4.17 The derivative is given by :

$$\frac{\partial y_k}{\partial a_j} = exp(a_k)\frac{\partial}{\partial a_j}\left(\frac{1}{\sum_n exp(a_n)}\right) + \left(\sum_n exp(a_n)\right)^{-1}\frac{\partial exp(a_k)}{\partial a_j}$$

Calculating the derivative of the first term:

$$\begin{aligned}
&exp(a_k)\frac{\partial}{\partial a_j}\left(\frac{1}{\sum_n exp(a_n)}\right)\\
&= exp(a_k)\frac{\partial\left(\sum_n exp(a_n)\right)\right)^{-1}}{\partial a_j}\\
&= exp(a_k)\left(-\left(\sum_n exp(a_n)\right)^{-2}\right)\frac{\partial exp(a_j)}{\partial a_j}\\
&= exp(a_k)\left(-\left(\sum_n exp(a_n)\right)^{-2}\right)\exp(a_j)\\
&= -\frac{exp(a_k)}{\left(\sum_n exp(a_n)\right)}\frac{exp(a_j)}{\left(\sum_n exp(a_n)\right)}\\
&= -y_k y_j
\end{aligned}$$

Calculating the derivative of the second term, it is 0 if $j \neq k$.

If $j = k$,

$$\begin{aligned}
\left(\sum_n exp(a_n)\right)^{-1}\frac{\partial exp(a_k)}{\partial a_j} &= \left(\sum_n exp(a_n)\right)^{-1}\frac{\partial exp(a_k)}{\partial a_k}\\
&= \left(\sum_n exp(a_n)\right)^{-1} exp(a_k)\\
&= \frac{exp(a_k)}{\sum_n exp(a_n)}\\
&= y_k
\end{aligned}$$

which can be written as $I_{kj}y_k$.

Therefore, the derivative becomes:

$$\frac{\partial y_k}{\partial a_j} = -y_k y_j + I_{kj} y_k$$

$$= y_k(I_{kj} - y_j)$$

which is the same as 4.106.