"REVIEW OF INDEPENDENT COMPONENT ANALYSIS ALGORITHMS AND ITS APPLICATION"

NARESH NIMJE

Department of Electronices & Tele-Communication Engineering

Dr. Rajendra Gode College of Engineering & Technology, Amravati, India

ABSTRACT: *The independent component analysis of a random vector consists of finding for a linear transformation that minimizes the statistical dependence between its components. In order to define suitable search criteria, the expansion of mutual information is utilized as a function of cumulates of increasing orders. An efficient algorithm is proposed, which allows the computation of the ICA of a data matrix within a polynomial time. The concept of lCA may actually be seen as an extension of the principal component analysis, which can only impose independence up to the second order and, consequently, defines directions that are orthogonal. Potential applications of ICA include data analysis and compression, Bayesian detection, localization of sources, and blind identification and deconvolution.*

Keywords: - independent component analysis, data analysis and compression, Bayesian detection, localization of sources, and blind identification and deconvolution.

.

1. INTRODUCTION

Nowadays, performing statistical analysis is only a few clicks away. However, before anyone carries out the desired analysis, some assumptions must be met. Of all the assumptions required, one of the most frequently encountered is about the normality of the distribution

ICA is very closely related to the method called blind source separation (BSS) or blind signal separation. A "source" means here an original signal, i.e. independent component, like the speaker in a cocktail party problem. "Blind" means that we no very little, if anything, on the mixing matrix, and make little assumptions on the source signals. ICA is one method, perhaps the most widely used, for performing blind source separation. An example application of ICA is to solve the "Cocktail Party Problem." This problem states we have multiple people talking at a party, but alas the party is very crowded. We try to record the happenings of the party with a microphone, but so many people are talking that no one is understood. However, if we record the party with multiple microphones, can we reconstruct individual voices. Ignoring delays due to distance to the microphones, ICA can solve this problem.

This paper presents an introduction to independent component analysis (ICA). unlike principal component analysis, which is based on the assumptions of uncorrelatedness and normality, ICA is rooted in the assumption of statistical independence.

2. ALGORITHM OF INDEPENDENT COMPONENT ANALYSIS

In this section discuss the some basic independent component analysis algorithm.

[1] FastICA FOR ONE UNIT:

To begin with, we shall show the one-unit version of FastICA. By a "unit" we refer to a computational unit, eventually an artificial neuron, having a weight vector $\mathbf{w}$ that the neuron is able to update by a learning rule[1]. The FastICA learning rule finds a direction, i.e. a unit vector $\mathbf{w}$ such that the projection $\mathbf{w}^T\mathbf{x}$ maximizes non-gaussianity. Non-gaussianity is here measured by the approximation of negentropy $J(\mathbf{w}^T\mathbf{x})$. the variance of $\mathbf{w}^T\mathbf{x}$ must here be constrained to unity; for whitened data this is equivalent to constraining the norm of $\mathbf{w}$ to be unity**.** The FastICA is based on a fixed-point iteration scheme for finding a maximum of the nongaussianity of $\mathbf{w}^T\mathbf{x}$, as measure. It can be also derived as an approximative Newton iteration[1]. Denote by *g* the derivative of the non quadratic function *G*.

The basic form of the FastICA algorithm is as follows:

1. Choose an initial (e.g. random) weight vector $\mathbf{w}$

2. Let $\mathbf{w}^+ = E\{\mathbf{x}g(\mathbf{w}^T\mathbf{x})\} - E\{g'(\mathbf{w}^T\mathbf{x})\}\mathbf{w}$
3. Let $\mathbf{w} = \mathbf{w}^+ / \|\mathbf{w}^+\|$
4. If not converged, go back to 2.

convergence means that the old and new values of $\mathbf{w}$ point in the same direction, i.e. their dot-product is (almost) equal to 1[1][2]. It is not necessary that the vector converges to a single point, since $\mathbf{w}$ and $-\mathbf{w}$ define the same direction. This is again because the independent components can be defined only up to a multiplicative sign. Note also that it is here assumed that the data is prewhitened.

[2] FastICA FOR SEVERAL UNITS:

The one-unit algorithm of the preceding subsection estimates just one of the independent components, or one projection pursuit direction. To estimate several independent components, we need to run the one-unit FastICA algorithm using several units (e.g. neurons) with weight vectors $\mathbf{w}_1, ..., \mathbf{w}_n$. To prevent different vectors from converging to the same maxima we must decorrelate the outputs $\mathbf{w}_1^T\mathbf{x}, ..., \mathbf{w}_n^T\mathbf{x}$ after every iteration[1][2]. A simple way of achieving decorrelation is a deflation scheme based on a Gram-Schmidt-like decorrelation. This means that we estimate the independent components one by one. When we have estimated *p* independent components, or *p* vectors $\mathbf{w}_1, ..., \mathbf{w}_p$, we run the one-unit fixed-point algorithm for $\mathbf{w}_{p+1}$, and after every iteration step subtract from $\mathbf{w}_{p+1}$ the ``projections'' $\mathbf{w}_{p+1}^T\mathbf{w}_j\mathbf{w}_j, j = 1, ..., p$ of the previously estimated *p* vectors, and then renormalize $\mathbf{w}_{p+1}$:

