JOURNAL OF ARTIFICIAL INTELLIGENCE AND COMPUTATIONAL THEORY
JAICT

Independent Component Analysis Based on Chatterjee's Correlation Coefficient

Hossein Nadeb

Department of Statistics, Yazd University, Yazd, Iran
honadeb@yahoo.com

Article Information

Received: June 22, 2026
Revised: July 18, 2026
Accepted: July 25, 2026

Keywords

Amari error, Clustering, Monte Carlo simulation, Time series

*Corresponding Author: Hossein Nadeb

Abstract

In conventional Independent Component Analysis (ICA) algorithms, the objective function is typically defined based on specific dependency criteria. The choice of these criteria is not merely a technical detail; it fundamentally influences the algorithm's convergence speed, robustness to outliers, and ability to separate sources under non-ideal conditions such as limited sample sizes or noisy mixtures. This article introduces a novel ICA algorithm that uses Chatterjee's correlation coefficient as the contrast function. The performance of the proposed method is evaluated and compared against two recent ICA algorithms, using the Amari error as a quantitative benchmark across multiple simulation scenarios. Furthermore, to demonstrate its practical utility, the proposed algorithm is applied to real-world time series data as a pre-processing step for clustering.

Please cite this paper as:

1 Introduction

Multivariate data present substantial analytical challenges due to their high dimensionality and the complex interdependencies among observed variables. It is therefore essential to decompose such data into interpretable components in order to reveal meaningful underlying structures. ICA has been extensively employed as an unsupervised tool for blind source separation in multivariate statistics, particularly for disentangling mixtures of independent source signals [21]. It was formalized and extended by [10, 13]. Its application to signal processing has been further investigated in [11, 19].

ICA is a matrix factorization technique that optimizes the mutual independence among the source signals represented by each factor matrix. The primary objective of ICA is to extract salient components from a dataset, which has been applied in diverse scientific fields, such as medical signal analysis, audio noise reduction, climate data analysis, and econometrics.

Fundamentally, ICA is a computational technique that decomposes a multivariate signal into its constituent additive sub-components. It seeks to identify the independent components by maximizing the statistical independence among the estimated sources. The specific formulation of an ICA algorithm depends on the chosen independence criterion.

In this study, we introduce a novel ICA algorithm based on Chatterjee's correlation coefficient. To assess its performance, we conduct a comparative evaluation between the proposed method and two recently developed existing approaches. Finally, we demonstrate the utility of our algorithm by applying it as a pre-processing step of time series clustering.

The rest of this paper is structured as follows. Section 2 provides an overview of the ICA framework. Section 3 reviews Chatterjee's correlation coefficient. In Section 4, we present our proposed ICA algorithm. Its comparative performance, measured in terms of the average Amari error, is evaluated via Monte Carlo simulations against two existing ICA algorithms in Section 5. Section 6 illustrates the application of our proposed algorithm to time series data clustering.

2 ICA

In this study, ICA is employed for signal analysis. Each signal is treated as a time-varying vector, denoted by $\boldsymbol{S}_i = (S_{1i}, \ldots, S_{ni})'$, for $i = 1, \ldots, d$, where n is the number of time steps, S_{ji} is the value of the signal $\boldsymbol{S}_i$ at time j, and d is the number of source signals. These source signals are collected into the matrix $\mathbf{S}_{n\times d} = (\boldsymbol{S}_1, \ldots, \boldsymbol{S}_d)$. The source signals may undergo mixing, with each source contributing differently to the observed outputs. Consequently, the d mixtures can be expressed as the mixed signal matrix $\mathbf{X}_{n\times d} = \mathbf{S}A$, where $A_{d\times d}$ is the matrix of mixing coefficients. The primary objective of the model is to recover the original source signals by estimating an unmixing coefficient matrix. This matrix is then applied to the mixed signals to produce a set of independent signals, given by $\mathbf{Y} = \mathbf{X}W$, where $W_{d\times d}$ is the inverse of A. Centring and whitening of the data is an essential step to prepare the data for ICA and is often carried out using the principal component analysis.

