

独立分量分析研究

包云飞, 张伟川, 武飞, 张立

(成都理工大学管理科学学院, 成都 610059)

摘要:概述独立分量分析的基本理论,总结三个等价目标函数:非高斯最大化、互信息最小、极大似然估计,总结三类常用优化方法:自适应法、固定点迭代法、Jacobi 法。介绍含噪独立分量分析、欠定独立分量分析、核独立分量分析等新近的方法,简述独立分量分析的一些应用。

关键词:独立分量分析;目标函数;含噪 ICA;欠定 ICA;核 ICA

中图分类号:N93

文献标识码:A

独立分量分析^[1](Independent component analysis,以下简称 ICA)是 20 世纪 90 年代发展起来的一种新的信号处理技术。其目的是从观测的混合信号中分离出分布未知但相互统计独立的源信号。由于 ICA 是基于高阶统计特性的分析方法,经 ICA 分解的各信号分量之间是相互独立的,在应用中的广泛性和技术上的更高要求,这也使得 ICA 成为国际上的研究热点。

目前 ICA 研究主要可分两大类,一是基本理论和算法研究。理论研究主要有线性 ICA、非线性 ICA、含噪 ICA、欠定 ICA、核 ICA 等的研究。算法方面,有一些算法是基于原有算法的优化改进,另外一些是基于二阶累积量、四阶累积量等高阶累积量方法及基于信息论准则的自适应算法等的研究。二是 ICA 的实际应用。随着 ICA 的研究发展,被大量应用在图像处理、信号处理、语音识别、故障检测、生物医学信号处理等领域。

1 ICA 的理论

1.1 ICA 的基本模型

若有 n 未知的源信号 s , 构成一个列向量 $s = (s_1, s_2, \dots, s_m)^T$, 假定在某时刻 t , 有 $x = (x_1, x_2, \dots, x_n)^T$ 为 n 维随机观测列向量,且满足下列方程:

$$x(t) = As(t) = \sum_{i=1}^m a_i s_i$$

其中 a_i 表示混合矩阵 A 的第 m 个行向量中的第 i 个。ICA 的目的就是求出一个解混矩阵 B , 使得 x 通过它后

得到 y 是 s 的最优逼近。用数学公式的可以表示为: $y(t) = Bx(t) = BAs(t)$ 。整个过程可以简单表示为图 1。

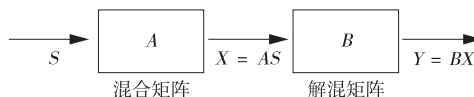


图 1 ICA 的基本模型

对于 ICA 来说,需要找到度量统计独立性的的一些目标函数^[1-5]。当取定目标函数,使之达到最大或者最小时,则认为达到了 ICA 解的要求,而这一过程可以通过优化方法实现。直观的说,ICA 其实主要包含两部分内容:目标函数和优化方法。芬兰学者 Hyvarinen^[5]等将这两部分表示为:

独立分量分析 = 目标函数 + 优化方法

1.2 ICA 目标函数

在假设统计独立的情况下,从高阶统计量分析、信息论以及参数估计推导信号独立的数学描述,构造三类基本目标函数^[2]:非高斯最大化、互信息最小、极大似然估计。

1.2.1 非高斯最大化

非高斯性的存在是 ICA 方法的前提条件,根据中心极限定理可知,当其非高斯性达到最大时,则实现了最佳分离。实际计算中,非高斯性程度常采用四阶累积量即 Kurtosis(峭度)^[6]来表示:

收稿日期:2012-04-09

基金项目:中国地质调查局项目(1212011120212);公益性行业科研专项(201011002-08);数学地质四川省重点实验室开放基金资助项目(SCSXDZ2009017, SCSXDZ2011015, SCSXDZ2009002);四川省教育厅自然科学重点项目(10ZA1116, 11ZA050)

作者简介:包云飞(1987-),男,甘肃武山人,硕士生,主要从事非线性算法理论方面的研究,(E-mail)byf_study@126.com

$$kurt(x) = E\{x^4\} - 3(E\{x^2\})^2$$

当随机变量为高斯分布时,峭度为零。而超高斯分布的峭度为正值,亚高斯分布的峭度为负值,且非高斯性越强,峭度的绝对值越大。

$$J(x) = H_c(x) - H(x) = D(p(x) \| p_c(x)) \quad (1)$$

另一种描述随机变量非高斯的方法是负熵^[1-2],其定义如式(1)。负熵的值总是非负的,只有当随机变量 x 为高斯分布时,负熵为零。在等协方差的限定条件下,随机变量的非高斯性越强,其负熵值越大。

