汕頭大學
SHANTOU UNIVERSITY

$y_0 = f(x+b_0)$

$y_{pre} = g(y_0 \cdot w_1 + b_1)$

$$J = \sum_{i=1}^{n} (y_{pre,i} - y_i)^2 .$$

$g_i(t_1, t_2, \dots, t_n)$

$f_i(h_1, h_2, \dots, h_m)$

w_1的第k列

$t_k = y_0 \cdot w_{1,k} + b_{1,k}$

$$\frac{\partial t_k}{\partial b_{1,p}} = \begin{cases} 1 & k=p \\ 0 & k \neq p \end{cases}$$

$$\frac{\partial g}{\partial t_p} = \begin{pmatrix} \frac{\partial g_1}{\partial t_p} \\ \frac{\partial g_2}{\partial t_p} \\ \vdots \\ \frac{\partial g_n}{\partial t_p} \end{pmatrix}$$

①
$$\frac{\partial J}{\partial b_{1,p}} = \sum_{i=1}^{n} 2(y_{pre,i} - y_i) \cdot \frac{\partial g_i}{\partial b_{1,p}}$$
$$= \sum_{i=1}^{n} 2(y_{pre,i} - y_i) \cdot \sum_{k=1}^{n} \frac{\partial g_i}{\partial t_k} \cdot \frac{\partial t_k}{\partial b_{1,p}}$$
$$= \sum_{i=1}^{n} 2(y_{pre,i} - y_i) \cdot \frac{\partial g_i}{\partial t_p}$$
$$= 2(y_{pre} - y) \cdot \frac{\partial g}{\partial t_p}$$

$$\Rightarrow \frac{\partial J}{\partial b_1} = 2(y_{pre} - y) \left(\frac{\partial g}{\partial t_1} \; \frac{\partial g}{\partial t_2} \; \cdots \; \frac{\partial g}{\partial t_n} \right)$$
$$= 2(y_{pre} - y) \begin{pmatrix} \frac{\partial g_1}{\partial t_1} & \frac{\partial g_1}{\partial t_2} & \cdots & \frac{\partial g_1}{\partial t_n} \\ \frac{\partial g_2}{\partial t_1} & \frac{\partial g_2}{\partial t_2} & \cdots & \frac{\partial g_2}{\partial t_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial g_n}{\partial t_1} & \frac{\partial g_n}{\partial t_2} & \cdots & \frac{\partial g_n}{\partial t_n} \end{pmatrix} = 2(y_{pre} - y) \cdot \frac{\partial g}{\partial t} .$$

汕頭大學

SHANTOU UNIVERSITY

② $$\frac{\partial J}{\partial w_{1,p,q}} = 2\sum_{i=1}^{n}(y_{pre,i}-y_i)\cdot\frac{\partial g_i}{\partial w_{1,p,q}}$$

$$= 2\sum_{i=1}^{n}(y_{pre,i}-y_i)\cdot\sum_{k=1}^{n}\frac{\partial g_i}{\partial t_k}\cdot\frac{\partial t_k}{\partial w_{1,p,q}}$$

$$= 2\sum_{i=1}^{n}(y_{pre,i}-y_i)\cdot\sum_{k=1}^{n}\frac{\partial g_i}{\partial t_k}\cdot\sum_{z=1}^{m}y_{0,z}\frac{\partial w_{1,z,k}}{\partial w_{1,p,q}}$$

$$= 2\sum_{i=1}^{n}(y_{pre,i}-y_i)\cdot\frac{\partial g_i}{\partial t_q}\cdot y_{0,p}$$

$$= 2\,(y_{pre}-y)\cdot\begin{pmatrix}\frac{\partial g_1}{\partial t_q}\cdot y_{0,p}\\ \frac{\partial g_2}{\partial t_q}\cdot y_{0,p}\\ \vdots\\ \frac{\partial g_n}{\partial t_q}\cdot y_{0,p}\end{pmatrix}$$

$$= 2\,y_{0,p}\cdot(y_{pre}-y)\cdot\frac{\partial g}{\partial t_q}$$

w_1的第k列

$$t_k = y_0\cdot w_{1,k} + b_{1,k}$$

$$\frac{\partial t_k}{\partial w_{1,p,q}} = \frac{\partial(y_0\cdot w_{1,k}+b_{1,k})}{\partial w_{1,p,q}}$$

