2025 USA-NA-AIO Round 2, Problem 2, Part 3

USAAIO 🛡️   **May 2025**

Part 3 (10 points, non-coding task)

Define function $\mathrm{Softmax} : \mathbb{R}^d \to \mathbb{R}^d$, with the ith output value as

$$\mathrm{Softmax}_i(\mathbf{z}) = \frac{\exp(z_i)}{\sum_{j=0}^{d-1} \exp(z_j)}.$$

At position l_1 in the attending sequence, its attention score to position l_2 in the being attended sequence for head h is denoted as $\alpha_{h,l_1 l_2}$.

We can write $\alpha_{h,l_1 l_2}$ in the following form:

$$\alpha_{h,l_1 l_2} = \mathrm{Softmax}_{l_2}\left(\boxed{???}\right),$$

What is the formula in the above red box (reasoning is not required)?

USAAIO 🛡️   **May 2025**

Misplaced '#'

$$\alpha_{h,l_1 l_2} = \mathrm{Softmax}_{l_2}\left(\frac{\mathbf{q}_{h,l_1}^\top \mathbf{K}_h^\top}{\sqrt{D_{qk}}}\right),$$

where

$$\mathbf{K}_h = \begin{bmatrix} \mathbf{k}_{h,0}^\top \\ \mathbf{k}_{h,1}^\top \\ \vdots \\ \mathbf{k}_{h,L_2-1}^\top \end{bmatrix} \in \mathbb{R}^{L_2 \times D_{qk}}.$$

"" END OF THIS PART ""

✦ Related topics

Topic	Replies	Activity
2025 USA-NA-AIO Round 2, Problem 2, Part 5	1	**May 2025**
2025 USA-NA-AIO Round 2, Problem 2, Part 1	2	**Dec 2025**
2025 USA-NA-AIO Round 2, Problem 2, Part 13	1	**May 2025**
2025 USA-NA-AIO Round 2, Problem 2, Part 4	1	**May 2025**

Skip to main content

Topic	Replies	Activity
2025 USA-NA-AIO Round 2, Problem 2, Part 2	2	Dec 2025