Outlier Detection for Compositional Data Using Robust Methods

Peter Filzmoser[1] **and Karel Hron**[2]

Outlier detection based on the Mahalanobis distance (MD) requires an appropriate transformation in case of compositional data. For the family of logratio transformations (additive, centered and isometric logratio transformation) it is shown that the MDs based on classical estimates are invariant to these transformations, and that the MDs based on affine equivariant estimators of location and covariance are the same for additive and isometric logratio transformation. Moreover, for 3-dimensional compositions the data structure can be visualized by contour lines, and in higher dimension the MDs of closed and opened data give an impression of the multivariate data behavior.

KEY WORDS: Mahalanobis distance, robust statistics, ternary diagram, multivariate outliers, logratio transformation.

INTRODUCTION

Outlier detection is one of the most important tasks in multivariate data analysis. The outliers give valuable information on data quality, and they are indicative of atypical phenomena. Although a comprehensive literature exists on outlier detection (e.g. Rousseeuw and Leroy, 2003; Maronna, Martin, and Yohai, 2006), also in the context of geochemical data (e.g. Filzmoser, Garett, and Reimann, 2005), further research is needed for outlier detection in the context of compositional data (e.g. Barceló, Pawlowsky, and Grunsky, 1996). Compositional or closed data sum up to a constant value (Aitchison, 1986). This constraint makes it necessary to first transform the data to an unconstrained space where standard statistical methods can be used. One of the most convenient transformations is the family of logratio transformations (see Aitchison, 1986).

[1] Dept. of Statistics and Probability Theory, Vienna University of Technology, Wiedner Hauptstr. 8-10, 1040 Vienna, Austria; e-mail: P.Filzmoser@tuwien.ac.at

[2] Dept. of Mathematical Analysis and Applications of Mathematics, Palacký University Olomouc, Tomkova 40, CZ-77100 Olomouc, Czech Rep.; e-mail: hronk@seznam.cz

However, it is not clear if different transformations will lead to different answers for identifying outliers. In this paper we will consider three well known transformations, the additive, the centered, and the isometric logratio transformation. The next section will provide a brief overview about their formal definitions, and the definitions of the inverse transformations. Furthermore, we will discuss multivariate outlier detection methods, as they are used for unconstrained multivariate data. Our focus here is on "standard" methods for outlier detection that are widely used and implemented in statistical software packages. The link between outlier detection and the different types of logratio transformations is made in the section afterwards. In contrast to Barceló, Pawlowsky, and Grunsky (1996) where only the additive logratio transformation is considered for outlier detection, this section provides theoretical results on the equivalence of the additive, the centered, and the isometric logratio transformation in the context of outlier identification. In the case of 3-dimensional compositional data we show how the multivariate data structure can be viewed in the ternary diagram and how multivariate outliers are highlighted. For higher dimensional compositions a plot is introduced that is useful for revealing multivariate outliers.

COMPOSITIONAL DATA AND TRANSFORMATIONS

Compositional or closed data are multivariate data with positive values that sum up to a constant, usually chosen as 1, i.e.

$$\mathbf{x} = (x_1, \ldots, x_D)',\ x_i > 0,\ \sum_{i=1}^{D} x_i = 1.$$

The set of all closed observations, denoted as $\mathcal{S}^D$, forms a simplex sample space, a subset of $\mathbb{R}^D$. Convenient operations on the simplex and their properties for dealing with compositions were summarized in Aitchison and Egozcue (2005). In practice, standard statistical methods can lead to questionable results if they are directly applied to the original, closed data. For this reason, the family of logratio one-to-one transformations from $\mathcal{S}^D$ to the real space was introduced (Aitchison, 1986). We will briefly review these transformations as well as their inverse counterparts:

Additive logratio (alr) transformation: This is a transformation from $\mathcal{S}^D$ to $\mathbb{R}^{D-1}$, and the result for an observation $\mathbf{x} \in \mathcal{S}^D$ are the transformed data

$\mathbf{x}^{(j)} \in \mathbb{R}^{D-1}$ with

$$\mathbf{x}^{(j)} = (x_1^{(j)}, \ldots, x_{D-1}^{(j)})' = \left(\log \frac{x_1}{x_j}, \ldots, \log \frac{x_{j-1}}{x_j}, \log \frac{x_{j+1}}{x_j}, \ldots, \log \frac{x_D}{x_j} \right)' . \qquad (1)$$

The index $j \in \{1, \ldots, D\}$ refers to the variable that is chosen as ratioing variable in the transformation. This choice usually depends on the context, but also on the suitability of the results for visualization and data exploration. The main advantage of the alr transformation is that it opens compositional data into an unconstrained form in the real space. The **inverse alr** transformation from $\mathbb{R}^{D-1}$ to $\mathcal{S}^D$, also called *additive logistic transformation*, is defined as

$$\begin{aligned} x_i &= \frac{\exp\left(x_i^{(j)}\right)}{\exp\left(x_1^{(j)}\right) + \ldots + \exp\left(x_{D-1}^{(j)}\right) + 1} \qquad \text{for } i = 1, \ldots, D,\ i \neq j, \\ x_j &= \frac{1}{\exp\left(x_1^{(j)}\right) + \ldots + \exp\left(x_D^{(j)}\right) + 1} \qquad \text{for } j \in \{1, \ldots, D\}. \end{aligned} \qquad (2)$$

Centered logratio (clr) transformation: Compositions $\mathbf{x} \in \mathcal{S}^D$ are transformed to data $\mathbf{y} \in \mathbb{R}^D$, with

$$\mathbf{y} = (y_1, \ldots, y_D)' = \left(\log \frac{x_1}{\sqrt[D]{\prod_{i=1}^D x_i}}, \ldots, \log \frac{x_D}{\sqrt[D]{\prod_{i=1}^D x_i}} \right)' . \qquad (3)$$

It is easy to see that this transformation results in collinear data because $\sum_{i=1}^D y_i = 0$. On the other hand, the clr transformation treats all components symmetrically by dividing by the geometric mean. The interpretation of the resulting values might thus be easier. The **inverse clr** transformation is

$$x_i = \frac{\exp(y_i)}{\exp(y_1) + \ldots + \exp(y_D)} \qquad \text{for } i = 1, \ldots, D. \qquad (4)$$

Isometric logratio (ilr) transformation: This transformation solves the problem of data collinearity resulting from the clr transformation, while preserving all its advantageous properties (Egozcue and others, 2003). It is based on the choice of an orthonormal basis on the hyperplane in $\mathbb{R}^D$ that is formed by the clr transformation, so that the compositions $\mathbf{x} \in \mathcal{S}^D$ result in noncollinear data $\mathbf{z} \in \mathbb{R}^{D-1}$. The explicit transformation formulas for one such chosen basis are

$$\mathbf{z} = (z_1, \ldots, z_{D-1})',\ z_i = \sqrt{\frac{i}{i+1}} \log \frac{\sqrt[i]{\prod_{j=1}^i x_j}}{x_{i+1}} \qquad \text{for } i = 1, \ldots, D-1. \qquad (5)$$

The **inverse ilr** transformation is then obtained using (4) in which the terms

$$y_i = \sum_{j=i}^{D} \frac{z_j}{\sqrt{j(j+1)}} - \sqrt{\frac{i-1}{i}} z_{i-1} \quad \text{with } z_0 = z_D = 0 \quad \text{for } i = 1, \ldots, D \tag{6}$$

are substituted.

For all logratio transformations, the problem of values $x_i = 0$ is solvable in many ways, e.g. Martín-Fernández, Barceló-Vidal, and Pawlowsky-Glahn (2003).

