

Covariance Structure of Compositional Tables

Kamila Fačevicová
Palacký University in Olomouc

Karel Hron
Palacký University in Olomouc

Abstract

Recent experience with interpretation of orthonormal coordinates in compositional data shows clearly a necessity of their better understanding in terms of logratios that form the primary source of information within the logratio methodology. This is even more crucial in the special case of compositional tables, where both balances and coordinates with odds ratio interpretation are involved. The aim of the paper is to provide a decomposition of covariance structure of orthonormal coordinates in compositional tables in terms of logratio variances, which could serve this purpose. For their better interpretability, the formulas are also accompanied with appropriate comments and graphical illustrations, and implications for the prominent case of 2×2 compositional tables are discussed.

Keywords: compositional tables, covariance structure, orthonormal coordinates.

1. Introduction

Although the logratio methodology seems to be nowadays a well-established approach to statistical analysis of compositional data, i.e. multivariate observations carrying relative information (Aitchison 1986; Pawlowsky-Glahn and Buccianti 2011), it is also suitable for more complex data structures, where not the absolute values but rather ratios are of primary interest. One of them are compositional tables (Egozcue, Díaz-Barrero, and Pawlowsky-Glahn 2008; Egozcue, Pawlowsky-Glahn, Templ, and Hron 2014; Fačevicová, Hron, Todorov, Guo, and Templ 2014a; Fačevicová, Hron, Todorov, and Templ 2014b), a continuous counterpart to well-known contingency tables (Agregsti 2002). Besides the difference in the nature of data, (cells of the contingency table are discrete counts, while parts of the compositional tables are continuous values), the main difference between is that, on the one hand, a contingency table collects results from n independent observations, while, on the other hand, a compositional table itself represents one observation. The analysis of relationships between row and column factors is thus based on sample of n compositional tables. Furthermore, by applying the Aitchison geometry (Egozcue and Pawlowsky-Glahn 2006), and following the principles of compositional data analysis, it is possible to decompose the original table into its independent and interactive parts (the latter capturing relations between both factors), while assuming geometric marginals instead of the standard arithmetic ones. Moreover, in Fačevicová *et al.* (2014b) orthonormal coordinates were assigned to both independence and interaction tables that enable to perform statistical analysis using standard methods and focus only on coordi-

nates of the interaction table. The problem of analysis of relationship between factors from a sample of tables, which would need to be handled using three-dimensional contingency tables or log-linear models in the standard case, thus can directly transferred to standard statistical treatment (like hypotheses testing) in coordinates; for example, independence of factors corresponds to zero coordinates of the interaction table.

While coordinates of the independence table can be interpreted through balances (Egozcue and Pawlowsky-Glahn 2005), coordinate representation of the interaction one needs to be formulated in sense of odds ratios (Fačevicová *et al.* 2014b). Although motivation for the latter coordinates is quite intuitive as odds ratios become popular to represent also contingency tables (Agresti 2002), interpretation of the coordinate system as a whole may seem to be too complex for practical purposes. On the way to enhance the interpretability, one possibility is to analyze covariance structure of coordinates (Fišerová and Hron 2011) to see which ratios contribute (in positive or negative sense) to values of individual variances and covariances. Nevertheless, a specific structure of compositional tables and their respective coordinates requires a deeper insight as balances form just one part of the coordinate system.

The aim of the presented paper is to analyze covariance structure of orthonormal coordinates for compositional tables in terms of elements of the variation matrix (Aitchison 1986), i.e., as linear combinations of variances of single logratios, which seems to be a necessary step in further development of any reasonable coordinate representation of compositional tables. The paper is organized as follows. In the next section, basics of compositional tables and their decomposition into independent and interactive parts are recalled. Section 3 is devoted to the covariance structure of coordinates itself, where the corresponding formulas (that might seem to be rather complex) are always illustrated with a graphical scheme to allow their better understanding. In Section 4 some implications for the special case of 2×2 compositional tables are briefly mentioned. Section 5 presents an illustrative example and Section 6 concludes.

2. Orthonormal coordinates of $I \times J$ compositional tables

A $I \times J$ compositional table \mathbf{x} is a special case of compositional data that are arranged into a form of table to identify relation between two (row and column) factors. They are formed by parts $x_{ij} > 0$ for $i = 1, 2, \dots, I$ and $j = 1, 2, \dots, J$ which carry only relative information. Consequently, their sum κ is arbitrary (like $\kappa = 1$ for the case of proportions), reached formally using the closure operation

$$\mathcal{C}(\mathbf{x}) = \left(\frac{\kappa x_{ij}}{\sum_{k,l=1}^{I,J} x_{kl}} \right)_{i,j=1}^{I,J}.$$

The sample space of representations of $I \times J$ compositional tables is the simplex, $(IJ - 1)$ -dimensional subset of \mathbf{R}^{IJ} defined as

$$\mathcal{S}^{IJ} = \left\{ \mathbf{x} = (x_{ij})_{i,j=1}^{I,J} \mid x_{ij} > 0, i = 1, 2, \dots, I, j = 1, 2, \dots, J; \sum_{i,j=1}^{I,J} x_{ij} = \kappa \right\}.$$

To follow specific features of compositional tables (as a special case of compositional data), the Aitchison geometry on the simplex is defined, see Egozcue and Pawlowsky-Glahn (2006) for details. This geometry has the same algebraic-geometrical structure as the standard Euclidean geometry in real space and is represented by operations of perturbation, power transformation, and the Aitchison inner product. According to Egozcue *et al.* (2008) these operations are defined for compositional tables \mathbf{x} and \mathbf{y} and $\alpha \in \mathbf{R}$ as

$$\mathbf{x} \oplus \mathbf{y} = \mathcal{C} (x_{ij} y_{ij})_{i,j=1}^{I,J}, \quad \alpha \odot \mathbf{x} = \mathcal{C} (x_{ij}^\alpha)_{i,j=1}^{I,J}$$

and

$$\langle \mathbf{x}, \mathbf{y} \rangle_A = \frac{1}{2IJ} \sum_{i,j} \sum_{k,l} \ln \frac{x_{ij}}{x_{kl}} \ln \frac{y_{ij}}{y_{kl}}.$$

Compositional table $\mathbf{n} = \mathcal{C}(x_{ij} = 1)_{i,j=1}^{I,J}$ stands for the neutral element in the $(IJ - 1)$ -dimensional vector space $(\mathcal{S}^{IJ}, \oplus, \odot)$.

