DSSV 2023 Data Science, Statistics & Visualisation
ECDA 2023 European Conference on Data Analysis

Compositional splines for representation of bivariate density functions

Stanislav Škorňa[1], Karel Hron[1], Jana Burkotová[1], and Jitka Machalová[1]

[1]Department of Mathematical Analysis and Applications of Mathematics, Faculty of Science, Palacký University Olomouc, 17. listopadu 1192/12, Olomouc 779 00, Czech Republic

As a result of aggregation of massive data, multivariate densities occur naturally and they are used to further process them using methods of functional data analysis and to analyze the association structure. Accordingly, proper spline (continuous) representation of the input discrete data (e.g. in form of histograms) is crucial for their statistical analysis.

Bayes Hilbert spaces methodology enables to capture specific properties of probability density functions and to construct so-called compositional splines which respect their decomposition into interactive and independent parts [1]. One-to-one transformation to standard L^2 space can be achieved using centered log-ratio (clr) transformation, where the original densities (as well as compositional splines) are represented by ZB-spline representation. The resulting spline functions fulfill zero-integral constraint, which must be considered already when building the basis of the ZB-spline representation. This can be done using the classical approach of standard B-spline basis with implemented zero-integral constraint [1, 2] or using ZB-spline basis, which satisfies the zero-integral constraint automatically [3]. In the contribution we focus on the latter case, provide an appropriate spline representation supported with a detailed simulation study and application for descriptive analysis of geochemical data.

[1] K. Hron, J. Machalová, and A. Menafoglio, "Bivariate densities in bayes spaces: orthogonal decomposition and spline representation," *Stat Papers*, 2022.

[2] J. Machalová, K. Hron, and G. Monti, "Preprocessing of centred logratio transformed density functions using smoothing splines," *Journal of Applied Statistics*, vol. 43:8, pp. 1419–1435, 2016.

[3] J. Machalová, R. Talská, K. Hron, and A. Gába, "Compositional splines for representation of density functions," *Computational Statistics*, vol. 36, p. 1031–1064, 2020.