

MODEL-BASED CLUSTERING VIA PARSIMONIOUS MIXTURES OF DIMENSION-WISE SCALED NORMAL MIXTURES

Antonio Punzo¹, Luca Bagnato² and Salvatore Daniele Tomarchio¹

¹ Dipartimento di Economia e Impresa, Università di Catania, (e-mail: antonio.punzo@unict.it, daniele.tomarchio@unict.it)

² Dipartimento di Scienze Economiche e Sociali, Università Cattolica del Sacro Cuore, (e-mail: luca.bagnato@unicatt.it)

ABSTRACT: Dimension-wise scaled normal mixtures (DSNMs; Punzo & Bagnato, 2022) are a recently defined family of d -variate continuous distributions that generalize the multivariate normal (MN) to allow for 1) a more general central symmetry, and 2) an excess kurtosis that can vary dimension-wise. DSNMs have the further interesting property, shared by the MN distribution too, that no correlation implies independence. These peculiarities are obtained in an MN scale mixture framework by introducing a d -variate mixing random variable with independent and similar components acting separately for each dimension.

We introduce parsimonious finite mixtures of DSNMs for model-based clustering in the presence of symmetric clusters with an amount of excess kurtosis that can vary in each dimension. For illustrative purposes, we describe two members of the DSNM mixture family obtained in the case of mixing random variables being either uniform or shifted exponential; these are examples of mixing distributions that guarantee a closed-form expression for the joint density of the DSNM. For the two described members, we introduce parsimony by putting convenient constraints on the conditional correlation and scale parameters, as well as on the tailedness parameters. This gives rise to a family of 60 interpretable models. Depending on the model under consideration, we describe and use one of two possible variants of the expectation-maximization algorithm to obtain maximum likelihood estimates of the parameters. Finally, we consider simulated and real data to appreciate the advantages of our mixture models over well-established mixtures of symmetric heavy-tailed distributions.

KEYWORDS: Central symmetry, Heavy-tailed distributions, Scale mixtures.

References

PUNZO, A., & BAGNATO, L. 2022. Dimension-wise scaled normal mixtures with application to finance and biometry. *Journal of Multivariate Analysis*, **191**, 105020.