盲信号分离

张贤达,保 铮
(西安电子科技大学雷达信号处理重点实验室,陕西西安 710071)

摘 要: 阵列处理和数据分析的一个典型问题是从混合的观测数据向量中恢复不可观测的各个源信号.盲信号分离是解决这一问题的一门新技术,近几年吸引了信号处理学界和神经网络学界众多学者的研究兴趣.本文将以独立分量分析和非线性主分量分析为主要对象,综述盲信号分离技术的理论、方法及应用等方面的发展,并作有关展望.

关键词: 信号分离;神经网络;独立分量分析;主分量分析
中图分类号: TN911.23 **文献标识码**: A **文章编号**: 0372-2112(2001)12A-1766-06

Blind Source Separation

ZHANG Xian-da, BAO Zheng
(*Key Laboratory for Radar Signal Processing, Xidian University, Xi'an, Shanxi* 710071, *China*)

Abstract: A typical problem in array processing and data analysis is to recover the unobserved source signals from their mixtures. Blind source separation (BSS) is a powerful methodology for solving this problem. In recent years, the BSS has received considerable attention from the signal processing community and the neural network community. This paper presents a survey and review on the BSS, focusing on the independent component analysis (ICA) and the principal component analysis (PCA).

Key words: source separation; neural network; independent component analysis; principal component analysis

1 引言

近几年来,盲信号分离已成为信号处理学界和神经网络学界共同感兴趣的研究热点领域,并获得了迅速的发展.简而言之,盲信号分离就是根据观测到的混合数据向量确定一变换,以恢复原始信号或信源.典型情况下,观测数据向量是一组传感器的输出,其中每个传感器接收到的是源信号的不同组合.术语"盲的"有两重含义:(1)源信号不能被观测,(2)源信号如何混合是未知的.显然,当从信源到传感器之间的传输很难建立其数学模型,或者关于传输的先验知识无法获得时,盲信号分离是一种很自然的选择.

盲信号分离的核心问题是分离(或解混合)矩阵的学习算法,它属于无监督的学习,其基本思想是抽取统计独立的特征作为输入的表示,而又不丢失信息.当混合模型为非线性时,一般是无法从混合数据中恢复源信号的,除非对信号和混合模型有进一步的先验知识可资利用.因此,在大多数的研究中,只讨论线性混合模型.

由于盲信号分离在阵列处理、多用户通信、生物医学工程等中有着重要的应用,近几年的发展极为迅速,本文将对盲信号分离的有关发展、重要理论和方法作一综述.

2 盲信号分离问题

盲信号分离可以用下面的混合方程描述:

$$\mathrm{x}(t)=\mathrm{As}(t) \tag{1}$$

式中 $\mathrm{s}(t)=[s_1(t),\cdots,s_n(t)]^T$ 为 n 个源信号构成的 n 维向量;$\mathrm{x}(t)=[x_1(t),\cdots,x_m(t)]^T$ 为 m 维观测数据向量,其元素是各个传感器得到的输出;$m\times n$ 维矩阵 A 称为混合矩阵,其元素表示信号的混合情况.式(1)的含义是 n 个源信号通过混合得到 m 维观测数据向量.盲信号分离问题的提法是:在混合矩阵 A 和源信号未知的情况下,只根据观测数据向量 $\mathrm{x}(t)$ 确定分离矩阵 W,使得变换后的输出

$$\mathrm{y}(t)=\mathrm{Wx}(t) \tag{2}$$

是源信号向量 $\mathrm{s}(t)$ 的拷贝或估计.

2.1 盲可辨识性

作为混合矩阵 A 结构信息未知的一种补偿,必须有关于源信号的某些附加假设.最基本的假设是源信号中至多只能有一个高斯信号,因为两个高斯信号是不能盲分离的[7].

由于信号传输(即信道)以及源信号知识的缺乏,盲信号分离存在两种不确定性或模糊性:分离后信号顺序排列和复振幅(幅值和初始相位)的不确定性.盲信号分离的不确定性主要表现为混合矩阵 A 的非完全辨识.既然 A 具有不确定性,所以不失一般性,假定源信号具有单位方差,即把源信号振幅的动态变化归并到混合矩阵 A 相应列的元素中.

归一化只是解决了混合矩阵 A 各元素的幅值的不确定性,各列的排列顺序和初始相位仍然保持不确定性.为了描述

收稿日期:2001-06-11;修回日期:2001-09-20
基金项目:国家自然科学基金资助项目(No.60072043)

和解决混合矩阵 A 的这些不确定性,Cardoso 等人[9]将两个矩阵的"本质相等"的概念引入到盲信号分离中.

定义 1 两个矩阵 M 和 N 称为本质相等,并记作 $M \doteq N$,若存在一矩阵 G 使得 $M \doteq NG$,其中 G 是一广义交换矩阵,并且其元素具有单位模.

