Süleyman Demirel Üniversitesi
Fen Bilimleri Enstitüsü Dergisi
Cilt 24, Sayı 2, 474-486, 2020

Süleyman Demirel University
Journal of Natural and Applied Sciences
Volume 24, Issue 2, 474-486, 2020

DOI: 10.19113/sdufenbed.699241

Evaluation of Independent Components Analysis from Statistical Perspective and Its Comparison with Principal Components Analysis

Nurbanu BURSA[*1], **Huseyin TATLIDIL**[2]

[1,2]Hacettepe University, Faculty of Science, Department of Statistics, 06800, Ankara, Turkey

(Alınış / Received: 05.03.2020, Kabul / Accepted: 11.05.2020, Online Yayınlanma / Published Online: 20.08.2020)

Keywords
Independent components analysis,
Multivariate statistical methods,
Blind source separation problem,
Principal components analysis

Abstract: One of the most important problems in statistics and related fields is that finding an appropriate representation of multivariate data. Here is meant by representation; to transform the data into a more visible (accessible) form. Independent Components Analysis (ICA) is a statistical method used to find the underlying components of multivariate data and makes its main structure more visible. In this respect, ICA can also be seen as an extension of the Principal Components Analysis (PCA). However, ICA, contrary to PCA, is based on statistical independence rather than unrelatedness and statistical independence is a much stronger feature than unrelatedness. In addition, while the normal distribution of the components obtained in PCA is desired, the independent components of ICA are requested not to distribute normally. In the study, although it is a multivariate statistical method, the subject of ICA, which is not well known in the field of statistics and which is mostly used in engineering, was discussed in detail and contributed to the limited statistical literature on the subject. In the application part, ICA was compared with a similar method, PCA. Both analyzes were applied to an artificial dataset and it was concluded that ICA was much more successful than PCA in detecting non-normal components.

Bağımsız Bileşenler Analizinin İstatistiksel Bakış Açısıyla Değerlendirilmesi ve Temel Bileşenler Analizi ile Karşılaştırılması

Anahtar Kelimeler
Bağımsız bileşenler analizi,
Çok değişkenli istatistiksel yöntemler,
Kör kaynak ayrıştırma problemi,
Temel bileşenler analizi

Özet: İstatistik ve ilgili alanlardaki en önemli problemlerden biri, çok değişkenli verinin uygun bir temsilinin bulunmasıdır. Burada temsilden kasıt; veriyi, esas yapısı daha görünür (ulaşılır) bir şekle dönüştürmektir. Bağımsız Bileşenler Analizi (BBA); çok değişkenli verinin altında yatan bileşenleri bularak esas yapısını daha görünür hale getiren istatistiksel bir yöntemdir. Bu açıdan BBA, Temel Bileşenler Analizi'nin (TBA) bir uzantısı olarak da görülebilir. Ancak BBA, TBA'nın aksine ilişkisizlik yerine istatistiksel bağımsızlığı temel alır ve istatistiksel bağımsızlık, ilişkisizliğe göre çok daha güçlü bir özelliktir. Ayrıca TBA'da elde edilen bileşenlerin normal dağılması istenirken, BBA'da tam tersi bağımsız bileşenlerin normal dağılmaması istenmektedir. Çalışmada, çok değişkenli istatistiksel bir yöntem olmasına rağmen istatistik alanında pek fazla bilinmeyen ve daha çok mühendislik alanında kullanılan BBA konusu ayrıntılı bir şekilde ele alınmış ve konuyla ilgili kısıtlı istatistik literatürüne katkıda bulunulmuştur. Uygulama bölümünde BBA, benzer bir yöntem olan TBA ile karşılaştırılmıştır. Her iki analiz yapay bir veri kümesine uygulanmış ve BBA'nın normal olmayan bileşenleri ortaya çıkarmada TBA'dan çok daha başarılı olduğu sonucuna ulaşılmıştır.

1. Introduction

Independent Components Analysis (ICA) is a statistical method that transforms the underlying structure of multivariate data into a more visible form. Thanks to ICA, multivariate data is expressed as linear or nonlinear combinations of statistically independent components [1].

Especially in recent years, ICA has become a standard statistical data analysis method that has been applied to many problems in signal processing and machine

*Corresponding author: nurbanubursa@hacettepe.edu.tr

learning, as it effectively finds solutions for many real-life problems [2]. Although it is a multivariate statistical method, ICA is widely used in fields such as signal processing and machine learning, which are the sub-branches of engineering, but it is still not much known in the field of statistics. For this reason, ICA is not widely used in studies as much as a similar analysis, Principal Components Analysis (PCA). Particularly, when the Turkish statistical literature in the field of statistics related to the subject is examined, it is noteworthy that there is no study except the study of Ozdamar [3] in which he analyzed EEG signals with ICA and that Bursa [4] proposed a new approach using ICA to solve the multicollinearity problem. It is considered that this situation occurs, due to the fact that,

- In the current studies, mostly in the field of engineering, the statistical structure behind ICA is not brought to the fore in full sense,
- Its differences and strengths are not explained well enough with respect to similar statistical methods like PCA and therefore it is perceived as an analysis used only in the field of engineering.

For this purpose, the study is based on ICA in order to overcome the shortcomings listed above and contribute to the limited statistical literature on ICA. In this study, it is aimed to provide the researchers with detailed information that may be required about the usage of ICA in other fields such as signal processing and machine learning, in addition to increasing the awareness of ICA, especially in statistics.

In the second section of the study, comprehensive information and examples about the development process and application areas of ICA are given. Then, in order to better understand the reason for the emergence of ICA, a problem that ICA offers solutions in the field of statistics is examined. In the third section, ICA model is known as the basic ICA model and the characteristics of this model are discussed. In the fourth section, algorithms and software used for the solution of ICA model are given. In the fifth section, PCA, which is one of the multivariate statistical methods and similar to ICA in terms of operation, is compared with ICA and an exemplary application is carried out on an artificial dataset, in which ICA performed better. Finally, in the sixth and last section, general evaluations are carried out and the contributions of the study are mentioned.

2. Development Process and Application Areas of Independent Components Analysis

This method, which was not known as ICA in the first years of its use, was used for the first time to encode the movements of the contracted muscles by Hérault, Jutten and Ans [5]. In the aforementioned study, it has been shown that the nervous system determines the angle and speed of the stimulus coming from the muscle by realizing an unsupervised learning with the contraction of the muscles. However, the publication of the results in French caused the subject to not become widespread in the international literature at that time and its effect was limited to French researchers only. This method was used as the Hérault-Jutten model in studies published in the following years by French researchers working on signal processing and artificial neural networks. Thanks to the studies published by Jutten and Hérault [6-8] in English in the following years, the subject has gradually become known worldwide.

The name ICA was firstly used in a study published by Comon [9]. Also, in this study, a comprehensive mathematical formulation of ICA is presented for the first time. Since the mid-1990s, the interest in ICA has increased; different ICA models and new algorithms to be used in solving these models have been developed and it has started to be shown that ICA can be used for different purposes other than blind source separation problem. As of the beginning of the 2000s, books about ICA have been published and special congresses have been organized on this subject.

Nowadays, ICA is widely used in,

- Medical imaging (Determining the sources of the signals emitted by the brain and separating the ultrasonography signals [10-12]),
- Geology (Investigating seismic waves and geological mapping [13-15]),
- Fault detection (Detecting faults encountered in quality control processes [16-18]),
- Image processing (Revealing the attributes of the images and deblurring the images [19-21]),
- Telecommunications (Separating radio waves or audio signals [22-24]),
- Econometrics (Determining the components that play a role in the formation of financial series [25-27]),
- Data and text mining (Reducing size [28-30]),
- Bioinformatics and genetics (Ascertaining essential components in gene expression [31-33]),
- Chemistry (Analyzing the components in the NIR spectroscopy used in the production and break down of foods [34-36]).

2.1. Statistical Problem Playing a Role in Independent Components Analysis: Linear Representation of Multivariate Data

The most important feature of ICA that differs from other methods is that it searches both statistically independent and non-normal components in multivariate data. Although the main starting point of ICA seems to be to find a solution to the blind source separation problem, it is essential to have a better representation of multivariate data [1, 4].

How to find a suitable representation of multivariate data is one of the most important problems in statistics and other related fields. Here, what is meant by the expression of representation is the transformation of the main structure of the data into a more visible (accessible) form. Good representation is the main goal of many methods such as data mining, descriptive data analysis and signal processing [1, 4].

