Търсене и извличане на информация

Зад 1

$\boxed{1.}$ $\dfrac{\partial(J_{w,c,\bar{c}_1,\ldots,\bar{c}_n}(U,V,W))}{\partial u_w} = -\dfrac{\partial(\log \sigma(v_c^T W u_w - q_c))}{\partial u_w}$

> За съкратен запис
> ще обозначим $J := J_{w,c,\bar{c}_1,\ldots,\bar{c}_n}(u,v,w)$

$$-\sum_{j=1}^{n} \dfrac{\partial(\log \sigma(q_{\bar{c}_j} - v_{\bar{c}_j}^T W u_w))}{\partial u_w}$$

свойство на сигмоид
$$\overset{=}{=} -(1 - \sigma(v_c^T W u_w - q_c)) \cdot \dfrac{\partial(v_c^T W u_w - q_c)}{\partial u_w}$$

$$-\sum_{j=1}^{n} (1 - \sigma(q_{\bar{c}_j} - v_{\bar{c}_j}^T \cdot W u_w)) \cdot \dfrac{\partial(q_{\bar{c}_j} - v_{\bar{c}_j}^T \cdot W u_w)}{\partial u_w}$$

$$\boxed{\dfrac{\partial(v_c^T W u_w - q_c)}{\partial u_w} = W^T v_c}$$ $$\boxed{\dfrac{\partial(q_{\bar{c}_j} - v_{\bar{c}_j}^T W u_w)}{\partial u_w} = -W^T \cdot v_{\bar{c}_j}}$$

$$\Rightarrow \dfrac{\partial J}{\partial u_w} = \sum_{j=1}^{n} (1 - \sigma(q_{\bar{c}_j} - v_{\bar{c}_j}^T W u_w)) \cdot W^T v_{\bar{c}_j} -$$
$$- (1 - \sigma(v_c^T W u_w - q_c)) \cdot W^T \cdot v_c$$

Можем да изнесем W^T, понеже $\sigma(X) \in \mathbb{R}$ за $X \in \mathbb{R}$ (скалар)

$$\Rightarrow \dfrac{\partial J}{\partial u_w} = W^T \left(\sum_{j=1}^{n} (1 - \sigma(q_{\bar{c}_j} - v_{\bar{c}_j}^T W u_w)) \cdot v_{\bar{c}_j} - (1 - \sigma(v_c^T W u_w - q_c)) \cdot v_c \right)$$

2. По условие $\overline{c_{i,j}}$ са индексите на и разл. случайно избрани думи, различни

от c_i.

$\Rightarrow$ В израза със сума не присъства векторът V_c.

$$\Rightarrow \frac{\partial\left(-\log \sigma(V_c^T W_{uw} - q_c) + \sum_{j=1}^{n} \text{...}\right)}{\partial V_c}$$

$$= \frac{\partial\left(\log \sigma(V_c^T W_{uw} - q_c)\right)}{\partial V_c} = (1 - \sigma(V_c^T W_{uw} - q_c)) \cdot W_{uw}$$

$$= \frac{\partial J}{\partial V_c}$$

3. Аналогично на 2. за непресичащите се вектори между $\#j = \overline{1, n}$: $V_{\overline{c_j}} \neq V_c$ и са независими

$$\Rightarrow \frac{\partial\left(\log \text{...} + \sum_{j=1}^{n} \text{...}\right)}{\partial V_{\overline{c_j}}} = -\sum_{j=1}^{n} \frac{\partial\left(\log \sigma(q_{\overline{c_j}} - V_{\overline{c_j}}^T W_{uw})\right)}{\partial V_{\overline{c_j}}}$$

$$\frac{\partial J}{\partial V_{\overline{c_j}}} = (1 - \sigma(q_{\overline{c_j}} - V_{\overline{c_j}}^T W_{uw})) \cdot W_{uw}$$

