

WILEY

*Independent Component Analysis*

生命科学  
与信息技术丛书

# 独立成分 分析

Aapo Hyvärinen  
[芬] Juha Karhunen 著  
Erkki Oja

周宗潭 董国华 徐 昕 胡德文 等译



电子工业出版社  
PUBLISHING HOUSE OF ELECTRONICS INDUSTRY

<http://www.phei.com.cn>

生命科学与信息技术丛书

# 独立成分分析

Independent Component Analysis

Aapo Hyvärinen

[芬] Juha Karhunen 著

Erkki Oja

周宗潭 董国华  
徐昕 胡德文



电子工业出版社

Publishing House of Electronics Industry

北京 · BEIJING

## 内 容 简 介

独立成分分析(ICA)已经成为神经网络、高级统计学和信号处理等研究领域中的重要方向之一。本书是国际上一本 ICA 的综合性著作,其中包括理解和使用该技术的相应数学基础知识。本书不仅介绍 ICA 的基本知识与概况,给出了重要的求解过程及算法,还涵盖了图像处理、无线通信、音频信号处理及更多其他应用。全书分四个部分共 24 章,第一部分介绍本书所用到的主要数学知识,第二部分是本书的重点,详细讲述了基本 ICA 模型及其求解过程,第三部分讨论基本 ICA 模型的多种扩展形式,第四部分讨论 ICA 方法在不同领域的应用。

本书可以作为不同工程应用领域的大学教师、研究生和科技工作者的 ICA 入门教材;而对于探索 ICA 技术的专业研究人员来说,本书也是一本极有价值的参考书。

Independent Component Analysis, 0-471-40540-X, Aapo Hyvärinen, Juha Karhunen, Erkki Oja.  
Original English language edition copyright © 2001, John Wiley & Sons, Inc. All rights reserved.  
This translation published under license.  
No part of this book may be reproduced in any form without the written permission of John Wiley & Sons, Inc.

本书简体中文版专有翻译版权由美国 John Wiley & Sons, Inc. 公司授予电子工业出版社。未经许可,不得以任何手段和形式复制或抄袭本书内容。

版权贸易合同登记号 图字:01-2006-0894

## 图书在版编目(CIP)数据

独立成分分析/(芬)海韦里恩,(芬)卡尔胡恩,(芬)奥亚著;周宗潭等译.

北京:电子工业出版社,2014.5

(生命科学与信息技术丛书)

书名原文:Independent Component Analysis

ISBN 978-7-121-22981-7

I. ①独… II. ①海… ②卡… ③奥… ④周… III. ①信号处理 IV. ①TN911.7

中国版本图书馆 CIP 数据核字(2014)第 078170 号

策划编辑:马 岚

责任编辑:李秦华

印 刷:三河市鑫金马印装有限公司

装 订:三河市鑫金马印装有限公司

出版发行:电子工业出版社

北京市海淀区万寿路 173 信箱 邮编 100036

开 本:720×1000 1/16 印张:28.25 字数:554 千字

印 次:2014 年 5 月第 1 次印刷

定 价:79.00 元

凡所购买电子工业出版社图书有缺损问题,请向购买书店调换。若书店售缺,请与本社发行部联系,联系及邮购电话:(010)88254888。

质量投诉请发邮件至 zlts@phei.com.cn,盗版侵权举报请发邮件至 dbqq@phei.com.cn。

服务热线:(010)88258888。

## 再 版 序

本书不仅是国际上比较全面的关于独立成分分析 (Independent Component Analysis, ICA) 的专著,也是国内正式翻译出版的关于 ICA 和盲源信号分离等技术的一本标准教材。中译本由电子工业出版社出版发行后反响很好,在同类教材中属于比较畅销的,某些时段还供不应求,以至于有读者来信联系译者索书,这些读者不仅有从事数学和信号处理研究的学者,还有更多来自无线通信、医学图像处理,特别是脑科学与认知研究等领域的工程师、研究生和专家教授。本书出版后,从国际和国内的引用、评价和应用案例的分析和统计来看,ICA 这项 20 世纪 90 年代才发展起来的应用数学方法正在由一项创新逐步变成为一部经典,而在这个过程中本书起到了重要的推广作用。