因此,盲信号分离问题也可叙述为:只根据传感器输出 $x(t)$ 辨识混合矩阵 A 的本质相等矩阵与/或恢复源信号.

文献[7]证明了信号的盲可分离性:对于各个元素相互独立,并且只有一个高斯分量的信号向量 s 而言,若 $y = Cs$(其中 C 是一任意可逆矩阵)的元素相互独立,则 y 是 s 的一个拷贝.

盲信号分离的性能在很大程度上与信号的随机性质有关.这些性质由信号的高阶统计量(累积量)决定.特别地,称为峰度的四阶累积量起着重要的作用.对信号分量 $s_i(t)$,归一化峰度定义为[52]

$$\kappa_4[s_i(t)] = \frac{E\{|s_i(t)|^4\}}{E^2\{|s_i(t)|^2\}} - 3 \tag{3}$$

对于高斯信号,其归一化峰度等于零.若 $\kappa_4[s_i(t)] > 0$,则称 $s_i(t)$ 为超高斯信号;若 $\kappa_4[s_i(t)] < 0$,则 $s_i(t)$ 为亚高斯信号.亚高斯和超高斯信号另有其它定义,参见文献[52].

2.2 等变化性

除了可辨识性外,盲信号分离还应该具有等变化性.考虑对观测数据向量 x 使用一代数运算即左乘一可逆矩阵 M,得 Mx.

定义 2[9] 令 $\hat{A} = \hat{A}(x)$ 是混合矩阵 A 的某个批处理估计器.若对任意可逆矩阵 M 恒有

$$\hat{A}(Mx) = M\hat{A}(x) \tag{4}$$

则称估计器 $\hat{A}$ 是等变化的,且式(3)称作等变化条件.

假定源信号的估计为 $\hat{s}(t) = \hat{A}^{-1}X(t)$,其中 $\hat{A} = \hat{A}(x)$ 是混合矩阵 A 的等变化估计器,则容易证明

$$\begin{aligned}\hat{s}(t) &= (\hat{A}(x))^{-1}x(t) = (\hat{A}(As))^{-1}As(t)\\ &= (A\hat{A}(s))^{-1}As(t) = [\hat{A}(s)]^{-1}s(t)\end{aligned} \tag{5}$$

即是说,如果信号分离算法具有等变化性,则该算法的信号分离性能与混合矩阵(即信号传输的信道)完全无关,只决定于原始信号,这一性能称为均匀性能.很显然,"一个批处理的信号分离算法的性能与源信号如何混合无关"是我们期望盲信号分离应该具有的基本性能.

等变化性也可用另一方式叙述.令 $W(t)$ 是一盲信号分离算法求得的分离矩阵,且 $C(t) = W(t)A$ 是分离与混合的合成系统,则算法是等变化的,若 $C(t)$ 满足[9]

$$C(t+1) = C(t) - \eta(t)H(C(t)s(t))C(t) \tag{6}$$

式中 $H(C(t)s(t))$ 是矩阵乘积 $C(t)s(t)$ 的矩阵函数,它与混合矩阵 A 和分离矩阵 $W(t)$ 无关.

盲信号分离已有许多的算法,这些算法大致可分为以下三类[4]:

(1) 信号经过变换后,使不同信号分量之间的相依性(dependency)最小化.这类方法称为独立分量分析,由 Comon 于 1994 年提出[16].

(2) 利用非线性传递函数对输出进行变换,使得输出分布包含在一个有限的超立方体中;然后熵的最大化将迫使输出尽可能在超立方体中均匀散布.这类方法称为熵最大化方法,是 Bell 与 Sejnowski 于 1995 年提出的[5].

(3) 非线性主分量分析是线性主分量分析方法的推广,由 Oja 与 Karhumen 等人于 1994 年提出[25,37].

业已证明[48],熵最大化与独立分量分析是等价的.因此,下面只讨论这两类方法.

3 独立分量分析

顾名思义,独立分量分析(ICA)的基本目的就是确定线性变换矩阵 W,使得变换后的输出分量 $y_i(t)$ 尽可能统计独立.因此,ICA 是冗余压缩的一种特殊情况.

与其它优化方法需要目标函数一样,信号分离也需要目标函数.在信号分离中,常采用"对比函数"作目标函数.

定义 3 输出向量 y 的对比函数记作 $\Phi[y]$,定义为将 y 的概率密度分布集合映射为一实值函数的算子 Φ,并且映射函数 Φ 满足下列条件:

(1) 若向量 y 的元素 y_i 改变排列位置,则函数 $\Phi[y]$ 保持不变,即对所有交换矩阵 P 恒有 $\Phi[Py] = \Phi[y]$;

(2) 若 y 的元素 y_i 改变"尺度"时函数 $\Phi[y]$ 保持不变,即对所有可逆对角矩阵 D 恒有 $\Phi[Dy] = \Phi[y]$.

选择不同的对比函数,可以得到不同的独立分量分析算法.

