MLE for the One-Hot Categorical

$$\underline{\theta}^* = \underset{\substack{\theta \in [0,1]^D \\ \sum_{d=0}^{D-1} \theta_d = 1}}{\operatorname{argmax}} \left(\ell(\mathcal{D}; \underline{\theta}) \right)$$

$$\Rightarrow \theta_e^* = \frac{N_e}{N}$$

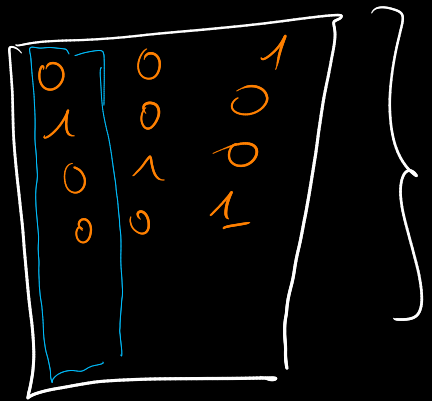
Dataset for the one-Hot

$$\mathcal{D} = \{ [0,0,1], [0,1,0], [0,0,1], [1,0,0], \vdots \}$$

?

View the dataset as a matrix

$\underline{D} =$



# of samples (N)

# of states (D)

reduce matrix to vector over 0th axis

$$\hookrightarrow \underline{N} = [N_0, N_1, N_2, \dots]$$

$$\underline{\theta} = \left[\frac{N_0}{N}, \frac{N_1}{N}, \frac{N_2}{N}, \dots \right]$$