一般认为,ICA 的基本数学框架是 1986 年建立的,1995 年后随着其快速算法的出现而开始快速发展。目前国际上采用最多的称为 FastICA 的快速独立成分分析算法是由 Aapo Hyvärinen 和 Erkki Oja 两位学者提出的,而这两位学者就是本书的主要作者。可以说,本书不仅仅是一般的关于 ICA 和盲源信号分离的标准教材,同时也是一本关于 ICA 理论框架、ICA 算法和 ICA 在不同领域应用方法等学术研究的专著,并具有大量探讨性和启发性的内容,这是其他同类教材所不能比拟的。因此,虽然目前关于独立成分分析的文献、教材和在线教程等资源已经非常丰富,但本书的地位仍然是不可替代的。

本书的中译版于 2007 年由电子工业出版社推出,对独立成分分析技术在国内的普及应用已经起到了重要的推动作用。目前国内通信、电子等工科研究生教育和许多科学研究领域(特别是脑科学研究领域)对于独立成分分析方法的需求仍然在不断增长,需要比较全面系统且具有启发性的教材。考虑到本书特点和权威性,出版社决定将本书再次出版以满足读者的需求。本书现有的中译本是由国内较早从事独立成分分析和脑科学与认知科学的几位学者翻译的,他们对原书中专业性较强内容的理解和把握是到位的,翻译质量也得到了大部分读者的认同。故再版中译本仍沿用原先的翻译文字,主要是根据现代读者的阅读习惯和教材重点突出的风格要求,对原书内容进行了重新排版和修订。

# 原作者序

非常感谢胡德文教授为把我们的《独立成分分析》一书翻译成中文(原书由 John Wiley & Son 于 2001 年出版)所做出的重要努力。

我们于 1999 年开始撰写本书,当时独立成分分析(ICA)还是一种相当新颖的方法,它的前景在一定程度上尚未可定:它是否能够成为实际应用中的一种重要技术、其理论是否已得到充分的发展,这些都还是不清楚的。

应该说,在应用和理论两个方面,我们都是相当幸运的。一方面,许多实际应用领域的相关工作已经证实 ICA 是非常有用的,而在这些领域中,神经信息学和生物信息学尤其值得特别提及;另一方面,本书中所阐述的理论也经受住了时间的考验,现在已经被广泛接受。

中国的科学家已经对 ICA 的研究做出了重要的贡献,我们希望本书能够吸引和激励更多的参与者。同样也希望对 ICA 应用感兴趣的中国学者和专家能够在本书中找到有用的原则性导引,并在理论上有所帮助。

Aapo Hyvärinen, Juha Karhunen, Erkki Oja

## 英文原文:

We are very grateful to Prof. Dewen Hu for his great efforts in translating our book "Independent component analysis" (originally published by John Wiley and Sons in 2001) into Chinese.

When we started writing the book in 1999, ICA was still a rather new method, and its future was somewhat uncertain. It was not clear whether it would become an important technique for practical applications, and whether the theory was sufficiently developed.

We have been quite fortunate in both respects. On the one hand, ICA has proven very useful in a large number of application fields, of which neuroinfor-

matics and bioinformatics deserve special mention. On the other hand, the theory that we expounded in the book has stood the test of time, and is now well—established.

Chinese scientists have made important contributions to research on ICA, and we hope this book will inspire many more. We equally hope that Chinese researchers and professionals interested in applications of ICA will find useful guidelines and theoretical understanding in this book.

Aapo Hyvärinen, Juha Karhunen, and Erkki Oja

# 序 言

独立成分分析(ICA)是一种统计和计算技术,用于揭示随机变量、测量数据或信号中的隐藏成分。对于通常以大量样本数据库形式给出的多元观测数据,ICA定义了一个生成模型。此模型假设观测数据变量是某些未知内在变量的线性或非线性混合,而且不仅内在变量是未知的,实现混合的系统也是未知的。我们还假定那些内在变量是非高斯且相互独立的,并称它们为观测数据的独立成分。这些独立成分(也可称为源或因子)可以通过 ICA 方法找到。

ICA 可以看成是主成分分析和因子分析的延展。但是,ICA 是一项更强有力的技术,当经典方法完全失效时,它仍然能够找出支撑观测数据的内在因子或源。

ICA 所分析的数据可能来源于许多不同的应用领域,包括数字图像、文档数据库,以及经济指标和心理学测量。在许多实例中,测量结果是以一组并行(Parallel)信号或时间序列的形式给出的,盲源分离(Blind source separation)这一术语可以用于刻画该类问题。盲源分离的典型例子有:多个麦克风拾取同时发出语音的混合信号、多个传感器记录的脑电波、手机中的射频干扰信号或从某些工业过程中得到的并行时间序列。

