

Kynčlová, Petra; Hron, Karel; Filzmoser, Peter

Correlation between compositional parts based on symmetric balances. (English)

Zbl 1369.86020

Math. Geosci. 49, No. 6, 777-796 (2017).

Summary: Correlation coefficients are most popular in statistical practice for measuring pairwise variable associations. Compositional data, carrying only relative information, require a different treatment in correlation analysis. For identifying the association between two compositional parts in terms of their dominance with respect to the other parts in the composition, symmetric balances are constructed, which capture all relative information in the form of aggregated logratios of both compositional parts of interest. The resulting coordinates have the form of logratios of individual parts to a (weighted) “average representative” of the other parts, and thus, they clearly indicate how the respective parts dominate in the composition on average. The balances form orthonormal coordinates, and thus, the standard correlation measures relying on the Euclidean geometry can be used to measure the association. Simulation studies provide deeper insight into the proposed approach, and allow for comparisons with alternative measures. An application from geochemistry (Kola moss) indicates that correlations based on symmetric balances serve as a sensitive tool to reveal underlying geochemical processes.

MSC:

86A32 Geostatistics

Cited in 2 Documents

Keywords:

correlation analysis; compositional data; sequential binary partitioning; symmetric balances; logratio transformations

Software:

R

Full Text: DOI

References:

- [1] Aitchison J (1986) The statistical analysis of compositional data. Chapman and Hall, London · Zbl 0688.62004
- [2] Buccianti, A; Pawlowsky-Glahn, V, New perspectives on water chemistry and compositional data analysis, Math Geol, 37, 703-727, (2005) · Zbl 1103.62111
- [3] Chayes, F, On correlation between variables of constant sum, J Geophys Res, 65, 4185-4193, (1960) · Zbl 1236.34119
- [4] Eaton M (1983) Multivariate statistics. A vector space approach. Wiley, New York · Zbl 0587.62097
- [5] Egozcue, J, Reply to “on the harker variation diagrams; \$...\$” by J.A. cortés, Math Geosci, 41, 829-834, (2009) · Zbl 1178.86018
- [6] Egozcue J, Pawlowsky-Glahn V (2015) Proceedings of the 6th international workshop on compositional data analysis. In: Thió-Henestrosa S, Martín Fernández J (eds) Changing the reference measure in the simplex and its weighting effects, University of Girona, Girona, pp 1-10
- [7] Egozcue, JJ; Pawlowsky-Glahn, V, Groups of parts and their balances in compositional data analysis, Math Geol, 37, 795-828, (2005) · Zbl 1177.86018
- [8] Egozcue, JJ; Pawlowsky-Glahn, V; Mateu-Figueras, G; Barceló-Vidal, C, Isometric logratio transformations for compositional data analysis, Math Geol, 35, 279-300, (2003) · Zbl 1302.86024
- [9] Egozcue JJ, Lovell D, Pawlowsky-Glahn V (2013) Testing compositional association. In: Hron K, Filzmoser P, Templ M (eds) Proceedings of the 5th International Workshop on Compositional Data Analysis. Vorau, Austria
- [10] Filzmoser P, Hron K (2015) Robust coordinates for compositional data using weighted balances. In: Nordhausen K, Taskinen S (eds) Modern nonparametric. Robust and multivariate Methods. Springer, Heidelberg · Zbl 1135.62040
- [11] Filzmoser, P; Hron, K; Reimann, C, Univariate statistical analysis of environmental (compositional) data: problems and possibilities, Sci Total Environ, 407, 6100-6108, (2009)
- [12] Filzmoser, P; Hron, K; Reimann, C, The bivariate statistical analysis of environmental (compositional) data, Sci Total Environ,

408, 4230-4238, (2010) · [Zbl 1191.39011](#)

- [13] Fišerová, E; Hron, K, On interpretation of orthonormal coordinates for compositional data, *Math Geosci*, 43, 455-468, (2011)
- [14] Johnson RA, Wichern DW (2007) *Applied multivariate statistical analysis*, 6th edn. Prentice Hall, Englewood · [Zbl 0987.62001](#)
- [15] Korhoňová, M; Hron, K; Klimčíková, D; Müller, L; Bednář, P; Barták, P, Coffee aroma—statistical analysis of compositional data, *Talanta*, 80, 710-715, (2009)
- [16] McKinley, J; Hron, K; Grunsky, E; Reimann, C; Caritat, P; Filzmoser, P; Boogaart, K; Tolosana-Delgado, R, The single component geochemical map: fact or fiction, *J Geochem Explor*, 162, 16-28, (2016)
- [17] Pawlowsky-Glahn V, Buccianti A (2011) *Compositional data analysis: theory and applications*. Wiley, Chichester · [Zbl 1103.62111](#)
- [18] Pawlowsky-Glahn, V; Egozcue, JJ, Geometric approach to statistical analysis on the simplex, *Stoch Environ Res Risk Assess (SERRA)*, 15, 384-398, (2001) · [Zbl 0987.62001](#)
- [19] Pawlowsky-Glahn V, Egozcue J, Tolosana-Delgado R (2015) *Modeling and analysis of compositional data*. Wiley, Chichester
- [20] Pearson K (1897) Mathematical contributions to the theory of evolution. On a form of spurious correlation which may arise when indices are used in the measurement of organs. *Proc R Soc Lond LX*:489-502 · [Zbl 28.0209.02](#)
- [21] R Development Core Team (2015) *R: a language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna
- [22] Reimann C, Åyräs M, VC, et al (1998) *Environmental geochemical Atlas of the Central Barents Region*. NGU-GTK-CKE Special publication, Geological Survey of Norway, Trondheim, Norway
- [23] Reimann C, Filzmoser P, Fabian K, Hron K, Birke M, Demetriades A, Dinelli E, Ladenberger A, The GEMAS Project Team (2012) The concept of compositional data analysis in practice. Total major element concentrations in agricultural and grazing land soils of Europe. *Sci Total Environ* 426:196-210

This reference list is based on information provided by the publisher or from digital mathematics libraries. Its items are heuristically matched to zbMATH identifiers and may contain data conversion errors. It attempts to reflect the references listed in the original paper as accurately as possible without claiming the completeness or perfect precision of the matching.