

深度学习数学推导

$L(a_i) = (Y_i \cdot \log a_i + (1 - Y_i) \log 1 - a_i)$
 $\frac{\partial L}{\partial w_2} = \frac{\partial L}{\partial a} \cdot \frac{\partial a}{\partial z} \cdot \frac{\partial z}{\partial w_2}$ Y : 真值
 3×1 $= (a - Y) \cdot \frac{\partial z}{\partial w_2}$
 $= (a - Y) \cdot y$
 $A = [G(a_1) \ G(a_2) \ \dots \ G(a_m)]^T$ $m \times 1$
 $Y_{true} = [Y_1 \ Y_2 \ \dots \ Y_m]^T$ $m \times 1$
 $\frac{\partial L}{\partial w_1} = \frac{\partial L}{\partial a} \cdot \frac{\partial a}{\partial z} \cdot \frac{\partial z}{\partial w_1}$ $m \times 3$
 $(A - Y_{true}) \cdot y$ $1/m \times m$
 $\begin{pmatrix} G(a_1) - Y_1 \\ G(a_2) - Y_2 \\ \vdots \\ G(a_m) - Y_m \end{pmatrix}$ $\begin{pmatrix} y_1 & y_2 & y_3 \\ y_2 & y_2 & y_3 \\ y_3 & y_2 & y_3 \end{pmatrix}$
 $y^T \cdot (A - Y_{true}) = \begin{pmatrix} y_1 & y_2 & \dots & y_m \\ y_2 & y_2 & \dots & y_m \\ y_3 & y_2 & \dots & y_m \end{pmatrix} \cdot \begin{pmatrix} G(a_1) - Y_1 \\ G(a_2) - Y_2 \\ \vdots \\ G(a_m) - Y_m \end{pmatrix}$ $m \times 3$
 $3 \times m$ $m \times 1$

深度学习逻辑回归反向传播数学推导 1

$L(a^i) = (y^i \cdot \log a^i + (1 - y^i) \cdot \log 1 - a^i)$
 $\frac{\partial L}{\partial w_2} = \frac{\partial L}{\partial a} \cdot \frac{\partial a}{\partial z} \cdot \frac{\partial z}{\partial w_2}$
 $\frac{\partial L}{\partial w_2} = (a - y) \cdot \frac{\partial z}{\partial w_2}$
 $\frac{\partial L}{\partial w_2} = (a - y) \cdot y$

$A = \begin{bmatrix} \sigma(a_1) & \sigma(a_2) & \dots & \sigma(a_m) \end{bmatrix}^T$ $m \times 1$
 $Y_{\text{true}} = \begin{bmatrix} y_1 & y_2 & \dots & y_m \end{bmatrix}^T$ $m \times 1$
 $(A - Y_{\text{true}}) \cdot y$ $m \times 1$
 $\begin{pmatrix} \sigma(a_1) - y_1 \\ \sigma(a_2) - y_2 \\ \vdots \\ \sigma(a_m) - y_m \end{pmatrix} \begin{pmatrix} y_1 & y_2 & y_3 \\ y_1^2 & y_2^2 & y_3^2 \\ \vdots & \vdots & \vdots \\ y_1^m & y_2^m & y_3^m \end{pmatrix}$

$\frac{\partial L}{\partial w_1} = \frac{\partial L}{\partial a} \cdot \frac{\partial a}{\partial z} \cdot \frac{\partial z}{\partial w_1} \cdot \frac{\partial y}{\partial w_1}$
 $\frac{\partial L}{\partial w_1} = (a - y) \cdot w_2^T \cdot X$
 $\frac{\partial L}{\partial w_1} = X^T \cdot (a - y) \cdot w_2^T$
 $\frac{\partial L}{\partial w_1} = \frac{1}{m} \sum_{i=1}^m (y^i \log a^i + (1 - y^i) \cdot \log 1 - a^i)$