Numerous studies have explored various ICA algorithms and their underlying interpretations. Most ICA methods seek to minimize a contrast function that quantifies the statistical dependence among the estimated components. The performance of these algorithms depends critically on both the choice of the contrast function and the optimization technique employed. The unmixing matrix W can be estimated using several different criteria for independence, each yielding slightly different solutions. Regardless of the specific approach, the general procedure involves obtaining an unmixing matrix and then projecting the whitened data onto it to recover the independent signals. Representative works in this area include those by [2, 3, 4, 6, 8, 12, 13, 14, 15, 16, 17, 20, 21, 22, 23, 24, 27].

3 Chatterjee's Correlation Coefficient

Let (X_1, X_2) be a pair of random variables with an absolutely continuous joint distribution function F and marginal distribution functions F_1 and F_2 for X_1 and X_2, respectively. [9] proposed a measure of association between X_1 and X_2 defined as

$$\xi(X_1, X_2) = \frac{\int \mathrm{Var}\left[\mathrm{E}\left[I(X_2 \geq t)\middle|X_1\right]\right] \mathrm{d}F_2(t)}{\int \mathrm{Var}\left[I(X_2 \geq t)\right] \mathrm{d}F_2(t)},$$

where I denotes the indicator function. This measure satisfies $0 \leq \xi(X_1, X_2) \leq 1$. It equals 0 if and only if X_1 and X_2 are independent, and equals 1 if and only if there exists a measurable function $g : \mathbb{R} \to \mathbb{R}$ such that $X_2 = g(X_1)$ almost surely.

Let $(X_{11}, X_{21}), \ldots, (X_{1n}, X_{2n})$ be a random sample of size n drawn from (X_1, X_2). Suppose that $(X_{1(1)}, X_{2(1)}), \ldots, (X_{1(n)}, X_{2(n)})$ is a rearrangement of the sample such that $X_{1(1)} \leq \cdots \leq$

$X_{1(n)}$ and $X_{2(1)} \leq \cdots \leq X_{2(n)}$, with possible ties. Define $R_i = \sum_{j=1}^{n} I\left(X_{2(j)} \leq X_{2(i)}\right)$ and $L_i = \sum_{j=1}^{n} I\left(X_{2(j)} \geq X_{2(i)}\right)$. [9] proposed an estimator for $\xi(X_1, X_2)$, denoted by $\xi_n(X_1, X_2)$, of the form

$$\xi_n(X_1, X_2) = 1 - \frac{n \sum_{i=1}^{n-1} |R_{i+1} - R_i|}{2 \sum_{i=1}^{n} L_i(n - L_i)}.$$

In the absence of ties among $X_{1(1)}, \ldots, X_{1(n)}$, this expression simplifies to

$$\xi_n(X_1, X_2) = 1 - \frac{3 \sum_{i=1}^{n-1} |R_{i+1} - R_i|}{n^2 - 1}.$$

[9] showed that, provided X_2 is not almost surely constant, $\xi_n(X_1, X_2)$ converges almost surely to $\xi(X_1, X_2)$. Because $\xi_n(X_1, X_2)$ relies solely on rank-based information, it remains resistant to the influence of outliers. Therefore, it can be used as a contrast function in ICA to avoid the adverse effects of outliers.

The measure $\xi(X_1, X_2)$ is not symmetric in X_1 and X_2. To assess whether X_1 is a function of X_2, one should use $\xi(X_2, X_1)$ rather than $\xi(X_1, X_2)$. To address this asymmetry, [9] proposed the symmetric dependence measure $\xi^*(X_1, X_2) = \max\left(\xi(X_1, X_2), \xi(X_2, X_1)\right)$. This quantity equals 0 if and only if X_1 and X_2 are independent, and equals 1 if and only if at least one variable is a measurable function of the other. Accordingly, in this paper we adopt $\xi^*(X_1, X_2)$ and estimate it using $\xi_n^*(X_1, X_2) = \max\left(\xi_n(X_1, X_2), \xi_n(X_2, X_1)\right)$.

