The Derivative of the Softmax Function

Non-Mandatory Homework

Consider the Softmax function:

$$\sigma(z_j) = \frac{\exp z_j}{\sum_{k=1}^{m} \exp z_k}$$

Prove that the derivative of $\sigma(z_j)$ w.r.t (with respect to) z_v can be formulated as:

$$\frac{\partial \sigma(z_j)}{\partial z_v} = \begin{cases} \sigma(z_j)\left(1 - \sigma(z_j)\right) & j = v \\ -\sigma(z_j)\sigma(z_v) & j \neq v \end{cases}$$