$X = \begin{bmatrix} x_1^1 & x_1^2 & \dots & x_1^m \\ x_2^1 & x_2^2 & \dots & x_2^m \\ \vdots & \vdots & \ddots & \vdots \\ x_m^1 & x_m^2 & \dots & x_m^m \end{bmatrix}$ $m \times m$
 $X^T = \begin{bmatrix} x_1^1 & x_2^1 & \dots & x_m^1 \\ x_1^2 & x_2^2 & \dots & x_m^2 \\ \vdots & \vdots & \ddots & \vdots \\ x_1^m & x_2^m & \dots & x_m^m \end{bmatrix}$ $m \times m$
 $\frac{\partial L}{\partial w_2} = \frac{1}{m} \sum_{i=1}^m (y^i \log a^i + (1 - y^i) \cdot \log 1 - a^i)$

$y^T \cdot (A - Y_{\text{true}}) = \begin{pmatrix} y_1^1 & y_1^2 & \dots & y_1^m \\ y_2^1 & y_2^2 & \dots & y_2^m \\ \vdots & \vdots & \ddots & \vdots \\ y_m^1 & y_m^2 & \dots & y_m^m \end{pmatrix} \cdot \begin{pmatrix} \sigma(a_1) - y_1 \\ \sigma(a_2) - y_2 \\ \vdots \\ \sigma(a_m) - y_m \end{pmatrix}$

$\frac{\partial L}{\partial w_2} = \frac{1}{m} \sum_{i=1}^m (y^i \log a^i + (1 - y^i) \cdot \log 1 - a^i)$

深度学习逻辑回归反向传播数学推导 2

softmax 反向传播

第 i 个样本 第 l 层输出

这里 l 是最后层

num_class = k, sample_count = m, l = layers_num

softmax $(z^{(l)}) = P(y_i = j | w^{(l)}, b^{(l)}) (y_i = 0, 1, \dots, k)$

$L(P) = -\frac{1}{m} \sum_{i=1}^m \left(\sum_{j=1}^k \log P(y_i = j | z_{ij}^{(l)}, y_{ij}^{(l)}) \right)$

$= -\frac{1}{m} \sum_{i=1}^m \left(\sum_{j=1}^k y_{ij}^{(l)} \log \frac{e^{w_{ij}^{(l)} \cdot x_i^{(l)} + b_j}}{\sum_{t=1}^k e^{w_{it}^{(l)} \cdot x_i^{(l)} + b_t}} \right)$

$= -\frac{1}{m} \sum_{i=1}^m \left(y_{ij}^{(l)} \cdot \log \frac{e^{w_{ij}^{(l)} \cdot x_i^{(l)} + b_j}}{\sum_{t=1}^k e^{w_{it}^{(l)} \cdot x_i^{(l)} + b_t}} + \sum_{t \neq j} y_{it}^{(l)} \cdot \log \frac{e^{w_{it}^{(l)} \cdot x_i^{(l)} + b_t}}{\sum_{t=1}^k e^{w_{it}^{(l)} \cdot x_i^{(l)} + b_t}} \right)$

设 $w_{ij}^{(l)} \cdot x_i^{(l)} + b_j = z_{ij}^{(l)}$
($j = 1, 2, 3, \dots, k$)

$= -\frac{1}{m} \sum_{i=1}^m \left[y_{ij}^{(l)} \cdot (z_{ij}^{(l)} - \log \sum_{t=1}^k e^{z_{it}^{(l)}}) + \sum_{t \neq j} y_{it}^{(l)} \cdot (z_{it}^{(l)} - \log \sum_{t=1}^k e^{z_{it}^{(l)}}) \right]$

$\frac{\partial L}{\partial z_j} = -\frac{1}{m} \sum_{i=1}^m \left[y_{ij}^{(l)} \cdot \left(1 - \frac{e^{z_{ij}^{(l)}}}{\sum_{t=1}^k e^{z_{it}^{(l)}}} \right) + \sum_{t \neq j} y_{it}^{(l)} \cdot \left(-\frac{e^{z_{it}^{(l)}}}{\sum_{t=1}^k e^{z_{it}^{(l)}}} \right) \right]$

