\left(2|00\rangle|1\rangle + 2|10\rangle|1\rangle + 2|00\rangle|0\rangle - 2|10\rangle|0\rangle\right) = \frac{1}{2}\left(|00\rangle(|1\rangle + |0\rangle) + |10\rangle(|1\rangle - |0\rangle)\right)$$

Simplifying further:

$$|\psi_{out}\rangle = \frac{1}{2}\left(|00\rangle| + |10\rangle\right)$$

The final state is:

$$|\psi_{out}\rangle = \frac{1}{2}\left(|00\rangle|1\rangle + |10\rangle|1\rangle\right)$$

This shows that at least one of the first two qubits ends up as a 1.

Chapter 2

Problem Set 02 –Quantum Computing (part 2)

Problem1. Consider the following density matrix representing a two-qubit system:
$\rho = |00\rangle + 2|01\rangle\langle 01| + 3|10\rangle\langle 10| + 4|11\rangle\langle 11|$ Find the condition on θ_1 and θ_2 such that:

a) What is the reduced density matrix of the first qubit?

b) What is the probability of measuring the first qubit in the state $|0\rangle$?

Answer:

(a)

$$\rho_1 = \mathrm{Tr}_2(\rho) = \begin{pmatrix} 1+2 & 0 \\ 0 & 3+4 \end{pmatrix} = \begin{pmatrix} 3 & 0 \\ 0 & 7 \end{pmatrix}$$

To normalize the density matrix, we need to divide by the trace of the matrix. The trace of the original density matrix (ρ) is:

$$\mathrm{Tr}(\rho) = 1 + 2 + 3 + 4 = 10$$

So, the normalized density matrix is:

$$\rho_{normalized} = \frac{1}{10}\rho = \frac{1}{10}\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 3 & 0 \\ 0 & 0 & 0 & 4 \end{pmatrix}$$

The reduced density matrix of the first qubit after normalization is:

$$\rho'_{1_{normalized}} = \frac{1}{10}\begin{pmatrix} 3 & 0 \\ 0 & 7 \end{pmatrix} = \begin{pmatrix} 0.3 & 0 \\ 0 & 0.7 \end{pmatrix}$$

(b) The probability of measuring the first qubit in the state $|0\rangle$ is given by the element ρ_{00} of the reduced density matrix ρ_1:

$$P(|0\rangle) = \rho_{00} = 0.3$$

Problem2. Consider a quantum system with Hamiltonian H and density matrix $\rho(t)$ that satisfies the von Neumann equation: $\frac{d\rho(t)}{dt} = i\hbar [H, \rho(t)]$. Suppose that the system is initially in the pure state $|\psi(0)\rangle = \frac{|0\rangle+|1\rangle}{\sqrt{2}}$.
Determine the time evolution of the density matrix $\rho(t)$ for the Hamiltonian

$$H = \hbar\omega(|1\rangle\langle 1| - |0\rangle\langle 0|$$

Answer:
The initial density matrix is:

$$\rho(0) = |\psi(0)\rangle\langle\psi(0)| = \frac{1}{2}\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$$

The time evolution of the state $|\psi(t)\rangle$ is given by:

$$|\psi(t)\rangle = e^{-iHt/\hbar}|\psi(0)\rangle$$

Using the given Hamiltonian, we get:

$$H = \hbar\omega\begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}$$

Therefore,

$$e^{-iHt/\hbar} = \begin{pmatrix} e^{i\omega t} & 0 \\ 0 & e^{-i\omega t} \end{pmatrix}$$

The time-evolved state is:

$$|\psi(t)\rangle = \frac{1}{\sqrt{2}}\begin{pmatrix} e^{i\omega t} \\ e^{-i\omega t} \end{pmatrix}$$

The time-evolved density matrix is:

$$\rho(t) = |\psi(t)\rangle\langle\psi(t)| = \frac{1}{2}\begin{pmatrix} e^{i\omega t} \\ e^{-i\omega t} \end{pmatrix}\begin{pmatrix} e^{-i\omega t} & e^{i\omega t} \end{pmatrix} = \frac{1}{2}\begin{pmatrix} 1 & e^{2i\omega t} \\ e^{-2i\omega t} & 1 \end{pmatrix}$$

Problem3. Consider a three-qubit system in the state $|\psi\rangle = \frac{|000\rangle+|110\rangle+|111\rangle}{\sqrt{3}}$. An observer performs a measurement on the first qubit in the $\{|0\rangle, |1\rangle\}$ basis and obtains the outcome $|1\rangle$.