OUTLIER DETECTION METHODS

In contrast to univariate outliers, multivariate outliers are not necessarily extreme along single coordinates. Rather, they could deviate from the multivariate data structure formed by the majority of observations. Basically, there are two different procedures to identify multivariate outliers: (a) methods based on projection pursuit and (b) methods based on the estimation of the covariance structure. The idea of (a) is to repeatedly project the multivariate data to the univariate space, because univariate outlier detection is much simpler (Gnanadesikan and Kettenring, 1972; Peña and Prieto, 2001; Maronna and Zamar, 2002). Such methods are usually computationally intensive, but they are particularly useful for high-dimensional data with low sample size. We will focus here on (b). The estimated covariance structure is used to assign a distance to each observation indicating how far the observation is from the center of the data cloud with respect to the covariance structure. This distance measure is the well-known Mahalanobis distance, defined for a sample $\mathbf{x}_1, \ldots, \mathbf{x}_n$ of n observations in the d-dimensional real space $\mathbb{R}^d$ as

$$\mathrm{MD}(\mathbf{x}_i) = \left[(\mathbf{x}_i - T)' C^{-1} (\mathbf{x}_i - T)\right]^{1/2} \qquad \text{for } i = 1, \ldots, n. \tag{7}$$

Here, T and C are location and covariance estimators, respectively.

The choice of the estimators T and C in (7) is crucial. In the case of multivariate normally distributed data, the arithmetic mean and the sample covariance matrix are the best choices, leading to the best statistical efficiency. In this case, the squared Mahalanobis distances approximate a chi-square distribution χ^2_d with d degrees of freedom. A certain cut-off value like the 97.5% quantile of χ^2_d can be taken as an indication of extremeness: data points with higher (squared) Mahalanobis distance than the cut-off value are considered as potential outliers (Rousseeuw and Van Zomeren, 1990).

Both the arithmetic mean and the sample covariance matrix are highly sensitive to outlying observations (Maronna, Martin, and Yohai, 2006). Therefore, using these estimators for outlier detection leads to questionable results. A number of robust counterparts have been proposed in the literature, like the MCD or S estimator (see Maronna, Martin, and Yohai, 2006). The resulting estimates of location and covariance also lead to robust estimates of the Mahalanobis distance (7). It is common to use the same cut-off value from the χ^2_d distribution (Rousseeuw and Van Zomeren, 1990), although other approximations could lead to more accurate cut-off values (Filzmoser, Garrett, and Reimann, 2005; Hardin and Rocke, 2005). Besides robustness properties the property of affine equivariance of the estimators T and C is important. The location estimator T and the covariance estimator C are called affine equivariant, if for any nonsingular $d \times d$ matrix $\mathbf{A}$ and for any vector $\mathbf{b} \in \mathbb{R}^d$ the conditions

$$\begin{aligned} T(\mathbf{A}\mathbf{x}_1 + \mathbf{b}, \ldots, \mathbf{A}\mathbf{x}_n + \mathbf{b}) &= \mathbf{A}T(\mathbf{x}_1, \ldots, \mathbf{x}_n) + \mathbf{b}, \\ C(\mathbf{A}\mathbf{x}_1 + \mathbf{b}, \ldots, \mathbf{A}\mathbf{x}_n + \mathbf{b}) &= \mathbf{A}C(\mathbf{x}_1, \ldots, \mathbf{x}_n)\mathbf{A}' \end{aligned}$$

are fulfilled. Thus, the estimators transform accordingly, and it can be easily seen that the Mahalanobis distances remain unchanged under regular affine transformations, i.e.

$$\mathrm{MD}(\mathbf{A}\mathbf{x}_i + \boldsymbol{b}) = \mathrm{MD}(\mathbf{x}_i) \qquad \text{for } i = 1, \ldots, n. \tag{8}$$

The identified outliers will thus be the same, independent of the choice of $\mathbf{A}$ and $\mathbf{b}$ for the transformation. The above mentioned robust MCD and S estimators share the property of affine equivariance.

PROPERTIES OF THE LOGRATIO TRANSFORMATIONS IN THE CONTEXT OF OUTLIER DETECTION

The usefulness of robust Mahalanobis distances for multivariate outlier detection has been demonstrated in the literature and in many applications (Maronna, Martin, and Yohai, 2006). However, this tool would not be appropriate for closed data but only for the data after transformation. Here the problem arises, which logratio transformation from the simplex to the real space is most suitable. An answer concerning the alr transformation is given by the following theorem. The proof to this theorem as well as the proofs to subsequent theorems can be found in the appendix.

Theorem 1 The Mahalanobis distances (MDs) for alr transformed data are invariant with respect to the choice of the ratioing variable if the location estimator T and the scatter estimator C are affine equivariant.

Theorem 1 thus guarantees that the identified outliers will not depend on the ratioing variable that has been chosen for the alr transformation, as long as the location and scatter estimators are taken to be affine equivariant.
A result for the clr transformation is given in the following theorem.

Theorem 2 The MDs for clr and alr transformed data are the same if the location estimator T is the arithmetic mean and the covariance estimator C is the sample covariance matrix.
The result of Theorem 2 is unsatisfactory from a robustness point of view, because the equality of the Mahalanobis distances is only valid for the non-robust estimators arithmetic mean and sample covariance matrix, but not for robust estimators like the MCD or S estimators which are not even computable for the clr transformed data. It should be noted that relations between the sample covariance matrices of alr and clr transformed data were already investigated in Aitchison (1986, Property 5.7), Aitchison (1992), Bohling and others (1998), and Barceló-Vidal, Martín-Fernández, and Pawlowsky-Glahn (1999). However, the results in the proof of this theorem are valuable also for finding the link to the ilr transformation, shown in the next theorem.

Theorem 3 The MDs for ilr transformed data are the same as in the case of alr transformation if the location estimator T and the covariance estimator C are affine equivariant.
This theorem completes the relations between the three mentioned transformations. In case of using the classical estimators arithmetic mean and sample covariance matrix, all three transformations lead to the same MDs. Since outlier detection is only reliable with robust estimates of location and covariance, the resulting robust MDs are the same for alr and ilr transformed data, if affine equivariant estimators are used. In the following we will use the MCD estimator for this purpose, because of the good robustness properties, and because of the fast algorithm for its computation (Rousseeuw and Van Driessen, 1999). The MCD (Minimum Covariance Determinant) estimator looks for a subset h out of n

observations with the smallest determinant of their sample covariance matrix. A robust estimator of location is the arithmetic mean of these observations, and a robust estimator of covariance is the sample covariance matrix of the h observations, multiplied by a factor for consistency at normal distribution. The subset size h can vary between half the sample size and n, and it will determine the robustness of the estimates, but also their efficiency. The clr transformation will not be considered in the following, since there exist no affine equivariant robust estimators of location and covariance that could be applied to the opened singular data.

NUMERICAL EXAMPLES

In this section we apply the theoretical results to real data examples. The first two examples are taken from Barceló, Pawlowsky, and Grunsky (1996), who applied outlier detection based on different additive logratio transformations combined with Box-Cox transformation. Since the closed data have 3 parts or components, we can even plot them in the ternary diagram. We additionally visualize the Mahalanobis distances in the ternary diagram which gives a better impression of the multivariate data structure.

