Meet The Jury Seminar

Correlation Analysis for Compositional Data

Prof. dr. Peter Filzmoser

Head of research unit Computational Statistics, Institute of Statistics and Mathematical Methods in Economics, Vienna University of Technology

November 26, 2019 Room B02.14

10.00-11.00h **Celestijnenlaan 200B, Heverlee**

Abstract

Many environmental data sets consist of concentration values of different chemical elements or compounds, and already their unit (mk/kg, ppm, etc.) reflects that the information refers to a certain total. Analyzing "relative" information is the purpose of compositional data analysis, and the most prominent methods in this area are based on the so-called log-ratio methodology.

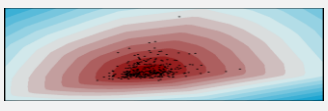
Particularly when analyzing the correlation between pairs of variables it is known that the traditional non-compositional methods can result in spurious correlations. This is caused by a tendency towards negative covariances, and thus the correlations loose their predicative value. There are different possibilities to avoid this problem. We will consider so-called symmetric pivot coordinates. Their purpose is to build an orthonormal coordinate system, expressing all relative information in terms of aggregated log-ratios, where the first two coordinates refer to all relative information of the two variables of interest. Since log-ratios could become instable in case of data quality problems, we propose an extension to weighted symmetric pivot coordinates, where appropriate weights can be assigned to the variables. The usefulness of these methods will be demonstrated with simulated data and data sets from environmental sciences.

Organisation

Data Science and Statistics section, Department of Mathematics, KU Leuven

Research Group ROBUST@Leuven

Financial support: Meet The Jury Fund



ROBUST@LEUVEN

YouReCa

KU LEUVEN