4 The Proposed ICA Algorithm

In this section, we employ $\xi_n^*(X_1, X_2)$ to develop an algorithm for ICA, which we call CICA. Let $\mathbf{X}_{n\times d}$ be a mixed-signal matrix. The goal of ICA is to find a matrix W such that the columns of $\mathbf{Y} = \mathbf{X}W$ exhibit minimal statistical dependence. For ICA problems, pairwise independence serves as a sufficient measure of statistical independence [10]. Consequently, a d-dimensional ICA problem can be solved successively by solving all corresponding 2-dimensional ICA subproblems. Hence, the idea of searching for a rotation angle that minimizes $\xi_n^*(X_1, X_2)$ for the demixed data is analogous to the approach taken in the RADICAL algorithm [16]. In Algorithm 1, we present the CICA procedure for the d-dimensional case.

5 Simulation Study

To conduct a simulation study, the Monte Carlo method was employed to compare the performance of CICA with the algorithms RLICA proposed by [23], and GDCICA$_g$ with $g(x) = \log(1/\sqrt{x}) + \sqrt{x} - 1$ proposed by [3].

The data generation process for constructing the matrix $\mathbf{X}$ followed these steps:

- Generate a matrix $S_{n\times d}$, where n denotes the number of observations and d denotes the number of variables ($n \geq d$). Each column consists of a random sample of size n drawn from a specific distribution. The study incorporated 18 distinct one-dimensional densities recommended by [5], including distributions such as Student's t, uniform, exponential, a mixture of two Laplace densities, and symmetric and nonsymmetric Gaussian mixtures. Figure 1 displays the density plots of these 18 distributions.

- Generate a random matrix

$$A_{d\times d} = \begin{pmatrix} U_{11} & U_{12} & \cdots & U_{1d} \\ \vdots & \vdots & \vdots & \vdots \\ U_{d1} & U_{d2} & \cdots & U_{dd} \end{pmatrix},$$

  where $U_{ij}, i, j = 1, 2, \cdots, d$, are independently drawn from an arbitrary random distribution.

- Construct the data matrix $\mathbf{X}_{n\times d}$ using the formula $\mathbf{X} = \mathbf{S}A$.

Algorithm 1 Procedure of CICA

Input: A matrix $\mathbf{X}_{n\times d}$ whose rows contain the mixed signals (centered).

1: Whiten the matrix $\mathbf{X}$ as $\mathbf{Y} = \mathbf{X} \times Q'$, where Q is a whitening matrix.

2: Initialize the unmixing matrix $\widehat{W}_{d\times d}$ and the source-signal matrix $\widehat{\mathbf{S}}_{n\times d}$ as zero matrices.

3: Let $R_{ij}(\theta)$ denote a $d \times d$ rotation matrix acting on elements (i, j), defined as follows:

$$R_{ij}(\theta) = \begin{pmatrix} 1 & \cdots & 0 & 0 & \cdots & 0 & 0 & \cdots & 0 \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & r_{ii} = \cos(\theta) & 0 & \cdots & 0 & r_{ij} = -\sin(\theta) & \cdots & 0 \\ 0 & \cdots & 0 & 1 & \cdots & 0 & 0 & \cdots & 0 \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & 0 & \cdots & 1 & 0 & \cdots & 0 \\ 0 & \cdots & r_{ji} = \sin(\theta) & 0 & \cdots & 0 & r_{jj} = \cos(\theta) & \cdots & 0 \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & 0 & \cdots & 0 & 0 & \cdots & 1 \end{pmatrix}; \tag{1}$$

For all $1 \leqslant i < j \leqslant d$, repeat the following steps:

4: Define $\xi_n^*(S_i(\theta), S_j(\theta))$ as a function of θ, where the (i, j)-th columns of $\mathbf{Y} \times R_{i,j}(\theta)$ constitute a sample from $(S_i(\theta), S_j(\theta))$.

5: Minimize $\xi_n^*(S_i(\theta), S_j(\theta))$ over $\theta \in [-\frac{\pi}{2}, \frac{\pi}{2}]$ and set

$$\theta_0^{(i,j)} = \underset{\theta}{\operatorname{argmin}}\ \xi_n^*(S_i(\theta), S_j(\theta)).$$

6: Update the (i, j)-th columns of $\widehat{W}$ as $R'_{i,j}(\theta_0^{(i,j)}) \times Q$, and update the (i, j)-th columns of $\widehat{\mathbf{S}}$ as $\mathbf{Y} \times R_{ij}(\theta_0^{(i,j)})$.