1.2.2 互信息最小

设 N 维随机列向量 $x(t)$ 的联合概率密度函数(简称 pdf)为 $p(x)$, 它的各分量 $x_i(t)$ 的 pdf 为 $p_i(x_i)$ 。用 $p(x_i)$ ($i = 1 \sim N$) 和 $\prod_{i=1}^N p(x_i)$ 二者的 KL 散度^[1-2] 衡量 $x(t)$ 各分量之间的统计独立性。这也称为 $x(t)$ 各分量间的互信息,并表示为:

$$I(x) = KL[p(x), \prod_{i=1}^N p(x_i)] = \int p(x) \log[p(x) / \prod_{i=1}^N p(x_i)] dx$$

可以看到三种完全等价的表述: $I(x) = 0, p(x) = \prod_{i=1}^N p(x_i), x(t)$ 各分量统计独立。从信息论可知互信息是用来度量随机变量之间独立性的基本准则。互信息总是非负的^[7], 当变量之间相互独立时,互信息为零。

1.2.3 极大似然估计

极大似然估计是一种常用的估计方法,能够有效的用来求解无噪声的 ICA 问题。若观察信号 x 由信号源 s 经过线性映射 A 产生,选择 A 使 $\log \hat{p}(x | A)$ 达到最大,则称作极大似然估计。期望值:

$$E[\log \hat{p}(x | A)] = \int p(x) \log \hat{p}(x | A) dx \stackrel{def}{=} L(A)$$

就是目标函数^[1,3,7],极大似然估计的含义就是选择 A 使 $L(A)$ 最大。

$$L(B) = \frac{1}{T} \sum_{t=1}^T \log p(x | B) = \log |B| + \frac{1}{T} \sum_{t=1}^T \sum_{i=1}^M \log p[\sum_j b_{ij} x_j(t)] \quad (2)$$

若某时刻解混过程为 $Y = Bx$, 给定解混矩阵 B , 且 y 各分量互独立时, x 的对数似然函数为 $\log p(x | B)$, 且有式(2), 其中 b_{ij} 是解混矩阵 B 的第 (i, j) 元素。选择适当的 B , 使 $L(B)$ 取最大值, 这就是最大似然估计^[7]。

1.3 ICA 的主要优化方法

目标函数都是基于源信号的独立性的等价描述,是相互统一的,其本质都归结为一个求解函数最优化的问题,可以简单的统一表示为^[3]:

$$L(B) = \min(\max) f(B, x) \quad (3)$$

1.3.1 自适应法(梯度法)

梯度算法^[2-3] 是对(3)式求导,令 $\partial f(B, x) / \partial B = B(k+1) - B(k) = \nabla B$ 。以 ∇B 为指导,迭代求解 $B(k)$ 。当 $\nabla B < \varepsilon$ (ε 是事先定义的一个阈值),迭代结束。一般情况还需要引进一个步长 μ_k 调节,公式为:

$$B(k+1) = B(k) + \mu_k \nabla B = B(k) + \mu_k (B^{-T}(k) - \psi(y(k))x(k)^T)$$

其中, $x(k)$ 是混合信号, $y(k)$ 是分离出的信号,这里的 k 表示第 k 次迭代。自适应算法最大的特点就是通过逐步调节 B 使目标函数达到最优。常用的自适应算法如:自然梯度法^[4,8]、相对梯度法^[9]、Infomax 法^[10] 等。

1.3.2 固定点迭代方法

自适应算法优点在于能在变化的环境中快速的自适应,其缺点是收敛速度往往较慢。因此产生了一种称为 FastICA 的定点迭代算法^[4]。1997 年芬兰学者 Hyvarinen 等人首次提出了基于峭度的 FastICA 算法^[11]。由于峭度不是鲁棒度量,且特异值对它的影响较大,而负熵可以很好的解决这一问题。1999 年又提出了基于负熵的改进 FastICA 算法^[12]。基于负熵的固定点迭代效果优于累积量,这也使得采用负熵的 FastICA 算法应用更加广泛。为提高算法的稳定性,还可以采用牛顿迭代法优化。