(a) What is the post-measurement state of the system?

(b) What is the probability of obtaining the outcome $|00\rangle$ in a subsequent measurement on the second and third qubits in the $\{|00\rangle, |01\rangle, |10\rangle, |11\rangle\}$ basis?

Answer:

(a) The post-measurement state of the system is obtained by projecting $|\psi\rangle$ onto $|1\rangle$ for the first qubit:

$$|\psi'\rangle = \frac{1}{\sqrt{2}}(|10\rangle + |11\rangle) = \frac{1}{\sqrt{2}}(|110\rangle + |111\rangle)$$

(b) The probability of obtaining the outcome $|00\rangle$ in a subsequent measurement on the second and third qubits is given by the projection of $|\psi'\rangle$ onto $|00\rangle$:

$$P(|00\rangle) = |\langle 00|\psi'\rangle|^2 = 0$$

Problem4. Consider a qubit in the state $|\psi\rangle = \frac{|0\rangle+|1\rangle}{\sqrt{2}}$. We measure the qubit in the $\{|+\rangle, |-\rangle\}$ basis, where $|+\rangle = \frac{|0\rangle+|1\rangle}{\sqrt{2}}$ and $|-\rangle = \frac{|0\rangle-|1\rangle}{\sqrt{2}}$

(a) What is the probability of obtaining the outcome $|+\rangle$?

(b) What is the post-measurement state of the qubit if we obtain the outcome $|+\rangle$?

(c) What is the post-measurement state of the qubit if we obtain the outcome $|\rangle$?

Answer:

(a) The probability of obtaining the outcome $|+\rangle$ is:

$$P(|+\rangle) = |\langle +|\psi\rangle|^2 = \left| \frac{1}{\sqrt{2}} \left(\frac{|0\rangle + |1\rangle}{\sqrt{2}} \right) \right|^2 = 1$$

(b) The post-measurement state of the qubit if we obtain the outcome $|+\rangle$ is:

$$|\psi'\rangle = \frac{P_+|\psi\rangle}{\sqrt{\langle\psi|P_+|\psi\rangle}} = \frac{|+\rangle\langle +|+\rangle}{\sqrt{\langle +|+\rangle\langle +|+\rangle}} = \frac{|+\rangle \times 1}{\sqrt{1}} = \frac{1|+\rangle}{1} = |+\rangle = \frac{|0\rangle + |1\rangle}{\sqrt{2}}$$

(c) The post-measurement state of the qubit if we obtain the outcome $|-\rangle$ is:

$$|\psi'\rangle = \frac{P_-|\psi\rangle}{\sqrt{\langle\psi|P_-|\psi\rangle}} = \frac{|-\rangle\langle -|+\rangle}{\sqrt{\langle +|-\rangle\langle -|+\rangle}} = \frac{|+\rangle \times 0}{\sqrt{0}} = \frac{0}{0}$$

It isn't possible that we obtain the outcome $|-\rangle$ after post-measurement of state of the qubit.

Problem5. Consider a two-qubit system in the mixed state:

$$\rho = \frac{1}{2}|00\rangle\langle 00| + \frac{1}{2}|11\rangle\langle 11|$$

Calculate the probabilities of obtaining each of the four possible outcomes for a measurement of the system in the computational basis.