$$= y_0\cdot\frac{\partial w_{1,k}}{\partial w_{1,p,q}}$$

$$= \sum_{z=1}^{m}y_{0,z}\cdot\frac{\partial w_{1,z,k}}{\partial w_{1,p,q}}$$

$$\frac{\partial w_{1,z,k}}{\partial w_{1,p,q}} = \begin{cases}1, & z=p\text{且}k=q\\ 0, & \text{否则}\end{cases}$$

$$\frac{\partial J}{\partial w_1} = \begin{pmatrix}\frac{\partial J}{\partial w_{1,1,1}} & \frac{\partial J}{\partial w_{1,1,2}} & \cdots & \frac{\partial J}{\partial w_{1,1,n}}\\ \vdots & & & \\ \frac{\partial J}{\partial w_{1,n,1}} & \frac{\partial J}{\partial w_{1,n,2}} & \cdots & \frac{\partial J}{\partial w_{1,n,n}}\end{pmatrix} = 2\begin{pmatrix}y_{0,1}(y_{pre}-y)\frac{\partial g}{\partial t_1} & \cdots & y_{0,1}(y_{pre}-y)\frac{\partial g}{\partial t_n}\\ \vdots & & \\ y_{0,n}(y_{pre}-y)\frac{\partial g}{\partial t_1} & \cdots & y_{0,n}(y_{pre}-y)\frac{\partial g}{\partial t_n}\end{pmatrix}$$

$$= 2\begin{bmatrix}y_{01}\cdot(y_{pre}-y)\\ y_{02}\cdot(y_{pre}-y)\\ \vdots\\ y_{0n}(y_{pre}-y)\end{bmatrix}\begin{bmatrix}\frac{\partial g}{\partial t_1} & \frac{\partial g}{\partial t_2} & \cdots & \frac{\partial g}{\partial t_n}\end{bmatrix}$$

$$= 2\,y_0^T(y_{pre}-y)\cdot\frac{\partial g}{\partial t}$$

$$\frac{\partial g_i}{\partial t} = \left(\frac{\partial g_i}{\partial t_1}\ \frac{\partial g_1}{\partial t_2} \cdots \frac{\partial g_i}{\partial t_n}\right) \qquad w_1^T = \begin{pmatrix} w_{1,1} \\ w_{1,2} \\ \vdots \\ w_{1,n}^T \end{pmatrix}$$

汕頭大學
SHANTOU UNIVERSITY

③
$$\frac{\partial J}{\partial b_{0,p}} = \sum_{i=1}^{n} 2\,(y_{pre,i} - y_i)\cdot \frac{\partial g_i}{\partial b_{0,p}}$$

$$= \sum_{i=1}^{n} 2\,(y_{pre,i} - y_i)\cdot \sum_{k=1}^{n} \frac{\partial g_i}{\partial t_k}\cdot \frac{\partial t_k}{\partial b_{0,p}}$$

$$= 2\sum_{i=1}^{n} (y_{pre,i} - y_i)\cdot \sum_{k=1}^{n} \frac{\partial g_i}{\partial t_k}\cdot \frac{\partial y_0}{\partial b_{0,p}}\cdot w_{1,k}$$

$$= 2\sum_{i=1}^{n} (y_{pre,i} - y_i)\cdot \sum_{k=1}^{n} \frac{\partial g_i}{\partial t_k}\cdot \sum_{r=1}^{m} \frac{\partial y_{0,r}}{\partial b_{0,p}}\cdot w_{1,r,k}$$

$$= 2\sum_{i=1}^{n} (y_{pre,i} - y_i)\cdot \sum_{k=1}^{n} \frac{\partial g_i}{\partial t_k}\cdot \sum_{r=1}^{m} \frac{\partial f_r}{\partial h_p}\cdot w_{1,r,k}$$

$$= 2\sum_{i=1}^{n} (y_{pre,i} - y_i)\cdot \sum_{k=1}^{n} \frac{\partial g_i}{\partial t_k}\ w_{1k}^T \cdot \begin{pmatrix} \frac{\partial f_1}{\partial h_p} \\ \frac{\partial f_2}{\partial h_p} \\ \vdots \\ \frac{\partial f_m}{\partial h_p} \end{pmatrix}$$

