Paper Number: 1437

Latent endmember identification by means of Gaussian anamorphosis

Tolosana-Delgado, R.[1], van den Boogaart, K.G. [12] and Mueller, U.[3]

[1]Helmholtz Zentrum Dresden-Rossendorf, Helmholtz Institute Freiberg for Resource Technology, Halsbrückerstr. 34, D-09599 Freiberg, Germany, r.tolosana@hzdr.de
[2]TU Bergakademie Freiberg, Institute for Stochastics, Prüferstr. 9, D-09699 Freiberg, Germany
[3]Edith Cowan University, 270 Joondalup Drive; Joondalup, WA6027, Australia

---

In a geological context compositions are often generated by variable mixtures of different source materials. These sources are called endmembers of the mixture [1,2]. This contribution presents a data-driven exploratory method of detection of the presence of endmembers, and preliminary identification of their possible composition.

The key assumptions of the method are: the P endmembers have a unknown fixed compositions $\mathbf{x}_1$, $\mathbf{x}_2$,…, $\mathbf{x}_P$; each composition $\mathbf{x}$ is a convex linear mixture of the endmembers, $\mathbf{x}=w_1\mathbf{x}_1+ w_2\mathbf{x}_2+\cdots+ w_P\mathbf{x}_P$; and the vector of mixture weights $\mathbf{w}=[w_1,w_2,\ldots,w_P]$ follows an additive logistic normal distribution [3]. With these assumptions, the method is as follows.

First, the initial compositions are logratio transformed, for instance with the so called additive logratio transform [3] or with the isometric logratio transform [4]. This is necessary in order to account for the fact that compositions convey only relative information; it is important to note that in the space of logratio scores the generating additive mixture of endmembers is no longer linear. Second, a multidimensional flow Gaussian anamorphosis [5,6] is applied to the logratio scores. This is a deformation of the space of the logratios that seeks to place the data in such a way that a kernel density estimate of the distribution of the data will produce a multivariate normal distribution. Because of the extreme character of the endmembers, they must be placed on areas of intense deformation. We propose to highlight these areas by: generating a regular grid on the Gaussian space covering a region with 99.9% probability; transforming the grid with the inverse flow anamorphosis, and using a kernel density estimate on the backtransformed grid.

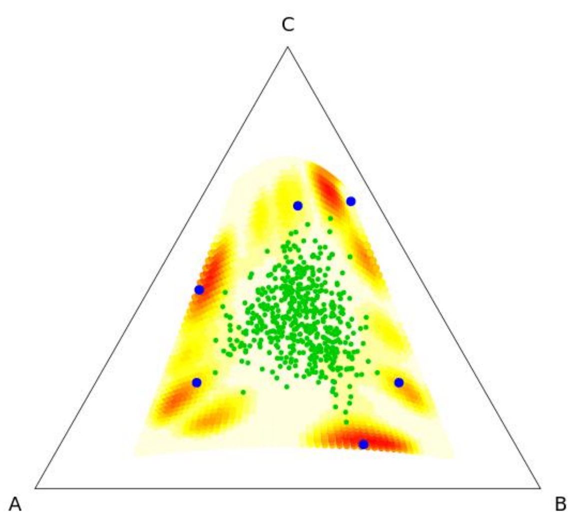

Figure 1: a simulated case with 6 endmembers (blue dots), 500 compositions of 3 parts (green dots) and identified areas of high deformation (background color scale: the intensity of the deformation is indicated by the hue of red: the darker the hue, the stronger the deformation)

Promising results with a simulated case study are shown in Fig. 1. A data set of 500 compositions of three parts was obtained by a linear mixture of six endmembers, according to the assumptions listed above. Out of the six endmembers used for the simulation, three were detected perfectly (dark red areas), two were identified adequately (orange areas), and the last one was only slightly suggested (intense yellow areas).

References:

[1] Weltje GJ (1997) Math Geol 29: 503-549

[2] Tolosana-Delgado R et al. (2011) Comp Geosci 37(5):677-691

[3] Aitchison J (1986) *The Statistical Analysis of Compositional Data*, Chapman Hall

[4] Egozcue JJ et al. (2003) Math Geosci 35: 279-300

[5] van den Boogaart KG et al (2015) In: *Proc Ann Conf IAMG* Freiberg (Germany), pp 1302-1311

[6] Tolosana-Delgado R et al. (2016) Math Geosci, submitted