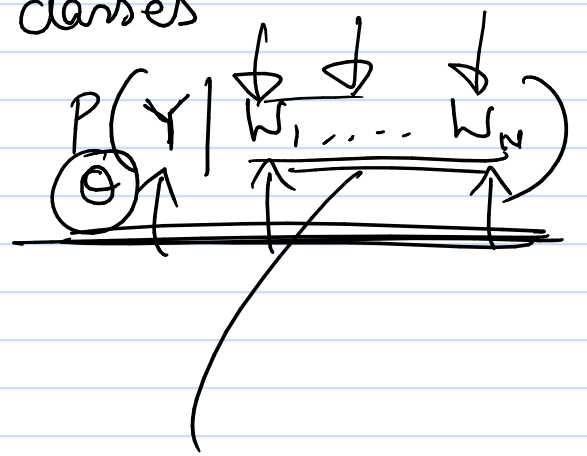


NB 2 classes



$$\propto \underbrace{P(w_1, \dots, w_N | Y)}_{\text{wavy underline}} \cdot \underbrace{P(Y)}_{\text{circled } \theta}$$

$$\prod_{i=1}^N \underbrace{P(w_i | Y)}_{\text{circled } \theta} \quad \text{circled } *$$

$[1 \ 0 \ 0 \ 0 \ 1 \ 0 \dots \ 1]$
 \vdots
 $N \times 1$

$\underbrace{\begin{matrix} 1 \\ \vdots \\ \vdots \end{matrix}}_{\text{circled } 1} \} \rightarrow \underbrace{P(w_i | Y)}_{\text{circled } \hat{P}(w_i | Y)} = \frac{\# w_i, Y \text{ spans}}{\# Y \text{ spans}}$

Generate text.

Language model.

Markov model.

$$P(\text{new word} \mid \underbrace{\text{previous words}}_{10})$$

$$P(W_{11} \mid W_{10}, W_9, \dots, W_1) = P(W_{11} \mid W_{10}, W_9, W_8)$$

$$\left. \begin{array}{l} \text{Doc}_A = w_1 \dots w_N \\ \vdots \\ \vdots \\ \vdots \end{array} \right\} = \text{Data.}$$

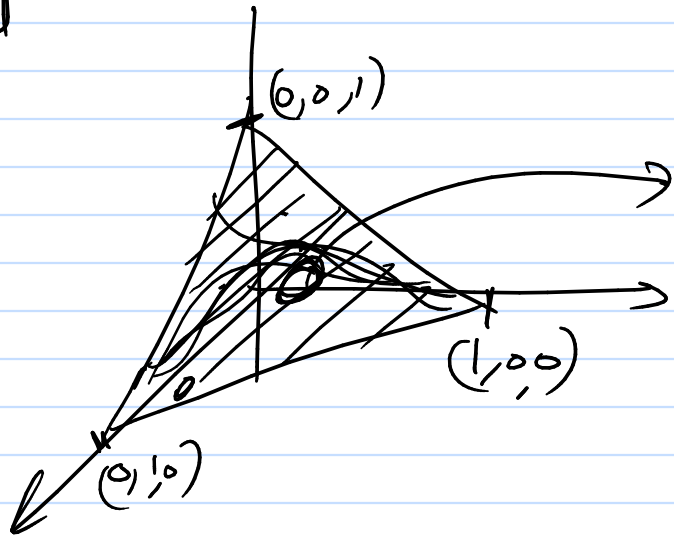
for every 3 word tuple n grams

$$\rightarrow \hat{P}(\text{next word} = w \mid \underline{\text{tuple}})$$

for every w in my vocab.

$$\rightarrow \frac{\#(w, \text{tuple})}{\#(\text{tuple})}$$

3d



$$p_1 + p_2 + p_3 = 1$$

$[p_1, p_2, p_3]$ $\begin{matrix} \text{+ve} \\ \text{3x1} \end{matrix}$

$\Theta \sim \text{Dir}(\alpha_1, \alpha_2, \alpha_3)$

~~α~~

$\theta_1 + \theta_2 + \theta_3 = 1$
 $\theta_1 \geq 0$ $\theta_2 \geq 0$ $\theta_3 \geq 0$