To explain the problem of linear representation of multivariate data on an example, consider a set of data that is observed together and consists of several variables. The number of variables is denoted by p, the number of observations; T, data; $x_i(t), i = 1, \dots, p$ and $t = 1, \dots, T$. The answer of the question "What could be the linear function that provides the conversion of the dataset, provided that $k \leq p$ from p dimensional space to k dimensional space and where the transformed variables (underlying factors or components that show the main structure in the data) give the hidden information in the data?" will give a good representation of multivariate data. Equation 1 can be used to express this question mathematically:

$$y_i(t) = \sum_j w_{ij} x_i(t),\ i = 1, \dots, k; j = 1, \dots, p \qquad (1)$$

In Equation 1, w_{ij}s are weight coefficients that state the representation of observed variables. It is possible to express the same problem with the matrix-vector representation as in Equation 2:

$$\begin{bmatrix} y_1(t) \\ y_2(t) \\ \vdots \\ y_p(t) \end{bmatrix} = \mathbf{W} \begin{bmatrix} x_1(t) \\ x_2(t) \\ \vdots \\ x_p(t) \end{bmatrix} \qquad (2)$$

Each $x_i(t),\ i = 1, \dots, T$ in Equation 2 is a sample of a random variable and **W** matrix, can be determined using the statistical properties of y_i transformed components [1].

One of the methods used to determine the **W** matrix is to limit the number of the components of y_i and to determine y_i so that y_i contains the information contained in the data as much as possible. This method also pioneers PCA and Factor Analysis (FA) [1].

Another method used to determine the **W** matrix is independence. According to this method, the **W** matrix is determined provided that the components of y_i are statistically independent. In other words, the value of any component does not provide information about the values of other components. In fact, the factors and components obtained in PCA and FA are also independent; however, this is partially true. Because these methods assume that the data is normally distributed. Independent components are easy to obtain if the data is normally distributed, as unrelated components are always independent in the normally distributed data. In reality, the data is not normally distributed usually. This constitutes the starting point of ICA [1, 4].

3. Independent Components Analysis

In terms of signal processing and machine learning, the operation of ICA is best explained by the cocktail party problem, which is an example of blind source separation. In the cocktail party problem, to distinguish the voices of more than one person in the same room is aimed. For this purpose, suppose that two people are in the same room and two microphones record the voice of these two people. The aim of the problem is to reveal the speech of each person independently from the mixed sounds recorded by the microphones. Thanks to ICA, each person's speech, or in other words, two original sounds (source signal) can be obtained from the mixture of two signals (from the microphone). Since separating these mixtures is only possible when the source signals are independent of each other, the basic assumption of ICA is that the source signals are independent of each other. Mixtures can consist of sounds, as in the cocktail party problem, or radio waves, brain signals, or images [37].

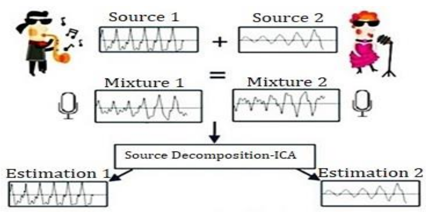

Figure 1. Cocktail party problem [4]

From a statistical point of view, ICA can be considered as a more advanced version of PCA and FA, which define the variability between the related variables in terms of fewer unobserved variables, which are called principal components and factors, respectively. While PCA and FA only use analysis of second-order statistics such as covariance or correlation matrix, ICA uses higher-order statistics of random variables such as kurtosis and skewness. Similar to FA, a hidden variable model is created in ICA too and it is assumed that the variables observed in this model are mixtures of variables that are mutually independent and not normally distributed [38].

3.1. Independent Components Analysis Model

The basic ICA model is expressed as in Equation 3:

$$x_i = a_{i1}s_1 + a_{i2}s_2 + \dots + a_{in}s_n,\ \ i = 1, \dots, n \quad (3)$$

Here, $x_1, x_2, \dots, x_n$ are random variables with the number of observations n; provided that $i, j = 1, \dots, n$ are real coefficients, a_{ij}'s are modeled as a linear combination of the random variables $s_1, s_2, \dots, s_n$ in number n. In Equation 3, since s_j independent components are hidden variables, they cannot be observed directly and by definition, they are statistically independent of each other. Also, a_{ij} mixture coefficients (or weights) are also unknown. The purpose of ICA is to estimate mixture coefficients and independent components using random variables x_i [1, 4].

It is also possible to express ICA model by vector-matrix notation. In Equation 5, it is seen that $\mathbf{x}$ random vector consists of $x_1, x_2, \dots, x_n$ mixtures, while $\mathbf{s}$ random vector consists of $s_1, s_2, \dots, s_n$. Beside, $\mathbf{A}$ matrix consists of a_{ij}. In Equation 4, $\mathbf{a}_i$'s correspond to the columns of the $\mathbf{A}$ matrix.

$$\mathbf{x} = \sum_{i=1}^{n} \mathbf{a}_i s_i \tag{4}$$

$$\mathbf{x} = \mathbf{A}\mathbf{s} \tag{5}$$

For the case where $\mathbf{x}^T$ random vectors in number p exist, ICA model is expressed as in Equation 6:

$$\mathbf{X}_{p\times n} = \mathbf{A}_{p\times p}\mathbf{S}_{p\times n} \tag{6}$$

As can be seen from the equations, there is no error term in the basic ICA model. For this reason, the basic ICA model is also called the noise-free ICA model by some researchers. The basic ICA model that has no error term is sufficient for most applications [1, 4].

In ICA model in Equation 6, $\mathbf{A}$ and $\mathbf{S}$ matrices are estimated only by using the $\mathbf{X}$ matrix. For this, primarily $\mathbf{W}$ is determined, which is the estimation of the inverse of the mixture matrix $(\mathbf{A}^{-1})$ and is called the decomposition matrix. Afterward, by using $\mathbf{W}$ matrix, $\mathbf{U}$ matrix is obtained as Equation 7. Provided that $k \le p$, $\mathbf{U}$; is the estimation of $k \times n$ dimensional $\mathbf{S}$ matrix, in other words, is the estimation of the independent components matrix.

$$\mathbf{U} = \mathbf{W}\mathbf{X} \tag{7}$$

While obtaining the mixture matrix and independent components, many different estimation methods such as maximization of non-normality, maximum likelihood method, minimization of mutual information, tensorial methods, and nonlinear PCA are used.

3.2. Assumptions, Constraints and Uncertainties of the Independent Components Analysis Model

The assumptions and constraints required to estimate ICA model and some uncertainties due to the structure of ICA model are listed below.

- *Independent components are assumed to be statistically independent of each other.*

In other words, the joint probability density function of the independent components should be written as $f(\mathbf{s}) = f(s_1, s_2, \dots, s_p) = \prod_j f(s_j)$, which is the product of the marginal probability density functions. This is the main assumption for the model to be estimable. For some applications, independence can also be addressed physically. If components are created by physically separate and non-interacting objects, they can be considered statistically independent [39].

- *The distribution of the independent components is not normal.*

It is the most important assumption that distinguishes ICA from other methods such as PCA and FA. Higher-order information is needed to estimate ICA model. However, as is known, the higher-order cumulant (such as kurtosis and skewness) of the normal distribution equals zero, and therefore it is not possible to apply ICA for normally distributed variables.

- *The mixture matrix is assumed to be an invertible square matrix.*

Although this assumption is not essential, it is generally preferred for ease of calculation [1].

- *Variances of independent components cannot be calculated.*

The reason is that both $\mathbf{S}$ and $\mathbf{A}$ is not known in the model. Since the mixture matrix $\mathbf{A}$ is unknown, when the independent component is multiplied by any α_i coefficient as in Equation 8, this coefficient will be neutralized by $\mathbf{a}_i$ column of $\mathbf{A}$ matrix. Consequently, the signs of the independent components are also indefinite.

$$\mathbf{x} = \sum \left(\frac{1}{\alpha_i}\mathbf{a}_i\right)(s_i\alpha_i) \tag{8}$$

- *The order of importance of independent components cannot be determined.*

That is because the model has two unknowns. Any component obtained can be considered the first component [40].

