J. Indones. Math. Soc.
Vol. 31, No. 02 (2025), pp. 1–20.

Tucker3 Tensor Decomposition for the Standardized Residual Hypermatrix on Three-Way Correspondence Analysis

Karunia Eka Lestari[1*], Mokhammad Ridwan Yudhanegara[2], Edwin Setiawan Nugraha [3], and Sisilia Sylviani [4]

[1]Department of Mathematics Education, University of Singaperbangsa Karawang, Indonesia.
[1]karunia@staff.unsika.ac.id, [2]mridwan.yudhanegara@staff.unsika.ac.id
[3]Study Program of Actuarial Science, President University, Indonesia.
edwin.nugraha@president.ac.id
[4]Department of Mathematics, Padjadjaran University, Indonesia.
sisilia.sylviani@unpad.ac.id

Abstract. This study investigates the theoretical and practical mathematical aspects of Tucker3 tensor decomposition from the three-way correspondence analysis point of view. Since the standardized residual hypermatrix represents the association of the three categorical variables, this study focused on (1) Tucker3 tensor decomposition for the standardized residual hypermatrix, (2) some mathematical properties of Tucker3 tensor decomposition, and (3) constructing the correspondence plot via Tucker3 tensor decomposition. Some mathematical results are presented in lemmas, theorems and algorithms, while a practical result is exhibited at the end of the discussion.

Key words and Phrases: Tucker3 decomposition, tensor and hypermatrix, three-way correspondence analysis, categorical data analysis

1. INTRODUCTION

Statistical and graphical analysis of associations between categorical variables has a long and interesting history. The contributions of several statisticians, including Karl Pearson and Ronald Aylmer Fisher, have left a trail on how the analysis of categorical data is carried out, including its graphical representation. These experts have produced various statistical techniques to measure, model, examine and

*Corresponding author
2020 Mathematics Subject Classification: 62H25
Received: 15-08-2023, accepted: 18-01-2025.

visualize how the variables are related. Some of these statistical techniques involve contingency tables of categorical data in their analysis.

There are various techniques for measuring associations in categorical data, as have been proposed in [1, 2, 3]. These techniques generally involve calculating statistical measures that quantify the magnitude of interrelation among variables. On the other side, obtaining a graphical representation of the data can help users better visually understand this association's nature.

Correspondence analysis (CA) is mostly used to explore associations between categorical variables statistically and graphically. CA provides intuitive visual observations of the associations between variables at the category level. CA is adequately flexible to be used on large data matrices since it only requires data input in the contingency table [4, 5]. If a contingency table consists of two categorical variables, the technique is well-known as two-way correspondence analysis (CA2) or simple correspondence analysis (SCA).

The data in a two-way contingency table where consisting I rows and J columns can be considered as a data matrix $\mathbf{N}$ of size $I \times J$. By performing a statistical procedure that involves matrix operations, one could obtain: (1) Row or column profiles that represent the marginal distribution for each row or column category; (2) Principal coordinates which are linear combinations of the eigenvectors and centered row or column profiles; and (3) CA plot resulting from principal coordinate mapping on d-dimensional plot. However, the CA plots that can be presented visually are limited to $d = 1, 2, 3$.

In practice, one can display the CA plot visualization in one, two or three dimensions according to the required data analysis needs. Nevertheless, this impacts the absorption level of information generated through the plot [6]. Thus, the main problem in CA is how the existing plot can represent the rows or columns in the contingency table in a low-dimensional space (dimensional reduction) while optimally absorbing as much information about their association structure as possible.

Beh & Lombardo [4] focused on solving dimension-reduction problems using matrix decomposition since the essence of all decomposition methods is to reduce dimensions. Several decomposition methods were developed for a two-way contingency table where consisting of I row categories and J column categories, including eigendecomposition [7], singular value decomposition [8], bivariate moment decomposition [9], and hybrid decomposition [10]. In the CA context, the decomposition is performed on the standard residual matrix $\mathbf{S}$, representing the linear associations between each row and column category.

In line with today's advances in technology and information, various data can be accessed easily by anyone at any time. It also influenced CA research and development. Problems are even more complex when dealing with data consisting of three categorical variables. When this problem is solved using CA2, it requires an extensive matrix calculation process, as well as raises two significant question, including: (1) how to present each separate plot in a whole plot comprehensively? and (2) does the absorption level of information after merging in a whole plot decrease?

Another alternative to solving such a problem is using multiple correspondence analysis (MCA). MCA's main idea is to group data through coding techniques to be arranged as a two-way contingency table. Thus, the procedure on CA2 can be applied later. In the context of MCA, "coding techniques" refer to the manners in which categorical data are prepared, restructured, and represented as a numerical matrix such that it can be analyzed mathematically. Since MCA involves the analysis of associations among categorical variables, coding techniques is fundamental for transforming qualitative data into a form that allows for mathematical operations [11, 12]. One of the coding techniques commonly used in MCA is the Burt matrix [11]. Nevertheless, the use of MCA also has several limitations, as stated by [13, 14], including (1) there is no information obtained regarding interactions between multiple variables; (2) the absorption level of information on lower dimensions is not optimal, and (3) impractical to use in big data. Therefore, Beh & Lombardo [4] proposed solving problems related to three categorical variables using three-way correspondence analysis (CA3) since it can provide a merge plot display of each categorical variable in a whole plot with the same dimensions and more optimal information absorption [4].

Analogous to CA2, to reveal the association between three categorical variables graphically, the determination of the principal coordinates and the seeking for low-dimensional space in CA3 also be solved by matrix decomposition. Rather, the decomposition method in CA2 is no longer appropriate when applied to CA3. As a consequence, particular attention will be paid to the three-way generalization of the singular value decomposition or, especially, the Tucker3 decomposition [15], which De Lathauwer et al. [16] have mentioned as a higher-order SVD (HOSVD) [16]. The decomposition on CA3 was undertaken on the standard residual hypermatrix $\boldsymbol{\mathcal{S}}$, which reflects the association between the three observed categorical variables. In the process, this method involves algebraic calculations such as tensor operations and hypermatrix properties. This research extends the field of CA3 by offering a structured tensor-based approach that simplifies, improves interpretinability, and allows for more effective visualization of three-way categorical data. Such approach harnesses the Tucker3 decomposition to address persistent challenges in standardized residual analysis within correspondence analysis and potentially expanding the applicability of CA3 across multiple disciplines. Thus, this study will focus on (1) Tucker3 tensor decomposition for $\boldsymbol{\mathcal{S}}$, (2) some mathematical properties of Tucker3 tensor decomposition, and (3) constructing the CA3 plot via Tucker3 tensor decomposition. These three core problems become urgent in this research, which yield theoretical and practical mathematical novelty.

2. THEORETICAL FRAMEWORK

2.1. **Hypermatrix and related tensor operations.**

The hypermatrix is a generalization of a matrix of order $n_1 \times n_2 \times \cdots n_d$, where $n_1, n_2, \cdots, n_d \in \mathbb{N}$. In data analysis, a hypermatrix can be viewed as a

representation of a d-order tensor for $d > 2$. Formally, a hypermatrix is defined as follows.

Definition 2.1. *Let* $n_1, \cdots n_d \in \mathbb{N}$.

$$f : \langle n_1 \rangle \times \cdots \times \langle n_d \rangle \to F$$

is an order-d of hypermatrix or d-hypermatrix

Hypermatrix is denoted by boldface Euler's letters, for example, $\boldsymbol{\mathcal{S}}$. The element of $\boldsymbol{\mathcal{S}}$ with order-d is denoted as $s_{m_1,\cdots,m_d}$ representing a value of the function $f(m_1, \cdots, m_d)$ with $m_1 \in \langle n_1 \rangle, \cdots, m_d \in \langle n_d \rangle$.Thus, a d-hypermatrix can be written as $\left[S_{m_1,\ldots,m_d}\right]_{m_1,\ldots,m_d=1}^{n_1,..,n_d}$ or $\boldsymbol{\mathcal{S}} = \left[s_{m_1,\ldots,m_d}\right]$. The set of all d-hypermatrix with domain $\langle n_1 \rangle \times \cdots \times \langle n_d \rangle$ is denoted by $\mathbb{F}^{n_1 \times \cdots \times n_d}$. If $n_1 = n_2 = ... = n_d = n$ then the hypermatrix $\boldsymbol{\mathcal{S}} \in \mathbb{F}^{n \times \cdots \times n}$ is called hypercubical or cubical of dimension n [17].