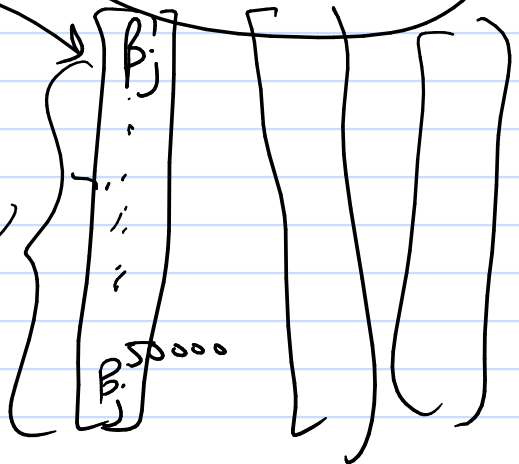
Z_i

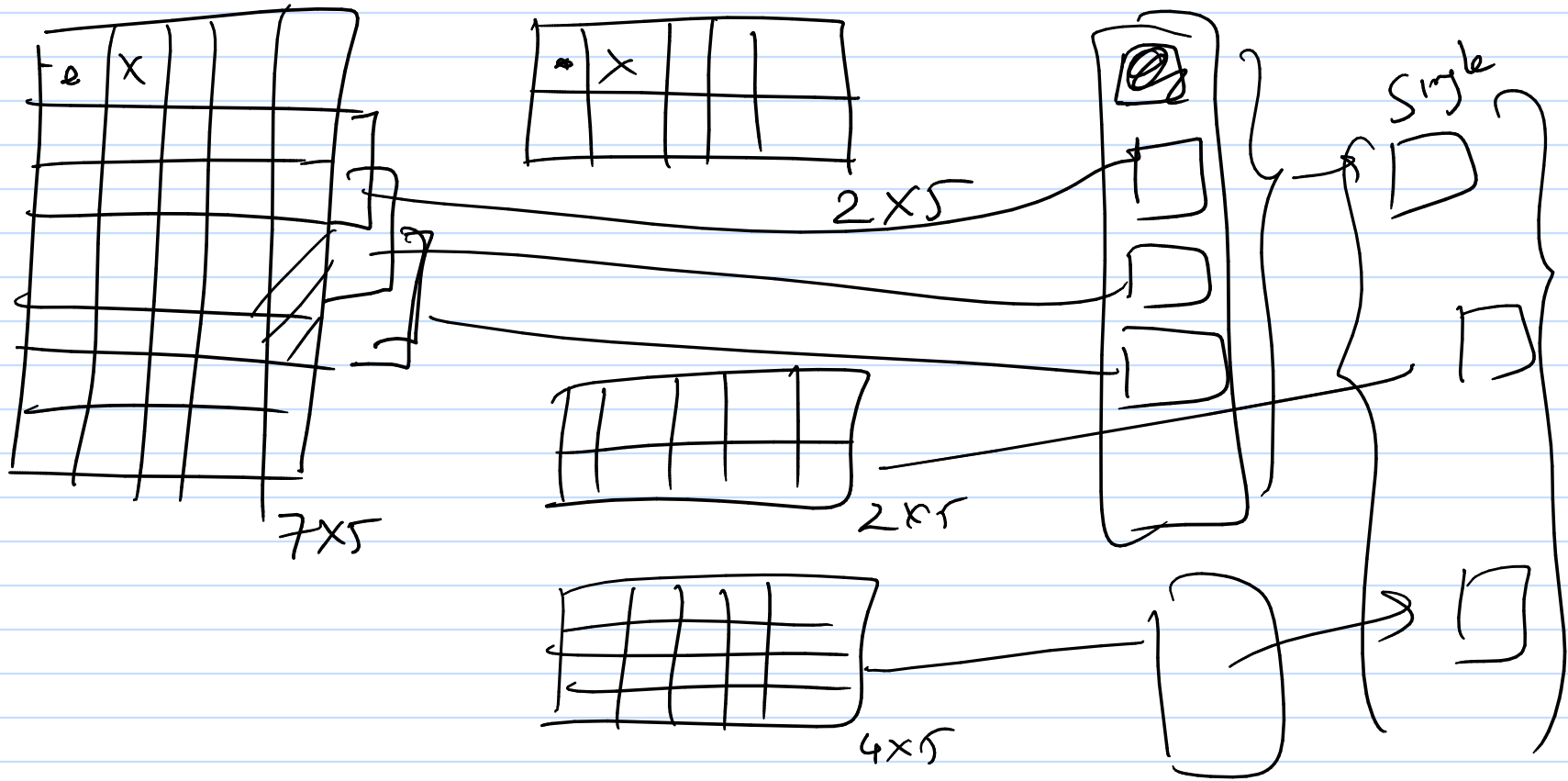


Boots Boots Maiz finance fm Mu Ban M

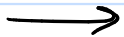
finance
Maiz
Boots

β_j $j=1 \dots T=2$



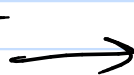