$$= 2\sum_{i=1}^{n} (y_{pre,i} - y_i)\cdot \sum_{k=1}^{n} \frac{\partial g_i}{\partial t_k}\cdot w_{1k}^T \cdot \frac{\partial f}{\partial h_p}$$

$$= 2\,(y_{pre} - y)\cdot \begin{pmatrix} \sum_{k=1}^{n} \frac{\partial g_1}{\partial t_k}\cdot w_{1k}^T \frac{\partial f}{\partial h_p} \\ \vdots \\ \sum_{k=1}^{n} \frac{\partial g_n}{\partial t_k}\cdot w_{1k}^T \cdot \frac{\partial f}{\partial h_p} \end{pmatrix}$$

$$= 2\,(y_{pre} - y) \begin{pmatrix} \frac{\partial g_1}{\partial t}\cdot w_1^T \frac{\partial f}{\partial h_p} \\ \vdots \\ \frac{\partial g_n}{\partial t}\cdot w_1^T \frac{\partial f}{\partial h_p} \end{pmatrix}$$

$$= 2\,(y_{pre} - y)\cdot \frac{\partial g}{\partial t}\, w_1^T \frac{\partial f}{\partial h_p}$$

$$\frac{\partial J}{\partial b_0} = \left[\frac{\partial J}{\partial b_{0,1}}\ \frac{\partial J}{\partial b_{0,2}} \cdots \frac{\partial J}{\partial b_{0,m}}\right] = 2\,(y_{pre} - y)\,\frac{\partial g}{\partial t}\cdot w_1^T \left[\frac{\partial f}{\partial h_1}\ \frac{\partial f}{\partial h_2} \cdots \frac{\partial f}{\partial h_m}\right]$$

$$= 2\,(y_{pre} - y)\,\frac{\partial g}{\partial t}\, w_1^T \cdot \frac{\partial f}{\partial h}$$

w_1的第k列

$$t_k = y_0 \cdot w_{1,k} + b_{1,k}$$

$$f_r(h_1, h_2, \cdots h_m)$$

$$h_s = x_s + b_{0,s} \Rightarrow \frac{\partial h_s}{\partial b_{0,p}} = \begin{cases} 1 & s=p \\ 0 & s\neq p \end{cases}$$

$$\frac{\partial y_{0,r}}{\partial b_{0,p}} = \frac{\partial f_r}{\partial b_{0,p}} = \sum_{s=1}^{m} \frac{\partial f_r}{\partial h_s}\,\frac{\partial h_s}{\partial b_{0,p}}$$

$$= \frac{\partial f_r}{\partial h_p}$$

$$\frac{\partial f}{\partial h_p} = \begin{pmatrix} \frac{\partial f_1}{\partial h_p} \\ \frac{\partial f_2}{\partial h_p} \\ \vdots \\ \frac{\partial f_m}{\partial h_p} \end{pmatrix}$$

$$\sum_{k=1}^{n} \left(\frac{\partial g_i}{\partial t_k}\ w_{1k}^T\right) \frac{\partial f}{\partial h_p}$$

$$= \frac{\partial g_i}{\partial t_1}\cdot w_{1,1}^T \frac{\partial f}{\partial h_p} + \frac{\partial g_i}{\partial t_2}\cdot w_{1,2}^T \frac{\partial f}{\partial h_p} + \cdots + \frac{\partial g_i}{\partial t_n}\cdot w_{1,n}^T \cdot \frac{\partial f}{\partial h_p}$$

$$= \left[\frac{\partial g_i}{\partial t_1}\cdot w_{1,1}^T + \frac{\partial g_i}{\partial t_2}\cdot w_{1,2}^T \cdots + \frac{\partial g_i}{\partial t_n}\cdot w_{1,n}^T\right] \frac{\partial f}{\partial h_p}$$

$$= \frac{\partial g_i}{\partial t}\, w_1^T\, \frac{\partial f}{\partial h_p}$$

汕頭大學

SHANTOU UNIVERSITY

softmax $= g$. 则 $g_p(t_1, t_2, \dots t_n) = \dfrac{e^{t_p}}{e^{t_1}+e^{t_2}+\dots+e^{t_n}} = \dfrac{e^{t_p}}{\sum_{i=1}^{n} e^{t_i}}$