1.3.3 Jacobi 法

Givens 旋转是 Jacobi 算法的基础^[1,3],它按照一定要求对一矢量进行正交归一变换,并且反复进行一系列坐标平面旋转来达到变换目的。对于矢量 z 旋转:

$$G_{ij} z = \begin{bmatrix} 1 & & & & & & & & & & \\ & \ddots & & & & & & & & & \\ & & \cos\theta_{ij} & \cdots & -\sin\theta_{ij} & & & & & & \\ & & \vdots & \ddots & \vdots & & & & & & \\ & & \sin\theta_{ij} & \cdots & -\cos\theta_{ij} & & & & & & \\ & & & & & \ddots & & & & & \\ & & & & & & \ddots & & & & \\ & & & & & & & & & & 1 \end{bmatrix} \begin{bmatrix} z_1 \\ \vdots \\ z_i \\ \vdots \\ z_j \\ \vdots \\ z_n \end{bmatrix}$$

式中 G_{ij} 除与 i, j 行列相交的 4 个元素外,其余的对角元素为 1,非对角元素为 0。对所有的 z_i, z_j 作多次迭代扫描,使得所有的 θ_{ij} 都小于预设的阈值,则可以求得 G 。此时的解混矩阵为 $B = W \cdot G[\theta(B)]$ 。特征矩阵联合近似对角化法^[13] 就是应用这种优化方法。

2 当前主要的 ICA 方法

2.1 含噪声的 ICA

一般 ICA 假设源信号是不含噪声的,但实际的信号往往是若干含噪声信号的混合。因此如何分离含有噪

声信号的 ICA 问题成为一个研究热点^[14-16]。处理该问题常用方法是利用贝叶斯(Bayes)方法或极大似然估计与混合高斯模型结合的方法,常见的如:Bayesian ensemble learning 方法与 mean field 方法^[17-18]。

Hyvarinen 等用稀疏编码方法去噪^[19],根据阈值判别噪声,对高斯噪声消除有较好的效果;Haritopoulos 等利用基于非线性 ICA 的自组织神经网络进行去噪^[20];文献[21]根据信号和噪声在不同尺度下表现不同,利用小波门限去噪的方法对有噪信号进行去噪处理后再利用 ICA 分离,有效的提升分离结果的峰值信号比。

2.2 欠定的 ICA

在实际分离中,源信号个数是未知的,而且通过传感器个数时常少于源信号个数,这样就产生了欠定的盲源分离问题。1991年 J. F. Cardoso 最早提出欠定盲信号分离问题^[22]。对于欠定的盲信号问题,经典的 ICA 算法不再适用。由于信号的稀疏性和独立性在某种意义上可以近似等价,这样产生了稀疏分量分析(SCA)。目前稀疏分量分析就是解决欠定盲源分离问题的主要方法。

1996年 X. R. Cao 对欠定盲信号分离做了分析^[23]。2001年, P. Bofill 提出一种基于 SCA 的欠定盲信号分离两步算法^[24]。2007年 A. Cichocki 等利用多层 NMF 实现了极度稀疏欠定混合盲信号的分离^[25]。文献[26]中提出了一种基于 l^∞ 范数的稀疏独立分量分析(SICA)算法用来做图像的特征提取。

2.3 核独立分量分析

核 ICA^[27]是一类再生核希尔伯特空间(P reproducing kernel Hilbert Space, RKHS)上借助对比函数完成独立分量分析的算法。通过非线性函数 $\varphi(\cdot)$ 将输入数据 $x_k \in R^n$, ($k = 1, 2, \dots, T$) 映射到高维特征空间上得到 $\varphi(x_1), \varphi(x_2), \dots, \varphi(x_T)$, 进而在该空间上对映射数据进行 ICA 处理^[28]。由于利用了“核技巧”(即在特征空间里,核可以用来计算内积的值),该算法可以适应许多种源信号的分布,具有更强的鲁棒性。快速独立分量分析^[29]是一种基于 KICA 的改进算法,具有更好的执行效率和适用性。

2.4 约束独立分量分析(cICA)

在进行独立分量分析时,有时会有一些不确定性对特定的应用会产生不利影响,为了消除这种不确定性的影响,需要提供一些譬如信号的正负特征、大致波形等的附加条件。基于这种思想,产生约束独立分量分析方法^[30]。后在此基础上产生非负 ICA 算法^[31]。

2.5 其他一些 ICA 模型

除了上面介绍的几种 ICA 拓展模型外还有非线性

ICA^[32],与线性 ICA 相比,它在应用中更符合实际。局部 ICA^[33]是先将数据按照一个标准分类,然后对每类数据进行 ICA 处理。分组 ICA^[34]与局部 ICA 类似,区别是各分组之间的分量是相互独立的。其它的如时空 ICA^[35]、判别独立分量分析^[36]等。还有大量的改进优化算法:改进的 FastICA 算法^[37]、改进的基于非高斯的故障检测的 ICA 算法^[38]等。还有些与其它原理相结合而产生的方法,如与小波结合的 ICA 算法^[15]、基于 EMD 方法的 ICA 算法^[39]等。