Visualizing Mahalanobis distances in the ternary diagram

Let $\mathbf{p}_1, \ldots, \mathbf{p}_n$ be the opened (alr or ilr) transformed data in the 2-dimensional real space (i.e. the original closed data were in the space $\mathcal{S}^3$). Using an estimation of location T and covariance C based on the data $\mathbf{p}_1, \ldots, \mathbf{p}_n$, the Mahalanobis distances can be computed. Moreover, any other point $\mathbf{p} \in \mathbb{R}^2$ can be assigned a Mahalanobis distance using the same estimates T and C, i.e. $\mathrm{MD}(\mathbf{p}) = [(\mathbf{p} - T)' C^{-1} (\mathbf{p} - T)]^{1/2}$. Now we are interested in those points $\mathbf{p}_c \in \mathbb{R}^2$ that have the same constant Mahalanobis distance c, i.e. $\mathrm{MD}(\mathbf{p}_c) = c$. Using polar coordinates, it is easy to see that

$$\mathbf{p}_c = \boldsymbol{\Gamma} \begin{pmatrix} \sqrt{a_1} & 0 \\ 0 & \sqrt{a_2} \end{pmatrix} \begin{pmatrix} c \cdot \cos(2\pi \cdot m) \\ c \cdot \sin(2\pi \cdot m) \end{pmatrix} + T, \tag{9}$$

where $\boldsymbol{\Gamma} = (\boldsymbol{\gamma}_1, \boldsymbol{\gamma}_2)$ is the matrix with the eigenvectors of C, a_1 and a_2 are the associated eigenvalues, and m is any number in the interval $[0, 1)$. In particular, distances $c = \sqrt{\chi^2_{2;q}}$ will be of interest, for certain quantiles q, like the quantile 97.5% indicating the outlier cut-off value.

The points $\mathbf{p}_c$ can be back-transformed to the original space $\mathcal{S}^3$ by applying the corresponding inverse transformation, i.e. formula (2) if an alr transformation has been applied, or formulas (4) and (6) in case of a clr and ilr transformation. The resulting back-transformed points can be drawn as contours in the ternary diagram.

Example 1 (Arctic Lake Sediment data)
This data set from Aitchison (1986, p. 359) describes 39 sediment samples of sand, silt and clay compositions in an Arctic lake, and comes originally from Coakley and Rust (1968). The ternary diagram shown in Figure 1 (lower left and right) reveals deviating data points. However, in this display it is not clear which data points belong to a joint data structure and which points are deviating from this structure. We open the data with alr transformation, using the second variable as ratioing variable (Figure 1, upper left and right). Now the real bivariate data structure is immediately visible. We compute the classical MDs using sample mean and covariance, and the robust MDs using the MCD estimator. The plots are overlaid using (9) with the ellipses corresponding to quantiles 0.75, 0.9, and 0.975 of $\sqrt{\chi^2_2}$ for the classical (left) and robust (right) estimators. While classical estimation only reveals two observations as outliers, robust estimation discovers the data structure of the majority of the data points in a much better way and thus highlights additional points as potential outliers. Back-transformation of the ellipses to the original data space results in the contours shown in Figure 1 lower left (classical) and lower right (robust). Of course, the same data points as in the above plots are flagged as outliers. Additionally, the robust contours make the main data structure visible (right ternary diagram). Note that the contours would be exactly the same if another variable had been used as ratio variable (Theorem 1), or if an ilr transformation had been used (Theorem 3), or if a clr transformation had been used for the classical case (Theorem 2).

Figure 1 about here

Barceló, Pawlowsky, and Grunsky (1996) also used these data for outlier detection. The authors used a very different procedure (alr and different Box-Cox transformations), and the observations 6, 7, 12 and 14 were identified as potential outliers. Our approach flagged the same observations as atypical, but also some additional data points. The visual impression in the transformed space (Figure 1, upper right) confirms our findings. It should be noted that the representation of the alr transformed data with orthogonal coordinates in Figure 1 (upper left and right) is not coherent with the Aitchison geometry of the simplex (Egozcue and others, 2003). Nevertheless, the results concerning outlier detection are correct.

Example 2 (Aphyric Skye Lavas data)
The data in Aitchison (1986, p. 360), adapted from Thompson, Esson, and Duncan (1972), represent percentages of $Na_2O + K_2O$ (A), Fe_2O_3 (F) and MgO (M) in 23 aphyric Skye lavas and define compositions with sum 100%. Here we apply the ilr transformation and compute classical and robust MDs. The graphical representation of the results is analogous to Figure 1: The upper row of Figure 2 shows the ilr transformed data with ellipses corresponding to classical (left) and robust (right) MDs, and the lower row of Figure 2 shows the original data in the ternary diagram, with the ellipses (classical: left; robust: right) back-transformed. Only the robust analysis identifies two potential outliers: the observations 2 and 3.

Figure 2 about here

For this data set, Barceló, Pawlowsky, and Grunsky (1996) did not report any outliers. Note that the two observations 2 and 3 identified as potential outliers with our method are really on the boundary. If another outlier cut-off value would be used, these observations could fall inside the boundary. In practice, a more detailed inspection of the two atypical data points is recomended.

Example 3 (Kola data)
This data set comes from a large geochemical mapping project, carried out from 1992 to 1998 by the Geological Surveys of Finland and Norway, and the Central Kola Expedition, Russia. An area covering 188000 km^2 at the peninsula Kola in Northern Europe was sampled. In total, around 600 samples of soil were taken in 4 different layers (moss, humus, B-horizon, C-horizon), and subsequently analyzed by a number of different techniques for more than 50 chemical elements. The project was primarily designed to reveal the environmental conditions in the area. More details can be found in Reimann and others (1998) which also includes maps of the single element distributions. The data are available in the library "mvoutlier" of the statistical software package R (R development core team, 2006). Here we use the 10 major elements Al, Ca, Fe, K, Mg, Mn, Na, P, Si and Ti of the C-horizon for multivariate outlier detection. We apply the ilr transformation to open the data.

For this example it is no longer possible to use ternary diagrams for graphical inspection. However, we still can compute the Mahalanobis distances and show them graphically, together with an outlier cut-off value. It could be interesting to see the effect of robust versus classical estimation of the Mahalanobis distances. Figure 3 shows the distance-distance plot introduced in Rousseeuw and Van Driessen (1999), comparing both measures. The robust Mahalanobis distances are based on MCD estimates. The outlier cut-off values are the quantiles 0.975 of $\sqrt{\chi^2_9}$, and are shown as the horizontal and vertical line. The dashed line indicates equal distance measures.

Using the outlier cut-off values, the plot can be subdivided into 4 quadrants: regular observations (lower left; symbol grey dot), outliers (upper right; symbol "+"), outliers only identified with the classical MD (empty), and outliers only identified with the robust MD (symbol triangle). Figure 3 (right) shows the map of the survey area. The same symbols as used on the left plot are plotted at the sample locations. The multivariate outliers marked with "+" are in the northern costal area and in the east around Monchegorsk, a big industrial center, and Apatity (compare Filzmoser, Garrett, and Reimann, 2005). However, the additional multivariate outliers identified with the robust method (symbol triangle) emphasize the atypical regions in a much clearer way, and additionally highlight an area left from the center of the survey area. This area is characterized by a felsic/mafic granulite belt (compare Reimann and others, 1998) which obviously has deviating multivariate data behavior.

Figure 3 about here

Figure 3 makes the necessity of robust estimation clear. Besides robust estimation it could also be interesting to see the effect of opening the data for outlier detection. Figure 4 is a modification of the distance-distance plot, we plot the robust Mahalanobis distances of the closed original data against the robust Mahalanobis distances of the ilr transformed data. The horizontal lines are the outlier cut-off values, namely the quantiles 0.975 of $\sqrt{\chi^2_{10}}$ and $\sqrt{\chi^2_9}$, respectively. As before we can split the plot into 4 quadrants, and we use different symbols in each quadrant. Additionally, for the observations identified as multivariate outliers by both distance measures (upper right; symbol "+") we use black and gray symbols, depending on which distance measure is larger.

Figure 4 (right) shows the same symbols in the map. We see that the multivariate outliers characterize much the same areas as in Figure 3, but the measure based on the closed data would miss many outliers in the center of the survey area (symbol triangle). The outliers only identified with the closed data (symbol open circle) seem to make no sense at all, because they form no spatial pattern in the map. Interestingly, the distinction in size of the outliers identified with both measures (symbol "+", black and gray) allows also a geographical distinction. The gray symbols are mainly around Monchegorsk and Apatity in the east, and they are over-emphasized by resulting in too large distances, if the data are not opened.

