

Egozcue, J. J.; Pawlowsky-Glahn, V.

Groups of parts and their balances in compositional data analysis. (English) Zbl 1177.86018
Math. Geol. 37, No. 7, 795–828 (2005).

Summary: Amalgamation of parts of a composition has been extensively used as a technique of analysis to achieve reduced dimension, as was discussed during the CoDaWork'03 meeting (Girona, Spain, 2003). It was shown to be a nonlinear operation in the simplex that does not preserve distances under perturbation. The discussion motivated the introduction in the present paper of concepts such as group of parts, balance between groups, and sequential binary partition, which are intended to provide tools of compositional data analysis for dimension reduction. Key concepts underlying this development are the established tools of subcomposition, coordinates in an orthogonal basis of the simplex, balancing element and, in general, the Aitchison geometry in the simplex. Main new results are: a method to analyze grouped parts of a compositional vector through the adequate coordinates in an *ad hoc* orthonormal basis; and the study of balances of groups of parts (inter-group analysis) as an orthogonal projection similar to that used in standard subcompositional analysis (intra-group analysis). A simulated example compares results when testing equal centers of two populations using amalgamated parts and balances; it shows that, in certain circumstances, results from both analysis can disagree.

MSC:

86A32 Geostatistics

Cited in 48 Documents

Keywords:

simplex; Euclidean geometry; log-ratio analysis; orthogonal projection; subcomposition; amalgamation

Full Text: [DOI](#)

References:

- [1] Aitchison, J., 1982, The statistical analysis of compositional data (with discussion): *J. R. Stat. Soc. B (Stat. Methodol.)*, v. 44, no. 2, p. 139–177. · [Zbl 0491.62017](#)
- [2] Aitchison, J., 1992, On criteria for measures of compositional difference: *Math. Geol.*, v. 24, no. 4, pp. 365–379. · [Zbl 0970.86531](#) · [doi:10.1007/BF00891269](#)
- [3] Aitchison, J., 2003a, The statistical analysis of compositional data (reprint): Blackburn Press, Caldwell, NJ, 416 p.
- [4] Aitchison, J., 2003b, Compositional data analysis: Where are we and where should we be heading? See Thió-Henestrosa and Martín-Fernández (2003).
- [5] Aitchison, J., Barceló-Vidal, C., Egozcue, J. J., and Pawlowsky-Glahn, V., 2002, A concise guide for the algebraic–geometric structure of the simplex, the sample space for compositional data analysis, in Bayer U., Burger H., and Skala W., eds., *Proceedings of IAMG'02–The Eighth Annual Conference of the International Association for Mathematical Geology*, Terra Nostro, no. 3, p. 387–392.
- [6] Barceló-Vidal, C., 2000, *Fundamentación matemática del análisis de datos composicionales*: Technical Report IMA 00-02-RR, Departament d'Informàtica i Matemàtica Aplicada, Universitat de Girona, Spain, 77 p.
- [7] Barceló-Vidal, C., Martín-Fernández, J. A., and Pawlowsky-Glahn, V., 2001, Mathematical foundations of compositional data analysis, in Ross G., ed., *Proceedings of IAMG'01–The Sixth Annual Conference of the International Association for Mathematical Geology*, CO-ROM, 20 p. · [Zbl 1101.86310](#)
- [8] Billheimer, D., Guttorp, P., and Fagan, W., 2001, Statistical interpretation of species composition: *J. Am. Stat. Assoc.*, v. 96, p. 1205–1214. · [Zbl 1073.62573](#) · [doi:10.1198/016214501753381850](#)
- [9] Egozcue, J. J., Pawlowsky-Glahn, V., Mateu-Figueras, G., and Barceló-Vidal, C., 2003, Isometric logratio transformations for compositional data analysis: *Math. Geol.*, v. 35, no. 3, p. 279–300. · [Zbl 1302.86024](#) · [doi:10.1023/A:1023818214614](#)
- [10] Mateu-Figueras, G., 2003, *Models de distribució sobre el símplex*. Ph.D. thesis, Universitat Politècnica de Catalunya, Barcelona, Spain.
- [11] Pawlowsky-Glahn, V., 2003, Statistical modelling on coordinates. See Thió-Henestrosa and Martín-Fernández (2003).
- [12] Pawlowsky-Glahn, V. and Egozcue, J. J., 2001, Geometric approach to statistical analysis on the simplex: *Stochastic Environ. Res. Risk Assess. (SERRA)*, v. 15, no. 5, p. 384–398. · [Zbl 0987.62001](#) · [doi:10.1007/s004770100077](#)
- [13] Pawlowsky-Glahn, V. and Egozcue, J. J., 2002, BLU estimators and compositional data: *Math. Geol.*, v. 34, no. 3, p. 259–274.

- [14] Thió-Henestrosa, S. and Martín-Fernández, J. A., eds., 2003, Compositional Data Analysis Workshop–CoDaWork'03, Proceedings, Universitat de Girona, CD-ROM, ISBN 84-8458-111-X, <http://ima.udg.es/Activitats/CoDaWork03/>.

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.