Oct 29, 2024

Gaussian Mixture Model

$\theta = \{\pi, \mu_1, \Sigma_1, \mu_2, \Sigma_2\}$

$\pi_1 + \pi_2 = 1$

$$P(x_i \mid \theta) = \sum_{k=1}^{K} \underbrace{P(z_k)}_{\pi_k} \underbrace{P_k(x_i \mid \theta)}_{\sim N(\mu_k, \Sigma_k)}$$

$$L(\theta) = \log \prod_{i=1}^{N} P(x_i \mid \theta)$$

$\hat{\theta} = \max_{\theta} L(\theta)$

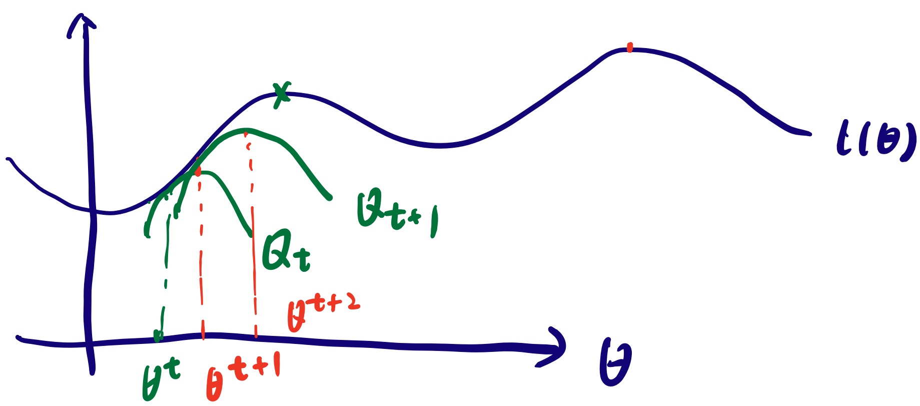
$$= \sum_{i=1}^{N} \log \left(\sum_{k=1}^{K} P(z_k) P_k(x_i \mid \theta) \right)$$

$\log \Pi = \Sigma \log$ easy to calculate

$\log \Sigma$ difficult to calculate

EM algorithm

Expectation Maximization



Q : Auxilary fuction, Q is alway below $l(\theta)$

no guarentee to global optimal

E : constructing Q, given θ

$$P(z_i = k \mid x_i, \theta) = r_{ik}$$

M : optimize Q function, reestimate θ.

$$\pi_k = \frac{1}{N} \sum_i r_{ik}$$

$$\mu_k = \frac{\sum_i r_{ik} x_i}{\sum_i r_{ik}}$$

$$\Sigma_k = \frac{\sum_i r_{ik} (x_i - \mu_k)(x_i - \mu_k)^T}{\sum_i r_{ik}}$$

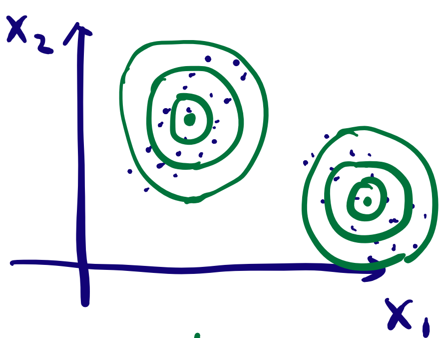
k-means clustering.

assume:

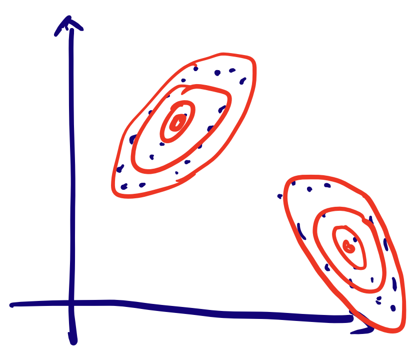
$$\Sigma = \sigma^2 I = \begin{bmatrix} \sigma^2 & & \\ & \ddots & \\ & & \sigma^2 \end{bmatrix}$$

$$\pi_k = \frac{1}{k}$$

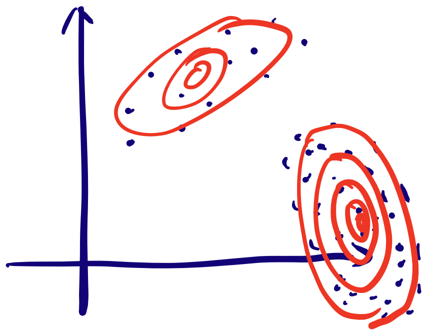
$$\mu_k = ?$$



k-means μ_k



GMM π_k μ_k Σ_k



$\pi_k = \{0.2, 0.8\}$