Figure 4 about here

CONCLUSIONS

Robust Mahalanobis distances are a very common tool for multivariate outlier detection. However, in case of compositional data the application of this tool to the closed data can lead to unrealistic results. As a way out, different data transformations like the alr, clr, or ilr transformation should be applied first. We have shown that all three transformations result in the same Mahalanobis distances if classical estimates are used. If a robust affine equivariant estimator (like the MCD estimator) is used, the Mahalanobis distances are the same for alr and ilr transformed data. The data used in Examples 1 and 2 allow a visualization of the Mahalanobis distances in the ternary plot as contour lines, making the multivariate data structure clearer visible. For data of higher dimension the visualization can be done by comparing Mahalanobis distances of the original (closed) and the opened data.

It should be noted that outlier detection based on robust Mahalanobis distances implicitly assumes that the majority of data points is elliptically symmetric. If the transformation for opening the data does not approach this elliptical symmetry, an additional data transformation should be applied. In fact, this was proposed in Barceló, Pawlowsky, and Grunsky (1996) who used a Box-Cox transformation of the data. However, nice theoretical properties are then lost, and it will again depend on the type of transformation which observations are identified as potential outliers. A way out of this situation is to use covariance estimators which are less sensitive to deviations from elliptical symmetry, like estimators based on spatial signs or ranks (Visuri, Koivunen, and Oja, 2000). For 3-dimensional compositional data the elliptical symmetry can be graphically inspected by visualizing the Mahalanobis distances in the transformed data space (Figures 1 and 2, upper right).

Finally, we would like to point out that the critical outlier cut-off value used in this paper only indicates *potential* outliers, but it should not be used to automatically declare these observations as outliers. These observations are different from the majority of data points. The reason for this difference could be a different process influencing the data (another data distribution), or atypically high or low values causing "extreme" observations (same data distribution). Filzmoser, Garrett, and Reimann (2005) discussed this issue, and introduced modified cut-off values to distinguish between these types of outliers.

ACKNOWLEDGEMENTS

The authors are grateful to the referees for helpful comments and suggestions. Moreover, Professor Kubáček (Palacky University, Olomouc) and Professor Antoch (Charles University of Prague) are thanked for their support and valuable comments.

REFERENCES

Aitchison, J., 1986, The statistical analysis of compositional data. Monographs on statistics and applied probability: Chapman and Hall, London, 416 p.

Aitchison, J., 1992, On criteria for measures of compositional difference: Math. Geol., vo. 24, no. 4, p. 365-379.

Aitchison, J., Egozcue, J.J., 2005, Compositional data analysis: Where are we and where should we be heading?: Math. Geol., vo. 37, no. 7, p. 829-850.

Barceló, C., Pawlowsky, V., Grunsky, E., 1996, Some aspects of transformations of compositional data and the identification of outliers: Math. Geol., vo. 28, no. 4, p. 501-518.

Barceló-Vidal, C.B., Martín-Fernandez, J.A., Pawlowsky-Glahn, V., 1999, Comment on "Singularity and nonnormality in the classification of compositional data" by G.C. Bohling, J.C. Davis, R.A. Olea, and J. Harff (Letter to the editor): Math. Geol., vo. 31, no. 5, p. 581-585.

Bohling, G.C., Davis, J.C., Olea, R.A., Harff, J., 1998, Singularity and nonnormality in the classification of compositional data: Math. Geol., vo. 30, no. 1, p. 5-20.

Coakley, J.P. , Rust, B.R., 1968, Sedimentation in an Arctic lake: Jour. Sed. Pet., vo. 38, no. 4, p. 1290-1300: Quoted in Aitchison, 1986, The statistical analysis of compositional data: Chapman and Hall, London, 416 p.

Egozcue, J.J., Pawlowsky-Glahn, V., Mateu-Figueras, G., Barceló-Vidal, C., 2003, Isometric logratio transformations for compositional data analysis: Math. Geol., vo. 35, no. 3, p. 279-300.

Filzmoser, P., Garrett, R.G., Reimann, C., 2005, Multivariate outlier detection in exploration geochemistry: Computers and Geosciences, vo. 31, p. 579-587.

Gnanadesikan, R., Kettenring, J.R., 1972, Robust estimates, residuals, and outlier detection with multiresponse data: Biometrics, vo. 28, p. 81-124.

Hardin, J., Rocke, D.M., 2005, The distribution of robust distances: Journal of Computational and Graphical Statistics, vo. 14, p. 928-946.

Harville, D.A., 1997, Matrix algebra from a statistican's perspective: Springer-Verlag, New York, 630 p.

Maronna, R., Martin, R.D., Yohai, V.J, 2006, Robust statistics: Theory and methods. John Wiley, New York, 436 p.

Maronna, R., Zamar, R., 2002, Robust estimates of location and dispersion for high-dimensional data sets: Technometrics, vo. 44, no. 4, p. 307-317.

Martín-Fernández, J.A., Barceló-Vidal, C., Pawlowsky-Glahn, V., 2003, Dealing with zeros and missing values in compositional data sets using nonparametric imputation: Math. Geol., vo. 35, no. 3, p. 253-278.

Peña, D., Prieto, F., 2001, Multivariate outlier detection and robust covariance matrix estimation: Technometrics, vo. 43, no. 3, p. 286-310.

R development core team, 2006, R: A language and environment for statistical computing: Vienna, `http://www.r-project.org`.

Reimann, C., Äyräs, M., Chekushin, V., Bogatyrev, I., Boyd, R., Caritat, P. d., Dutter, R., Finne, T., Halleraker, J., Jæger, O., Kashulina, G., Lehto, O., Niskavaara, H., Pavlov, V., Räisänen, M., Strand, T., Volden, T., 1998, Environmental geochemical atlas of the Central Barents Region: Geological Survey of Norway (NGU), Geological Survey of Finland (GTK), and Central Kola Expedition (CKE), Special Publication, Trondheim, Espoo, Monchegorsk, 745 p.

Rousseeuw, P.J., Leroy, A.M., 2003, Robust regression and outlier detection: Wiley & Sons, New York, 360 p.

Rousseeuw, P., Van Driessen, K., 1999, A fast algorithm for the minimum covariance determinant estimator: Technometrics, vo. 41, p. 212-223.

Rousseeuw, P.J., Van Zomeren, B.C., 1990, Unmasking multivariate outliers and leverage points: Journal of the American Statistical Association, vo. 85, no. 411, p. 633-651.

Thompson, R.N., Esson, J., Duncan, A.C., 1972, Major element chemical variation in the Eocene lavas of the Isle of Skye, Scotland: Jour. Petrology, vo. 13, no. 2, p. 219-253: Quoted in Aitchison, J., 1986, The statistical analysis of compositional data: Chapman and Hall, London, 416 p.

Visuri, S., Koivunen, V., Oja, H., 2000, Sign and rank covariance matrices: Journal of Statistical Planning and Inference, vo. 91, p. 557-575.