A hypermatrix $\boldsymbol{\mathcal{S}} = \left[s_{m_1,\ldots,m_d}\right]$ can be transformed into matrix $\mathbf{S}_{(d)}$. This process is called matricization. In some references, it is also called unfolding or flattening [18, 19]. The rearrangement of the $\boldsymbol{\mathcal{S}}$ elements into columns in the matrix $\mathbf{S}_{(d)}$ is represented by mode-d fibers. Figure 1 illustrates the matricization process of $\boldsymbol{\mathcal{S}} \in \mathbb{R}^{2\times2\times2}$ for generating $\mathbf{S}_{(1)}, \mathbf{S}_{(2)}$, and $\mathbf{S}_{(3)}$. Moreover, a mode-1 fibers $\mathbf{S}_{(1)}$ is called a column fiber, mode-2 fiber $\mathbf{S}_{(2)}$ is a row fiber, and mode-3 fiber $\mathbf{S}_{(3)}$ is a tube fiber.

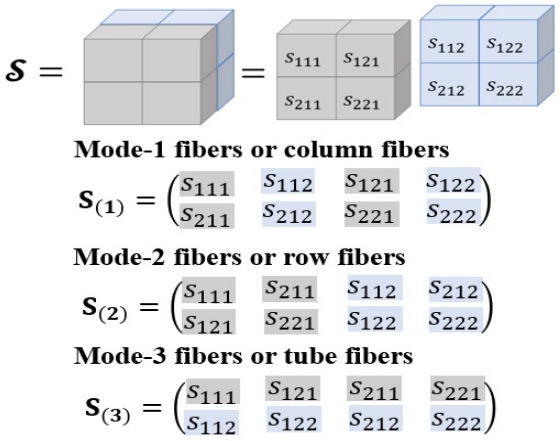

FIGURE 1. Illustration of matricization for hypermatrix $\boldsymbol{\mathcal{S}} \in \mathbb{R}^{2\times2\times2}$ the column, row, or tube fibers are aligned to form $\mathbf{S}_{(1)}, \mathbf{S}_{(2)}$ or $\mathbf{S}_{(3)}$ matrices

Several references use the rearranging of the columns for different mode-d fibers. As a comparison, see [16, 20]. Basically, the permutation of a particular column is not important as long as it is consistent in related calculations; for more details, see [21]. The structure, operations, and properties inherent in a hypermatrix can be found in [17]. The following discussion will focus on the structure and

operation of the hypermatrix required in the Tucker3 decomposition on CA3. The tensor multiplication or the d-mode product of $\boldsymbol{\mathcal{S}}\in\mathbb{R}^{n_1\times\cdots\times n_d}$ is defined as follows.

Definition 2.2. *Suppose* $\boldsymbol{\mathcal{S}}\in\mathbb{R}^{n_1\times\cdots\times n_d}$ *and* $\mathbf{U}\in\mathbb{R}^{j\times n_d}$. *The d-mode product of hypermatrix* $\boldsymbol{\mathcal{S}}$ *with a matrix* $\mathbf{U}$ *is denoted by* $\boldsymbol{\mathcal{S}}\times_d\mathbf{U}$, *and is defined as*

$$\boldsymbol{\mathcal{S}}\times_d\mathbf{U}=\sum_{n_d=1}^{N_d}s_{n_1\cdots n_d}u_{jn_d}$$

where $n_1\times\cdots\times n_{d-1}\times J\times n_{d+1}\times\cdots\times n_d$ *is the size of* $\boldsymbol{\mathcal{S}}\times_d\mathbf{U}$.

Definition 2.2 asserts that each mode-d fiber is multiplied by a matrix $\mathbf{U}$. Kolda & Bader [18] apply this idea to $S_{(d)}$ such that it yields an equivalent,

$$\boldsymbol{\chi}=\boldsymbol{\mathcal{S}}\times_d\mathbf{U}\Leftrightarrow\mathbf{X}_{(d)}=\mathbf{U}\mathbf{S}_{(d)}. \tag{1}$$

If $\mathbf{V}\in\mathbb{R}^{Q\times J}$ for $Q\in\mathbb{N}$, then

$$\boldsymbol{\mathcal{S}}\times_{n_1}\mathbf{U}\times_{n_2}\mathbf{V}=\begin{cases}\boldsymbol{\mathcal{S}}\times_{n_1}(\mathbf{V}\mathbf{U}), & for\ n_1=n_2\\ \boldsymbol{\mathcal{S}}\times_{n_2}\mathbf{V}\times_{n_1}\mathbf{U}, & for\ n_1\neq n_2\end{cases} \tag{2}$$

The tensor product of matrices $\mathbf{U}$ and $\mathbf{V}$ is defined as the following Kronecker product.

Definition 2.3. *Given* $\boldsymbol{U}\in\mathbb{R}^{P\times I}$ *and* $\boldsymbol{V}\in\mathbb{R}^{Q\times J}$ *where* $I,J,P,Q\in\mathbb{N}$. *The Kronecker product of* $\boldsymbol{U}$ *and* $\boldsymbol{V}$ *is denoted by* $\boldsymbol{U}\otimes\boldsymbol{V}\in\mathbb{R}^{PQ\times IJ}$ *and is defined as*

$$\mathbf{U}\otimes\mathbf{V}=\begin{bmatrix}u_{11}V & u_{12}V & \cdots & u_{11}V\\ u_{21}V & u_{22}V & & u_{21}V\\ \vdots & & \ddots & \vdots\\ u_{P1}V & u_{p2}V & \cdots & u_{P1}V\end{bmatrix} \tag{3}$$

2.2. **Recognition and notation of three-way correspondence.**

In the late 1990s, CA3 was first introduced by Carlier and Kroonenberg as a generalization of CA2 [22]. CA3 has been proposed to analyze the $I\times J\times K$ contingency table using a three-mode principal component (Tucker3 model) or parallel factor analysis model (PARAFAC), further studied in [15, 22, 23, 24, 25]. CA3 is commonly used for categorical data or categorized numerical data. In 2017, Lombardo & Beh have conducted a CA3 for ordinal-nominal variables [26]. CA3 can also be used for nonlinear associations or as a data grouping and reduction technique. D'Ambra et al. used CA3 to reduce dimensions in an ordinal three-way contingency table [27].

The CA3 input is a three-way contingency table consisting of I row, J column, and K tube categories. The rows, columns, and tubes of such table represent X,Y and Z categorical variable, respectively. Therefore, the $I\times J\times K$ three-way contingency table represents the three categorical variables, as shown in Table 1.

The elements of the data hypermatrix $\boldsymbol{\mathcal{N}}=\left[n_{ijk}\right]$ are the frequencies for each combination of i-th row, j-th column, and k-th tube (cell frequencies), where $n_{ijk}\in\mathbb{N}$ for $i=1,\cdots,I,j=1,\cdots,J$, and $k=1,\cdots,K$. The univariate marginal frequencies of the i-th row, j-th column, and k-th tube are reffered to as

TABLE 1. The $I \times J \times K$ three-way contingency table

XYZ	Z_1				Z_2				$\cdots$	Z_K			
	Y_1	Y_2	$\cdots$	Y_J	Y_1	Y_2	$\cdots$	Y_J	$\cdots$	Y_1	Y_2	$\cdots$	Y_J
X_1	n_{111}	n_{121}	$\cdots$	n_{1J1}	n_{112}	n_{122}	$\cdots$	n_{1J2}	$\cdots$	n_{11K}	n_{12K}	$\cdots$	n_{1JK}
X_2	n_{211}	n_{221}	$\cdots$	n_{2J1}	n_{212}	n_{222}	$\cdots$	n_{2J2}	$\cdots$	n_{21K}	n_{22K}	$\cdots$	n_{2JK}
$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$	$\cdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$
X_I	n_{I11}	n_{I21}	$\cdots$	n_{IJ1}	n_{I12}	n_{I22}	$\cdots$	n_{IJ2}	$\cdots$	n_{I1K}	n_{I2K}	$\cdots$	n_{IJK}

slices, respectively, $n_{i..} = \sum_{j=1}^{J}\sum_{k=1}^{K} n_{ijk}, n_{.j.} = \sum_{i=1}^{I}\sum_{k=1}^{K} n_{ijk}$ and $n_{..k} = \sum_{i=1}^{I}\sum_{j=1}^{J} n_{ijk}$. Similarly, the bivariate marginal frequencies of the i-th row, j-th column, and k-th tube are reffered to as fibers determined by $n_{i.k} = \sum_{j=1}^{J} n_{ijk}, n_{.jk} = \sum_{i=1}^{I} n_{ijk}$, and $n_{ij.} = \sum_{k=1}^{K} n_{ijk}$. The frequency of all observations is the grand total of $\boldsymbol{\mathcal{N}}$ be $N = \sum_{i=1}^{I}\sum_{j=1}^{j}\sum_{k=1}^{K} n_{ijk}$.

