Issues on Fitting Univariate and Multivariate Hyperbolic Distributions

Thomas Tran
PhD student
Department of Statistics
University of Auckland

thtam@stat.auckland.ac.nz

– Fitting Hyperbolic Distributions –
Acknowledgement

- I am grateful my supervisor for his guidance
- The financial assistance of my department and the workshop organizers is highly appreciated
Outline

• Generalized inverse Gaussian distributions, normal mean-variance mixture distributions, generalized hyperbolic distributions, hyperbolic distribution

• Fitting the univariate hyperbolic distribution using functions in the HyperbolicDist package

• EM and MCECM algorithms

• Fitting multivariate generalized hyperbolic distributions
References

Grounding work:


Recently, among others:


Practical work


- The R package `HyperbolicDist, QRMlib` and others
Generalized Inverse Gaussian distributions

\[ h(w \mid \lambda, \chi, \psi) = \frac{(\psi/\chi)^{\lambda/2}}{2 K_\lambda(\sqrt{\psi\chi})} w^{\lambda-1} \exp \left( -\frac{1}{2} (\psi w^{-1} + \psi w) \right) \]

\[ E(W^\alpha) = \left( \frac{\chi}{\psi} \right)^{\frac{\alpha}{2}} \frac{K_{\lambda+\alpha}(\sqrt{\chi\psi})}{K_\lambda(\sqrt{\chi\psi})} \]

Where \( w > 0, \chi > 0 \) and \( \psi > 0 \)

- \( K_\lambda \) denotes the modified Bessel function of the third kind with order \( \lambda \)

- Representable in 4 parameterizations with \( \lambda \) specifies the order of the modified Bessel function. The skewness and kurtosis parameters can be: \( (\chi, \psi), (\delta, \gamma), (\alpha, \beta), (\omega, \zeta) \). Function \textit{gigChangePars}
Normal Mean-Variance Mixtures distributions

Univariate generalized hyperbolic distribution

\[ f(x) = \int_0^\infty \frac{1}{(2\pi)^{\frac{1}{2}}w^{\frac{1}{2}}} \exp \left[ -\frac{(x - M)^2}{2w} \right] h(w)dw \]

Multivariate generalized hyperbolic distribution

\[ g(x) = \int_0^\infty \frac{1}{(2\pi)^{\frac{d}{2}}|\Sigma|^{\frac{1}{2}}w^{\frac{d}{2}}} \exp \left[ -\frac{(x - M)'\Sigma^{-1}(x - M)}{2w} \right] h(w)dw \]

where \( \Sigma = AA' \) and \( \Sigma \) has rank \( d \). In both cases, \( M = \mu + w\gamma \) and \( h(w) \) is the density of the generalized inverse Gaussian distribution.
Multivariate hyperbolic distribution

From the mean-variance mixture, a multivariate generalized hyperbolic distribution can be derived

\[
g(x) = c \frac{K_{\lambda - \frac{d}{2}} \left( \sqrt{[\chi + (x - \mu)\Sigma^{-1}(x - \mu)]}[\psi + \gamma'\Sigma^{-1}\gamma] \right)}{\left( \sqrt{[\chi + (x - \mu)\Sigma^{-1}(x - \mu)]}[\psi + \gamma'\Sigma^{-1}\gamma] \right)^{\frac{d}{2} - \lambda}} e^{(x - \mu)'\Sigma^{-1}\gamma}
\]

Where the normalizing constant is

\[
c = \frac{(\sqrt{\chi\psi})^{-\lambda}\psi^\lambda(\psi + \gamma'\Sigma^{-1}\gamma)^{\frac{d}{2} - \lambda}}{(2\pi)^{\frac{d}{2}}|\Sigma|^{\frac{1}{2}}K_{\lambda}(\sqrt{\chi\psi})}
\]

- If \( \lambda = \frac{d+1}{2} \) we have a \( d \)-dimensional hyperbolic distribution. This is represented in the \((\lambda, \chi, \psi, \Sigma, \gamma, \mu)\) parameterizations. The parameterizations suggested in Blæsild (1985) is \((\lambda, \alpha, \delta, \Delta, \beta, \mu)\)
Univariate Generalized Hyperbolic Distributions

Univariate generalized hyperbolic distribution

\[
f(x|\lambda, \alpha, \beta, \delta, \mu) = \frac{(\gamma/\delta)^\lambda}{\sqrt{2\pi}K_\lambda(\delta\gamma)} \frac{K_{\lambda-1/2}(\alpha\sqrt{\delta^2 + (x-\mu)^2})}{(\sqrt{\delta^2 + (x-\mu)^2}/\alpha)^{1/2-\lambda}} e^{\beta(x-\mu)}
\]

where \( K_\lambda \) denotes the third modified Bessel function of order \( \lambda \), \( \gamma^2 = \alpha^2 - \beta^2 \) and \(-\infty < x < \infty\)