Output: The unmixing matrix $\widehat{W}$ and the matrix of estimated source signals $\widehat{\mathbf{S}}$.

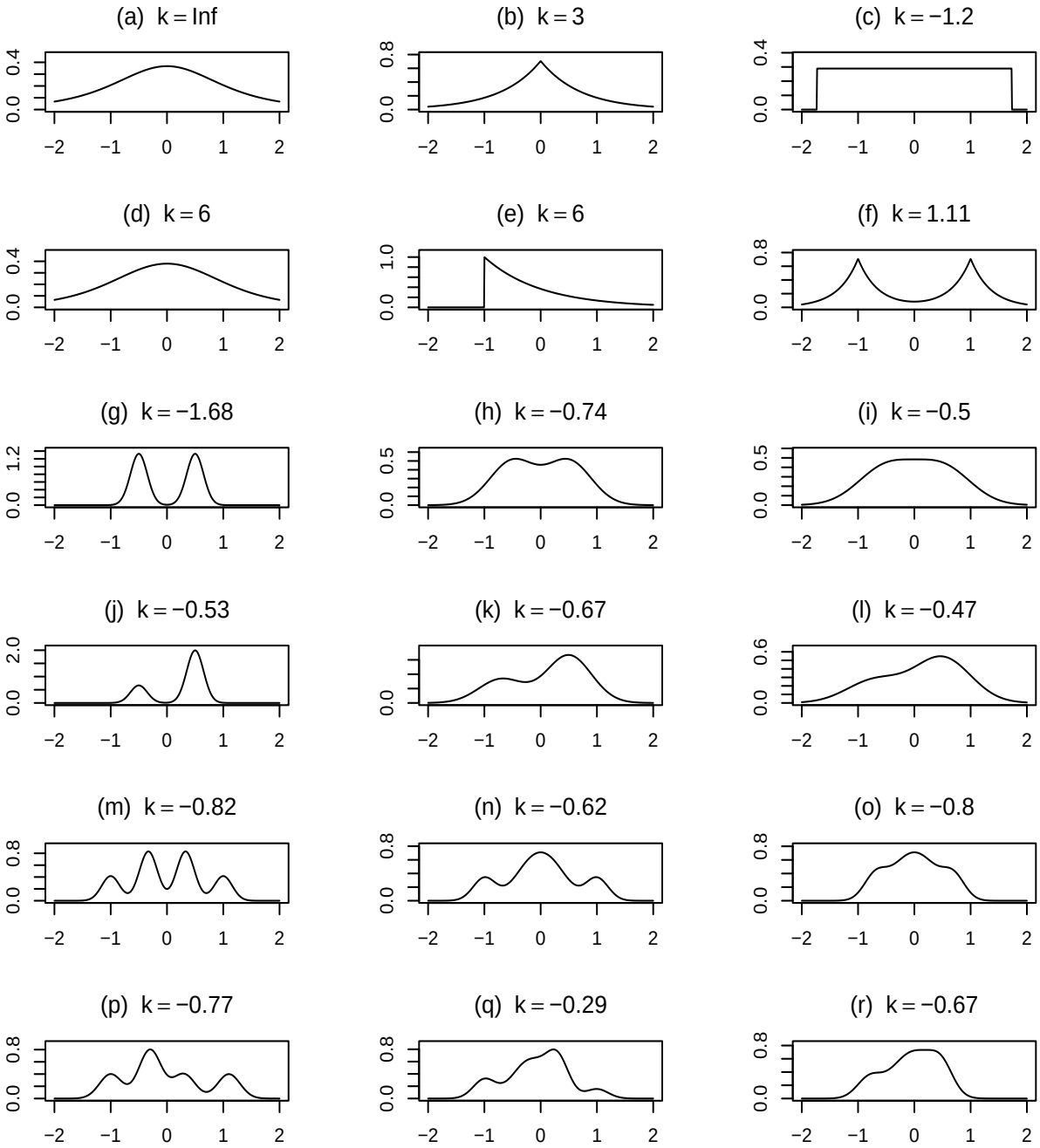

Figure 1: Density plots of 18 different source distributions: (a) Student's t with 3 degrees of freedom; (b) double exponential; (c) uniform; (d) Student's t with 5 degrees of freedom; (e) exponential; (f) mixture of two double exponentials; (g)-(h)-(i) symmetric mixtures of two Gaussians: multimodal, transitional and unimodal; (j)-(k)-(l) nonsymmetric mixtures of two Gaussians: multimodal, transitional and unimodal; (m)-(n)-(o) symmetric mixtures of four Gaussians: multimodal, transitional and unimodal; (p)-(q)-(r) nonsymmetric mixtures of four Gaussians: multimodal, transitional and unimodal.