Each compositional table \mathbf{x} can be expressed as

$$\mathbf{x} = \langle \mathbf{x}, \mathbf{e}_1 \rangle_A \odot \mathbf{e}_1 \oplus \dots \oplus \langle \mathbf{x}, \mathbf{e}_{IJ-1} \rangle_A \odot \mathbf{e}_{IJ-1},$$

where $(\mathbf{e}_1, \dots, \mathbf{e}_{IJ-1})$ form an orthonormal basis in $(IJ-1)$ -dimensional simplex (with respect to the Aitchison geometry), resulting in $(IJ - 1)$ -dimensional real vector of orthonormal coordinates

$$\mathbf{z} = h(\mathbf{x}) = (z_1, \dots, z_{IJ-1}) = (\langle \mathbf{x}, \mathbf{e}_1 \rangle_A, \dots, \langle \mathbf{x}, \mathbf{e}_{IJ-1} \rangle_A).$$

Consequently, the following relations between the Aitchison and the Euclidean geometries can be derived,

$$h(\alpha \odot \mathbf{x} \oplus \beta \odot \mathbf{y}) = \alpha \cdot h(\mathbf{x}) + \beta \cdot h(\mathbf{y}), \quad \langle \mathbf{x}, \mathbf{y} \rangle_A = \langle h(\mathbf{x}), h(\mathbf{y}) \rangle_E,$$

i.e. h is an isometric mapping from \mathcal{S}^{IJ} to \mathbf{R}^{IJ-1} (we refer also to isometric logratio (ilr) transformation (Egozcue, Pawlowsky-Glahn, Mateu-Figueras, and Barceló-Vidal 2003)).

Within the framework of the Aitchison geometry it is possible to decompose the original compositional table into its independent and interactive parts, $\mathbf{x} = \mathbf{x}_{ind} \oplus \mathbf{x}_{int}$, see Egozcue *et al.* (2008) for details. The independent part (independence table) is compositional table with elements

$$\mathbf{x}_{ind} = \left(x_{ij}^{ind} = \left(\prod_{k=1}^I \prod_{l=1}^J x_{kl} x_{il} \right)^{\frac{1}{IJ}} \right)_{i,j=1}^{I,J} \quad (1)$$

and the interactive part (interaction table)

$$\mathbf{x}_{int} = \left(x_{ij}^{int} = \left(\prod_{k=1}^I \prod_{l=1}^J \frac{x_{ij}}{x_{kl} x_{il}} \right)^{\frac{1}{IJ}} \right)_{i,j=1}^{I,J}. \quad (2)$$

According to Fačevicová *et al.* (2014b), the independence and the interaction tables can be expressed in $I + J - 2$ and $(I - 1)(J - 1)$ nonzero orthonormal coordinates, respectively, the remaining coordinates (up to the total number of $IJ - 1$ variables) being zero. The coordinates of the independence table can be expressed as balances

$$z_i^r = \sqrt{\frac{(I-i)J}{I-i+1}} \ln \frac{(x_{i1} \dots x_{iJ})^{1/J}}{(x_{i+1,1} \dots x_{IJ})^{1/(IJ-iJ)}}, \quad i = 1, \dots, I-1, \quad (3)$$

and

$$z_j^c = \sqrt{\frac{I(J-j)}{J-j+1}} \ln \frac{(x_{1j} \dots x_{Ij})^{1/I}}{(x_{1,j+1} \dots x_{IJ})^{1/(IJ-Ij)}}, \quad j = 1, \dots, J-1, \quad (4)$$

representing the row and column information (logratios), respectively, conveyed by the independence table; coordinates of the interaction table can be chosen as

$$z_{rs}^{int} = \frac{1}{\sqrt{r \cdot s \cdot (r-1) \cdot (s-1)}} \ln \prod_{i=1}^{r-1} \prod_{j=1}^{s-1} \frac{x_{ij} x_{rs}}{x_{is} x_{rj}}, \quad (5)$$

with an odds ratio structure. These two sets of coordinates together form an orthonormal coordinate representation of the original compositional table \mathbf{x} . Covariance structure in terms of elements of the variation matrix (Aitchison 1986), especially for coordinates of the interaction table, will be studied in detail in the next section.

3. Covariance structure of coordinates of the compositional table

In the following, covariance structure of the above mentioned coordinate representation will be expressed as linear combinations of variances of logratios. At first, covariance structure of the interaction table is introduced, followed by the independence table structure, and finally also mutual relations between both tables (expressed through the corresponding covariances) are analyzed.

Variances of logratios form the elemental information on variability in compositional tables and are summarized in $IJ \times IJ$ variation matrix

$$\mathbf{T} = \begin{pmatrix} \text{var} \left(\ln \frac{x_{11}}{x_{11}} \right) & \text{var} \left(\ln \frac{x_{11}}{x_{12}} \right) & \cdots & \text{var} \left(\ln \frac{x_{11}}{x_{IJ}} \right) \\ \text{var} \left(\ln \frac{x_{12}}{x_{11}} \right) & \text{var} \left(\ln \frac{x_{12}}{x_{12}} \right) & \cdots & \text{var} \left(\ln \frac{x_{12}}{x_{IJ}} \right) \\ \vdots & \vdots & \ddots & \vdots \\ \text{var} \left(\ln \frac{x_{IJ}}{x_{11}} \right) & \text{var} \left(\ln \frac{x_{IJ}}{x_{12}} \right) & \cdots & \text{var} \left(\ln \frac{x_{IJ}}{x_{IJ}} \right) \end{pmatrix}. \quad (6)$$

As it is usual within the logratio methodology, all coordinates are logcontrasts, i.e. they can be expressed in form

$$z = \sum_{i=1}^I \sum_{j=1}^J a_{ij} \ln x_{ij} = \mathbf{a}' \ln \mathbf{x}, \quad \text{where} \quad \sum_{i=1}^I \sum_{j=1}^J a_{ij} = 0.$$

Also the covariance structure can be derived accordingly (Aitchison 1986).

Proposition 3.1 *Variances and covariances for logcontrasts $\mathbf{a}' \ln \mathbf{x}$ and $\mathbf{b}' \ln \mathbf{x}$ of a IJ -part compositional table \mathbf{x} are*

$$\text{var}(\mathbf{a}' \ln \mathbf{x}) = -\frac{1}{2} \mathbf{a}' \mathbf{T} \mathbf{a}, \quad (7)$$

$$\text{cov}(\mathbf{a}' \ln \mathbf{x}, \mathbf{b}' \ln \mathbf{x}) = -\frac{1}{2} \mathbf{a}' \mathbf{T} \mathbf{b}. \quad (8)$$

Since the possible logcontrast representation of coordinates (2), (3) and (4), Equations (7) and (8) are crucial to derive of their covariance structure. As the interaction table is usually of main interest for the analysis, we start with variances of its respective coordinates.