3 ICA 的应用

作为目前一种很有效技术,ICA 已经在信号处理、图像处理、生物医学信号处理、金融、故障检测等领域取得了广泛的应用。

在信号处理方面,文献[14-15]利用 ICA 方法实现了信号分离;文献[19,39]中应用基于 ICA 的技术对信号的降噪;文献[37]利用 ICA 方法做信号检测。在图像处理^[38]方面,文献[16,19]中应用 ICA 方法对图像进行去噪;文献[31]中进行了图像的特征提取;文献[40]用 ICA 来做三维图像融合。

近年来 ICA 方法在医学图像信号中的应用^[16]发展比较突出。在其他领域 ICA 的应用也十分迅速。文献[17]应用 ICA 算法来分类,文献[28]利用 ICA 进行机械故障检测,文献[21,35]做数据处理。除此之外,ICA 技术在其它领域还有广泛的应用。

4 结束语

ICA 技术作为一种有效信号处理技术,由于发展迅速,产生了大量的理论和算法,大量应用在了各种实际工作中,并取得了很好的成果。这些也展示了 ICA 技术的特点和应用价值。本文对 ICA 理论和算法的总结,旨在希望可以有更多的人了解到 ICA。有些需要完善和扩充的理论和算法,希望将来的工作者补充和完善。

参考文献:

- [1] 杨福生,洪波.独立分量分析的原理与应用[M].北京:清华大学出版社,2006.
- [2] 马建仓,牛奕龙,陈海洋.盲信号处理[M].北京:国防工业出版社,2006.
- [3] 蔡荣太,王延杰.独立分量分析的理论框架[J].软件时空,2007,3(23):197-305.
- [4] 史习智.盲信号处理—理论与实践[M].上海:上海交通大学出版社,2008.
- [5] Aapo Hyvarinen, Juha Karhunen, Oja Erkki. Independent

- Component Analysis[M].New York:John Wiley,2001.
- [6] 杨竹青,李勇,胡德文.独立分量分析方法综述[J].自动化学报,2002,28(5):762-772.
- [7] 田其冲,郑卫国,孙大雷.独立分量分析及其应用[J].电脑知识与技术,2009,18(5):5010-5012.
- [8] Yang H, Amari S. Adaptive on-line learning algorithms for blind separation-maximum entropy and minimum mutual information[J]. Neural Computations, 1997 (9): 1457-1482.
- [9] Cardoso J F. Equivariant adaptive separation[J]. IEEE Trans. on Signal Processing, 1996, SP-44(12):3017-3029.
- [10] Bell A, Sejnowski T. An information maximization approach to blind separation and blind deconvolution [J]. Neural Computation, 1995, 7(6):1129-1159.
- [11] Hyvarinen A, Oja E. A Fast Fixed-point Algorithm for Independent Component Analysis [J]. Neural Computation, 1997, 9(7):1483-1492.
- [12] Hyvarinen A. Fast and robust fixed-point algorithms for independent component analysis [J]. IEEE Transactions on Neural Networks, 1999, 10(3):626-634.
- [13] Poletto F, Rocca F, Bertelli L. Drill-bit signal separation for RVSP using statistical independence [J]. Geo-physics, 2000, 65:1645-1659.
- [14] Paraschiv-Ionescu A, Utten C, Aminian K, et al. Wavelet de-noising for highly noisy source separation [A]. In IEEE International Conference on Acoustics, Speech and Signal Processing (ASSP) [C]. USA, 2002, 1: 201-204.
- [15] Heinz Mathisa, Marcel Joho. Blind signal separation in noisy environments using a three-step quantizer [J]. Neuro-computing, 2002, 49(1-4):61-78.
- [16] Zhou Weidong, Jia Lei. Removal of noise and ECG artifact from EEG based on wavelet transform and independent component analysis [J]. Journal of Shandong University (Health Science), 2003, 41(2):116-119.
- [17] Opperand M, Winther O. Gaussian Processes for Classification: Mean-Field Algorithms [J]. Neural Computation, 2000, 12(11):2655-2684.
- [18] Petersen B. Mean Field ICA [D]. Denmark, Technical University of Denmark, 2005.
- [19] Hyvarinen A, Hoyer P, Oja E. Sparse code shrinkage for image denoising [J]. IEEE World Congress on Computational Intelligence, 1998, 2(10):859-864.
- [20] Haritopoulos M, Yin H J, Allinson N M. Image denoising using self-organizing map-based nonlinear independent component analysis [J]. Neural Network, 2002, 22 (15): 1085-1098.
- [21] 张朝柱,张健沛,孙晓东.基于 curvelet 变换和独立分量分析的含噪盲源分离[J].计算机应用,2008,28(5):1208-1210.
- [22] Cardoso J F. Super-symmetric decomposition of fourth-order cumulant tensor. Blind identification of more sources than sensors [C]. International Conference on Acoustics, Speech, and Signal Processing, 1991, 5: 3109-3112.
- [23] Cao X R, X Liu X R. General approach to blind source separation [J]. IEEE Transactions on Signal processing, 1996, 44(3):562-571.
- [24] Bofill P, Zibulevsky M. Underdetermined blind source separation using sparser representations [J]. Signal Process, 2001, 81(11):2353-2362.
- [25] Cichocki A, Zdunek R, Amari S. Hierarchical ALS Algorithms for Nonnegative Matrix and 3D Tensor Factorization [J]. Springer, Heidelberg, 2007, 4666:169-176.
- [26] 尚丽,杜吉祥.基于 l^{∞} 范数的稀疏 ICA 和 FCM 的自然图像特征提取 [J]. 2010, 31(4):783-787.
- [27] Bach F, Jordan M I. Kernel independent component analysis [J]. Journal of Machine Learning Research, 2002, 3(7):1-48.
- [28] Achmad Widodo, Bo-Suk Yang. Application of Nonlinear Feature Extraction and Support Vector Machines for Fault Diagnosis of Induction Motors [J]. Expert Systems with Applications, 2007, 1(33):241-250.
- [29] Shen H, Jegelka S, Gretton A. Fast kernel-based independent component analysis [J]. IEEE Trans on Signal Process. 2009, 57(9):202-857.
- [30] Lu W, Rajapakse J C. Constrained independent component analysis [C]. Advances in Information Processing Systems. 2000:570-576.
- [31] Plumbley M D. Algorithms for nonnegative independent component analysis [J]. IEEE Transactions on Neural Networks, 2003, 14:534-543.
- [32] Jutten C, Karhunen J. Advances in nonlinear Blind source Separation [C]. 4th International Symposium on Independent Component Analysis and Blind Source Separation, 2003:254-256.
- [33] Cichocki A, Amari S. Adaptive Blind Signal and Image Processing: Learning Algorithms and Applications [M].