In this study, we set $d = 2$, $n = 250$, and each $U_{ij}, i, j = 1, 2, \cdots, d$, follows a uniform distribution over $(0, 2)$.

To compare the performances, we used the Amari error criterion proposed by [1]. Let A denote the mixing matrix, and let the unmixing matrix be defined by $W = A^{-1}$, with $\widehat{W}$ as its estimator. The Amari error is computed using the following expression:

$$\text{Amari error} = \frac{1}{2d(d-1)} \sum_{i,j=1}^{d} \left(\frac{|a_{ij}|}{\max_i |a_{ij}|} + \frac{|a_{ij}|}{\max_j |a_{ij}|} \right) - \frac{1}{d-1},$$

where $a_{ij} = [\widehat{W}A]_{ij}$ denotes the (i, j)-th entry of the estimated unmixing matrix $\widehat{W}$. The Amari error lies within the range $[0, d-1]$, and equals zero if and only if $\widehat{W}$ and W are identical up to permutation and scaling of the columns. Moreover, this metric is invariant to column permutation and scaling in both A and $\widehat{W}$.

The average Amari errors were computed over 100 replications, and the results are presented in Table 1. The findings indicate that across the 18 cases, the CICA method consistently outperforms its competitors in 6 instances, specifically for distributions b, d, f, i, l, and r.

Furthermore, two distributions were selected from among all distributions labeled a to r. The average Amari errors for these selections were computed and are discussed in the final row of Table 1. The results reveal that when a distribution is randomly chosen, CICA yields better performance than its competitors.

The average Amari errors were also computed for performance evaluation across all methods under various distributions, categorized into different classes. First, the distributions were classified into unimodal distributions (a, b, d, e, i, l, o, r), multimodal distributions (f, g, j, m, p), and transitional distributions (c, h, k, n, q). Second, the perspective shifted to symmetric distributions (a, b, c, d, f, g, h, i, m, n, o) versus nonsymmetric distributions (e, j, k, l, p, q, r). Third, the distributions were grouped into positively kurtotic distributions (a, b, d, e, f) and negatively kurtotic distributions (c, g, h, i, j, k, l, m, n, o, p, q, r). Each of these classes formed a partition of the set of all distributions. The detailed results of this analysis are presented in Table 2. It is observed that the CICA method demonstrates superior performance for unimodal distributions and the positively kurtotic class.

6 Real Data

Let $\mathbf{X}$ be a time series matrix, and let $\hat{\mathbf{S}}$ and $\hat{A}$ be the estimates of the source signal matrix and the mixing coefficient matrix obtained by applying an ICA algorithm to $\mathbf{X}$, respectively. Thus, the time series matrix $\mathbf{X}$ can be approximated by $\hat{\mathbf{S}}\hat{A}$, that is, $\mathbf{X} \approx \hat{\mathbf{S}}\hat{A}$. This means that the columns of $\hat{\mathbf{S}}$ contain the independent time series, while their dependencies are captured in the weight matrix $\hat{A}$. On the other hand, the i-th column of $\mathbf{X}$ is constructed as the product of $\hat{\mathbf{S}}$ and the i-th column of $\hat{A}$. Therefore, similar columns in $\hat{A}$ correspond to similar trends in the original time series. Consequently, detecting similar columns in $\hat{A}$ implies detecting similar columns in $\mathbf{X}$. Hence, for clustering the time series matrix $\mathbf{X}$, we can apply a clustering algorithm directly to $\hat{A}$.