ICA 技术也是一项相对较新的发明。它是 20 世纪 80 年代初期首先在神经网络建模领域中引入的。到 20 世纪 90 年代中期,几个研究小组引入了一些极为成功的新算法,类似鸡尾酒会效应问题的演示,也给人们留下了深刻的印象:ICA 可以从混合信号中找出每个人的语音波形。因此,无论是在神经网络领域(特别是无监督学习),还是在更为一般的高级统计学和信号处理领域中,ICA 都成了激动人心的新话题之一。ICA 在生物信号处理、语音信号分离、无线通信、故障诊断、特征提取、金融时间序列分析和数据挖掘等现实领域中的应用报道也正在不断涌现。

在过去的 20 年内,ICA 的文章大量发表在信号处理、人工神经网络、统计学、信息论及各种应用领域的期刊和会议论文集中。最近还举办了几场关于 ICA 的专题会议和学术研讨<sup>[70,348]</sup>,出版了一些关于 ICA、盲源分离以及其他相关主题的论文合集<sup>[315,173,150]</sup>和专著<sup>[105,267,149]</sup>。这些现有的文献对于其预期读者来说是非常有用的,但它们大都只集中于论述 ICA 方法的某些特定主题;而简短的专业论文和书籍章节中通常不包含数学和统计学的预备知识。因此,对更广泛的读者群来说,完全理解 ICA 这个技术性相当强的主题变得非常困难。

直到目前为止还没有一本内容广泛、细节丰富的教材:既能覆盖数学背景知

识、基本原理和求解算法,又能兼顾 ICA 实际应用的现状。本书的出版目的就是希望填补这个空缺,并充当 ICA 的入门导引。

本书的预期读者可能来自多个学科:诸如统计学、信号处理、神经网络、应用数学、神经和认知科学、信息论、人工智能和工程领域。不论是研究人员、学生还是工程实践领域的工作者都可以使用本书。我们还做了各种努力使本书能够自足(self-contained),以便那些仅有大学微积分、矩阵代数、概率论和统计学基本背景知识的读者阅读。本书也适合作为 ICA 方面的研究生教材,许多章节后都有习题和计算机练习作业。

## 本书的范围和内容

本书提供了把 ICA 作为统计与计算技术的全面导引,重点是数学原理和基本算法。大部分素材基于作者自己的研究小组中开展的原创性研究工作,这从不同主题所占比重的角度可以自然地反映出来。本书覆盖面较宽,特别涵盖了那些能同时适用于大规模问题求解的算法;即算法在观测变量数目和数据点个数都非常巨大时仍然有效。当 ICA 大量地用于真正的现实应用问题(而不是直到最近仍占主流的玩偶型问题或小规模探讨性研究)时,这些算法就能在不久的将来得到越来越多的应用。卷积型混合、时滞及其他一些不同于 ICA 的盲源分离技术中涉及到的更为专门的信号处理方法,本书相应地强调得略少一些。

由于 ICA 是一个进展迅速的研究领域,不可能在一本教材中囊括每项已报道过的研究进展。我们已尽力在适当的地方涵盖了其他研究者的核心贡献,并给出了一个庞大的文献目录,以便进一步参考。对于可能被遗漏的那些重要贡献,我们深表歉意。

为便于阅读,本书分成了四个部分。

- 第一部分给出了**数学预备知识**,引入了书中后面部分会用到的基本数学概念。该部分从第 2 章的概率论速成教程开始。假设读者已熟悉本章中的大部分内容,不过本章还引入了专为 ICA 预备的一些概念,如高阶累积量和多变量概率理论。在第 3 章中,接着讨论最优化理论和梯度法,它们在构造 ICA 算法时将会用到。第 4 章回顾和综述了估计理论。信息论是 ICA 的一个补充性理论框架,它包含在第 5 章中。第 6 章是第一部分的总结,其中讨论了与 ICA 相关的主成分分析、因子分析和去相关方法。

