IE 506: Machine Learning - Principles and Techniques **Jan-May 2023**

End-term Project Report : Constrained Independent Component Analysis

Team Name: ASPIRE *Team Members: 21i190010,22D1143*

Abstract

In this paper, the author proposes a novel ICA algorithm called CICA that addresses the permutation problem by using a constrained optimization approach with a unit variance constraint on the demixing matrix. The algorithm uses an augmented Lagrangian method to solve the constrained optimization problem and includes a normalization step for the demixing matrix. Experimental results show that CICA can accurately recover independent components from mixed signals, outperforming other ICA algorithms in terms of separation accuracy and computational efficiency

1 Introduction

Here the author introduces a novel algorithm for extracting independent sources from mixed signals in the presence of constraints. Independent component analysis (ICA) is a popular technique for blind source separation in various fields such as signal processing, neuroscience, and image analysis. However, traditional ICA methods may not perform well when sources are highly correlated or have similar distributions.

To overcome these limitations, the proposed algorithm incorporates constraints on the demixing matrix to ensure the independence of the extracted sources. Specifically, the algorithm formulates the problem as a constrained optimization problem and solves it using an augmented Lagrangian method. The authors demonstrate the effectiveness of the proposed algorithm on both synthetic and real-world datasets.

Overall, this paper provides a valuable contribution to the field of blind source separation and has potential applications in various domains such as speech recognition, medical signal processing, and image analysis.

Structure of the project report: We present a literature review and describe the proposed approach in detail, including the methodology and algorithm used in sections 2 and 3 respectively. In section 4 we present the experimental results of applying the proposed approach to real-world datasets, including performance metrics and analysis. we disuss in section 5,6 the implications of the results, limitations of the proposed approach, and future directions of research . In conclusion, section 7 summarizes the paper's main contribution and last References.

2 Literature Survey

1. We studied the Non-negative matrix factorization technique known as the NMF technique.[5] So Non-negative matrix factorization (NMF) is a mathematical technique used to extract meaningful information from high-dimensional datasets. It factorizes a non-negative matrix into two non-negative matrices that, when multiplied together, approximate the original matrix. Also it can be viewed as a special case of ICA with the constraint that the output signal space is non-negative.NMF has become a popular technique in many fields, such as image processing, bio-informatics, and natural language processing, as it can identify patterns and features in the data that may not be immediately apparent from the raw data. It is also useful for data compression, feature selection, and signal processing.

2.we studied Constrained ICA using Projection pursuit(CICA-PP) so Constrained Independent Component Analysis using Projection Pursuit (CICA-PP) is a statistical method for blindly separating mixed signals into their underlying independent components.[2] CICA-PP combines the principles of two established methods: Constrained Independent Component Analysis (CICA) and Projection Pursuit (PP). CICA aims to decompose a set of mixed signals into independent components by imposing constraints on the mixing matrix, while PP aims to find the directions in the signal space that maximize some predefined measures of interestingness.In CICA-PP, the mixing matrix is assumed to be unknown but constrained by some prior knowledge, and the goal is to find the independent components that have maximal projection onto a predefined set of directions. These directions are selected based on the predefined measures of interestingness, which could be related to higher-order statistical properties of the signal or some other criterion.CICA-PP can be applied in various fields, such as signal processing, image processing, and data mining, where the goal is to identify underlying sources from observed mixed signals.In CICA-PP, the mixing matrix is assumed to be unknown but constrained by some prior knowledge, and the goal is to find the independent components that have maximal projection onto a predefined set of directions. These directions are selected based on the predefined measures of interestingness, which could be related to higher-order statistical properties of the signal or some other criterion.CICA-PP can be applied in various fields, such as signal processing, image processing, and data mining, where the goal is to identify underlying sources from observed mixed signals.

3. A fast fixed-point algorithm for independent component analysis" by Aapo Hyv arinen (1999). [3] This paper proposed a fast fixed-point algorithm for independent component analysis, which can solve the optimization problem of ICA in a few iterations. The algorithm is based on an iterative scheme that updates the demixing matrix using a gradient descent method with a step size that is adaptively adjusted. The paper demonstrated the efficiency and accuracy of the algorithm in separating mixtures of speech signals, and compared it with other ICA algorithms.