ICA is frequently used as a pre-processing step for clustering time series data. For instance, we refer to [18, 24]. We employ the CICA algorithm as a pre-processing step for time series clustering. Our objective is to cluster 26 countries—Australia, Austria, Belgium, Burkina Faso, Cameroon, Canada, Chile, China, Denmark, Finland, France, Germany, Ghana, India, Indonesia, Italy, Japan, South Korea, Malaysia, Pakistan, Qatar, Saudi Arabia, Singapore, Sweden, Turkey, and the United States—based on the patterns observed in their GDP per capita time series over the past 46 years. The data were sourced from the World Bank's repository (`https://data.worldbank.org`) for the period from 1975 to 2020.

First, the GDP per capita time series were standardized. Then, the CICA algorithm was applied to the standardized GDP per capita matrix. The resulting mixing coefficient matrix was then used as input to the PAM clustering algorithm, as introduced by [25]. The first step involved determining the appropriate number of clusters based on the Silhouette criterion, using the `NbClust R` package. This algorithm suggests 8 clusters for the considered data. The clustering results and the trend plots for all

Table 1: Average Amari errors for $d = 2$ and $n = 250$ based on 100 replications for each distribution (a) through (r). The smallest (best) value in each row is shown in bold.

Distribution	RLICA	GDCICA$_g$	CICA
a	0.5317	**0.5245**	0.5368
b	0.5426	0.5441	**0.5237**
c	0.5418	**0.5035**	0.5210
d	0.5135	0.5152	**0.5061**
e	0.5602	**0.5471**	0.5581
f	0.5115	0.5170	**0.4932**
g	**0.5426**	0.5828	0.5493
h	0.5157	**0.5147**	0.5332
i	0.5557	0.5713	**0.5428**
j	**0.5000**	0.5036	0.5071
k	**0.5020**	0.5032	0.5302
l	0.5232	0.5132	**0.5083**
m	0.5484	**0.5022**	0.5691
n	**0.5020**	0.5231	0.5289
o	0.5284	**0.5211**	0.5298
p	0.5465	**0.5388**	0.5530
q	**0.5024**	0.5275	0.5081
r	0.4946	0.5039	**0.4745**
mean	0.5257	**0.5253**	0.5263
rand	0.5303	0.5389	**0.5157**

Table 2: Average Amari errors for $d = 2$ and $n = 250$ for each class of distributions. The smallest (best) value in each row is shown in bold.

Class	RLICA	GDCICA$_g$	CICA
unimodal	0.5312	0.5299	**0.5225**
multimodal	0.5298	**0.5289**	0.5343
transitional	**0.5128**	0.5144	0.5243
symmetric	0.5304	**0.5290**	0.5304
nonsymmetric	**0.5184**	0.5195	0.5199
positive kurtosis	0.5319	0.5296	**0.5236**
negative kurtosis	**0.5233**	0.5237	0.5273

clusters are illustrated in Figure 2. This figure indicates that the pre-processing technique employed by the proposed algorithm for clustering countries is highly effective. This is evidenced by the clear observation that countries within the same cluster exhibit similar trends in their standardized GDP per capita.