Theorem 3.2 *Consider an arbitrary coordinate z_{rs} , for $r = 2, \dots, I$ and $s = 2, \dots, J$ of the interaction table \mathbf{x}_{int} from (5). Its variance is formed by three parts,*

$$\text{var}(z_{rs}) = A_1 - B_1 - C_1. \quad (9)$$

The first part, increasing the variance, is

$$\begin{aligned} A_1 &= \frac{1}{rs(s-1)} \sum_{i=1}^{r-1} \sum_{j,j'=1}^{s-1} \text{var} \left(\ln \frac{x_{ij}}{x_{rj'}} \right) + \frac{1}{rs(r-1)} \sum_{i,i'=1}^{r-1} \sum_{j=1}^{s-1} \text{var} \left(\ln \frac{x_{ij}}{x_{i's}} \right) + \\ &+ \frac{r-1}{rs} \sum_{j=1}^{s-1} \text{var} \left(\ln \frac{x_{rj}}{x_{rs}} \right) + \frac{s-1}{rs} \sum_{i=1}^{r-1} \text{var} \left(\ln \frac{x_{is}}{x_{rs}} \right). \end{aligned} \quad (10)$$

The variance of the coordinate is reduced by parts

$$B_1 = \frac{1}{2} \frac{1}{rs(r-1)(s-1)} \sum_{i,i'=1}^{r-1} \sum_{j,j'=1}^{s-1} \text{var} \left(\ln \frac{x_{ij}}{x_{i'j'}} \right) + \frac{1}{rs} \sum_{i=1}^{r-1} \sum_{j=1}^{s-1} \text{var} \left(\ln \frac{x_{ij}}{x_{rs}} \right) \quad (11)$$

and

$$\begin{aligned}
 C_1 = & \frac{1}{2} \frac{r-1}{rs(s-1)} \sum_{j,j'=1}^{s-1} \text{var} \left(\ln \frac{x_{rj}}{x_{rj'}} \right) + \frac{1}{2} \frac{s-1}{rs(r-1)} \sum_{i,i'=1}^{r-1} \text{var} \left(\ln \frac{x_{is}}{x_{i's}} \right) + \\
 & + \frac{1}{rs} \sum_{i=1}^{r-1} \sum_{j'=1}^{s-1} \text{var} \left(\ln \frac{x_{is}}{x_{rj'}} \right). \tag{12}
 \end{aligned}$$

Proof: When parts of the compositional table \mathbf{x} are rearranged in form of composition $\mathbf{x}_r = (x_{11}, x_{12}, \dots, x_{1J}, x_{21}, \dots, x_{IJ})$, coordinate z_{rs} of the interaction table can be expressed as $z_{rs} = \mathbf{a}' \ln \mathbf{x}_r$, where for elements of the coefficient vector $\mathbf{a} = (a_{11}, a_{12}, \dots, a_{1J}, a_{21}, \dots, a_{IJ})$ the following relations hold,

$$\begin{aligned}
 a_{ij} &= 1/\sqrt{rs(r-1)(s-1)} & \text{for } i = 1, \dots, r-1 & \quad j = 1, \dots, s-1 \\
 a_{ij} &= -(r-1)/\sqrt{rs(r-1)(s-1)} & \text{for } i = r & \quad j = 1, \dots, s-1 \\
 a_{ij} &= -(s-1)/\sqrt{rs(r-1)(s-1)} & \text{for } i = 1, \dots, r-1 & \quad j = s \\
 a_{ij} &= (r-1)(s-1)/\sqrt{rs(r-1)(s-1)} & \text{for } i = r & \quad j = s \\
 a_{ij} &= 0 & & \text{otherwise.}
 \end{aligned}$$

Equation (9) is then consequence of Proposition 3.1.

◇

From Theorem 3.2 it is clear that variance of the coordinate z_{rs} is formed by nine groups of logratio variances. Four of them increase the overall variability and the other five reduce it. The first four groups are represented by A_1 , which is formed by logratios of “inner” parts of the partial table or part x_{rs} , with its last row and column (i.e. r -th row and s -th column of the original table \mathbf{x}) except of the part x_{rs} itself:

- variances of logratios between an inner part of the partial table and a part from its last row (except of x_{rs}),
- variances of logratios between an inner part of the partial table and a part from its last column (except of x_{rs}),
- variances of logratios between a part from the last row (except of x_{rs}) and x_{rs} itself,
- variances of logratios between a part from the last column (except of x_{rs}) and x_{rs} itself.

The variance of z_{rs} is reduced by B_1 and C_1 , formed by variances of logratios corresponding to remaining possible relations between parts of the above defined groups (inner tables, last row/column without x_{rs} , part x_{rs} itself). Concretely, B_1 consists of

- variances of logratios between inner parts of the partial table,
- variances of logratios between an inner part and x_{rs} .

Similarly, C_1 is formed by

- variances of logratios between parts from the last row (except of x_{rs}),
- variances of logratios between parts from the last column (except of x_{rs}),
- variances of logratios between parts from the last row and the last column (except of x_{rs}).

The above relations can be expressed also graphically, as shown in Figure 1.

Covariances between coordinates of the interaction table are derived in the next theorem.

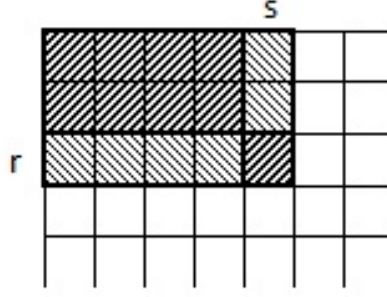


Figure 1: Variance of coordinate z_{rs} is increased by variances of logratios between a part from area highlighted by (/) and a part from the second area highlighted by (\) – A_1 . The variance of z_{rs} is reduced by variances of logratios between two parts from area (/) – B_1 or two parts from (\) – C_1 .

Theorem 3.3 Consider two coordinates of the interaction table $z_{r_1 s_1}, z_{r_2 s_2}$, for $r_1, r_2 = 2, \dots, I$ and $s_1, s_2 = 2, \dots, J$. Then for their covariance the following holds,

$$\text{cov}(z_{r_1 s_1}, z_{r_2 s_2}) = K(A_2 + B_2 - C_2 - D_2), \quad (13)$$

where

$$\begin{aligned} A_2 = & (s_2 - 1) \sum_{i_1=1}^{r_1-1} \sum_{i_2=1}^{r_2-1} \sum_{j_1=1}^{s_1-1} \text{var} \left(\ln \frac{x_{i_1 j_1}}{x_{i_2 s_2}} \right) + (r_2 - 1) \sum_{i_1=1}^{r_1-1} \sum_{j_1=1}^{s_1-1} \sum_{j_2=1}^{s_2-1} \text{var} \left(\ln \frac{x_{i_1 j_1}}{x_{r_2 j_2}} \right) + \\ & + (r_1 - 1)(s_1 - 1)(s_2 - 1) \sum_{i_2=1}^{r_2-1} \text{var} \left(\ln \frac{x_{r_1 s_1}}{x_{i_2 s_2}} \right) + (r_1 - 1)(r_2 - 1)(s_1 - 1) \sum_{j_2=1}^{s_2-1} \text{var} \left(\ln \frac{x_{r_1 s_1}}{x_{r_2 j_2}} \right), \end{aligned} \quad (14)$$

$$\begin{aligned} B_2 = & (s_1 - 1) \sum_{i_1=1}^{r_1-1} \sum_{i_2=1}^{r_2-1} \sum_{j_2=1}^{s_2-1} \text{var} \left(\ln \frac{x_{i_1 s_1}}{x_{i_2 j_2}} \right) + (s_1 - 1)(s_2 - 1)(r_2 - 1) \sum_{i_1=1}^{r_1-1} \text{var} \left(\ln \frac{x_{i_1 s_1}}{x_{r_2 s_2}} \right) + \\ & + (r_1 - 1) \sum_{j_1=1}^{s_1-1} \sum_{i_2=1}^{r_2-1} \sum_{j_2=1}^{s_2-1} \text{var} \left(\ln \frac{x_{r_1 j_1}}{x_{i_2 j_2}} \right) + (r_1 - 1)(r_2 - 1)(s_2 - 1) \sum_{j_1=1}^{s_1-1} \text{var} \left(\ln \frac{x_{r_1 j_1}}{x_{r_2 s_2}} \right), \end{aligned} \quad (15)$$

