Univariate Mixture model - gaussian [Without $Z_{ij}$]

$$P(\vec{x_i}|\theta) = \sum_{i=1}^{k} P_j \, P(\vec{x_i}|\theta_j)$$

$P_j$ - mixing weight

$\theta_j$ - parameters of pdf

$$\sum_{j=1}^{k} P_j = 1$$

$$\theta = \{\theta_1, \theta_2 \cdots \theta_k, P_1 \cdots P_k\}$$

The PDF is a gaussian distribution

$$P(x_i|\theta_j) = \frac{1}{\sqrt{2\pi\sigma_j}} e^{-1/2\left[\frac{x-\mu}{\sigma_j}\right]^2}$$

The likelihood function is

$$P(\vec{x}|\theta_j) = \prod_{i=1}^{N} \sum_{j=1}^{k} P_j \cdot P(x_i|\theta_j)$$

The log likelihood function is

$$\mathcal{L}\phi(x, \theta_j, \lambda_1 \cdots \lambda_k, \Lambda) = \sum_{i=1}^{N} \log \sum_{j=1}^{k} P_j \cdot P(x_i|\theta_j) + \lambda_1 P_1 + \lambda_2 P_2 + \cdots + \lambda_k P_k + \Lambda\left(1 - \sum_{j=1}^{k} P_j\right)$$

To derive the maximum log likelihood, differentiate w.r.t all parameters & equate to zero.

① $\dfrac{\partial\phi(x, \theta_j, \lambda_1 \cdots \lambda_k, \Lambda)}{\partial P_j} = 0$

② $\dfrac{\partial\phi(x, \theta_j, \lambda_1 \cdots \lambda_k, \Lambda)}{\partial \theta_j} = 0$

Scanned by CamScanner

③ $\frac{\partial \phi}{\partial \lambda_j}(x, \theta_j, \lambda_1 \cdots \lambda_K, \Lambda) = 0$

④ $\frac{\partial \phi}{\partial \Lambda}(x, \theta_j, \lambda_1 \cdots \lambda_K, \Lambda) = 0$

Take ③ and ④

Solve $\frac{\partial \phi}{\partial \lambda_j}(x, \theta, \lambda_1 \cdots \lambda_K, \Lambda) = P_j$

$\neq 0$

As it contradicts constraint

$\lambda_1 = \lambda_2 = \cdots \lambda_K = 0$

$\frac{\partial \phi}{\partial \Lambda}(x, \theta, \lambda_1 \cdots \lambda_K, \Lambda) = 1 - \sum_{j=1}^{K} P_j = 0$

$\boxed{\sum_{j=1}^{K} P_j = 1} \rightarrow ⑤$

10

Take ① $\frac{\partial \phi}{\partial P_j}(x, \theta, \lambda_1 \cdots \lambda_K, \Lambda) = \sum_{i=1}^{N} \frac{1}{\sum_{j=1}^{K} P_j \cdot P(x_i|\theta_j)} \frac{\partial}{\partial P_j}\left\{\sum_{j=1}^{K} P_j P(x_i|\theta_j)\right\} - \Lambda$

$\Rightarrow \sum_{i=1}^{N} \frac{P(x_i|\theta_j)}{\sum_{j=1}^{K} P_j \cdot P(x_i|\theta_j)} - \Lambda = 0$

Multiplying and dividing by $P_j$

$\frac{1}{P_j}\left\{\sum_{i=1}^{N} \frac{P_j \cdot P(x_i|\theta_j)}{\sum_{j=1}^{K} P_j \cdot P(x_i|\theta_j)}\right\} - \Lambda$

By Bayes rule, $\frac{1}{P_j} \sum_{i=1}^{N} P(j|x_i) - \Lambda = 0$

$$\Rightarrow \hat{P}_j = \frac{1}{\Lambda} \sum_{i=1}^{N} P(j|x_i) \longrightarrow (6)$$

