Ted Themistokleous
Pawel Jaworski

ECE 657 A2

Q1) Show $W(n+1) = W(n) + \mu_w e(n) \Psi(n)$

$$c_k(n+1) = c_k(n) + \mu_c \frac{e(n) w_k(n)}{\sigma_k^2(n)} \phi[x(n), c_k(n), \sigma_k][x(n) - c_k(n)]$$

$$\sigma_k(n+1) = \sigma_k(n) + \mu_\sigma \frac{e(n) w_k(n)}{\sigma_k^3(n)} \phi[x(n), c_k(n), \sigma_k] \| x(n) - c_k(n) \|^2$$

where

$$\Psi(n) \triangleq [\phi(x(n), c_1, \sigma_1), \phi(x(n), c_2, \sigma_2), \ldots \phi(x(n), c_N, \sigma_N)]^T$$

For gaussian kernel $\phi(\cdot,\cdot,\cdot)$

$$J(n) \triangleq \frac{1}{2}\left[y_d(n) - \sum_{k=1}^{N} w_k(n) \exp\left(\frac{-\|x(n) - c_k(n)\|^2}{2\sigma_k^2(n)}\right)\right]^2$$

for $J(n)$ with $\phi = \exp\left(\frac{-\|x(n) - c_k(n)\|^2}{2\sigma_k^2(n)}\right)$

① $W(n+1) = W(n) - \mu_w \frac{\partial}{\partial w} J(n) \Big|_{w = W(n)}$

② $c(n+1) = c_k(n) - \mu_c \frac{\partial}{\partial c_k} J(n) \Big|_{c_k = c_k(n)}$

③ $\sigma_k(n+1) = \sigma_k(n) - \mu_\sigma \frac{\partial}{\partial \sigma_k} J(n) \Big|_{\sigma_k = \sigma_k(n)}$

For update eq

μ_σ, μ_w, μ_c learning rate params

∴ Follow through differentiation for each update eqn.

①
$$W(n+1) = W(n) - \mu_w \frac{\partial}{\partial w}\left[\frac{1}{2}\left[y_d(n) - \sum_{k=1}^{N} w_k(n) e^{\frac{-\|x(n) - c_k(n)\|^2}{2\sigma_k^2(n)}}\right]\right]^2$$

$$= W(n) - \mu_w \frac{1}{\cancel{2}} \cancel{2} \underbrace{\left[y_d(n) - \sum_{k=1}^{N} w_k(n) e^{\frac{-\|x(n) - c_k(n)\|^2}{2\sigma_k^2(n)}}\right]}_{e(n)} \left[\underbrace{-\sum_{k}^{N} e^{\frac{-\|x(n) - c_k(n)\|^2}{2\sigma_k^2(n)}}}_{-\phi(x(n), c_k(n), \sigma_k(n))}\right]$$ by chain rule

$$= W(n) + \mu_w e(n) \left(\sum_{k=1}^{N} \phi(x(n), c_k(n), \sigma_k(n))\right)$$

$$= W(n) + \mu_w e(n) \underbrace{[\phi(x(n), c_1, \sigma_1), \phi(x(n), c_2, \sigma_2), \ldots, \phi(x(n), c_N, \sigma_N)]^T}_{\Psi(n)}$$

$$W(n+1) = W(n) + \mu_w e(n) \Psi(n)$$

Thus our weight update is just our μ_w rate to scale our output error from each of our weights and each radial node output $\phi(x(n), c_k, \sigma_k)$.

② $c_k(n+1) = c_k(n) - \mu_c \frac{\partial}{\partial c_k} J(n)\Big|_{c_k = c_k(n)}$

$$= c_k(n) - \frac{\mu_c}{2}\frac{\partial}{\partial c_k}\left[y_d(n) - \sum_{k=1}^{N} w_k(n)\exp\left(-\frac{\|x(n)-c_k(n)\|^2}{2\sigma_k^2(n)}\right)\right]^2$$

$$= c_k(n) - \frac{2\mu_k}{2}\Bigg[\underbrace{y_d(n) - \sum_{k=1}^{N} w_k(n)\, e^{-\frac{\|x(n)-c_k(n)\|^2}{2\sigma_k^2(n)}}}_{e(n)}\Bigg]\Bigg[+\frac{w_k(n)}{2\sigma_k^2(n)}\underbrace{e^{-\frac{\|x(n)-c_k(n)\|^2}{2\sigma_k^2(n)}}}_{\phi(x(n),\, c_k(n),\, \sigma_k(n))}\Bigg]2\left[x(n)-c_k(n)\right]$$ by applying chain rule twice

$$= c_k(n) + \mu_k \frac{e(n)\, w_k(n)}{\sigma_k^2(n)}\, \phi(x(n), c_k(n), \sigma_k(n))\left[x(n)-c_k(n)\right]$$

③ $\sigma_k(n+1) = \sigma_k(n) - \mu_\sigma \frac{\partial}{\partial \sigma_k} J(n)\Big|_{\sigma_k = \sigma_k(n)}$

$$= \sigma_k(n) - \frac{\mu_\sigma}{2}\frac{\partial}{\partial \sigma_k}\left[y_d(n) - \sum_{k=1}^{N} w_k(n)\, e^{-\frac{\|x(n)-c_k(n)\|^2}{2\sigma_k^2(n)}}\right]^2$$

$$= \sigma_k(n) - \mu_\sigma \frac{2}{2}\Bigg[\underbrace{y_d(n) - \sum_{k=1}^{N} w_k(n)\, e^{-\frac{\|x(n)-c_k(n)\|^2}{2\sigma_k^2(n)}}}_{e(n)}\Bigg]\Bigg[-w_k(n)\underbrace{e^{-\frac{\|x(n)-c_k(n)\|^2}{2\sigma_k^2(n)}}}_{\phi(x(n),\, c_k(n),\, \sigma_k(n))}\Bigg]\frac{-2\|x(n)-c_k(n)\|^2}{2\sigma_k^3(n)}$$ by chain rule twice

$$= \sigma_k(n) + \mu_\sigma \frac{e(n)\, w_k(n)\, \phi(x(n), c_k(n), \sigma_k(n))}{\sigma_k^3(n)}\|x(n)-c_k(n)\|^2$$