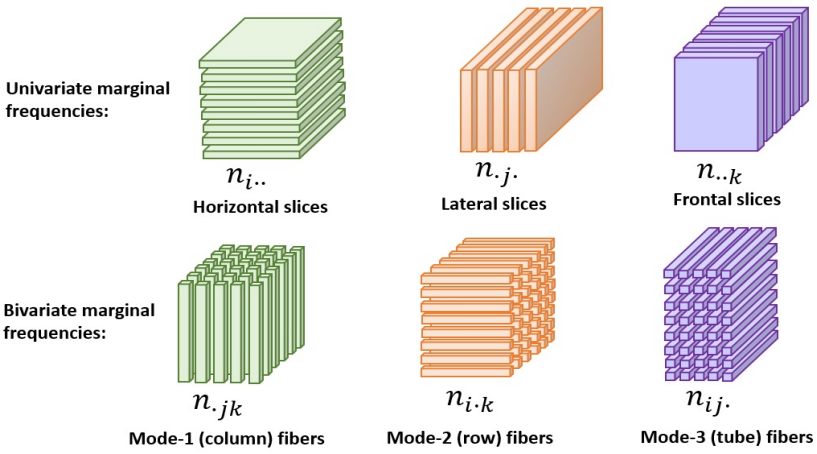

FIGURE 2. Illustration of univariate and bivariate marginal frequencies of data hypermatrix $\boldsymbol{\mathcal{N}}$ from the three-way contingency table

A three-way contingency table can be constructed from two categorical variables observed in different conditions, times or spaces [28]. The characteristics of such data are in the form of arrays or data cubes, also called boxes, by Kroonenberg [25].

The cell frequencies n_{ijk} in the contingency table can be converted to the relative frequencies p_{ijk} by dividing with the grand total N, that is, $p_{ijk} = \frac{nijk}{N}$. The relative frequency hypermatrix is called the correspondence hypermatrix and is denoted by $\boldsymbol{\mathcal{P}}$,

$$\boldsymbol{\mathcal{P}} = \left[p_{ijk}\right] = \left[\tfrac{n_{ijk}}{N}\right] \tag{4}$$

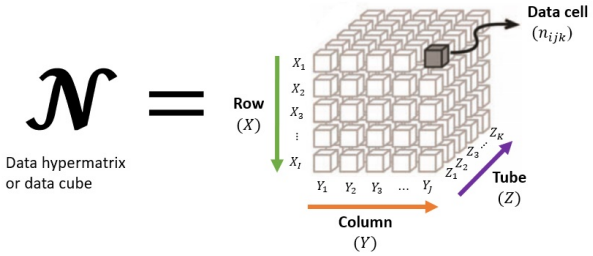

FIGURE 3. The data structure derived from the $I \times J \times K$ three-way contingency table as an input on CA3 represented by data hypermatrix $\boldsymbol{\mathcal{N}}$ (modified from [4])

The total dependence of the $I \times J \times K$ table with relative frequency p_{ijk} is measured by inertia $\mathbf{\Phi}^2$. An alogous to the CA2, $\mathbf{\Phi}^2$ is based on derivation from the three-way model of independence, such that

$$\begin{aligned} \mathbf{\Phi}^2 &= \sum_{ijk} \frac{\left(p_{ijk} - p_{i..}p_{.j.}p_{..k}\right)^2}{p_{i..}p_{.j.}p_{..k}} \\ &= \sum_{i.j.k} p_{i..}p_{.j.}p_{..k} \left[\frac{\left(p_{ijk} - p_{i..}p_{.j.}p_{..k}\right)}{p_{i..}p_{.j.}p_{..k}}\right]^2 \\ &= \sum_{i.j.k} p_{i..}p_{.j.}p_{..k} \left(\Pi_{ijk}\right)^2 \end{aligned} \tag{5}$$

The total dependency of cell Π_{ijk} can be divided into the partial contribution of the two-way interaction and the three-way interaction, such that

$$\Pi_{ijk} = \frac{p_{ij.} - p_{i..}p_{.j.}}{p_{i..}p_{.j.}} + \frac{p_{i.k} - p_{i..}p_{..k}}{p_{i..}p_{..k}} + \frac{p_{.jk} - p_{.j.}p_{..k}}{p_{.j.}p_{..k}} + \frac{p_{ijk} - \alpha p_{ijk}}{p_{i..}p_{.j.}p_{..k}} \tag{6}$$

Here, $\alpha p_{ijk} = p_{ij.}p_{..k} + p_{i.k}p_{.j.} + p_{.jk}p_{i..} - 2p_{i..}p_{.j.}p_{..k}$, reflects the three-way interaction measure for the (i, j, k)-cell. The element Π_{ijk} describes the dependency or association measure of the (i, j, k)-cell, which can be written as

$$\Pi_{ijk} = \frac{P\left(ij \mid k\right)}{P\left(ij\right)} \cdot \frac{P\left(ij\right)}{P\left(i\right)P\left(j\right)} - 1 \tag{7}$$

If the conditional probabilities for each k are equal, then $P(ij \mid k) = P(ij)$ and the first ratio is 1. Consequently, $\Pi_{ijk} = \Pi_{ij}$ and the three-way contingency table can be analyzed by CA2. The element of the bivariate marginal total is defined as the sum of the weights over the third index. Therefore, the elements of $I \times J$ bivariate marginal to be

$$\Pi_{ij.} = \sum_k p_{..k}\Pi_{ijk} = \sum_k p_{..k}\frac{p_{ijk} - p_{i..}p_{.j.}p_{..k}}{p_{i..}p_{.j.}p_{..k}} = \frac{p_{ij.} - p_{i..}p_{.j.}}{p_{i..}p_{.j.}} \tag{8}$$

The elements of other bivariate marginals, $\Pi_{i.k}$ and $\Pi_{.jk}$, are defined in the same way. The univariate marginal total is defined as the sum of the weights of the two indices, and the value is 0 (by definition Π_{ijk}). Thus, the row univariate marginal is defined by

$$\Pi_{i..} = \sum_j \sum_k p_{.j}p_{..k}\Pi_{ijk} = \sum_j p_{.j.}\frac{p_{ij.} - p_{i..}p_{.j.}}{p_{i..}p_{.j.}} = \sum_j \frac{p_{ij.}}{p_{i..}} - \sum_j \frac{p_{i..}p_{.j.}}{p_{.j.}} = 1 - 1 = 0$$

The elements of other univariate marginals, $\Pi_{.j.}$ and $\Pi_{..k}$, are defined similarly. Furthermore, the inertia of $\mathbf{\Phi}^2$, which measures the total dependencies of the three-way contingency table, can be partitioned as

$$\begin{aligned}
\mathbf{\Phi}^2 &= \sum_{ij} p_{i..}p_{.j}\left(\frac{p_{ij.} - p_{i..}p_{.j.}}{p_{i..}p_{.j.}}\right)^2 + \sum_{ik} p_{i..}p_{..k}\left(\frac{p_{i.k} - p_{i..}p_{..k}}{p_{i..}p_{..k}}\right)^2 \\
&\quad + \sum_{jk} p_{.j}p_{..k}\left(\frac{p_{.jk} - p_{.j}p_{..k}}{p_{.j}p_{..k}}\right)^2 + \sum_{ijk} p_{i..}p_{.j}p_{..k}\left(\frac{p_{ijk} - \alpha p_{ijk}}{p_{i..}p_{.j}p_{..k}}\right)^2 \\
&= \mathbf{\Phi}^2_{IJ} + \mathbf{\Phi}^2_{IK} + \mathbf{\Phi}^2_{JK} + \mathbf{\Phi}^2_{IJK}
\end{aligned} \tag{9}$$

This partition helps present the appropriate measure for each interaction such that its contribution to the total dependency can be known. Considering the symmetrical association structure between the three categorical variables from the three-way contingency table (under the assumption that X, Y, and Z are completely independent), the (i, j, k)-th elements of the table can be expressed in terms of deviations from the three-way independence model using the three-way Pearson's ratio as follows.

$$\boldsymbol{\mathcal{S}} = [s_{ijk}] = \left[\frac{p_{ijk} - p_{i..}p_{.j.}p_{..k}}{p_{i..}p_{.j.}p_{..k}}\right] \tag{10}$$

2.3. **Tucker3 tensor decomposition.**

Since the last six decades, Tucker3 tensor decomposition, also known as three-mode principal component analysis (3MPCA), has been considered an ingenious method to solve order-d tensor decomposition problems with d>2. The Tucker3 decomposition was first introduced by Tucker in 1963 and rewritten by Levin [29] and Tucker [15]. Tucker's 1966 article was more comprehensive than the early literature and widely cited. The article discusses some of the mathematical notes on three-mode factor analysis. Various analyzes in different fields have been carried out using the Tucker3 decomposition [30, 31, 32, 33, 34]. The Tucker3 decomposition has some terminology, as summarized in the table below.