更为自信的读者可以按照其意愿跳过第一部分中部分或全部导论性的章节,而直接从第二部分中 ICA 的原理开始学习。

- 第二部分涵盖并解决了 **ICA 基本模型**: 即 ICA 中经典的线性瞬态无噪声混合模型, 它构成了 ICA 理论的核心。第 7 章引入了该基本模型, 并讨论了混合矩阵的可辨识性问题。后面各章讨论对该模型进行估计的不同方法。一个核心的原则是非高斯性, 它与 ICA 的关系在第 8 章中首次进行了讨论。接下来我们回顾了极大似然原理(第 9 章)和极小互信息原理(第 10 章), 并展示了这些基本原理之间的联系。第 11 章讨论了使用高阶累积张量的代数方法, 第 12 章回顾了基于非线性去相关的早期 ICA 工作以及非线性主成分方法, 这两章中的内容可能不太适合于(本书)导论教程的性质。在这些章节中, 我们还讨论了与每一种基本原理相对应的计算独立成分和混合矩阵的实用算法。在第 13 章中, 接着讨论了主要与数据预处理和数据降维相关的一些实际问题与具体考虑, 其中包括引导使用者如何把 ICA 真正确实地应用于自己的具体问题。第 14 章综述和比较了各种 ICA 方法, 并对第二部分进行了总结。
- 第三部分中给出了 ICA 基本模型的各种**扩展**。相比第二部分而言, 这部分内容具有更多的探索性质(而不够成熟), 因为大多数扩展是最近才引入的, 尚存在许多未解决的问题。如果要开设一门 ICA 的导论课程, 可能只需在该部分挑选出几章即可。我们首先在第 15 章中讨论了在 ICA 模型中引入显式观测噪声的问题, 然后在第 16 章中探讨独立成分个数多于混合观测量个数的情况。第 17 章进一步将模型进行推广: 混合过程可以是非常一般的非线性形式。第 18 章讨论了估计另外一种线性混合模型的方法, 该混合模型与 ICA 模型类似, 但是假设的前提条件差别较大: 即成分不是非高斯的, 但具有某种时间依赖性。第 19 章讨论了混合系统中包含卷积的情况。ICA 更进一步的扩展, 特别是不要求成分严格独立的模型, 将在第 20 章中给出。
- 第四部分探讨 ICA 方法的一些**应用**: 特征提取(第 21 章)与图像处理和视觉研究都有关系; 脑成像的应用(第 22 章)集中讨论人脑的电/磁活动测量。第 23 章讨论 ICA 在无线通信中的应用。ICA 在计量经济学和音频信号处理中的应用, 以及 ICA 对其他应用领域的启示在第 24 章中探讨。

书中那些非常复杂和入门课程中可以跳过的章节, 我们都已用星号标出。本书中给出的一些算法, 其对应软件是公开的, 可以通过互联网在我们的网页或其他 ICA 研究者提供的网页中获得。包含真实数据的数据库也同时提供下载, 以便测试算法。我们还为本书专门制作了网页, 其中包含了一些适当的提示, 地址是:

[www.cis.hut.fi/projects/ica/book](http://www.cis.hut.fi/projects/ica/book)

建议读者通过访问这些网页以获得更多的信息。

本书由三位作者合作完成, Aapo Hyvärinen 负责第 5 章、第 7 章、第 8 章、第 9 章、第 10 章、第 11 章、第 13 章、第 14 章、第 15 章、第 16 章、第 18 章、第 20 章、第 21 章和第 22 章; Juha Karhunen 负责第 2 章、第 4 章、第 17 章、第 19 章和第 23 章; Erkki Oja 负责第 3 章、第 6 章和第 12 章。第 1 章和第 24 章则由三位作者联合撰写。

## 致谢

感谢许多 ICA 研究者, 他们的原创性贡献形成了 ICA 的基础, 并使得本书的问世成为可能。我们特别对丛书主编 Simon Haykin 致以谢意, 多年以来, 他在信号处理和神经网络方面的文章和著作一直鼓舞着我们。

