Towards compositional multifractal analysis

R. Tolosana-Delgado[1], K.G. van den Boogaart[1,2], and J. McKinley[3]
[1]Helmholtz Zentrum Dresden Rossendorf, Helmholtz Institute Freiberg for Resource Technology, Germany; *r.tolosana@hzdr.de*
[2]Institute for Stochastics, Technische Universitaet Bergakademie Freiberg, Germany
[3]Queen's University Belfast, UK

Abstract

Traditionally, multifractal analysis on a set E seeks to characterize the way a *measure* $\nu(A)$ of a subset $A \in E$ depends on another reference measure $\lambda(A)$, which is some sort of power law,

$$\frac{\nu(A)}{\lambda(A)} \propto \lambda(A)^{-\alpha}.$$

In the study of geochemical data, $\nu(A)$ is typically the *total mass* of a certain chemical element on a region A, while $\lambda(A)$ is the *area* of that region. Often, the ratio $\nu(A)/\lambda(A) = c_\nu(A)$ is understood as an average *concentration* of the element on the area. In these cases, $-\alpha$ is often estimated by the slope of the regression curve of $\log c_\nu(A)$ vs. $\log \lambda(A)$.

Note that both the mass and the area are measures, as required by fractal theory, but the concentration (the *data* actually available) is not. From the point of view of Compositional Data Analysis (CoDA), concentrations of individual elements cannot be univocally interpreted, and they have to be related to concentrations of other elements. It thus appears sensible to consider possible alternatives to conventional fractal analysis which take some or all principles of CoDA into account.

A prior step of such a construction is the obtention of meaningful measures out of the concentrations of the several chemical elements considered, $\{Z_1, Z_2, \ldots, Z_D\}$. Denoting by $z_i(x_n)$ the regionalized concentration of element Z_i at location x_n, one can construct an estimate of the measure of the mass of element Z_i at an area A as $\nu_i(A) \approx \sum_{x_n \in A} z_i(x_n)\rho(x_i)\Delta x_i$, where $\rho(x_i)$ is the average areal density of the D chemical elements considered and Δx_i is the area attached to location x_i. However, for the sake of simplicity and because of lack of information, we have considered these quantities as constants, such that $\nu_i(A) \propto \sum_{x_n \in A} z_i(x_n)$.

Once these are available, the question is which two measures are we going to contrast. Contrary to classical multifractal analysis, we propose to study the exponents α_{ij} of

$$\frac{\nu_i(A)}{\nu_j(A)} \propto \nu_j(A)^{-\alpha_{ij}},$$

which could be estimated by looking at plots of $\log(\nu_i(A)/\nu_j(A))$ vs. $\log(\nu_j(A))$, i.e. making use of logarithms and pairwise logratio transforms. However, given that both ordinate and abscisa are now random, it might be more appropriate to estimate α_{ij} via Total Least Squares methods instead of via conventional regression. The proposed method is, partly, scale invariant and subcompositionally coherent. Scale invariance requires all data to be scaled by the same constant, and subcompositional coherence that the condition of constant density $\rho(x_i)$ is preserved. The result is a whole matrix $\mathbf{A}$ of $D \times D$ exponents to be mapped. First experiments show that maps of such exponents inform of the locations at which component X_i systematically and strongly depart from proportionality with respect to component X_j. Depending on which regression method is used, though, the slopes α_{ij} and α_{ji} are not necessarily related, and the matrix does not exhibit any particular mathematical structure.

To check the appropriateness and interpretability of the method proposed, the Tellus data set has been used, covering Northern Ireland (UK) at a sampling density of 1 sample per 2 km^2. Exponents α_{ij} were calculated by taking samples within concentric circles of decreasing area A around each sample, in a sort of moving window. Results have been obtained for those elements without missing values, unveiling some interesting patterns.