ELSEVIER

Neurocomputing 41 (2001) 145–152

NEUROCOMPUTING

www.elsevier.com/locate/neucom

Independent component ordering in ICA time series analysis☆

Yiu-ming Cheung*, Lei Xu

Department of Computer Science and Engineering, The Chinese University of Hong Kong, Shatin, N.T., Hong Kong

Received 2 May 2000; revised 30 October 2000; accepted 8 November 2000

Abstract

Independent component analysis (ICA) has provided a new tool to analyze time series, which also gives rise to a question — how to order independent components? In the literature, some methods (Back and Trappenberg, Proceedings of International Joint Conference on Neural Networks, Vol. 2, 1999, pp. 989–992; Hyvärinen, Neural Computing Surveys 2 (1999) 94; Back and Weigend, Int. J. Neural Systems 8(4) (1997) 473) have been suggested to decide the order based on each individual component without considering their interactions on the observed series. In this paper, we propose an alternative approach to order the components according to their joint contributions in data reconstruction, which naturally leads the component ordering to a typical combinatorial optimization problem, whereby the underlying optimum ordering can be found in an exhaustive way. To save computing costs, we also present a fast approximate search algorithm. The accompanying experiments have shown the outperformance of this new approach in comparison with an existing method.

Keywords: Independent component analysis; Independent component ordering; Data reconstruction

1. Introduction

In the past, time-series analysis had been extensively applied to system identification, adaptive signal prediction and so on. In the literature, *principal component*

☆The work described in this paper was fully supported by a grant from the Research Grant Council of the Hong Kong SAR (Project No. CUHK4383/99E).

* Corresponding author. Tel.: + 852-2609-8440; fax: + 852-2603-5302.

E-mail addresses: ymcheung@cse.cuhk.edu.hk (Y.M. Cheung), lxu@cse.cuhk.edu.hk (L. Xu).

0925-2312/01/$ - see front matter
PII: S0925-2312(00)00358-1

analysis (PCA) is a prevalent analysis tool, whose advantages are generally two-fold:

- The order of uncorrelated principal components is explicitly given in terms of their variances.
- The underlying structure of series can be revealed in using the first few principal components.

However, the PCA technique only uses second-order statistics information, which makes the principal components de-correlated but not really independent. Recently, it has been realized that *independent component analysis* (ICA), rather than PCA, is more appropriate for time-series analysis. The reason is that the extraction of independent components by ICA involves higher-order statistics, which makes the components reveal more useful information than principal components. The paper [3] has shown the advantage of ICA in analyzing a Japanese stock price in comparison with PCA. In the community, ICA has been providing a new tool for time-series analysis as shown in [7,8].

In analyzing time series by ICA, we often need to select several dominant components from the component pool, which can bring in at least two advantages:

1. *Robust Performance* — Those dominant components reveal the major information about the stochastic mechanism that gives rise to the observed series. Without considering trivial details, the model built based on these components has a robust performance in series forecasting;
2. *Reduction of Analysis Complexity* — To well understand the mechanism in the background of the series, sometimes we need to interpret the physical meaning of the independent components, which is a non-trivial task. Hence, studies concentrating on those dominant components make the analysis more easy.

One possible way to choose dominant components is by two separating steps as follows:

Step 1. List all the independent components in an appropriate order;

Step 2. Select the first few components in the order as dominant ones.

An empirical study in [5] has been done towards Step 2. In this paper, we will discuss the component ordering in Step 1 only.

In the literature, some methods have been suggested to determine the component order. For example, the components are sorted according to their non-Gaussianity [6]. Back and Trappenberg [2] suggested to select a subset of the components based on the mutual information between the observations and the individual components, which also provides another way to order the components. Furthermore, the paper [3] decides the component order according to the L_∞ norm of each individual component. In these existing methods, the component order is determined based on each individual component only.

