①

CLASSMATE Page No.
Date / /

1) a) $A = \begin{bmatrix} 1 & 2 & 1 \\ 1 & 3 & 4 \\ 2 & 3 & -1 \end{bmatrix}$

rank:

$= \begin{bmatrix} 1 & 2 & 1 \\ 0 & 1 & 3 \\ 2 & 3 & -1 \end{bmatrix}$ $r_2 = r_2 - r_1$

$= \begin{bmatrix} 1 & 2 & 1 \\ 0 & 1 & 3 \\ 0 & -1 & -3 \end{bmatrix}$ $r_3 = r_3 - 2r_1$

$= \begin{bmatrix} 1 & 2 & 1 \\ 0 & 1 & 3 \\ 0 & 0 & 0 \end{bmatrix}$ $r_3 = r_2 + r_3$

no. of linearly independent rows = 2

Hence, rank = 2

Inverse:

$|A| = 1(-3-12) - 2(-1-8) + 1(3-6)$

$= -15 + 18 - 3$

$= 0$

$\Rightarrow A^{-1}$ does not exist.

b) $A = \begin{bmatrix} 0 & -2 & 4 \\ 1 & 1 & -1 \\ 2 & 4 & -5 \end{bmatrix}$

Rank

$= \begin{bmatrix} 1 & 1 & -1 \\ 0 & -2 & 4 \\ 2 & 4 & -5 \end{bmatrix}$ as R_1 has first element as 0. swap $R_1 \leftrightarrow R_2$

$= \begin{bmatrix} 1 & 1 & -1 \\ 0 & -2 & 4 \\ 0 & 2 & -3 \end{bmatrix}$ $R_3 = R_3 - 2R_1$

$= \begin{bmatrix} 1 & 1 & -1 \\ 0 & -2 & 4 \\ 0 & 0 & 1 \end{bmatrix}$ $R_3 = R_3 + R_2$

No. of linearly independent rows $= 3$

Hence, Rank $= 3$

Inverse:

$|A| = 0 + 2(-5+2) + 4(4-2)$

$= -6 + 8 = 2 \Rightarrow$ ~~rank~~ Inverse exists

Calculating adj(A):

$A^T = \begin{bmatrix} 0 & 1 & 2 \\ -2 & 1 & 4 \\ 4 & -1 & -5 \end{bmatrix}$

$\begin{vmatrix} 1 & 4 \\ -1 & -5 \end{vmatrix} = -1$ $\begin{vmatrix} -2 & 4 \\ 4 & -5 \end{vmatrix} = -6$ $\begin{vmatrix} -2 & 1 \\ 4 & -1 \end{vmatrix} = -2$

$\begin{vmatrix} 1 & 2 \\ -1 & -5 \end{vmatrix} = -3$ $\begin{vmatrix} 0 & 2 \\ 4 & -5 \end{vmatrix} = -8$ $\begin{vmatrix} 0 & 1 \\ 4 & -1 \end{vmatrix} = -4$

$\begin{vmatrix} 1 & 2 \\ 1 & 4 \end{vmatrix} = 2$ $\begin{vmatrix} 0 & 2 \\ -2 & 4 \end{vmatrix} = 4$ $\begin{vmatrix} 0 & 1 \\ -2 & 1 \end{vmatrix} = 2$

$$\text{Adj}(A) = \begin{bmatrix} -1 & -(-6) & -2 \\ -(-3) & -8 & -(-4) \\ 2 & -(4) & 2 \end{bmatrix}$$

$$A^{-1} = \frac{1}{|A|}\text{Adj}(A) = \begin{bmatrix} -0.5 & 3 & -1 \\ 1.5 & -4 & 2 \\ 1 & -2 & 1 \end{bmatrix}$$

c) $A = \begin{bmatrix} 1 & 2 & 1 & 0 \\ 2 & 5 & 5 & 1 \\ -2 & -3 & 0 & 3 \\ 3 & 4 & -2 & -3 \end{bmatrix}$

<u>rank</u>

$r_2 = r_2 - 2r_1$

$$= \begin{bmatrix} 1 & 2 & 1 & 0 \\ 0 & 1 & 3 & 1 \\ -2 & -3 & 0 & 3 \\ 3 & 4 & -2 & -3 \end{bmatrix}$$

CLASSTIME Page No.
Date / /

$r_3 = r_3 + 2r_1$

$$= \begin{bmatrix} 1 & 2 & 1 & 0 \\ 0 & 1 & 3 & 1 \\ 0 & 1 & 2 & 3 \\ 3 & 4 & -2 & -3 \end{bmatrix}$$

$r_4 = r_4 - 3r_1$

$$= \begin{bmatrix} 1 & 2 & 1 & 0 \\ 0 & 1 & 3 & 1 \\ 0 & 1 & 2 & 3 \\ 0 & -2 & -5 & -3 \end{bmatrix}$$

$r_3 = r_3 - r_2$

$$= \begin{bmatrix} 1 & 2 & 1 & 0 \\ 0 & 1 & 3 & 1 \\ 0 & 0 & -1 & 2 \\ 0 & -2 & -5 & -3 \end{bmatrix}$$

$r_4 = r_4 + 2r_2$

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

$r_4 = r_4 + r_3$

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

no. of linearly independent rows = 4

Hence, rank = 4.

Inverse

$|A| =$ mul(diagonal elements)

$= 1 \times 1 \times -1 \times 1$

$= -1$

$\Rightarrow$ inverse exists.