1. Let $\mathbf{w}_{p+1} = \mathbf{w}_{p+1} - \sum_{j=1}^{p} \mathbf{w}_{p+1}^T\mathbf{w}_j\mathbf{w}_j$
2. Let $\mathbf{w}_{p+1} = \mathbf{w}_{p+1} / \sqrt{\mathbf{w}_{p+1}^T\mathbf{w}_{p+1}}$

In certain applications, however, it may be desired to use a symmetric decorrelation, in which no vectors are ``privileged'' over others. This can be accomplished, e.g., by the classical method involving matrix square roots,

$$\text{Let } \mathbf{W} = (\mathbf{W}\mathbf{W}^T)^{-1/2}\mathbf{W}$$

where $\mathbf{W}$ is the matrix $(\mathbf{w}_1, ..., \mathbf{w}_n)^T$ of the vectors, and the inverse square root $(\mathbf{W}\mathbf{W}^T)^{-1/2}$ is obtained from the eigen value decomposition of $\mathbf{W}\mathbf{W}^T = \mathbf{F}\Lambda\mathbf{F}^T$ as $(\mathbf{W}\mathbf{W}^T)^{-1/2} = \mathbf{F}\Lambda^{-1/2}\mathbf{F}^T$. A simpler alternative is the following iterative algorithm

1. Let $\mathbf{W} = \mathbf{W} / \sqrt{\|\mathbf{W}\mathbf{W}^T\|}$

Repeat 2. until convergence:

2. Let $\mathbf{W} = \frac{3}{2}\mathbf{W} - \frac{1}{2}\mathbf{W}\mathbf{W}^T\mathbf{W}$

[3] FastICA AND MAXIMUM LIKELIHOOD :

This algorithm gives new version of FastICA that shows explicitly the connection to the well-known infomax or maximum likelihood algorithm introduced.. If we express FastICA using the intermediate formula , and write it in matrix form, we show that FastICA takes the following form:[3][4]

$$\mathbf{W}^+ = \mathbf{W} + \Gamma[\text{diag}(-\beta_i) + E\{g(\mathbf{y})\mathbf{y}^T\}]\mathbf{W}.$$

Where $\mathbf{y} = \mathbf{W}\mathbf{x}$, $\beta_i = E\{y_i g(y_i)\}$, and $\Gamma = \text{diag}(1/(\beta_i - E\{g'(y_i)\}))$. The matrix $\mathbf{W}$ needs to be orthogonalized after every step. In this matrix version, it is natural to orthogonalize $\mathbf{W}$ symmetrically.

The above version of FastICA could be compared with the stochastic gradient method for maximizing likelihood

$$\mathbf{W}^+ = \mathbf{W} + \mu[\mathbf{I} + g(\mathbf{y})\mathbf{y}^T]\mathbf{W}.$$

where μ is the learning rate, not necessarily constant in time. we show that FastICA can be considered as a fixed-point algorithm for maximum likelihood

FastICA is that it can estimate both sub- and super-gaussian independent components, which is in contrast to ordinary ML algorithms, which only work for a given class of distributions

[4] JADE (ICA): Joint Approximation Diagonalisation of Eigenmatrices (JADE) is an algorithm for independent component analysis that separates observed mixed signals into latent source signals by exploiting fourth order moments. The fourth order moments are a measure of non-Gaussianity, which is used as a proxy for defining independence between the source signal[5]s. The motivation for this measure is that Gaussian distributions possess zero excess kurtosis, and with non-Gaussianity being a canonical assumption of ICA, JADE seeks an orthogonal rotation of the observed mixed vectors to estimate source vectors which possess high values of excess kurtosis.

Let $\mathbf{X} = (x_{ij}) \in \mathbb{R}^{m\times n}$ denote an observed data matrix whose n columns correspond to observations of m-variate mixed vectors. It is assumed that $\mathbf{X}$ is prewhitenend, that is, its rows have a sample mean equaling zero and a sample covariance is the $m \times m$ dimensional identity matrix, that is,

$$\frac{1}{n}\sum_{j=1}^{n} x_{ij} = 0 \quad \text{and} \quad \frac{1}{n}\mathbf{X}\mathbf{X}' = \mathbf{I}_m$$

Applying JADE to $\mathbf{X}$ entails

1. computing fourth-order cumulants of $\mathbf{X}$ and then
2. optimizing a contrast function to obtain a $m \times m$ rotation matrix O

to estimate the source components given by the rows of the $m \times n$ dimensional matrix $\mathbf{Z} := \mathbf{O}^{-1}\mathbf{X}$.