- New York: Wiley, 2003.
- [34] Calhoun V, Adali T, Pearlson G. et al. Group ICA of functional MRI data: separability, stationarity, inference [C]. In Processing of ICA 2001, San Diego, 2001: 155-160.
- [35] Stone J V, Porrill J, Porter N R. Spatiotemporal ICA of fMRI data [R]. Computational Neuros. science Report , 202, May 2000.
- [36] Chandra Shekhar Dhir, Soo-Young Lee. Discriminant Independent Component Analysis [J]. IEEE Transactions on Neural Networks, 2011, 22(6): 845-856.
- [37] 汪斌, 王年, 蒋云志, 等. 改进 FastICA 算法在谐波检测中的应用 [J]. 电力自动化设备, 2011, 31(3): 135-138.
- [38] Zhang Yingwei, Zhang Yang. Fault detection of non-Gaussian processes based on modified independent component analysis [J]. Chemical Engineering Science, 2010, 65: 4630-4639.
- [39] 李洪, 孙云莲. 基于 EMD 虚拟通道的 ICA 算法在信号消噪中的应用 [J]. 北京邮电大学学报, 2007, 30(5): 33-36.
- [40] Guong Manh Do, Bahram Javidi. Multifocus Holographic 3-D Image Fusion Using Independent Component Analysis [J]. IEEE Journal of Display Technology, 2007, 3(3): 326-332.

Research on Independent Component Analysis

BAO Yun-fei, ZHANG Wei-chuan, WU Fei, ZHANG Li

(College of Management Science, Chengdu University of Technology, Chengdu 610059, China)

Abstract: Summarized the basic theory of Independent component analysis. Summarize three equivalent target function: non-gaussian maximization, the minimum mutual information, maximum likelihood estimate. Summarize three kinds of common optimization method: adaptive method, fixed-point iteration, jacobi method. Introduced the latest methods: Containing noise of ICA, Underdetermined ICA, Kernel ICA etc. The application of ICA is briefly introduced.

Key words: ICA; target function; containing noise of ICA; underdetermined ICA; kernel ICA