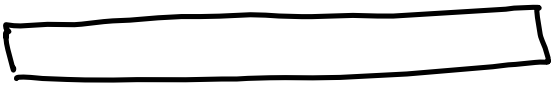
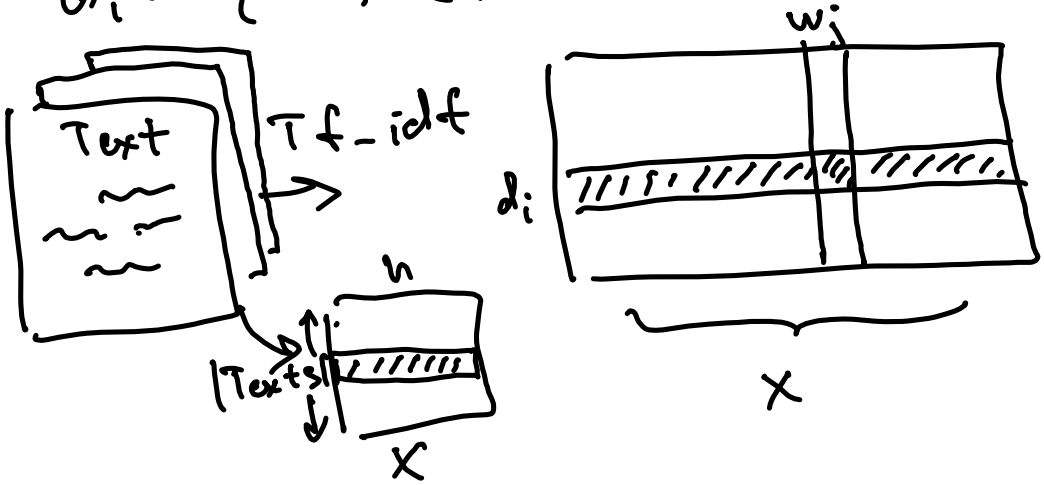
Word2Vec

$w \rightarrow$  $\in \mathbb{R}^h$

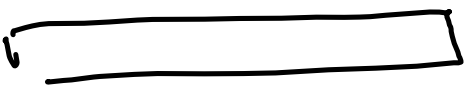
$D = \{d_1, \dots, d_{|D|}\}$

$d_i = \{w_1, w_2, \dots, w_{|d_i|}\}$



tezgiino →  v_1

wine →  v_2

coat →  v_3.

$\rho(v_1, v_2) < \rho(v_1, v_3)$

$$\mathbb{P}(w_1, w_2, \ldots, w_T) = \mathbb{P}(w_1) \cdot \mathbb{P}(w_2 | w_1)$$

$$\cdot \mathbb{P}(w_3 | w_2, w_1) \cdot \ldots \cdot \mathbb{P}(w_T | w_1, \ldots, w_{T-1})$$

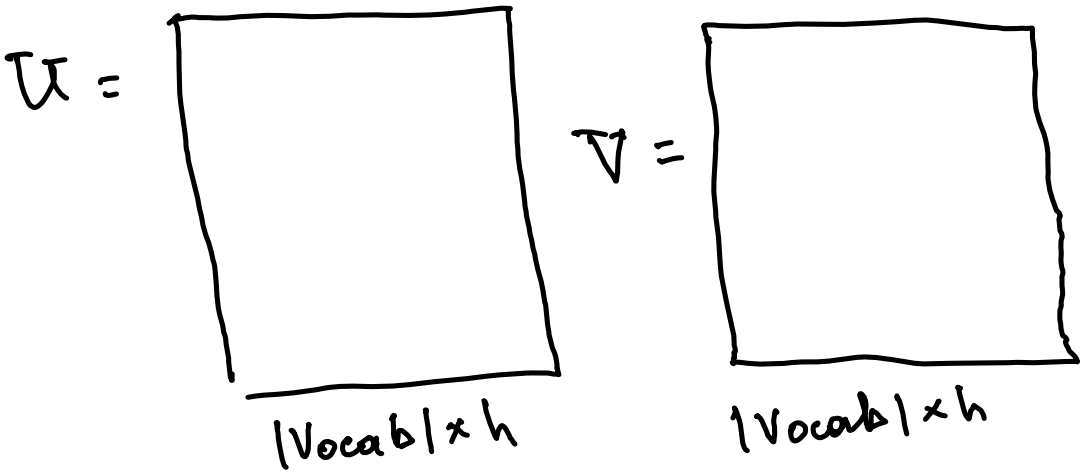
$$\simeq \mathbb{P}(w_1 | w_3) \cdot \cancel{\mathbb{P}(w_3)} \cdot \mathbb{P}(w_2 | w_3) \cdot$$

$$\cdot \mathbb{P}(w_4 | w_3) \cdot \mathbb{P}(w_5 | w_3) \cdot$$

$$L = \prod_{t=1}^{T} \prod_{\substack{-m \le j \le m \\ j \neq 0}} \mathbb{P}(w_{t+j} | w_j)$$

$$\mathbb{P}(\text{out} | \text{center}) = \frac{\exp(u_{out}^T \cdot v_{center})}{\sum_{u \in U} \exp(u^T v_{center})}$$

$$\text{softmax}: \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix} \rightarrow \begin{bmatrix} \frac{e^{a_1}}{e^{a_1} + e^{a_2} + \ldots + e^{a_n}} \\ \frac{e^{a_2}}{e^{a_1} + \ldots + e^{a_n}} \\ \vdots \end{bmatrix}$$



$$= J$$

$$-\log \mathbb{P}(\text{милый} \mid \text{кот}) =$$

$$= -\log\left(\frac{\exp(u^T_{\text{милый}} v_{\text{кот}})}{\sum_{u \in U} \exp(u^T v_{\text{кот}})}\right) =$$

$$= -u^T_{\text{милый}} v_{\text{кот}} + \log\left(\sum_{u \in U} \exp(u^T v_{\text{кот}})\right)$$

$$\nabla_{v_{\text{кот}}} J = -u_{\text{милый}} + \frac{\sum_u \exp(u^T v_{\text{кот}}) \cdot u}{\underbrace{\sum \exp(u^T v_{\text{кот}})}_{F(u)}}$$

$$J = -u_{\text{правильный}}^{T} v_{\text{центр}} + \log\left(\sum_{u \in \{u_{i_1}, u_{i_2}, \ldots, u_{i_k}\}} \exp(u^T v_{\text{центр}})\right)$$

↑

Skip-Gram NS.

(CBOW)

$$\text{Text} = w_1\ w_2\ \ldots\ w_p$$

$$\text{Vec}_{\text{Text}} = \frac{V_{w_1} \cdot idf(w_1) + V_{w_2} \cdot idf(w_2) + \ldots}{idf(w_1) + idf(w_2) + \ldots}$$