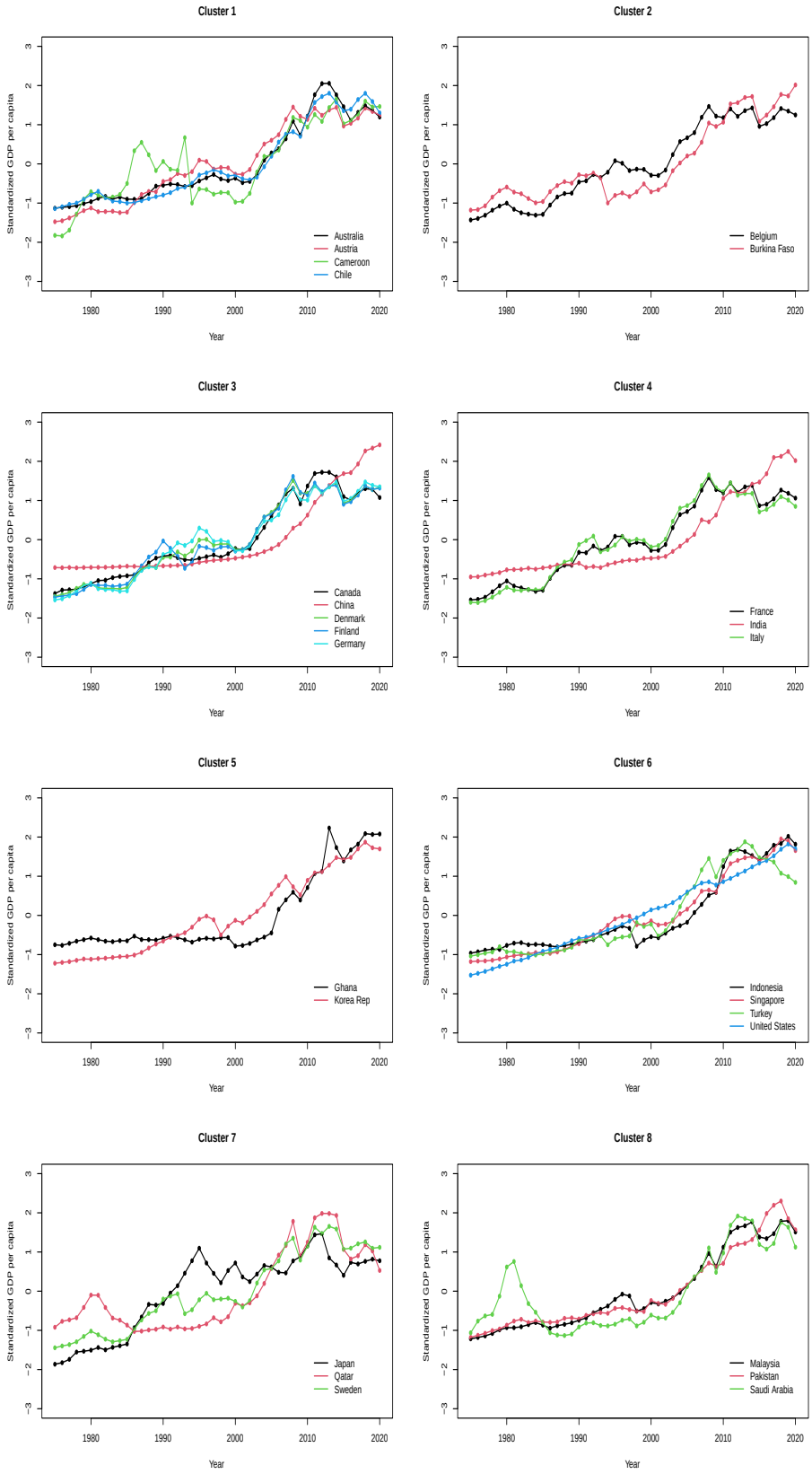

Figure 2: Trend plots of standardized GDP per capita time series in 8 clusters.

Acknowledgments

The author would like to thank three anonymous reviewers for their helpful comments and suggestions, which led to the improved presentation of this article significantly.

References

[1] Amari, S., Cichocki, A., and Yang, H., "A new learning algorithm for blind source separation", Advances in Neural Information Processing Systems, Vol. 8, pp. 757-763, 1996.

[2] Asadi, F., Torabi, H., and Nadeb, H., "An algorithm for independent component analysis using a general class of copula-based dependence criteria", Journal of Mahani Mathematical Research, Vol. 14, No. 1, pp. 527-550, 2025.

[3] Asadi, F., Torabi, H., and Nadeb, H., "A new approach for independent component analysis and its application for clustering the economic data", International Journal of Computational Economics and Econometrics, Vol. 15, No. 1-2, pp. 147-171, 2025.

[4] Asadi, F., Torabi, H., and Nadeb, H., "New algorithms for independent component analysis based on a general class of dependence criteria", International Journal of Computational Economics and Econometrics, Vol. 20, No. 1, pp. 8, 2025.

[5] Bach, F.R., and Jordan, M.I., "Kernel independent component analysis", Journal of Machine Learning Research, Vol. 3, pp. 1-48, 2002.

[6] Bingham, E., Hyvärinen, A., "A fast fixed-point algorithm for independent component analysis of complex valued signals", International Journal of Neural Systems, Vol. 10, No. 1, pp. 1-8, 2000.

