1.
$$P_i = \frac{e^{S_i}}{\sum_j e^{S_j}} \quad , \quad i = 1, 2 \ldots c$$

by
$$\left(\frac{f(x)}{h(x)}\right)' = \frac{f(x)'h(x) - f(x)h'(x)}{[h(x)]^2}$$

$$\boxed{\frac{\partial e^x}{\partial x} = e^x}$$

① when $i = j$ $(S_i = S_j)$

$$\frac{\partial P_i}{\partial S_j} = \frac{e^{S_i} \cdot \sum_j e^{S_j} - e^{S_j} \cdot e^{S_i}}{[\sum_j e^{S_j}]^2} = \frac{e^{S_i}(\sum_j e^{S_j} - e^{S_j})}{[\sum_j e^{S_j}]^2}$$

$$= \frac{e^{S_i}}{\sum_j e^{S_j}} \cdot \frac{\sum_j e^{S_j} - e^{S_j}}{\sum_j e^{S_j}} = P_i(1 - P_j)$$

② when $i \neq j$

$$\frac{\partial P_i}{\partial S_j} = \frac{0 \cdot \sum_j e^{S_j} - e^{S_j} \cdot e^{S_i}}{[\sum_j e^{S_j}]^2}$$

$$= -\frac{e^{S_j}}{\sum_j e^{S_j}} \cdot \frac{e^{S_i}}{\sum_j e^{S_j}} = -P_j P_i$$

2.

$$\begin{cases} P_{y_i} = \dfrac{e^{s_{y_i}}}{\sum_j e^{s_j}} \\ s_j = \underline{W_j}^T \underline{x_j} \end{cases} \qquad \frac{\partial P_{y_i}}{\partial s_j} = \begin{cases} P_{y_i}(1-P_j) & y_i = j \\ -P_{y_i} P_j & y_i \neq j \end{cases}$$

$$\nabla_{\underline{w_j}} L_i = \frac{\partial L_i}{\partial s_j} \frac{\partial s_j}{\partial \underline{W_j}}$$

① $$\frac{\partial L_i}{\partial s_j} = \frac{\partial(-\log P_{y_i})}{\partial s_j} = -\frac{1}{P_{y_i}} \frac{\partial P_{y_i}}{\partial s_j} = \begin{cases} P_j - 1 \\ P_j \end{cases}$$

② $$\frac{\partial s_j}{\partial \underline{W_j}} = \frac{\partial(\underline{w_j}^T \underline{x_j})}{\partial \underline{W_j}} = \underline{x_j}$$

① · ② $$\Rightarrow \nabla_{w_j} L_i = \begin{cases} (P_j - 1)\underline{x_j} & y_i = j \\ P_j \underline{x_i} & y_i \neq j \end{cases}$$

3.

SVM:

$$\nabla_{\mathbf{w}_j} L_i = 1_{(\mathbf{w}_j^T \mathbf{x}_i - \mathbf{w}_{y_i}^T \mathbf{x}_i + 1 > 0)} \mathbf{x}_i,$$

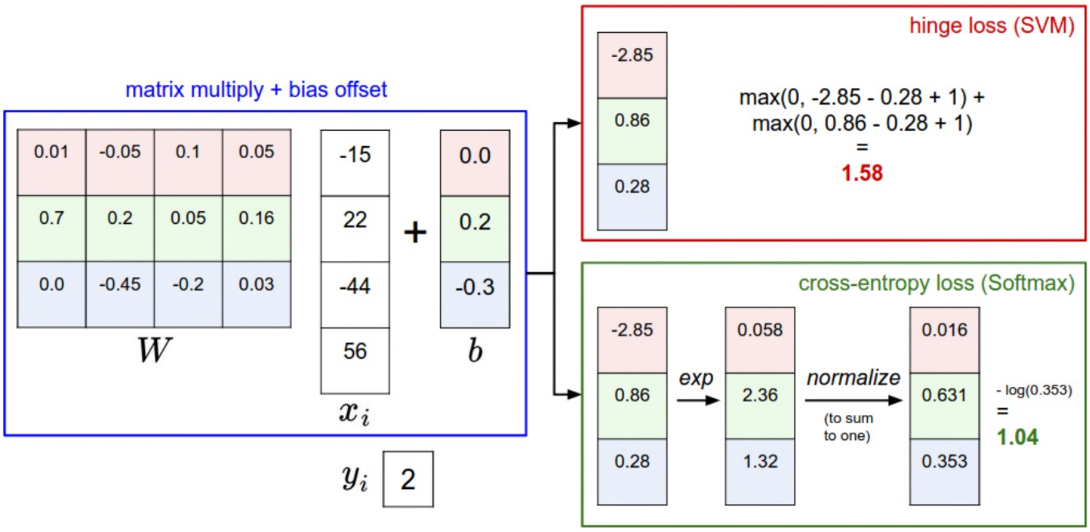
$$\nabla_{\mathbf{w}_{y_i}} L_i = -\left(\sum_{j \neq y_i} 1_{(\mathbf{w}_j^T \mathbf{x}_i - \mathbf{w}_{y_i}^T \mathbf{x}_i + 1 > 0)} \right) \mathbf{x}_i.$$

softmax:

$$\nabla_{\mathbf{w}_j} L_i = P_j \mathbf{x}_i,$$

$$\nabla_{\mathbf{w}_{y_i}} L_i = (P_j - 1) \mathbf{x}_i.$$

Recall:



SVM: Scores.

Softmax: Scores $\longrightarrow$ Probability. (confidence)