4. Approaches and Algorithms Used in Obtaining Independent Components

Most of ICA algorithms perform the estimation of the model in two stages: 'preprocessing' and 'estimation'. During the preprocessing stage, some transformations are applied to the data, thus, time is saved by providing ease of calculation. In the estimation stage, independent components are revealed from the data prepared by using one of the selected optimization methods [41].

4.1. Preparation of Data to be Used Before Independent Components Analysis

In algorithms to be used for ICA, it is desired that the observed (mixture) values and independent components have a zero mean for ease of operation. If this feature is not provided, it can be achieved through some preprocessing operations. For this, the values observed before applying ICA are subjected to 0 mean by subtracting the differences from their means with Equation 9.

$$\mathbf{x} = \mathbf{x} - \mathrm{E}\{\mathbf{x}\} \tag{9}$$

In this way, independent components are made to have a mean of 0. The mixture matrix after this preprocessing operation remains the same as it is seen in Equation 10.

$$\mathrm{E}\{\mathbf{s}\} = \mathrm{E}\{\mathbf{A^{-1}x}\} = \mathbf{A^{-1}}\mathrm{E}\{\mathbf{x}\} \tag{10}$$

If desired, after the mixture matrix and independent components are estimated from 0 mean data, it can be switched to the original independent components by adding the term $\mathbf{A^{-1}}\mathrm{E}\{\mathbf{x}\}$ to the 0 mean independent components [1, 3, 4].

The whitening process is performed to make the data unrelated and have unit variance. If a random vector ($\mathbf{x}$) with $\mathbf{0}$ mean is white, its components are unrelated and the covariance matrix $\mathbf{\Sigma}$ is equal to the unit matrix ($\mathbf{I}$).

$$\mathbf{\Sigma} = \mathrm{E}\{\mathbf{xx^T}\} = \mathbf{I} \tag{11}$$

It is possible to whiten non-white data using the transformation matrix, thereby freeing them from the effects of first and second-ordered statistics. For instance, an observed vector $\mathbf{x}$ can be transformed into a white vector $\mathbf{z}$ with a linear transformation [1].

$$\mathbf{z} = \mathbf{Vx} \tag{12}$$

There are many different transformations for whitening. The most used among these is the decomposition of the eigenvalues of the covariance matrix. In this method, $\mathbf{\Lambda}$ is the orthogonal matrix of eigenvalues $\mathrm{E}\{\mathbf{xx^T}\}$, $\mathbf{D} = diag(d_1, \dots, d_n)$ is a diagonal matrix consisting of eigenvalues.

$$\mathrm{E}\{\mathbf{xx^T}\} = \mathbf{\Lambda D \Lambda}^T \tag{13}$$

In line with Equation 13, the whitening matrix $\mathbf{V}$ can be created as in Equation 14.

$$\mathbf{V}=\mathbf{\Lambda D}^{-1/2}\mathbf{\Lambda}^T \tag{14}$$

Another method other than the decomposition of eigenvalues for whitening is to use ICA. After data in ICA model is whitened, it can be expressed as

$$\mathbf{z} = \mathbf{Vx} = \mathbf{VAs} = \widetilde{\mathbf{A}}\mathbf{s} \tag{15}$$

The new mixture matrix $\widetilde{\mathbf{A}} = \mathbf{VA}$ which is obtained as a result of whitening is orthogonal. The orthogonality feature is seen in Equation 16.

$$\mathrm{E}\{\mathbf{zz}^T\} = \widetilde{\mathbf{A}}\mathrm{E}\{\mathbf{ss}^T\}\widetilde{\mathbf{A}}^T = \widetilde{\mathbf{A}}\widetilde{\mathbf{A}}^T = \mathbf{I} \tag{16}$$

Thanks to this feature, only the orthogonal matrix space will be explored while searching for the mixture matrix. In other words, instead of n^2 estimating elements of the original $\mathbf{A}$ matrix, it will be sufficient to determine $n(n-1)/2$ elements of the orthogonal $\widetilde{\mathbf{A}}$ matrix. In this way, the complexity of the problem will also be reduced for ICA algorithms [1].

4.2. Estimating Independent Components

In ICA, the aim is to estimate the decomposition matrix ($\mathbf{W}$), which is the inverse of the mixture matrix, and then apply this matrix on whitened data to obtain independent components.

After the preprocessing stage of the data is completed, the $\mathbf{W}$ matrix can be estimated using three different independence approaches. The first of these approaches is based on the assumption of normality and determines the independent components so as to maximize their non-normality. In the second approach, the aim is to minimize mutual information, while in the last approach the maximum likelihood estimate is used. After determining which approach to use when obtaining independent components, it is decided with which algorithm to provide the optimization of the approach. In the study, the method of maximizing non-normality is discussed only. In this method, independent components are revealed by maximizing their non-normality. Kurtosis and negentropy values are used to measure how far the variables are from normality.

- Kurtosis

In this method, independent components are determined so that their kurtosis values are at maximum. For example, assume that from ICA model denoted by $\mathbf{X} = \mathbf{AS}$, two independent components $\mathbf{s_1}$ and $\mathbf{s_2}$ are revealed; and $\mathbf{s_1}$, $\mathbf{s_2}$ and $\mathbf{S}$ have unit variance. The kurtosis of the independent components can be written as in Equation 18, by using the addition property of the kurtosis coefficient:

$$\mathbf{U} = \mathbf{WX} = \mathbf{WAS} = \mathbf{QS} = \mathbf{q}_1\mathbf{s}_1 + \mathbf{q}_2\mathbf{s}_2 \tag{17}$$

$$\begin{aligned} kurt(\mathbf{S}) &= kurt(\mathbf{q}_1\mathbf{s}_1) + kurt(\mathbf{q}_2\mathbf{s}_2) \\ &= \mathbf{q}_1{}^4 kurt(\mathbf{s}_1) + \mathbf{q}_2{}^4 kurt(\mathbf{s}_2) \end{aligned} \quad (18)$$

Since independent components have unit variance, for $\mathbf{q}_1$ and $\mathbf{q}_2$, there is also $\mathrm{E}(\mathbf{U}^2) = \mathbf{q}_1{}^2 + \mathbf{q}_2{}^2$ constraint. This constraint means that $\mathbf{Q}$ should be a unit circle in two-dimensional space. So the goal of ICA turns into maximizing the equation $kurt(\mathbf{S}) = \mathbf{q}_1{}^4 kurt(\mathbf{s}_1) + \mathbf{q}_2{}^4 kurt(\mathbf{s}_2)$ on the unit circle. The optimal, that is, the maximum solution is obtained when one of $\mathbf{Q}$ is 0 and the other is -1 or +1, and these optimal solutions give independent components $\pm\mathbf{s}_i$.

If the data matrix is a white matrix ($\mathbf{Z}$), then this time independent components are obtained by maximizing the kurtosis of the independent components in the equation $\mathbf{U} = \mathbf{WZ}$. In this case, as $\mathbf{W}$ and $\mathbf{Z}$ have unit variance, the kurtosis of the independent components will be denoted by $kurt(\mathbf{U}) = \mathrm{E}\{(\mathbf{WZ})^4\} - 3$ [1, 41].

The most basic algorithm used to maximize kurtosis is the gradient algorithm. Kurtosis gradient value of $\mathbf{U}$ is $\frac{\partial |kurt(\mathbf{W}^T\mathbf{Z})|}{\partial \mathbf{W}} = c\mathrm{E}\{\mathbf{Z}(\mathbf{W}^T\mathbf{Z})^3\}$ and here c is a constant. At each iteration, the weight vector which was randomly determined at the beginning is renewed as $\mathbf{w}_{new} = \mathbf{w}_{old} + \eta \mathrm{E}\big(\mathbf{Z}(\mathbf{w}_{old}\mathbf{Z})\big)$. Here η is the learning coefficient. Since the kurtosis is optimized on the unit circle, the obtained weight vector is then divided by its norm to update $\mathbf{w}_{new}^n = \mathbf{w}_{new}/\|\mathbf{w}_{new}\|$. The algorithm continues to operate until convergence is achieved [41].

