

一种新的基于归一化峭度的 ICA 算法

崔志涛, 温洁嫦, 胡南辉

(广东工业大学 应用数学学院, 广东 广州 510006)

摘要: 基于峭度的独立分量分析中, 采用梯度法寻找代价函数的极值点时, 搜索的方向对算法的收敛起着关键的作用, 对 Kuicnet 算法中原来的梯度做一个倍数替换, 得到一个新的 ICA 算法. 新算法在迭代过程中适时对梯度方向作出调整, 加快收敛速度. 数值仿真说明了算法的有效性.

关键词: 独立分量分析; 峭度; 快速 ICA 算法.

中图分类号: TN911

文献标识码: A

文章编号: 1007-7162(2007) 04-0038-04

1 ICA 数学模型

独立分量分析 (Independent Component Analysis ICA) 最早由法国学者 C. Jutten 和 J. H. Erault^[1] 于 20 世纪 80 年代首次提出, 是当今信号处理领域的一个研究热点, 在无线通讯、阵列信号处理、声纳信号处理、生物医学信号处理等领域具有广阔的应用前景^[2-4].

基本 ICA 数学模型不考虑噪声的影响, m 个观测信号 $x_i(k)$, ($i = 1, 2, \dots, m$) 由 n 个相互统计独立、零均值的未知源信号 $s_j(k)$ ($j = 1, 2, \dots, m$) 线性混合而成, 写成分量形式为

$$x_i(k) = \sum_{j=1}^n h_{ij} s_j(k), \quad (i = 1, 2, \dots, m), \quad (1)$$

矩阵的形式为

$$\mathbf{x}(k) = \mathbf{H}\mathbf{s}(k), \quad (k = 1, 2, \dots, N), \quad (2)$$

其中, $\mathbf{x}(k) = [x_1(k), x_2(k), \dots, x_m(k)]^T$, $\mathbf{s}(k) = [s_1(k), s_2(k), \dots, s_n(k)]^T$ 分别表示观测信号和未知的源信号矢量, k 是时间指标, $\mathbf{H}$ 为未知 $m \times n$ 满秩矩阵. 为使模型有确定的解, 一般还假设观测信号数目等于源信号的数目 (即 $m = n$), 源信号中至多有一个服从高斯分布. 通过学习找到一个满秩分离矩阵 $\mathbf{w}$, 使输出信号 $\mathbf{y}(k) = \mathbf{w}^T \mathbf{x}(k)$ 的各分量 y_j ($j = 1, 2, \dots, n$) 之间尽可能相互独立, 依此作为对未知源信号的一个估计. 解混模型用下式表示

$$\mathbf{y}(k) = \mathbf{w}^T \mathbf{x}(k) = \mathbf{w}^T \mathbf{H} \mathbf{s}(k) = \mathbf{G} \mathbf{s}(k), \quad (3)$$

$\mathbf{G}$ 称为全局矩阵, 各行各列只有一个非零元素 (通常各行各列只有一个元素远离 0 其余元素都接近于 0 便可认为分离成功). 分离的信号 $\mathbf{y}(k) = [y_1(k),$

$y_2(k), \dots, y_n(k)]^T$ 是源信号 $\mathbf{s}(k)$ 的一个排列, 估计信号 $\mathbf{y}(k)$ 和源信号相比有一定尺度的伸缩, 但这并不是关键所在, 源的信息主要蕴含在信号的波形当中.

2 基于峭度的 ICA 原理

在非高斯最大化的独立分量分析算法中很多采用梯度寻优原理来获得分离矩阵, 在寻优过程中梯度的方向才是关键所在, 梯度的大小对算法的收敛是没有影响的, 因此文献 [5] 中的学习参数得以消去, 使之成为一种新的 ICA 算法.

峭度 kurtosis 是衡量随机变量非高斯性强弱的一种最简单的度量, 定义一个零均值的随机变量 $\mathbf{y}$ 的峭度:

$$\text{kurt}(\mathbf{y}) = E\{\mathbf{y}^4\} - 3(E\{\mathbf{y}^2\})^2, \quad (4)$$

归一化以后具有形式:

$$\text{kurt}^*(\mathbf{y}) = \frac{E\{\mathbf{y}^4\}}{(E\{\mathbf{y}^2\})^2}. \quad (5)$$

根据峭度的数值随机变量可划分为高斯型 ($\text{kurt}(\mathbf{y}) = 0$)、非高斯型 ($\text{kurt}(\mathbf{y}) \neq 0$), 其中峭度值大于零的又称为超高斯型变量, 峭度值小于零的称为亚高斯型随机变量. 此外, 峭度的两个重要线性性质是

1) 对于常数 α 有 $\text{kurt}(\alpha \mathbf{y}) = \alpha^4 \text{kurt}(\mathbf{y})$;

2) 对于两个独立的随机变量 $\mathbf{y}_1, \mathbf{y}_2$, 有 $\text{kurt}(\mathbf{y}_1 + \mathbf{y}_2) = \text{kurt}(\mathbf{y}_1) + \text{kurt}(\mathbf{y}_2)$, 在方差相等的条件下, 峭度的绝对值越大, 随机变量的非高斯性越强.

