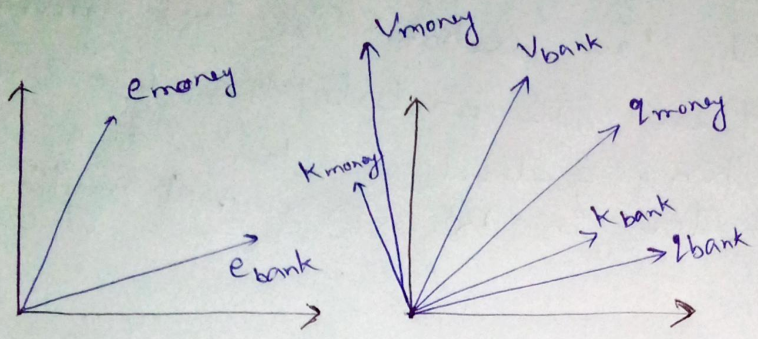
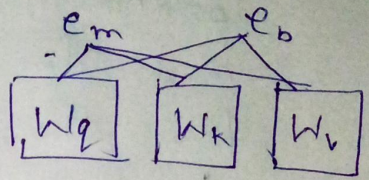
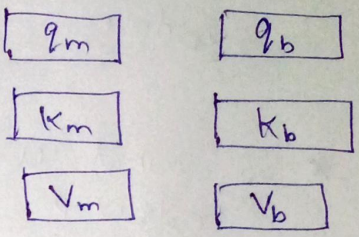
11 Mar 2025

SELF ATTENTION GEOMETRIC INTUITION



"money bank"





$W_q = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}$ $W_K = \begin{bmatrix} 3 & 4 \\ 5 & 1 \end{bmatrix}$ $W_V = \begin{bmatrix} 4 & 1 \\ 2 & 1 \end{bmatrix}$

- All values are hypothetical

Dot product

$S_{21} = 10$ $S_{22} = 32$

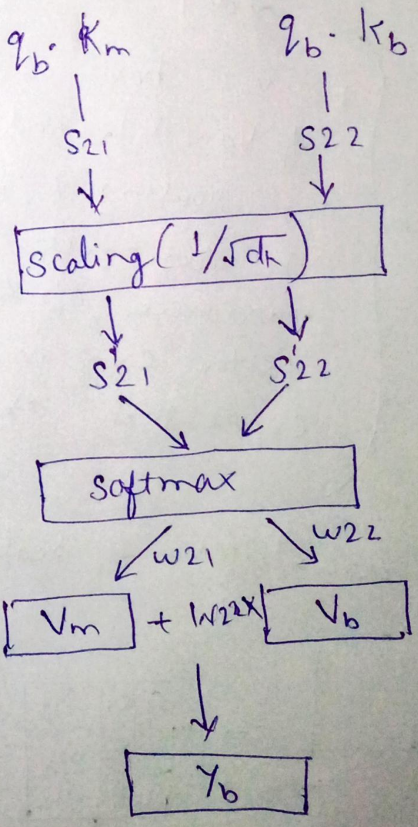
Scaling

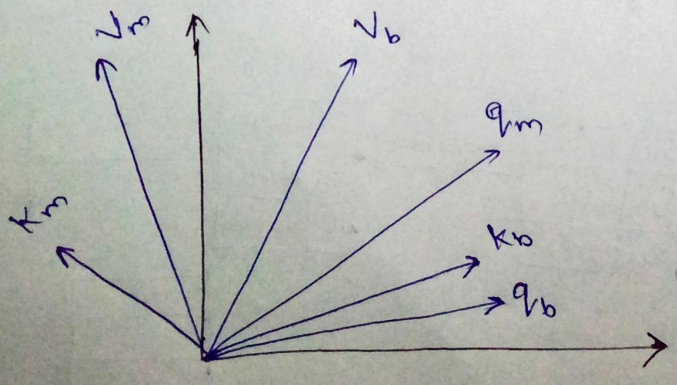
$S'_{21} = \frac{10}{\sqrt{2}} = 7.09$ $S'_{22} = \frac{32}{\sqrt{2}} = 22.69$

Softmax

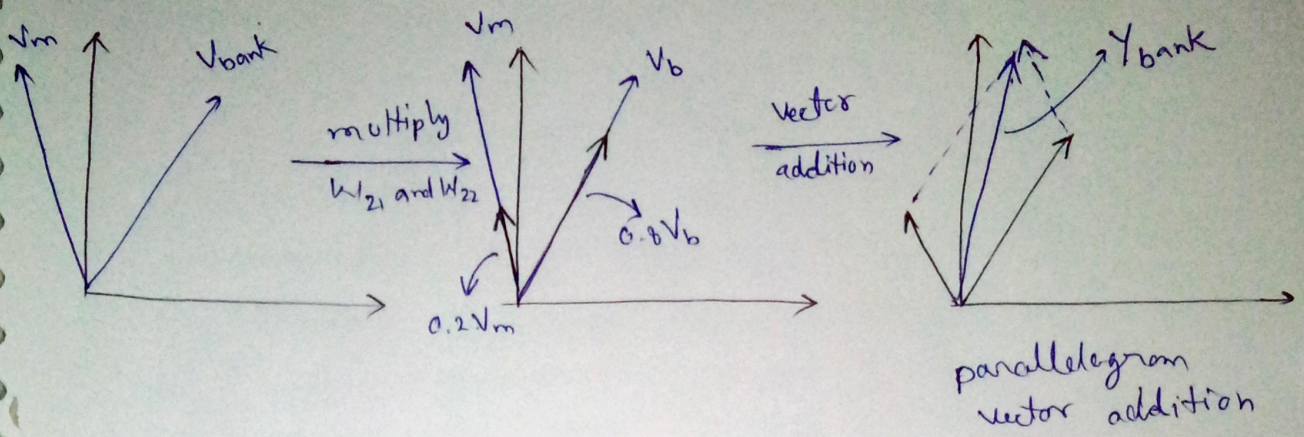
$W_{21} = 0.2$ $W_{22} = 0.8$

all above values are assumption.





Here $d_k = 2$



— Now compare the Y_{bank} with original embedding vectors