Cost Function:

$$F(heta,\hat{ heta}) = lpha L(x,y, heta) + (1-lpha)U(x,\hat{ heta})$$

Where heta and $\hat{ heta}$ are L and U parameters; and lpha is a trade-off value

$$L(x,w) = rac{1}{2} \sum_{n=1}^N (y_n - w^T \kappa(x_n,C))^2.$$

Donde $\mu, \Sigma \in C$

$$U(x) = \sum_{n=1}^N \sum_{k=1}^K z_{nk} \ln
ho_{nk}$$

Donde $\ln
ho_{nk} = \mathbb{E}[\ln \pi_k] + \frac{1}{2}\mathbb{E}[\ln |\Lambda_k|] - \frac{D}{2}\ln(2\pi) - \frac{1}{2}\mathbb{E}_{\mu_k,\Lambda_k}[(x_n-\mu_k)^T\Lambda(x_n-\mu_k)]$

$$F = rac{1}{2}lpha(y-w^T\kappa(x,C))^2 + (1-lpha)(z\ln
ho)$$

Derivando F respecto a w, μ, Σ :

Derivada respecto a w:

$$rac{\partial F}{\partial w} = lpha \kappa(.) (y_n - w^T \kappa(.))$$

$$x \in \mathbb{R}^{N*D}$$

$$\mu \in \mathbb{R}^{K*D}$$

$$\Lambda \in \mathbb{R}^{K*D*D}$$

$$x_n \in \mathbb{R}^{D*1}$$

$$\mu_k \in \mathbb{R}^{D*1}$$

$$\Lambda_k \in \mathbb{R}^{D*D}$$

$$\kappa(x_n, \mu_k, \Lambda_k) = \sqrt{(x_n - \mu_k)^T \Lambda(x_n - \mu_k)}$$

$$\sqrt{(D*1)^T (D*D)(D*1)} \Rightarrow (1*1)$$

Derivada respecto a μ :

$$\frac{\partial F}{\partial \mu} = \frac{\partial L}{\partial \mu} + \frac{\partial U}{\partial \mu}$$

$$\frac{\partial L}{\partial \mu} = \frac{\partial}{\partial \mu} \left(\frac{1}{2} \alpha (y - w^T \kappa(x, C))^2 \right)$$

$$\frac{\partial L}{\partial \mu} = \frac{1}{2} \alpha 2 (y_n - w^T \sqrt{(x_n - \mu_k)^T \Lambda(x_n - \mu_k)} \left(-w^T \frac{1}{2 \sqrt{(x_n - \mu_k)^T \Lambda(x_n - \mu_k)}} (-2\Lambda(x_n - \mu_k)) \right)$$

$$\frac{\partial L}{\partial \mu} = \alpha (y_n - w^T \kappa(.)) \left(\frac{1}{2} \frac{w^T}{\kappa(.)} 2\Lambda(x_n - \mu_k) \right)$$

$$egin{aligned} rac{\partial U}{\partial \mu} &= rac{\partial}{\partial \mu} igg((1-lpha) z_{nk} \ln igg(\mathbb{E}[\ln \pi_k] + rac{1}{2} \mathbb{E}[\ln |\Lambda_k|] - rac{D}{2} \ln (2\pi) - rac{1}{2} \mathbb{E}_{\mu_k, \Lambda_k} [(x_n - \mu_k)^T \Lambda (x_n - \mu_k)] igg) igg) \ & rac{\partial U}{\partial \mu} = (1-lpha) z_{nk} rac{1}{
ho_{nk}} (\Lambda (x_n - \mu_k)) \end{aligned}$$

$$rac{\partial F}{\partial \mu} = lpha(y_n - w^T \kappa(.\,)) \left(rac{1}{2} rac{w^T}{\kappa(.\,)} 2 \Lambda(x_n - \mu_k)
ight) + (1 - lpha) z_{nk} rac{1}{
ho_{nk}} (\Lambda(x_n - \mu_k))$$