则 $$\frac{\partial g}{\partial t} = \begin{pmatrix} \frac{\partial g_1}{\partial t_1} & \frac{\partial g_1}{\partial t_2} & \cdots & \frac{\partial g_1}{\partial t_n} \\ \frac{\partial g_2}{\partial t_1} & \frac{\partial g_2}{\partial t_2} & & \frac{\partial g_2}{\partial t_n} \\ \vdots & \vdots & & \\ \frac{\partial g_n}{\partial t_1} & \frac{\partial g_n}{\partial t_2} & & \frac{\partial g_n}{\partial t_n} \end{pmatrix}$$

$$= \begin{pmatrix} \frac{e^{t_1}\left(\sum_{i=1}^{n} e^{t_i} - e^{t_1}\right)}{\left(\sum_{i=1}^{n} e^{t_i}\right)^2} & \frac{-e^{t_1}\cdot e^{t_2}}{\left(\sum_{i=1}^{n} e^{t_i}\right)^2} & & \\ \frac{-e^{t_2}\cdot e^{t_1}}{\left(\sum_{i=1}^{n} e^{t_i}\right)^2} & \frac{e^{t_2}\left(\sum_{i=1}^{n} e^{t_i} - e^{t_2}\right)}{\left(\sum_{i=1}^{n} e^{t_i}\right)^2} & & \\ & & & \\ & & & \frac{e^{t_n}\left(\sum_{i=1}^{n} e^{t_i} - e^{t_n}\right)}{\left(\sum_{i=1}^{n} e^{t_i}\right)^2} \end{pmatrix}$$

$$= \begin{pmatrix} g_1 & & & \\ & g_2 & & \\ & & \ddots & \\ & & & g_n \end{pmatrix} - \begin{pmatrix} g_1^2 & g_1 g_2 & \cdots & g_1 g_n \\ g_1 g_2 & g_2^2 & & \vdots \\ \vdots & & & \\ g_1 g_n & & \cdots & g_n^2 \end{pmatrix}$$

$$= \mathrm{diag}(g) - \boxed{g^T \cdot g = \begin{pmatrix} g_1 \\ g_2 \\ \vdots \\ g_n \end{pmatrix} (g_1\, g_2 \cdots g_n)}$$

$f_i(h_1, h_2, \ldots, h_m)$

若 $f = \tanh$

$$\tanh x = \frac{e^x - e^{-x}}{e^x + e^{-x}}$$

$$f_1 = \tanh(h_1)$$
$$f_2 = \tanh(h_2)$$
$$\vdots$$
$$f_n = \tanh(h_n)$$

$$\frac{\partial f}{\partial h} = \begin{pmatrix} \frac{\partial f_1}{\partial h_1} & \frac{\partial f_1}{\partial h_2} & \cdots & \frac{\partial f_1}{\partial h_n} \\ \frac{\partial f_2}{\partial h_1} & \frac{\partial f_2}{\partial h_2} & \cdots & \frac{\partial f_2}{\partial h_n} \\ \vdots & & & \\ \frac{\partial f_n}{\partial h_1} & \frac{\partial f_n}{\partial h_2} & \cdots & \frac{\partial f_n}{\partial h_n} \end{pmatrix}$$

$$= \begin{pmatrix} \frac{\partial f_1}{\partial h_1} & & & \\ & \frac{\partial f_2}{\partial h_2} & & \\ & & \ddots & \\ & & & \frac{\partial f_n}{\partial h_n} \end{pmatrix}$$

$$= \begin{pmatrix} 1-f_1^2 & & & \\ & 1-f_2^2 & & \\ & & \ddots & \\ & & & 1-f_n^2 \end{pmatrix}.$$

$$(\tanh x)' = \frac{(e^x + e^{-x})^2 - (e^x - e^{-x})^2}{(e^x + e^{-x})^2}$$

$$= 1 - (\tanh x)^2.$$