5994730249 Woojae Jeong

1. $a = h(s) = \frac{1}{\sum_{m=1}^{M} e^{s_m}} \begin{bmatrix} e^{s_1} \\ \vdots \\ e^{s_m} \end{bmatrix}$

$C = -\sum_{i=1}^{n} y_i \ln a_i$

$\dot{A} = \frac{\partial h(s)}{\partial s} = \begin{bmatrix} \frac{\partial a_1}{\partial s_1} & \cdots & \frac{\partial s_m}{\partial s_1} \\ \vdots & \ddots & \\ \vdots & & \ddots \end{bmatrix}$

$\frac{\partial}{\partial a_i} \ln(a_i) = \frac{1}{a_i} \Rightarrow \frac{\partial}{\partial s_j} \ln(a_i) = \frac{1}{a_i} \frac{\partial a_i}{\partial s_j}$

$\frac{\partial a_i}{\partial s_j} = a_i \frac{\partial}{\partial s_j} \ln(a_i)$, $\ln(a_i) = \ln\left(\frac{e^{s_i}}{\sum_{m=1}^{M} e^{s_m}}\right)$

$= s_i - \ln\left(\sum_{m=1}^{M} e^{s_m}\right)$

$\frac{\partial}{\partial s_j}(a_i) = \frac{\partial s_i}{\partial s_j} - \frac{\partial}{\partial s_j} \ln\left(\sum_{m=1}^{M} e^{s_m}\right)$, $\frac{\partial s_i}{\partial s_j} = \begin{cases} 1, & i=j \\ 0, & i \neq j \end{cases}$

$\frac{\partial}{\partial s_j} \log(a_i) = 1\{i=j\} - \frac{1}{\sum_{m=1}^{M} e^{s_m}} \left(\frac{\partial}{\partial s_j} \sum_{m=1}^{M} e^{s_m}\right)$

$= 1\{i=j\} - \left(\frac{1}{\sum_{m=1}^{M} e^{s_m}} \cdot e^{s_j}\right) = 1\{i=j\} - a_j$

$\therefore \frac{\partial a_i}{\partial s_j} = a_i(1\{i=j\} - a_j)$

$\frac{\partial C}{\partial a_j} = -\frac{\partial}{\partial a_j} \sum_{i=1}^{M} y_i \ln(a_i) = -\sum_{i=1}^{M} y_i \frac{\partial}{\partial a_j} \ln(a_i) = -\sum_{i=1}^{M} y_i \cdot \frac{1}{a_i} \{i=j\}$

$\delta = \begin{bmatrix} a_1(1-a_1), & -a_2 a_1, & \cdots & -a_m a_1 \\ -a_1 a_2, & a_2(1-a_2), & \cdots & \\ \vdots & & \ddots & \\ \vdots & & & \ddots \end{bmatrix} \begin{bmatrix} -\frac{y_1}{a_1} \\ -\frac{y_2}{a_2} \\ \vdots \\ \vdots \end{bmatrix}$

$\delta_j = \left[a_j y_i - y_i\{i=j\}\right] + \sum_{i=2}^{M} a_j y_i$

$= a_j \sum_{i=1}^{M} y_i - y_j$ Since y is one hot shot $\sum_{i=1}^{M} y_i = 1$ $\therefore \delta = a - y$