TABLE 2. Some terminology of the Tucker3 decomposition; some specific for three-way and others for N-way

Terminology	**Proposed by**
Three-mode factor analysis (3MFA/Tucker3)	Tucker [15]
Three-mode principal component analysis (3MPCA)	Kroonenberg and De Leeuw [35]
n-Mode component analysis	Kapteyn et al. [36]
Higher-order singular value decomposition (HOSVD)	De Lathauwer et al. [16]
N-mode singular value decomposition (N-mode SVD)	Vasilescu and Terzopoulos [37]

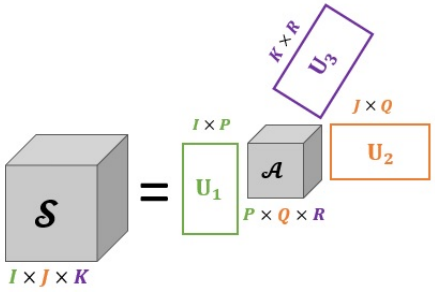

FIGURE 4. Tucker3 tensor decomposition model for order-3 hypermatrix

In the Tucker3 tensor decomposition, a hypermatrix $\boldsymbol{\mathcal{S}}$ is decomposed into three matrices: $\mathbf{U}_1, \mathbf{U}_2$, and $\mathbf{U}_3$, representing row, column and tube profiles, and one hypermatrix $\boldsymbol{\mathcal{A}}$ as core, representing the interactions of row, column and tube. Van Loan [38] elaborates on the HOSVD to represent the Tucker3 decomposition, as asserted in Definition 2.4.

Definition 2.4. (Van Loan, 2015):
Let $\boldsymbol{\mathcal{S}} \in \mathbb{R}^{I \times J \times K}$. *Suppose* $\mathbf{S}_{(1)} = \mathbf{U}_1 \mathbf{D}_1 \mathbf{V}_1^T$, $\mathbf{S}_{(2)} = \mathbf{U}_2 \mathbf{D}_2 \mathbf{V}_2^T$, *and* $\mathbf{S}_{(3)} = \mathbf{U}_3 \mathbf{D}_3 \mathbf{V}_3^T$, *respectively are the SVD of* $\mathbf{S}_{(1)}, \mathbf{S}_{(2)}, \mathbf{S}_{(3)}$, *where* $\mathbf{U}_1 \in \mathbb{R}^{I \times P}, \mathbf{U}_2 \in \mathbb{R}^{J \times Q}, \mathbf{U}_3 \in \mathbb{R}^{K \times R}$ *and* $\boldsymbol{\mathcal{A}} = \boldsymbol{\mathcal{S}} \times_1 \mathbf{U}_1^T \times_2 \mathbf{U}_2^T \times_3 \mathbf{U}_3^T \in \mathbb{R}^{P \times Q \times R}$. *Then*

i) $\boldsymbol{\mathcal{S}} = \boldsymbol{\mathcal{A}} \times_1 \mathbf{U}_1 \times_2 \mathbf{U}_2 \times_3 \mathbf{U}_3$ *is the Tucker3 tensor decomposition of* $\mathcal{S}$,
ii) $\mathbf{S}_{(1)} = \mathbf{U}_1 \mathbf{A}_{(1)} (\mathbf{U}_3 \otimes \mathbf{U}_2)^T$, $\mathbf{S}_{(2)} = \mathbf{U}_2 \mathbf{A}_{(2)} (\mathbf{U}_3 \otimes \mathbf{U}_1)^T$ *and* $\mathbf{S}_{(3)} = \mathbf{U}_3 \mathbf{A}_{(3)} (\mathbf{U}_2 \otimes \mathbf{U}_1)^T$ *are unfolding matrix of* $\boldsymbol{\mathcal{S}}$.

3. MAIN RESULTS

3.1. **Tucker3 tensor decomposition for the standardized residual hypermatrix.**

The standard residual hypermatrix $\boldsymbol{\mathcal{S}}$ is calculated using Pearson's three-way ratio in Equation (10). This equation indicates the difference between the joint

relative frequencies and the univariate marginal relative frequencies for each category. Such calculation allows for an inefficient numerical process (much rounding occurs) which can make the resulting CA3 plot inaccurate. In order to minimize the numerical process, the hypermatrix $\boldsymbol{\mathcal{S}}$ is calculated directly from the elements of the data hypermatrix $\boldsymbol{\mathcal{N}}$ as formulated in Lemma 3.1.

Lemma 3.1. *If* $\boldsymbol{\mathcal{N}} = \left[n_{ijk}\right]$ *is the* $I \times J \times K$ *hypermatrix derive from a three-ways contingency table. with* $n_{i..} = \sum_{j=1}^{J}\sum_{K=1}^{K} n_{ijk}, n_{.j.} = \sum_{i=1}^{I}\sum_{K=1}^{K} n_{ijk}$, $n_{..k} = \sum_{i=1}^{I}\sum_{j=1}^{J} n_{ijk}$, *and grand total* $N = \sum_{i=1}^{I}\sum_{j=1}^{J}\sum_{k=1}^{K} n_{ijk}$ *, then the element of* $\boldsymbol{\mathcal{S}} = \left[s_{ijk}\right]$ *are*

$$s_{ijk} = \frac{n^2 \cdot n_{ijk} - n_{i..}n_{.j.}n_{..k}}{n_{i..}n_{.j.}n_{..k}}$$

Proof. Consider $\boldsymbol{\mathcal{S}} = \left[s_{ijk}\right] = \left[\frac{p_{ijk} - p_{i..}p_{.j.}p_{..k}}{p_{i..}p_{.j.}p_{..k}}\right]$. Since $s_{ijk} = \frac{p_{ijk} - p_{i..}p_{.j.}p_{..k}}{p_{i..}p_{.j.}p_{..k}}$

$$\begin{aligned}
s_{ijk} &= \frac{\frac{n_{ijk}}{N} - \left(\sum_{j=1}^{J}\sum_{k=1}^{K} p_{ijk}\right)\left(\sum_{i=1}^{I}\sum_{k=1}^{K} p_{ijk}\right)\left(\sum_{i=1}^{I}\sum_{j=1}^{J} p_{ijk}\right)}{\left(\sum_{j=1}^{J}\sum_{k=1}^{K} p_{ijk}\right)\left(\sum_{i=1}^{I}\sum_{k=1}^{K} p_{ijk}\right)\left(\sum_{i=1}^{I}\sum_{j=1}^{J} p_{ijk}\right)} \\
&= \frac{\frac{n_{ijk}}{N} - \left(\sum_{j=1}^{J}\sum_{k=1}^{K} \frac{n_{ijk}}{N}\right)\left(\sum_{i=1}^{I}\sum_{k=1}^{K} \frac{n_{ijk}}{N}\right)\left(\sum_{i=1}^{I}\sum_{j=1}^{J} \frac{n_{ijk}}{N}\right)}{\left(\sum_{j=1}^{J}\sum_{k=1}^{K} \frac{n_{ijk}}{N}\right)\left(\sum_{i=1}^{I}\sum_{k=1}^{K} \frac{n_{ijk}}{N}\right)\left(\sum_{i=1}^{I}\sum_{j=1}^{J} \frac{n_{ijk}}{N}\right)} \\
&= \frac{\frac{n_{ijk}}{N} - \frac{1}{N^3}\left(\sum_{j=1}^{J}\sum_{k=1}^{K} n_{ijk}\right)\left(\sum_{i=1}^{I}\sum_{k=1}^{K} n_{ijk}\right)\left(\sum_{i=1}^{I}\sum_{j=1}^{J} n_{ijk}\right)}{\frac{1}{N^3}\left(\sum_{j=1}^{J}\sum_{k=1}^{K} n_{ijk}\right)\left(\sum_{i=1}^{I}\sum_{k=1}^{K} n_{ijk}\right)\left(\sum_{i=1}^{I}\sum_{j=1}^{J} n_{ijk}\right)} \\
&= \frac{\frac{n_{ijk}}{N} - \frac{1}{N^3}\left(n_{i..}n_{.j.}n_{..k}\right)}{\frac{1}{N^3}\left(n_{i..}n_{.j.}n_{..k}\right)} \\
&= \frac{N^2 n_{ijk} - n_{i..}n_{.j.}n_{..k}}{n_{i..}n_{.j.}n_{..k}}
\end{aligned}$$

□

The vectors aligned from the columns of the hypermatrix $\boldsymbol{\mathcal{S}}$ are not always orthogonal. Consequently, the association values of the row, column, and tube categories cannot always be plotted to Cartesian coordinates where the coordinate axes are mutually perpendicular [39, 40]. For this reason, new bases, which are linear combinations of the row, column, and tube components with the core elements, called principal coordinates, are constructed. In order to find these new bases, the standardized residual hypermatrix $\boldsymbol{\mathcal{S}}$ is decomposed using Tucker3.

3.2. **Some mathematical properties of Tucker3 tensor decomposition.**

There is a huge amount of literature available on Tucker3 tensor decomposition and its properties. In particular, one may refer to Van Loan [38] for an extensive and excellent discussion on Tucker3. This study shall only touch on some

of the more pertinent mathematical properties related to three-way correspondence analysis.

