

$$S = \begin{bmatrix} s_{00} & s_{01} & \dots \\ s_{10} & & \\ \vdots & & \\ \vdots & & \end{bmatrix} \quad y = [y_0, \dots, y_n]$$

$N \times C$ array
 N pieces of data
 C class scores
 for each

\uparrow correct
 class label
 for each
 datum

$$L_i = \sum_{\substack{k=1 \\ k \neq y_i}}^C \max(0, \Delta + s_k^{(i)} - s_{y_i})$$

$$\max(0, x) = \begin{cases} 0, & 0 > x \\ x, & 0 < x \end{cases}$$

↓

$$\frac{d \max(0, x)}{dx} = \begin{cases} 0 \\ \frac{dx}{dx} = 1 \end{cases} \rightarrow \begin{cases} 0 & 0 > x \\ 1 & 0 < x \end{cases}$$

$$\mathbb{I}(x) \equiv \begin{cases} 0, & 0 > x \\ 1, & 0 < x \end{cases}$$

$$\frac{\partial L_i}{\partial s_j} = \mathbb{I}(0, \Delta + s_j^{(i)} - s_{y_i}) \quad (j \neq y_i)$$

$$\frac{\partial L_i}{\partial s_{y_i}} = -\mathbb{I}(0, \Delta + s_j^{(i)} - s_{y_i})$$