- Has 5 parameters

- Can be represented by 4 parameterizations All have \( \lambda, \delta, \mu \) as the order of the Bessel function, scale and location parameter respectively. The skewness and kurtosis parameters can be: \( (\alpha, \beta) \), \( (\zeta, \rho) \), \( (\xi, \chi) \), and \( (\bar{\alpha}, \bar{\beta}) \). Function \( ghypChangePars \)
The Hyperbolic Distribution

- For $\mu = 0$ and $\delta = 1$ the density of the hyperbolic distribution is
  \[
  \frac{1}{2\sqrt{1 + \pi^2} K_1(\zeta)} \times \exp \left( -\zeta \left[ \sqrt{1 + \pi^2} \sqrt{1 + x^2} - \pi x \right] \right)
  \]
  where $K_1$ denotes the modified Bessel function of third kind with index 1.

- Four different parameterizations. All have $\mu$ as location and $\delta$ as scale parameter. The skewness and kurtosis parameters can be $(\pi, \zeta)$; $(\alpha, \beta)$; $(\phi, \gamma)$; $(\xi, \chi)$ each parameterizations can be useful for different purposes. Function `hyperbChangePars`
The Hyperbolic Distribution

Plot of the log densities of different values of $\xi$ and $\chi$
Fitting the Univariate Hyperbolic Distribution

For fitting purpose of our talk we use the \((\pi, \zeta)\) parameterizations

\[
c(\pi, \zeta, \delta, \mu) \times \exp \left\{ -\zeta \left[ \sqrt{1 + \pi^2} \sqrt{1 + \left( \frac{x - \mu}{\delta} \right)^2} - \pi \left( \frac{x - \mu}{\delta} \right) \right] \right\}
\]

where

\[
c(\pi, \zeta, \delta, \mu) = \left[ 2\delta \sqrt{1 + \pi^2} K_1(\zeta) \right]^{-1}
\]

as in the \((\pi, \zeta)\) the only restriction is \(\zeta > 0\)
Fitting the Univariate Hyperbolic Distribution

\texttt{hyperbFitStart(x, startValues = "BN",}
\texttt{ startMethodSL = "Nelder-Mead",}
\texttt{ startMethodMoM = "Nelder-Mead",...)}

This function returns 4 different starting values for the optimization process carried out by the \texttt{hyperbFit} function.

- BN: Barndorf Neilsen 1977
- SL: Fitted Skewed Laplace
- FN: Fitted normal
- MoM: Method of Moment
- US: User supplied
Fitting the Univariate Hyperbolic Distribution

- Fitting functions:

```r
hyperbFit(x, startValues = "BN", method = "Nelder-Mead", ..
```

- Nelder-Mead: A non-derivative optimization method suggested by Nelder and Mead (1965)
- BFGS: A derivative optimization method pioneered by Broyden-Fletcher-Goldfarb-Shanno
Fitting the Univariate Hyperbolic Distribution

• Three questions raised in Barndorff-Nielsen and Blæsild (1981) regarding to:
  – Sample size?
  – Parameterization?
  – Fitting methods?

Fitting the Univariate Hyperbolic Distribution

Xi fitting
\[ \xi = 0.3y = -0.04 \]

Chi fitting
\[ \chi = 0.3y = -0.04 \]

Sample Size

<table>
<thead>
<tr>
<th>Sample Size</th>
</tr>
</thead>
<tbody>
<tr>
<td>500</td>
</tr>
<tr>
<td>1000</td>
</tr>
<tr>
<td>1500</td>
</tr>
<tr>
<td>2000</td>
</tr>
<tr>
<td>2500</td>
</tr>
<tr>
<td>3000</td>
</tr>
</tbody>
</table>

\[ \xi^* - \xi \]

\[ \chi^* - \chi \]

SL_BFGS
FN_BFGS
SL_Nelder-Mead
FN_Nelder-Mead

ξ = 0.3
χ = −0.04
Fitting the Univariate hyperbolic distribution

---

Fitting Hyperbolic Distributions – 17
Fitting the Univariate hyperbolic distribution
Fitting the Univariate hyperbolic distribution
Conclusions on fitting Univariate Hyperbolic Distribution

- None of the methods outperforms one another

- Better fit when sample size becomes larger. The “fairly” large sample size is about 2000 or more observations

- The fitting of $\xi$ appears to be fairly accurate even with small sample size when the hyperbolic distributions are at the left and right boundary of the shape triangle is notable
**Expectation - Maximization algorithm**

Proposed by Dempster *et al.* (1977). If we were able to have a complete data $Y$ we wish

$$L(\theta|Y) \propto \log f(Y|\theta) \quad \theta \in \Theta \subseteq \mathbb{R}^d$$

