Modern Artificial Intelligence

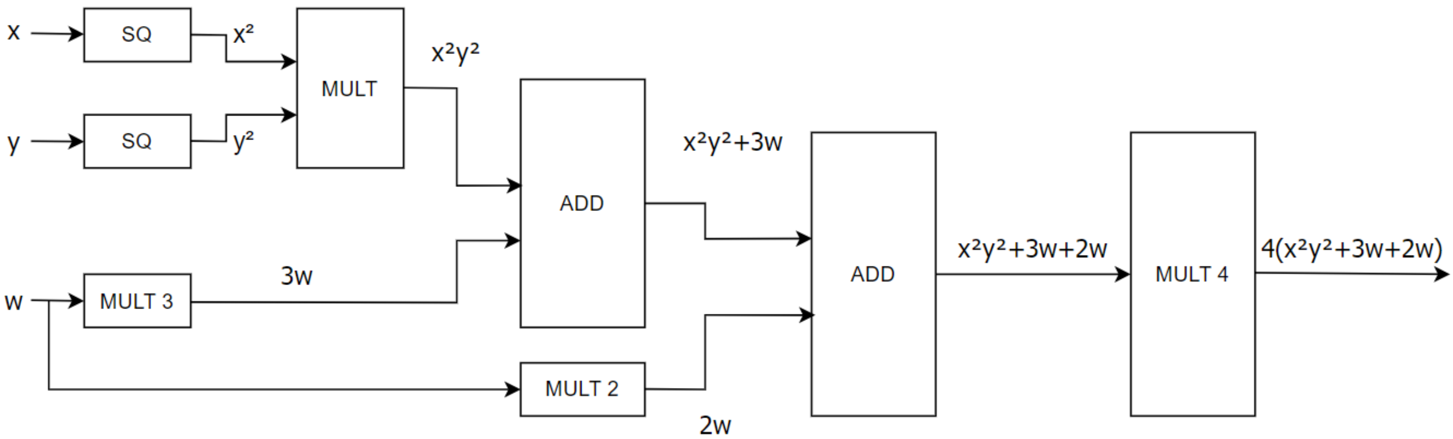
Individual Assignment, CSCI 323

Name: Jeslyn Ho Ka Yan ID: 8535383

Question 1

Loss(*x, y, z, w*) = 4(*x*²*y*² + 3min{*w, z*} + 2*w*).

x = 4, y = -2, z = 5 and w = -1.



Gradient for X

$$\frac{\boldsymbol{dL}}{\boldsymbol{dx}} = \frac{dL}{d(x^2y^2+3w+2w)} \times \frac{d(x^2y^2+3w+2w)}{d(x^2y^2+3w)} \times \frac{d(x^2y^2+3w)}{d(x^2y^2)} \times \frac{d(x^2y^2)}{dx^2} \times \frac{dx^2}{dx}$$

$$\frac{\boldsymbol{dL}}{\boldsymbol{dx}} = 4 \times 1 \times 4 \times 2 = 32$$

Gradient for Y

$$\frac{\boldsymbol{dL}}{\boldsymbol{dy}} = \frac{dL}{d(x^2y^2+3w+2w)} \times \frac{d(x^2y^2+3w+2w)}{d(x^2y^2+3w)} \times \frac{d(x^2y^2+3w)}{d(x^2y^2)} \times \frac{d(x^2y^2)}{dy^2} \times \frac{dy^2}{dy}$$

$$\frac{\boldsymbol{dL}}{\boldsymbol{dy}} = 4 \times 1 \times 1 \times 4 = 16$$

Gradient for w

$$\frac{\boldsymbol{dL}}{\boldsymbol{dw}} = \left\{\frac{dL}{d(x^2y^2+3w+2w)} \times \frac{d(x^2y^2+3w+2w)}{d(x^2y^2+3w)} \times \frac{d(x^2y^2+3w)}{d(3w)} \times \frac{d(3w)}{dw}\right\}$$

$$+\{\frac{dL}{d(x^2y^2+3w+2w)} \times \frac{d(x^2y^2+3w+2w)}{d(2w)} \times \frac{d(2w)}{dw}\}$$

$$\frac{\boldsymbol{dL}}{\boldsymbol{dw}} = \{4 \times 1 \times 1 \times 3\} + \{4 \times 1 \times 2\} = 20$$

Gradient for Z, not required since min(w,z) =w, so z is ignored.

