EM Algorithm

Junrui Lin

2023 Jan

Method of Optimization
2023 Jan

Junrui Lin
Email: jl12680@nyu.edu

1 Setup

In finite mixtures model, we assume there is K types of heterogeneity (group heterogeneity).

$$F(y|x) = \sum_{k=1}^{K} \pi_k(x) F_k(y|x)$$

Where π_k is the mixing probabilities, and $F_k(y|x)$ are the mixture component. When we do MLE, we choose θ to maximize the log-likelihood

$$\log F(y|\theta, \pi) = \sum_{i=1}^{N} \log \left(\sum_{k=1}^{K} \pi_k F_k(y_{i1}, ..., y_{iT}|\theta) \right)$$

But the objective function is hard to maximize by FOC, we use EM algorithm.

2 Computation

Pretends that we know π and θ, and iterate until $\pi^{(t)}, \theta^{(t)}$ converge. Initial guess $\pi^{(0)}, \theta^{(0)}$
Define

$$P\bigg(Z_i = k | \{y_{it}\}_{t=1}^{T}, \theta^{(0)}, \pi^{(0)} \bigg) = \frac{\pi_k^{(0)} F_k\bigg(\{y_{it}\}_{t=1}^{T}, \theta^{(0)} \bigg)}{\sum_{j=1}^{K} \pi_j^{(0)} F_j\bigg(\{y_{it}\}_{t=1}^{T}, \theta^{(0)} \bigg)}$$

Expectation

$$Q(\theta, \pi | \theta^{(0)}, \pi^{(0)}) := \sum_{i=1}^{N} \sum_{k=1}^{K} P\bigg(Z_i = k | \{y_{it}\}_{t=1}^{T}, \theta^{(0)}, \pi^{(0)} \bigg) \log \bigg[\pi_k F_k\bigg(\{y_{it}\}_{t=1}^{T}, \theta \bigg) \bigg]$$

Maximization

$$(\theta^{(1)}, \pi^{(1)}) = \arg\max_{\theta,\pi} Q(\theta, \pi|\theta^{(0)}, \pi^{(0)})$$

Repeat the process until converge.

3 Proof of Convergence

$$\log F(y_i|\theta,\pi) = \log F(y_i, z_i|\theta,\pi) - \log F(z_i|y_i,\theta,\pi)$$

$$\begin{aligned}\log F(y_i|\theta,\pi) = &\sum_{k=1}^{K} P(z_i = k|y_i, \theta^{(t)}, \pi^{(t)}) \log F(y_i, z_i = k|\theta,\pi) - \\ &\sum_{k=1}^{K} P(z_i = k|y_i, \theta^{(t)}, \pi^{(t)}) \log P(z_i = k|y_i, \theta, \pi) \\ \log F(y_i|\theta,\pi) - \log F(y_i|\theta^{(t)}, \pi^{(t)}) = &\sum_{k=1}^{K} P(z_i = k|y_i, \theta^{(t)}, \pi^{(t)}) \log P(z_i = k|y_i, \theta, \pi) - \\ &\sum_{k=1}^{K} P(z_i = k|y_i, \theta^{(t)}, \pi^{(t)}) \log P(z_i = k|y_i, \theta^{(t)}, \pi^{(t)}) + \\ &\sum_{k=1}^{K} P(z_i = k|y_i, \theta^{(t)}, \pi^{(t)}) \log P(z_i = k|y_i, \theta^{(t)}, \pi^{(t)}) - \\ &\sum_{k=1}^{K} P(z_i = k|y_i, \theta^{(t)}, \pi^{(t)}) \log P(z_i = k|y_i, \theta, \pi)\end{aligned}$$

Then use the inequality: $\forall p, q \in \triangle_K \Rightarrow -\sum_{l=1}^{K} p_l \log p_l \leq -\sum_{l=1}^{K} p_l \log q_l$

$$\sum_{k=1}^{K} P(z_i = k|y_i, \theta^{(t)}, \pi^{(t)}) \log P(z_i = k|y_i, \theta, \pi) - \sum_{k=1}^{K} P(z_i = k|y_i, \theta^{(t)}, \pi^{(t)}) \log P(z_i = k|y_i, \theta^{(t)}, \pi^{(t)}) \leq 0$$

$$\sum_{k=1}^{K} P(z_i = k|y_i, \theta^{(t)}, \pi^{(t)}) \log P(z_i = k|y_i, \theta^{(t)}, \pi^{(t)}) - \sum_{k=1}^{K} P(z_i = k|y_i, \theta^{(t)}, \pi^{(t)}) \log P(z_i = k|y_i, \theta, \pi) \geq 0$$

$$\log F(y_i|\theta,\pi) - \log F(y_i|\theta^{(t)}, \pi^{(t)}) \geq Q(\theta, \pi|\theta^{(t)}, \pi^{(t)}) - Q(\theta^{(t)}, \pi^{(t)}|\theta^{(t)}, \pi^{(t)})$$