APPENDIX

Proof of Theorem 1 Let $\mathbf{X}_{n,D}$ be a data matrix with closed observations $\mathbf{x}_i = (x_{i1}, \ldots, x_{iD})'$ with $\sum_{j=1}^{D} x_{ij} = 1$ and $x_{ij} > 0$ for $i = 1, \ldots, n$, i.e. $\mathbf{x}_i \in \mathcal{S}^D$. Let $\mathbf{X}^{(l)}_{n,D-1}$ be matrix resulting from alr transformation of $\mathbf{X}$ using column l, i.e. the rows of $\mathbf{X}^{(l)}$ are

$$\mathbf{x}_i^{(l)} = \left(\log \frac{x_{i1}}{x_{il}}, \ldots, \log \frac{x_{i,l-1}}{x_{il}}, \log \frac{x_{i,l+1}}{x_{il}}, \ldots, \log \frac{x_{iD}}{x_{il}}\right)' \tag{10}$$

(compare with (1)). Similarly, let $\mathbf{X}^{(k)}$ be the alr transformed data matrix from $\mathbf{X}$ using column k, with $k \neq l$. Then, using $\log \frac{x_{ij}}{x_{il}} = \log x_{ij} - \log x_{il}$, it can be easily shown that $\mathbf{X}^{(l)} = \mathbf{X}^{(k)}\mathbf{B}_{kl}$ or $\mathbf{x}_i^{(l)} = \mathbf{B}'_{kl}\mathbf{x}_i^{(k)}$ with the $(D-1) \times (D-1)$ matrix

$$\mathbf{B}_{kl} = \left(\begin{array}{cccccccccc}
1 & & & & & & 0 & & & \\
 & \ddots & & & & & \vdots & & & \\
 & & 1 & & & & 0 & & & \\
-1 & \ldots & -1 & -1 & -1 & \ldots & -1 & \ldots & & -1 \\
 & & & 1 & 0 & & 0 & & & \\
 & & & & \ddots & \ddots & \vdots & & & \\
 & & & & & 1 & 0 & & & \\
 & & & & & & 0 & 1 & & \\
 & & & & & & \vdots & & \ddots & \\
 & & & & & & 0 & & & 1
\end{array}\right).$$

The undisplayed entries in this matrix are zero. The l-th row includes only entries of -1. The main diagonal is 1, except for entry l where it is -1 and the entries $l+1$ to $k-1$ which are 0. Finally, all entries to the left of the main diagonal zeros are 1. An example of such a matrix for $D = 7$, $k = 5$ and $l = 2$ is

$$\mathbf{B}_{5,2} = \begin{pmatrix}
1 & 0 & 0 & 0 & 0 & 0 \\
-1 & -1 & -1 & -1 & -1 & -1 \\
0 & 1 & 0 & 0 & 0 & 0 \\
0 & 0 & 1 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 1 & 0 \\
0 & 0 & 0 & 0 & 0 & 1
\end{pmatrix}.$$

The matrix $\mathbf{B}_{kl}$ is evidently nonsingular, so its inverse $\mathbf{B}_{kl}^{-1}$ and the inverse of the transposed matrix $(\mathbf{B}'_{kl})^{-1}$ exist. Thus, for T and C affine equivariant, we have

$$T(\mathbf{x}_1^{(l)}, \ldots, \mathbf{x}_n^{(l)}) = T(\mathbf{B}'_{kl}\mathbf{x}_1^{(k)}, \ldots, \mathbf{B}'_{kl}\mathbf{x}_n^{(k)}) = \mathbf{B}'_{kl}T(\mathbf{x}_1^{(k)}, \ldots, \mathbf{x}_n^{(k)}),$$

$$C(\mathbf{x}_1^{(l)},\dots,\mathbf{x}_n^{(l)}) = C(\mathbf{B}_{kl}'\mathbf{x}_1^{(k)},\dots,\mathbf{B}_{kl}'\mathbf{x}_n^{(k)}) = \mathbf{B}_{kl}'C(\mathbf{x}_1^{(k)},\dots,\mathbf{x}_n^{(k)})\mathbf{B}_{kl},$$

and consequently $\mathrm{MD}^2(\mathbf{x}_i^{(l)}) =$

$$\begin{aligned}
&= [\mathbf{x}_i^{(l)} - T(\mathbf{x}_1^{(l)},\dots,\mathbf{x}_n^{(l)})]'[C(\mathbf{x}_1^{(l)},\dots,\mathbf{x}_n^{(l)})]^{-1}[\mathbf{x}_i^{(l)} - T(\mathbf{x}_1^{(l)},\dots,\mathbf{x}_n^{(l)})] = \\
&= [\mathbf{B}_{kl}'\mathbf{x}_i^{(k)} - \mathbf{B}_{kl}'T(\mathbf{x}_1^{(k)},\dots,\mathbf{x}_n^{(k)})]'[\mathbf{B}_{kl}'C(\mathbf{x}_1^{(k)},\dots,\mathbf{x}_n^{(k)})\mathbf{B}_{kl}]^{-1} \\
&\quad [\mathbf{B}_{kl}'\mathbf{x}_i^{(k)} - \mathbf{B}_{kl}'T(\mathbf{x}_1^{(k)},\dots,\mathbf{x}_n^{(k)})] = \\
&= [\mathbf{x}_i^{(k)} - T(\mathbf{x}_1^{(k)},\dots,\mathbf{x}_n^{(k)})]'\mathbf{B}_{kl}\mathbf{B}_{kl}^{-1}[C(\mathbf{x}_1^{(k)},\dots,\mathbf{x}_n^{(k)})]^{-1}(\mathbf{B}_{kl}')^{-1} \\
&\quad \mathbf{B}_{kl}'[\mathbf{x}_i^{(k)} - T(\mathbf{x}_1^{(k)},\dots,\mathbf{x}_n^{(k)})] = \mathrm{MD}^2(\mathbf{x}_i^{(k)}). \qquad \square
\end{aligned}$$

Proof of Theorem 2 Let the composition $\mathbf{x} = (x_1,\dots,x_D)' \in \mathcal{S}^D$, i.e. $\sum_{i=1}^{D} x_i = 1$, $x_i > 0$, be given. First we provide a matrix transformation between alr and clr transformations of $\mathbf{x}$. Without loss of generality, the last variable D is used for the alr transformation. Using an alternative representation of (1),

$$\mathbf{x}^{(D)} = (\log x_1 - \log x_D,\dots,\log x_{D-1} - \log x_D)',$$

and another presentation of (3),

$$\mathbf{y} = (y_1,\dots,y_D)',\ y_i = \frac{D-1}{D}\log x_i - \frac{1}{D}\sum_{j=1, j\neq i}^{D} \log x_j,\ i = 1,\dots,D,$$

it is easy to show that $\mathbf{x}^{(D)} = \mathbf{F}\mathbf{y}$ and $\mathbf{y} = \mathbf{F}^*\mathbf{x}^{(D)}$, where

$$\mathbf{F}_{D-1,D} = \begin{pmatrix} 1 & & -1 \\ & \ddots & \vdots \\ & 1 & -1 \end{pmatrix} \quad \text{and} \quad \mathbf{F}^*_{D,D-1} = \begin{pmatrix} \frac{D-1}{D} & -\frac{1}{D} & \cdots & -\frac{1}{D} \\ -\frac{1}{D} & \frac{D-1}{D} & \ddots & \vdots \\ \vdots & \ddots & \ddots & -\frac{1}{D} \\ \vdots & & \ddots & \frac{D-1}{D} \\ -\frac{1}{D} & \cdots & \cdots & -\frac{1}{D} \end{pmatrix}$$

(see also Aitchison, 1986, Section 5.1). Moreover, $\mathbf{F}\mathbf{F}^* = \mathbf{I}_{D-1}$ (identity matrix of order $D-1$), $\mathbf{F}^*\mathbf{F}$ is symmetric, $\mathbf{F}\mathbf{F}^*\mathbf{F} = \mathbf{F}$, and $\mathbf{F}^*\mathbf{F}\mathbf{F}^* = \mathbf{F}^*$. Thus, $\mathbf{F}^*$ fulfills all properties of the Moore–Penrose inverse matrix $\mathbf{F}^+$ of $\mathbf{F}$, i.e.

$$\mathbf{F}\mathbf{F}^+\mathbf{F} = \mathbf{F},\ \mathbf{F}^+\mathbf{F}\mathbf{F}^+ = \mathbf{F}^+,\ (\mathbf{F}\mathbf{F}^+)' = \mathbf{F}\mathbf{F}^+,\ (\mathbf{F}^+\mathbf{F})' = \mathbf{F}^+\mathbf{F},$$

and in our case additionally $\mathbf{F}\mathbf{F}^+ = \mathbf{I}$. Analogous conclusions can be obtained also for other choices of the ratioing variable for the alr transformation, but the structures of the matrices are different.