$$\begin{aligned} C_2 = & \sum_{i_1=1}^{r_1-1} \sum_{i_2=1}^{r_2-1} \sum_{j_1=1}^{s_1-1} \sum_{j_2=1}^{s_2-1} \text{var} \left(\ln \frac{x_{i_1 j_1}}{x_{i_2 j_2}} \right) + (r_2 - 1)(s_2 - 1) \sum_{i_1=1}^{r_1-1} \sum_{j_1=1}^{s_1-1} \text{var} \left(\ln \frac{x_{i_1 j_1}}{x_{r_2 s_2}} \right) + \\ & + (r_1 - 1)(s_1 - 1) \sum_{i_2=1}^{r_2-1} \sum_{j_2=1}^{s_2-1} \text{var} \left(\ln \frac{x_{r_1 s_1}}{x_{i_2 j_2}} \right) + (r_1 - 1)(r_2 - 1)(s_1 - 1)(s_2 - 1) \text{var} \left(\ln \frac{x_{r_1 s_1}}{x_{r_2 s_2}} \right), \end{aligned} \quad (16)$$

$$\begin{aligned} D_2 = & (s_1 - 1)(s_2 - 1) \sum_{i_1=1}^{r_1-1} \sum_{i_2=1}^{r_2-1} \text{var} \left(\ln \frac{x_{i_1 s_1}}{x_{i_2 s_2}} \right) + (s_1 - 1)(r_2 - 1) \sum_{i_1=1}^{r_1-1} \sum_{j_2=1}^{s_2-1} \text{var} \left(\ln \frac{x_{i_1 s_1}}{x_{r_2 j_2}} \right) + \\ & + (r_1 - 1)(s_2 - 1) \sum_{j_1=1}^{s_1-1} \sum_{i_2=1}^{r_2-1} \text{var} \left(\ln \frac{x_{r_1 j_1}}{x_{i_2 s_2}} \right) + (r_1 - 1)(r_2 - 1) \sum_{j_1=1}^{s_1-1} \sum_{j_2=1}^{s_2-1} \text{var} \left(\ln \frac{x_{r_1 j_1}}{x_{r_2 j_2}} \right) \end{aligned} \quad (17)$$

$$\text{and } K = \frac{1}{2} \frac{1}{\sqrt{r_1 r_2 s_1 s_2 (r_1 - 1)(r_2 - 1)(s_1 - 1)(s_2 - 1)}}.$$

Proof: The covariances are obtained using the general formula (8), where the corresponding

coefficient vectors \mathbf{a}^1 and \mathbf{a}^2 have elements

$$a_{ij}^k = \begin{cases} 1/\sqrt{r_k s_k (r_k - 1)(s_k - 1)} & \text{for } i = 1, \dots, r_k - 1 \quad j = 1, \dots, s_k - 1 \\ -(r_k - 1)/\sqrt{r_k s_k (r_k - 1)(s_k - 1)} & \text{for } i = r_k \quad j = 1, \dots, s_k - 1 \\ -(s_k - 1)/\sqrt{r_k s_k (r_k - 1)(s_k - 1)} & \text{for } i = 1, \dots, r_k - 1 \quad j = s_k \\ (r_k - 1)(s_k - 1)/\sqrt{r_k s_k (r_k - 1)(s_k - 1)} & \text{for } i = r_k \quad j = s_k \\ 0 & \text{otherwise,} \end{cases} \quad (18)$$

and $k = 1, 2$.

◇

Similarly as for the case of variances, there is a group of logratio variances that increases the overall covariance between coordinates (A_2 and B_2) and the remaining variances reduce it (C_2 and D_2). Specifically, for construction of logratios in A_2 the following parts are employed,

- an inner part of the first partial table and a part from the last column of the second partial table (except of x_{r_2, s_2}),
- an inner part of the first partial table and a part from the last row of the second partial table (except of x_{r_2, s_2}),
- the part x_{r_1, s_1} and a part from the last column of the second partial table (except of x_{r_2, s_2}),
- the part x_{r_1, s_1} and a part from the last row of the second partial table (except of x_{r_2, s_2}),

where we always deal with two “virtual” tables corresponding to the coordinates of interest. Similarly, B_2 is formed by variances of logratios of

- a part from the last column of the first partial table (except of x_{r_1, s_1}) and an inner part of the second partial table,
- a part from the last column of the first partial table (except of x_{r_1, s_1}) and the part x_{r_2, s_2} ,
- a part from the last row of the first partial table (except of x_{r_1, s_1}) and an inner part of the second partial table,
- a part from the last row of the first partial table (except of x_{r_1, s_1}) and the part x_{r_2, s_2} .

On the other hand, the covariance is reduced by C_2 , involving logratios between

- an inner part of the first partial table and an inner part of the second partial table,
- an inner part of the first table and the part x_{r_2, s_2} ,
- the part x_{r_1, s_1} and an inner part of the second partial table,
- parts x_{r_1, s_1} and x_{r_2, s_2} ,

and by D_2 consisting of logratios, formed by

- a part from the last column of the first partial table (except of x_{r_1, s_1}) and a part from the last column of the second partial table (except of x_{r_1, s_1}),
- a part from the last column of the first partial table (except of x_{r_1, s_1}) and a part from the last row of the second partial table (except of x_{r_2, s_2}),
- a part from the last row of the first partial table (except of x_{r_1, s_1}) and a part from the last column of the second partial table (except of x_{r_2, s_2}),

- a part from the last row of the first partial table (except of x_{r_1, s_1}) and a part from the last row of the second partial table (except of x_{r_2, s_2}).

Also covariance between two coordinates of the interaction table could be supported by its graphical representation, see Figure 2.

Since coordinates of the independence table (3), (4) are balances obtained from sequential binary partitions, dividing rows and columns of the original table, respectively (Egozcue and Pawłowsky-Glahn 2005), their variances and covariances are obtained as direct consequence of Fišerová and Hron (2011).

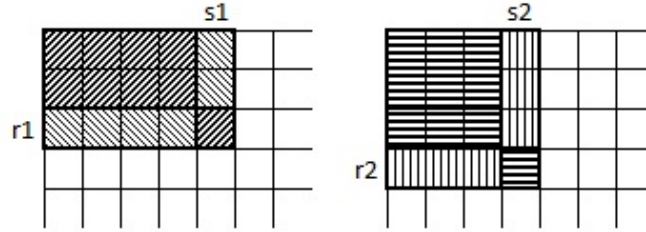


Figure 2: Covariance between coordinates z_{r_1, s_1} and z_{r_2, s_2} is increased by variances of logratios between a part of the first partial table from area highlighted by (/) and a part of the second partial table from area highlighted by (|) – A_2 . The second group of variances increasing the covariance between coordinates are connected to logratios between parts from (\) and (–) areas – B_2 . The covariance is reduced by variances of logratios between parts from (/) and (–) area – C_2 or two parts from (\) and (|) – D_2 .