3 Methods and Approaches

Lagrange Multiplier Methods - The general form of the constrained nonlinear optimization problem is given as follows:

$$minimize f(X)$$

$$\text{subject to} \quad \text{g(X)} \leq 0 \ \text{h(X)} = 0 \tag{1}$$

where f(X): objective function, g(X) inequality constraints and h(X) equality constraints.
Now the augmented lagrangian function for problem 1 is defined as:

$$L(X,\mu,\lambda) = f(X) + (1/2\gamma)\sum_{i=1}^{m}[[max(0,\bar{g}(X))]^2 - \mu_i^2] + \lambda h(X) + (1/2)\gamma||h(X)||^2 \tag{2}$$

Now the iterative equations for X, μ and λ are given :

$$\begin{aligned} X(k+1) &= X(k) - \Delta_x L(X(k),\mu(k),\lambda(k)) \\ \mu(k+1) &= \mu(k) + \gamma p(X(k)) = max(0,\bar{g}(X(k))) \\ \lambda(k+1) &= \lambda(k) + \gamma h(X(k)) \end{aligned} \tag{3}$$

3a The Ordering of independent components based on CICA [1] is given by

$$\text{minimize M(W)}$$
$$\text{subject to g(W)} \leq\!=0,\ \text{g(W)} = [\text{g}_i(W)...g_{m-1}(W)]^T$$

where g(W) are inequality constraints,

$$g_i(W) = I(u_{i+1}) - I(u_i) \tag{4}$$

The augmented Lagrangian function is defined as:

$$L(W, \mu) = M(W) + (1/2\gamma)\sum_{i=1}^{m}[[max(0, \bar{g}(X))]^2 - \mu_i^2] \tag{5}$$

$\text{w}_{ij} are changed by using the given iterative equation$

$$\Delta w_{ij} \propto \Delta_{w_{ij}} L(W(k), \mu(k)) = minM(W(k)) + (max0, g_{i-1}^{-}(W(k) - max0, \bar{g}_i(W(k))I^{'}(u_i(k))x_j \tag{6}$$

The iterative formula for determining each multiplier

$$\mu_i(k+1) = max0, \mu_i(k) + \gamma[I(\mu_{i+1}(k)) - I(\mu_i(k)] \tag{7}$$

The iterative procedure for W is given as:

$$\Delta W \propto \Delta_w L(W, \mu) = W^{-T} + \Psi(u)x^T \tag{8}$$

$$where\Psi(u) = \begin{bmatrix} \Phi_1(u_1) - \mu_1 I^{'}(\mu_1) \\ \Phi_2(u_2) + (\mu_1 - \mu_2 I^{'}(\mu_2) \\ . \\ . \\ . \\ \Phi_{M-1}(u_{M-1}) + (\mu_{M-2} - \mu_{M-1} I^{'}(\mu_{M-1}) \\ \Phi_m(u_M) + \mu_{M-1} I^{'}(\mu_M) \end{bmatrix} \tag{9}$$

and now we find the kurtosis value to order the signals

$$kurtosis : I_{kur}(u_i) = (E(\mu_i^4)/E((\mu_i^2)^2) - 3 \tag{10}$$

$$I^{'}_{kur}(u_i) = (4\mu_i^3/E((\mu_i^2)^2) - 4E(\mu_i^4)\mu_i/E((\mu_i^2)^3)$$

3b Normalization of Demixing Matrix

The mixing and demixing matrix's norm is ambiguous according to the ICA notion. The rows of the demixing matrix W can be normalized to produce a normalized demixing channel by incorporating a constraint term into the ICA energy function..
Now constraint ICA problem defines as:

$$minimizeM(W)$$

$$\text{subject to h(W)} = [\text{h}_1(W)...h_M(W]^T = 0 \tag{12}$$

where h(W) set of equality constraints.

$$h_i(w_i) = w_i^T w_i - 1(energyfunction) \tag{13}$$

The augmented Lagrangian function is defined as:

$$L(W, \lambda) = M(W) + \lambda^T[DiagWW^T - 1] + (1/2)\gamma||[DiagWW^T - 1]||^2 \tag{14}$$

The iterative formula for determining each multiplier

$$\lambda_i(k+1) = \lambda_i(k) + \gamma(w_i^T w_i - 1) \tag{15}$$

The iterative procedure for W is given as:

$$\Delta W \propto \Delta_w L(W, \lambda) = W^{-T} + \Phi(u)x^T + \Omega(W) \tag{16}$$

where $\Omega_i(w_i) = 2\lambda_i w_i^T$

let's suppose that c is the normalised source with unit variance.

$$S.t \qquad E[cc^T] = 1 \tag{17}$$

,

3.1 Work done before mid-term project review

•Reading and understanding of the research paper based on Constrained Independent Component analysis.
•Understanding the previous work based on independent component analysis and reading other research papers related to this.
•Understanding the FAST ICA algorithm [4] and building the code from the information given in the paper from scratch.