- Negentropy

As a measure of non-normality, because kurtosis is sensitive to outliers, negentropy is preferred instead of kurtosis, although it is often more complex and difficult to calculate. The most important reason behind the difficulty in calculating the negentropy is the necessity to have a nonparametric estimate of the probability density function. Therefore, different approaches have been introduced in which higher-order cumulants are used to calculate the approximate value of negentropy.

$$J(u) \approx \tfrac{1}{12}\mathrm{E}(u^3)^2 + \tfrac{1}{48} kurt(u)^2 \quad (19)$$

In Equation 19, u random variable is assumed to have 0 mean and unit variance. However, since there is a kurtosis in this equation, another approach based on maximum entropy has been proposed by Hyvärinen [42] as it may be affected by outliers:

$$J(u) \approx \sum_{i=1}^{p} k_i \left(\mathrm{E}\big(G_i(u)\big)\right) - \mathrm{E}\big(G_i(v)\big)^2 \quad (20)$$

In Equation 20, k_i represents positive constants, v represents a randomly distributed random variable with zero mean and unit variance, and G_i represents quadratic functions. As G_i function, functions in Equation 21 are selected provided that $1 \le a_1 \le 2$ (a_1 usually equals to 1).

$$G_1(u) = \tfrac{1}{a_1}\mathrm{logcosh}a_1 u,\; G_2(u) = -\exp(-u^2/2) \quad (21)$$

Equation 20 can be maximized by using the gradient algorithm so as to maximize negentropy as it was in kurtosis.

FastICA Algorithm

The fast fixed-point algorithm or FastICA algorithm developed by Hyvärinen [42] reveals the independent components in ICA model. FastICA can be used to optimize all of the estimation methods, where non-normality is maximized, the amount of mutual information is minimized, or the maximum likelihood estimator is used.

When obtaining independent components, other than FastICA algorithm, SOBI (Second-Order Blind Identification) which uses second-order statistics such as delayed correlation matrix, FOBI (Fourth-Order Blind Identification) which uses fourth-ordered statistics, Infomax which is based on maximization of entropy and JADE (Joint Approximation Diagonalization of Eigenmatrices) which is based on common diagonalization algorithms are also widely used [38]. Studies comparing the mentioned ICA algorithms in terms of efficiency and speed are available in the literature [43-46].

FastICA has been preferred in the application section of the study since it is one of the most preferred algorithms in applications. There are packages that can be used in both MATLAB and R environment for applications of other algorithms and various variations of FastICA. (For R, see. JADE package [48], fICA package [49], ICA package (It is also the R package used in obtaining independent components in the application section of the study) [50], fastICA package [51]).

FastICA is faster than gradient-based algorithms since it has a quadratic convergence rate. Also, it is very easy to apply since the algorithm does not contain any values that need to be determined beforehand, such as the learning coefficient. FastICA can estimate all independent components individually (deflation approach) or simultaneously (symmetric approach). While doing so, it usually uses the negentropy values [52]. Since FastICA, which is based on negentropy, is used in the application section of the study, only the algorithm steps of negentropy based FastICA are included in this section.

- *Steps of Negentropy Based FastICA Algorithm Regarding the Determination of Multiple Independent Components Simultaneously*

Provided that $1 \leq a_1 \leq 2$ (a_1 usually equals to 1), derivatives of G function depicted in Equation 21 is denoted by g as in Equation 22 and Equation 23 and one of these functions is selected.

$$g_1(u) = \tanh a_1 u,\ \ g_1'(u) = a_1(1 - \tanh a_1 u) \qquad (22)$$

$$g_2(u) = u\exp(-u^2/2),\ \ g_2'(u) = (1 - u^2)\exp(-u^2/2) \quad (23)$$

To estimate multiple independent components simultaneously, in each iteration $\mathbf{w}_1, \ldots, \mathbf{w}_n$ vectors need to be orthogonalized. The purpose of orthogonalizing vectors is to prevent different vectors from converging to the same maximum point.

The steps of the algorithm are as follows [1]:

1) By centralizing and then whitening the dataset, obtain $\mathbf{z}$.
2) Determine (m), the number of independent components to be estimated.
3) Randomly select a $\mathbf{W} = (\mathbf{w}_1, \ldots, \mathbf{w}_m)^T$ starting matrix consisting of vectors with unit norms.
4) Calculate $\mathbf{w}_i = E\big(\mathbf{z}g(\mathbf{w}_i\mathbf{z})\big) - E\big(g'(\mathbf{w}_i\mathbf{z})\big)\mathbf{w}_i$, $i = 1, \ldots, m$ by using Equation 23 or Equation 24.
5) Perform the orthogonalization $\mathbf{W}^{\circ} = (\mathbf{W}\mathbf{W}^T)^{-1/2}\mathbf{W}$.
6) Perform the normalization $\mathbf{W}^n = \mathbf{W}^{\circ}/\|\mathbf{W}^{\circ}\|$.
7) If convergence cannot be achieved, go back to step 4.

5. Comparison of Independent Components Analysis with Principal Components Analysis and An Exemplary Application

In the third part, it was mentioned that ICA can be considered as an advanced version of PCA which is widely used in the field of statistics. However, as it can be understood from what has been described so far, although there are similarities between the two analyzes, these analyses have many different aspects. In this context, similar and different aspects of these analyzes are gathered under this heading and listed as follows:

- Both PCA and ICA is a statistical transformation technique.
- PCA is based on normal distribution since it uses the information (covariance or correlation values) included in the second-order statistics of the measured data. ICA, on the other hand, uses the information in its higher-order statistics (kurtosis) as it takes advantage of the non-normality features of the data. Therefore, PCA can only eliminate correlations between variables, but not higher-order dependencies. In contrast, ICA eliminates both correlations and high-level dependencies and is a stronger technique compared to PCA [4].
- If the data is normally distributed, PCA is sufficient, but if the data is not normally distributed, the use of ICA is necessary because higher-order statistics will be required [4].
- The principal components obtained in PCA (eigenvectors of the correlation or covariance matrix) are orthogonal to each other. However, the components obtained in ICA are not orthogonal to each other as seen in Figure 2 [4].
- In PCA, the principal components that best explain the variance in the data are determined and the first principal component explains the greatest part of the variance. Then, the second principal component explains the second largest part of the variance. Moreover, since the first principal component and the second principal component are orthogonal to each other, they are unrelated. In other words, PCA maximizes the variance and therefore the principal components obtained in PCA correspond to the directions in which the variance is at maximum. On the other hand, ICA maximizes the non-normality of the components to be revealed and ensures that independent components are obtained. Therefore, ICA vectors correspond to the axes of the data as demonstrated in Figure 2 [4].

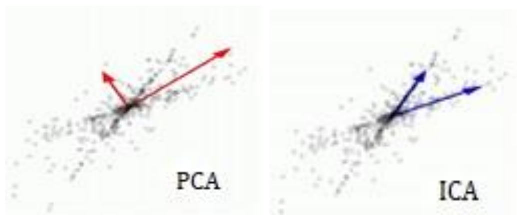

Figure 2. An example for components obtained with Principal Components Analysis and Independent Components Analysis

- As prioritization can be done for the principal components obtained in PCA, some components are more important than others. However, since such a ranking cannot be made in ICA, all the independent components obtained are considered to be of equal importance [4].
- While the main purpose of PCA is to provide the data representation in lower sizes, the main purpose of ICA is to ensure that the data is represented by independent orthogonal vectors as much as possible. In short, while PCA gathers and compresses the data, ICA enables the data to be decomposed [2].

This section also includes an exemplary application where PCA and ICA are compared on the same dataset and ICA yields better results. In application, six distinct and non-normally distributed series (signals) of 550 observations ($t = 1,2, \dots ,550$) are generated in the R software [4, 47]. The functions that the series are produced are given below from Equation 24 to Equation 29. In addition, the graphics of these series can be seen from Figure 3 to Figure 8. Studies of Mutihac and Vun Helle [53] are taken as the basis for the functions used in the production of the series.

$$s_1(t) = 2sin(t/180)cos(t/12) + 0.25(uniform{\sim}[0,1]) \quad (24)$$

$$s_2(t) = sign\left(sin(12t + 8cos(2/33))\right) + 0.15(uniform{\sim}[0,1]) \quad (25)$$

$$s_3(t) = (rem(t, 85) - 13)/32 + 0.15(uniform{\sim}[0,1]) \quad (26)$$

$$s_4(t) = ((rem(t, 29) - 15)/11)^5 + 0.20(uniform{\sim}[0,1]) \quad (27)$$

$$s_5(t) = 4exp(-t/144)cos(38t) + 0.15(uniform{\sim}[0,1]) \quad (28)$$

$$s_6(t) = 3\big((uniform{\sim}[0,1]) < 0.7\big)log(uniform{\sim}[0,1]) \quad (29)$$