本书的某些部分, 是建立在与赫尔辛基技术大学 (Helsinki University of Technology) 中我们研究小组的其他成员紧密合作基础上的。第 21 章主要基于我们与 Patrik Hoyer 的合作研究, 该章所有实验也都是他完成的。第 22 章基于 Ricardo Vigário 的实验和素材。13. 2. 2 节基于我们与 Jaakko Särelä 及 Ricardo Vigário 的联合工作。16. 2. 3 节由 Razvan Cristescu 提供。20. 3 节基于我们和 Ella Bingham 的联合工作, 14. 4 节基于我们和 Xavier Giannakopoulos 的联合工作, 20. 2. 3 节则基于我们和 Patrik Hoyer 及 Mika Inki 的联合工作。第 19 章部分基于 Kari Torkkola 提供的素材。第 17 章中的许多内容基于我们和 Harri Valpola 及 Petteri Pajunen 的联合工作, 24. 1 节是我们和 Kimmo Kiviluoto 及 Simona Malaroiu 联合完成的。

在本书写作的不同阶段, 有几个人友善地同意阅读并评述了该书的部分或全部内容。为此, 我们诚挚地感谢 Ella Bingham、Jean-François Cardoso、Adrian Flanagan、Mark Girolami、Antti Honkela、Jarmo Hurri、Petteri Pajunen、Tapani Ristaniemi 和 Kari Torkkola。Leila Koivisto 在编辑上给予了我们帮助, Antti Honkela、Mika Ilmoniemä、Merja Oja 和 Tapani Raiko 在某些插图上提供了帮助。

我们在 ICA 方面的研究工作, 以及本书写作是在芬兰赫尔辛基技术大学的神经网络研究中心 (Neural Networks Research Centre) 进行的。研究工作的部分资助来自于“BLISS”项目 (欧盟) 和“信息处理新原理”项目 (芬兰科学院), 在此深表谢意。另外, A. H. (即 A. Hyvärinen, 原书只用了姓名的缩写) 要表达对赫尔辛基大学心理学系 Göte Nyman 和 Jukka Häkkinen 的谢意, 他们提供的友好服务, 使得本书的部分内容能够完成。

Aapo Hyvärinen, Juha Karhunen, Erkki Oja  
埃斯波 (Espoo), 芬兰 (Finland)

# 目 录

|                     |    |
|---------------------|----|
| 第 1 章 引论 .....      | 1  |
| 1.1 多元数据的线性表示 ..... | 1  |
| 1.2 盲源分离 .....      | 3  |
| 1.3 独立成分分析 .....    | 6  |
| 1.4 ICA 的历史 .....   | 10 |

## 第一部分 数学预备知识

|                      |    |
|----------------------|----|
| 第 2 章 随机向量和独立性 ..... | 14 |
| 2.1 概率分布和概率密度 .....  | 14 |
| 2.2 期望和矩 .....       | 18 |
| 2.3 不相关性和独立性 .....   | 22 |
| 2.4 条件密度和贝叶斯法则 ..... | 26 |
| 2.5 多元高斯密度 .....     | 29 |
| 2.6 变换的密度 .....      | 32 |
| 2.7 高阶统计量 .....      | 33 |
| 2.8 随机过程 .....       | 39 |
| 2.9 小结与文献引述 .....    | 47 |
| 习题 .....             | 47 |
| 计算机练习 .....          | 51 |
| 第 3 章 梯度和最优化方法 ..... | 53 |
| 3.1 向量和矩阵梯度 .....    | 53 |
| 3.2 无约束优化和学习规则 ..... | 58 |
| 3.3 约束优化的学习规则 .....  | 68 |
| 3.4 小结与文献引述 .....    | 70 |
| 习题 .....             | 70 |
| 计算机练习 .....          | 72 |
| 第 4 章 估计理论 .....     | 73 |
| 4.1 基本概念 .....       | 73 |
| 4.2 估计器的性质 .....     | 75 |