Question 2

To make the datasets linearly separable, we use a quadratic feature transformation: $\boldsymbol{\phi}(\boldsymbol{x}) = [\mathbf{1}, \boldsymbol{x}^2]$

The first component "1" provides a bias term for flexibility. The second component, $\boldsymbol{x}^2$, adds non-linearity, allowing us to separate data that cannot be split by a straight line in the original space. This transformation makes it possible to use a linear classifier to perfectly separate the classes in the new feature space.

Orginal Datasets

- $D_1 = \{(-1, +1), (0, -1), (1, +1)\}$
- $D_2 = \{(-1, -1), (0, +1), (1, -1)\}$

Trasnformed Dataset using $\boldsymbol{\phi}(\boldsymbol{x}) = [\mathbf{1}, \boldsymbol{x}^2]$

- $D_1 = \{(1, 1), (1, 0), (1, 1)\}$
- $D_2 = \{(1, 1), (1, 0), (1, 1)\}$

Feature Vectors

- When $\boldsymbol{x} = -\mathbf{1}$ $\boldsymbol{or}$ $\boldsymbol{x} = \mathbf{1}$: $\boldsymbol{\phi}(\boldsymbol{x}) = [\mathbf{1}, \mathbf{1}]$
- When $\boldsymbol{x} = \mathbf{0}$: $\boldsymbol{\phi}(\boldsymbol{x}) = [\mathbf{1}, \mathbf{0}]$

f(x)= $\boldsymbol{w_0} + \boldsymbol{w_1 x^2}$

Weight Vector For D_1

$\mathbf{f}(\boldsymbol{x}) > \mathbf{0}$ $\boldsymbol{for}$ $[\mathbf{1}, \mathbf{1}]$ $(+\mathbf{1}$ $\boldsymbol{Label})$ $\mathbf{f}(\boldsymbol{x}) = \boldsymbol{w_0} + \boldsymbol{w_1} \times \mathbf{1} = \boldsymbol{w_0} + \boldsymbol{w_1} > \mathbf{0}$

$\mathbf{f}(\boldsymbol{x}) < \mathbf{0}$ $\boldsymbol{for}$ $[\mathbf{1}, \mathbf{0}]$ $(-\mathbf{1}$ $\boldsymbol{Label})$ $\mathbf{f}(\boldsymbol{x}) = \boldsymbol{w_0} + \boldsymbol{w_1} \times \mathbf{0} = \boldsymbol{w_0} < \mathbf{0}$

Choose $\boldsymbol{w_0} = -\mathbf{1}$ $\boldsymbol{and}$ $\boldsymbol{w_1} = \mathbf{2}$

For $[1,1]$: $-1 + 2 = 1 > 0$ Correct for +1 label

For $[1,0]$: $-1 < 0$ Correct for -1 label

For $(-\mathbf{1}, -\mathbf{1})$:

$(-\mathbf{1}) + (\mathbf{2}) \times \mathbf{1} = \mathbf{1} > \mathbf{0}$ (correct for +1 label)

For $(\mathbf{0}, \mathbf{1})$:

$(-\mathbf{1}) + (\mathbf{2}) \times \mathbf{0} = -\mathbf{1} < \mathbf{0}$ (correct for -1 label)

For $(\mathbf{1}, -\mathbf{1})$:

$(-\mathbf{1}) + (\mathbf{2}) \times \mathbf{1} = \mathbf{1} > \mathbf{0}$ (correct for +1 label)