当使用不同的神经网络时,信号分离输出的数学模型各异.例如,若使用前馈神经网络实现盲信号分离,则网络输出可写作

$$y(t) = W(t)x(t) \tag{7}$$

而若使用递归神经网络,则网络输出为

$$y(t) = [I + \hat{W}(t)]^{-1}x(t) \tag{8}$$

上述两式中,W 和 $\hat{W}$ 分别为前馈和递归神经网络的突触权矩阵.下面以前馈神经网络为分析对象.

最小互信息的基本思想是选择变换矩阵 W,使输出 $y = Wx$ 各分量之间的相依性最小化.相依性用 y 的联合概率密度函数 $p(y;W) = p(y_1,\cdots,y_n)$ 和边缘概率密度函数乘积 $\tilde{p}(y;W) = p(y_1)\cdots p(y_n)$ 之间的 Kullback-Leibler 散度表示:

$$\begin{aligned}I(W) &= D[p(y;W)\,\|\,\tilde{p}(y;W)]\\ &= \int p(y_1,\cdots,y_n)\frac{p(y_1,\cdots,y_n)}{p(y_1)\cdots p(y_n)}dy_1\cdots dy_n\end{aligned} \tag{9}$$

互信息是非负的,即 $I(W) \geq 0$;从式(9)容易看出,当且仅当 y 的各分量独立时,互信息等于零.Comon 证明了[16],$I(W)$ 是独立分量分析的对比函数,即

$$I(W) = 0 \text{ iff } W = DPA^{-1} \tag{10}$$

式中,A 为混合矩阵,D 和 P 分别是可逆的对角矩阵和任意的交换矩阵.

互信息与熵之间存在以下关系:

$$I(W) = -H(y;W) + \sum_{i=1}^{n} H(y_i;W) \tag{11}$$

式中

$$H(y;W) = -\int p(y;W)\log p(y;W)\,dy$$

为向量 y 各元素的联合熵,而

$$H(y_i;W)=-\int p(y_i;W)\log p(y_i;W)\,dy_i$$

为边缘熵. 输出向量 y 各分量的联合熵为

$$H(y;W)=H(x)+E\{\log|\det(W)|\} \tag{12}$$

令 z 是输出向量 y 的分量形式的非线性变换,即 $z_i=g_i(y_i)$,其中 $g_1=\cdots=g_n$ 为非线性变换,其目的是引出 y 各分量的高阶统计量. 向量 z 各分量的联合熵 $H(z;W)$ 可写作

$$H(z;W)=H(y;W)+\sum_{i=1}^{n}E\{\log g_i'(y_i)\} \tag{13}$$

式中 $g_i'(y_i)$ 为 $g_i(y_i)$ 的一阶导数. 将式(12)代入式(13),求得 $H(z;W)$ 相对于 W 的梯度为

$$\frac{\partial H(z;W)}{\partial W}=W^{-T}(t)-E\{\phi(y)x^T(t)\} \tag{14}$$

式中 $W^{-T}=(W^{-1})^T$,且

$$\phi(y)=[\phi_1(y_1),\cdots,\phi_n(y_n)]^T=\left[-\frac{g_1''(y_1)}{g_1'(y_1)},\cdots,-\frac{g_n''(y_n)}{g_n'(y_n)}\right]^T \tag{15}$$

下面是几种典型的独立分量分析算法.

(1) 随机梯度算法:用瞬时或随机梯度代替式(14)的真实梯度,即得到由 Bell 和 Sejnoeski[5]提出的随机梯度算法为

$$W(t+1)=W(t)+\eta(t)(W^{-T}(t)-\phi(y(t))x^T(t)) \tag{16}$$

式中 $\eta(t)$ 为学习速率或步长. 这一算法的主要缺点是:收敛速度慢,同时由于涉及分离矩阵 $W(t)$ 的求逆,一旦 $W(t)$ 在更新过程中条件数变差,算法就可能发散.

(2) 自然梯度算法:当式(16)中的随机梯度 $\frac{\partial H(z;W)}{\partial W}$ 用自然梯度 $\frac{\partial H(z;W)}{\partial W}W^TW$ 代替后,即得自然梯度算法如下:

$$W(t+1)=W(t)+\eta(t)(I-\phi(y(t))y^T(t))W(t) \tag{17}$$

式中,非线性变换函数

$$\phi_i(y_i)=\alpha_i(\kappa_3^i,\kappa_4^i)y_i^2+\beta_i(\kappa_3^i,\kappa_4^i)y_i^3 \tag{18}$$

这里 $\kappa_3^i=E\{y_{i,k}^3\}$ 及 $\kappa_4^i=E\{y_{i,k}^4\}-3$ 分别表示 y_i 的偏度和峰度,而

$$a_i(\kappa_3^i,\kappa_4^i)=-\frac{1}{2}\kappa_3^i+\frac{9}{4}\kappa_3^i\kappa_4^i \tag{19a}$$

$$\beta_i(\kappa_3^i,\kappa_4^i)=-\frac{1}{6}\kappa_4^i+\frac{3}{2}(\kappa_3^i)^2+\frac{3}{4}(\kappa_4^i)^2 \tag{19b}$$

