

Chapter 6 . Ex 21

解: a) $\text{softmax}(x) = \frac{e^{x_i}}{\sum_j e^{x_j}} \stackrel{\text{def}}{=} z(x)$

则 $\frac{\partial z}{\partial x} = z \cdot (1 - z)$

已知三层神经网络向前计算为

$$a_j = \sum_{i=1}^d w_{ji} x_i$$

$$y_j = f(a_j)$$

$$o_k = \sum_{j=1}^c w_{kj} y_j$$

$$z_k = \frac{e^{o_k}}{\sum_{m=1}^c e^{o_m}}$$

目标函数为 $J = \frac{1}{2} \sum_{k=1}^c (t_k - z_k)^2$

现在要求 $\frac{\partial J}{\partial w_{kj}}$ 与 $\frac{\partial J}{\partial w_{ji}}$

$$\frac{\partial J}{\partial w_{kj}} = \frac{\partial J}{\partial o_k} \cdot \frac{\partial o_k}{\partial w_{kj}} = \left(\sum_{s=1}^c \frac{\partial J}{\partial z_s} \cdot \frac{\partial z_s}{\partial o_k} \right) \cdot \frac{\partial o_k}{\partial w_{kj}}$$

$$\frac{\partial J}{\partial \theta_k} = \begin{cases} \frac{e^{\theta_s} \cdot (-1) \cdot e^{\theta_k}}{\left(\sum_{m=1}^C e^{\theta_m}\right)^2} = -z_s z_k, & s \neq k \\ \frac{e^{\theta_s}}{\left(\sum_{m=1}^C e^{\theta_m}\right)} + (-1) \frac{e^{\theta_s} \cdot e^{\theta_s}}{\left(\sum_{m=1}^C e^{\theta_m}\right)^2} = z_k - z_k^2, & s = k \end{cases}$$

$$\therefore \frac{\partial J}{\partial z_s} = (-1) (t_s - z_s)$$

$$\therefore \frac{\partial J}{\partial \theta_k} = \sum_{s \neq k} (-1) (t_s - z_s) (-z_s z_k) + (-1) (t_k - z_k) (z_k - z_k^2)$$

$$\text{or } \frac{\partial \theta_k}{\partial w_{kj}} = y_j$$

$$\therefore \frac{\partial J}{\partial w_{kj}} = y_j \sum_{s \neq k} (t_s - z_s) (z_s z_k) - y_j (t_k - z_k) (z_k - z_k^2)$$

$$\text{or } \frac{\partial J}{\partial w_{ji}} = \frac{\partial J}{\partial y_i} \cdot \frac{\partial y_i}{\partial a_j} \cdot \frac{\partial a_j}{\partial w_{ji}}$$

$$\text{or } \frac{\partial J}{\partial y_i} = \sum_{s=1}^C \frac{\partial J}{\partial z_s} \cdot \frac{\partial z_s}{\partial y_i}$$

$$= - \sum_{s=1}^C (t_s - z_s) \frac{\partial z_s}{\partial y_i}$$

$$= - \sum_{s=1}^C (t_s - z_s) \left[\frac{\partial z_s}{\partial \theta_s} \frac{\partial \theta_s}{\partial y_i} + \sum_{r \neq s} \frac{\partial z_s}{\partial \theta_r} \frac{\partial \theta_r}{\partial y_i} \right]$$

$$= -\sum_{s=1}^c (t_s - z_s) (z_s - z_s^2) w_{sj}$$

$$+ \sum_{s=1}^c \sum_{r \neq s} z_s z_r w_{rj} (t_s - z_s)$$

又易知 $\frac{\partial a_j}{\partial w_{ji}} = p_i$

$$\therefore \frac{\partial J}{\partial w_{ji}} = -p_i f'(a_j) \sum_{s=1}^c (t_s - z_s) w_{sj} (z_s - z_s^2)$$

综上所述学习规则为：

$$\Delta w_{ji} = -\eta \frac{\partial J}{\partial w_{ji}}$$

$$\Delta w_{kj} = -\eta \frac{\partial J}{\partial w_{kj}}$$

其中 η 为学习率