Theorem 3.2. *If* $\boldsymbol{\mathcal{S}} = \left[s_{ijk}\right] \in \mathbb{R}^{n\times n\times n}$ *is 3-hypercubical, such that* $\boldsymbol{\mathcal{S}} = \boldsymbol{\mathcal{A}}\times_1\mathbf{U}_1\times_2\mathbf{U}_2\times_3\mathbf{U}_3$ *is the Tucker3 tensor decomposition of* $\boldsymbol{\mathcal{S}}$*, then*

$$\mathbf{S}_{(1)} = \mathbf{U}_3\mathbf{U}_2\mathbf{U}_1\mathbf{A}_{(1)} \text{ and } \mathbf{A}_{(1)} = \left(\mathbf{U}_3\mathbf{U}_2\mathbf{U}_1\right)^T\mathbf{S}_{(1)}$$

Proof. Consider $\boldsymbol{\mathcal{S}} = \left[s_{ijk}\right] \in \mathbb{R}^{n\times n\times n}$ is 3-hypercubical. The matrices $\mathbf{U}_1, \mathbf{U}_2$ and $\mathbf{U}_3$ are obtained by determining the SVD of $\mathbf{S}_{(1)}, \mathbf{S}_{(2)}$, and $\mathbf{S}_{(3)}$, which are of size n×m with $m \geq n$ such that $\mathbf{U}_1, \mathbf{U}_2$ and $\mathbf{U}_3$ are the symmetric matrices of size n×n. Since $\boldsymbol{\mathcal{S}} = \boldsymbol{\mathcal{A}}\times_1\mathbf{U}_1\times_2\mathbf{U}_2\times_3\mathbf{U}_3$, based on Equation 1 and Equation 2, such that
$\iff \boldsymbol{\mathcal{S}} = \boldsymbol{\mathcal{A}}\times_1\left(\mathbf{U}_3\mathbf{U}_2\mathbf{U}_1\right)$
$\iff \mathbf{S}_{(1)} = \mathbf{U}_3\mathbf{U}_2\mathbf{U}_1\mathbf{A}_{(1)}$
Since $\mathbf{U}_1, \mathbf{U}_2$ and $\mathbf{U}_3$ are the orthogonal matrices, where $\mathbf{U}_1\mathbf{U}_1^T = \mathbf{I} = \mathbf{U}_2\mathbf{U}_2^T = \mathbf{I} = \mathbf{U}_3\mathbf{U}_3^T$ thus

$$\begin{aligned}
&\iff && \mathbf{S}_{(1)} = \mathbf{U}_3\mathbf{U}_2\mathbf{U}_1\mathbf{A}_{(1)}\\
&\iff && \mathbf{U}_3^T\mathbf{S}_{(1)} = \mathbf{U}_3^T\mathbf{U}_3\mathbf{U}_2\mathbf{U}_1\mathbf{A}_{(1)}\\
&\iff && \mathbf{U}_2^T\mathbf{U}_3^T\mathbf{S}_{(1)} = \mathbf{U}_2\mathbf{U}_1\mathbf{A}_{(1)}\\
&\iff && \mathbf{U}_1^T\mathbf{U}_2^T\mathbf{U}_3^T\mathbf{S}_{(1)} = \mathbf{U}_1^T\mathbf{U}_1\mathbf{A}_{(1)}\\
&\iff && \left(\mathbf{U}_3^T\mathbf{U}_2^T\mathbf{U}_1^T\right)\mathbf{S}_{(1)} = \mathbf{A}_{(1)}
\end{aligned}$$

□

Based on Theorem 3.2, in the same way, it can also be proven that
i)$\mathbf{S}_{(2)} = \mathbf{U}_1\mathbf{U}_3\mathbf{U}_2\mathbf{A}_{(2)}$ and $\mathbf{A}_{(2)} = \left(\mathbf{U}_1\mathbf{U}_3\mathbf{U}_2\right)^T\mathbf{S}_{(2)}$
ii)$\mathbf{S}_{(3)} = \mathbf{U}_2\mathbf{U}_1\mathbf{U}_3\mathbf{A}_{(3)}$ and $\mathbf{A}_{(3)} = \left(\mathbf{U}_2\mathbf{U}_1\mathbf{U}_3\right)^T\mathbf{S}_{(3)}$
Core hypermatrix $\boldsymbol{\mathcal{A}}$ is obtained by rearranging the columns of matrices $\mathbf{A}_{(1)}$, $\mathbf{A}_{(2)}$, and $\mathbf{A}_{(3)}$ into the elements in the corresponding core $\boldsymbol{\mathcal{A}}$. It implies that the core hypermatrix $\boldsymbol{\mathcal{A}}$ is not unique.

Theorem 3.3. *Let* $\boldsymbol{\mathcal{S}} \in \mathbb{R}^{I\times J\times K}$ *where* $I, J, K \in \mathbb{N}$*. Given* $\mathbf{S}_{(1)} = \mathbf{U}_1\mathbf{D}_1\mathbf{V}_1^T, \mathbf{S}_{(2)} = \mathbf{U}_2\mathbf{D}_2\mathbf{V}_2^T$ *and* $\mathbf{S}_{(3)} = \mathbf{U}_3\mathbf{D}_3\mathbf{V}_3^T$ *are the SVD of* $\mathbf{S}_{(1)}, \mathbf{S}_{(2)}, \mathbf{S}_{(3)}$*, respectively. Suppose* $\mathbf{U}_1 \in \mathbb{R}^{I\times P}, \mathbf{U}_2 \in \mathbb{R}^{J\times K}, \mathbf{U}_3 \in \mathbb{R}^{K\times R}, \boldsymbol{\mathcal{A}}_1 \in \mathbb{R}^{P\times Q\times R}$*. If* $\boldsymbol{\mathcal{S}} = \boldsymbol{\mathcal{A}}\times_1\mathbf{U}_1\times_2\mathbf{U}_2\times_3\mathbf{U}_3$ *is the Tucker3 tensor decomposition of* $\boldsymbol{\mathcal{S}}$*, then* $\boldsymbol{\mathcal{A}}$ *is not unique.*

Proof. Suppose $\boldsymbol{\mathcal{S}} = \boldsymbol{\mathcal{A}}\times_1\mathbf{U}_1\times_2\mathbf{U}_2\times_3\mathbf{U}_3$ is the Tucker3 tensor decomposition of $\boldsymbol{\mathcal{S}}$. Based on Definition 2.4, then $\boldsymbol{\mathcal{A}} = \boldsymbol{\mathcal{S}}\times_1\mathbf{U}_1^T\times_2\mathbf{U}_2^T\times_3\mathbf{U}_3^T$. Let $\mathbf{B} \in \mathbb{R}^{I\times I}$ is an orthogonal matrix such that,

$$\boldsymbol{\mathcal{A}} = \boldsymbol{\mathcal{S}}\times_1\mathbf{U}_1^T\times_2\mathbf{U}_2^T\times_3\mathbf{U}_3^T = \left(\boldsymbol{\mathcal{S}}\times_1\mathbf{B}\right)\times_1\mathbf{U}_1^T\mathbf{B}^T\times_2\mathbf{U}_2^T\times_3\mathbf{U}_3^T.$$

By applying Definition 2.4, obtained

$$\begin{aligned}\mathbf{S}_{(1)} &= \mathbf{U}_1\mathbf{A}_{(1)}\left(\mathbf{U}_3\otimes\mathbf{U}_2\right)^T \\ &= \mathbf{U}_1\mathbf{B}\mathbf{B}^T\mathbf{A}_{(1)}\left(\mathbf{U}_3\otimes\mathbf{U}_2\right)^T \iff \mathbf{A}_{(1)} = \mathbf{U}_1^T\mathbf{B}^T\mathbf{B}\mathbf{S}_{(1)}\left(\mathbf{U}_3\otimes\mathbf{U}_2\right)\end{aligned}$$

□

Finding a unique core hypermatrix $\boldsymbol{\mathcal{A}}$, such as a super diagonal hypermatrix, is almost impossible, even for a symmetric hypermatrix [18]. Another alternative that can be considered to increase this uniqueness is to make a lot of zero-value core elements. Such simplifying the core structure in some way might be useful in computation. By applying Theorem 3.2, it yields a core structure that has this property. Moreover, the Tucker3 tensor decomposition procedure for the standardized residual hypermatrix $\boldsymbol{\mathcal{S}}$ is generated in Algorithm 1.

Algorithm 1 Tucker3 tensor decomposition for the standardized residual hypermatrix

Step 1 : Start
Step 2 : Read the $I\times J\times K$ three-way contingency table
Step 3 : Create data hypermatrix, $\boldsymbol{\mathcal{N}}$
Step 4 : Compute the standardized residual hypermatrix $\boldsymbol{\mathcal{S}}$ using Lemma 3.1
Step 5 : Compute univariate marginal frequencies
Step 6 : Compute bivariate marginal frequencies
Step 7 : Apply matricization process
Step 8 : Find mode-1, mode-2, and mode-3 fibers, $\mathbf{S}_{(1)}, \mathbf{S}_{(2)}$ and $\mathbf{S}_{(3)}$
Step 9 : Apply SVD on $\mathbf{S}_{(1)}, \mathbf{S}_{(2)}$ and $\mathbf{S}_{(3)}$
Step 10: Apply Tucker3 tensor decomposition
Step 11: Print $\mathbf{U}_1$, $\mathbf{U}_2$, $\mathbf{U}_3$, and core hypermatrix $\boldsymbol{\mathcal{A}}$
Step 12: End

Figure 5 illustrates the Tucker3 tensor decomposition process for the standard residual hypermatrix formulated in Algorithm 1.