[5] KernelICA : Kernel independent component analysis (Kernel ICA) is an efficient algorithm for independent component analysis which estimates estimation of the ICA data model. . In FastICA, convergence speed is optimized by the choice of the matrices Γ and $\mathbf{diag}(-\beta_i)$. Another advantage of source components by optimizing a generalized variance contrast function[6], which is based on representations in a reproducing kernel Hilbert space. Those contrast functions use the notion of mutual information as a measure of statistical independence.

3. APPLICATION OF INDEPENDENT COMPONENT ANALYSIS

1. Independent Component Analysis used for Decision Tree: Decision trees are typically used as computationally efficient representations in classification or regression. For example binary trees allows one to reach a leaf node in log2 n decisions when the tree contains n nodes in total. If the decisions are easy to compute ,this may result in a very light computational process[7]. Typical problems involving decision trees deal with multivariate data where the components are usually attribute values with a clear interpretation. Decisions can be simply implemented as threshold values on one of the attributes. An example is an attribute height(k) which is the height of person k. A decision could be a threshold height (k) > 175cm, which would split the data in two classes based on height.

One of the key properties of Independent Component Analysis is the ability to find linear combinations of multivariate data that have certain information-theoretic properties[7]. Often these linear combinations represent underlying sources, which cannot be directly observed. If the data contains such hidden sources, decision based on thresholding observed data components may not be very effective. In, single-component ICA was used to implement binary decisions in a decision tree. The key idea is that the independent components have more structure than the observed components, and therefore can be expected to be better candidates for linear threshold decisions. Since only single ICA component needs to be computed at each tree node, computational complexity resulting from a large number of ICA components can be avoided.

[2] ICA for Text Mining: Independent component analysis (ICA) was originally developed for signal processing applications.Recently it has been found out that ICA is a powerful tool for analyzing text document data as well, if the text documents are presented in a suitable numerical form.

This opens up new possibilities for automatic analysis of large textual data bases: finding the topics of documents and grouping them accordingly[8].

First approaches of using ICA in the context of text data considered the data static. In our recent study, we concentrated on text data whose topic changes over time. Examples of dynamically evolving text are chat line discussions or newsgroup documents. The dynamical text stream can be seen as a time series, and methods of time series processing may be used to extract the underlying characteristics here the topics of the data[8].

[3] Face Recognition using Independent Component Analysis: A number of current face recognition algorithms use face representations found by unsupervised statistical methods[9]. Typically these methods find a set of basis images and represent faces as a linear combination of those images. Principal component analysis (PCA) is a popular example of such methods. The basis images found by PCA depend only on pair wise relationships between pixels in the image database. In a task such as face recognition, in which important information may be contained in the high-order relationships among pixels, it seems reasonable to expect that better basis images may be found by methods sensitive to these high-order statistics. Independent component analysis (ICA), a generalization of PCA, is one such method[9]. We used a version of ICA derived from the principle of optimal information transfer through sigmoidal neurons. ICA was performed on face images in the FERET database under two different architectures, one which treated the images as random variables and the pixels as outcomes, and a second which treated the pixels as random variables and the images as outcomes. The first architecture found spatially local basis images for the faces. The second architecture produced a factorial face code. Both ICA representations were superior to representations based on PCA for recognizing faces across days and changes in expression. A classifier that combined the two ICA representations gave the best performance[9].

[4] Mobile Phone Communications Using Independent Component Analysis:

In commercial cellular networks, like the systems based on direct sequence code division multiple accesses (DSCDMA), many types of interferences can appear, starting from multi-user interference inside each sector in a cell to interoperator interference[11]. Also unintentional jamming can be present due to co-existing systems at the same band, whereas intentional jamming arises mainly in military applications. Independent Component Analysis (ICA) use as an advanced pre-processing tool for blind suppression of interfering signals in direct sequence spread spectrum communication systems utilizing antenna arrays. The role of ICA is to provide an interference-mitigated signal to the conventional

detection. Several ICA algorithms exist for performing Blind Source Separation (BSS)[11]. ICA has been used to extract interference signals, but very less literature is available on the performance, that is, how does it behave in communication environment? This needs an evaluation of its performance in communication environment. This chapter evaluates the performance of some major ICA algorithms like Bell and Sejnowski’s infomax algorithm, Cardoso’s Joint Approximate Diagonalization of Eigen matrices (JADE), Pearson-ICA, and Comon’s algorithm in a communication blind source separation problem. Independent signals representing Sub-Gaussian, Super-Gaussian, and mix users, are generated and then mixed linearly to simulate communication signals. Separation performance of ICA algorithms is measured by performance index.