由概率论中的中心极限定理: 几个独立的随机变量的线性组合具有比其中的任何一个变量都要强的高斯性, 通过适当的调整线性组合的系数可以使新的随机变量的高斯性变弱, 非高斯性变强, 当分离

结果的非高斯性达到最大时,便可认为实现了对混合信号的分离.在实际的操作当中首先对观测数据进行白化^[6]处理.用白化矩阵 $V=D^{-1/2}E^T$ 左乘式 (2)的两端,其中 $E=[e_1, e_2, \dots, e_n]$ 为 x 的自相关矩阵, $E=E\{xx^T\}$, e_k ($k=1, 2, \dots, n$)为矩阵 E 的特征向量, D 是由该阵的特征值构成的对角矩阵,于是式 (2)变为

$$z = Vx = VHs \tag{6}$$

白化后 z 的各分量互不相关,且 $E\{zz^T\}=1$ 为了简化,这里以二维为例说明算法的原理.考察观测数据的线性组合

$$\begin{aligned} y &= w^T z = w^T Vx = b^T x = \\ w^T V H s &= q^T s = q_1 s_1 + q_2 s_2 \end{aligned} \tag{7}$$

其中 $b^T = w^T V$, $q^T = w^T V H$. 由峭度的性质 2)可以得到

$$\text{kurt}(y) = \text{kurt}(q^T s) = q_1^4 \text{kurt}(s_1) + q_2^4 \text{kurt}(s_2). \tag{8}$$

限定输出分量 y 具有单位方差,即

$$E\{y^2\} = q_1^2 + q_2^2 = w^T E\{zz^T\} w = w^T w = \|w\|_2 = 1 \tag{9}$$

建立目标函数:

$$\begin{aligned} F(b) &= F(q) = F(w) = \text{kurt}(y) = \\ \text{kurt}(q^T s) &= q_1^4 \text{kurt}(s_1) + q_2^4 \text{kurt}(s_2), \end{aligned} \tag{10}$$

问题转换为在式 (9)的约束下,通过调整 b_1 和 b_2 来改变 q_1 和 q_2 ,使得式 (10)达到最大或最小,从而使 y 的非高斯性最大化,此时的 y 便是 s_1 或者 s_2 的一个估计.这是一个条件极值问题,约束问题的几何图形如图 1所示.

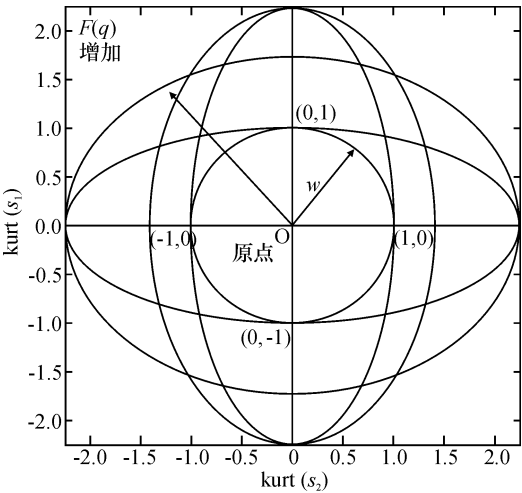


图 1 约束问题 (9)、(10)的几何图形

通过搜索寻优的方式改变 w 的方向以寻找 $F(w)$ 即 $F(q)$ 极值.从图 1可以看出 w 收敛于 $[0$

$1]^T$ 、 $[0-1]^T$ 、 $[1,0]^T$ 、 $[-1,0]^T$ 的其中之一时为目标函数的极值.在这些位置上有 $y=\pm s_1$ 或者 $y=\pm s_2$ 独立分量符号顺序的不确定性再一次得到证明.在每次寻优过程中保证 w 相互正交,此时的 q_1, q_2 也是正交向量:

$$\begin{aligned} q_i^T q_j &= w_i^T V H (V H)^T w_j = w_i^T H w_j = w_i^T w_j = 0 \end{aligned}$$

从而保证了每次分离出不同的源信号,而且从理论上讲算法能够给出每一个非高斯信号的精确估计(精确到峭度相等)^[7].