The $\mathbf{X}$ matrix observed by using these produced series as independent components is created as in Equation 30. While creating the observed $\mathbf{X}$ matrix, the $\mathbf{A}$ mixture matrix in Equation 31 is used. Observation graphs of the created $\mathbf{X}$ matrix can be seen starting from Figure 9.

$$\mathbf{X} = \mathbf{AS} \quad (30)$$

$$\mathbf{A} = \begin{bmatrix} 5.09 & 2.46 & 0.94 & 9.90 & 4.68 & 8.74 \\ 9.56 & 9.32 & 3.16 & 4.33 & 8.37 & 1.47 \\ 4.05 & 1.07 & 5.14 & 9.83 & 0.64 & 8.60 \\ 5.53 & 0.10 & 5.03 & 6.26 & 2.87 & 9.64 \\ 4.07 & 1.36 & 2.56 & 6.06 & 6.28 & 0.92 \\ 1.90 & 4.62 & 3.59 & 1.61 & 1.97 & 1.47 \end{bmatrix} \quad (31)$$

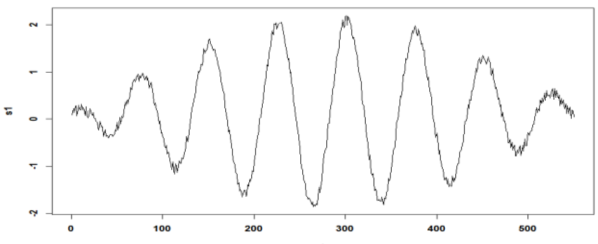

Figure 3. $\mathbf{s}_1$ series produced for **S** matrix

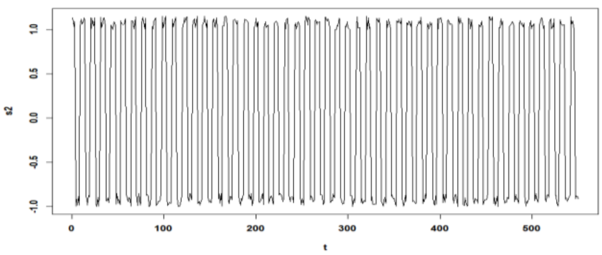

Figure 4. $\mathbf{s}_2$ series produced for **S** matrix

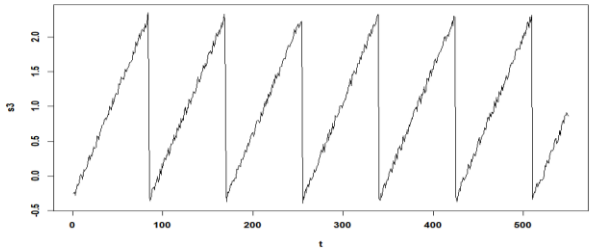

Figure 5. $\mathbf{s}_3$ series produced for **S** matrix

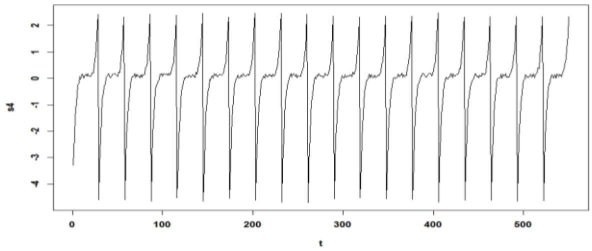

Figure 6. $\mathbf{s}_4$ series produced for **S** matrix

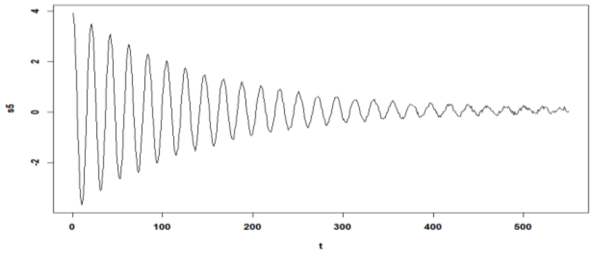

Figure 7. $\mathbf{s}_5$ series produced for **S** matrix

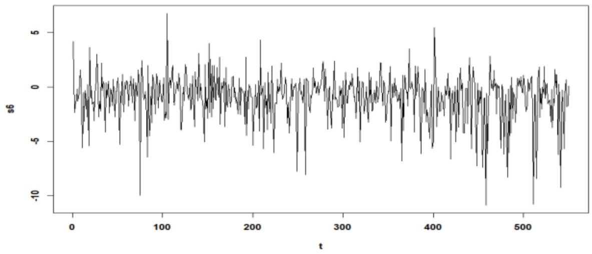

Figure 8. $\mathbf{s}_6$ series produced for **S** matrix

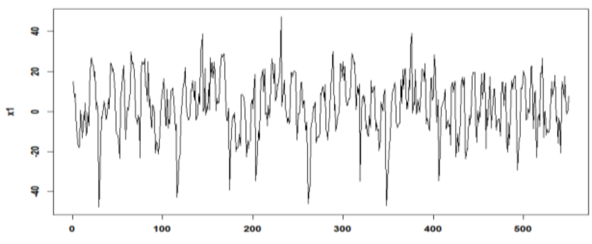

Figure 9. Observed $\mathbf{x}_1$ series in the generated **X** matrix

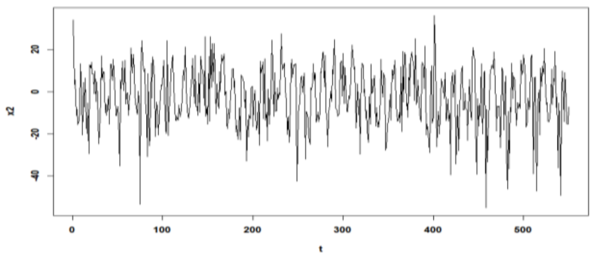

Figure 10. Observed $\mathbf{x}_2$ series in the generated **X** matrix

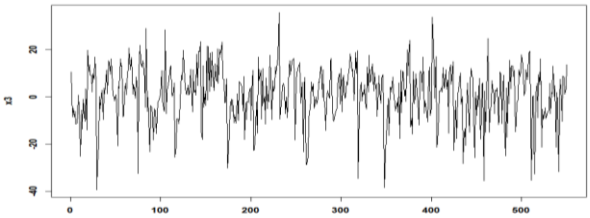

Figure 11. Observed $\mathbf{x}_3$ series in the generated **X** matrix

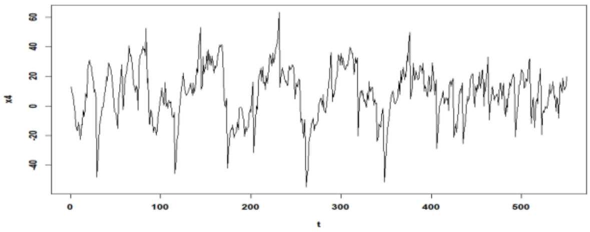

Figure 12. Observed $\mathbf{x}_4$ series in the generated **X** matrix

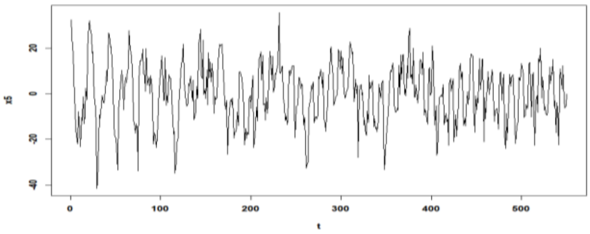

Figure 13. Observed $\mathbf{x}_5$ series in the generated **X** matrix

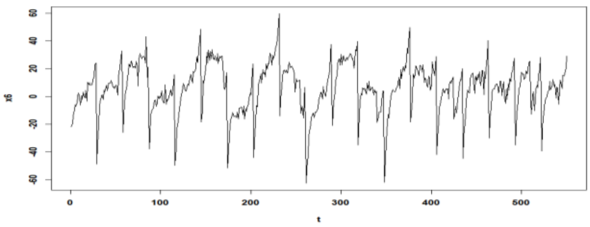

Figure 14. Observed $\mathbf{x}_6$ series in the generated **X** matrix

These created observations are then tried to be decomposed into their original components by ICA and PCA. The estimates of the **A** matrix obtained as a result of ICA using the FastICA algorithm and PCA are demonstrated in Equation 32 and Equation 33, respectively.