Theorem 3.4 Consider coordinates of the independence table z_k^r for $k = 1, \dots, I - 1$ and z_l^c for $l = 1, \dots, J - 1$, then their variances are

$$\begin{aligned} \text{var}(z_k^r) &= K \sum_{i'=k+1}^I \sum_{j,j'=1}^J \text{var} \left(\ln \frac{x_{kj}}{x_{i'j'}} \right) - \frac{K}{2} (I - k) \sum_{j,j'=1}^J \text{var} \left(\ln \frac{x_{kj}}{x_{kj'}} \right) - \\ &\quad - \frac{K}{2(I - k)} \sum_{i,i'=k+1}^I \sum_{j,j'=1}^J \text{var} \left(\ln \frac{x_{ij}}{x_{i'j'}} \right), \end{aligned} \quad (19)$$

where $K = \frac{1}{J(I-k+1)}$, for balances between rows, and

$$\begin{aligned} \text{var}(z_l^c) &= K \sum_{i,i'=I}^I \sum_{j'=l+1}^J \text{var} \left(\ln \frac{x_{il}}{x_{i'j'}} \right) - \frac{K}{2} (J - l) \sum_{i,i'=1}^I \text{var} \left(\ln \frac{x_{il}}{x_{i'l}} \right) - \\ &\quad - \frac{K}{2(J - l)} \sum_{i,i'=1}^I \sum_{j,j'=l+1}^J \text{var} \left(\ln \frac{x_{ij}}{x_{i'j'}} \right), \end{aligned} \quad (20)$$

where $K = \frac{1}{I(J-l+1)}$, for balances between columns.

The variances of these coordinates are enlarged by variances of logratios between a part from the k -th row/ l -th column and any part from the subsequent rows/columns. On the other hand, the variances of z_k^r and z_l^c are reduced by variances of logratios between parts from the same row/column.

According to relation (8) there are three main options how to get covariance between coordinates of the independence table, depending on concrete balances of interest. All these possible covariances are summarized in the following theorem.

Theorem 3.5 Consider three coordinates of the independence table $z_{k_1}^r$, $z_{k_2}^r$ and z_k^r , for $k_1, k_2, k = 1, \dots, I - 1$, $k_1 \neq k_2$, computed using expression (3), and three coordinates $z_{l_1}^c$, $z_{l_2}^c$

and z_l^c , for $l_1, l_2, l = 1, \dots, J - 1, l_1 \neq l_2$, computed from (4). Then

$$\begin{aligned} \text{cov}(z_{k_1}^r, z_{k_2}^r) &= \frac{K}{(I - k_2)} \sum_{i'=k_2+1}^I \sum_{j,j'=1}^J \text{var} \left(\ln \frac{x_{k_1 j}}{x_{i' j'}} \right) + \frac{K}{(I - k_1)} \sum_{i=k_1+1}^I \sum_{j,j'=1}^J \text{var} \left(\ln \frac{x_{i j}}{x_{k_2 j'}} \right) - \\ &\quad - K \sum_{j,j'=1}^J \text{var} \left(\ln \frac{x_{k_1 j}}{x_{k_2 j'}} \right) - \frac{K}{(I - k_1)(I - k_2)} \sum_{i=k_1+1}^I \sum_{i'=k_2+1}^I \sum_{j,j'=1}^J \text{var} \left(\ln \frac{x_{i j}}{x_{i' j'}} \right), \end{aligned} \tag{21}$$

where $K = \frac{1}{2J} \sqrt{\frac{(I - k_1)(I - k_2)}{(I - k_1 + 1)(I - k_2 + 1)}}$, for row balances,

$$\begin{aligned} \text{cov}(z_{l_1}^c, z_{l_2}^c) &= \frac{K}{(J - l_2)} \sum_{i,i'=1}^I \sum_{j',j=l_2+1}^J \text{var} \left(\ln \frac{x_{i l_1}}{x_{i' j'}} \right) + \frac{K}{(J - l_1)} \sum_{i,i'=1}^I \sum_{j=l_1+1}^J \text{var} \left(\ln \frac{x_{i j}}{x_{i' l_2}} \right) - \\ &\quad - K \sum_{i,i'=1}^I \text{var} \left(\ln \frac{x_{i l_1}}{x_{i' l_2}} \right) - \frac{K}{(J - l_1)(J - l_2)} \sum_{i,i'=1}^I \sum_{j=l_1+1}^J \sum_{j'=l_2+1}^J \text{var} \left(\ln \frac{x_{i j}}{x_{i' j'}} \right), \end{aligned} \tag{22}$$

where $K = \frac{1}{2I} \sqrt{\frac{(J - l_1)(J - l_2)}{(J - l_1 + 1)(J - l_2 + 1)}}$, for column balances, and

$$\begin{aligned} \text{cov}(z_k^r, z_l^c) &= \frac{K}{(J - l)} \sum_{i'=1}^I \sum_{j=1}^J \sum_{j'=l+1}^J \text{var} \left(\ln \frac{x_{k j}}{x_{i' j'}} \right) + \frac{K}{(I - k)} \sum_{i=1}^I \sum_{i'=k+1}^I \sum_{j=1}^J \text{var} \left(\ln \frac{x_{i l}}{x_{i' j'}} \right) - \\ &\quad - K \sum_{i'=1}^I \sum_{j=1}^J \text{var} \left(\ln \frac{x_{k j}}{x_{i' l}} \right) - \frac{K}{(I - k)(J - l)} \sum_{i=k+1}^I \sum_{i'=1}^I \sum_{j=1}^J \sum_{j'=l+1}^J \text{var} \left(\ln \frac{x_{i j}}{x_{i' j'}} \right), \end{aligned} \tag{23}$$

where $K = \frac{1}{2} \sqrt{\frac{(I - k)(J - l)}{IJ(I - k + 1)(J - l + 1)}}$, between row and column balances.

To complete the covariance structure of coordinates of the compositional table \mathbf{x} , covariances between coordinates of the interaction and independence tables are necessary. They are provided in the last theorem.

Theorem 3.6 Consider coordinate of the interaction table z_{rs} , for $r = 2, \dots, I$ and $s = 2, \dots, J$, and two coordinates of the independence table, z_k^r , for $k = 1, \dots, I - 1$, and z_l^c , for $l = 1, \dots, J - 1$. Then for covariances between coordinates of the interaction and independence tables the following hold,

$$\text{cov}(z_{rs}, z_k^r) = K \cdot (A_3 - B_3), \tag{24}$$

where

$$\begin{aligned} A_3 &= \frac{1}{J(I - k)} \sum_{i=1}^{r-1} \sum_{i'=k+1}^I \sum_{j=1}^{s-1} \sum_{j'=1}^J \text{var} \left(\ln \frac{x_{i j}}{x_{i' j'}} \right) + \frac{s-1}{J} \sum_{i=1}^{r-1} \sum_{j'=1}^J \text{var} \left(\ln \frac{x_{i s}}{x_{k j'}} \right) + \\ &\quad + \frac{r-1}{J} \sum_{j=1}^{s-1} \sum_{j'=1}^J \text{var} \left(\ln \frac{x_{r j}}{x_{k j'}} \right) + \frac{(r-1)(s-1)}{J(I - k)} \sum_{i'=k+1}^I \sum_{j'=1}^J \text{var} \left(\ln \frac{x_{rs}}{x_{i' j'}} \right), \end{aligned} \tag{25}$$