偏度和峰度用下面的公式更新:

$$\kappa_{3,k+1}^i=\kappa_{3,k}^i-u\cdot T\cdot(\kappa_{3,k}^i-y_{i,k}^3) \tag{20}$$

$$\kappa_{4,k+1}^i=\kappa_{4,k}^i-u\cdot T\cdot(\kappa_{4,k}^i-y_{i,k}^4+3) \tag{21}$$

自然梯度算法最早是 Cichocki 等人 1994 年提出的[13],后来 Amari 等人[2,3,48]从理论上证明了它的有效性.

(3) EASI 算法:1996 年,Cardoso 和 Laheld[9]提出了一种借助独立性的等变化自适应分离(简称 EASI)算法:

$$\begin{aligned}W(t+1)=W(t)+\eta(t)[&I-\phi(y(t))y^T(t)\\&+y(t)\phi^T(y(t))-y(t)y^T(t)]W(t)\end{aligned} \tag{22}$$

相对于一般的随机梯度 $\frac{\partial H(z;W)}{\partial W}$,Cardoso 和 Laheld 将梯度 $\frac{\partial H(z;W)}{\partial W}W^T$ 称为相对梯度. 一般的(随机)梯度计算代价函数 $H(z;W)$ 相对于权矩阵的绝对变化量 $\frac{\partial H(z;W)}{\partial W}$. 与之不同,相对梯度或自然梯度则分别计算权矩阵的相对变化量 $\frac{\partial H(z;W)}{\partial W}W^T$ 和 $\frac{\partial H(z;W)}{\partial W}W^TW$. 业已证明[2],自然梯度是一般梯度在 Riemann 空间的推广,而一般梯度则是自然梯度在均匀 Euclidean 空间的特殊表现.

(4) 迭代求逆算法:与自然梯度算法和 EASI 算法不同,若取两个非线性变换函数 $f(y)=[f_1(y_1),\cdots,f_n(y_n)]^T$ 和 $g(y)=[g_1(y_1),\cdots,g_n(y_n)]^T$,则有迭代求逆算法[17]

$$W(t+1)=W(t)+\eta(t)[I-f(y(t))g(y(t))^T]W(t) \tag{23}$$

式中,非线性变换函数 $f_i(y_i)$ 和 $g_i(y_i)$ 分别取

$$f_i(y_i)=|y_i|^2y_i \tag{24}$$

$$g_i(y_i)=sgn(Re(y_i))+j\cdot sgn(Im(y_i)) \tag{25}$$

式中 $sgn(\cdot)$ 为符号函数.

容易证明,自然梯度算法、EASI 算法和迭代求逆算法都具有等变化性.

上述三种梯度算法只适合于亚高斯或超高斯信号单独存在的情况. 在亚高斯和超高斯信号同时存在时,可以使用广义 ICA 算法[31]和灵活 ICA 算法[12]等自适应算法. 一般说来,它们的计算比较复杂.

在自适应的 ICA 算法中,学习速率 $\eta(t)$ 的选择对算法的收敛起着关键的作用. 最简单的做法是使用固定的学习速率,与一般梯度算法相同,其缺点是:若学习速率大,则算法收敛快,但信号的分离精度(即稳态性能)差;反之,则稳态性能好,但算法收敛慢. 更好的做法是采用随时间变化的学习速率即变步长. 变步长又分为非自适应的和自适应的. 例如,基于退火规则的学习速率[3]和指数衰减的学习速率[47]就是两种典型的非自适应步长.

自适应步长又称学习速率的学习,是 Amari[1]早在 1967 年提出的一种方法. 下面是几种针对 ICA 提出的自适应变步长算法:

® 每个权系数都用各自的步长更新[15];

® 基于辅助变量的变步长算法[35];

® 梯度步长算法[20].

(5) 固定点算法(离线算法):在无须自适应分离的情况下,Hyvarinen[22]的固定点算法是一种快速和数值稳定的独立分量分析算法. 这种算法的对比函数定义为

$$J_G(w)=[E\{G(w^Tx)\}-E\{G(v)\}]^2 \tag{26}$$

式中 G 是一任意非二次型函数,v 是一零均值和单位方差的高斯随机变量,w 是一权向量,满足 $E\{(w^Tx)^2\}=1$. 基于"逐个分离"的原则,n 个源信号的独立分量分析变成了 n 个优化问题的求解:

$$w_i=\arg\min\sum_{i=1}^{n}J_G(w_i),\ i=1,\cdots,n \tag{27}$$

约束条件为

$$E\{(w_k^Tx)(w_j^Tx)\}=\delta_{jk} \tag{28}$$

上述优化问题的解可以使用下面的固定点算法获得:

$$w_{p+1}=w_{p+1}-\sum_{j=1}^{p}w_{p+1}^TCw_jw_j \tag{29}$$

$$w_{p+1} = w_{p+1} / \sqrt{w_{p+1}^T C w_{p+1}} \quad (30)$$

式中 $C = E\{xx^T\}$是观测数据的协方差矩阵. 式(29)和(30)需要迭代计算,直到 w_{p+1}收敛. 注意,固定点算法需要先对数据 x 进行白化.