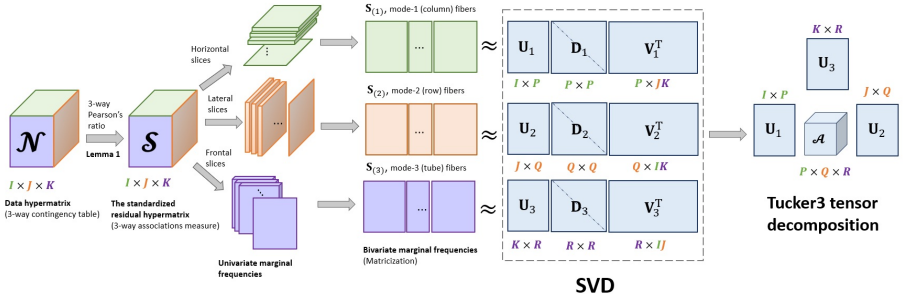

FIGURE 5. Tucker3 tensor decomposition for the standardized residual hypermatrix on CA3

3.3. **Constructing the CA3 plot via Tucker3 tensor decomposition.**

Considering Tucker3 decomposition, Beh & Lombardo [4] define the row principal coordinates (**F**) as

$$\mathbf{F} = \mathbf{U_1 A}_{(P\times QR)} \left(= f_{i,qr} = \sum_{p=1}^{P} u_{1ip} a_{pqr} \right) \tag{11}$$

Equation 11 confirmed that the row principal coordinates are a linear combination of the row components (categories) and the core elements. Instead, the (j,k) pair of the column-tube categories will be represented by a single-point coordinate [4]. Hence, the tube-column principal coordinates (**H**) are defined as

$$\mathbf{H} = (\mathbf{U}_2 \otimes \mathbf{U}_3)\, \mathbf{A}_{QR\times P} \left(= h_{jk,p} = \sum_{q=1}^{Q} \sum_{r=1}^{R} a_{pqr} u_{2jq} u_{3kr} \right) \tag{12}$$

Depicting the row (**F**) and column-tube (**H**) principal coordinates yields the correspondence plot (CA3 plot). A coordinate point on the plot represents each row and column-tube category in the contingency table. The construction of the principal coordinates in CA3 is formulated in Algorithm 2.

Algorithm 2 Constructing the CA3 plot via Tucker3 tensor decomposition

Step 1 : Start
Step 2 : Read the $I\times J\times K$ three-way contingency table
Step 3 : Create data hypermatrix, $\boldsymbol{\mathcal{N}}$
Step 4 : Compute the standardized residual hypermatrix $\boldsymbol{\mathcal{S}}$
Step 5 : Compute univariate marginal frequencies
Step 6 : Compute bivariate marginal frequencies
Step 7 : Apply matricization process
Step 8 : Find mode-1, mode-2, and mode-3 fibers, $\mathbf{S}_{(1)}, \mathbf{S}_{(2)}$ and $\mathbf{S}_{(3)}$
Step 9 : Apply SVD on $\mathbf{S}_{(1)}, \mathbf{S}_{(2)}$ and $\mathbf{S}_{(3)}$
Step 10: Apply Tucker3 tensor decomposition
Step 11: Print $\mathbf{U}_1$, $\mathbf{U}_2$, $\mathbf{U}_3$, and core hypermatrix $\boldsymbol{\mathcal{A}}$
Step 12: Compute the row principal coordinates, **F**
Step 13: Compute the column-tube principal coordinates, **H**
Step 14: Print the CA3 plot
Step 15: End

Another CA3 output is inertia which reflects the amount of information in each dimension on the CA3 plot. The last two equations calculate the total inertia (τ_{num}) from the contingency table and the category contribution.

$$\tau_{num} = \sum_{j=1}^{J} \sum_{k=1}^{K} \sum_{p=1}^{P} p_{.j.} p_{..k} h_{jk,p}^2 = \sum_{i=1}^{I} \sum_{q=1}^{Q} \sum_{r=1}^{R} f_{i,qr}^2 \tag{13}$$

The contribution percentage of the jk-th column-tube category to the p-dimensional ($C_{jk,p}$) determined by:

$$C_{jk,p} = 100 \times \frac{h^2_{jk,p}}{I_{tot}} \tag{14}$$

4. APPLICATION IN PRACTICAL DATA ANALYSIS

4.1. **Data Description.**

As an illustration of the application in data analysis, this study take into consideration the olive data that was considered by Agresti [41] and rediscussed by Beh & Lombardo [4], as summarized in the three-way contingency table of Table 3. The row categorical variable is given by the preferences for black olives of Armed Forces personnel with six levels in increasing order: A, B, C, D, E, and F. The column categorical variable represents position with three categories: South-West (SW), North-West (NW), and North-East (NE). The tube categorical variable deputized the location consisting of Urban and Rural categories. Therefore, the data provide a three-way cross-classification of a number of black olives of Armed Forces personnel based on the its preference, geographical position and location.

TABLE 3. The three-way contingency table for olive data [41, 4]

Preferensi	Urban			Rural		
	SW	NW	NE	SW	NW	NE
A	12	20	18	11	30	23
B	9	15	17	9	22	18
C	23	12	18	26	21	20
D	21	17	18	19	17	18
E	19	16	6	17	8	10
F	30	28	25	24	12	15

Instead, seeking the asymmetric association structure as discussed by Beh & Lombardo [4], this study interested is in the symmetrical association between the preference for black olives of Armed Forces personnel with the geographical position and location. Indeed, the dimensional approach of interpretation is valid for both asymmetric and symmetric plots. However, the asymmetric plot work well when total inertia is high, but are problematic when total inertia is low because the profile points in principal coordinates occupy a small space around the origin [14]. Since asymmetric plot interpret the distance between column and row points, the column profiles must be presented in row space or vice-versa. Eaton & Tayler [7] also described that it was 'extremely dangerous' to interpret the proximity of principal coordinates from different variables (row-to-column). Rather than defining the column coordinates in row space or vice-versa, consider instead defining the position of the row and column categories so that the strength of the association

that exists between the variables is reflected [4]. In a symmetric plot the separate configurations of row profiles and column profiles are overlaid in a joint display, even though they emanate, strictly speaking, from different spaces [14]. Therefore, both row and column points are displayed in principal coordinates. For this reason, we chose to focus on symmetrical association rather than asymmetrical. The association statistical tests using partitioning Pearson's phi-squared as in Equation 9 are recorded in the following table.

TABLE 4. Partitioning of the Pearson's phi-squared statistics for olive data in Table 3

Preferences	$\mathbf{\Phi}^2$	$\mathbf{\Phi}^2_{IJ}$	$\mathbf{\Phi}^2_{IK}$	$\mathbf{\Phi}^2_{JK}$	$\mathbf{\Phi}^2_{IJK}$
Index	0.078	0.048	0.019	0.000	0.011
Explained inertia	100%	61.49	24.12	0.60	13.79
df	27	10	5	2	10
p-value	0.004	0.001	0.032	0.859	0.728

Table 4 records the three-way association term from the partition of Pearson phi-squared statistic is $\mathbf{\Phi}^2 = 0.078$ with 27 degrees of freedom and p-value =0.004. It implies that for a 95% confidence interval, there is strong evidence to suggest that the preference for black olives of Armed Forces personnel is strongly symmetrically associated with the geographical position and location.

When considering the two-way association, one can see that the association between preference and geographical position is statistically significant since the $\mathbf{\Phi}^2_{IJ}$ statistic of 0.048 has a p-value =0.001 is less than the significant level $\alpha = 0.05$. This association accounts for 61.49% of the association that exists in the olive data table. Similarly, the association between preference and geographical location is statistically significant since the p-value of $\mathbf{\Phi}^2_{IK}$ is less than $\alpha = 0.05$. Their association contributes to 24.12% of the association that exists among three categorical variables. Meanwhile, the association between the geographical position and location is not statistically significant since the statistic $\mathbf{\Phi}^2_{JK}$ has p-value =0.859. Additionally, despite the residual three-way interaction $\mathbf{\Phi}^2_{IJK}$ contributing 13.79% to the total association, it is not statistically significant and will be ignored along with the other non-significant marginal partitioning $\mathbf{\Phi}^2_{JK}$.