$$\hat{\mathbf{A}}_{\mathbf{ICA}} = \begin{bmatrix} 9.83 & 9.40 & 3.51 & 4.74 & 8.41 & 2.05 \\ 5.02 & 2.08 & 0.87 & 9.78 & 4.52 & 8.67 \\ 3.73 & 0.80 & 4.87 & 9.22 & 0.36 & 8.08 \\ -5.32 & 0.29 & -4.64 & -5.81 & -2.67 & -9.27 \\ -4.65 & -1.51 & -3.13 & -7.11 & -7.34 & -1.03 \\ -2.59 & -6.22 & -4.99 & -2.53 & -2.53 & -2.50 \end{bmatrix} \quad (32)$$

$$\hat{\mathbf{A}}_{\mathbf{PCA}} = \begin{bmatrix} -0.44 & -0.30 & -0.32 & -0.52 & -0.35 & -0.47 \\ 0.17 & 0.68 & 0.11 & -0.29 & 0.32 & -0.56 \\ -0.25 & 0.31 & 0.65 & -0.31 & -0.48 & 0.31 \\ 0.60 & -0.36 & 0.10 & -0.65 & 0.17 & 0.21 \\ -0.15 & -0.48 & 0.67 & 0.17 & 0.30 & -0.42 \\ 0.58 & -0.04 & 0.08 & 0.28 & -0.66 & -0.38 \end{bmatrix} \quad (33)$$

Original components estimated by ICA are given starting from Figure 15 and original components estimated by PCA are given starting from Figure 21. When the graphs are analyzed together, it is seen that ICA finds almost the same components (series) as the components of the **S** matrix and therefore is more successful than PCA in the determination of original components that do not exhibit a normal distribution.

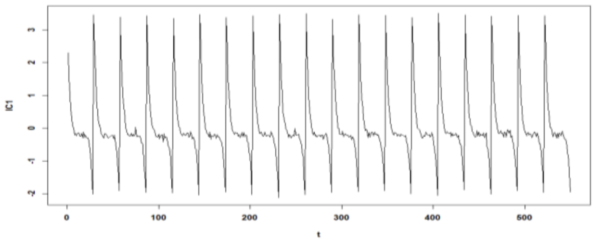

Figure 15. Independent component-1 estimated by ICA (IC1)

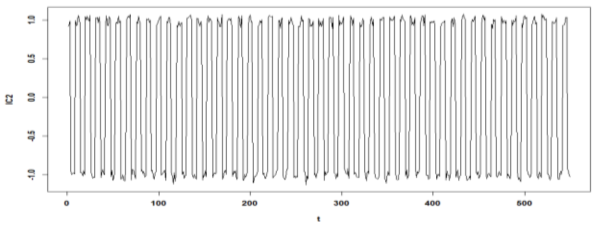

Figure 16. Independent component-2 estimated by ICA (IC2)

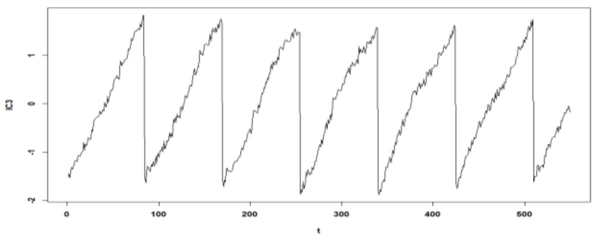

Figure 17. Independent component-3 estimated by ICA (IC3)

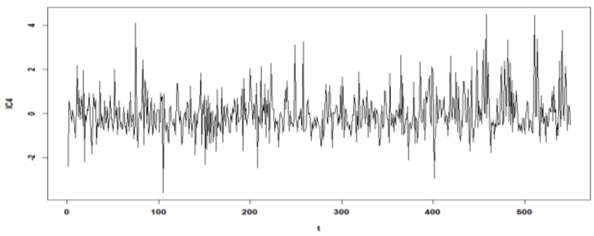

Figure 18. Independent component-4 estimated by ICA (IC4)

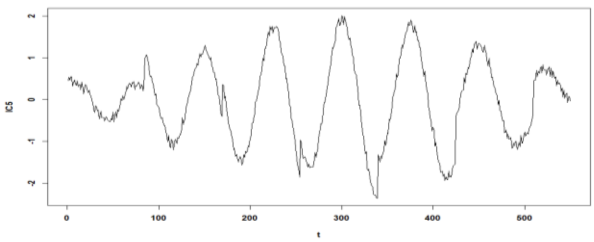

Figure 19. Independent component-5 estimated by ICA (IC5)

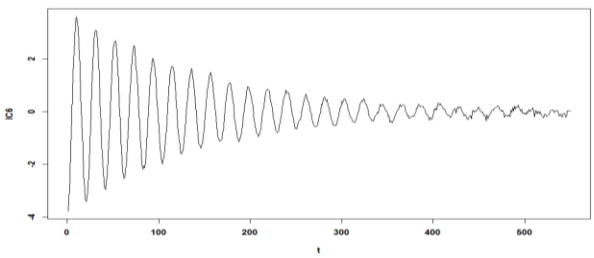

Figure 20. Independent component-6 estimated by ICA (IC6)

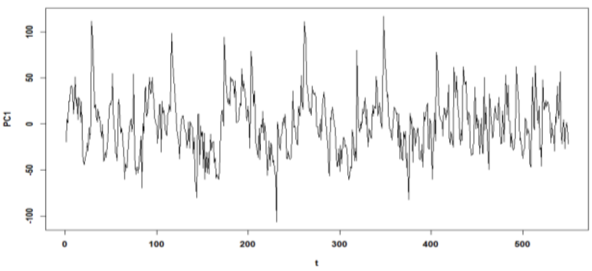

Figure 21. Principal component-1 estimated by PCA (PC1)

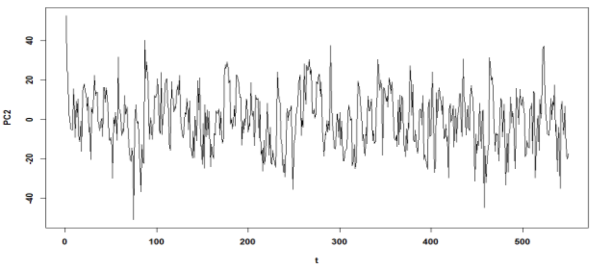

Figure 22. Principal component-2 estimated by PCA (PC2)

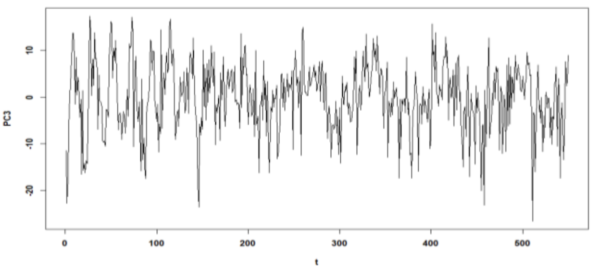

Figure 23. Principal component-3 estimated by PCA (PC3)

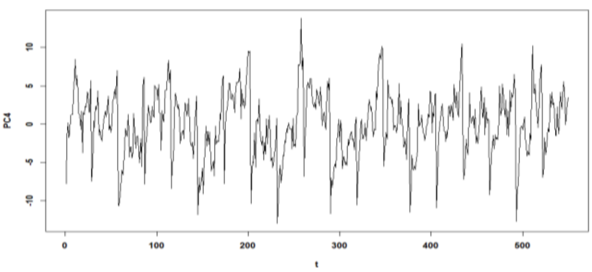

Figure 24. Principal component-4 estimated by PCA (PC4)

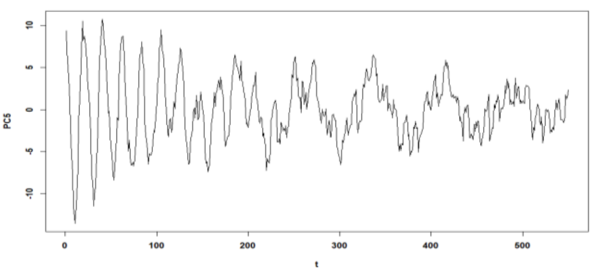

Figure 25. Principal component-5 estimated by PCA (PC5)

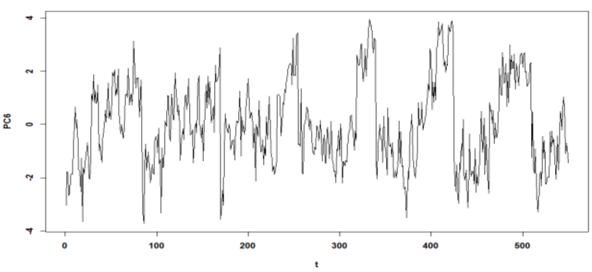

Figure 26. Principal component-6 estimated by PCA (PC6)

It is also seen in the correlation matrices in Table 1 and Table 2 that ICA is more successful than PCA in the determination of original components.