3 基于归一化峭度的独立分量分析算法

以峭度为代价函数的独立分量分析,可以归结为下面带约束的最优化问题:

$$F(w) = \sum_{j=1}^n \text{kurt}(s_j) w_j^4, \quad \text{s.t. } \|w\|_2 = 1 \tag{11}$$

S. Amari^[8]证明了如果源信号是相互独立且为非高斯分布,式 (11)的每个局部极值点都是全局极值点,即代价函数不存在伪平衡点,那么不管初始条件如何总能保证从混合信号中恢复出源信号.

在求解该问题时, Scott C. Douglas 和 S. Y. Kung^[5]以式 (5)为目标函数提出了基于神经网络的 Kuicnet算法,学习过程中采用最速上升或者最速下降法来寻找代价函数的极值

$$\begin{aligned} w(k+1) &= w(k) + \beta [E\{\frac{Z^T(k)}{m}\} - \\ &\frac{E\{|y(k)|^4\}}{E\{|y(k)|^2\}^2} E\{y(k) z^T(k)\}]. \end{aligned} \tag{12}$$

β 为学习速率,基于提取信号 $y(k)$ 的峭度值,如果 $\text{kurt}(y(k))$ 为正,则 $\beta > 0$ 若 $\text{kurt}(y(k))$ 为负,则 $\beta < 0$ $z(k)$ 为预白化信号,括号里面为式 (5)的梯度.算法的收敛与快慢依赖于良好的学习速率 β 的选取,为了使算法计算量小,收敛速度快,利用 A. Hyvarinen等在文献[9]中提出的理论

$$\begin{aligned} w(k+1) &= \frac{\nabla_w \text{kurt}(y(k))}{\|\text{kurt}(y(k))\|_2}, \nabla_w \text{kurt}(y(k)) = \\ \partial \text{kurt}(y(w)) / \partial w. \end{aligned}$$

梯度乘以一个常数不会改变方向,在式 (12) 两边同时乘以 $\frac{1}{\beta E\{y(k)^4\}}$,同时考虑 $E\{(y(k))^2\}=1, y(k)=w^T x(k)$ 得到

$$\frac{1}{\beta E\{(y(k))^4\}} \Delta w(k) = \frac{E\{(y(k))^3 z^T(k)\}}{E\{(y(k))^4\}} - w(k), \tag{13}$$

令

$$\mathbf{w}^*(k+1) = \left\{ \frac{E\{(\mathbf{y}(k))^3 \mathbf{z}^T(k)\}}{E\{(\mathbf{y}(k))^4\}} - \mathbf{w}(k) \right\}$$

(超高斯: $l = 1$; 亚高斯: $l = -1$), (14)

单位化得到

$$\mathbf{w}(k+1) = \frac{\mathbf{w}^*(k+1)}{\|\mathbf{w}^*(k+1)\|_2}. \quad (15)$$

这里利用一个向量的倍数与原来向量共线的性质以及梯度方向对算法收敛的决定作用, 将学习参数 β 消去得到一个新的基于归一化峭度的 ICA 算法, 参数 l 基于输出信号 $\mathbf{y}(k)$ 的峭度适时做出调整. 算法收敛时有 $\mathbf{w}(k+1) = \mathbf{w}(k)$, 代入式 (12) 得到

$$E\{|\mathbf{y}(k)|^2 \mathbf{y}(k) \mathbf{z}^T(k)\} - \frac{E\{|\mathbf{y}(k)|^4\}}{E\{|\mathbf{y}(k)|^2\}} E\{\mathbf{y}(k) \mathbf{z}^T(k)\} = 0 \quad (16)$$

也即 $\tilde{\mathbf{N}}_w \text{kurt}(\mathbf{y}(k)) = 0$ 此时的 $\mathbf{y}(k)$ 为代价函数 (5) 的局部极值点.

利用式 (14)、(15) 分离源信号, 具体实现为

- ① $i = 1$;
- ② 随机选取范数为 1 的初始化向量 $\mathbf{w}(0)$, 令 $k = 1$;
- ③ $\mathbf{w}_i(k) = 1 \left(\frac{E\{\mathbf{w}_i^T(k-1) \mathbf{z}^3 \mathbf{z}^T\}}{E\{\mathbf{w}_i^T(k-1) \mathbf{z}^4\}} - \mathbf{w}_i(k-1) \right)$;
- ④ 归一化 $\mathbf{w}_i(k) = \frac{\mathbf{w}_i(k)}{\|\mathbf{w}_i(k)\|_2}$;
- ⑤ 如果 $|\mathbf{w}_i^T(k) \mathbf{w}_i(k-1)|$ 收敛于 1, 输出 $\mathbf{w}_i(k)$, 否则 $k = k + 1$ 返回③继续迭代;
- ⑥ 如果 $i <$ 源信号个数, 返回②, 直到将所有源信号分离. 为保证分离信号分别为不同的源信号, 需要在④中插入一个施密特正交: $\mathbf{w}_i(k) = \mathbf{w}_i(k) - \sum_{j=1}^{i-1} \frac{\mathbf{w}_i^T(k) \mathbf{w}_j(j)}{\mathbf{w}_i^T \mathbf{w}_j(j)} \mathbf{w}_j(j)$.

