

Softmax の 不安定性

$$\hat{y} = \text{softmax}(x)$$

$$= \left(\frac{e^{x_1}}{\sum_{i=1}^n e^{x_i}}, \dots \right)$$

e^{x_i} sometimes became very large value beyond normal range

$$\hat{y} = \left(\frac{e^{x_1} / e^{\alpha}}{\sum_{i=1}^n e^{x_i} / e^{\alpha}}, \dots \right)$$

by finding largest x_i then divide

numerator and denominator by e^{α}

it can be stable computation of softmax