4. За съкратен запис $\{v_c^T W u_w - q_c\} = g_1(w)$

$$\frac{\partial J}{\partial W} = \text{~~~~~} \{q_{\bar c_j} - v_{\bar c_j}^T W u_w\} = g_2(w)$$

$$= - \frac{\partial \log(g_1(w))}{\partial W} - \sum_{j=1}^{n} \frac{\partial \log(g_2(w))}{\partial W}$$

$$= -(1 - \sigma(g_1(w))) \cdot \frac{\partial(g_1(w))}{\partial W} - \sum_{j=1}^{u} (1 - \sigma(g_2(w))) \frac{\partial g_2(w)}{\partial W}$$

$$g_1^*(w) = v_c^T W u_w = \sum_{j=1}^{M} \left(\sum_{i=1}^{M} v_{c,i} \cdot W_{i,j} \right) \cdot u_{w,j}$$

$$= \sum_{i=1}^{M} \sum_{j=1}^{M} v_{c,i} \cdot W_{i,j} \cdot u_{w,j}$$

$$\frac{\partial g_1^*(w)}{\partial W_{i,j}} = v_{c,i} \cdot u_{w,j} \implies \frac{\partial g_1^*(w)}{\partial W} = v_c \cdot u_w^T$$

$$\mathbb{R}^{M \times 1} \quad \mathbb{R}^{1 \times M}$$

Аналогично за $\frac{\partial g_2(w)}{\partial W} = - v_{\bar c_j} \cdot u_w^T$

$$= \frac{\partial g_1(w)}{\partial W}$$

$$\implies \frac{\partial J}{\partial W} = \left(\sum_{j=1}^{n} (1 - \sigma(q_{\bar c_j} - v_{\bar c_j}^T \cdot W u_w)) \cdot v_{\bar c_j} - (1 - \sigma(v_c^T W u_w - q_c)) \cdot v_c \right) u_w^T$$

2 зад.

1. Нека дефинираме $t := \bar{V} W u_w - \bar{q} \in \mathbb{R}^{n+1}$, където

$$t_1 \xi = v_c^T \cdot W \cdot u_w - q_c$$

$$\overline{i=2,n+1} : t_i = v_{\overline{c_{i-1}}}^T \cdot W \cdot u_w - q_{\overline{c_{i-1}}}$$

$\mathcal{J}$ има вида:

$$\mathcal{J} = - \sum_{i=1}^{n+1} \left[(\bar{\delta}_c)_i \log \sigma(t_i) + (1 - (\bar{\delta}_c)_i) \log \sigma(-t_i) \right]$$

$$\frac{\partial \mathcal{J}}{\partial u_w} \in \mathbb{R}^M, \text{ понеже } \mathcal{J} \in \mathbb{R} \text{ и } u_w \in \mathbb{R}^M$$

$$\frac{\partial \mathcal{J}}{\partial u_w} = \underbrace{\left(\frac{\partial t}{\partial u_w} \right)^T}_{\in \mathbb{R}^{M \times (n+1)}} \cdot \underbrace{\frac{\partial \mathcal{J}}{\partial t}}_{\in \mathbb{R}^{n+1}}$$

За I:
$$\mathcal{I} = \frac{\partial (\bar{V} W u_w - \bar{q})}{\partial u_w} =$$

$$\boxed{\mathcal{I} = \bar{V} W} \in \mathbb{R}^{(n+1) \times M}$$
$$\bar{V} \in \mathbb{R}^{(n+1) \times M} \quad W \in \mathbb{R}^{M \times M}$$

За II: $\frac{\partial \mathcal{J}}{\partial t} = \begin{pmatrix} \frac{\partial \mathcal{J}}{\partial t_1} \\ \vdots \\ \frac{\partial \mathcal{J}}{\partial t_{n+1}} \end{pmatrix}$, където ще намерим $\forall i = \overline{1, n+1} : \frac{\partial \mathcal{J}}{\partial t_i}$.

$$\frac{\partial \mathcal{J}}{\partial t_i} = \cancel{\text{------}} (\bar{\delta}_c)_i (1 - \sigma(t_i)) \cancel{\ast} (1 - (\bar{\delta}_c)_i)(1 - \sigma(-t_i))$$