Table 1. Pearson correlation coefficients between the independent components and the original components

	$\mathbf{s}_1$	$\mathbf{s}_2$	$\mathbf{s}_3$	$\mathbf{s}_4$	$\mathbf{s}_5$	$\mathbf{s}_6$
IC1	-0.003	0.013	-0.038	**-0.999**	0.005	0.009
IC2	0.005	**0.999**	0.009	-0.005	-0.001	0.004
IC3	0.269	-0.001	**0.993**	0.002	0.045	-0.002
IC4	0.009	-0.004	-0.019	0.041	-0.036	**-0.999**
IC5	**0.963**	-0.007	-0.108	-0.006	-0.006	-0.013
IC6	0.013	-0.001	-0.008	0.001	**-0.998**	0.019

Table 2. Pearson correlation coefficients between the principal components and the original components

	$\mathbf{s}_1$	$\mathbf{s}_2$	$\mathbf{s}_3$	$\mathbf{s}_4$	$\mathbf{s}_5$	$\mathbf{s}_6$
PC1	-0.483	-0.423	-0.429	-0.499	-0.293	-0.354
PC2	-0.313	**0.617**	-0.328	-0.470	0.108	0.503
PC3	-0.333	-0.330	0.132	0.253	-0.473	**0.660**
PC4	-0.449	0.354	-0.407	**0.654**	-0.157	-0.320
PC5	**-0.490**	-0.306	0.079	0.126	**-0.796**	0.104
PC6	-0.342	0.334	**0.721**	-0.146	-0.140	-0.269

When Table 1 is examined, it is observed that the components estimated by ICA have almost 100% relationship with the components in the **S** matrix. According to Table 1, the fifth independent component (IC5) estimated by ICA corresponds to $\mathbf{s}_1$, second independent component (IC2) corresponds to $\mathbf{s}_2$, third independent component (IC3) corresponds to $\mathbf{s}_3$, first independent component (IC1) corresponds to $\mathbf{s}_4$, sixth independent component (IC6) corresponds to $\mathbf{s}_5$, and fourth independent component (IC4) corresponds to $\mathbf{s}_6$. In addition, it is obviously seen that Figure 15 is the opposite of Figure 6, since IC1 has a negative relationship with the corresponding component. Similar comments are also valid for IC4 and IC6.

When Table 2 is examined, it is evident that the components estimated by PCA do not have very high level relations with the components in the **S** matrix. For example, the first principal component (PC1) estimated by PCA; has almost the same amount of relationships with $\mathbf{s}_1$, $\mathbf{s}_2$, $\mathbf{s}_3$, and $\mathbf{s}_4$. For this reason, it cannot be determined which original components correspond to the basic components obtained as a result of PCA.

6. Conclusion

The main aim of the study is to increase the awareness of ICA, which is not known by statisticians although it is a multivariate statistical method and to contribute to the limited literature on the subject. In this context, firstly, the historical process and application areas of ICA are discussed in detail. Afterward, a comprehensive resource is created on the subject by including the reason for the emergence of ICA, the basic ICA model, the features of the model, the algorithms used in the solution of the model, and related software. In the application part, the comparison of ICA and PCA, which is a method similar to ICA and is frequently used in statistics, is evaluated on an artificial dataset. As a result of the application, it is concluded that ICA is much more successful than PCA in finding original components (sources) that are not distributed normally. As is known, ICA is a powerful technique that gives better results in cases where PCA fails. This is because, unlike PCA, it is based on statistical independence rather than unrelatedness, and statistical independence is a much stronger technique than unrelatedness.

When evaluated in general, it is thought that the study will meet the detailed resource needs of researchers on ICA who work or will work in a wide range of fields from telecommunications to chemistry, geology to the economy, lead new studies in the field of statistics related to ICA, and additionally, it will contribute to the widespread use of this alternative and stronger technique in researches that do not yield results with PCA.

Acknowledgment

This study was produced from a part of the doctoral thesis prepared by the first author under the supervision of the second author. Therefore, thanks to thesis jury members for their valuable contributions and suggestions.

References

[1] Hyvärinen, A., Karhunen, J., Oja, E. 2001. Independent Component Analysis. John Wiley&Sons, New York, 504p.

[2] Shlens, J. 2014. A Tutorial on Independent Component Analysis. https://arxiv.org/pdf/1404.2986.pdf (Accessed Date: 01.21.2019).

[3] Ozdamar, E.O. 2009. EEG Analizinde Bağımsız Bileşenler. Mimar Sinan University, Graduate School of Science and Engineering, Doctoral Thesis, 125p, Istanbul.

[4] Bursa, N. 2019. Bağımsız Bileşenler Analizi ile Çoklu Bağlantı Sorununa Bir Yaklaşım. Hacettepe University, Graduate School of Science and Engineering, Doctoral Thesis, 151p, Ankara.

[5] Hérault, J., Jutten, C., Ans, C. 1998. Détection de Grandeurs Primitives dans un Message Composite par une Architecture de Calcul Neuromimétique en Apprentissage non Suprévise.http://documents.irevues.inist.fr/bitstream/handle/2042/10937/AR12_9.pdf?sequence=1 (Accessed Date: 05.23.2019).

[6] Jutten C., Hérault, J. 1991. Blind Separation of Sources, Part I: An Adaptive Algorithm Based on Neuromimetric Architecture. Signal Processing, 24(1), 1-10.

[7] Jutten, C., Hérault, J. 1991. Blind Separation of Sources, Part II: Problems Statement. Signal Processing, 24(1), 11-20.

[8] Jutten, C., Hérault, J. 1991. Blind Separatrion of Sources, Part III: Stability Analysis. Signal Processing, 24(1), 21-29.

[9] Comon, P. 1994. Independent Component Analysis, a New Concept?, Signal Processing, 36(3), 287-314.

[10] Artoni, F., Delorme, A., Makeig, S. 2019. A Visual Working Memory Dataset Collection with Bootstrap Independent Component Analysis for Comparison of Electroencephalographic Preprocessing Pipelines. Data in Brief, 22, 787-793.

[11] Tierney, J.E., Wilkes, D.M., Byram, B.C. 2019. Independent Component Analysis-Based Tissue Clutter Filtering for Plane Wave Perfusion Ultrasound Imaging. Medical Imaging: Ultrasonic Imaging and Tomography, 17-18 February, San Diego, 2.

[12] Baker, B.T., Abrol, A., Silva, R.F., Damaraju, E., Sarwate, A.D., Calhoun, V.D., Plis, S.M. 2019. Decentralized Temporal Independent Component Analysis: Leveraging Fmrı Data in Collaborative Settings. Neurolmage, 186, 557-569.

[13] Albert, S.A., Bowman, D.C. 2018. Tracking Scattered Signals in the Acoustic Coda Using Independent Component Analysis in a Topographically Complex Setting. Geophysical Journal International, 216(2), 767-776.

[14] de Lauro, E., Petrosino, S., Falanga, M. 2018. Independent Component Analysis as a Monitoring Tool in Geophysical Environment: The Case of Campi Flegrei (Italy). IEEE Workshop on Environmental, Energy, and Structural Monitoring Systems, 21-22 June, Salerno, 1-6.

[15] Cohen-Waeber, J., Bürgmann, R., Chaussard, E., Giannico, C., Ferretti, A. 2018. Spatiotemporal Patterns of Precipitation-Modulated Landslide Deformation from Independent Component Analysis of InSAR Time Series. Geophysical Research Letters, 45(4), 1878-1887.

[16] Garcia-Bracamonte, J.E., Rangel-Magdaleno, J., Ramirez-Cortes, J.M., Gomez-Gill, P., Paregrina-Barreto, H. 2018. Induction Motors Fault Detection Using Independent Component Analysis on Phase Current Signals, IEEE International Instrumentation and Measurement Technology Conference, 14-17 May, Houstan, 1-6.