I will be solving using gaussian elimination method as this is a way ~~method~~. matrix.

$$= \left[\begin{array}{cccc|cccc} 1 & 2 & 1 & 0 & 1 & 0 & 0 & 0 \\ 2 & 5 & 5 & 1 & 0 & 1 & 0 & 0 \\ -2 & -3 & 0 & 3 & 0 & 0 & 1 & 0 \\ 3 & 4 & -2 & -3 & 0 & 0 & 0 & 1 \end{array}\right]$$

$$R_2 = R_2 - 2R_1$$

$$= \left[\begin{array}{cccc|cccc} 1 & 2 & 1 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 3 & 1 & -2 & 1 & 0 & 0 \\ -2 & -3 & 0 & 3 & 0 & 0 & 1 & 0 \\ 3 & 4 & -2 & -3 & 0 & 0 & 0 & 1 \end{array}\right]$$

$$R_3 = R_3 + 2R_1$$

$$= \left[\begin{array}{cccc|cccc} 1 & 2 & 1 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 3 & 1 & -2 & 1 & 0 & 0 \\ 0 & 1 & 2 & 3 & 2 & 0 & 1 & 0 \\ 3 & 4 & -2 & -3 & 0 & 0 & 0 & 1 \end{array}\right]$$

$$R_4 = R_4 - 3R_1$$

$$= \left[\begin{array}{cccc|cccc} 1 & 2 & 1 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 3 & 1 & -2 & 1 & 0 & 0 \\ 0 & 1 & 2 & 3 & 2 & 0 & 1 & 0 \\ 0 & -2 & -5 & -3 & -3 & 0 & 0 & 1 \end{array}\right]$$

$R_1 = R_1 - 2R_1$

$$= \left[\begin{array}{cccc|cccc} 1 & 0 & -5 & -2 & 5 & -2 & 0 & 0 \\ 0 & 1 & 3 & 1 & -2 & 1 & 0 & 0 \\ 0 & 1 & 2 & 3 & 2 & 0 & 1 & 0 \\ 0 & -2 & -5 & -3 & -3 & 0 & 0 & 1 \end{array}\right]$$

$R_3 = R_3 - R_2$

$$= \left[\begin{array}{cccc|cccc} 1 & 0 & -5 & -2 & 5 & -2 & 0 & 0 \\ 0 & 1 & 3 & 1 & -2 & 1 & 0 & 0 \\ 0 & 0 & -1 & 2 & 4 & -1 & 1 & 0 \\ 0 & -2 & -5 & -3 & -3 & 0 & 0 & 1 \end{array}\right]$$

$R_4 = R_4 + 2R_2$

$$= \left[\begin{array}{cccc|cccc} 1 & 0 & -5 & -2 & 5 & -2 & 0 & 0 \\ 0 & 1 & 3 & 1 & -2 & 1 & 0 & 0 \\ 0 & 0 & -1 & 2 & 4 & -1 & 1 & 0 \\ 0 & 0 & 1 & -1 & -7 & 2 & 0 & 1 \end{array}\right]$$

$R_1 = R_1 - 5R_3$

$$= \left[\begin{array}{cccc|cccc} 1 & 0 & 0 & -12 & -15 & 3 & -5 & 0 \\ 0 & 1 & 3 & 1 & -2 & 1 & 0 & 0 \\ 0 & 0 & -1 & 2 & 4 & -1 & 1 & 0 \\ 0 & 0 & 1 & -1 & -7 & 2 & 0 & 1 \end{array}\right]$$

$R_2 = R_2 + 3R_3$

$$= \left[\begin{array}{cccc|cccc} 1 & 0 & 0 & -12 & -15 & 3 & -5 & 0 \\ 0 & 1 & 0 & 7 & 10 & -2 & 3 & 0 \\ 0 & 0 & -1 & 2 & 4 & -1 & 1 & 0 \\ 0 & 0 & 1 & -1 & -7 & 2 & 0 & 1 \end{array}\right]$$

$R_4 = R_4 + R_3$

$$= \left[\begin{array}{cccc|cccc} 1 & 0 & 0 & -12 & -15 & 3 & -5 & 0 \\ 0 & 1 & 0 & 7 & 10 & -2 & 3 & 0 \\ 0 & 0 & -1 & 2 & 4 & -1 & 1 & 0 \\ 0 & 0 & 0 & 1 & -3 & 1 & 1 & 1 \end{array}\right]$$

$R_3 = -R_3$

$$= \left[\begin{array}{cccc|cccc} 1 & 0 & 0 & -12 & -15 & 3 & -5 & 0 \\ 0 & 1 & 0 & 7 & 10 & -2 & 3 & 0 \\ 0 & 0 & 1 & -2 & -4 & 1 & -1 & 0 \\ 0 & 0 & 0 & 1 & -3 & 1 & 1 & 1 \end{array}\right]$$

CLASSTIME Page No.
Date / /

$R_1 = R_1 + 12R_4$

$$= \left[\begin{array}{cccc|cccc} 1 & 0 & 0 & 0 & -51 & 15 & 7 & 12 \\ 0 & 1 & 0 & 7 & 10 & -2 & 3 & 0 \\ 0 & 0 & 1 & -2 & -4 & 1 & -1 & 0 \\ 0 & 0 & 0 & 1 & -3 & 1 & 1 & 1 \end{array}\right]$$