3.2 Work done after mid-term project review

• Implementation of the CICA algorithm on different mixing matrices and found kurtosis values for ordering.
• Comparison of the two experiments of CICA algorithm.
•Implementation of the complete code of the sub-process involved in FAST ICA and found the correct kurtosis values of signal for ordering the signal that was remaining at the time of mid-term.

4 Data set Details

We consider three signals:
• Non-Gaussian • Randomly Gaussian 1 • Randomly Gaussian 2
For our experiment, we consider two mixing matrices A and B which are given respectively

$$A = \begin{pmatrix} 0.5 & 1 & 0.2 \\ 1 & 0.5 & 0.4 \\ 0.5 & 0.8 & 1 \end{pmatrix}, B = \begin{pmatrix} 0.7 & 0.3 & 0.2 \\ 0.9 & 0.5 & 0.4 \\ 0.5 & 0.8 & 0.6 \end{pmatrix}$$

Now we performed two preprocessing steps here:
The first preprocessing step is Centering of the data, In which the mean is subtracted from each variable for centering the data. The second preprocessing step involves the whitening of the data which is done basically by using singular value decomposition of the covariance matrix of mixed signals.
The mixed signal is created by the linear combination of mixing matrix and source signals then we find the whitening matrix, and the whitening data is prepared by using the whitening matrix and mixed signals. The demixing matrix is obtained by using the CICA algorithm now we find the recovered signals

5 Experiments

For the first experiment, we basically used 3 independent source signals Non-Gaussian,Randomly Gaussian 1 signal, and Randomly Gaussian 2 and these signals were mixed by using a mixing matrix A we trained the signals for finding the demixing matrix by using the CICA algorithm and we separated the required independent component by using kurtosis values. We perform the 2^{nd} experiment by using the same signal and different mixing matrix B and we again trained the signal by using the same algorithm we find the different demixing matrix and we separate the required independent component by kurtosis values and finally, we compare both the results.
The above experiments are performed basically by using an optimization problem which includes the minimization of the augmented lagrangian function and required lagrangian multiplier, demixing matrix is updated by an iterative equation that we have shown in the earlier section 3.

6 Results

First of all, we plot the independent signals that we have described in the data set detail.

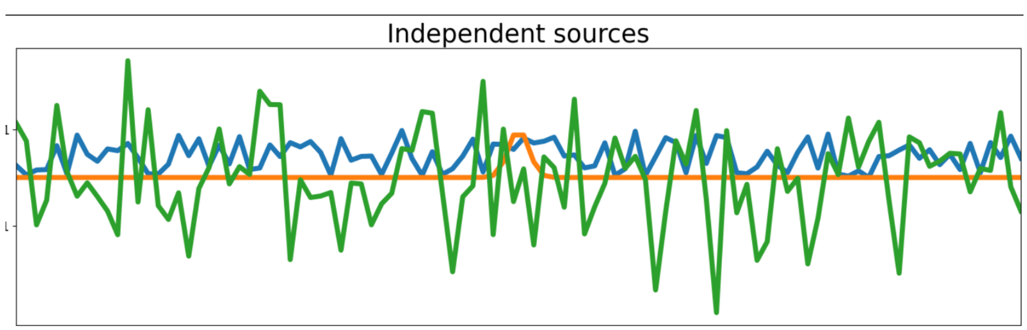

Figure 1: Independent signals

Now by using mixing matrix A and B the mixed signals are shown respectively.

