# David Freilich

## 1. Derivation of Soft-max

$$y_j(x) = \text{softmax}(f(x)) = \frac{e^{f(x_i)}}{\sum_k e^{f(x_k)}} \qquad f(x) = Wx + b$$

① $\frac{\partial y_i}{\partial f_j} = \partial \frac{e^{f(x)}}{\sum_k e^{f(x_k)}}$

② Through quotient rule, get: $\frac{e^{f(x_i)} \sum e^{f(x_k)} - e^{f(x_i)} e^{f(x_j)}}{\left(\sum_k e^{f(x_k)}\right)^2}$

③ @ if $i = j$, get: $\frac{e^{f(x_j)}}{\sum_k e^{f(x_k)}} \cdot \frac{\sum_k e^{f(x_k)} - e^{f(x_j)}}{\sum_k e^{f(x_k)}} = y_i \cdot (1 - y_i)$

ⓑ if $i \neq j$, $\frac{\partial e^{f(x_i)}}{\partial e^{f(x_j)}} = 0$, so get: $\frac{0 \sum e^{f(x_k)} - e^{f(x_i)} e^{f(x_j)}}{\left(\sum_k e^{f(x_k)}\right)^2} = \frac{-e^{f(x_i)} e^{f(x_j)}}{\sum_k e^{f(x_k)}} = -y_i \cdot y_j$

## 2. Derivation of Loss

① $L(X, Y; W, b) = -\frac{1}{N} \sum_j \sum_i Y'_{j,i} \log Y_{j,i}$ (ground truth)

$Y' = 1$ iff image$_j \in$ class$_i$

$Y_{j,i}$ = prob that image$_j \in$ class$_i$

② $\frac{\partial L}{\partial f_i} = -\sum_j \frac{\partial y'_j \cdot \log(y_j)}{\partial x_i} = -\sum_j y'_j \cdot \frac{\partial \log(y_j)}{\partial f_j}$

③ through log derivation, we get: $-\sum_j y'_j \cdot \frac{1}{y_j} \cdot \frac{\partial y_j}{\partial f_i} = -\left(\sum_j \frac{y'_j}{y_j} \cdot \frac{\partial y_j}{\partial f_i}\right)$

④ Split summation, taking out case $j = i$ (truth) so we get: $-\frac{y'_i}{y_i} \frac{\partial y_i}{\partial f_i} - \sum_{j \neq i} \frac{y'_j}{y_j} \cdot \frac{\partial y_j}{\partial f_i}$

⑤ Fill in $\frac{\partial y_i}{\partial x_i}$ softmax, so we get: $-\frac{y'_i}{y_i} \cdot y_i(1 - y_i) - \sum_{j \neq i} \frac{y'_j}{y_j} \cdot (-y_i y_j)$

⑥ $\frac{\partial L}{\partial f_j} = -\frac{y'_i}{y_i}(y_i(1-y_i) - \sum_{j \neq i} \frac{y'_j}{y_j}(-y_i y_j) = -y'_i + y'_i y_i + \sum_{j \neq i} y'_j \cdot y_i$

⑦ Combine $j=i$ w/ summation, so we get:

$$\frac{\partial L}{\partial f_j} = -y'_i + \sum_{j=1} y'_j y_i = -y'_i + y_i \sum_{j=1} y'_j$$

⑧ Since $y'$ is one-hot-encoded, $\sum y'_j = 1$, so we get:

$$\frac{\partial L}{\partial f_j} = y_i - y'_i$$

3. Derive weight gradients

$f(x) = Wx+b$

$$\frac{\partial L}{\partial W} = \frac{\partial L}{\partial f_j} \cdot \frac{\partial f_j}{\partial W_j} = (y_i - y'_i) \cdot \frac{\partial (Wx+b)}{\partial W} = (y_i - y'_i)(x)$$

$$\frac{\partial L}{\partial b_i} = \frac{\partial L}{\partial f_j} \cdot \frac{\partial f_j}{\partial b_j} = (y_i - y'_i) \cdot \frac{\partial (Wx+b)}{\partial b} = (y_i - y'_i)$$