

4.18

$$\begin{aligned}
E(\mathbf{w}_1, \dots, \mathbf{w}_K) &= - \sum_{n=1}^N \sum_{k=1}^K t_{nk} \ln y_{nk} \\
\Rightarrow \frac{\partial E(\mathbf{w}_1, \dots, \mathbf{w}_K)}{\partial \mathbf{w}_j} &= \frac{\partial}{\partial \mathbf{w}_j} \left( - \sum_{n=1}^N \sum_{k=1}^K t_{nk} \ln y_{nk} \right) \\
&= \frac{\partial}{\partial \mathbf{w}_j} \left( - \sum_{n=1}^N \sum_{k=1}^K t_{nk} \ln y_{nk} \right) \\
&= - \sum_{n=1}^N \sum_{k=1}^K \frac{\partial (t_{nk} \ln y_{nk})}{\partial \mathbf{w}_j} \\
&= - \sum_{n=1}^N \sum_{k=1}^K t_{nk} \frac{\partial \ln y_{nk}}{\partial \mathbf{w}_j} \\
&= - \sum_{n=1}^N \sum_{k=1}^K t_{nk} \left( \frac{\partial \ln y_{nk}}{\partial y_{nk}} \right) \left( \frac{\partial y_{nk}}{\partial a_{nj}} \right) \left( \frac{\partial a_{nj}}{\partial \mathbf{w}_j} \right) \\
&= - \sum_{n=1}^N \sum_{k=1}^K t_{nk} \left( \frac{\partial \ln y_{nk}}{\partial y_{nk}} \right) \left( \frac{\partial y_k(\phi_n)}{\partial a_{nj}} \right) \left( \frac{\partial \mathbf{w}_j^T \phi_n}{\partial \mathbf{w}_j} \right) \\
&= - \sum_{n=1}^N \sum_{k=1}^K t_{nk} \left( \frac{1}{y_{nk}} \right) (y_{kn}(I_{jk} - y_{nj})) (\phi_n) \\
&= - \sum_{n=1}^N \sum_{k=1}^K t_{nk} (I_{jk} - y_{nj}) (\phi_n) \\
&= - \sum_{n=1}^N \left( \sum_{k=1}^K t_{nk} I_{jk} - \sum_{k=1}^K t_{nk} y_{nj} \right) (\phi_n) \\
&= - \sum_{n=1}^N \left( t_{nj} - y_{nj} \sum_{k=1}^K t_{nk} \right) (\phi_n) \\
&= - \sum_{n=1}^N (t_{nj} - y_{nj}(1)) (\phi_n) \\
&= - \sum_{n=1}^N (y_{nj} - t_{nj}) (\phi_n)
\end{aligned}$$

which is the same as 4.109.