Assignment 5:

Q1: $$P(C_k \mid \phi) = y_k(\phi) = \frac{\exp(a_k)}{\sum_j \exp(a_j)}$$

$$\frac{\partial y_k}{\partial a_j} = \frac{d}{da_j}\left[\frac{\exp(a_k)}{\sum_j \exp(a_j)}\right]$$

$$= \left[\frac{\left(\sum_j \exp(a_j)\right)\cdot \frac{d}{da_j}\left[\exp(a_k)\right] - \exp(a_k)\cdot \frac{d}{da_j}\left[\sum_j \exp(a_j)\right]}{\left(\sum_j \exp(a_j)\right)^2}\right]$$

$$= \frac{\partial y_k}{\partial a_j} = \frac{\left[\left(\sum_j \exp(a_j)\right)\cdot \exp(a_k)\cdot I_{kj} - \exp(a_k)\cdot \exp(a_j)\right]}{\left(\sum_j \exp(a_j)\right)^2}$$

$$\begin{Bmatrix} I_{kj} = 1 & \text{if } j=k, \\ 0 & \text{otherwise} \end{Bmatrix}$$

$$\Rightarrow \frac{\partial y_k}{\partial a_j} = \frac{\exp(a_k)}{\sum_j \exp(a_j)} \cdot \frac{\left[\sum_j \exp(a_j) \cdot I_{kj} - \exp(a_j)\right]}{\sum_j \exp(a_j)}$$

$$= \frac{\partial y_k}{\partial a_j} = y_k \cdot \left[I_{kj} - \frac{\exp(a_j)}{\sum_j \exp(a_j)} \right]$$

$$\boxed{= \frac{\partial y_k}{\partial a_j} = y_k \cdot \left(I_{kj} - y_j \right)}$$

$$E(w_1, \ldots, w_k) = -\ln P(T \mid w_1, \ldots, w_K) = -\sum_{n=1}^{N} \sum_{k=1}^{K} t_{nk} \ln y_{nk}$$

$$= E(w_1, \ldots, w_k) = \frac{d}{dw_j}\left[-\sum_{n=1}^{N} \sum_{k=1}^{K} t_{nk} \ln y_{nk}\right], \quad a_{jk} = w_j^T \phi_n$$

$$= -\sum_{n=1}^{N} \sum_{k=1}^{K} t_{nk} \cdot \left(\frac{1}{y_{nk}}\right) \cdot \frac{\partial y_{nk}}{\partial w_j}$$

$$\Rightarrow \frac{\partial y_{nk}}{\partial w_j} = \sum_j \frac{\partial y_{nk}}{\partial (w_j^T \phi_n)} \cdot \frac{\partial (w_j^T \phi_n)}{\partial w_j} = \sum_j \left[y_{nk}(I_{kj} - y_{nj})\right] \cdot \phi_n$$

$$\nabla_{w_j} E = -\sum_{n=1}^{N} \sum_{k=1}^{K} t_{nk} \cdot \left(\frac{1}{y_{nk}}\right) \cdot \frac{\partial y_{nk}}{\partial w_j}$$

$$= -\sum_{n=1}^{N} \sum_{k=1}^{K} t_{nk} \cdot \left(\frac{1}{y_{nk}}\right) \cdot \sum_j \left[y_{nk}(I_{kj} - y_{nj})\right] \phi_n$$

$$\left\{\begin{array}{ll} I_{kj} = 1 & \text{if } j = k, \\ 0 & \text{otherwise} \end{array}\right\} \quad \sum_{k=1}^{K} \rightarrow K$$

$$= -\sum_{n=1}^{N} K t_{nk} \cdot (I_{kj} - y_{nj}) \phi_n$$

$$= -\sum_{n=1}^{N} (t_{nj} - y_{nj})\,\phi_n$$

$$\boxed{\Rightarrow \nabla_{w_j} E(w_1, \ldots, w_k) = -\sum_{n=1}^{N} (y_{nj} - t_{nj})\,\phi_n}$$