Model Based Clustering:

clustering set of data points by fitting a mixture model, where cluster corresponds to a component of the mixture

Mixture Model:

Each sample is from one & only one component

$$f(x) = \pi_1 f_1(x_1 \theta_i) + \dots + \pi_K f_K(x_K \theta_K)$$

$$x_1, x_2, \dots, x_K \underset{iid}{\sim} f(x)$$

$$z_1, z_2, \dots, z_n \sim Dir(\pi_1, \dots, \pi_K)$$

$$x_i \sim f_\theta(\cdot)$$

$$f_\theta(x) = \pi \phi(x_i, \mu_1, \sigma_1^2)$$

$$+ (1-\pi)\phi(x_i, \mu_2, \sigma_2^2)$$

$$(x_i, z_i) \qquad z_i = \begin{cases} 1 & w/p \ \pi \\ 2 & w/p \ 1-\pi \end{cases}$$

$$X_i \mid Z_i = \begin{cases} N(-, \mu_1, \sigma_1^2), & \text{if } z_i = 1 \\ N(-, \mu_2, \sigma_2^2), & \text{if } z_i = 2 \end{cases}$$

$$\prod_{i=1}^{n} \left[\pi \phi(x_i, \mu_1, \sigma_1^2)\right]^{\{z_i=1\}} \left[(1-\pi)\phi(x_i, \mu_2, \sigma_2^2)\right]^{\{z_i=2\}}$$

$$\prod_{i:z_i=1} \pi \phi(x_i, \mu_1, \sigma_1^2) \prod_{i:z_i=2} (1-\pi)\phi(x_i, \mu_2, \sigma_2^2)$$

$$\prod_{i=1}^{n} \prod_{K=1}^{K} \left[\pi_K f_K(x_i, \theta_K) \right]^{\mathbb{I}\{z_i = K\}}$$

But we don't observe z_i so instead use iterative methods to esitimate

$$x \sim P_\theta(\cdot) \qquad P_\theta = \sum_{z=1}^{K} P_\theta(x, z) \rightarrow P_\theta(z) P(x \mid z)$$

$$\hat{\theta}_{mle} = \arg\max_{\theta} \sum_{i=1}^{n} \log P_\theta(x_i)$$

$$g(\theta) = \sum_{i=1}^{n} \sum_{k=1}^{K} p_{ik}\left[\log \pi_k + -\frac{1}{2}\log \sigma_k^2 - \frac{(x_i - \mu_k)}{2\sigma_k^2}\right]$$

$$\theta = \begin{cases} \pi_1, \dots, \pi_K \\ (\mu_k, \sigma_k^2)_{k=1}^{K} \end{cases}$$

$$\sum_{i=1}^{n} \sum_{k=1}^{K} p_{ik} \log \pi_k = \sum_{k=1}^{K} \left(\sum_{i=1}^{n} p_{ik} \right) \log \pi_k$$

$$= p_{i1} \log \pi_1 + p_{i2} \log \pi_2 + \dots + p_{iK} \log \pi_K$$

ex: $\underset{11}{\underline{3.2}} \log \pi_1 + \underset{11}{\underline{4.8}} \log \pi_2 + \underset{11}{\underline{3}} \log \pi_3$

$p_1\, p_2\, p_3$ is prob vector

$$\Rightarrow p_1 \log \pi_1 + p_2 \log \pi_2 + p_3 \log \pi_3$$

$$\Rightarrow p_1 \log \frac{\pi_1}{p_1} + p_2 \log \frac{\pi_2}{p_2} + p_3 \log \frac{\pi_3}{p}$$

$$= -KL(p_{1:3} \,||\, \pi_{1:3})$$

$\downarrow$

KL Divergence