EM estimation for the mixture of von Mises distribution

The von Mises (vM) distribution is a probability distribution defined to describe periodic data such as angles and directions on a circle. Its probability density distribution (pdf) is given by:

$$f_{vM}(\theta|\mu,\kappa) = \frac{1}{2\pi I_0(\kappa)} e^{\kappa\cos(\theta-\mu)}, \kappa \geq 0 \text{ and } 0 \leq \mu < 2\pi, \tag{1}$$

where θ is expected to be a random variable representing angles, κ is the shape parameter, μ is the location parameter, and $I_0(\cdot)$ is the modified Bessel function. The mathematical form of α-order Bessel function with zero-order and firs-order cases are defined as follows:

$$I_\alpha(x) = \sum_{m=0}^{\infty} \frac{(\frac{x}{2})^{2m+\alpha}}{m!\Gamma(m+\alpha+1)}; I_0(x) = \sum_{\kappa=0}^{\infty} \frac{\left(\frac{x}{2}\right)^{2\kappa}}{\kappa!\kappa!}; I_1(x) = \sum_{\kappa=0}^{\infty} \frac{\left(\frac{x}{2}\right)^{2k+1}}{\kappa!(\kappa+1)!}. \tag{2}$$

Then I extend the discussion to mixture of vM distribution. Assume the mixture distribution has m components with parameters $\boldsymbol{\kappa} = \{\kappa_1, ..., \kappa_m\}$, $\boldsymbol{\mu} = \{\mu, ..., \mu\}$. Given angles $\boldsymbol{\theta} = \{\theta_1, ..., \theta_n\}, i.i.d.$, and a matrix of corresponding latent class variables $\mathbf{Z} = (z_{ij})_{n\times m}$ with z_{ij}representing the j-th observation in component i, the complete data set is written as $\{\theta, \boldsymbol{Z}\}$. I introduce $\boldsymbol{\pi} = \{\pi_1, ..., \pi_m\}$ as the responsibility (i.e. proportion of distributional shape explained) of component j on the angle data $\boldsymbol{\theta}$. The probability mass function of a mixture of vM distribution can be written as:

$$p(\boldsymbol{\theta}, \mathbf{Z}|\boldsymbol{\pi}, \boldsymbol{\mu}, \boldsymbol{\kappa}) = \prod_{i=1}^{n}\prod_{j=1}^{m} \pi_j^{z_{ij}} f_{vM}(\theta_i|\mu_j,\kappa_j)^{z_{ij}} \tag{3}$$

E step:

$$\begin{aligned}
\mathbb{E}(z_{ij}) &= \mathbb{E}_z\left[\ln p(\theta_i, z_{ij}|\pi_j, \mu_j, \kappa_j)\right] && (4)\\
&= \sum_{i=1}^{n}\sum_{j=1}^{m} \gamma_{ij}\left\{\ln \pi_j + \ln f_{vM}(\theta_i|\mu_j, k_j)\right\} && (5)\\
&= \sum_{i=1}^{n}\sum_{j=1}^{m} \gamma_{ij}\left\{\ln \pi_j + \kappa\cos(\theta_i - \mu_j) - \ln 2\text{pi} - \ln I_0(\kappa)\right\}, && (6)\\
&\quad \text{where} \quad \gamma_{ij} = \mathbb{E}(z_{ij}) = \frac{\pi_j f_{vM}(\theta_i|\mu_j,\kappa_j)}{\sum_{s=1}^{m}\pi_s f_{vM}(\theta_i|\mu_s,\kappa_s)}. && (7)
\end{aligned}$$

M step: Maximize the log-likelihood of $\mathbb{E}(z_{ij})$ with respect to $\{\boldsymbol{\pi}, \boldsymbol{\mu}, \boldsymbol{\kappa}\}$, and the solutions are:

$$\begin{aligned}
\pi_j &= \frac{1}{n}\sum_{i=1}^{n}\gamma_{ij} && (8)\\
\mu_j &= \arctan\left(\frac{\sum_{i=1}^{n}\gamma_{ij}\sin\theta_i}{\sum_{i=1}^{n}\gamma_{ij}\cos\theta_i}\right) && (9)\\
\kappa_j &= A^{-1}(\kappa_j), \quad \text{where} \quad A(\kappa_i) = \frac{I_1(\kappa_j)}{I_0(\kappa_j)} = \frac{\sum_{i=1}^{n}\gamma_{ij}\cos(\theta_i-\mu_j)}{\sum_{i=1}^{n}\gamma_{ij}} && (10)
\end{aligned}$$

The main computing burden lies in the computation of the M step over the parameter κ due to inverse matrix computing. Use a combined form of **Theorem 2 & Theorem 4** in Ruiz-Antolín and Segura (2016), a tight bound of the ratio of modified Bessel function is given by:

$$\frac{x}{\nu - 1/2 + \sqrt{(\nu + 1/2)^2 + x^2}} < \frac{I_\nu(x)}{I_{\nu-1}(x)} < \frac{x}{\nu - 1/2 + \sqrt{(\nu - 1/2)^2 + x^2}}, \; \nu \geq 0. \tag{11}$$

Then the approximation of $\frac{I_1(\kappa_j)}{I_0(\kappa_j)}$ has the specific numerical form:

$$\frac{x}{\frac{1}{2} + \sqrt{\frac{9}{4} + x^2}} < \frac{I_1(x)}{I_0(x)} < \frac{x}{\frac{1}{2} + \sqrt{\frac{1}{4}^2 + x^2}}, \; \nu \geq 0. \tag{12}$$

This result greatly facilities the computing of the M step. By constructing a table containing the approximate values of the ratio of modified Bessel functions, the estimation of parameter κ is transformed to searching the closest value of the original parameter "x" given the ratio of modified Bessel functions "y". A primal sample code is available at https://github.com/alanaguo/Sample-code/blob/main/delta_radian%20%26%20fit_MoV.jl.

References

Ruiz-Antolín, D., Segura, J., 2016. A new type of sharp bounds for ratios of modified bessel functions. Journal of Mathematical Analysis and Applications 443, 1232–1246. URL: https://www.sciencedirect.com/science/article/pii/S0022247X16302402.