Let us consider now alr and clr transformed data matrices $\mathbf{X}_{n,D-1}^{(D)}$ and $\mathbf{Y}_{n,D}$ with rows $\mathbf{x}_i^{(D)}$ and $\mathbf{y}_i$, for $i = 1, \ldots, n$, respectively. We use the notations $\bar{\mathbf{x}}^{(D)}$ and $\bar{\mathbf{y}}$ for the corresponding arithmetic mean vectors, and $\mathbf{S}_{\mathbf{x}^{(D)}}$ and $\mathbf{S}_{\mathbf{y}}$ for the sample covariance matrices. For the latter we find the relation

$$\begin{aligned}\mathbf{S}_{\mathbf{y}} &= \frac{1}{n}\sum_{i=1}^{n}(\mathbf{y}_i-\bar{\mathbf{y}})(\mathbf{y}_i-\bar{\mathbf{y}})' = \frac{1}{n}\sum_{i=1}^{n}(\mathbf{F}^+\mathbf{x}_i^{(D)}-\mathbf{F}^+\bar{\mathbf{x}}^{(D)})(\mathbf{F}^+\mathbf{x}_i^{(D)}-\mathbf{F}^+\bar{\mathbf{x}}^{(D)})' \\ &= \mathbf{F}^+\frac{1}{n}\sum_{i=1}^{n}(\mathbf{x}_i^{(D)}-\bar{\mathbf{x}}^{(D)})(\mathbf{x}_i^{(D)}-\bar{\mathbf{x}}^{(D)})'(\mathbf{F}^+)' = \mathbf{F}^+\mathbf{S}_{\mathbf{x}^{(D)}}(\mathbf{F}^+)'.\end{aligned}$$

Furthermore, $\mathrm{MD}^2(\mathbf{x}_i^{(D)}) =$

$$\begin{aligned}&= (\mathbf{x}_i^{(D)}-\bar{\mathbf{x}}^{(D)})'\mathbf{S}_{\mathbf{x}^{(D)}}^{-1}(\mathbf{x}_i^{(D)}-\bar{\mathbf{x}}^{(D)}) = (\mathbf{F}\mathbf{y}_i-\mathbf{F}\bar{\mathbf{y}})'\mathbf{S}_{\mathbf{x}^{(D)}}^{-1}(\mathbf{F}\mathbf{y}_i-\mathbf{F}\bar{\mathbf{y}}) \\ &= (\mathbf{y}_i-\bar{\mathbf{y}})'\mathbf{F}'\mathbf{S}_{\mathbf{x}^{(D)}}^{-1}\mathbf{F}(\mathbf{y}_i-\bar{\mathbf{y}}), \quad i=1,\ldots,n.\end{aligned}$$

We denote $\mathbf{S}_{\mathbf{y}}^* = \mathbf{F}'\mathbf{S}_{\mathbf{x}^{(D)}}^{-1}\mathbf{F}$. Then, using the above mentioned properties of the Moore–Penrose inverse, property $\mathbf{F}\mathbf{F}^+ = \mathbf{I}$, and basic matrix algebra, we can compute

$$\begin{aligned}\mathbf{S}_{\mathbf{y}}\mathbf{S}_{\mathbf{y}}^*\mathbf{S}_{\mathbf{y}} &= \mathbf{F}^+\mathbf{S}_{\mathbf{x}^{(D)}}(\mathbf{F}^+)'\mathbf{F}'\mathbf{S}_{\mathbf{x}^{(D)}}^{-1}\mathbf{F}\mathbf{F}^+\mathbf{S}_{\mathbf{x}^{(D)}}(\mathbf{F}^+)' = \mathbf{F}^+\mathbf{S}_{\mathbf{x}^{(D)}}(\mathbf{F}^+)' = \mathbf{S}_{\mathbf{y}}, \\ \mathbf{S}_{\mathbf{y}}^*\mathbf{S}_{\mathbf{y}}\mathbf{S}_{\mathbf{y}}^* &= \mathbf{F}'\mathbf{S}_{\mathbf{x}^{(D)}}^{-1}\mathbf{F}\mathbf{F}^+\mathbf{S}_{\mathbf{x}^{(D)}}(\mathbf{F}^+)'\mathbf{F}'\mathbf{S}_{\mathbf{x}^{(D)}}^{-1}\mathbf{F} = \mathbf{F}'\mathbf{S}_{\mathbf{x}^{(D)}}^{-1}\mathbf{F} = \mathbf{S}_{\mathbf{y}}^*, \\ (\mathbf{S}_{\mathbf{y}}\mathbf{S}_{\mathbf{y}}^*)' &= [\mathbf{F}^+\mathbf{S}_{\mathbf{x}^{(D)}}(\mathbf{F}^+)'\mathbf{F}'\mathbf{S}_{\mathbf{x}^{(D)}}^{-1}\mathbf{F}]' = (\mathbf{F}^+\mathbf{F})' = \mathbf{F}^+\mathbf{F} = \\ &= \mathbf{F}^+\mathbf{S}_{\mathbf{x}^{(D)}}(\mathbf{F}^+)'\mathbf{F}'\mathbf{S}_{\mathbf{x}^{(D)}}^{-1}\mathbf{F} = \mathbf{S}_{\mathbf{y}}\mathbf{S}_{\mathbf{y}}^*, \\ (\mathbf{S}_{\mathbf{y}}^*\mathbf{S}_{\mathbf{y}})' &= [\mathbf{F}'\mathbf{S}_{\mathbf{x}^{(D)}}^{-1}\mathbf{F}\mathbf{F}^+\mathbf{S}_{\mathbf{x}^{(D)}}(\mathbf{F}^+)']' = [(\mathbf{F}^+\mathbf{F})']' = (\mathbf{F}^+\mathbf{F})' = \\ &= \mathbf{F}'\mathbf{S}_{\mathbf{x}^{(D)}}^{-1}\mathbf{F}\mathbf{F}^+\mathbf{S}_{\mathbf{x}^{(D)}}(\mathbf{F}^+)' = \mathbf{S}_{\mathbf{y}}^*\mathbf{S}_{\mathbf{y}}.\end{aligned}$$

This shows that $\mathbf{S}_{\mathbf{y}}^* = \mathbf{S}_{\mathbf{y}}^+$ is the Moore–Penrose inverse of $\mathbf{S}_{\mathbf{y}}$, and consequently

$$\mathrm{MD}^2(\mathbf{x}_i^{(D)}) = (\mathbf{y}_i-\bar{\mathbf{y}})'\mathbf{F}'\mathbf{S}_{\mathbf{x}^{(D)}}^{-1}\mathbf{F}(\mathbf{y}_i-\bar{\mathbf{y}}) = (\mathbf{y}_i-\bar{\mathbf{y}})'\mathbf{S}_{\mathbf{y}}^+(\mathbf{y}_i-\bar{\mathbf{y}}) = \mathrm{MD}^2(\mathbf{y}_i)$$

for $i = 1, \ldots, n$. Here we have directly used the Moore-Penrose inverse matrix $\mathbf{S}_{\mathbf{y}}^+$ in the expression of $\mathrm{MD}^2(\mathbf{y}_i)$ since in most statistical software packages it is directly computable. Another equivalent possibility to prove above mentioned property is presented in Aitchison (1986, Property 5.6). Using Theorem 1 and the notation of (10), we obtain

$$\mathrm{MD}^2(\mathbf{x}_i^{(l)}) = \mathrm{MD}^2(\mathbf{x}_i^{(D)}) = \mathrm{MD}^2(\mathbf{y}_i) \quad \text{for} \quad l = 1, \ldots, D-1,$$