$R_2 = R_2 - 7R_4$

$$= \left[\begin{array}{cccc|cccc} 1 & 0 & 0 & 0 & -51 & 15 & 7 & 12 \\ 0 & 1 & 0 & 0 & 31 & -9 & -4 & -7 \\ 0 & 0 & 1 & -2 & -4 & 1 & -1 & 0 \\ 0 & 0 & 0 & 1 & -3 & 1 & 1 & 1 \end{array}\right]$$

$R_3 = R_3 + 2R_4$

$$= \left[\begin{array}{cccc|cccc} 1 & 0 & 0 & 0 & -51 & 15 & 7 & 12 \\ 0 & 1 & 0 & 0 & 31 & -9 & -4 & -7 \\ 0 & 0 & 1 & 0 & -10 & 3 & 1 & 2 \\ 0 & 0 & 0 & 1 & -3 & 1 & 1 & 1 \end{array}\right]$$

$$A^{-1} = \begin{bmatrix} -51 & 15 & 7 & 12 \\ 31 & -9 & -4 & -7 \\ -10 & 3 & 1 & 2 \\ -3 & 1 & 1 & 1 \end{bmatrix}$$

CLASSTIME | Page No.
Date / /

$$A^T = \begin{bmatrix} 1 & 2 & -2 & 3 \\ 2 & 5 & -3 & 4 \\ 1 & 5 & 0 & -2 \\ 0 & 1 & 3 & -3 \end{bmatrix}$$

$$\begin{vmatrix} 5 & -3 & 4 \\ 5 & 0 & -2 \\ 1 & 3 & -3 \end{vmatrix} \quad \begin{vmatrix} 2 & -3 & 4 \\ 1 & 0 & -2 \\ 0 & 3 & -3 \end{vmatrix} \quad \begin{vmatrix} 2 & 5 & 4 \\ 1 & 5 & -2 \\ 0 & 1 & -3 \end{vmatrix} \quad \begin{vmatrix} 2 & 5 & -3 \\ 1 & 5 & 0 \\ 0 & 1 & 3 \end{vmatrix}$$

$= 30-39+60$ $\quad = 12-9+12$ $\quad = -26+15+4$ $\quad = 30-15-3$

$= 51$ $\quad = 15$ $\quad = -7$ $\quad = 12$

$$\begin{vmatrix} 2 & -2 & 3 \\ 5 & 0 & -2 \\ 1 & 3 & -3 \end{vmatrix} \quad \begin{vmatrix} 1 & -2 & 3 \\ 1 & 0 & -2 \\ 0 & 3 & -3 \end{vmatrix} \quad \begin{vmatrix} 1 & 2 & 3 \\ 1 & 5 & -2 \\ 0 & 1 & -3 \end{vmatrix} \quad \begin{vmatrix} 1 & 2 & -2 \\ 1 & 5 & 0 \\ 0 & 1 & 3 \end{vmatrix}$$

$= 12+2(-13)+45$ $\quad = 6-6+9$ $\quad = -13+6+13$ $\quad = 15-6-2$

$= 31$ $\quad = 9$ $\quad = 4$ $\quad = 7$

$$\begin{vmatrix} 2 & -2 & 3 \\ 5 & -3 & 4 \\ 1 & 3 & -3 \end{vmatrix} \quad \begin{vmatrix} 1 & -2 & 3 \\ 2 & -3 & 4 \\ 0 & 3 & -3 \end{vmatrix} \quad \begin{vmatrix} 1 & 2 & 3 \\ 2 & 5 & 4 \\ 0 & 1 & -3 \end{vmatrix} \quad \begin{vmatrix} 1 & 2 & -2 \\ 2 & 5 & -3 \\ 0 & 1 & 3 \end{vmatrix}$$

$= -6-38+54$ $\quad = -3-12+18$ $\quad = -19+12+6$ $\quad = 18-16$

$= 10$ $\quad = 3$ $\quad = -1$ $\quad = 2$

2)

$$\begin{vmatrix} 2 & -2 & 3 \\ 5 & -3 & 4 \\ 5 & 0 & -2 \end{vmatrix} \quad \begin{vmatrix} 1 & -2 & 3 \\ 2 & -3 & 4 \\ 1 & 0 & -2 \end{vmatrix} \quad \begin{vmatrix} 1 & 2 & 3 \\ 2 & 5 & 4 \\ 1 & 5 & -2 \end{vmatrix} \quad \begin{vmatrix} 1 & 2 & -2 \\ 2 & 5 & -3 \\ 1 & 5 & 0 \end{vmatrix}$$

$= 12-60+45 \qquad = 6-16+9 \qquad = -30+16+15 \qquad = 15+6-10$

$= -3 \qquad\qquad = -1 \qquad\qquad = 1 \qquad\qquad = -1$

$$Adj(A) = \begin{bmatrix} 51 & -(15) & -7 & -(12) \\ -(31) & 9 & -(4) & 7 \\ 10 & -(3) & -1 & -(2) \\ -(-3) & -1 & -(1) & -1 \end{bmatrix}$$

$$A^{-1} = \frac{1}{|A|} Adj(A) = \frac{1}{-1} Adj(A)$$

$$= \begin{bmatrix} -51 & 15 & 7 & 12 \\ 31 & -9 & 4 & -7 \\ -10 & 3 & 3 & 2 \\ -3 & 1 & 1 & 1 \end{bmatrix}$$

