$$\frac{\text{Def.}}{\text{The Standard (unit) softmax function is}}$$

$$T : \mathbb{R}^{n} \longrightarrow (0,1)^{n}$$
where $T_{i}(z) = \frac{e^{z_{i}}}{\sum_{i} e^{z_{k}}}$.

Note. $Z = (z_1,...,z_n) \in \mathbb{R}^n$, and $T_i = \mathbb{R}^n \longrightarrow \mathbb{R}$ is just the ith component function (where i = 1,...,n).

Rmk. The <u>Jacobian matrix</u> of of is just the n×n matrix of partial derivatives:

$$\frac{5^{\prime}}{9\alpha^{\prime}} \qquad \frac{5^{\prime}}{9\alpha^{\prime}}$$

$$\vdots \qquad \vdots \qquad \vdots \\
\frac{5^{\prime}}{9\alpha^{\prime}} \qquad \frac{5^{\prime}}{9\alpha^{\prime}}$$

So, we just have to find $\frac{\partial U_i}{\partial z_j}$ for each $i, j \in \{1, ..., n\}$.

We have to use the quotient rule!

Quotient Rule. For a function $9: \mathbb{R} \longrightarrow \mathbb{R}$ defined by $9(x) = \frac{f(x)}{g(x)}, 9'(x) = \frac{g(x) \cdot f'(x) - f(x) \cdot g'(x)}{[g(x)]^2}$.

In this situation,
$$f(z_i) = e^{z_i}$$
 and $g(z_i) = \sum_{k=1}^{n} e^{z_k}$. Note, $f'(z_i) = e^{z_i} = g'(z_i)$. So,
$$\frac{\partial \mathcal{T}_i}{z_i} = \frac{\left[\sum_{k=1}^{n} e^{z_k}\right] \cdot e^{z_i} - e^{z_i} \cdot e^{z_i}}{\left[\sum_{k=1}^{n} e^{z_k}\right]^2}$$

$$= \left(\frac{e^{\frac{2}{\epsilon}}}{\sum_{\kappa=1}^{n} e^{\frac{2}{\kappa}}}\right) \cdot \left(\frac{\left[\sum_{\kappa=1}^{n} e^{\frac{2}{\kappa}}\right] - e^{\frac{2}{\epsilon}}}{\sum_{\kappa=1}^{n} e^{\frac{2}{\kappa}}}\right)$$

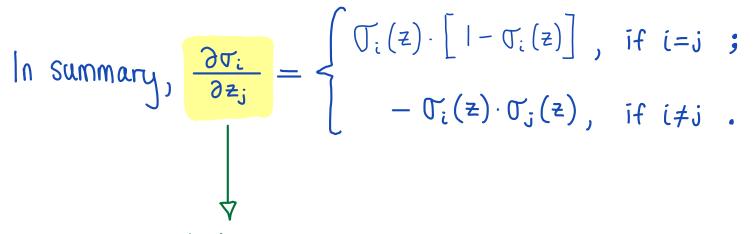
Case #2: $i \neq j$.

In this situation, $f(z_j) = e^{z_i}$ and $g(z_j) = \sum_{k=1}^{n} e^{z_k}$. Note, $f'(z_j) = 0$ and $g'(z_j) = e^{z_j}$. So,

$$\frac{\partial \mathcal{T}_{i}}{\mathcal{Z}_{j}} = \frac{\left[\sum\limits_{\kappa=1}^{n} e^{\mathbf{z}_{\kappa}}\right] \cdot 0 - e^{\mathbf{z}_{i}} \cdot e^{\mathbf{z}_{j}}}{\left[\sum\limits_{\kappa=1}^{n} e^{\mathbf{z}_{\kappa}}\right]^{2}}$$

$$= -\left(\frac{e^{\frac{2}{2}i}}{\sum_{\kappa=1}^{n}e^{\frac{2}{2}\kappa}}\right) \cdot \left(\frac{e^{\frac{2}{2}i}}{\sum_{\kappa=1}^{n}e^{\frac{2}{2}\kappa}}\right)$$

$$= - \mathcal{O}_{i}(z) \cdot \mathcal{O}_{j}(z)$$



The (i,i)th entry
of the Jacobian matrix
of T.