which completes the proof. □

Proof of Theorem 3 Let $\mathbf{x}^{(D)}$, $\mathbf{y}$, and $\mathbf{z}$ be alr (last variable is chosen as ratio variable), clr and ilr transformations, respectively, for composition $\mathbf{x} \in \mathcal{S}^D$, see (1), (3), and (5). Then, from the proof of Theorem 2, $\mathbf{x}^{(D)} = \mathbf{F}\mathbf{y}$ and $\mathbf{y} = \mathbf{F}^{+}\mathbf{x}^{(D)}$. The relations $\mathbf{y} = \mathbf{V}\mathbf{z}$, $\mathbf{V}'\mathbf{V} = \mathbf{I}_{D-1}$ for a $D \times (D-1)$ matrix $\mathbf{V}$ with orthogonal basis vectors in its columns, follow immediately from the properties of isometric logratio transformation. Consequently,

$$\mathbf{x}^{(D)} = \mathbf{F}\mathbf{V}\mathbf{z} \quad \text{and} \quad \mathbf{z} = \mathbf{V}'\mathbf{F}^{+}\mathbf{x}^{(D)}$$

are relations between alr and ilr transformations, with $(D-1) \times (D-1)$ matrices $\mathbf{F}\mathbf{V}$ and $\mathbf{V}'\mathbf{F}^{+}$. The second relation was derived from $\mathbf{y} = \mathbf{F}^{+}\mathbf{x}^{(D)}$, multiplied with $\mathbf{V}'$ from the left and using the above described properties. By substitution into the first relation we obtain

$$\mathbf{x}^{(D)} = \mathbf{F}\mathbf{V}\mathbf{V}'\mathbf{F}^{+}\mathbf{x}^{(D)},$$

and comparing both sides it immediately follows that $\mathbf{F}\mathbf{V}\mathbf{V}'\mathbf{F}^{+} = \mathbf{I}$. Thus, $\mathbf{V}'\mathbf{F}^{+}$ is the inverse matrix of the nonsingular matrix $\mathbf{F}\mathbf{V}$ (Harville, 1997, p. 80, Lemma 8.3.1). Using (8) and Theorem 1 results in

$$\mathrm{MD}^2(\mathbf{z}) = \mathrm{MD}^2(\mathbf{V}'\mathbf{F}^{+}\mathbf{x}^{(D)}) = \mathrm{MD}^2(\mathbf{x}^{(j)}) \quad \text{for} \quad j = 1, \ldots, D.$$

□

Figure Captions

Figure 1 ilr transformed Aphyric Skye Lavas data with classical (upper left) and robust (upper right) MDs and their transformation into the ternary diagram (classical: lower left; robust: lower right).

Figure 2 alr transformed Arctic Lake Sediment data with classical (upper left) and robust (upper right) MDs and their transformation into the ternary diagram (classical: lower left; robust: lower right).

Figure 3 Comparison of classical and robust Mahalanobis distances of the ilr transformed Kola data (left) and presentation of the regular observations and identified outliers in the map (right).

Figure 4 Comparison of robust Mahalanobis distances with original and ilr transformed Kola data (left) and presentation of the regular observations and identified outliers in the map (right).

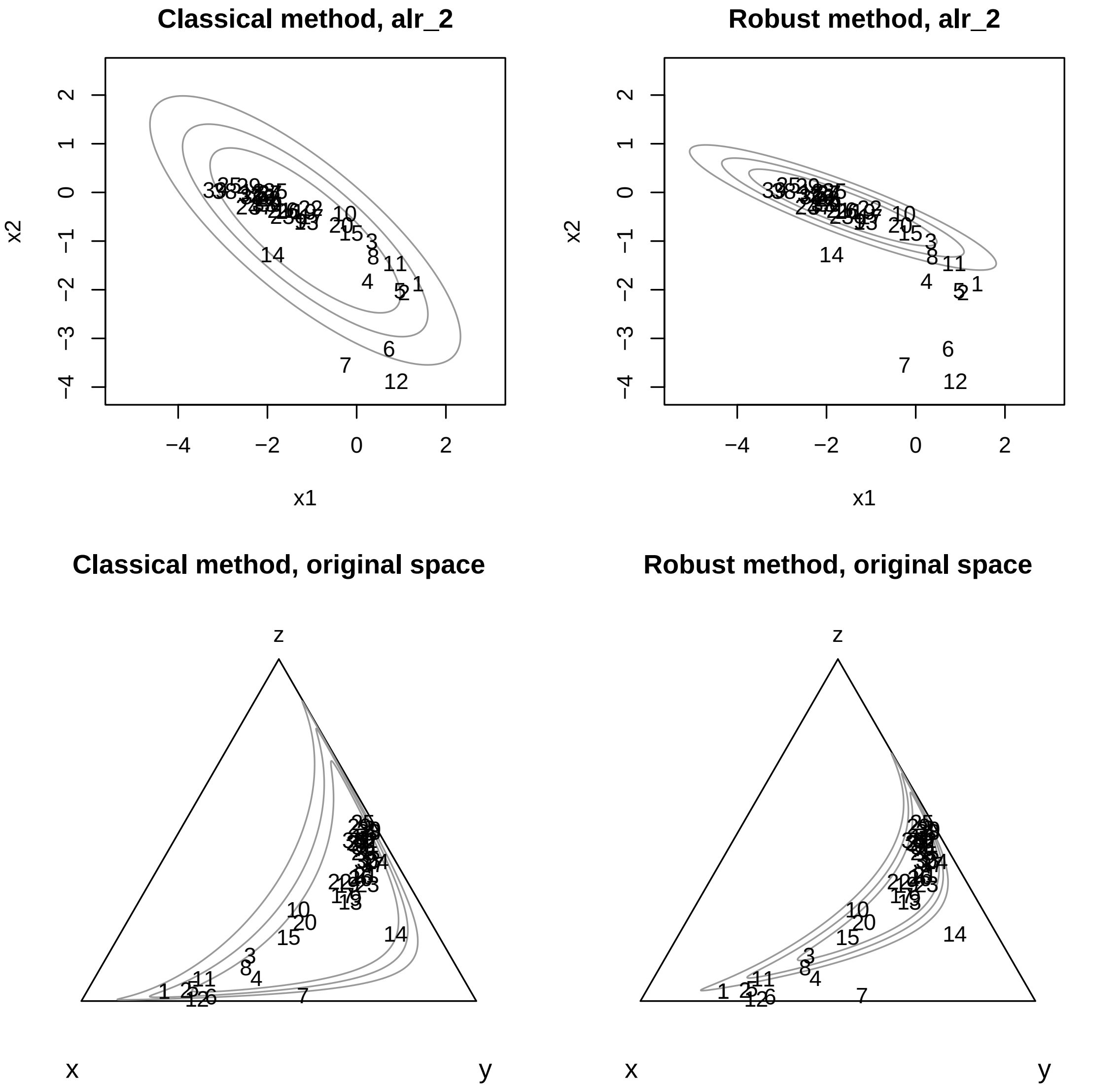
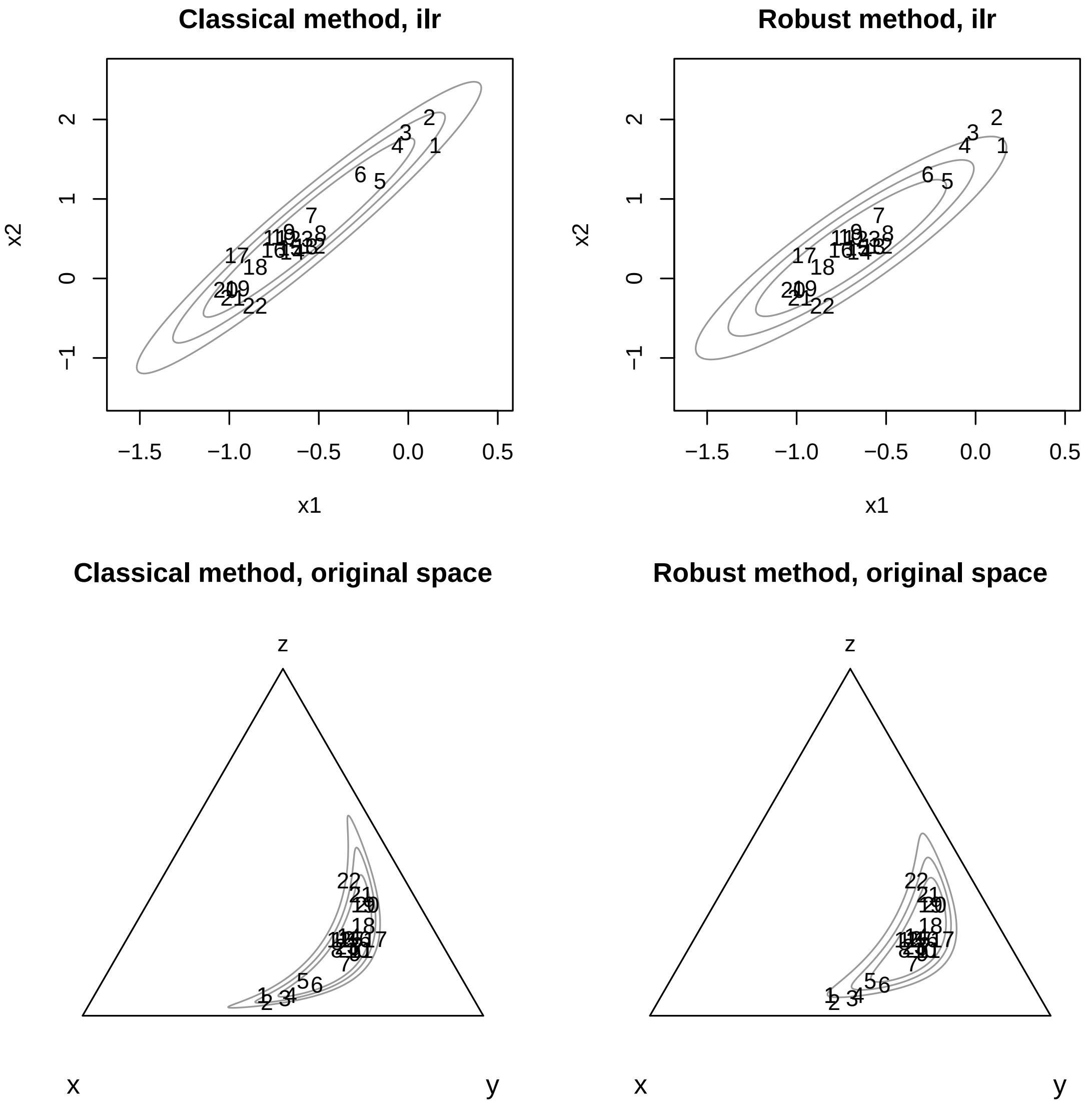