2) <u>Given :</u>

$$\frac{O_1 + O_2 + O_3 + O_4 + O_5}{5} = 3 \ (\mu) \quad \text{--- (1)}$$

$$O_1^2 + O_2^2 + O_3^2 = 14 \quad \text{--- (2)}$$

$$O_3^2 + O_4^2 + O_5^2 = 50 \quad \text{--- (3)}$$

$$O_1^2 + O_2^2 = 5 \quad \text{--- (4)}$$

Sobstitute (4) in (2)

$$5 + O_3^2 = 14$$

$$O_3 = 3 \quad \text{--- (5)}$$

substite O_3 in (2) & add with (3)

$$O_1^2 + O_2^2 + O_3^2 + O_4^2 + O_5^2 = 50 + 14 - 9$$

$$= 55 \quad \text{--- (6)}$$

W.K.T

$$\sigma^2 = \frac{\Sigma(x-\mu)^2}{N}$$

$$= \frac{\Sigma(x^2 - 2\mu x + \mu^2)}{N}$$

$$= \frac{\Sigma x^2}{N} - 2\mu\left(\frac{\Sigma x}{N}\right) + \mu^2 \quad \left[\because \frac{\Sigma x}{N} = \mu\right]$$

$$= \frac{\Sigma x^2}{N} - 2\mu^2 + \mu^2$$

$$= \frac{\Sigma x^2}{N} - \mu^2 \quad \text{--- (7)}$$

Substitute ① & ⑥ in ⑦.

$$\sigma^2 = \frac{55}{5} - 9$$

$$= 11 - 9$$

$$\sigma = \sqrt{2}$$

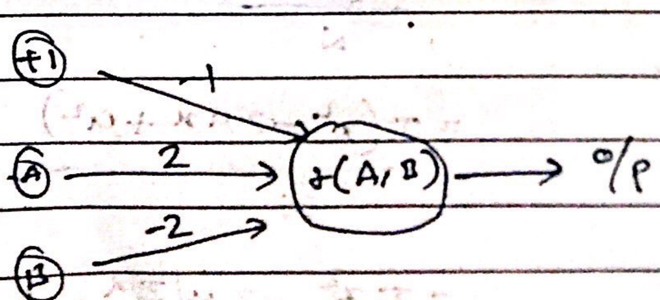
3) a) True.

- A decision tree forms a subset of all the features at each node
- if more nodes are available, the noise (with very varying parameters) will also fit.

b) True (Generally)

- this is ~~called~~ result of "curse of dimensionality"
- but there can be few corner cases in which this can be invalid.

c) Yes, it's possible.



$$\underbrace{\begin{bmatrix} 1 & 1 & 1 \\ 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix}}_{1_{bias}} \times \underbrace{\begin{bmatrix} -1 \\ 2 \\ -2 \end{bmatrix}}_{4\times3} = \begin{bmatrix} -1 \\ -1 \\ +1 \\ -3 \end{bmatrix} \quad \boxed{f_{threshold}\ 0} \Rightarrow \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \end{bmatrix}$$

d) the given table can be transformed into the following form.

A \ B	0	1
0	0.32	0.48
1	0.08	0.12

(i) $P(A=0) = P(A=0, B=0) + P(A=0, B=1)$

$= 0.8$

$P(A=1) = P(A=1, B=0) + P(A=1, B=1)$

$= 0.2$

$P(B=0) = P(A=0, B=0) + P(A=1, B=0)$

$= 0.4$

$P(B=1) = P(A=0, B=1) + P(A=1, B=1)$

$= 0.6$

(ii) $P(A=0, B=0) = P(A=0) P(B=0)$

$= 0.8 \times 0.4$

$= 0.32$

$P(A=1, B=1) = P(A=1) P(B=1)$

$= 0.2 \times 0.6 = 0.12$

$P(A=0, B=1) = P(A=0) P(B=1)$

$= 0.8 \times 0.6 = 0.48$

$P(A=1, B=0) = P(A=1) P(B=0)$

$= 0.2 \times 0.4$

$= 0.08$

as these form the given joint distribution,

A & B are independent.

CLASSTIME Page No.
Date / /

5) (A) (i) cases of overfit
- if k is too less
- using test set alone for hyperparameter tuning

(ii) reducing overfitting
- Cross Validation
- choose k using elbow methow
- Include preprocessing step to remove noise (outlyers)

(B) (i) cases of overfit
- too ~~fit~~ little training data
- noisy training data
- high depth
- high number of leaves
- basically a complex tree

(ii) reducing overfit
- limit max depth
- threshold for node splitting criteria
2 use pruning
- do not grow the split becomes insignificant.

6)

CLASSTIME Page No.
Date / /

6) <u>Given</u>

$$X = \begin{bmatrix} 252 & 4 & 155 & 175 \\ 175 & 10 & 186 & 200 \\ 82 & 131 & 230 & 100 \\ 115 & 138 & 80 & 88 \end{bmatrix}_{4\times4}$$

let us split bias & actual weights weights.