However, we notice that the observed series are actually influenced by several components, whose individually decided optimum order in a sense is generally no longer optimum as a whole from the viewpoint of analyzing the observed series. It

may therefore be more helpful to consider the joint contributions of the components to the time series in performing ordering. In this paper, we propose an alternative ordering scheme to arrange the components according to their joint contributions in data reconstruction. Under the circumstances, the component ordering naturally leads to a typical combinatorial optimization problem, whereby the optimum order is decided under the reconstruction criterion.

Moreover, to avoid exhaustive search, we also propose a sub-optimum search algorithm called *Testing-and-Acceptance* (TnA), which determines the order by sequentially testing each independent component in the series' reconstruction before accepting it. Although this algorithm may introduce some artificial ordering, it can still give satisfactory results in most cases.

2. ICA in time-series analysis

2.1. ICA model

We suppose that the observed k time series $x_1(t), x_2(t), \ldots, x_k(t)$ are an instantaneous linear mixture of unknown mutually independent components $y_1(t), y_2(t), \ldots, y_k(t)$. The observed series can therefore be modeled in a matrix form

$$\boldsymbol{x}(t) = \boldsymbol{A}\boldsymbol{y}(t), \quad 1 \le t \le N, \tag{1}$$

where $\boldsymbol{y}(t) = [y_1(t), y_2(t), \ldots, y_k(t)]^{\mathrm{T}}$, $\boldsymbol{x}(t) = [x_1(t), x_2(t), \ldots, x_k(t)]^{\mathrm{T}}$, and $\boldsymbol{A}$ is an unknown $k \times k$ mixing matrix.

2.2. Extraction of independent components

Many existing ICA techniques, e.g. INFORMAX [4], MMI [1], etc., can recover those independent components $\boldsymbol{y}(t)$ in Eq. (1) up to an unknown constant and a permutation of indices through a de-mixing matrix $\boldsymbol{W}$ with

$$\hat{\boldsymbol{y}}(t) = \boldsymbol{W}\boldsymbol{x}(t) = \boldsymbol{W}\boldsymbol{A}\boldsymbol{y}(t) = \boldsymbol{P}\boldsymbol{\Lambda}\boldsymbol{y}(t), \quad 1 \le t \le N, \tag{2}$$

where $\hat{\boldsymbol{y}}(t) = [\hat{y}_1(t), \hat{y}_2(t), \ldots, \hat{y}_k(t)]^{\mathrm{T}}$ is an estimate of $\hat{\boldsymbol{y}}(t)$, $\boldsymbol{P}$ is a permutation matrix and $\boldsymbol{\Lambda}$ is a diagonal matrix. $\boldsymbol{W}$ can be tuned by one of those existing ICA algorithms. This paper chooses *Learned Parametric Mixture Based ICA algorithm* (LPM) [9,10] based upon the fact that it can separate any combination of super-Gaussian and sub-Gaussian signals. The details of LPM can be referred to in the relevant papers.

3. Independent component ordering under data reconstruction criterion

Given k independent components $\{\hat{y}_j(t)\}_{j=1}^{k}$ by Eq. (2), we determine a list L, which shows the components' order according to their subscripts in the positions of L.

3.1. Independent component ordering

Considering the ith time series $\{x_i(t)\}_{t=1}^{N}$, we denote the contribution of component $\hat{y}_j$ to the reconstruction of x_i as

$$\hat{u}_{ij}(t) = W_{ij}^{-1}\hat{y}_j(t), \quad 1 \le j \le k, \tag{3}$$

where W_{ij}^{-1} denotes the (i,j)th element of the matrix $\boldsymbol{W}^{-1}$. Thus, the reconstruction of x_i by using the first m independent components under a specific list L_i is given by

$$\hat{x}_{L_i}^{m}(t) = \sum_{r=1}^{m} \hat{u}_{iq(r)}(t), \quad \text{with } q(r)\text{: } r\text{th element of } L_i.$$

