学籍番号　　　　　　　　　　　　　　氏名

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 $\boldsymbol{x}_n \in \mathbb{R}^D$ で表し、$\boldsymbol{X} = [\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_N]^\top \in \mathbb{R}^{N\times D}$ とする。トランスフォーマーは入力トークン系列 $\boldsymbol{X}$ を次式に従ってクエリ ($\boldsymbol{Q}$)、キー ($\boldsymbol{K}$)、バリュー ($\boldsymbol{V}$) の系列へと変換する:

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

計算した結果 $\boldsymbol{Q}$ と $\boldsymbol{K}$ と $\boldsymbol{Q}\boldsymbol{K}^\top$ 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 $\mathrm{SM}(\boldsymbol{Q}\boldsymbol{K}^\top/\sqrt{D_k})$ 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。

$N=4$

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

$D_Q = D_K = 2$　　　ひとつめ

$$\frac{QK^\top}{\sqrt{2}} = \begin{bmatrix} 1/2 & -1/2 & -1/2 & 1/2 \\ & \cdots & & \end{bmatrix}$$

$$\mathrm{SM}(\cdot) = \frac{\exp(\frac{1}{2})}{\exp(\frac{1}{2}) + \exp(-\frac{1}{2}) + \exp(-\frac{1}{2}) + \exp(\frac{1}{2})}$$

①

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

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

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

$y_2 = \mathrm{SM}(\quad)\ \sqrt{\ } = [\ \bigcirc\ \ \bigcirc\ ]$

$[\ 11 \quad 5/2\ ]$

@期末　Teams で アンケート.