$$Y = \begin{bmatrix} 1 \\ 1 \\ 0 \\ 0 \end{bmatrix}_{4\times1}$$

$$b0 = [-0.00469 \quad 0.00797]$$

$$b1 = [-0.00588 \quad -0.00232]$$

$$W0 = \begin{bmatrix} -0.00256 & 0.0089 \\ 0.00146 & 0.00322 \\ 0.00816 & 0.00258 \\ -0.00597 & -0.00876 \end{bmatrix}_{4\times2}$$

$$W1 = \begin{bmatrix} -0.00647 & 0.00540 \\ 0.00374 & -0.00005 \end{bmatrix}_{2\times2}$$

$$\sigma(x) = \frac{1}{1+e^{-x}} \qquad softmax(x) = \left[\frac{e_i^{x}}{\Sigma e^{x} \ldots} \right]$$

$$Cross\ entropy(x) = -\sum_{c=1}^{c} x(y_c) \cdot \log(x(y_c))$$

Derivatives

Sigmoid: $\sigma = \frac{1}{1+e^{-x}}$

$$\frac{\partial \sigma}{\partial x} = \frac{1}{1+e^{-x}} \times \left(1 - \frac{1}{1+e^{-x}}\right)$$

$$= \sigma \times (1-\sigma)$$

Softmax:

$$softmax = \frac{e_i^{x}}{\Sigma e_i^{x}}$$

In our case two nodes are there in o/p

$= [0, 1]$

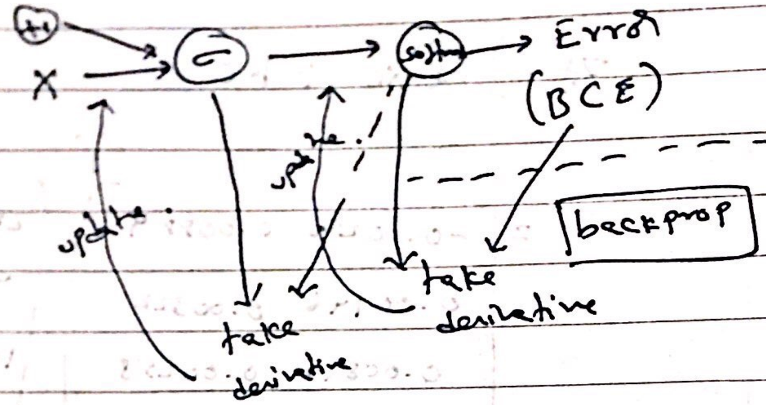
$$softmax = \frac{e_p^{x}}{e_0^{x} + e_1^{x}}$$

$$\frac{\partial softmax}{\partial x_1} = \frac{e_1^{x} \times (e_2^{x})}{(e_1^{x} + e_2^{x})^2}$$

$$\frac{\partial (softmax)}{\partial x_2} = \frac{e_2^{x} \times e_1^{x}}{(e_1^{x} + e_2^{x})^2}$$

CLASSTIME Page No.
Date / /

Architecture.



err : Binary cross entropy

activation: layer1 – sigmoid

final layer – softmax.

$$\text{Error} = \frac{-1}{n}\left(y_i \times \log x_i + (1-y_i) \log(1-x_i) \right)$$

$o/p = z_2$

$$\frac{\partial E}{\partial o/p} = \frac{\partial}{\partial o/p}\left(-1 \times \left(y_i \times \log(z_2) + (1-y_i)\log(1-z_2)\right)\right)$$

$$= -1\left(y_1 \times \left(\frac{1}{z_2}\right) + (1-y_1) \times \frac{1}{(1-z_2)}\right)$$

CLASSTIME	Page No.
Date / /	

$$a_0 = X^T$$

$$Z_1 = W_0^T x_0 + b_0$$

$$= \begin{bmatrix} -0.00256 & 0.00889 \\ 0.00146 & 0.00322 \\ 0.00816 & 0.00258 \\ -0.00597 & -0.00876 \end{bmatrix}^T_{4\times 2} \begin{bmatrix} 252 & 175 & 82 & 115 \\ 4 & 10 & 131 & 131 \\ 155 & 186 & 230 & 80 \\ 175 & 200 & 100 & 88 \end{bmatrix}_{4\times 4}$$

$$+ \begin{bmatrix} -0.00469 \\ 0.00997 \end{bmatrix}$$

$$= [(4\times 4)\times(4\times 2)] + (2\times 1)$$

$$\downarrow$$

$$[4\times 2] + (2\times 1)$$

$$= [4\times 2]^T \longrightarrow 2\times 4$$

$$= \begin{bmatrix} -0.42392 & -0.11433 & 1.25645 & 0.02883 \\ 1.12803 & 0.3238 & 0.87617 & 0.9102 \end{bmatrix}$$

$a_1 = \sigma(z_1)$

$$= \begin{bmatrix} 0.3955791 & 0.471 & 0.778 & 0.507 \\ 0.755 & 0.580 & 0.706 & 0.713 \end{bmatrix}$$

$z_2 = W_1^T a_1 + b_1$

$= (2\times2) \times (2\times4) + (1\times2)$

$= (2\times4)$

$$= \begin{bmatrix} -0.0056 & -0.0067 & -0.0082 & -0.0064 \\ -0.0002 & -0.0001 & 0.0018 & 0.0003 \end{bmatrix}_{2\times4}$$

$a_2 = \text{softmax}(z_2)$

$$= \begin{bmatrix} 0.4986 & 0.4982 & 0.4974 & 0.4982 \\ 0.5013 & 0.5017 & 0.5025 & 0.5017 \end{bmatrix}$$

Cross entropy derivative can be also taken as $\dfrac{\text{softmax} - \text{true}}{m \text{ (training samples)}}$

delta 2 = CE(a2, y in one hot encoding form)

- a2 is from softmax.