4.2. **Visualizing the association by the CA3 plot.**

This subsection aims to visually display the associations that exist among three categorical variables. By considering olive data in Table 3, the associations among preference, position, and location are represented by the standardized residual hypermatrix $\boldsymbol{\mathcal{S}}$ determined by Lemma 3.1. Furthermore, $\boldsymbol{\mathcal{S}}$ is decomposed using the Tucker3 tensor decomposition as formulated in Algorithm 1 and illustrated in Figure 5. This decomposition involves a matricization process that is undertaken by utilizing Theorem 3.2. Meanwhile, Theorem 3.3 asserted that despite the core hypermatrix $\boldsymbol{\mathcal{A}}$ obtained from Tucker3 is not unique, Theorem 3.2 guaranteed the resulting core consists of many zero-value elements since it involves the SVD of

TABLE 5. Total inertia τ_{num} and contributions of the column-tube categories to the τ_{num}

Category	Axis 1	Axis 2	Axis 3
SW-Urban	0.1217	0.0023	0.0013
NW-Urban	0.0051	0.2516	0.0941
NE-Urban	0.0077	0.0444	0.3591
SW-Rural	0.0899	0.1221	0.0001
NW-Rural	0.1824	0.0260	0.0214
NE-Rural	0.0505	0.0219	0.0072
τ_{num}	75.48	16.11	8.14
Cumulative	75.48	91.58	99.72

each fiber matrix. The last step of Algorithm 1 produces three matrices representing row, column, and tube profiles and one core hypermatrix representing the three-way interaction of the three. Finally, the CA3 plot is obtained by applying Algorithm 2, as depicted in Figure 6.

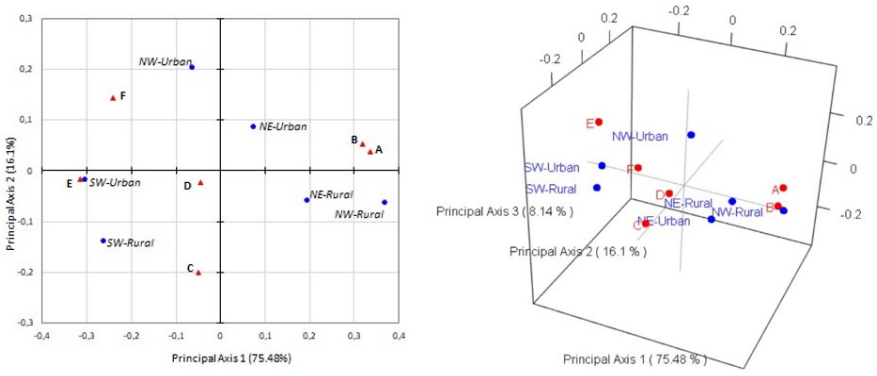

FIGURE 6. Tucker3 tensor decomposition for the standardized residual hypermatrix on CA3

Figure 6 shows the two-dimensional plot visually reflects 91.58% of the association that exists among preference, position, and location in olive data. In comparison, the three-dimensional plot visually reflects 99.72% of such associations. Therefore, using the two-dimensional plot has kept the visual display quality high, making it even simpler and easier to interpret. In the two-dimensional plot, the NW-Urban category is relatively far from the origin, meaning that this category contributes the most to the association; their contribution value in Table 5 is 0.2516. On the other hand, in a three-dimensional plot, NE-Urban is the category that contributes the most, with a contribution value of 0.3591. The contribution of each column-tube category to the association depicted along the first three axes is

summarized in Table 5. Moreover, a row category of preference E is strongly associated with a column-tube category of SW-Urban since their coordinates are close together. The complete association structure between three categorical variables is visualized in Figure 7.

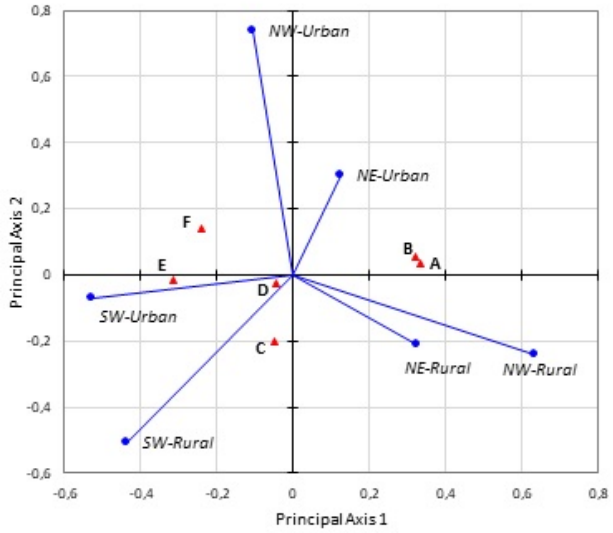

FIGURE 7. Column-tube interactive biplot from the CA3 for olive data in Table 3

The interactive biplot in Figure 7 shows that the projection of a row point coordinate of preference E on the arrow of SW-Urban is short. It indicates there is a strong association between preference E and SW-Urban. Similarly, the point projection of preference D on the SW-Rural arrow is shorter, which implies that their association is stronger. The NW-Urban is the column-tube category that contributes the most to the complete association structure since it has the most extended vector length. Finally, it can conclude that the symmetric association between the preference for black olives of Armed Forces personnel with the geographical position and location is statistically significant, which the preference for black olives depends, in a statistically significant manner, both on position and location, while the position does not statistically significant depending on the location.

5. CONCLUDING REMARKS

This study investigates the decomposition of the Tucker3 tensor in the CA3 context. The decomposition is performed on a standardized residual hypermatrix

$\boldsymbol{\mathcal{S}}$ that deputizes the association between the three categorical variables. As a first step, the study formulates how to get a more precise hypermatrix $\boldsymbol{\mathcal{S}}$ by calculating its elements directly from the data hypermatrix $\boldsymbol{\mathcal{N}}$ as in Lemma 3.1. Furthermore, Algorithm 1 is applied to decompose the hypermatri $\boldsymbol{\mathcal{S}}$ using Tucker3 tensor decomposition. This decomposition requires a process matricization, which is the rearrangement of the $\boldsymbol{\mathcal{S}}$ elements into columns in the matrix $\mathbf{S}_{(d)}$, referred to as mode-d fibers. This process uses Theorem 3.2, which yields the three matrices, each representing row, column, and tube profiles, and a core hypermatrix reflecting the interactions between them. Considering the importance of the core hypermatrix $\boldsymbol{\mathcal{A}}$ in Tucker3, the study explores some of the properties of this core as formulated in Theorem 3.3. Finally, Algorithm 2 is applied to obtain the row and column-tubes principal coordinates plotted to the two- and three-dimensional CA3 plots. As an illustration of the application in data analysis, it conducts the entire procedure to the olive data in Table 3. The results show that for the 95% confidence interval, there is strong evidence to suggest that the preference for black olives of Armed Forces personnel is strongly symmetrically associated with geographical position and location. Future research will examine the convergence of the Tucker3 algorithm and the reconstruction of three-way tables by three component matrices and core hypermatrix, including their properties to find the core with more zero elements.

Acknowledgement. The authors thank the research assistant team Mr Doni Ramdan, Mrs Julia Permata, and Mr Dian Pebriana, for contributing. The authors thank the anonymous reviewers for providing constructive comments to improve the earlier version of the manuscript. LPPM Universitas Singaperbangsa Karawang supported this research through the HIPKA grant scheme.

REFERENCES

[1] P. B. Imrey, G. G. Koch, M. E. Stokes, J. N. Darroch, D. H. Freeman, and H. D. Tolley, "Categorical data analysis: some reflection on the log linear model and logistic regression, par 1: Historical and methodological overview," *International Statistical Review / Revue Internationale de Statistique*, vol. 49, no. 3, pp. 265–283, 1981. https://doi.org/10.2307/1402613.

[2] A. Agresti, *Analysis of Ordinal Categorical Data*. Jhon Wileys & Sons, 2010. https://doi.org/10.1002/9780470594001.

[3] S. E. Fienberg and A. Rinaldo, "Three centuries of categorical data analysis: Log-linear models and maximum likelihood estimation," *Journal of Statistical Planning and Inference*, vol. 137, no. 11, pp. 3430–3445, 2007. https://doi.org/10.1016/j.jspi.2007.03.022.

[4] E. Beh and R. Lombardo, "Correspondence analysis theory, practice, and new strategies," *Theory, paractice and new strategies*, 2014. https://doi.org/10.1002/9781118762875.

[5] K. E. Lestari, U. S. Pasaribu, S. W. Indratno, and H. Garminia, "Generating roots of cubic polynomials by cardano's approach on correspondence analysis," *Heliyon*, vol. 6, no. 6, pp. 1–8, 2020. https://doi.org/10.1016/j.heliyon.2020.e03998.

[6] K. E. Lestari, M. R. Utami, and M. R. Yudhanegara, "Exploratory analysis on adaptive reasoning of undergraduate student in statistical inference.," *International Journal of Instruction*, vol. 15, no. 4, pp. 535–554, 2022. https://doi.org/10.29333/iji.2022.15429a.