$$\begin{aligned} B_3 &= \frac{1}{J} \sum_{i=1}^{r-1} \sum_{j=1}^{s-1} \sum_{j'=1}^J \text{var} \left(\ln \frac{x_{i j}}{x_{k j'}} \right) + \frac{s-1}{J(I - k)} \sum_{i=1}^{r-1} \sum_{i'=k+1}^I \sum_{j'=1}^J \text{var} \left(\ln \frac{x_{i s}}{x_{i' j'}} \right) + \\ &\quad + \frac{r-1}{J(I - k)} \sum_{i'=k+1}^I \sum_{j=1}^{s-1} \sum_{j'=1}^J \text{var} \left(\ln \frac{x_{r j}}{x_{i' j'}} \right) + \frac{(r-1)(s-1)}{J} \sum_{j'=1}^J \text{var} \left(\ln \frac{x_{rs}}{x_{k j'}} \right), \end{aligned} \tag{26}$$

for $K = \frac{1}{2} \frac{1}{\sqrt{rs(r-1)(s-1)}} \sqrt{\frac{J(I-k)}{I-k+1}}$, and

$$\text{cov}(z_{rs}, z_l^c) = K \cdot (A_4 - B_4), \quad (27)$$

where

$$\begin{aligned} A_4 = & \frac{1}{I(J-l)} \sum_{i=1}^{r-1} \sum_{i'=1}^I \sum_{j=1}^{s-1} \sum_{j'=l+1}^J \text{var} \left(\ln \frac{x_{ij}}{x_{i'j'}} \right) + \frac{s-1}{I} \sum_{i=1}^{r-1} \sum_{i'=1}^I \text{var} \left(\ln \frac{x_{is}}{x_{i'l}} \right) + \\ & + \frac{r-1}{I} \sum_{i'=1}^I \sum_{j=1}^{s-1} \text{var} \left(\ln \frac{x_{rj}}{x_{i'l}} \right) + \frac{(r-1)(s-1)}{I(J-l)} \sum_{i'=1}^I \sum_{j'=l+1}^J \text{var} \left(\ln \frac{x_{rs}}{x_{i'j'}} \right), \end{aligned} \quad (28)$$

$$\begin{aligned} B_4 = & \frac{1}{I} \sum_{i=1}^{r-1} \sum_{i'=1}^I \sum_{j=1}^{s-1} \text{var} \left(\ln \frac{x_{ij}}{x_{i'l}} \right) + \frac{s-1}{I(J-l)} \sum_{i=1}^{r-1} \sum_{i'=1}^I \sum_{j'=l+1}^J \text{var} \left(\ln \frac{x_{is}}{x_{i'j'}} \right) + \\ & + \frac{r-1}{I(J-l)} \sum_{i'=1}^I \sum_{j=1}^{s-1} \sum_{j'=l+1}^J \text{var} \left(\ln \frac{x_{rj}}{x_{i'j'}} \right) + \frac{(r-1)(s-1)}{I} \sum_{i'=1}^I \text{var} \left(\ln \frac{x_{rs}}{x_{i'l}} \right), \end{aligned} \quad (29)$$

for $K = \frac{1}{2} \frac{1}{\sqrt{rs(r-1)(s-1)}} \sqrt{\frac{I(J-l)}{J-l+1}}$.

Proof: The assertion of the theorem is a direct consequence of Proposition 3.1 and Equations (3), (4) and (5). ◇

Similarly as for the case of interaction table, also the above results could be interpreted graphically. Because Theorems 3.4 and 3.5 represent a special case of balances, that were in detail analyzed in (Fišerová and Hron 2011), in Figure 3 we focus just on covariances, resulting from Theorem 3.6.

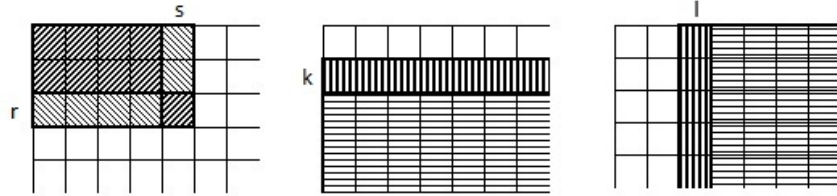


Figure 3: Covariance between a coordinate of the interaction table, z_{rs} (left), and coordinates of the independence table, z_k^r (middle) or z_l^c (right), is increased by variances of logratios between parts from areas (/) and (-), or (\) and (|), respectively, and reduced by variances of logratios between parts from areas (/) and (|), or (\) and (-), respectively.

4. Implications for 2 x 2 compositional tables

In practice, 2×2 compositional (and also contingency) tables represent a prominent special case that requires a special treatment (Fačevicová *et al.* 2014a; Agresti 2002). From Equations (3), (4) and (5) it is easy to see that for coordinate representation of the compositional table

$$\mathbf{x} = \mathcal{C} \begin{pmatrix} x_{11} & x_{12} \\ x_{21} & x_{22} \end{pmatrix}$$

it is sufficient to consider the following coordinates,

$$z_1^{ind} = \frac{1}{2} \ln \frac{x_{11}x_{12}}{x_{21}x_{22}}, \quad z_2^{ind} = \frac{1}{2} \ln \frac{x_{11}x_{21}}{x_{12}x_{22}} \quad \text{and} \quad z^{int} = \frac{1}{2} \ln \frac{x_{11}x_{22}}{x_{12}x_{21}}. \quad (30)$$

By applying the above theorems their covariance structure can be easily derived,

$$\begin{aligned} \text{var}(z^{int}) &= \frac{1}{4} \left[\text{var} \left(\ln \frac{x_{11}}{x_{21}} \right) + \text{var} \left(\ln \frac{x_{11}}{x_{12}} \right) + \text{var} \left(\ln \frac{x_{21}}{x_{22}} \right) + \text{var} \left(\ln \frac{x_{12}}{x_{22}} \right) \right. \\ &\quad \left. - \text{var} \left(\ln \frac{x_{11}}{x_{22}} \right) - \text{var} \left(\ln \frac{x_{12}}{x_{21}} \right) \right], \\ \text{var}(z_1^{ind}) &= \frac{1}{4} \left[\text{var} \left(\ln \frac{x_{11}}{x_{21}} \right) + \text{var} \left(\ln \frac{x_{11}}{x_{22}} \right) + \text{var} \left(\ln \frac{x_{12}}{x_{21}} \right) + \text{var} \left(\ln \frac{x_{12}}{x_{22}} \right) \right. \\ &\quad \left. - \text{var} \left(\ln \frac{x_{11}}{x_{12}} \right) - \text{var} \left(\ln \frac{x_{21}}{x_{22}} \right) \right], \\ \text{var}(z_2^{ind}) &= \frac{1}{4} \left[\text{var} \left(\ln \frac{x_{11}}{x_{12}} \right) + \text{var} \left(\ln \frac{x_{11}}{x_{22}} \right) + \text{var} \left(\ln \frac{x_{21}}{x_{12}} \right) + \text{var} \left(\ln \frac{x_{21}}{x_{22}} \right) \right. \\ &\quad \left. - \text{var} \left(\ln \frac{x_{11}}{x_{21}} \right) - \text{var} \left(\ln \frac{x_{12}}{x_{22}} \right) \right], \\ \text{cov}(z^{int}, z_1^{ind}) &= \frac{1}{4} \left[\text{var} \left(\ln \frac{x_{11}}{x_{21}} \right) - \text{var} \left(\ln \frac{x_{22}}{x_{12}} \right) \right], \\ \text{cov}(z^{int}, z_2^{ind}) &= \frac{1}{4} \left[\text{var} \left(\ln \frac{x_{11}}{x_{12}} \right) - \text{var} \left(\ln \frac{x_{21}}{x_{22}} \right) \right], \\ \text{cov}(z_1^{ind}, z_2^{ind}) &= \frac{1}{4} \left[\text{var} \left(\ln \frac{x_{11}}{x_{21}} \right) - \text{var} \left(\ln \frac{x_{12}}{x_{21}} \right) \right]. \end{aligned}$$