$$= \left[\begin{bmatrix} 0.4986 & 0.4986 & 0.4974 & 0.4982 \\ 0.5013 & 0.5017 & 0.5025 & 0.5017 \end{bmatrix} - \begin{bmatrix} 0 & 0 & 1 & 1 \\ 1 & 1 & 0 & 0 \end{bmatrix} \right] \times \frac{1}{4}$$

for derivatives of Error on weights, we have to use chain rule.

$$\frac{\partial E}{\partial w} = \frac{\partial E}{\partial a} \times \frac{\partial a}{\partial h} \times \frac{\partial h}{\partial w}$$

$$\text{delta1} = W^T \text{delta2} \times \frac{\partial}{\partial w}(a_1)$$

a_1 already came from sigmoid, so, we need to apply derivative of sigmoid.

$$= \begin{bmatrix} -3.53 & -3.68 & 2.37 & 3.72 \\ 8.72 & 1.14 & -9.80 & -9.72 \end{bmatrix} e^{-4}$$

(5)

(11)

CLASSTIME	Page No.
Date / /	

Deltao = Deltao + $\text{delta1}^T \text{ao}$

Delta1 = Delta1 + $\text{delta2}^T \text{a1}$

Delta, actual error that we will backpropagate

we need to normalize the error.

$$\text{Delta0} = \begin{bmatrix} -0.08 & -0.07 & -0.03 & -0.07 \\ 0.02 & -0.02 & -0.004 & 0.01 \end{bmatrix}$$

$$\text{Delta1} = \begin{bmatrix} -0.05 & -0.01 \\ 0.05 & 0.01 \end{bmatrix}$$

<u>weights updation:</u>

$$w_0 = w_0 - lr \times \frac{1}{m} \text{Delta0}$$

$$w_1 = w_1 - lr \times \frac{1}{m} \text{Delta1}$$

$\boxed{lr = 0.1}$ $\boxed{m = 4}$

$$w_0 = \begin{bmatrix} -0.00256 & 0.00889 \\ -0.00496 & 0.00322 \\ 0.00816 & 0.00259 \\ -0.00597 & -0.00876 \end{bmatrix} - 0.1 \times \frac{1}{4} \begin{bmatrix} -0.03 & 0.013 \\ 0.03 & -0.013 \\ -0.014 & 0.002 \\ -0.03 & 0.01 \end{bmatrix}$$

$$= \begin{bmatrix} -0.00031 & 0.00831 \\ -0.00053 & 0.0038 \\ 0.0090 & 0.00246 \\ 0.00604 & -0.00925 \end{bmatrix} = \begin{bmatrix} -0.00031629 & 0.008378 \\ -0.0005387 & 0.00417 \\ 0.009020 & 0.002459 \\ -0.004041 & -0.00925 \end{bmatrix}$$

CLASSTIME Page No.
Date /

$$w_{12} = \begin{bmatrix} -0.003 & 0.002 \\ 0.005 & -0.0017 \end{bmatrix} - 0.1 \begin{bmatrix} -0.0133 & 0.0133 \\ -0.0029 & 0.0029 \end{bmatrix}$$

$$= \begin{bmatrix} -0.0051 & 0.004 \\ 0.004 & -0.0003 \end{bmatrix}$$

Calculation of bias:

bias has to be summed up from every input to the next node.

eg: $\Delta 1 = \begin{bmatrix} d_1 & d_2 \\ d_3 & d_4 \\ d_5 & d_6 \end{bmatrix}_{3\times2} \Rightarrow [\Sigma d_1 d_3 d_5 \quad d_2 d_4 d_6]_{1\times2}$

bias = bias $- lr \times \frac{1}{m} \times$ delta.sum()

$b_0 = [-0.00469 \quad 0.00797] - lr \times \frac{1}{4} \times [-9.259, 6.7] \times 10^{-4}$

$= [-0.004687768 \quad 0.007969985]$

$b_1 = [-0.00588, -0.00232] - 0.1 \times \frac{1}{4} \times [-0.00183, 0.00183]$

$= [-0.00583413, 0.00236597]$.

CLASSTIME Page No.
Date / /

Things to note:

- forward propogation done
- derivatives calculated
- weights updated seperately
- bias updated seperately
- Included Regularization
- learning rate = 0.1
- Converges in 4 iterations.

∴ used
Sigmoid → softmax → binary cross entropy.

CLASSTIME Page No.
Date / /

4)

$$A = \begin{bmatrix} 1 & 1 & 1 & 0 & 0 \\ -3 & -3 & -3 & 0 & 0 \\ 2 & 2 & 2 & 0 & 0 \\ 0 & 0 & 0 & -1 & -1 \\ 0 & 0 & 0 & 2 & 2 \\ 0 & 0 & 0 & -1 & -1 \end{bmatrix}$$

