

Softmax Operator

Rylan W. Yancey

10/1/2023

Definition & Gradient

The Softmax Function is defined as follows:

$$\sigma(\vec{z})_i = \frac{e^{z_i}}{\sum_{j=0}^K e^{z_j}}$$

Z is some vector of length K.

Not sure how to write this gracefully so for now heres some links:

<https://www.mldawn.com/the-derivative-of-softmaxz-function-w-r-t-z/>

<https://eli.thegreenplace.net/2016/the-softmax-function-and-its-derivative>