下面是固定点算法中采用的对比函数 $G(u)$的三种选择,$g(u)$为其一阶导数:

$$G_1(u) = \frac{1}{a_1}\log\cosh(a_1 u) \quad (31a)$$

$$g_1(u) = \tanh(a_1 u) \quad (31b)$$

$$G_2(u) = -\frac{1}{a_2}\exp(-a_2 u^2/2) \quad (32a)$$

$$g_2(u) = u\exp(-a_2 u^2/2) \quad (32b)$$

$$G_3(u) = \frac{1}{4}u^4 \quad (33a)$$

$$g_3(u) = u^3 \quad (33b)$$

其中 G_1 适合于亚高斯和超高斯信号并存的一般情况;当独立的源信号为峰度值很大的超高斯信号或数值稳定性非常重要时,G_2 可能是更好的选择;分离亚高斯信号时,选用 G_3.

固定点算法的核心是式(29)中执行的压缩映射. 在 ICA 中运用压缩映射技术逐个分离信号的思想最早是文献[19]提出的,虽然该文献对基于压缩映射的自适应 ICA 算法作了介绍,但算法的有效性仍有待证明. 固定点 ICA 算法也称快速 ICA 算法,可通过文献[51]获得.

在盲信号分离的一般研究中,都假定混合矩阵 A 是非奇异的,因此信号的混合是良态的. 然而,在病态情况下,信号是通过奇异的混合矩阵混合在一起的. 最近,Li 等人[32]研究了奇异混合的源信号的盲抽取. 这一方法的基础是单个信号的盲抽取和基于压缩映射的逐个抽取.

相对于线性混合模型的线性盲信号分离,非线性混合模型下的非线性盲信号分离更加困难. 已有很多作者研究了非线性盲信号分离问题,并提出了一些算法[6,21,32,38,40,41,49]. 这些算法大致可分为以下几类:

(1)无模型方法[21,33,38]:这类方法使用 Kohonen 的自组织映射(SOM)从非线性混合数据中抽取独立的信号,其缺点是:网络结构的复杂度是指数增加的.

(2)感知器模型法:基于二层感知器模型,使用梯度下降法使互信息最小化[6]或使用自然梯度得到 BP 算法[49].

(3)径向基函数网络法:使用径向基函数神经网络逼近非线性混合的逆映射[41].

4　非线性主分量分析

一般的主分量分析(PCA)是一种信号分析技术,它能在最小重构误差的意义下提供数据的最佳表示. PCA 有两个主要的性质[45]:它求数据空间内最大方差的非相关的方向,以及给出在最小二乘意义下的最佳线性投影.

使用线性变换矩阵 Q 对观测数据向量 $x(t)+n(t)$进行变换,PCA 的目的就是使

$$y(t) = Qx(t) + Qn(t) \quad (34)$$

的主要分量 $y_i(t)$, $i = 1, \cdots, n$ 是不相关的,即

$$R_{yy} = E\{y(t)y^H(t)\} = QR_{xx}Q^H + \sigma_n^2 I = \Lambda_1 + \sigma_n^2 I \quad (35)$$

式中 Λ_1 为一对角矩阵,它由 R_{xx}的 n 个特征值(称为主特征值)组成,而噪声 $n(t)$的特征值 σ_n^2 称为次特征值. 由此可见,PCA 只使用协方差描述数据,好象这些信号都是高斯信号似的. 因此,为了对非高斯信号进行分离,需要先对观测数据作非线性变换,引入高阶统计量,然后再作主分量分析. 这就是非线性 PCA 的基本思想.

由于使用高阶统计量,所以事先对数据进行预处理,使它的一阶和二阶矩标准化是非常有用的. 令 $x(t)$为原始观测数据,则使其各分量白化成单位方差的白噪声的步骤如下:

$$\bar{x}(t) = x(t) - E\{x(t)\} \quad (36)$$

$$v(t) = E\{\bar{x}(t)\bar{x}^T\}^{-1/2}x(t) \quad (37)$$

式(36)为零均值化,式(37)使协方差矩阵为单位矩阵.

非线性 PCA 算法的代价函数为[37]

$$J_1(W) = E\{\|v - Wg(W^T v)\|^2\} \quad (38)$$

式中 $v = Qx$ 为白化后的观测数据向量,而 $g(\cdot)$为非线性变换函数.