Moreover, we denote $Q(x_i,\hat{x}_{L_i}^{m})$ to be the corresponding reconstruction error under a given measure. Therefore, the cumulative data reconstruction error with $1 \le m \le k$ is given by

$$J_{L_i} = \sum_{m=1}^{k} J_{L_i}(m) = \sum_{m=1}^{k} Q(x_i,\hat{x}_{L_i}^{m}). \tag{4}$$

In this paper, we let Q be the *Relative Hamming Distance* (RHD) function based on the consideration that the trend of a time series may be mostly controlled by the underlying independent components. That is,

$$Q(x_i,\hat{x}_{L_i}^{m}) = RHD(x_i,\hat{x}_{L_i}^{m}) = \frac{1}{N-1}\sum_{t=1}^{N-1}[R_i(t) - \hat{R}_{L_i}^{m}(t)]^2, \tag{5}$$

where

$$R_i(t) = \text{sign}[x_i(t+1) - x_i(t)], \qquad \hat{R}_{L_i}^{m}(t) = \text{sign}[\hat{x}_{L_i}^{m}(t+1) - \hat{x}_{L_i}^{m}(t)], \tag{6}$$

and

$$\text{sign}(r) = \begin{cases} 1 & \text{if } r > 0, \\ 0 & \text{if } r = 0, \\ -1 & \text{otherwise.} \end{cases} \tag{7}$$

Hence, the optimum order list L_i^{*} under this Q-measure criterion is

$$L_i^{*} = \text{argmin}_{L_i} J_{L_i}.$$

3.2. Sub-optimum search approach

Since the exhaustive search for L_i^{*} out of all $k!$ possible values is quite time-consuming, we here propose a fast approximate search algorithm called *Testing-and-Acceptance* (TnA) to find a sub-optimum order of independent components.

The basic procedure of the TnA algorithm is given as follows: from the set of k independent components, we pick $\hat{y}_r$ as the last one in the ordering, which makes the

Table 1
The RHD reconstruction error between the USD–AUD series and the reconstructions by, respectively, using the first m independent components as listed in $L_1^*(\hat{L}_1^*)$ and L_1

m	$J_{L_1^\uparrow}(m) = J_{L_1^\uparrow}(m)$	$J_{L_1}(m)$
1	0.3735	0.5221
2	0.3483	0.3852
3	0.2565	0.2565
4	0.1350	0.1350
5	0.0963	0.0963
6	0	0

RHD error between x_i and the corresponding reconstruction from those $\{\hat{y}_j\}_{j=1,j\neq r}^k$ minimized. Then, we remove this independent component from the component set. Next, we repeat the same operations on the remaining $\{\hat{y}_j\}_{j=1,j\neq r}^k$, and select the second-last component, …, and so forth. As a result, we get the k independent components in an order. Note that this TnA algorithm only involves $k(k+1)/2-1$ reconstruction steps in comparison with the exhaustive search that requires $(k+1)!$ steps. Hence, TnA approach can save considerable computing costs when k is large. The specification of this algorithm is given as follows:

Step 1. Given a component subscript set $Z = \{j \mid 1 \leq j \leq k\}$, let $n = 1$, and the order list of independent components $L_i = (\,)$.

Step 2. For each $j \in Z$, let

$$v_{ij}(t) = \sum_{m \neq j,\, m \in Z} \hat{u}_{im}(t), \quad 1 \leq t \leq N \tag{8}$$

and select the subscript $\beta = \text{argmin}_{j \in Z}\, \text{RHD}(x_i, v_{ij})$ as the nth element of L_i. Then, let $Z^{\text{new}} = Z^{\text{old}} - \{\beta\}$.

Step 3. If $Z \neq \{\}$, let $n^{\text{new}} = n^{\text{old}} + 1$, and go to Step 2; otherwise go to Step 4.

Step 4. Let $\hat{L}_i^* = L_i^{-1}$, where $\hat{L}_i^*$ is an estimate of L_i^*, and L^{-1} denotes the inverse order of L, e.g., if $L = (3,2,4,1)$, $L^{-1} = (1,4,2,3)$. Then stop.