Sub in (5)

$$\sum_{j=1}^{K} \frac{1}{\Lambda} \sum_{i=1}^{N} P(j|x_i) = 1$$

$$\frac{1}{\Lambda} \sum_{i=1}^{N} \sum_{j=1}^{K} P(j|x_i) = 1$$

$$\frac{1}{\Lambda} N = 1$$

$$\boxed{\Lambda = N} \longrightarrow (7)$$

Sub in (6)

$$\boxed{\hat{P}_j = \frac{1}{N} \sum_{i=1}^{N} P(j|x_i)} \longrightarrow (8)$$

Solve (2)

$$\frac{\partial \phi(x, \theta, \lambda_1 \ldots \lambda_K, \Lambda)}{\partial \theta_j} = \sum_{i=1}^{N} \frac{1}{\sum_{j=1}^{K} P_j \cdot P(x_i|\theta_j)} \frac{\partial}{\partial \theta_j} \sum_{j=1}^{K} P_j \cdot P(x_i|\theta_j)$$

$$= \frac{\sum_{i=1}^{N} P_j \frac{\partial}{\partial \theta_j} \sum_{j=1}^{K} P(x_i|\theta_j)}{\sum_{j=1}^{K} P_j \cdot P(x_i|\theta_j)}$$

Multiplying and dividing by $P(x_i|\theta_j)$

$$= \frac{\sum_{i=1}^{N} P_j \cdot P(x_i|\theta_j)}{\sum_{j=1}^{K} P_j \cdot P(x_i|\theta_j)} \frac{\partial}{\partial \theta_j} \frac{\sum_{j=1}^{K} P(x_i|\theta_j)}{P(x_i|\theta_j)}$$

$$= \sum_{i=1}^{N} P(j|x_i) \frac{\partial}{\partial \theta_j} \log P(x_i|\theta_j)$$

$$\log P(x_i|\theta_j) = \log \left\{ \frac{1}{\sqrt{2\pi}\,\sigma_j} e^{-1/2 \left(\frac{x_i-\mu_j}{\sigma_j}\right)^2} \right\}$$

$$= \log (2\pi)^{-1/2} + \log(\sigma_j)^{-1/2} + \cancel{\log} \left[ \frac{-1}{2} \frac{(x_i-\mu_j)^2}{\sigma_j^2} \right]$$

$$= \frac{-1}{2} \log 2\pi - \frac{1}{2} \log \sigma_j - \frac{1}{2} \frac{(x_i-\mu_j)^2}{\sigma_j^2}$$

Deriving mean, $\theta_j = (\mu_j, \sigma_j^2)$

$$\frac{\partial}{\partial \mu_j} \log P(x_i|\theta_j) = \frac{\partial}{\partial \mu_j} \left[ \frac{-1}{2} \frac{(x_i-\mu_j)^2}{\sigma_j^2} \right]$$

$$= 2/_2 \frac{(x_i-\mu_j)}{\sigma_j^2}$$

$$\frac{\partial \phi}{\partial \mu_j} = \sum_{i=1}^{N} P(j|x_i) \left[ \frac{(x_i-\mu_j)}{\sigma_j^2} \right]$$

$$0 = \sum_{j=1}^{N} P(j|x_i) x_i - \sum_{i=1}^{N} P(j|x_i) \mu_j$$

$$\sum_{i=1}^{N} P(j|x_i) x_i = \sum_{i=1}^{N} P(j|x_i) \mu_j$$

$$\boxed{\hat{\mu}_j = \frac{\sum_{i=1}^{N} P(j|x_i) x_i}{P(j|x_i)}}$$

$$\frac{\partial}{\partial \sigma_j} \log P(x_i|\theta_j) = \frac{-1}{2\sigma_j} + \left(\frac{-1}{2\sigma_j^3}(x_i-\mu_j)^2\right)$$

$$\frac{\partial \phi}{\partial \sigma_j} = \sum_{i=1}^{N} P(j|x_i)\left\{\frac{-1}{\sigma_j} + \frac{1}{\sigma_j^3}(x_i-\mu_j)^2\right\}$$

$$0 = \sum_{i=1}^{N} P(j|x_i)\left\{-\sigma_j^2 + (x_i-\mu_j)^2\right\}$$

$$\sum_{i=1}^{N} P(j|x_i)\sigma_j^2 = \sum_{i=1}^{N} P(j|x_i)(x_i-\mu_j)^2$$

$$\boxed{\sigma_j^2 = \frac{\sum_{i=1}^{N} P(j|x_i)(x_i-\mu_j)^2}{\sum_{i=1}^{N} P(j|x_i)}}$$

(2) Multivariate Mixture model (Without $z_{ij}$)

(3) Binomial Mixture Model – With and without $z_{ij}$
– Univariate
– Multivariate

(4) Bernoulli Mixture Model

(5) Multinomial Mixture Model