

5.31 From 5.139, we have,

$$\begin{aligned}
\tilde{E}(\mathbf{w}) &= E(\mathbf{w}) + \lambda \Omega(\mathbf{w}) \\
\Rightarrow \frac{\partial \tilde{E}}{\partial \sigma_j} &= \frac{\partial E}{\partial \sigma_j} + \frac{\partial \lambda \Omega(\mathbf{w})}{\partial \sigma_j} \\
&= 0 + \lambda \frac{\partial \Omega(\mathbf{w})}{\partial \sigma_j} \\
&= \lambda \frac{\partial \Omega(\mathbf{w})}{\partial \sigma_j} \\
\frac{\partial \Omega(\mathbf{w})}{\partial \sigma_j} &= \frac{\partial}{\partial \sigma_j} \left(- \sum_i \ln \left(\sum_{j=1}^M \pi_j \mathcal{N}(w_i | \mu_j, \sigma_j^2) \right) \right) \\
&= - \sum_i \frac{\partial}{\partial \sigma_j} \ln \left(\sum_{j=1}^M \pi_j \mathcal{N}(w_i | \mu_j, \sigma_j^2) \right) \\
&= - \sum_i \frac{1}{\left(\sum_{j=1}^M \pi_j \mathcal{N}(w_i | \mu_j, \sigma_j^2) \right)} \frac{\partial}{\partial \sigma_j} \left(\sum_{j=1}^M \pi_j \mathcal{N}(w_i | \mu_j, \sigma_j^2) \right)
\end{aligned}$$

Since this partial derivative is w.r.t a specific j , it becomes:

$$= - \sum_i \frac{1}{\left(\sum_{j=1}^M \pi_j \mathcal{N}(w_i | \mu_j, \sigma_j^2) \right)} \left(\pi_j \frac{\partial \mathcal{N}(w_i | \mu_j, \sigma_j^2)}{\partial \sigma_j} \right)$$

Simplifying the derivative inside the braces:

$$\begin{aligned}
&\frac{\partial \mathcal{N}(w_i | \mu_j, \sigma_j^2)}{\partial \sigma_j} \\
&= \frac{\partial}{\partial \sigma_j} \left(\frac{1}{(2\pi\sigma_j^2)^{(1/2)}} \exp \left(-\frac{1}{2} \left(\frac{w_i - \mu_j}{\sigma_j} \right)^2 \right) \right) \\
&= -\frac{1}{(2\pi)^{(1/2)}\sigma_j^2} \exp \left(-\frac{1}{2} \left(\frac{w_i - \mu_j}{\sigma_j} \right)^2 \right)
\end{aligned}$$

$$\begin{aligned}
& + \frac{1}{(2\pi\sigma_j^2)^{(1/2)}} \left(\frac{\partial}{\partial\sigma_j} \exp \left(-\frac{1}{2} \left(\frac{w_i - \mu_j}{\sigma_j} \right)^2 \right) \right) \\
& = -\frac{1}{(2\pi)^{(1/2)}\sigma_j^2} \exp \left(-\frac{1}{2} \left(\frac{w_i - \mu_j}{\sigma_j} \right)^2 \right) \\
& + \frac{1}{(2\pi\sigma_j^2)^{(1/2)}} \exp \left(-\frac{1}{2} \left(\frac{w_i - \mu_j}{\sigma_j} \right)^2 \right) \frac{\partial}{\partial\sigma_j} \left(-\frac{1}{2} \left(\frac{w_i - \mu_j}{\sigma_j} \right)^2 \right) \\
& = -\frac{1}{(2\pi)^{(1/2)}\sigma_j^2} \exp \left(-\frac{1}{2} \left(\frac{w_i - \mu_j}{\sigma_j} \right)^2 \right) \\
& + \frac{1}{(2\pi\sigma_j^2)^{(1/2)}} \exp \left(-\frac{1}{2} \left(\frac{w_i - \mu_j}{\sigma_j} \right)^2 \right) \left(\frac{(w_i - \mu_j)^2}{\sigma_j^3} \right) \\
& = -\frac{1}{\sigma_j} \mathcal{N}(w_i|\mu_j, \sigma_j^2) + \mathcal{N}(w_i|\mu_j, \sigma_j^2) \left(\frac{(w_i - \mu_j)^2}{\sigma_j^3} \right) \\
& = \mathcal{N}(w_i|\mu_j, \sigma_j^2) \left(\frac{(w_i - \mu_j)^2}{\sigma_j^3} - \frac{1}{\sigma_j} \right)
\end{aligned}$$

Substituting this result in the partial derivative above, we get:

$$\begin{aligned}
\frac{\partial \tilde{E}}{\partial\sigma_j} & = -\lambda \sum_i \frac{1}{\left(\sum_{j=1}^M \pi_j \mathcal{N}(w_i|\mu_j, \sigma_j^2) \right)} \left(\pi_j \mathcal{N}(w_i|\mu_j, \sigma_j^2) \left(\frac{(w_i - \mu_j)^2}{\sigma_j^3} - \frac{1}{\sigma_j} \right) \right) \\
& = -\lambda \sum_i \left(\frac{\pi_j \mathcal{N}(w_i|\mu_j, \sigma_j^2)}{\left(\sum_{j=1}^M \pi_j \mathcal{N}(w_i|\mu_j, \sigma_j^2) \right)} \left(\frac{(w_i - \mu_j)^2}{\sigma_j^3} - \frac{1}{\sigma_j} \right) \right) \\
& = \lambda \sum_i \gamma_j(w_i) \left(\frac{1}{\sigma_j} - \frac{(w_i - \mu_j)^2}{\sigma_j^3} \right)
\end{aligned}$$

which is the same as the result in 5.143.