Unsupervised deep embedding for clustering analysis (DEC)

(F) Unsupervised deep embedding for clustering Analysis.

Problem : n points $\{x_i \in X\}_{i=1}^{n}$ ⇒ k clusters.

cluster centroids : $\{\mu_j\}$; $j = 1 \dots k$, $\mu_j \in Z$

data space ↔ embedding (low dim)

Non linear mapping $f_\theta : X \to Z$

Solution : clustering with KL divergence :

① mapping & initialisation μ_j f_θ

Soft Assignment :

$$q_{ij} = \frac{\left(1 + \|z_i - \mu_j\|^2/\alpha\right)^{-\left(\frac{\alpha+1}{2}\right)}}{\sum_{j'} \left(1 + \|z_i - \mu_{j'}\|^2/\alpha\right)^{-\left(\frac{\alpha+1}{2}\right)}}$$

($z_i = f_\theta(x_i)$; α : inverse Hyperparam = 1)

KL Divergence Minimization :-

choice option for P ??

Loss, $L = KL(P\|Q) = \sum_i \sum_j p_{ij} \log \frac{p_{ij}}{q_{ij}}$

$$\Rightarrow p_{ij} = \frac{q_{ij}^2/f_j}{\sum_{j'} q_{ij'}^2/f_{j'}} \quad // \quad f_j = \sum_i q_{ij}$$

(b) DEC cont.

Optimization:

$$\frac{\partial L}{\partial z_i} = \frac{\alpha+1}{\alpha} \sum_j \left(1 + \frac{\|z_i - \mu_j\|^2}{\alpha}\right)^{-1} \times (p_{ij} - q_{ij})(z_i - \mu_j)$$

$$\frac{\partial L}{\partial \mu_j} = -\frac{\alpha+1}{\alpha} \sum_i \left(1 + \frac{\|z_i - \mu_j\|^2}{\alpha}\right)^{-1} \times (p_{ij} - q_{ij})(z_i - \mu_j)$$

Accuracy Metric (classification setting)

$$ACC = \max_m \frac{\sum_{i=1}^{n} \mathbb{1}\{l_i = m(c_i)\}}{n}$$ // Latent class

$l_i \Rightarrow$ true label

$m \Rightarrow$ mapping of cluster c_i to label l_i