$= -\frac{1}{m} \sum_{i=1}^m \left[y_{ij}^{(l)} (1 - \hat{y}_{ij}^{(l)}) - \sum_{t \neq j} y_{it}^{(l)} \cdot \hat{y}_{it}^{(l)} \right]$

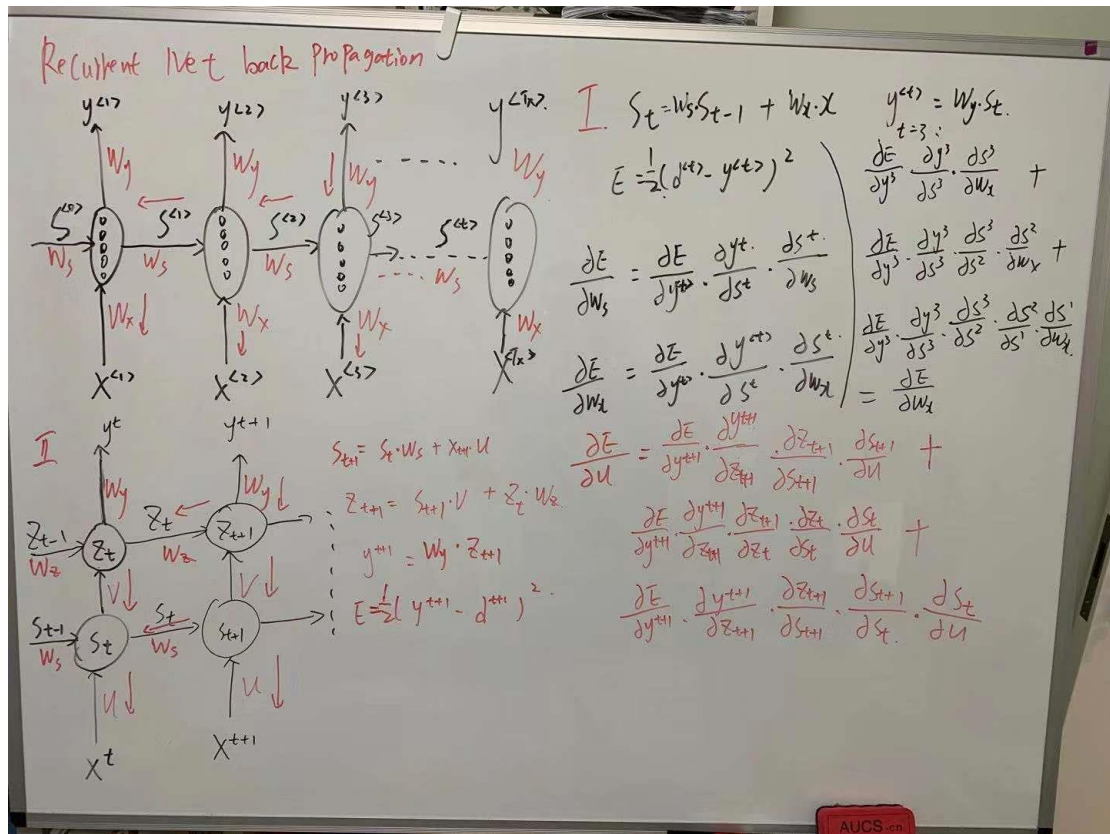
$\frac{\partial L}{\partial w_{ij}^{(l)}} = \frac{\partial L}{\partial z_j} \cdot \frac{\partial z_j}{\partial w_{ij}^{(l)}} \cdot x_i^{(l)}$

$\frac{\partial L}{\partial b_j} = \sum \left(\frac{\partial L}{\partial z_j} \right) \cdot 1 / m$

$= -\frac{1}{m} \sum_{i=1}^m \left[y_{ij}^{(l)} - y_{ij}^{(l)} \cdot \hat{y}_{ij}^{(l)} - \sum_{t \neq j} y_{it}^{(l)} \cdot \hat{y}_{it}^{(l)} \right]$

$= -\frac{1}{m} \sum_{i=1}^m \left(y_{ij}^{(l)} - \hat{y}_{ij}^{(l)} \right)$

Softmax 反向传播推导



循环神经网络反向传播