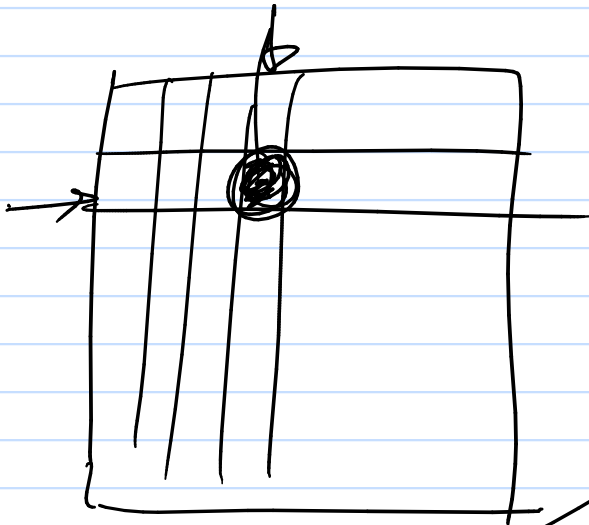


Corpus



Vocab = 100

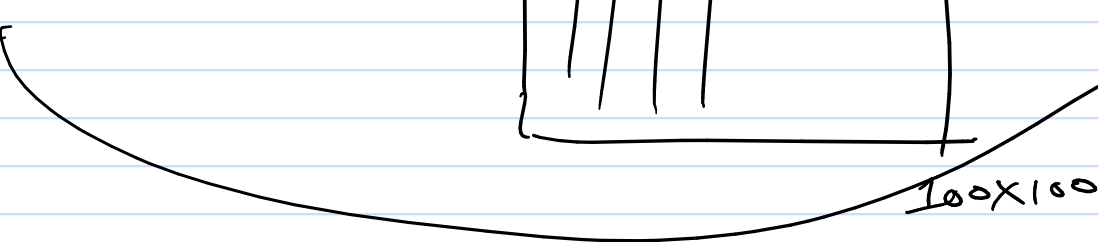
① Count based



PCA

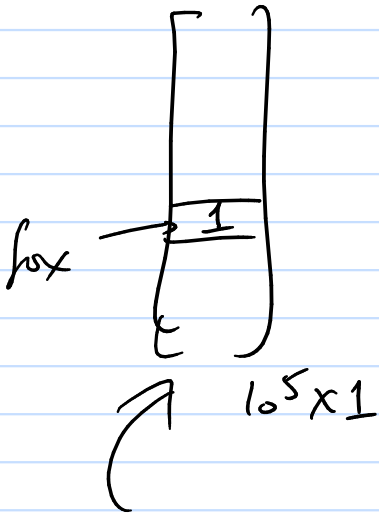


$W =$

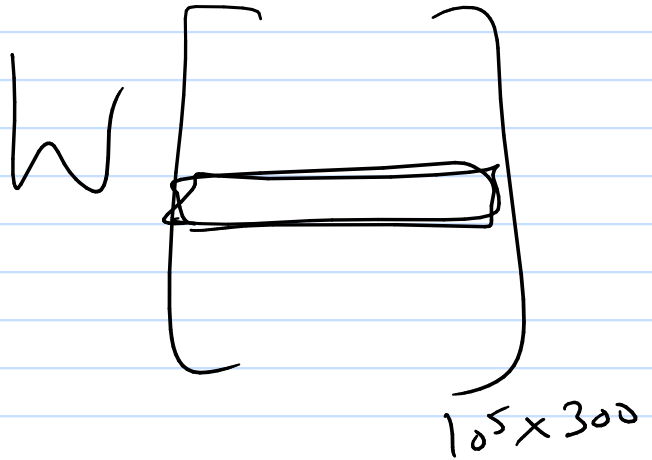


100x100

100xd



v



$v^T W$