与独立分量分析算法为梯度型(即 LMS)算法不同,非线性 PCA 可以使用递推最小二乘(RLS)算法自适应更新. 众所周知,RLS 算法的收敛性比 LMS 算法好.

为了运用 RLS 算法,文献[26]定义加权误差平方和为代价函数:

$$J_2(W(t)) = \sum_{i=1}^{t}\lambda^{t-i}\|v(i) - W(i)g(W^T(i-1)v(i))\|^2 \quad (39)$$

式中 $0<\lambda<1$ 为遗忘因子. 对非线性变换后的输出 $z(t) = g(W^T(t-1)v(t)) = g(y(t))$运用 Yang 的投影逼近子空间跟踪(PAST)算法[46],即可得到非线性 PCA 算法的 RLS 更新规则如下:

$$\begin{aligned}
z(t) &= g(W^T(t-1)v(t)) \\
h(t) &= P(t-1)z(t) \\
m(t) &= h(t)/(\lambda + z^T(t)h(t)) \\
P(t) &= \frac{1}{\lambda}Tri[P(t-1) - m(t)h^T(t)] \\
e(t) &= v(t) - W(t-1)z(t) \\
W(t) &= W(t-1) + e(t)m^T(t)
\end{aligned}$$

式中 $Tri[\cdot]$表示只取变元矩阵的上三角,其转置拷贝为下三角部分,以使 P 为对称矩阵. 上述算法的初始值可取 $P(0) = W(0) = I$,遗忘因子的典型取值为 $\lambda = 0.995$,而非线性函数 $g(y) = [g(y_1), \cdots, g(y_n)]^T$,其元素可以有不同取法,例如 $g(y) = y - \tanh(y)$对超高斯信号,或 $g(y) = \tanh(y)$对亚高斯信号.

最近,文献[50]建立了线性 PCA 与广义特征值分解之间的关系,探讨这一关系在盲信号分离中的应用可能是令人感兴趣的. 应该说,很多线性 PCA 算法都可以推广到非线性 PCA 中,但比较它们之间的性能是必要的.

由于梯度下降型盲自适应信号分离算法只使用当前时刻的观测数据对分离矩阵进行更新,所以更新量受到一定限制. 相反,在非线性 PCA 的 RLS 算法中,却是通过遗忘因子的作

用,利用了当前及以前一段时刻的观测数据来更新分离矩阵,所以更新幅度和质量都会优于梯度型算法.此外,固定点算法充分利用了所有的观测数据,因此信号分离的质量比梯度型和 RLS 型自适应算法都好.我们认为,要想从根本上提高盲信号分离的自适应 ICA 算法的性能,就应该发展其 RLS 型算法,在分离矩阵的每次更新中使用更多的观测数据.

5 展望

虽然盲信号分离在最近几年已获得了长足的发展,但是还有许多问题有待进一步研究和解决.尤其是需要发展以下算法:

® 能够在非平稳环境(其性质未知)下工作;

® 能够在奇异混合情况下工作;

® 能够在非线性混合情况下工作;

® 信号个数未知时能够工作;

® 信号个数动态变化时能够工作;

® 亚高斯和超高斯信号并存情况下能够有效工作的自适应算法;

® 独立分量分析的 RLS 算法.

此外,模糊系统理论在盲信号分离中的应用可能也是一个有前途的研究方向.当然在发展上述算法的过程中,学习算法的全局稳定性和收敛分析也需要同时考虑.

没有应用的土壤,任何一种新技术的发展都是极其有限的.和任何一种信号处理和神经网络的新技术一样,盲信号分离的发展也离不开实际的应用.事实上,盲信号分离已经获得了一些重要应用.例如,在通信[39,44],语音信号处理[42,43],包括 ECG,MEG 和 EEG 在内的生物医学信号处理[34],监视[9],图像恢复[29]中就有成功的应用.

6 结论

盲信号分离的开拓性研究起源于 Jutten 与 Herault 1991 年发表的论文[24].Comon[16]是提出盲信号分离的独立分量分析方法的第一人.正是他们的工作极大地推动了盲信号分离的研究工作,使得在短短的几年内涌现了大量的有效算法.应该说,盲信号分离不仅对信号处理的研究,而且也对神经网络理论的发展起到了积极的推动作用.

参考文献:

[1] Amari S. A theory of adaptive pattern classifiers [J]. IEEE Trans. Electronic Computers, 1967, 16: 299 - 307.

[2] Amari S. Natural gradient works efficiently in learning [J]. Neural Computation, 1998, 10: 251 - 276.

[3] Amari S, Cichocki A. Adaptive blind signal processing: Neural network approaches [J]. Proc. IEEE, 1998, 86: 2026 - 2048.

[4] Basak J, Amari S. Blind separation of uniformly distributed signals: A general approach [J]. IEEE Trans. Neural Networks, 1999, 10: 1173 - 1185.