|              |                       |            |
|--------------|-----------------------|------------|
| 4.3          | 矩方法 .....             | 79         |
| 4.4          | 最小二乘估计 .....          | 81         |
| 4.5          | 极大似然法 .....           | 84         |
| 4.6          | 贝叶斯估计 .....           | 88         |
| 4.7          | 小结与文献引述 .....         | 92         |
|              | 习题 .....              | 94         |
|              | 计算机练习 .....           | 97         |
| <b>第 5 章</b> | <b>信息论 .....</b>      | <b>99</b>  |
| 5.1          | 熵 .....               | 99         |
| 5.2          | 互信息 .....             | 103        |
| 5.3          | 极大熵 .....             | 104        |
| 5.4          | 负熵 .....              | 106        |
| 5.5          | 通过累积量逼近熵 .....        | 106        |
| 5.6          | 用非多项式函数近似熵 .....      | 108        |
| 5.7          | 小结与文献引述 .....         | 114        |
|              | 习题 .....              | 114        |
|              | 计算机练习 .....           | 114        |
|              | 本章附录:有关证明 .....       | 115        |
| <b>第 6 章</b> | <b>主成分分析和白化 .....</b> | <b>118</b> |
| 6.1          | 主成分 .....             | 118        |
| 6.2          | 在线学习的 PCA .....       | 124        |
| 6.3          | 因子分析 .....            | 129        |
| 6.4          | 白化 .....              | 130        |
| 6.5          | 正交化 .....             | 132        |
| 6.6          | 小结与文献引述 .....         | 133        |
|              | 习题 .....              | 134        |

## 第二部分 独立成分分析基本模型

|              |                        |            |
|--------------|------------------------|------------|
| <b>第 7 章</b> | <b>什么是独立成分分析 .....</b> | <b>138</b> |
| 7.1          | 动机 .....               | 138        |
| 7.2          | 独立成分分析的定义 .....        | 142        |
| 7.3          | ICA 的实例 .....          | 145        |
| 7.4          | ICA 比白化更加强大 .....      | 147        |
| 7.5          | 高斯变量为何不能适用 .....       | 150        |
| 7.6          | 小结与文献引述 .....          | 151        |

|                                      |            |
|--------------------------------------|------------|
| 习题 .....                             | 152        |
| 计算机练习 .....                          | 153        |
| <b>第 8 章 极大化非高斯性的 ICA 估计方法 .....</b> | <b>154</b> |
| 8.1 非高斯就是独立的 .....                   | 154        |
| 8.2 用峭度来度量非高斯性 .....                 | 158        |
| 8.3 用负熵度量非高斯性 .....                  | 166        |
| 8.4 估计多个独立成分 .....                   | 175        |
| 8.5 ICA 与投影寻踪 .....                  | 178        |
| 8.6 小结与文献引述 .....                    | 180        |
| 习题 .....                             | 180        |
| 计算机练习 .....                          | 182        |
| 本章附录:有关证明 .....                      | 182        |
| <b>第 9 章 ICA 的极大似然估计方法 .....</b>     | <b>185</b> |
| 9.1 ICA 模型中的似然度 .....                | 185        |
| 9.2 极大似然估计算法 .....                   | 188        |
| 9.3 信息极大原理 .....                     | 193        |
| 9.4 例子 .....                         | 194        |
| 9.5 小结与文献引述 .....                    | 196        |
| 习题 .....                             | 197        |
| 计算机练习 .....                          | 198        |
| 本章附录:有关证明 .....                      | 198        |
| <b>第 10 章 极小化互信息的 ICA 估计方法 .....</b> | <b>199</b> |
| 10.1 用互信息定义 ICA .....                | 199        |
| 10.2 互信息和非高斯性 .....                  | 200        |
| 10.3 互信息和似然估计 .....                  | 201        |
| 10.4 极小化互信息的算法 .....                 | 201        |
| 10.5 例子 .....                        | 202        |
| 10.6 小结与文献引述 .....                   | 203        |
| 习题 .....                             | 203        |
| 计算机练习 .....                          | 203        |
| <b>第 11 章 基于张量的 ICA 估计方法 .....</b>   | <b>204</b> |
| 11.1 累积张量的定义 .....                   | 204        |
| 11.2 由张量特征值得到独立成分 .....              | 205        |
| 11.3 用幂法计算张量分解 .....                 | 206        |