Moreover, from the above covariance structure it is also interesting to see that coordinates (30) are uncorrelated (or even independent under the assumption of normality) if, and only if

$$\text{var} \left(\ln \frac{x_{11}}{x_{21}} \right) = \text{var} \left(\ln \frac{x_{12}}{x_{22}} \right), \quad \text{var} \left(\ln \frac{x_{11}}{x_{12}} \right) = \text{var} \left(\ln \frac{x_{21}}{x_{22}} \right), \quad \text{var} \left(\ln \frac{x_{11}}{x_{22}} \right) = \text{var} \left(\ln \frac{x_{12}}{x_{21}} \right). \quad (31)$$

In other words, it means that zero covariances can be easily expressed in terms of logratio variances. Consequently, the above relations could be used, e.g., by designing simulation settings for 2×2 compositional tables using elements of the variation matrix as a source of elemental information in covariance structure of compositional tables.

Following Fačevicová *et al.* (2014a), it is possible to assign also another system of orthonormal coordinates to a 2×2 compositional table. Specifically, we get

$$z_1^{ind} = \frac{1}{\sqrt{2}} \ln \frac{x_{12}}{x_{21}}, \quad z_2^{ind} = \frac{1}{\sqrt{2}} \ln \frac{x_{11}}{x_{22}}, \quad z^{int} = \frac{1}{2} \ln \frac{x_{11}x_{22}}{x_{12}x_{21}}, \quad (32)$$

for the interaction and independent tables, respectively, and the covariance structure changes as follows,

$$\begin{aligned} \text{var}(z^{int}) &= \frac{1}{4} \left[\text{var} \left(\ln \frac{x_{11}}{x_{12}} \right) + \text{var} \left(\ln \frac{x_{11}}{x_{21}} \right) + \text{var} \left(\ln \frac{x_{12}}{x_{22}} \right) + \text{var} \left(\ln \frac{x_{21}}{x_{22}} \right) \right. \\ &\quad \left. - \text{var} \left(\ln \frac{x_{11}}{x_{22}} \right) - \text{var} \left(\ln \frac{x_{12}}{x_{21}} \right) \right], \\ \text{var}(z_1^{ind}) &= \frac{1}{2} \text{var} \left(\ln \frac{x_{12}}{x_{21}} \right), \\ \text{var}(z_2^{ind}) &= \frac{1}{2} \text{var} \left(\ln \frac{x_{11}}{x_{22}} \right), \\ \text{cov}(z^{int}, z_1^{ind}) &= \frac{1}{4\sqrt{2}} \left[\text{var} \left(\ln \frac{x_{11}}{x_{12}} \right) + \text{var} \left(\ln \frac{x_{11}}{x_{21}} \right) - \text{var} \left(\ln \frac{x_{12}}{x_{22}} \right) - \text{var} \left(\ln \frac{x_{21}}{x_{22}} \right) \right], \\ \text{cov}(z^{int}, z_2^{ind}) &= \frac{1}{4\sqrt{2}} \left[\text{var} \left(\ln \frac{x_{11}}{x_{21}} \right) + \text{var} \left(\ln \frac{x_{21}}{x_{22}} \right) - \text{var} \left(\ln \frac{x_{11}}{x_{12}} \right) - \text{var} \left(\ln \frac{x_{12}}{x_{22}} \right) \right], \\ \text{cov}(z_1^{ind}, z_2^{ind}) &= \frac{1}{4} \left[\text{var} \left(\ln \frac{x_{11}}{x_{21}} \right) + \text{var} \left(\ln \frac{x_{12}}{x_{22}} \right) - \text{var} \left(\ln \frac{x_{11}}{x_{12}} \right) - \text{var} \left(\ln \frac{x_{21}}{x_{22}} \right) \right]. \end{aligned}$$

Now, although coordinates of the independent table are formed just by (scaled) logratios, the covariance structure becomes more complex than before. For example, coordinates (32) are *mutually* uncorrelated (independent) if, and only if

$$\text{var} \left(\ln \frac{x_{11}}{x_{12}} \right) = \text{var} \left(\ln \frac{x_{11}}{x_{21}} \right) = \text{var} \left(\ln \frac{x_{12}}{x_{22}} \right) = \text{var} \left(\ln \frac{x_{21}}{x_{22}} \right). \quad (33)$$

In other words, it means that $\text{var} \left(\ln \frac{x_{12}}{x_{21}} \right)$ and $\text{var} \left(\ln \frac{x_{11}}{x_{22}} \right)$ are influential just for variances of coordinates z_1^{ind} , z_2^{ind} , z^{int} , forming also natural constraints for their possible values.

5. Numerical example

To illustrate the presented theoretical outputs, let us consider the sample of eighteen 2×3 compositional tables, each reflecting population structure in European country according to age and BMI index ((weight in kg)/(height in m)²), with values 25 – 44, 45 – 64, 65 – 84 and under- or normal weight and overweight or obesity, respectively. The data set is an aggregated version of data from Fačevicová *et al.* (2014b). Table 1 shows an example of compositional table from the sample.