[17] Yu, J., Yoo, J. , Jang, J., Park, J.H., Kim, S. 2018. A Novel Hybrid of Auto-Associative Kernel Regression and Dynamic Independent

Component Analysis for Fault Detection in Nonlinear Multimode Processes. Journal of Process Control, 68, 129-144.

[18] Li, Z., Yan, X. 2018. Adaptive Selective Ensemble-Independent Component Analysis Models for Process Monitoring. Industrial & Engineering Chemistry Research, 57(24), 8240-8252.

[19] Lahaw, Z.B., Essaidani, D., Seddik, H. 2018. Robust Face Recognition Approaches Using PCA, ICA, LDA Based on DWT and SVM Algorithms. 41st International Conference on Telecommunications and Signal Processing, 4-6 July, Atheans, 1-5.

[20] Wang, Y., Guo, Y. 2019. A Hierarchical Independent Component Analysis Model for Longitudinal Neuroimaging Studies. Neurolmage, 189, 380-400.

[21] Koush, Y., Masala, N., Scharnowski, F., De Ville, D.V. 2019. Data-Driven Tensor Independent Component Analysis for Model-Based Connectivity Neurofeedback. Neurolmage, 184, 214-226.

[22] Jiang, C., Zhang, H., Ren, Y., Han, Z., Chen, K., Hanzo, L. 2017. Machine Learning Paradigms for Next-Generation Wireless Networks. IEEE Wireless Communications, 24(2), 98-105.

[23] Wang, C., Xu, Y., Tang, M., Wang, L. 2018. Blind Source Separation Based on Variational Bayesian Independent Component Analysis. IEEE 3rd Advanced Information Technology. Electronic and Automation Control Conference, 12-14 October, Chongqing, 1614-1618.

[24] Aveta, F., Refai, H., Lo Presti, P., Tedder, A.S., Schoenholz, B.L. 2018. Independent Component Analysis for Processing Optical Signals in Support of Multi-User Communication. Free-Space Laser Communication and Atmospheric Propagation XXX, 29-30 January, San Francisco, 1-9.

[25] Gouriéroux, C., Monfort, A., Renne, J.P. 2017. Statistical Inference for Independent Component Analysis: Application to Structural VAR Models. Journal of Econometrics, 196(1), 111-126.

[26] Chowdhury, U.N., Chakravarty, S.K., Hossain, M.T. 2018. Short-Term Financial Time Series Forecasting Integrating Principal Component Analysis and Independent Component Analysis with Support Vector Regression. Journal of Computer and Communications, 6(3), 51-67.

[27] Chen, Y., Niu, L., Chen, R.B., He. Q. 2019. Sparse-Group Independent Component Analysis with Application to Yield Curves Prediction. Computational Statistics & Data Analysis, 133, 76-89, 2019.

[28] Witten, I.H., Frank, E., Hall, M.A., Pal, C.J. 2016. Data Mining: Practical Machine Learning Tools and Techniques, Elsevier, San Francisco, 525p.

[29] Thomas, M.C., Zhu, W., Romagnoli, J.A. 2018. Data Mining and Clustering in Chemical Process Databases for Monitoring and Knowledge Discovery. Journal of Process Control, 67, 160-175.

[30] Gultepe, E., Makrehchi, M. 2018. Improving Clustering Performance Using Independent Component Analysis and Unsupervised Feature Learning. Human-centric Computing and Information Sciences, 8(25), 1-19.

[31] Zhou W., Altman, R.B. 2018. Data-Driven Human Transcriptomic Modules. BMC Bioinformatics, 19(327), 1-25.

[32] Kamal, M.S., Trivdedi, M.C., Alam, J.B., Dey, N., Ashour, A.S., Shi, F., Tavares, J.M.R. 2018. Big DNA Datasets Analysis Under Push Down Automata. Journal of Intelligent & Fuzzy Systems, 35(2), 1555-1565.

[33] Ghosh, M., Adhikary, S., Kanti Ghosh, K., Sardar, A., Begum, S., Sarkar, R. 2019. Genetic Algorithm Based Cancerous Gene Identification from Microarray Data Using Ensemble of Filter Methods. Medical & Biological Engineering & Computing, 57(1), 159-176.

[34] Liu, Y., Xu, H., Xia, Z., Gong, Z. 2018. Multi-Spectrometer Calibration Transfer Based On Independent Component Analysis. Analyst, 143(5), 1274-1280.

[35] Alves, F.C.G.B.S., Coqueiro, A., Março, P.H., Valderrama, P. 2019. Evaluation of Olive Oils from the Mediterranean Region by UV–Vis Spectroscopy and Independent Component Analysis. Food Chemistry, 273, 124-129.

[36] Delaporte, G., Cladiére, M., Bouveresse, D.J.R., Camel, V. 2019. Untargeted Food Contaminant Detection Using UHPLC-HRMS Combined with Multivariate Analysis: Feasibility Study on Tea. Food Chemistry, 277, 54-62.

[37] Stone, J.V. 2004. Independent Component Analysis: A Tutorial Introduction. MIT Press, London, 191p.

[38] Nordhausen K., Oja, H. 2018. Independent Component Analysis: A Statistical Perspective. Wiley Interdisciplinary Reviews: Computational Statistics, 10(5), 1-23.

[39] Hyvärinen, A. 2013. Independent Component Analysis: Recent Advances. Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences, 371, 1-19.

[40] Hyvärinen, A., Oja, E. 2000. Independent Component Analysis: Algorithms and Applications. Neural Networks, 13(4-5), 411-430.

[41] Tharwat, A. 2018. Independent Component Analysis: An Introduction, Applied Computing

and Informatics. https://doi.org/10.1016/j.aci.2018.08.006 (Accessed Date: 01.28.2018).

[42] Hyvärinen, A. 1999. Fast and Robust Fixed-Point Algorithms for Independent Component Analysis. IEEE Transactions on Neural Networks, 10(3), 626-634.

[43] Matthias, K., Haueisen, J., Ivanova, G. 2009. Independent Component Analysis: Comparison of Algorithms for the Investigation of Surface Electrical Brain Activity. Medical & Biological Engineering & Computing, 47(4), 413-423.

[44] Naik, G.R. 2011. A Comparison of ICA Algorithms in Surface EMG Signal Processing. International Journal of Biomedical Engineering and Technology, 6(4), 363-374.

[45] Dharmaprani, D., Nguyen, H.K., Lewis, T.W., DeLosAngeles, D., Willoughby, J.O., Pope, K.J. 2016. A Comparison of Independent Component Analysis Algorithms and Measures to Discriminate Between EEG and Artifact Components. 38th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, 16-20 August, Orlando, 825-828.

[46] Sahonero-Alvarez, G., Calderon, H. 2017. A Comparison of SOBI, FastICA, JADE and Infomax Algorithms. 8th International Multi-Conference on Complexity, Informatics and Cybernetics, 21-24 March, Orlando, 17-22.

[47] R Core Team. 2019. R: A Language and Environment for Statistical Computing. https://www.r-project.org/ (Accessed Date: 03.01.2020).

[48] Miettinen, J., Nordhausen, K., Taskinen, S. 2017. Blind Source Separation Based on Joint Diagonalization in R: The Packages JADE and BSSasymp, Journal of Statistical Software, 76(2), 1-31.

[49] Miettinen, J., Nordhausen, K., Oja, H., Taskinen, S. 2017. fICA: Classical, Reloaded and Adaptive Fastıca Algorithms. https://cran.r-project.org/web/packages/fICA/index.html (Accessed Date: 08.11.2019).

[50] Helwig. N.E. 2015. Ica: Independent Component Analysis. https://cran.r-project.org/web/packages/ica/index.html (Accessed Date: 08.11.2019).

[51] Marchini, J.L., Heaton, C., Ripley. B.D. 2017. fastICA: FastICA Algorithms to Perform ICA and Projection Pursuit. https://cran.r-project.org/web/packages/fastICA/index.html (Accessed Date: 09.11.2019).

[52] Naik, G.R., Kumar, D.K. 2011. An Overview of Independent Component Analysis and Its Applications. Informatica, 35(2011), 63-81.

[53] Mutihac, R., Van Hulle, M.M. 2004. Comparison of Principal Component Analysis and Independent Component Analysis for Blind Source Separation. Romanion Reports in Physics, 56(1), 20-32.