|               |                                          |     |
|---------------|------------------------------------------|-----|
| 11.4          | 特征矩阵的联合近似对角化 .....                       | 208 |
| 11.5          | 加权相关矩阵方法 .....                           | 209 |
| 11.6          | 小结与文献引述 .....                            | 210 |
|               | 习题 .....                                 | 211 |
|               | 计算机练习 .....                              | 211 |
| <b>第 12 章</b> | <b>基于非线性去相关和非线性 PCA 的 ICA 估计方法</b> ..... | 212 |
| 12.1          | 非线性相关和独立性 .....                          | 212 |
| 12.2          | Hérault-Jutten 算法 .....                  | 214 |
| 12.3          | Cichocki-Unbenauen 算法 .....              | 216 |
| 12.4          | 估计函数方法 .....                             | 217 |
| 12.5          | 通过独立性的等变自适应分离 .....                      | 218 |
| 12.6          | 非线性主成分 .....                             | 220 |
| 12.7          | 非线性 PCA 指标和 ICA .....                    | 223 |
| 12.8          | 非线性 PCA 指标的学习规则 .....                    | 225 |
| 12.9          | 小结与文献引述 .....                            | 230 |
|               | 习题 .....                                 | 231 |
| <b>第 13 章</b> | <b>实际的考虑</b> .....                       | 232 |
| 13.1          | 时间滤波作为预处理 .....                          | 232 |
| 13.2          | 用 PCA 进行预处理 .....                        | 235 |
| 13.3          | 应该估计多少个成分 .....                          | 238 |
| 13.4          | 算法选择 .....                               | 238 |
| 13.5          | 小结与文献引述 .....                            | 239 |
|               | 习题 .....                                 | 239 |
|               | 计算机练习 .....                              | 239 |
| <b>第 14 章</b> | <b>基本 ICA 方法的综述和比较</b> .....             | 240 |
| 14.1          | 目标函数和算法 .....                            | 240 |
| 14.2          | ICA 估计原理的联系 .....                        | 240 |
| 14.3          | 统计最优非线性函数 .....                          | 242 |
| 14.4          | ICA 算法的实验比较 .....                        | 246 |
| 14.5          | 参考文献 .....                               | 251 |
| 14.6          | 基本 ICA 方法小结 .....                        | 252 |
|               | 本章附录:有关证明 .....                          | 253 |

### 第三部分 ICA 的扩展及其相关方法

|        |                            |     |
|--------|----------------------------|-----|
| 第 15 章 | 有噪声的 ICA 模型 .....          | 256 |
| 15.1   | 定义 .....                   | 256 |
| 15.2   | 传感器噪声和信号源噪声 .....          | 257 |
| 15.3   | 噪声成分数目较少的情况 .....          | 257 |
| 15.4   | 混合矩阵的估计 .....              | 258 |
| 15.5   | 估计无噪声的独立成分 .....           | 262 |
| 15.6   | 通过稀疏编码收缩而去噪 .....          | 265 |
| 15.7   | 小结 .....                   | 265 |
| 第 16 章 | 具有超完备基的 ICA 模型 .....       | 267 |
| 16.1   | 独立成分的估计 .....              | 267 |
| 16.2   | 估计混合矩阵 .....               | 269 |
| 16.3   | 小结 .....                   | 273 |
| 第 17 章 | 非线性 ICA .....              | 275 |
| 17.1   | 非线性 ICA 与 BSS .....        | 275 |
| 17.2   | 后非线性混合的分离 .....            | 278 |
| 17.3   | 采用自组织映射的非线性 BSS .....      | 280 |
| 17.4   | 非线性 BSS 的一种生成拓扑映射方法 .....  | 281 |
| 17.5   | 非线性 BSS 的一种集成学习方法 .....    | 287 |
| 17.6   | 其他方法 .....                 | 295 |
| 17.7   | 小结 .....                   | 297 |
| 第 18 章 | 使用时间结构的方法 .....            | 299 |
| 18.1   | 通过自协方差实现分离 .....           | 299 |
| 18.2   | 利用方差的非平稳性实现分离 .....        | 304 |
| 18.3   | 统一的分离原理 .....              | 309 |
| 18.4   | 小结 .....                   | 311 |
| 第 19 章 | 卷积性混合和盲去卷积 .....           | 312 |
| 19.1   | 盲去卷积 .....                 | 312 |
| 19.2   | 卷积性混合的盲分离 .....            | 317 |
| 19.3   | 小结 .....                   | 324 |
|        | 本章附录:离散时间滤波器和 $z$ 变换 ..... | 325 |
| 第 20 章 | ICA 的其他扩展 .....            | 328 |
| 20.1   | 混合矩阵的先验信息 .....            | 328 |