[7] Calhoun, V.D., and Adali, T., "Multisubject independent component analysis of fMRI: a decade of intrinsic networks, default mode, and neurodiagnostic discovery", IEEE Reviews In Biomedical Engineering, Vol. 5, pp. 60-73, 2012.

[8] Cardoso, J.F., "High-order contrasts for independent component analysis", Neural Computation, Vol. 11, No. 1, pp. 157-192, 1999.

[9] Chatterjee, S., "A new coefficient of correlation", Journal of the American Statistical Association, Vol. 116, No. 536, pp. 2009-2022, 2020.

[10] Comon, P., "Independent component analysis, a new concept?", Signal Processing, Vol. 36, No. 6, pp. 287-314, 1994.

[11] Comon, P., and Jutten, C., "Handbook of Blind Source Separation: Independent component analysis and applications", Academic Press, United States, 2010.

[12] Gaeta, M., "Source separation without prior knowledge: The maximum likelihood solution", In Proceeding. EUSIPCO'90, pp. 621–624, 1990.

[13] Hyvärinen A., Karhunen, J., and Oja, E., "Independent component analysis and blind source separation", New York, NY, USA: Wiley, pp. 20-60, 2001.

[14] Hyvärinen, A., "Fast and robust fixed-point algorithms for independent component analysis", IEEE Transactions on Neural Networks, Vol. 10, No. 3, pp. 626-634, 1999.

[15] Langlois, D., Chartier, S., and Gosselin, D., "An introduction to independent component analysis: Infomax and FASTICA algorithms", Tutorials in Quantitative Methods for Psychology, Vol. 6, No. 1, pp. 31-38, 2010.

[16] Learned-Miller, E.G., and Fisher III, J.W., "ICA using spacings estimates of entropy", Journal of Machine Learning Research, Vol. 4, pp. 1271-1295, 2003.

[17] Lee, T.W., Girolami, M., and Sejnowski, T.J., "Independent component analysis using an extended infomax algorithm for mixed subgaussian and supergaussian sources", Neural Computation, Vol. 11, No. 2, pp. 417-441, 1999.

[18] Nascimento, M., Silva, F.F.E., Safadi, T., Nascimento, A.C.C., Ferreira, T.E.M., Barroso, L.M.A., ..., and Serão, N.V.L., "Independent component analysis (ICA) based-clustering of temporal RNA-seq data", PLoS ONE, Vol. 12, No. 7, pp. 1-12, 2017.

[19] Nordhausen, K., and Oja, H., "Independent component analysis: statistical perspective", Wiley Interdisciplinary Reviews: Computational Statistics, Vol. 10, No. 5, pp. 757-763, 2018.

[20] Pearlmutter, B.A, and Parra, L.C., "Maximum likelihood blind source separation: A context-sensitive generalization of ICA", In Advances in Neural Information Processing Systems, Vol. 9, pp. 613-619, 1996.

[21] Pfister, N., Weichwald, S., Bühlmann, P., and Schölkopf, B., "Robustifying independent component analysis by adjusting for group-wise stationary noise", Journal of Machine Learning Research, Vol. 20, No. 147, pp. 1-50, 2019.

[22] Pham, D.T, Garrat, P., and Jutten, C., "Separation of a mixture of independent sources through a maximum likelihood approach", In European Signal Processing Conference, 771–774, 1992.

[23] Rahmanishamsi, J., and Dolati, A., "Rank based least-squares independent component analysis", Journal of Statistical Research of Iran, Vol. 14, No. 2, pp. 247-266, 2017.

[24] Rahmanishamsi, J., Dolati, A., and Aghabozorgi, M.R., "A copula based ICA algorithm and its application to time series clustering", Journal of Classification, Vol. 35, No. 2, pp. 230-249, 2018.

[25] Kaufman, L., and Rousseeuw, P.J., "Clustering by means of medoids", In Proceedings of the Statistical Data Analysis Based on the L1 Norm Conference, Vol. 31, pp. 405-416. Neuchatel, Switzerland, 1987.

[26] Stone, J.V., "Independent Component Analysis: A Tutorial Introduction", The MIT Press, USA, 2004.

[27] Tharwat, A., "Independent component analysis: An introduction", Applied Computing and Informatics, Vol. 17, No. 2, pp. 222-249, 2021.