We assume $Y = (Y_{obs}, Y_{mis})$. Each iteration of the EM algorithm comprises of two steps E and M. The $(t + 1)$st E-step finds:

$$Q(\theta|\theta^{(t)}) = \int L(\theta|Y)f(Y_{mis}|Y_{obs}, \theta = \theta^{(t)})dY_{mis}$$

$$= E[L(\theta|Y)|Y_{obs}, \theta = \theta^{(t)}]$$

The $(t + 1)$st M-step then finds $\theta^{(t+1)}$ that maximizes

$$Q(\theta^{(t+1)}|\theta^{(t)}) \geq Q(\theta|\theta^{(t)}) \quad \theta \in \Theta$$
Fitting the Multivariate Hyperbolic Distribution

The complete data $Y$ can also be thought of as the augmented data $Y_{aug}$. The $(t + 1)$ E step can be written as:

$$Q(\theta|\theta^{(t)}) = E[L(\theta|Y_{aug})|Y_{obs}, \theta = \theta^{(t)}]$$

We augmented the data to fit our model which is $f(x|\theta = \lambda, \chi, \psi, \mu, \Sigma, \gamma)'$. Using the mean variance mixture feature we could also alter our model to fit the data by partitioning $\theta = (\theta_1, \theta_2)$ where:

$$\theta_1 = (\mu, \Sigma, \gamma)'$$
$$\theta_2 = (\lambda, \chi, \psi)'$$


**Fitting the Multivariate Hyperbolic Distribution**

We could construct the augmented log likelihood

\[
\ln [L(\theta|X_1 \ldots X_n, W_1 \ldots W_n)] = \sum_{i=1}^{n} \ln (f_{X|W_i} X_i|W_i; \theta_1) + \sum_{i=1}^{n} \ln (h_W|W_i; \theta_2) 
\]

_E step._ We calculate

\[
Q(\theta; \theta^t) = E[\ln L(\theta; X_1, \ldots, X_n, W_1, \ldots, W_n|X_1, \ldots, X_n; \theta^t)]
\]

_M step._

\[
Q(\theta^{(t+1)}|\theta^{(t)}) \geq Q(\theta|\theta^{(t)})
\]
Fitting the Multivariate Hyperbolic Distribution

In practice the calculation of \((t + 1)\)st E step reduces to find the values of the following quantities:

\[
\delta_i^{[t]} = E[W_i^{-1}|X_i; \theta^{[t]}] \\
\eta_i^{[t]} = E[W_i|X_i; \theta^{[t]}] \\
\zeta_i^{[t]} = E[ln(W_i)|X_i; \theta^{[t]}]
\]

Where \(W_i|X_i\) follows a GIG with;

\[
\tilde{\lambda} = \lambda - \frac{1}{2}d \\
\tilde{\chi} = (X_i - \mu)'\Sigma^{-1}(X_i - \mu) + \chi \\
\tilde{\psi} = \psi + \gamma'\Sigma^{-1}\gamma
\]
Fitting the Multivariate Hyperbolic Distribution

In the M step we maximize the augmented log likelihood by maximizing each of its two components \( Q_1(\mu, \Sigma, \gamma; \theta_1^{[t]}) \) and \( Q_2(\lambda, \chi, \psi; \theta_2^{[t]}) \) separately.

- \( Q_1 \) can be maximized analytically as the first derivative of \( Q_1 \) wrt each parameter can be achieved in closed form

- \( Q_2 \) must be evaluated numerically
MCECM algorithm

Meng and Rubin (1993) proposed a variant of the EM called MCECM algorithm. At $(t)$th iteration

- The E step remains the same as with the EM algorithm

- The maximization method for $Q_1$ is unchanged and carried out first. This results in $(\theta_1^{[t+1]} = \mu^{[t+1]}, \Sigma^{[t+1]}, \gamma^{[t+1]})$.

- $Q_2$ is then numerically maximized as $Q_2(\lambda, \chi, \psi; \theta_2^{[t]}, \theta_1^{[t+1]})$. Compared with the calculation of $Q_2$ in the EM algorithm, it is now included $\theta_1^{[t+1]}$.
Results and conclusions

Comparison of convergence speed between EM and MCECM algorithm for 4 dimensional hyperbolic distribution

<table>
<thead>
<tr>
<th>conv.code</th>
<th>logLik</th>
<th>number.iter</th>
</tr>
</thead>
<tbody>
<tr>
<td>MCECM.BFGS</td>
<td>−12545.18</td>
<td>75.00</td>
</tr>
<tr>
<td>MCECM.NM</td>
<td>−12545.18</td>
<td>75.00</td>
</tr>
<tr>
<td>EM.BFGS</td>
<td>−12546.14</td>
<td>88.00</td>
</tr>
<tr>
<td>EM.NM</td>
<td>−12546.14</td>
<td>88.00</td>
</tr>
</tbody>
</table>

- The MCECM appears to converge faster than the EM
- The estimation of $\lambda$ causes both algorithms become unstable
Results

Generated data

var 1

var 2

var 3

– Fitting Hyperbolic Distributions –
Results

Fitted data

var 1

var 2

var 3

– Fitting Hyperbolic Distributions –