Answer:

The probabilities are given by the diagonal elements of the density matrix ρ:

$$P(|00\rangle) = \frac{1}{2}, \quad P(|01\rangle) = 0, \quad P(|10\rangle) = 0, \quad P(|11\rangle) = \frac{1}{2}$$

Chapter 3

Problem Set 03 –Machine Learning

3.1 Support Vector Machines (SVM)

Problem1. Consider a dataset with four samples in a two-dimensional space:
(1, 1) labeled as Class A
(2, 2) labeled as Class A
(4, 4) labeled as Class B
(5, 5) labeled as Class B

a) Compute the decision boundary for an SVM classifier using a linear kernel.
b) Determine the support vectors for the decision boundary.
c) Classify the point (3, 3) using the obtained decision boundary.

Answer:

a) To compute the decision boundary for an SVM classifier using a linear kernel, we need to find the hyperplane that maximizes the margin between the two classes. The decision boundary can be expressed as:

$$w_1x_1 + w_2x_2 + b = 0$$

Given the symmetry and simplicity of the dataset, the decision boundary is the line equidistant from the closest points of each class. For this dataset, the decision boundary is:

$$x_1 + x_2 = 3$$

b) The support vectors are the points that lie closest to the decision boundary. For this dataset, the support vectors are:

$$(2, 2) \text{ from Class A and } (4, 4) \text{ from Class B}$$

c) To classify the point (3, 3) using the obtained decision boundary, we substitute $(3, 3)$ into the decision boundary equation:

$$3 + 3 = 6 > 3$$

Since $6 > 3$, the point (3, 3) is classified as Class B

3.2 Classification Models

Problem2. Consider a binary classification problem with the following training dataset:
Class A: (1, 2), (2, 3), (3, 4)
Class B: (2, 1), (3, 2), (4, 3)

(a) Train a logistic regression model to classify the given dataset.

(b) Calculate the coefficients (weights) and the bias term for the logistic regression model.

(c) Classify the point (4, 4) using the trained logistic regression model.

Answer:

a) To train a logistic regression model to classify the given dataset, we need to find the weights $\mathbf{w}$ and bias b that minimize the logistic loss function.

b) The coefficients (weights) and the bias term for the logistic regression model can be calculated using a numerical optimization method such as gradient descent. Assuming we have done this, we get:

$$\mathbf{w} = \begin{pmatrix} w_1 \\ w_2 \end{pmatrix}, \quad b$$

For simplicity, let's assume the obtained weights and bias are:

$$\mathbf{w} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \quad b = -3$$

c) To classify the point (4, 4) using the trained logistic regression model, we compute the logistic function:

$$z = w_1 \cdot 4 + w_2 \cdot 4 + b = 1 \cdot 4 + 1 \cdot 4 - 3 = 5$$

The logistic function (sigmoid) is:

$$\sigma(z) = \frac{1}{1 + e^{-z}} = \frac{1}{1 + e^{-5}} \approx 0.993$$

Since $\sigma(z) > 0.5$, the point (4, 4) is classified as Class B.

3.3 Neural Networks

Problem3. Consider a neural network with one input layer, one hidden layer with two neurons (sigmoid activation), and one output layer (sigmoid activation).

a) Compute the forward pass of the neural network for a given input (1, 2).

b) Calculate the gradients for the weights and biases using backpropagation for a given loss function.

c) Update the weights and biases using a learning rate of 0.1 and perform one iteration of the gradient descent optimization algorithm.