Derivada respecto a Λ :

$$egin{aligned} rac{\partial F}{\partial \Lambda} &= rac{\partial L}{\partial \Lambda} + rac{\partial U}{\partial \Lambda} \ rac{\partial L}{\partial \Lambda} &= rac{\partial}{\partial \Lambda} \left(rac{1}{2}lpha(y-w^T\kappa(x,C))^2
ight) \ \\ rac{\partial L}{\partial \Lambda} &= lpha\left(y-w^T\kappa(.)
ight) \left(rac{1}{2}rac{w^T}{\kappa(.)}
ight) \left(\Lambda(x_n-\mu_k)^T(x_n-\mu_k)\Lambda
ight) \end{aligned}$$

$$egin{aligned} rac{\partial U}{\partial \Lambda} &= rac{\partial}{\partial \Lambda} igg((1-lpha) z_{nk} \ln igg(\mathbb{E}[\ln \pi_k] + rac{1}{2} \mathbb{E}[\ln |\Lambda_k|] - rac{D}{2} \ln (2\pi) - rac{1}{2} \mathbb{E}_{\mu_k, \Lambda_k} [(x_n - \mu_k)^T \Lambda (x_n - \mu_k)] igg) igg) \ &rac{\partial U}{\partial \Lambda} &= (1-lpha) z_{nk} rac{1}{
ho_{nk}} igg(rac{1}{2} ext{tr} \left(\Lambda_k^{-1} rac{\partial \Lambda_k}{\partial \Lambda}
ight) - rac{1}{2} (\Lambda (x_n - \mu_k)^T (x_n - \mu_k) \Lambda) igg) \end{aligned}$$

$$oxed{rac{\partial U}{\partial \Lambda}} = (1-lpha)rac{z_{nk}}{2
ho_{nk}}ig(ext{tr}\left(\Lambda_k^{-1}
ight) - (\Lambda(x_n-\mu_k)^T(x_n-\mu_k)\Lambda)ig)$$

$$rac{\partial F}{\partial \Lambda} = lpha \left(y - w^T \kappa(.)
ight) \left(\Lambda (x_n - \mu_k)^T (x_n - \mu_k) \Lambda
ight) + (1 - lpha) rac{z_{nk}}{2
ho_{nk}} \left(\operatorname{tr} \left(\Lambda_k^{-1}
ight) - (\Lambda (x_n - \mu_k)^T (x_n - \mu_k) \Lambda)
ight)$$

Derivada respecto a π_k :

$$egin{aligned} rac{\partial F}{\partial \pi_k} &= rac{\partial L}{\partial \pi_k} + rac{\partial U}{\partial \pi_k} \ \hline rac{\partial L}{\partial \pi_k} &= 0 \end{aligned}$$

$$oxed{rac{\partial U}{\partial \pi_k}} = (1-lpha) rac{z_{nk}}{
ho_{nk} \pi_k}$$

$$rac{\partial F}{\partial \pi_k} = (1-lpha)rac{z_{nk}}{
ho_{nk}\pi_k}$$

$$\frac{\partial}{\partial \Lambda} \frac{1}{2} \mathbb{E}[\ln |\Lambda_k|]$$

$$\mathbb{E}(q) = \int q(x)p(x)dx$$

$$egin{aligned} q(x) &\Rightarrow q(\Lambda) = \ln |\Lambda_k| \ p(x) &\Rightarrow p(\Lambda) = ? \ &rac{\partial}{\partial \Lambda} rac{1}{2} \mathbb{E}[\ln \pi_k] \end{aligned}$$

$$\int rac{\partial}{\partial w} \mathrm{ln} \, |w^T \Lambda_k w| p(\Lambda) d\Lambda \ \int \mathrm{tr}()$$

$$egin{aligned} y &= w^T arphi(x) = \sum_n \kappa(x_n, x | heta) t_n \ p(heta|t) \propto p(t| heta) p(heta) \ y(x) &= \sum_m w_n \kappa(x, u_n | \Sigma_m) = w^T \kappa(x) \Rightarrow \kappa(x) \in \mathbb{R}^M \ p(w|t, \kappa(\mu, | \Sigma)) \propto p(t|\kappa(\mu, \Sigma), w) p(w) \ p(w) &= \mathcal{N}(w|m_0, S_0) \ p(t|w, \mu, \Sigma) &= \mathcal{N}(t|\kappa w) \Rightarrow \kappa \in \mathbb{R}^{N*N} \ p(w, \mu, \Sigma) &= p(\mu, \Sigma | w) p(w) \ p(w, \mu, \Sigma) &= p(\mu, \Sigma) p(w) \end{aligned}$$

1

Double-click (or enter) to edit