SVD(A) will be in form $U \Sigma V^T$

U = eigenvectors$(A \cdot A^T)$

V = eigenvectors$(A^T A)$

Calculation of U

$$A \cdot A^T = \underset{6\times5}{\begin{bmatrix} 1 & 1 & 1 & 0 & 0 \\ -3 & -3 & -3 & 0 & 0 \\ 2 & 2 & 2 & 0 & 0 \\ 0 & 0 & 0 & -1 & -1 \\ 0 & 0 & 0 & 2 & 2 \\ 0 & 0 & 0 & -1 & -1 \end{bmatrix}} \cdot \underset{5\times6}{\begin{bmatrix} 1 & -3 & 2 & 0 & 0 & 0 \\ 1 & -3 & 2 & 0 & 0 & 0 \\ 1 & -3 & 2 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 & 2 & -1 \\ 0 & 0 & 0 & -1 & 2 & -1 \end{bmatrix}}$$

$$= \begin{bmatrix} 3 & -9 & 6 & 0 & 0 & 0 \\ -9 & 27 & -18 & 0 & 0 & 0 \\ 6 & -18 & 12 & 0 & 0 & 0 \\ 0 & 0 & 0 & 2 & -4 & 2 \\ 0 & 0 & 0 & -4 & 8 & -4 \\ 0 & 0 & 0 & 2 & -4 & 2 \end{bmatrix}$$

to find eigen values.

$$|A-\lambda I| = 0$$

$$A-\lambda I = \begin{bmatrix} 3-\lambda & -9 & 6 & 0 & 0 & 0 \\ -9 & 27-\lambda & -18 & 0 & 0 & 0 \\ 6 & -18 & 12-\lambda & 0 & 0 & 0 \\ 0 & 0 & 0 & 2-\lambda & -4 & 2 \\ 0 & 0 & 0 & -4 & 8-\lambda & -4 \\ 0 & 0 & 0 & 2 & -4 & 2-\lambda \end{bmatrix}$$

Gaussian elimination

$$R_2 = R_2 - \frac{9}{\lambda-3} R_1$$

$$= \begin{bmatrix} -\lambda+3 & -9 & 6 & 0 & 0 & 0 \\ 0 & \frac{-\lambda^2+30\lambda}{\lambda-3} & \frac{-18\lambda}{\lambda-3} & 0 & 0 & 0 \\ 6 & -18 & -\lambda+12 & 0 & 0 & 0 \\ 0 & 0 & 0 & -\lambda+2 & -4 & 2 \\ 0 & 0 & 0 & -4 & -\lambda+8 & -4 \\ 0 & 0 & 0 & 2 & -4 & -\lambda+2 \end{bmatrix}$$

$R_3 = R_3 - \frac{-6}{\lambda-3} R_1$

$$= \begin{bmatrix} -\lambda+3 & -9 & 6 & 0 & 0 & 0 \\ 0 & \frac{-\lambda^2+30\lambda}{\lambda-3} & \frac{-18\lambda}{\lambda-3} & 0 & 0 & 0 \\ 0 & \frac{-18\lambda}{\lambda-3} & \frac{-\lambda^2+15\lambda}{\lambda-3} & 0 & 0 & 0 \\ 0 & 0 & 0 & -\lambda+2 & -4 & 2 \\ 0 & 0 & 0 & -4 & -\lambda+8 & -4 \\ 0 & 0 & 0 & 2 & -4 & -\lambda+2 \end{bmatrix}$$

$R_3 = R_3 - \frac{18}{\lambda-30} R_2$

$R_5 = R_5 - \frac{4}{\lambda-2} R_4$

$R_6 = R_6 - \frac{-2}{\lambda-2} R_4$

$$= \begin{bmatrix} 0 & \frac{-\lambda^2+30\lambda}{\lambda-3} & \frac{-18\lambda}{\lambda-3} & 0 & 0 & 0 \\ 0 & 0 & \frac{-\lambda^2+42\lambda}{\lambda^2-30} & 0 & 0 & 0 \\ 0 & 0 & 0 & -\lambda+2 & -4 & 2 \\ 0 & 0 & 0 & 0 & \frac{-\lambda^2+10\lambda}{\lambda-2} & \frac{-4\lambda}{\lambda-2} \\ 0 & 0 & 0 & 0 & \frac{-4\lambda}{\lambda-2} & \frac{-\lambda^2+4\lambda}{\lambda-2} \end{bmatrix}$$

$$R_6 = R_6 - \frac{4}{\lambda - 10} R_5$$

finally,

$$= \begin{vmatrix} -\lambda+3 & -9 & 6 & 0 & 0 & 0 \\ 0 & \frac{-\lambda^2+30\lambda}{\lambda-3} & \frac{-18\lambda}{\lambda-3} & 0 & 0 & 0 \\ 0 & 0 & \frac{-\lambda^2+42\lambda}{\lambda-30} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \frac{-\lambda^2+10\lambda}{\lambda-2} & \frac{-4\lambda}{\lambda-2} \\ 0 & 0 & 0 & 0 & 0 & \frac{-\lambda^2+12\lambda}{\lambda-10} \end{vmatrix}$$

$$= (-\lambda+3)\left(\frac{-\lambda^2+30\lambda}{\lambda-3}\right)\left(\frac{-\lambda^2+42\lambda}{\lambda-30}\right)(-\lambda+2)$$
$$\left(\frac{-\lambda^2+10\lambda}{\lambda-2}\right)\left(\frac{-\lambda^2+12\lambda}{\lambda-10}\right)$$

$$= \lambda^6 - 54\lambda^5 + 504\lambda^4$$

$$= \lambda^4(\lambda^2 - 54\lambda + 504)$$

$$= \lambda\lambda\lambda\lambda(\lambda-12)(\lambda-42)$$

$$\Rightarrow (\lambda = 12, 0, 42)$$

CLASSTIME | Page No.
Date / /

Sub $\lambda = 0$ in $(A - \lambda I)$

$$\begin{bmatrix} 3 & -9 & 6 & 0 & 0 & 0 \\ -9 & 27 & -18 & 0 & 0 & 0 \\ 6 & -18 & 12 & 0 & 0 & 0 \\ 0 & 0 & 0 & 2 & -4 & 2 \\ 0 & 0 & 0 & -4 & 8 & -4 \\ 0 & 0 & 0 & 2 & -4 & 2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