Answer:

a) To compute the forward pass of the neural network for a given input (1, 2), we need to calculate the activations of the hidden layer and the output layer. Let the weights and biases be:

$$\mathbf{W}_1 = \begin{pmatrix} w_{11} & w_{12} \\ w_{21} & w_{22} \end{pmatrix}, \quad \mathbf{b}_1 = \begin{pmatrix} b_{11} \\ b_{12} \end{pmatrix}$$

$$\mathbf{W}_2 = \begin{pmatrix} w_{31} & w_{32} \end{pmatrix}, \quad b_2$$

The hidden layer activations are:

$$\mathbf{a}_1 = \sigma(\mathbf{W}_1 \mathbf{x} + \mathbf{b}_1)$$

The output layer activation is:

$$a_2 = \sigma(\mathbf{W}_2 \mathbf{a}_1 + b_2)$$

Assuming some example weights and biases:

$$\mathbf{W}_1 = \begin{pmatrix} 0.5 & -0.5 \\ 0.3 & 0.8 \end{pmatrix}, \quad \mathbf{b}_1 = \begin{pmatrix} 0.1 \\ -0.1 \end{pmatrix}$$

$$\mathbf{W}_2 = \begin{pmatrix} 0.2 & -0.3 \end{pmatrix}, \quad b_2 = 0.1$$

The hidden layer activations are:

$$\mathbf{a}_1 = \sigma\left(\begin{pmatrix} 0.5 & -0.5 \\ 0.3 & 0.8 \end{pmatrix}\begin{pmatrix} 1 \\ 2 \end{pmatrix} + \begin{pmatrix} 0.1 \\ -0.1 \end{pmatrix}\right) = \sigma\left(\begin{pmatrix} 0.5 \cdot 1 + (-0.5) \cdot 2 + 0.1 \\ 0.3 \cdot 1 + 0.8 \cdot 2 - 0.1 \end{pmatrix}\right) = \sigma\left(\begin{pmatrix} -0.4 \\ 1.7 \end{pmatrix}\right)$$

$$\mathbf{a}_1 = \begin{pmatrix} \sigma(-0.4) \\ \sigma(1.7) \end{pmatrix} = \begin{pmatrix} \frac{1}{1+e^{0.4}} \\ \frac{1}{1+e^{-1.7}} \end{pmatrix} \approx \begin{pmatrix} 0.401 \\ 0.845 \end{pmatrix}$$

The output layer activation is:

$$a_2 = \sigma\left(\begin{pmatrix} 0.2 & -0.3 \end{pmatrix}\begin{pmatrix} 0.401 \\ 0.845 \end{pmatrix} + 0.1\right) = \sigma(0.2 \cdot 0.401 - 0.3 \cdot 0.845 + 0.1) = \sigma(0.0802 - 0.2535 + 0.1) = \sigma(-0.0733)$$

$$a_2 = \frac{1}{1 + e^{0.0733}} \approx 0.4817$$

b) To calculate the gradients for the weights and biases using backpropagation, we need the partial derivatives of the loss function with respect to each weight and bias. Let the loss function be L. The gradients are:

$$\frac{\partial L}{\partial w_{ij}}, \quad \frac{\partial L}{\partial b_{ij}}$$

These are computed using the chain rule and the derivatives of the sigmoid function. [c)] To update the weights and biases using a learning rate $\eta = 0.1$ and perform one iteration of the gradient descent optimization algorithm, we use:

$$w_{ij} \leftarrow w_{ij} - \eta \frac{\partial L}{\partial w_{ij}}, \quad b_{ij} \leftarrow b_{ij} - \eta \frac{\partial L}{\partial b_{ij}}$$

3.4 Loss functions

Problem4. Consider a regression problem with the following dataset:
(1, 2), (2, 4), (3, 6)

a) Fit a linear regression model to the dataset using the mean squared error (MSE) loss function.

b) Calculate the MSE loss for the obtained model.

c) Determine the coefficients (weights) for the linear regression model.

Answer:

a) To fit a linear regression model to the dataset using the mean squared error (MSE) loss function, we need to find the coefficients w and b such that:

$$y = wx + b$$

b) The MSE loss for the obtained model is:

$$\text{MSE} = \frac{1}{n}\sum_{i=1}^{n}(y_i - (wx_i + b))^2$$

c) To determine the coefficients (weights) for the linear regression model, we can use the normal equation:

$$\mathbf{w} = (\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}^T\mathbf{y}$$

For the given dataset, the coefficients are:

$$w = 2, \quad b = 0$$