Weight Vector For D_2

$\mathbf{f}(x) < \mathbf{0}\ \boldsymbol{for}\ [\mathbf{1}, \mathbf{1}]\ (-\mathbf{1}\ \boldsymbol{Label})$ $\quad$ $\mathbf{f}(x) = \boldsymbol{w_0} + \boldsymbol{w_1} \times \mathbf{1} = \boldsymbol{w_0} + \boldsymbol{w_1} < \mathbf{0}$

$\mathbf{f}(x) > \mathbf{0}\ \boldsymbol{for}\ [\mathbf{1}, \mathbf{0}]\ (+\mathbf{1}\ \boldsymbol{Label})$ $\quad$ $\mathbf{f}(x) = \boldsymbol{w_0} + \boldsymbol{w_1} \times \mathbf{0} = \boldsymbol{w_0} > \mathbf{0}$

Choose $\boldsymbol{w_0} = \mathbf{1}\ \boldsymbol{and}\ \boldsymbol{w_1} = -\mathbf{2}$

For $[1,1]$: $1 + (-2) = -1 < 0$ Correct for -1 label

For $[1,0]$: $1 > 0$ Correct for +1 label

For $(-\mathbf{1}, -\mathbf{1})$:

$(\mathbf{1}) + (-\mathbf{2}) \times \mathbf{1} = -\mathbf{1} > \mathbf{0}$ (correct for -1 label)

For $(\mathbf{0}, \mathbf{1})$:

$(\mathbf{1}) + (-\mathbf{2}) \times \mathbf{0} = \mathbf{1} < \mathbf{0}$ (correct for +1 label)

For $(\mathbf{1}, -\mathbf{1})$:

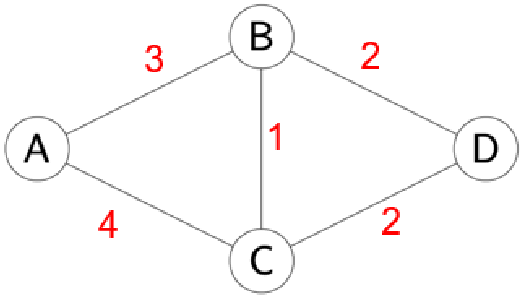
$(\mathbf{1}) + (-\mathbf{2}) \times \mathbf{1} = -\mathbf{1} > \mathbf{0}$ (correct for -1 label)

Therefore, the weight vectors listed below meet the classification requirements for the corresponding dataset:

Weight Vectors for D1: $\boldsymbol{w_0} = -\mathbf{1}\ \boldsymbol{and}\ \boldsymbol{w_1} = \mathbf{2}$

Weight Vectors for D2: $\boldsymbol{w_0} = \mathbf{1}\ \boldsymbol{and}\ \boldsymbol{w_1} = -\mathbf{2}$

Question 3



Start state: A, end state: D

Intialization

- **Start:** A
- **Goal:** D
- **Frontier:** (A, cost=0)
- **Explored:** {}

UCS Table

Step	Node Popped	Frontier (cost)	Explored	Notes
0	-	A (0)	-	Start state
1	A	B (3), C (4)	{A}	Expand A: A→B (3), A→C (4)
2	B	C (4), D (5)	{A, B}	Expand B: B→D (3+2=5), B→C (3+1=4)
3	C	D (5), D (6)	{A, B, C}	Expand C: C→D (4+2=6), but keep lowest
4	D	-	{A,B,C,D}	Goal found!

Step 1: Expand A, add B (3), C (4)
Step 2: Expand B, add D (5)
Step 3: Expand C, D (6) ignored (D at 5 already)
Step 4: Expand D (goal found)

Path Reconstruction

- From A → B → D (because cost to B is 3, then to D is 2, total 5)

Final Answer

- Minimum cost path: A → B → D
- Total minimum cost: 5