Table 1: Structure of population in Austria in 2008 according to age and BMI index (in proportions).

| AUT | 25 – 44 | 45 – 64 | 65 – 84 |
|-----------------|---------|---------|---------|
| under or normal | 0.249 | 0.144 | 0.074 |
| over or obesity | 0.171 | 0.221 | 0.140 |

Firstly, each table from the sample has been expressed in coordinates and, consequently, their descriptive statistics has been calculated. The sample mean is

$$\bar{\mathbf{z}} = (0.409, 0.294, -0.450, 0.578, 0.637),$$

but for our purposes the covariance structure of the sample is of primary interest. The variation matrix (6), as a source of elemental information in compositional tables, equals

$$\mathbf{T} = \begin{pmatrix} 0 & 0.037 & 0.083 & 0.024 & 0.030 & 0.069 \\ 0.037 & 0 & 0.030 & 0.077 & 0.050 & 0.051 \\ 0.083 & 0.030 & 0 & 0.127 & 0.098 & 0.065 \\ 0.024 & 0.077 & 0.127 & 0 & 0.019 & 0.078 \\ 0.030 & 0.050 & 0.098 & 0.019 & 0 & 0.040 \\ 0.069 & 0.051 & 0.065 & 0.078 & 0.040 & 0 \end{pmatrix}.$$

For example, using this matrix and equation (9), variance of the first coordinate of the interaction table, z_{22} , can be obtained as

$$\begin{aligned} \text{var}(z_{rs}) &= \frac{1}{4} \text{var} \left(\ln \frac{x_{11}}{x_{21}} \right) + \frac{1}{4} \text{var} \left(\ln \frac{x_{11}}{x_{12}} \right) + \frac{1}{4} \text{var} \left(\ln \frac{x_{21}}{x_{22}} \right) + \frac{1}{4} \text{var} \left(\ln \frac{x_{12}}{x_{22}} \right) \\ &\quad - \frac{1}{8} \text{var} \left(\ln \frac{x_{11}}{x_{11}} \right) - \frac{1}{4} \text{var} \left(\ln \frac{x_{11}}{x_{22}} \right) - \frac{1}{8} \text{var} \left(\ln \frac{x_{21}}{x_{21}} \right) - \frac{1}{8} \text{var} \left(\ln \frac{x_{12}}{x_{12}} \right) \\ &\quad - \frac{1}{4} \text{var} \left(\ln \frac{x_{12}}{x_{21}} \right) \\ &= \frac{1}{4} t_{14} + \frac{1}{4} t_{12} + \frac{1}{4} t_{45} + \frac{1}{4} t_{25} - \frac{1}{8} t_{11} - \frac{1}{4} t_{15} - \frac{1}{8} t_{44} - \frac{1}{8} t_{22} - \frac{1}{4} t_{24} \\ &= 0.0057. \end{aligned}$$

By comparing with the corresponding elements of the variation matrix we can conclude that none of logratios contributes exceptionally (in the positive sense) to variability of the coordinate. In the negative sense, the logratio $\ln(\text{underweight or normal weight in age 45-64}/\text{overweight or obesity in age 25-44})$ shows a dominant effect. Similarly, also other variances and covariances can be derived (and further analysed for structural patterns), resulting in a covariance matrix

$$\text{var}(\mathbf{z}) = \begin{pmatrix} 0.006 & 0.003 & -0.010 & 0.007 & 0.004 \\ 0.003 & 0.013 & -0.007 & -0.004 & -0.0005 \\ -0.010 & -0.007 & 0.051 & -0.021 & -0.012 \\ 0.007 & -0.004 & -0.021 & 0.051 & 0.026 \\ 0.004 & -0.0005 & -0.012 & 0.026 & 0.026 \end{pmatrix}.$$

Finally, note that by considering both markedly nonzero means of coordinates of the interaction table (first two elements of the vector $\bar{\mathbf{z}}$) and their corresponding small variances, we can conclude that, based on the considered sample, age and BMI index are not dependent.

6. Discussion

Recent experience with orthonormal coordinates for compositional data (Reimann, Filzmoser, Fabian, Hron, Birke, Demetriades, Dinelli, and Ladenberger 2012; Filzmoser and Walczak 2014) shows clearly the necessity of their better understanding in terms of logratios, which could be achieved also by decomposing the corresponding covariance structure. This is even more crucial for compositional tables, where both balances and coordinates with odds ratio interpretation are involved. Obviously, due to complex character of the above formulas for covariance structure in compositional tables, they will be rather rarely used for practical computations. Therefore, the formulas are also accompanied with comments and graphical illustrations to better understand their logical structure that is much more important for the aim of the paper. Consequently, similarly as for the case of balances (Fišerová and Hron 2011), we are convinced that decomposition of variances and covariances as linear combinations of logratio variances enhances interpretability of coordinates of compositional tables, using logratios as the primary source of information in compositional data.

Acknowledgments Authors gratefully acknowledge the support of the Operational Program Education for Competitiveness - European Social Fund (project CZ.1.07/2.3.00/20.0170 of the Ministry of Education, Youth and Sports of the Czech Republic) and the grant PrF_2014_028 Mathematical Models of the Internal Grant Agency of the Palacký University in Olomouc.

References

- Agresti A (2002). *Categorical Data Analysis (2 ed.)*. J. Wiley & Sons, New York.
- Aitchison J (1986). *The Statistical Analysis of Compositional Data*. Chapman and Hall, London.
- Egozcue J, Díaz-Barrero J, Pawłowsky-Glahn V (2008). "Simplicial Geometry for Compositional Data." In *Daunis-i-Estadella J, Martín-Fernández JA (eds) Proceedings of CODA-WORK'08, The 3rd Compositional Data Analysis Workshop*. University of Girona, Spain.
- Egozcue J, Pawłowsky-Glahn V (2005). "Groups of Parts and Their Balances in Compositional Data Analysis." *Mathematical Geology*, **37**, 795–828.
- Egozcue J, Pawłowsky-Glahn V (2006). "Simplicial Geometry for Compositional Data." In *Buccianti A, Mateu-Figueras G, Pawłowsky-Glahn V (eds) Compositional data analysis in the geosciences: From theory to practice*. Geological Society, London.

- Egozcue J, Pawlowsky-Glahn V, Mateu-Figueras G, Barceló-Vidal C (2003). “Isometric Logratio Transformations for Compositional Data Analysis.” *Mathematical Geology*, **35**, 279–300.
- Egozcue J, Pawlowsky-Glahn V, Templ M, Hron K (2014). “Independence in Contingency Tables using Simplicial Geometry.” *Communications in Statistics - Theory and Methods*, to appear.
- Fačevicová K, Hron K, Todorov V, Guo D, Templ M (2014a). “Logratio Approach to Statistical Analysis of 2×2 Compositional Tables.” *Journal of Applied Statistics*, **41**, 944–958.
- Fačevicová K, Hron K, Todorov V, Templ M (2014b). “Compositional Tables Analysis in Coordinates.” *Submitted*.
- Filzmoser P, Walczak B (2014). “What Can Go Wrong at the Data Normalization Step for Identification of Biomarkers?” *Journal of Chromatography A*, **1362**, 194–205.
- Fišerová E, Hron K (2011). “On Interpretation of Orthonormal Coordinates for Compositional Data.” *Math Geol*, **43**, 455–468.
- Pawlowsky-Glahn V, Buccianti A (2011). *Compositional Data Analysis: Theory and Applications*. Wiley, Chichester.
- Reimann C, Filzmoser P, Fabian K, Hron K, Birke M, Demetriades A, Dinelli E, Ladenberger A (2012). “The Concept of Compositional Data Analysis in Practice – Total Major Element Concentrations in Agricultural and Grazing Land Soils of Europe.” *Science of the Total Environment*, **426**, 196–210.

Affiliation:

Kamila Fačevicová

Department of Mathematical Analysis and Applications of Mathematics, Department of Geoinformatics

Faculty of Science, Palacký University

Olomouc, Czech Republic

E-mail: kamila.facevicova@gmail.com