Figure 1: alr transformed Arctic Lake Sediment data with classical (upper left) and robust (upper right) MDs and their transformation into the ternary diagram (classical: lower left; robust: lower right).


Classical method, ilr

Robust method, ilr

Classical method, original space

Robust method, original space

Figure 2: ilr transformed Aphyric Skye Lavas data with classical (upper left) and robust (upper right) MDs and their transformation into the ternary diagram (classical: lower left; robust: lower right).

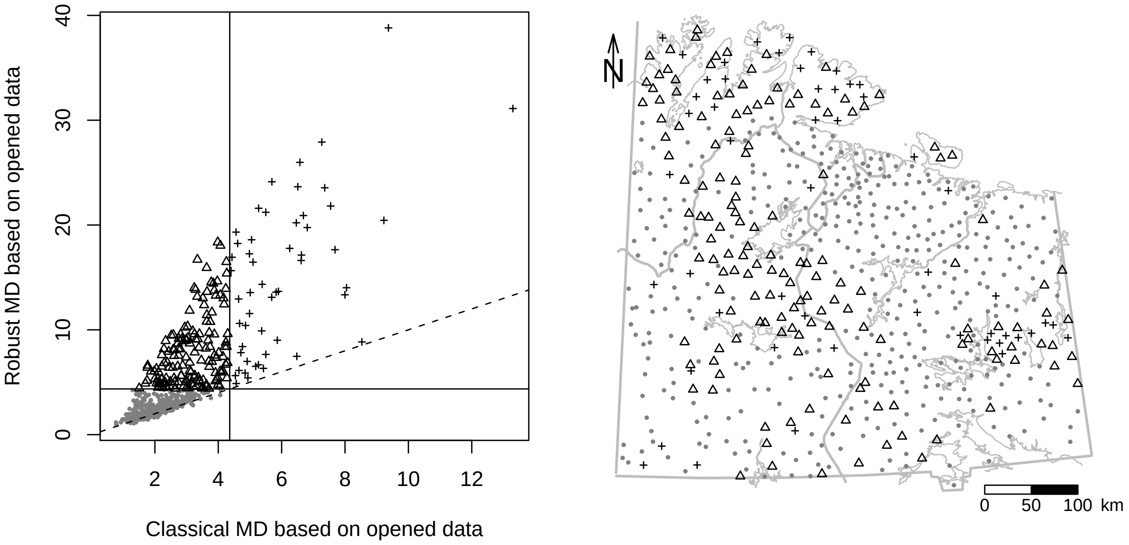


Figure 3: Comparison of classical and robust Mahalanobis distances of the ilr transformed Kola data (left) and presentation of the regular observations and identified outliers in the map (right).

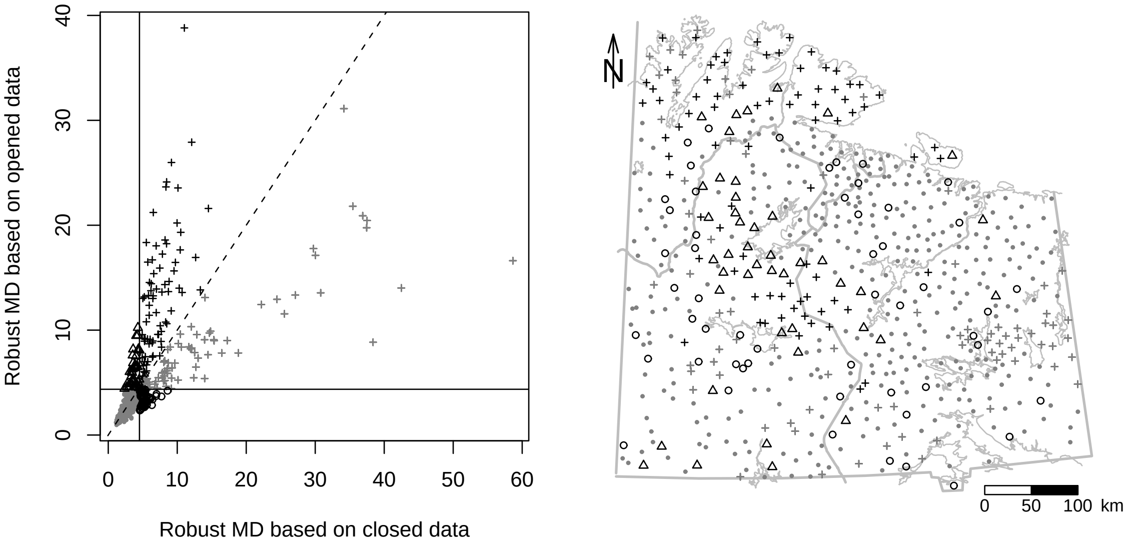


Figure 4: Comparison of robust Mahalanobis distances with original and ilr transformed Kola data (left) and presentation of the regular observations and identified outliers in the map (right).