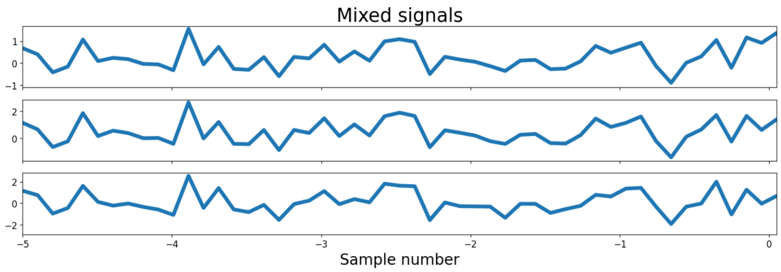

Figure 2: Mixed signal by using mixing matrix A

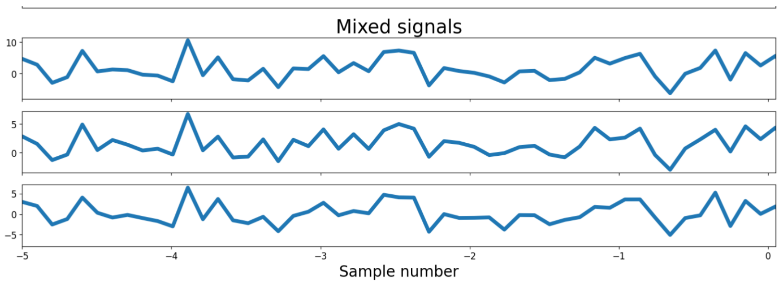

Figure 3: Mixed signal by using mixing matrix B

Finally, the recovered signals from both experiments are shown respectively.

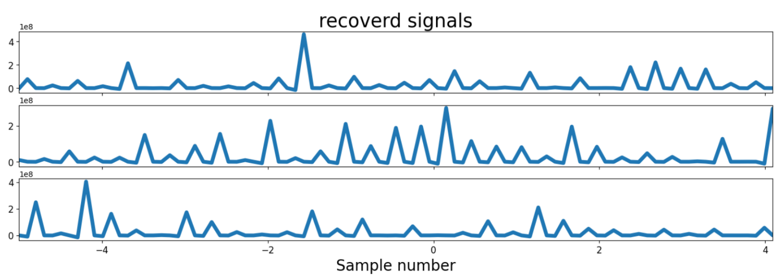

Figure 4: Recovered signal by using mixing matrix A

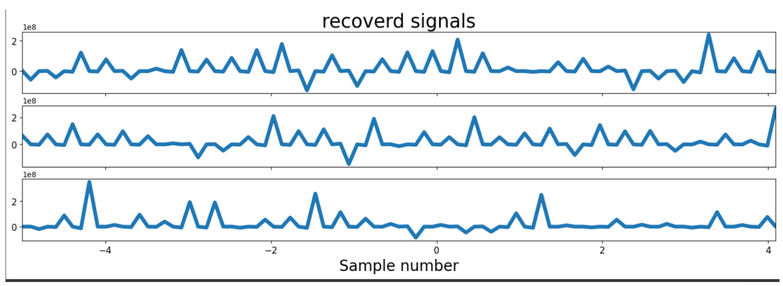

Figure 5: Recovered signal by using mixing matrix B

Now kurtosis values for the first and second experiments are given respectively and the ordering is done based on the kurtosis values $18.152861113446473, 5.204297240488279, 14.176862736964559$

$$2.7465396384861798, 3.687730735041881, 9.401504451456674$$

Hence for the first experiment, random non-Gaussian will come first and random Gaussian 2 comes second, and random Gaussian 1 comes last in order. Similarly, for the second experiment, random Gaussian 2 will come first and random Gaussian 1 comes second, and random Non-Gaussian comes last in order.As we can see from the graph the mixed signals somehow look similar but there is a slight difference in the scale.

7 Future Work

The paper focuses on the use of CICA for signal separation but it could explore the use of the algorithm for other applications, such as feature extraction. The proposed CICA algorithm can currently handle linear and non-linear constraints but it could explore the use of other types of constraints, such as convex constraints. Here CICA algorithm assumes that all the data is available at once but it could explore the development of online versions of the algorithm that can handle streaming data in a more efficient manner.

8 Conclusion

The problem addressed in this project is the task of independent component analysis, which involves decomposing a set of mixed signals into their independent components. For decomposing a set of mixed signals into their independent components we worked on the Fast ICA and CICA algorithm. The proposed Constrained Independent Component Analysis (CICA) algorithm can take into account linear and nonlinear constraints during the separation process and the algorithm is based on the Lagrangian optimization framework that enables the incorporation of constraints into the separation process. We have observed that the proposed constrained ICA algorithm can achieve better performance because FICA uses fixed point iteration but CICA is based on an iterative approach. By using the CICA algorithm the signals are recovered almost the same as the original signals but By FAST ICA the recovered signal have differences from the original signals.

References

[1] Huang, De-Shuang, and Jian-Xun M. A new constrained independent component analysis method. 2007.

[2] Aapo Hyvärinen. New approximations of differential entropy for independent component analysis and projection pursuit. 1997.

[3] Aapo Hyvärinen. Fast and robust fixed-point algorithms for independent component analysis. 1999.

[4] Stone and James V. Independent component analysis: an introduction. 2002.

[5] Zhang and Junying. Pattern expression nonnegative matrix factorization: algorithm and applications to blind source separation. 2008.