[5] Bell A J, Sejnowski T J. An information - maximization approach to blind separation and blind deconvolution [J]. Neural Computation, 1995, 7: 1129 - 1159.

[6] Burel G. Blind separation of sources: A nonlinear neural algorithm [J]. Neural Networks, 1992, 5: 937 - 947.

[7] Cao X R, Liu R W. A general approach to blind source separation [J]. IEEE Trans. Signal Processing, 1996, 44: 562 - 571.

[8] Cardoso J F. Blind signal separation: Statistical principles [J]. Proc. IEEE, 1998, 86(10): 2009 - 2025.

[9] Cardoso J F, Laheld B. Equivariant adaptive source separation [J]. IEEE Trans. Signal Processing, 1996, 44: 3017 - 3029.

[10] Cardoso J F, Souloumiac A. Blind beamforming for non-Gaussian signals [J]. IEE Proc.-F, 1993, 140: 362 - 370.

[11] Cardoso J F, Souloumiac A. Jacobi angles for simultaneous diagonalization [J]. SIAM J. Matrix Anal. Appl., 1996, 17: 161 - 164.

[12] Choi S, Cichocki A, Amari S I. Flexible independent component analysis [J]. Journal of VLSI Signal Processing, 2000, 26: 25 - 38.

[13] Cichocki A, Unbehauen R, Moszczynski R, Rummaert E. A new on-line adaptive learning algorithm for blind separation of source signals [A]. In: Proc. ISANN94, 406 - 411, 1994.

[14] Cichocki A, Unbehauen R, Rummert E. Robust learning algorithm for blind separation of signals [J]. Electronics Lettres, 1994, 30(17): 1386 - 1387.

[15] Cichocki A, Amari S I, Adachi M, Kasprzak W. Self-adaptive neural networks for blind separation of sources [A]. In Proc. 1996 International Symp. on Circuits and Systems, May 1996, 2: 157 - 160.

[16] Comon P. Independent component analysis, a new concept? [J] Signal Processing, 1994, 36: 287 - 314.

[17] Cruces-Alvarez S, Cichocki A, Castedo-Ribas L. An iterative inversion approach to blind source separation [J]. IEEE Trans. Neural Networks, 2000, 11: 1423 - 1437.

[18] Deco G, Brauer W. Nonlinear higher-order statistics decorrelation by volume-conserving neural architectures [J]. Neural Networks, 1995, 8: 525 - 535.

[19] Delfosse N, Loubaton P. Adaptive blind separation of independent sources: A deflation approach [J]. Signal Processing, 1995, 45: 59 - 83.

[20] Douglas S C, Cichocki A. Adaptive step size techniques for decorrelation and blind source separation [A]. In Proc. 32nd Asilomar Conf. on Signals, Systems and Computers, Pacific Grove, CA, Nov. 1998, 2: 1191 - 1195.

[21] Herrmann M, Yang H H. Perspectives and limitations of self-organizing maps in blind separation of source signals [A]. In: Progress in Neural Information Processing Systems: Proc. ICONIP'96, 1996, 1211 - 1216.

[22] Hyvarinen A. Fast and robust fixed-point algorithms for independent component analysis [J]. IEEE Trans. Neural Networks, 1999, 10(3): 626 - 634.

[23] Hyvarinen A, Pajunen P. Nonlinear independent component analysis: Existence and uniqueness results [J]. Neural Networks, 1999, 12: 429 - 439.

[24] Jutten C, Herault J. Blind separation of sourcesPart I: An adaptive algorithm based on neuromimetic architecture [J]. Signal Processing, 1991, 24: 1 - 10.

[25] Karhunen J, Joutsensalo J. Representation and separation of signals using nonlinear PCA type learning [J]. Neural Networks, 1994, 7: 113 - 127.

[26] Karhunen J ,Pajunen P. Blind source separation using least-squares type adaptive algorithms [A]. In Proc. ICASSP '97 ,April 21 - 24 ,3361 - 3364 ,1997.

[27] Karhunen J ,Oja E ,Wang L ,Eigario R ,Joutsensalo J. A class of neural networks for independent component analysis [J]. IEEE Trans. Neural Networks ,1997 ,8 :486 - 504.

[28] Karhunen J ,Pajunen P ,Oja E. The nonlinear PCA criterion in blind source separation :Relations with other approaches [J]. Neural Computing ,1998 ,22 :5 - 20.

[29] Kundur D ,Hatzmakos D. A novel blind deconvolution scheme for image restoration using recursive filtering [J]. IEEE Trans. Signal Processing ,1998 ,46 :375 - 390.

[30] Lainiotis D G,Papaparaskeva. A partitional adaptive approach to nonlinear channel equalization [J]. IEEE Trans. Commun. 1998 ,46 :1325 - 1336.

[31] Leen T W ,Girolami M ,Sejnowski T J. Independent component analysis using an extended infomax algorithm for mixed sub-Gaussian and super-Gaussian sources [J]. Neural Computations ,1999 ,11 :409 - 433.

