Mixture models and K-means clustering : similarity

(K-means : same cluster size
Mixtures : different cluster sizes determined by $\hat{w}_k$)

Mixture of Gaussians:

$$p(x) = \sum_{k=1}^{K} w_k \frac{1}{\sqrt{2\pi}\sigma} \exp^{-\frac{1}{2\sigma^2}(x_i - \mu_k)^2}$$

EM algorithm (model-based)

$\underline{\theta}_k = (\mu_k, \sigma_k^2)$

E-step: compute:

$$\hat{w}_{ik}^{(t)} = \frac{w_k \, g_k(x_i \mid \underline{\theta}_k)}{\sum_{k=1}^{K} w_k \, g_k(x_i \mid \underline{\theta}_k)}$$

M-step: update:

$$\hat{\mu}_k^{(t+1)} = \sum_{k=1}^{K} \hat{w}_{ik}^{(t)} x_i \Big/ \sum_{k=1}^{K} \hat{w}_{ik}^{(t)}$$

$$\sigma_k^{2(t+1)} = \sum_{k=1}^{K} \hat{w}_{ik}^{(t)} (x_i - \hat{\mu}_k^{(t+1)}) \Big/ \sum_{k=1}^{K} \hat{w}_k^{(t)}$$

1. update weights (mixing propotions) $\hat{w}_k$
2. update estimates $\hat{\mu}_k$ and σ_k^2 (mean and variance)

K-means clustering (nonparametric)

1. assign $c_i^{(t)} = \arg\min_k \| x_i - \mu_k^{(t)} \|$
2. update: $\mu_k^{(t+1)} = \sum_{i=1}^{n} x_i \, \mathbb{1}(c_i^{(t)} = k) \Big/ \sum_{i=1}^{n} \mathbb{1}(c_i^{(t)} = k)$

1. update nearest center c_i
2. update cluster means μ_k