

**5.32** From 5.139, we have,

$$\begin{aligned}
\tilde{E}(\mathbf{w}) &= E(\mathbf{w}) + \lambda \Omega(\mathbf{w}) \\
\implies \frac{\partial \tilde{E}}{\partial \eta_j} &= \frac{\partial E}{\partial \eta_j} + \frac{\partial \lambda \Omega(\mathbf{w})}{\partial \eta_j} \\
&= 0 + \lambda \frac{\partial \Omega(\mathbf{w})}{\partial \eta_j} \\
&= \lambda \frac{\partial \Omega(\mathbf{w})}{\partial \eta_j}
\end{aligned}$$

Solving the partial derivative:

$$\begin{aligned}
\frac{\partial \Omega(\mathbf{w})}{\partial \eta_j} &= \frac{\partial}{\partial \eta_j} \left( - \sum_i \ln \left( \sum_{k=1}^M \pi_k \mathcal{N}(w_i | \mu_k, \sigma_k^2) \right) \right) \\
&= - \sum_i \frac{\partial}{\partial \eta_j} \ln \left( \sum_{k=1}^M \pi_k \mathcal{N}(w_i | \mu_k, \sigma_k^2) \right) \\
&= - \sum_i \frac{1}{\left( \sum_{j=1}^M \pi_j \mathcal{N}(w_i | \mu_j, \sigma_j^2) \right)} \frac{\partial}{\partial \eta_j} \left( \sum_{k=1}^M \pi_k \mathcal{N}(w_i | \mu_k, \sigma_k^2) \right)
\end{aligned}$$

Note: Here,  $\eta_j$  affects all the  $\pi_k$ s. It does not affect the  $\sigma_j$  as can be seen in the errata of the book, that for equation 5.144,  $\eta$  should be replaced with  $\xi$ .

$$\begin{aligned}
&\frac{\partial}{\partial \eta_j} \left( \sum_{k=1}^M \pi_k \mathcal{N}(w_i | \mu_k, \sigma_k^2) \right) \\
&= \sum_{k=1}^M \left( \mathcal{N}(w_i | \mu_k, \sigma_k^2) \frac{\partial \pi_k}{\partial \eta_j} \right) \\
&= \sum_{k=1}^M \mathcal{N}(w_i | \mu_k, \sigma_k^2) (\pi_j (I_{jk} - \pi_k)) \\
&= \sum_{k=1}^M \mathcal{N}(w_i | \mu_k, \sigma_k^2) \pi_j I_{jk} - \sum_{k=1}^M \mathcal{N}(w_i | \mu_k, \sigma_k^2) \pi_k \pi_j \\
&= \pi_j \sum_{k=1}^M \mathcal{N}(w_i | \mu_k, \sigma_k^2) I_{jk} - \pi_j \sum_{k=1}^M \mathcal{N}(w_i | \mu_k, \sigma_k^2) \pi_k
\end{aligned}$$

$$= \pi_j \mathcal{N}(w_i | \mu_j, \sigma_j^2) - \pi_j \sum_{k=1}^M \mathcal{N}(w_i | \mu_k, \sigma_k^2) \pi_k$$

Substituting this result back into the expression for the partial derivative of  $\Omega$ ,

$$\begin{aligned} \frac{\partial \Omega(\mathbf{w})}{\partial \eta_j} &= - \sum_i \frac{1}{\left( \sum_{j=1}^M \pi_j \mathcal{N}(w_i | \mu_j, \sigma_j^2) \right)} \\ &\quad \left( \pi_j \mathcal{N}(w_i | \mu_j, \sigma_j^2) - \pi_j \sum_{k=1}^M \mathcal{N}(w_i | \mu_k, \sigma_k^2) \pi_k \right) \\ &= - \sum_i (\gamma_j(w_i) - \pi_j) \\ &= \sum_i (\pi_j - \gamma_j(w_i)) \end{aligned}$$

With one final substitution into the expression for the partial derivative of the error function, we get:

$$\Rightarrow \frac{\partial \tilde{E}}{\partial \eta_j} = \lambda \sum_i (\pi_j - \gamma_j(w_i))$$