A formulation for continuous mixtures of multivariate normal distributions

Adelchi Azzalini
Università di Padova, Italia

joint work with Reinaldo B. Arellano-Valle
Pontificia Universidad Católica de Chile

LinStat2020 Conference — 31st August 2021

Continuous mixtures of multivariate normal variables

Start from $X \sim \mathrm{N}_d(0, \Sigma)$, $\Sigma > 0$, and let $\xi \in \mathbb{R}^d$ be a fixed vector.

Then combine X with another (scalar, indept) random variable, such as:

1. $\sqrt{V}\, X$ for some random $V > 0 \implies$ *scale mixture* of X (a wide subset of the elliptical class of distributions);
2. $U\,\xi + X$ for some random $U \implies$ *mean mixture*;
3. $V\,\xi + \sqrt{V}\, X \implies$ *variance-mean mixture* (if $V \sim \mathrm{GIG}$, get the Generalized Hyperbolic distribution).

Considerable literature exists, especially for the forms 1 and 3.
(impossible to review it here; see the paper for references)

In recent years, mixtures of other distributions have been considered, especially with $X \sim \mathrm{SkewNormal}$.

A two-component formulation

- Re-start from $X \sim \mathrm{N}_d(0, \Sigma)$, but now use *two* mixing terms
- Let U, V two scalar variables, such that X, U, V are all independent.
- Given a real-valued function r, a positive-valued function s and $\gamma \in \mathbb{R}^d$, we consider

$$
\begin{aligned}
Y &= \xi + r(U,V)\,\gamma + s(U,V)\,X \\
&= \xi + R\,\gamma + S\,X
\end{aligned}
$$

called a *generalized mixture of normals* (GMNs).

Aims

- The basic aim is to highlight a common framework: many existing mixtures (not only of normals) are GMNs
- Hence obtain a better understanding of their nature and connections
- Also, can we carry out a unified treatment of the properties?
- Finally, the GMNs scheme might suggest new constructions

First general implications

$$Y = \xi + R\,\gamma + S\,X$$

- The relevant ingredients are ξ, γ, Σ and the distribution of $(R, S) \sim H$
- So write $Y \sim \mathrm{GMN}_d(\xi, \Sigma, \gamma, H)$
- An affine transformation of Y is still of type GMN
- Hence the GMN class is closed under marginalization
- $\mathbb{E}\{Y\} = \xi + \mathbb{E}\{R\}\,\gamma$, if the moment exists
- $\mathrm{var}\{Y\} = \mathrm{var}\{R\}\,\gamma\,\gamma^\top + \mathbb{E}\{S^2\}\,\Sigma$, if the moments exist

Results after some algebra (...sometimes a lot of it)

- If Y is partitioned into Y_1 and Y_2, the conditional distribution of Y_1 given Y_2, which is still of GMN type
- A collection of facts about quadratic forms of Y
- An expression for the Mardia's measures of multivariate asymmetry and kurtosis, involving moments of (R, S) up to the fourth order

Classical mixtures of normals

- $(R, S) = (0, S) \quad \Longrightarrow$ scale/variance mixtures of normals
- $(R, S) = (R, 1) \quad \Longrightarrow$ mean mixtures of normals
- $(R, S) = (V, \sqrt{V})$ with $V > 0 \quad \Longrightarrow$ variance-mean mixtures

Reproduce a wide range of classical constructions

Aside: the skew-normal (SN) distribution – 1

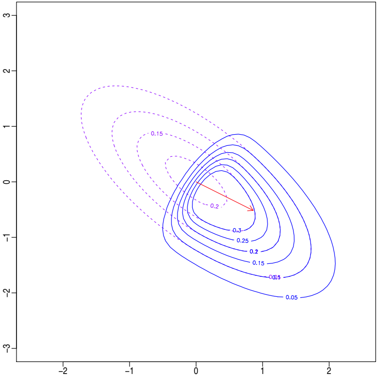
$$\mathrm{SN.density}(x) = 2\,\mathrm{Normal.density}(x)\,\Phi(\eta^\top x)$$

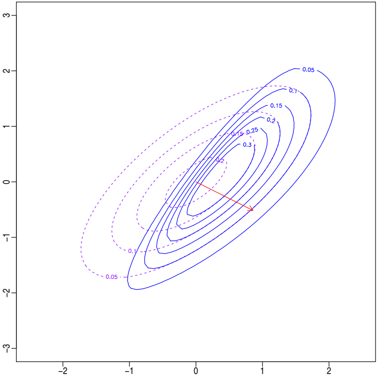
- Extends the Normal distribution with extra parameter $\eta \in \mathbb{R}^d$
- If $\eta = 0$ back to Normal, otherwise the density is asymmetric
- Numerous formal properties are available
- Representation as mean mixture of Normals:

$$\xi + R\gamma + X \equiv \xi + U\gamma + X$$

where $R \equiv U \sim \chi_1$ and γ is related to η.

Aside: the skew-normal (SN) distribution – 2





Mixtures of SN distributions

- If $Z \sim \mathrm{SN}$, its scale mixture takes the form, dropping ξ,

$$\sqrt{V}\,Z = \sqrt{V}\,(U\gamma + X) = U\sqrt{V}\gamma + \sqrt{V}\,X$$

  hence this is a GMN with $R = \sqrt{V}\,U$, $S = \sqrt{V}$.

  Can regulate both skewness and tail weight
  (Example: a popular instance is the skew-t distribution)

- Scale mixing can be combined with shape mixing

$$U\gamma \Longrightarrow U\,(\tilde{U}\tilde{\gamma}), \qquad \sqrt{V}\,U\gamma \Longrightarrow \sqrt{V}(U\,\tilde{U})\tilde{\gamma})$$

  hence GMN with $R = \sqrt{V}(U\,\tilde{U}) = \sqrt{V}\,U^*$ and $S = \sqrt{V}$.

An exploration with mean mixtures

$$Y = \xi + R\gamma + Z$$

- If $R \sim \chi_1$, get the SN distribution
- Another option: take $R \sim \chi_2$ (Rayleigh distribution)
- More generally, $R \sim \chi_\nu$ (Nakagami m-distribution)

Final comments

- The GMN class includes many forms of mixtures, not all
- Extensions are possible in a number of directions
- However, these will come at some cost,
  like loss of some formal properties, technical difficulties, etc.
- GMN seeks a balance between generality and tractability

The end

The full story:

R. B. Arellano-Valle and A. Azzalini
A formulation for continuous mixtures of multivariate normal distributions.
J. Multivariate Analysis, vol. 185, September 2021

Thanks for your attention.