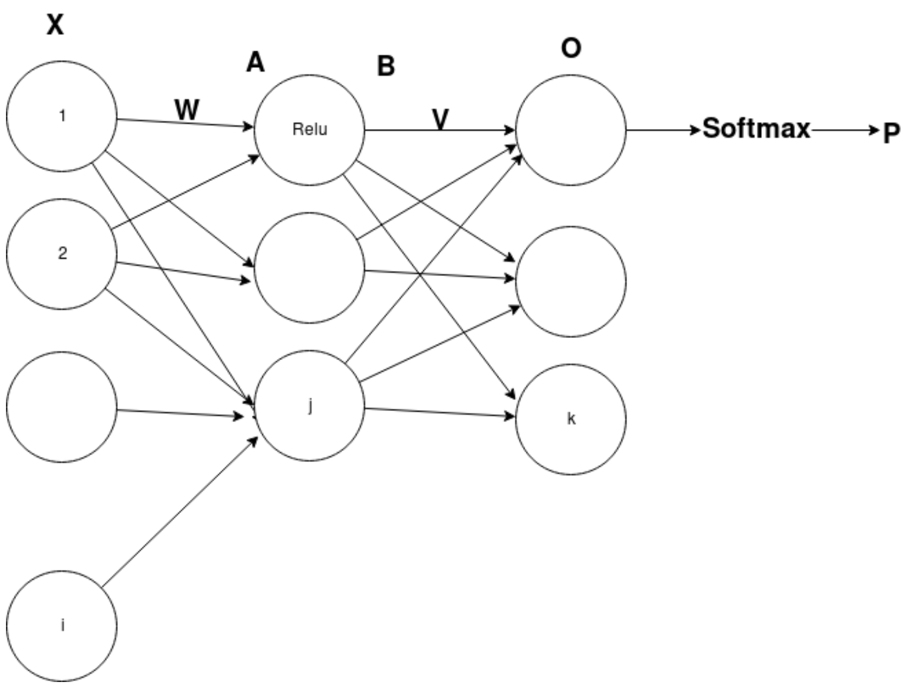
eq

February 2019

1 Introduction



$$J = \log Pc,\ c = \text{true label}$$

$$Pc = \frac{e^{O_c}}{\sum\limits_k e^{O_k}}$$

$$\frac{\partial J}{\partial V_{kj}} = \frac{\partial J}{\partial O_k}\frac{\partial O_k}{\partial V_{kj}}$$

$$\frac{\partial O_k}{\partial V_{kj}} = B_j$$

$$O_k = \sum_j V_{kj} Bj$$

$$\frac{\partial J}{\partial O_k} = \frac{\partial J}{\partial P_c} \frac{\partial P_c}{\partial O_k}$$

$$\frac{\partial J}{\partial P_c} = \frac{1}{P_c}$$

$$det(S) = \sum_k e^{O_k}$$

$$\frac{\partial P_c}{\partial O_k} = \begin{cases} \frac{-e^{O_c} e^{O_k}}{s^2}, & \text{if } c \neq k \\ \frac{e^{O_c} S - e^{O_c} e^{O_k}}{s^2} = \frac{e^{O_c}(S - e^{O_k})}{S^2}, & \text{if } c = k \end{cases}$$

$$\frac{\partial J}{\partial V_{kj}} = \frac{1}{P_c} B_j = B_j . \frac{S}{e^{O_c}} = \begin{cases} \frac{-B_j e^{O_k}}{S} = -B_j P_k, & \text{if } c \neq k \\ \frac{B_j (S - e^O_c}{S} = B_j (1 - P_k), & \text{if } c = k \end{cases}$$

$$= B_j [\mathbb{I}(k = c) - P_k]$$

$$= Bj[y_k - P_k]$$

$$\frac{\partial J}{\partial W_{ji}} = \frac{\partial J}{\partial B_j} . \frac{\partial B_j}{\partial A_j} . \frac{\partial A_j}{\partial W_{ji}}$$

$$\frac{\partial A}{\partial W_{ji}} = x_i \tag{1}$$

$$\frac{\partial B_j}{\partial A_j} = \text{Relu Activation} \tag{2}$$

$$= \begin{cases} 1 & \text{if } A_j > 0 \\ 0 & \text{else} \end{cases}$$

$$\frac{\partial J}{\partial B_j} = \sum_k \frac{\partial J}{\partial O_k} . \frac{\partial O_k}{\partial B_j} \tag{3}$$

$$= \sum_k \frac{\partial J}{\partial P_c} . \frac{\partial P_c}{\partial O_k} . \frac{\partial O_k}{\partial B_j}$$

$$= \frac{1}{P_c} \sum_k \frac{\partial P_c}{\partial O_k} . \frac{\partial O_k}{\partial B_j}$$

$$= \frac{1}{P_c} \sum_k P_c [y_k - P_k] V_{kj}$$

$$\sum_k (y_k - P_k) V_{kj}$$

Combining (1), (2), (3)

$$\frac{\partial J}{\partial W_{ji}} = \begin{cases} 0 & \text{if } A_j < 0 \\ \sum_k x_i.(y_k - P_k)V_{kj} & A_j > 0 \end{cases}$$