Note that the $\hat{L}_i^*$ obtained by the TnA arranges the k independent components in descending order.

4. Simulation results

We demonstrated the performance of our proposed approach on the foreign exchange rate series, presenting the results of the L_∞ norm method [3] as well for comparison.

4.1. Experimental data

In the experiments, we used six foreign exchange rates, which are *US Dollar* (USD) versus *Australian Dollar* (AUD), *French Franc*, *Swiss Franc*, *German Mark*,

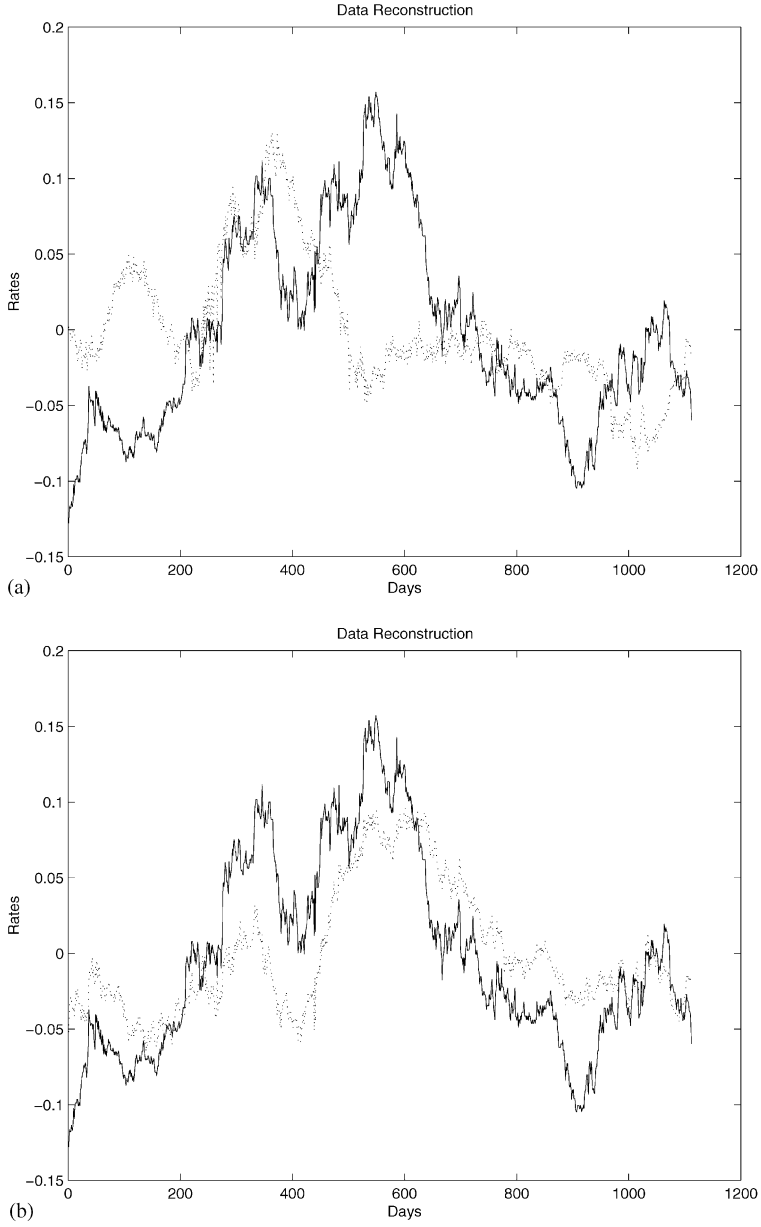

Fig. 1. The observed series of USD—AUD (——) and the reconstructed series (– – – –) by only using (a) $\{\hat{y}_1(t)\}_{t=1}^N$; (b) $\{\hat{y}_5(t)\}_{t=1}^N$. Note that the observed series has been normalized to zero mean in accordance with the requirement of the LPM algorithm.

British Pound, and *Japanese Yen*, respectively, from 26 November 1991 to 10 August 1995.