[7] M. L. Eaton and D. E. Tyler, "On wielandt's inequality and its application to the asymptotic distribution of the eigenvalues of a random symmetric matrix," *The Annals of Statistics*, pp. 260–271, 1991. https://doi.org/10.1214/aos/1176347980.

[8] C. Eckart and G. Young, "The approximation of one matrix by another of lower rank," *Psychometrika*, vol. 1, no. 3, pp. 211–218, 1936. https://doi.org/10.1214/aos/1176347980.

[9] P. L. Emerson, "Numerical construction of orthogonal polynomials from a general recurrence formula," *Biometrics*, vol. 24, no. 3, pp. 695–701, 1968. https://doi.org/10.2307/2528328.

[10] E. Beh, "Theory & methods: Partitioning pearson's chi-squared statistic for singly ordered two-way contingency tables," *Australian & New Zealand Journal of Statistics*, vol. 43, no. 3, pp. 327 – 333, 2002. https://doi.org/10.1111/1467-842X.00179.

[11] C. Burt, "The factorial analysis of qualitative data," *British Journal of Statistical Psychology*, vol. 3, no. 3, pp. 166–185, 1950. https://doi.org/10.1111/j.2044-8317.1950.tb00296.x.

[12] K. Lestari, U. Pasaribu, S. Indratno, and H. Garminia, "The comparative analysis of dependence for three-way contingency table using burt matrix and tucker3 in correspondence analysis," in *Journal of Physics: Conference Series*, vol. 1245, p. 012056, IOP Publishing, 2019a. https://doi.org/10.1088/1742-6596/1245/1/012056.

[13] M. S. Bartlett, "Contingency table interactions," *Supplement to the Journal of the Royal Statistical Society*, vol. 2, no. 2, pp. 248–252, 1935. https://doi.org/10.2307/2983639.

[14] M. Greenacre, *Correspondence analysis in practice*. chapman and hall/crc, 2017. https://doi.org/10.1201/9781315369983.

[15] L. R. Tucker, "Some mathematical notes on three-mode factor analysis," *Psychometrika*, vol. 31, no. 3, pp. 279–311, 1966. https://doi.org/10.1007/BF02289464.

[16] L. De Lathauwer, B. De Moor, and J. Vandewalle, "A multilinear singular value decomposition," *SIAM journal on Matrix Analysis and Applications*, vol. 21, no. 4, pp. 1253–1278, 2000. https://doi.org/10.1137/S0895479896305696.

[17] L. Hogben, "Handbook of linear algebra," 2013. https://doi.org/10.1201/b16113.

[18] T. G. Kolda and B. W. Bader, "Tensor decompositions and applications," *SIAM review*, vol. 51, no. 3, pp. 455–500, 2009. https://doi.org/10.1137/07070111X.

[19] K. Lestari, U. Pasaribu, and S. Indratno, "Graphical depiction of three-way association in contingency table using higher-order singular value decomposition tucker3," in *Journal of Physics: Conference Series*, vol. 1280, p. 022035, IOP Publishing, 2019. https://doi.org/10.1088/1742-6596/1280/2/022035.

[20] H. A. Kiers, "Towards a standardized notation and terminology in multiway analysis," *Journal of Chemometrics: A Journal of the Chemometrics Society*, vol. 14, no. 3, pp. 105–122, 2000. https://doi.org/10.1002/1099-128X.

[21] T. G. Kolda, "Multilinear operators for higher-order decompositions," tech. rep., Sandia National Laboratories (SNL), Albuquerque, NM, and Livermore, California, 2006. https://doi.org/10.2172/923081.

[22] A. Carlier and P. M. Kroonenberg, "Decompositions and biplots in three-way correspondence analysis," *Psychometrika*, vol. 61, pp. 355–373, 1996. https://doi.org/10.1007/BF02294344.

[23] P. M. Kroonenberg and C. W. Snyder Jr, "Individual differences in assimilation resistance and affective responses in problem solving," *Multivariate Behavioral Research*, vol. 24, no. 3, pp. 257–284, 1989. https://doi.org/10.1207/s15327906mbr2403_1.

[24] P. Kroonenberg and J. Ten Berge, "The equivalence of tucker3 and parafac models with two components," *Chemometrics and Intelligent Laboratory Systems*, vol. 106, no. 1, pp. 21–26, 2011. https://doi.org/10.1016/j.chemolab.2010.05.022.

[25] P. M. Kroonenberg, *Applied multiway data analysis*. John Wiley & Sons, 2008. https://www.wiley.com/en-us/Applied+Multiway+Data+Analysis-p-9780470164976.

[26] R. Lombardo and E. J. Beh, "Three-way correspondence analysis for ordinal-nominal variables," in *SIS 2017 Statistics and Data Science: New Challenges, New Generations*, pp. 28–30, 2017. http://meetings3.sis-statistica.org/index.php/sis2017/sis2017/paper/view/499/0.

[27] L. D'Ambra, B. Simonetti, and E. J. Beh, "A dimensional reduction method for ordinal three-way contingency table," in *Compstat 2006-Proceedings in Computational Statistics: 17th Symposium Held in Rome, Italy, 2006*, pp. 271–283, Springer, 2006. https://doi.org/10.1007/978-3-7908-1709-6_21.

[28] K. E. Lestari, M. R. Utami, and M. R. Yudhanegara, "Sequential exploratory design by performing correspondence analysis to investigate procedural fluency of undergraduate student," in *AIP Conference Proceedings*, vol. 2588, AIP Publishing, 2023. https://doi.org/10.1063/5.0111974.

[29] J. Levin, "Three-mode factor analysis.," *Psychological Bulletin*, vol. 64, no. 6, pp. 442–452, 1965. https://doi.org/10.1037/h0022603.

[30] A. K. Smilde, R. Bro, and P. Geladi, *Multi-way analysis: applications in the chemical sciences*. John Wiley & Sons, 2004. https://doi.org/10.1002/0470012110.

[31] R. Ellis, P. Kroonenberg, B. Harch, and K. Basford, "Non-linear principal components analysis: an alternative method for finding patterns in environmental data," *Environmetrics: The official journal of the International Environmetrics Society*, vol. 17, no. 1, pp. 1–11, 2006. https://doi.org/10.1002/env.738.

[32] P. M. Kroonenberg, R. A. Harshman, and T. Murakami, "Analysing three-way profile data using the parafac and tucker3 models illustrated with views on parenting," *Applied Multivariate Research*, vol. 13, no. 1, pp. 5–41, 2009. https://doi.org/10.22329/amr.v13i1.2833.

[33] M. Gallo and A. Buccianti, "Weighted principal component analysis for compositional data: application example for the water chemistry of the arno river (tuscany, central italy)," *Environmetrics*, vol. 24, no. 4, pp. 269–277, 2013. https://doi.org/10.1002/env.2214.

[34] L. Li, R. Bai, J. Lu, S. Zhang, and C.-C. Chang, "A watermarking scheme for color image using quaternion discrete fourier transform and tensor decomposition," *Applied Sciences*, vol. 11, no. 11, p. 5006, 2021. https://doi.org/10.3390/app11115006.

[35] P. M. Kroonenberg and J. De Leeuw, "Principal component analysis of three-mode data by means of alternating least squares algorithms," *Psychometrika*, vol. 45, pp. 69–97, 1980. https://doi.org/10.1007/BF02293599.

[36] A. Kapteyn, H. Neudecker, and T. Wansbeek, "An approach to n-mode components analysis," *Psychometrika*, vol. 51, pp. 269–275, 1986. https://doi.org/10.1007/BF02293984.

[37] M. A. O. Vasilescu and D. Terzopoulos, "Multilinear analysis of image ensembles: Tensorfaces," in *Computer Vision—ECCV 2002: 7th European Conference on Computer Vision Copenhagen, Denmark, May 28–31, 2002 Proceedings, Part I 7*, pp. 447–460, Springer, 2002. https://doi.org/10.1007/3-540-47969-4_30.

[38] C. F. Van Loan, "Structured matrix computations from structured tensors," 2015. Note CIME-EMS Summer School, Cornell University, Italia, June 22-26, 2015.

[39] I. Ginanjar, U. Pasaribu, and A. Barra, "Simplification of correspondence analysis for more precise calculation which one qualitative variables is two categorical data," *ARPN Journal of Engineering and Applied Sciences*, vol. 11, no. 3, pp. 1983–1991, 2016. http://www.arpnjournals.org/jeas/research_papers/rp_2016/jeas_0216_3592.pdf.

[40] K. Lestari, U. Pasaribu, S. Indratno, and H. Garminia, "The reliability of crash car protection level based on the circle confidence region on the correspondence plot," in *IOP Conference Series: Materials Science and Engineering*, vol. 598, p. 012061, IOP Publishing, 2019c. https://doi.org/10.1088/1757-899X/598/1/012061.

[41] A. Agresti, *Categorical Data Analysis*. John Wiley & Sons, 2nd ed., 1990. https://doi.org/10.1002/0471249688.