Input Sentence:

[0.3,0.4,0.7]	This
[0.1,0.2,0.7]	is
[0.2,0.2,0.8]	a
[0.3,0.7,0.8]	book

W^Q

1	1	0
1	1	0
1	1	0

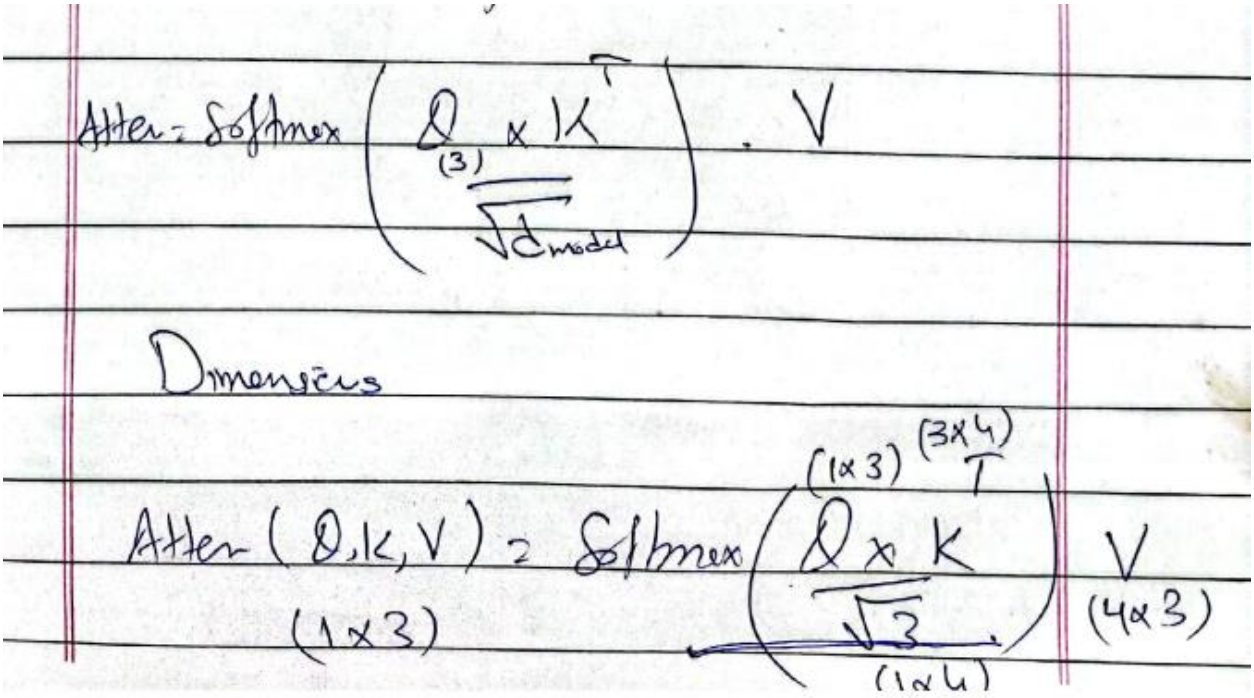
W^K

1	0	1
1	0	1
1	0	1

W^V

0	1	1
0	1	1
0	1	1

Compute self-attention for the 3rd word.



$$\underset{1\times3}{Q} = \underset{1\times3}{X_3} \times \underset{3\times3}{W^Q} \rightarrow [1.2, 1.2, 0]$$

$$K = X * W^K \rightarrow$$

$$\begin{bmatrix} 0.3 & 0.4 & 0.7 \\ 0.1 & 0.2 & 0.7 \\ 0.2 & 0.2 & 0.8 \\ 0.3 & 0.7 & 0.8 \end{bmatrix}_{(4\times3)} \quad \begin{bmatrix} 1 & 0 & 1 \\ 1 & 0 & 1 \\ 1 & 0 & 1 \end{bmatrix}_{(3\times3)}$$

$$K = \underset{(4\times3)}{\begin{bmatrix} 1.4 & 0 & 1.4 \\ 1 & 0 & 1 \\ 1.2 & 0 & 1.2 \\ 1.8 & 0 & 1.8 \end{bmatrix}}$$

$$V = \begin{bmatrix} 0 & 1.4 & 1.4 \\ 0 & 1 & 1 \\ 0 & 1.2 & 1.2 \\ 0 & 1.8 & 1.8 \end{bmatrix} \quad (4\times 3)$$

$$Q \times K^T = [1.2, 1.2, 0] \begin{bmatrix} 1.4 & 1 & 1.2 & 1.8 \\ 0 & 0 & 0 & 0 \\ 1.4 & 1 & 1.2 & 1.8 \end{bmatrix}$$

1×3 $\qquad$ 3×4

$$= [1.68, 1.2, 1.44, 2.16]$$

$$\frac{Q \times K^T}{\sqrt{3}} = [0.97, 0.69, 0.83, 1.25]$$

$$\text{Softmax}\left(\frac{Q \times K^T}{\sqrt{3}}\right) = [0.25, 0.19, 0.22, 0.33]$$

$$\text{Softmax}\left(\frac{Q \times K^T}{\sqrt{3}}\right) \cdot V$$

$(1\times4)\quad(4\times3)$

$$= [0.25, 0.19, 0.22, 0.33] \begin{Bmatrix} 0 & 1.4 & 1.4 \\ 0 & 1 & 1 \\ 0 & 1.2 & 1.2 \\ 0 & 1.8 & 1.8 \end{Bmatrix}$$

(1×4) $\qquad$ (4×3)

$$= [0,\ 1.398,\ 1.398]$$