Cost Function:

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

Where θ and $\hat{\theta}$ are L and U parameters; and α is a trade-off value

---

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

Donde $\mu, \Sigma \in C$

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

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

$$F = \frac{1}{2}\alpha(y - w^T\kappa(x,C))^2 + (1-\alpha)(z\ln\rho)$$

Derivando F respecto a w, μ, Σ:

Derivada respecto a w:

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

$$\boxed{\frac{\partial L}{\partial w} = \alpha\kappa(.)(y_n - w^T\kappa(.))}$$

$$\boxed{\frac{\partial U}{\partial w} = 0}$$

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

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

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

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

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

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

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

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

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

Derivada respecto a μ:

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

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

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

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

$$\frac{\partial U}{\partial \mu} = \frac{\partial}{\partial \mu}\left((1-\alpha)z_{nk}\ln\left(\mathbb{E}[\ln \pi_k] + \frac{1}{2}\mathbb{E}[\ln|\Lambda_k|] - \frac{D}{2}\ln(2\pi) - \frac{1}{2}\mathbb{E}_{\mu_k,\Lambda_k}[(x_n-\mu_k)^T\Lambda(x_n-\mu_k)]\right)\right)$$

$$\boxed{\frac{\partial U}{\partial \mu} = (1-\alpha)z_{nk}\frac{1}{\rho_{nk}}(\Lambda(x_n-\mu_k))}$$

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

Derivada respecto a Λ:

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

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

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

$$\frac{\partial U}{\partial \Lambda} = \frac{\partial}{\partial \Lambda}\left((1-\alpha)z_{nk}\ln\left(\mathbb{E}[\ln \pi_k] + \frac{1}{2}\mathbb{E}[\ln|\Lambda_k|] - \frac{D}{2}\ln(2\pi) - \frac{1}{2}\mathbb{E}_{\mu_k,\Lambda_k}[(x_n-\mu_k)^T\Lambda(x_n-\mu_k)]\right)\right)$$

$$\frac{\partial U}{\partial \Lambda} = (1-\alpha)z_{nk}\frac{1}{\rho_{nk}}\left(\frac{1}{2}\mathrm{tr}\left(\Lambda_k^{-1}\frac{\partial \Lambda_k}{\partial \Lambda}\right) - \frac{1}{2}(\Lambda(x_n-\mu_k)^T(x_n-\mu_k)\Lambda)\right)$$

$$\boxed{\frac{\partial U}{\partial \Lambda} = (1-\alpha)\frac{z_{nk}}{2\rho_{nk}}\left(\mathrm{tr}\left(\Lambda_k^{-1}\right) - (\Lambda(x_n-\mu_k)^T(x_n-\mu_k)\Lambda)\right)}$$

$$\frac{\partial F}{\partial \Lambda} = \alpha\left(y - w^T\kappa(.)\right)\left(\frac{w^T}{2\kappa(.)}\right)\left(\Lambda(x_n-\mu_k)^T(x_n-\mu_k)\Lambda\right) + (1-\alpha)\frac{z_{nk}}{2\rho_{nk}}\left(\mathrm{tr}\left(\Lambda_k^{-1}\right) - (\Lambda(x_n-\mu_k)^T(x_n-\mu_k)\Lambda)\right)$$

Derivada respecto a π_k :

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

$$\boxed{\frac{\partial L}{\partial \pi_k} = 0}$$

$$\boxed{\frac{\partial U}{\partial \pi_k} = (1-\alpha)\frac{z_{nk}}{\rho_{nk}\pi_k}}$$

$$\frac{\partial F}{\partial \pi_k} = (1-\alpha)\frac{z_{nk}}{\rho_{nk}\pi_k}$$

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

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

$q(x) \Rightarrow q(\Lambda) = \ln |\Lambda_k|$

$p(x) \Rightarrow p(\Lambda) = ?$

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

$\int \frac{\partial}{\partial w} \ln |w^T \Lambda_k w| p(\Lambda) d\Lambda$

$\int \mathrm{tr}()$

$y = w^T \varphi(x) = \sum_n \kappa(x_n, x|\theta) t_n$

$p(\theta|t) \propto p(t|\theta) p(\theta)$

$y(x) = \sum_m w_n \kappa(x, u_n|\Sigma_m) = w^T \kappa(x) \Rightarrow \kappa(x) \in \mathbb{R}^M$

$p(w|t, \kappa(\mu, |\Sigma)) \propto p(t|\kappa(\mu, \Sigma), w) p(w)$

$p(w) = \mathcal{N}(w|m_0, S_0)$

$p(t|w, \mu, \Sigma) = \mathcal{N}(t|\kappa w) \Rightarrow \kappa \in \mathbb{R}^{N*N}$

$p(w, \mu, \Sigma) = p(\mu, \Sigma|w) p(w)$

$p(w, \mu, \Sigma) = p(\mu, \Sigma) p(w)$

```
1
```

Double-click (or enter) to edit