Loss Function

① Structured SVM:

$$\boxed{L_i = \sum_{j \neq y_i} \max(0, S_j - S_{y_i} + 1)}$$

$$S = f(x_i, W) = W x_i \qquad \frac{\partial L}{\partial W} = \frac{\partial L}{\partial S} \frac{\partial S}{\partial W}$$

$$\frac{1}{N} \sum_{i=1} L_i$$

$S_j - S_{y_i} + 1 > 0$ 对于 $j \neq y_i$!

对于 $j = y_i$ 其余的 $-\text{sum}$

需要调整的!

② Softmax Function:

$$\boxed{L_i = -\ln \frac{e^{S_{y_i}}}{\sum_j e^{S_j}}}$$

$$S = f(x_i, W) = W x_i$$

$$\frac{\partial L}{\partial W} = \frac{\partial L}{\partial S} \frac{\partial S}{\partial W}$$

$$L_i = F(S)$$

$$S = G(W)$$

$$\underline{L_i = F(G(W))}$$

对单个样本 i,

$$L_i = -\ln P_{y_i} \quad \text{其中 } P_{y_i} = \frac{e^{S_{y_i}}}{\sum_{j=1}^{c} e^{S_j}}$$

计算 $\frac{\partial L_i}{\partial S_j}$

若 $j = y_i$ $\frac{\partial L_i}{\partial S_{y_i}} = -\frac{1}{P_{y_i}} \cdot \left(\frac{e^{S_{y_i}} \sum e^{S_j} - (e^{S_{y_i}})^2}{(\sum e^{S_j})^2} \right)$

$$= -\frac{1}{P_{y_i}} \cdot \frac{e^{S_{y_i}}}{\sum e^{S_j}} \cdot \frac{\sum e^{S_j} - e^{S_{y_i}}}{\sum e^{S_j}}$$

$$= -\frac{1}{P_{y_i}} \cdot P_{y_i} (1 - P_{y_i})$$

$$= P_{y_i} - 1$$

用原函数表示更方便

Probs对正确部分值减1

$$= \frac{\partial L}{\partial S}$$

若 $j \neq y_i$ $\quad \dfrac{\partial L_i}{\partial s_j} = -\dfrac{1}{p_{y_i}} \dfrac{-e^{s_{y_i}} e^{s_j}}{(\sum e^{s_j})^2} = \dfrac{1}{p_{y_i}} \cdot p_{y_i} \cdot p_j = p_j$

$\dfrac{\partial s_j}{\partial w_j} = x_i \quad (\because s_j = x_i w_j)$