4.2. Experimental results

To save space, we chose the series of USD versus AUD as an example. The order given by the exhaustive search (ExS), TnA and L_∞ norm method, respectively, is

$$\begin{aligned}
&\text{ExS:} && L_1^* = (5,3,1,6,4,2),\\
&\text{TnA:} && \hat{L}_1^* = (5,3,1,6,4,2),\\
&\text{L}_\infty\text{norm:} && L_1 = (1,5,3,6,4,2).
\end{aligned}$$

In this example, TnA yielded the same component order as ExS but using much less computing costs. Table 1 shows the RHD reconstruction error between the USD–AUD series and the reconstructions by, respectively, using the first m independent components as listed in L_1^* ($\hat{L}_1^*$) and L_1. We found that the reconstruction error by using the component $\hat{y}_5$ only is 0.3735, which is much smaller than that by solely using $\hat{y}_1$. Fig. 1 shows the reconstructed series by using $\hat{y}_5$ and $\hat{y}_1$, respectively. It can be seen that the reconstruction from the sole component $\hat{y}_5$ has a similar variation of the USD–AUD series, but that from $\hat{y}_1$ has totally lost the trend of the observed series. In other words, the independent component $\hat{y}_5$ can dominate the trend of USD–AUD series, but $\hat{y}_1$ cannot. Similarly, we can also show that the component order $(5,3)$ is better than $(1,5)$. That is, our proposed method provides a better component order under the series reconstruction criterion.

5. Conclusion

We have proposed a technique to order the independent components under a data reconstruction criterion, which leads the component ordering to a typical combinatorial optimization problem. Hence, the optimum order can be determined through an exhaustive search in principle. To save searching costs, we have also presented a sub-optimum ordering algorithm TnA. The accompanying experiments have shown that our approach outperforms the L_∞ norm method in the reconstruction criterion.

References

[1] S. Amari, A. Cichocki, H. Yang, A new learning algorithm for blind signal separation, Adv. Neural Inform. Processing System 8 (1996) 757–763.
[2] A.D. Back, T.P. Trappenberg, Input variable selection using independent component analysis, Proceedings of International Joint Conference on Neural Networks, Vol. 2, 1999, pp. 989–992.
[3] A.D. Back, A.S. Weigend, A first application of independent component analysis to extracting structure from stock returns, Int. J. Neural Systems 8 (4) (1997) 473–484.

[4] A.J. Bell, T.J. Sejnowski, An information-maximization approach to blind separation and blind deconvolution, Neural Computation 7 (1995) 1129–1159.
[5] Y.M. Cheung, L. Xu, An empirical method to select dominant independent components in ICA time series analysis, Proceedings of 1999 International Joint Conference on Neural Networks (IJCNN'99), Vol. 6, Washington, DC, July 10–16, 1999, pp. 3883–3887.
[6] A. Hyvärinen, Survey on independent component analysis, Neural Computing Surveys 2 (1999) 94–128.
[7] K. Kiviluoto, E. Oja, Independent component analysis for parallel financial time series, Proceedings of the Fifth International Conference on Neural Information Processing (ICONIP'98), 1998, pp. 895–898.
[8] J. Moody, L. Wu, What is the 'true price'? — State space models for high frequency FX data, in: A.S. Weigend, Y. Abu-Mostafa, A.-Paul N. Refenes (Eds.), Decision Technologies for Financial Engineering — Proceedings of the Fourth International Conference on Neural Networks in the Capital Markets (NNCM'96), World Scientific, Singapore, 1997, pp. 346–358.
[9] L. Xu, C.C. Cheung, S.I. Amari, Learned parametric mixture based ICA algorithm, Neurocomputing 22 (1998) 69–80.
[10] L. Xu, C.C. Cheung, H.H. Yang, S.I. Amari, Independent component analysis by the information-theoretic approach with mixture of density, Proceedings of 1997 IEEE International Conference on Neural Networks (IJCNN'97), Vol. 3, 1997, pp. 1821–1826.