

問い: 下記の問いに答えよ。

1. 長さ N の文章をトランスフォーマーにより処理する。 n 番目の単語に対応するトークンを $\mathbf{x}_n \in \mathbb{R}^D$ で表し、 $\mathbf{X} = [\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_N]^\top \in \mathbb{R}^{N \times D}$ とする。トランスフォーマーは入力トークン系列 \mathbf{X} を次式に従ってクエリ (\mathbf{Q})、キー (\mathbf{K})、バリュー (\mathbf{V}) の系列へと変換する:

$$\mathbf{Q} = \mathbf{X}\mathbf{W}^{(q)}, \mathbf{K} = \mathbf{X}\mathbf{W}^{(k)}, \mathbf{V} = \mathbf{X}\mathbf{W}^{(v)}.$$

計算した結果 \mathbf{Q} と \mathbf{K} と $\mathbf{Q}\mathbf{K}^\top$ は次式のとおりとなった。 $\exp(1/2) \simeq 3/2$ と近似したうえで $\text{SM}(\mathbf{Q}\mathbf{K}^\top / \sqrt{D_k})$ を計算せよ。ただし D_k はキーベクトルの次元である。行列の各要素は整数もしくは分数で記すこと。

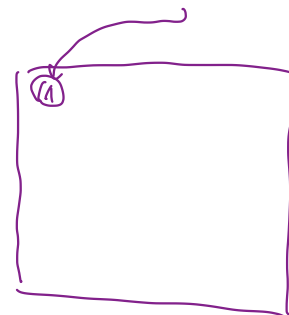
$N=4$

$$\mathbf{Q} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ -1 & 0 \\ 0 & -1 \end{bmatrix}, \mathbf{K} = \begin{bmatrix} 1/\sqrt{2} & 1/\sqrt{2} \\ -1/\sqrt{2} & 1/\sqrt{2} \\ -1/\sqrt{2} & -1/\sqrt{2} \\ 1/\sqrt{2} & -1/\sqrt{2} \end{bmatrix}, \mathbf{Q}\mathbf{K}^\top = \begin{bmatrix} 1/\sqrt{2} & -1/\sqrt{2} & -1/\sqrt{2} & 1/\sqrt{2} \\ 1/\sqrt{2} & 1/\sqrt{2} & -1/\sqrt{2} & -1/\sqrt{2} \\ -1/\sqrt{2} & 1/\sqrt{2} & 1/\sqrt{2} & -1/\sqrt{2} \\ -1/\sqrt{2} & -1/\sqrt{2} & 1/\sqrt{2} & 1/\sqrt{2} \end{bmatrix}$$

$D_q = D_k = 2$

$$\frac{\mathbf{Q}\mathbf{K}^\top}{\sqrt{2}} = \begin{bmatrix} 1/2 & -1/2 & -1/2 & 1/2 \\ \vdots & \vdots & \vdots & \vdots \end{bmatrix}$$

$$\text{SM}(\cdot) \approx \frac{\exp(1/2)}{\exp(1/2) + \exp(-1/2) + \exp(-1/2) + \exp(1/2)}$$



また、 \mathbf{V} を計算したところ次式のとおりとなった。 \mathbf{y}_2 を求めよ。

$$\mathbf{V} = \begin{bmatrix} 12 & -3 \\ 14 & 16 \\ 7 & -10 \\ 6 & -3 \end{bmatrix}$$

ただし $\mathbf{Y} = \text{SM}(\mathbf{Q}\mathbf{K}^\top / \sqrt{D_k})\mathbf{V}$ であり、 $\mathbf{Y} = [\mathbf{y}_1, \mathbf{y}_2, \dots, \mathbf{y}_N]^\top$ である。

$$\mathbf{y}_2 = \begin{bmatrix} \text{SM}(\cdot) \end{bmatrix} \mathbf{V} = \begin{bmatrix} 11 & 5/2 \end{bmatrix}$$

④ 期末 Teams v. P-アート.