[32] Li Y,Wang J ,Zurada J M. Blind extraction of singularly mixed source signals [J]. IEEE Trans. Neural Networks ,2000 ,11 :1413 - 1422.

[33] Lin J K,Grier D G,Cowan J D. Source separation and density estimation by faithful equivariant SOM [M]. In :Advances in Neural Information Processing Systems. Cambridge ,MA :MIT Press ,1997 ,9.

[34] Makeig S ,Bell A ,Jung T P ,et al. Independence component analysis in electro-encephalographic data [M]. In :Advances in Neural Information Processing Systems ,Mozer M et al. Eds. Cambridge :MIT Press ,1996 ,8 :145 - 151.

[35] Murata N ,Muller K R ,Ziehe A ,Amari S. Adaptive on-line learning in changing environments [M]. Advances in NIPS '9 , Cambridge ,MA : MIT Press ,1997 ,599 - 605.

[36] Oja E. Principal components ,minor components ,and linear neural networks [J]. Neural Networks ,1992 ,5 :927 - 935.

[37] Oja E. The nonlinear PCA learning rule in independent component analysis [J]. Neuralcomputing ,1997 ,17 :25 - 46.

[38] Pajunen P ,Hyvarinen A ,Karhunen J. Nonlinear blind source separation by self-organizing maps [A]. In : Progress in Neural Information Processing :Proc. ICONIP '96 ,1996 ,1211 - 1216.

[39] Papadias C B ,Paulraj A. A constant modulus algorithm for multi-user signal separation in presence of delay spread using antenna arrays [J]. IEEE Signal Processing Letters ,1997 ,4 :178 - 181.

[40] Taleb A ,Jutten C ,Olympieff S. Source separation in post nonlinear mixtures: An entropy-based algorithm [A]. In: Proc. ESANN '98 , 1998 ,2089 - 2092.

[41] Tan Y,Wang J ,Zurada J M. Nonlinear blind source separation using a radial basis function network [J]. IEEE Trans. Neural Networks ,2001 , 12 :124 - 134.

[42] Tong L ,Liu R W ,Soon V C ,Huang Y F. Indeterminacy and identifiability of blind identification [J]. IEEE Trans. Circuits and Systems , 1991 ,38 :499 - 509.

[43] Torkkola K. Blind separation for audio signalAre we there yet [A] ? In Proc. 1st ICA '99 Congr. ,Assosis ,France ,Jan 1999 ,239 - 244.

[44] Veen A J ,Talvar S ,Paulraj A. A subspace approach to blind space-time signal processing for wireless communication systems [J]. IEEE Trans. Signal Processing ,1997 ,45 :173 - 190.

[45] Weingessel A ,Hornik K. Local PCA algorithms. IEEE Trans [J]. Neural Networks ,2000 ,11 :1242 - 1250.

[46] Yang B. Projection approximation subspace tracking [J]. IEEE Trans. Signal Processing ,1995 ,43 :95 - 107.

[47] Yang H H. Series updating rule for blind separation derived from the method of scoring [J]. IEEE Trans. Signal Processing ,1999 ,47 :2279 - 2285.

[48] Yang H H ,Amari S. Adaptive on-line learning algorithms for blind separation :Maximum entropy and minimum mutual information [J]. Neural Computation ,1997 ,7 :1457 - 1482.

[49] Yang H H ,Amari S ,Cichocki A. Information-theoretic approach to blind separation of sources in nonlinear mixture [J]. Signal Processing ,1998 ,64 :291 - 300.

[50] Zhang Q ,Leung Y W. A class of learning algorithms for principal component analysis and minor component analysis [J]. IEEE Trans. Neural Networks ,2000 ,11 :200 - 204.

[51] The fast-ICA MATLAB package [E]. Available at http :/ / www. cis. hut. fi/ projects/ ica/ fastica/ 1998.

[52] 张贤达 ,保铮. 通信信号处理 [M]. 北京:国防工业出版社 , 2000 ,367 - 383 ,388 - 389.

作者简介:

张贤达 男. 1946 年出生于江西省. 现任西安电子科技大学特聘教授(教育部"长江学者"奖励计划)和博士生导师,并任清华大学教授和博士生导师. 研究方向为信号处理及其在雷达和通信中的应用. 曾获国家自然科学奖 1 项、部级科技进步奖 2 项及国防发明专利 1 项. 已发表论文 80 余篇,其中 IEEE 汇刊 22 篇(含正式录用),出版学术专著 5 部:《现代信号处理》、《时间序列分析——高阶统计量方法》、《信号处理中的线性代数》、《非平稳信号分析与处理》以及《通信信号处理》.

保 铮 男. 1927 年出生于江苏省. 教授,博士生导师,中国科学院院士,中国电子科学学会会士. 研究领域为雷达信号处理及现代信号处理.