5.

(a) $E(w) = -\ln p(t|w)$

$$= -\ln \left[\prod_{n=1}^{N} \prod_{k=0}^{k-1} p(C_k|\phi(x_n))^{1\{t_n=k\}} \right]$$

$$= - \sum_{n=1}^{N} \sum_{k=0}^{k-1} 1\{t_n=k\} \log \frac{\exp\{w_k^T \phi(x_n)\}}{\sum_0^{k-1} \exp\{w_k^T \phi(x_n)\}}$$

Thus, $\nabla_j E(w) = \frac{\partial E(w)}{\partial w_j} = \partial \left[- \sum_{n=1}^{N} \sum_0^{k-1} 1\{t_n=k\} \log \frac{\exp\{w_k^T \phi(x_n)\}}{\sum_0^{k-1} \exp\{w_k^T \phi(x_n)\}} \right] / \partial w_j$

$$= \partial \left[- \sum_1^{N} \sum_0^{k-1} 1\{t_n=k\} \left[w_k^T \phi(x_n) - \log \sum_{k=0}^{k-1} \exp\{w_k^T \phi(x_n)\} \right] \right] / \partial w_j$$

$$= - \sum_1^{N} \left[\phi\{x_n\} \left(1\{t_n=j\} - \frac{\exp\{w_j^T \phi(x_n)\}}{\sum_0^{k-1} \exp\{w_j^T \phi(x_n)\}} \right) \right]$$

(b) $E^\lambda(w) = E(w) + \frac{\lambda}{2} \sum_{k=0}^{k-1} w_k^T w_k$

$$\nabla_{w_j} E^\lambda(w) = - \sum_1^{N} \left[\phi\{x_n\} \left(1\{t_n=j\} - \frac{\exp\{w_j^T \phi(x_n)\}}{\sum_0^{k-1} \exp\{w_k^T \phi(x_n)\}} \right) \right] + \lambda w_j$$