$R_1 = \frac{R_1}{3}$

$\cancel{R_2 = 9R_1}$

$R_2 = R_2 + 9R_1$

$R_3 = R_3 - 6R_1$

$R_4 \Leftrightarrow R_2$

$R_2 = \frac{R_2}{2}$

$R_5 = R_5 + 4R_2$

$R_6 = R_6 - 2R_1$

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

$x_1 - 3x_2 + 2x_3 = 0$

$x_4 - 2x_5 + x_6 = 0$

$$X = \begin{bmatrix} 3x_2 - 2x_3 \\ x_2 \\ x_3 \\ 2x_5 - x_6 \\ x_5 \\ x_6 \end{bmatrix}$$

$$\Rightarrow x_2 \times \begin{bmatrix} 3 \\ 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} + x_3 \times \begin{bmatrix} -2 \\ 0 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} + x_5 \begin{bmatrix} 0 \\ 0 \\ 0 \\ 2 \\ 0 \\ 1 \end{bmatrix} + x_6 \begin{bmatrix} 0 \\ 0 \\ 0 \\ -1 \\ 0 \\ 1 \end{bmatrix}$$

CLASSTIME	Page No.
Date / /	

solving gives

$$\text{Vector} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ -1 \\ 0 \\ 1 \end{bmatrix}$$

Similarly for $\lambda_2 = 42$:

$$V = \begin{bmatrix} \frac{1}{2} \\ \frac{-3}{2} \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

for $\lambda_3 = 12$

$$V = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \\ -2 \\ 1 \end{bmatrix}$$

So finally

$$U = \begin{bmatrix} -0.26 & 0 & -0.94 & 0 & -0.17 & 0 \\ 0.8 & 0 & -0.3 & 0 & 0.5 & 0 \\ -0.5 & 0 & -0.0 & 0 & 0.8 & 0 \\ 0 & -0.4 & 0 & -0.9 & 0 & 0 \\ 0 & 0.8 & 0 & -0.3 & 0 & 0.4 \\ 0 & -0.4 & 0 & -0.1 & 0 & 0.8 \end{bmatrix}$$

Similarly for v

~~$v = |A - \lambda I| = 0$~~

$$A^T A = \begin{bmatrix} 14 & 14 & 14 & 0 & 0 \\ 14 & 14 & 14 & 0 & 0 \\ 14 & 14 & 14 & 0 & 0 \\ 0 & 0 & 0 & 6 & 6 \\ 0 & 0 & 0 & 6 & 6 \end{bmatrix}$$

doing $|A - \lambda I| = 0$

gives

$$-\lambda^5 + 54\lambda^4 - 504\lambda^3 = 0$$

$$\lambda = [0, 0, 0, 42, 12]$$

gives, for $\lambda = 0$ $\begin{bmatrix} -1 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} \begin{bmatrix} -1 \\ 0 \\ 1 \\ 0 \\ 0 \end{bmatrix} \begin{bmatrix} 0 \\ 0 \\ 0 \\ -1 \\ 1 \end{bmatrix}$

for $\lambda = 42 = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 0 \\ 0 \end{bmatrix}$

for $\lambda = 12 = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \\ 1 \end{bmatrix}$

So finally,

$$V = \begin{bmatrix} 0.82 & -0.4 & -0.4 & 0 & 0 \\ -0.5 & -0.5 & -0.5 & 0 & 0 \\ 0 & -0.7 & 0.7 & 0 & 0 \\ 0 & 0 & 0 & 0.7 & -0.7 \\ 0 & 0 & 0 & -0.7 & -0.7 \end{bmatrix}$$

CLASSTIME Page No.
Date / /

$$\Sigma = \begin{bmatrix} \sqrt{42} & 0 & 0 & 0 & 0 & 0 \\ 0 & \sqrt{12} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

(ii) first principle Component

$$U\Sigma = \begin{bmatrix} 0.26 & 0 & 0 & 0 & 0.8 & -0.4 \\ -0.8 & 0 & 0 & 0 & 0.1 & 0.4 \\ 0.5 & 0 & 0 & 0 & -0.1 & -0.3 \\ 0 & 0.4 & 0.5 & 0 & 0 \\ 0 & 0.8 & 0.5 & 0 & 0 \\ 0 & 0.4 & 0.5 & 0 & 0 \end{bmatrix}$$

$$\times \begin{bmatrix} \sqrt{42} & 0 & 0 & 0 & 0 & 0 \\ 0 & \sqrt{12} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

$$= \begin{bmatrix} 1.6 & 0 & 0 & 0 & 0 \\ -5.1 & 0 & 0 & 0 & 0 \\ 3.4 & 0 & 0 & 0 & 0 \\ 0 & 1.4 & 0 & 0 & 0 \\ 0 & -2.8 & 0 & 0 & 0 \\ 0 & 1.4 & 0 & 0 & 0 \end{bmatrix}$$

~~The first Principle Components are~~

~~1.6 3.4 -5.19~~

CLASSTIME	Page No.
Date / /	

first Principle Component is

$$\begin{bmatrix} 1.6848 \\ -5.19048 \\ 3.4344 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$