[5] Predicting Stock Market Prices Using Independent Component Analysis: In developing a stock price forecasting model, the first step is usually feature extraction. Nonlinear independent component analysis (NLICA) is a novel feature extraction technique to find independent sources given only observed data that are mixtures of the unknown sources, without prior knowledge of the mixing mechanisms[10]. It assumes that the observed mixtures are the nonlinear combination of latent source signals[10]. This study propose a stock price forecasting model which first uses NLICA as preprocessing to extract features from forecasting variables. The features, called independent components (ICs), are served as the inputs of support vector regression (SVR) to build the prediction model. Experimental results on Nikkei 225 closing cash index show that the proposed method can produce the best prediction performance compared to the SVR models that use linear ICA, principal component analysis (PCA) and kernel PCA as feature extraction, and the single SVR model without feature extraction.

[6] Removing Artifacts, Such As Eye Blinks, From EEG Data Using Independent Component Analysis:

Independent Component Analysis is a powerful tool for eliminating several important types of non-brain artifacts from EEG data[12]. EEGLAB allows the user to reject many such artifacts in an efficient and user-friendly manner. The quality of the data is critical for obtaining a good ICA decomposition. ICA can separate out certain types of artifacts -- only those associated with fixed scalp-amp projections. These include eye movements and eye blinks, temporal muscle activity and line noise[12]. ICA may not be used to efficiently reject other types of artifacts -- those associated with a series of one-of-a-kind scalp maps.

to be removed by the user before performing ICA decomposition[12].

4. CONCLUSION

This paper surveyed contrast functions and algorithms for ICA. ICA is a general concept with a wide range of applications in neural computing, signal processing, and statistics. ICA gives a representation, transformation, of multidimensional data that seems to be well suited for subsequent information processing.

5. REFERENCES

[1] Hyyarinen A. Oja. E. (2000), "Independent Component analysis: Algorithms and Application", Neural Network 13(4-5): 411-430.

[2] Aapo Hyvärinen, "Survey on Independent Component Analysis", Neural Computing Surveys 2, 94-128, 1999.

[3] Petr Tichavský, Senior Member, IEEE, Zbynˇek Koldovský, Member, IEEE, and Erkki Oja, Fellow, IEEE, "Performance Analysis of the FastICA Algorithm and Cramér–Rao Bounds for Linear Independent Component Analysis*",* IEEE TRANSACTIONS ON SIGNAL PROCESSING, VOL. 54, NO. 4, APRIL 2006

[4] Aapo Hyvärinen," Independent component analysis: recent advance ", Phil. Trans. R. Soc. A 2013 **371**, 20110534.

[5] D.N. Rutledge, D. Jouan-Rimbaud Bouveresse Corrigendum to "Independent Components Analysis with the JADE algorithm"**,** TrAC Trends in Analytical Chemistry, Volume 67, April 2015, Page 220.

[6]Francis R. Bach, Michael I. Jordan, "Kernel Independent Component Analysis", Journal of Machine Learning Research 3 (2002) 1-48.

[7] P. Pajunen and M. Girolami. Implementing Decisions in Binary Decision Trees using Independent Component Analysis. In Proceedings of the 2nd Int. Workshop on Independent Component Analysis and Blind Source Separation, pages 483{487, Espoo, Finland, June 2000.

For example, if the subject were to scratch their EEG cap for several seconds, the result would be a long series of slightly different scalp maps associated with the channel and wire movements, etc. Therefore, such types of "non-stereotyped" or "paroxysmal" noise need

[8] Bingham, E. Topic identi_cation in dynamical text by extracting minimum complexity time components. In: Proc. 3rd International Conference on Independent Component Analysis and Blind Signal Separation (ICA2001), December 9{13, 2001, San Diego, CA, USA, pp. 546{551.

[9] marian stewart bartlett, javier r. movellan, terrence j. sejnowski," face recognition by independent component analysis" IEEE Trans Neural Netw. 2002; 13(6): 1450–1464.

[10] Chi-Jie Lu, Wu, Jui-Yu; Cheng-Ruei Fan; Chih-Chou Chiu," Forecasting Stock Price Using Nonlinear Independent Component Analysis And Support Vector Regression, Industrial Engineering and Engineering Management, 2009. IEEM 2009. IEEE International Conference on 8-11 Dec. 2009.

[11] Sargam Parmar, Bhuvan Unhelkar, "Independent Component Analysis Algorithms in wireless communication systems", Handbook of Research in Mobile Business, Second Edition: Technical, Methodological and Social Perspectives.

[12] QUICK TUTORIAL ON REJECTING DATA ARTIFACTS WITH EEGLAB.

5. AUTHOR PROFILE

Naresh Nimje Currently pursing Bachelor of Engineering in Electronics and Tele-Communication from Dr. Rajendra Gode Institute of Technology & Research, Amravati, India.