|                      |                           |            |
|----------------------|---------------------------|------------|
| 20.2                 | 放宽独立性假设 .....             | 335        |
| 20.3                 | 复值数据的处理 .....             | 339        |
| 20.4                 | 小结 .....                  | 343        |
| <b>第四部分 ICA 的应用</b>  |                           |            |
| <b>第 21 章</b>        | <b>基于 ICA 的特征提取 .....</b> | <b>346</b> |
| 21.1                 | 线性表示 .....                | 346        |
| 21.2                 | ICA 和稀疏编码 .....           | 350        |
| 21.3                 | 从图像中估计 ICA 的基向量 .....     | 351        |
| 21.4                 | 压缩稀疏编码用于图像去噪 .....        | 353        |
| 21.5                 | 独立子空间和拓扑 ICA .....        | 356        |
| 21.6                 | 与神经生理学的联系 .....           | 358        |
| 21.7                 | 小结 .....                  | 358        |
| <b>第 22 章</b>        | <b>ICA 在脑成像中的应用 .....</b> | <b>359</b> |
| 22.1                 | 脑电图和脑磁图 .....             | 359        |
| 22.2                 | EEG 和 MEG 中的伪迹鉴别 .....    | 361        |
| 22.3                 | 诱发磁场分析 .....              | 363        |
| 22.4                 | ICA 使用于其他的测量技术中 .....     | 366        |
| 22.5                 | 小结 .....                  | 366        |
| <b>第 23 章</b>        | <b>无线通信 .....</b>         | <b>368</b> |
| 23.1                 | 多用户检测和 CDMA 通信 .....      | 368        |
| 23.2                 | CDMA 信号模型和 ICA .....      | 372        |
| 23.3                 | 衰落信道的估计 .....             | 375        |
| 23.4                 | 卷积 CDMA 信号的盲分离 .....      | 380        |
| 23.5                 | 采用复值 ICA 改进多用户检测 .....    | 384        |
| 23.6                 | 小结与文献引述 .....             | 389        |
| <b>第 24 章</b>        | <b>ICA 的其他应用 .....</b>    | <b>391</b> |
| 24.1                 | 金融方面的应用 .....             | 391        |
| 24.2                 | 音频分离 .....                | 395        |
| 24.3                 | 更多的应用领域 .....             | 397        |
| <b>参考文献 .....</b>    |                           | <b>398</b> |
| <b>中英文术语对照 .....</b> |                           | <b>423</b> |

# 第 1 章 引 论

独立成分分析(ICA)是从多元(多维)统计数据中寻找其内在因子或成分的一种方法。ICA 有别于其他方法的地方是,它寻找的是既统计独立又非高斯的成分。这里我们简要地介绍 ICA 的基本概念、应用及其估计原理。

## 1.1 多元数据的线性表示

### 1.1.1 一般统计框架

寻找多元数据的一种好的表示法,一直是统计学及相关领域中长期存在的问题。在这里,表示这个词指的是我们以某种方式对数据进行变换,使得其本质结构更显著或更容易理解。

在神经计算中,这个基本问题属于无监督学习的范畴,因为该表示必须从数据自身学习得出,而不需要从一个指导“教师”那里获取任何外部输入。获得一个合理的表示也是数据挖掘和探索性数据分析的许多技术的核心目标。在信号处理中,同样的问题出现在特征提取问题以及下面即将考虑的盲源分离问题中。

假设数据由已经得到其观测的一组变量构成。记变量数目为  $m$ , 观测数目为  $T$ 。这样可将数据记为  $x_i(t)$ , 其中, 下标取值  $i=1, \dots, m$  而  $t=1, \dots, T$ 。维数  $m$  和  $T$  可能非常大。

该问题的一种非常通用的表述如下:从  $m$  维空间到  $n$  维空间的什么函数可使得变换后的变量能够凸显原本隐藏在大量数据集中的信息。也就是说,变换后的变量应是内在因子或成分,它们描述了数据的本质结构。我们当然希望这些成分对应于数据生成过程中的某些物理原因。

大多数情况下我们只考虑线性函数,因为这样可使表示的解释与计算更加简单。这样,每个成分,比方说  $y_i$ , 可表示成观测变量的一个线性组合:

$$y_i(t) = \sum_j w_{ij} x_j(t), i = 1, \dots, n, j = 1, \dots, m \quad (1.1)$$

式中,  $w_{ij}$  是定义上述表示的某个系数。于是,原来的问题可以重新表述成如何确定系数  $w_{ij}$  的另外一个问题。利用线性代数,我们可以将公式(1.1)中的线性变换