стр 4

$$* \quad \frac{\partial J}{\partial t_i} = \left[(\bar{\delta_c})_i \cdot (1 - \sigma(t_i)) - (1 - \bar{\delta_c})_i \right] \left[\sigma(t_i) \right] =$$

$$= -\left\{ (\bar{\delta_c})_i - (\bar{\delta_c})_i \sigma(t_i) - \sigma(t_i) + (\bar{\delta_c})_i \sigma(t_i) \right\}$$

$$= \left[(\bar{\delta_c})_i - \sigma(t_i) \right] = \left[(\bar{\bar{\delta_c}} - \sigma(t))_i \right]$$

$$\Rightarrow \frac{\partial J}{\partial t} = -\bar{\delta_c} + \sigma(t) = I$$

от I и II $\Rightarrow$

$$\boxed{\frac{\partial J}{\partial u_w} = (\bar{V}W)^T (\bar{\delta_c} - \sigma(t)) = -W^T \cdot \bar{V}^T \cdot \left[\bar{\delta_c} - \sigma(\bar{V}Wu_w - \bar{q}) \right]}$$

$$\boxed{2.} \quad \frac{\partial J}{\partial \bar{V}} \in \mathbb{R}^{(n+1) \times M} = \left(\frac{\partial t}{\partial \bar{V}} \right)^T \cdot \frac{\partial J}{\partial t} = \left(\frac{\partial t}{\partial \bar{V}} \right)^T \cdot (-1)(\bar{\delta_c} - \sigma(\bar{V}Wu_w - \bar{q}))$$

$$\frac{\partial t}{\partial \bar{V}} \in \mathbb{R}^{(n+1) \times (n+1) \times M} \quad \frac{\partial \cancel{(\bar{V}Wu_w - \bar{q})}}{\partial \bar{V}}:$$

За да изразим лесно $\frac{\partial J}{\partial \bar{V}}$, нека изразим

$$\frac{\partial J}{\partial \bar{V}_i}, \quad t_i \in \overline{1, n+1}.$$

$$\frac{\partial J}{\partial \bar{V}} = \begin{pmatrix} \frac{\partial J}{\partial \bar{V}_1} \\ \vdots \\ \frac{\partial J}{\partial \bar{V}_{n+1}} \end{pmatrix} \cdot \left| \frac{\partial J}{\partial \bar{V}_i} = \frac{\partial J}{\partial t_i} \cdot \frac{\partial t_i}{\partial \bar{V}_i} = (-\bar{\delta_c} + \sigma(t))_i \cdot (Wu_w)^T \right.$$

$$\Rightarrow \frac{\partial J}{\partial \bar{V}} = \begin{pmatrix} (-\bar{\delta_c} + \sigma(t))_1 \\ \vdots \\ (-\bar{\delta_c} + \sigma(t))_{n+1} \end{pmatrix} \left\{ (Wu_w)^T \right. \quad \boxed{\frac{\partial J}{\partial \bar{V}} = (-\bar{\delta_c} + \sigma(t))(Wu_w)^T}$$

$$= (\sigma(\bar{V}Wu_w - \bar{q}) - \bar{\delta_c}) u_w^T W^T \quad \text{стр 5}$$

$$\frac{\partial J}{\partial \bar{V}} = \left(\sigma(\bar{V}WU_w - \bar{q}) - \bar{\delta}_c \right) U_w^T W^T$$

3.

$$\frac{\partial J}{\partial W} = \bar{V}^T \underbrace{\left(\sigma(\bar{V}WU_w - \bar{q}) - \bar{\delta}_c \right)}_{\in \mathbb{R}^{(n+1)\times 1}} \cdot \underbrace{U_w^T}_{\in \mathbb{R}^{1 \times M}}$$

$\overbrace{}^{\in \mathbb{R}^{M\times(n+1)}}$

$\in \mathbb{R}^{M \times M}$, понеже

⊛ $J \in \mathbb{R}$ и $W \in \mathbb{R}^{M \times M}$

за проверка.