4 计算机仿真

仿真 1

选取 $s_1 = \sin 20t$ $s_2 = 5 \cos t$ $s_3 = \text{rand}(1, 2000)$ - 0.2 为源信号, 采样点数为 2000 个, 随机选取混合矩阵:

$$\mathbf{A} = \begin{bmatrix} 0.0153 & 0.9318 & 0.8462 \\ 0.7468 & 0.4660 & 0.5252 \\ 0.4451 & 0.4186 & 0.2026 \end{bmatrix}$$

仿真图形见图 2~4 3 个权向量的收敛迭代次数分别为 7、8、13 次.

仿真 2

两个语音信号和一个随机噪声, 仿真见图 5~7 3 个权向量的迭代收敛次数分别为 5、6、8.

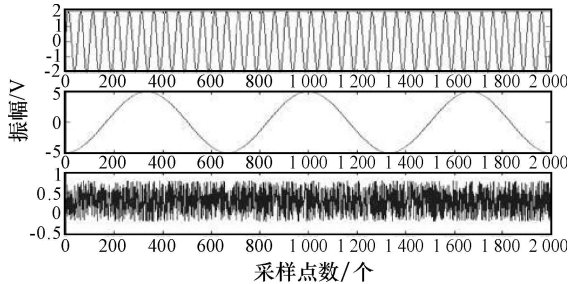


图 2 源信号的波形

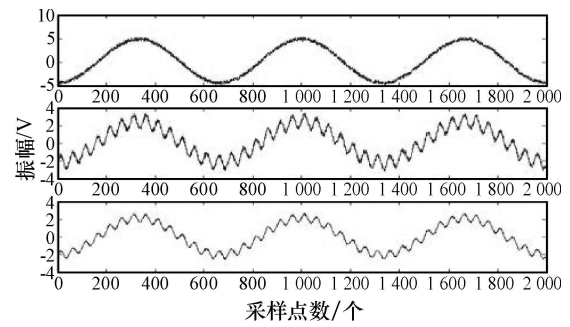


图 3 混合信号波形

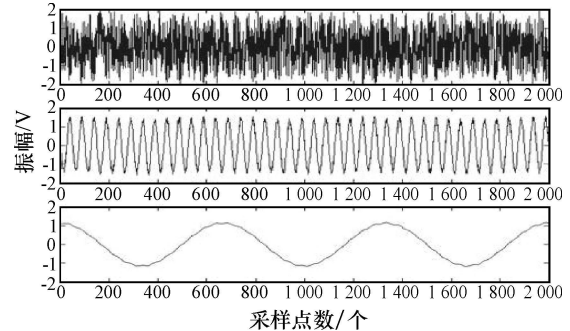


图 4 还原信号的波形

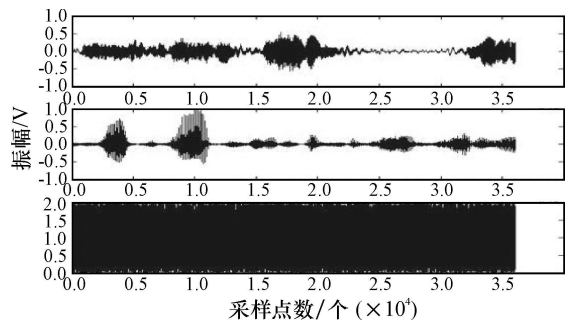


图 5 源声音信号图

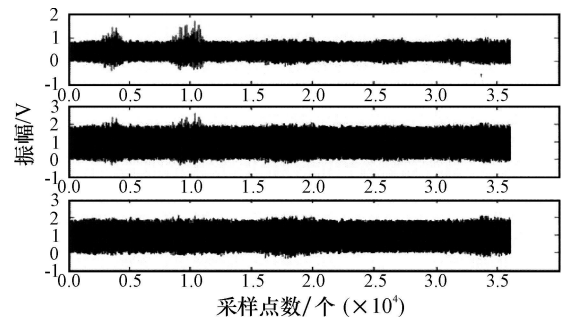


图 6 混合声音信号图

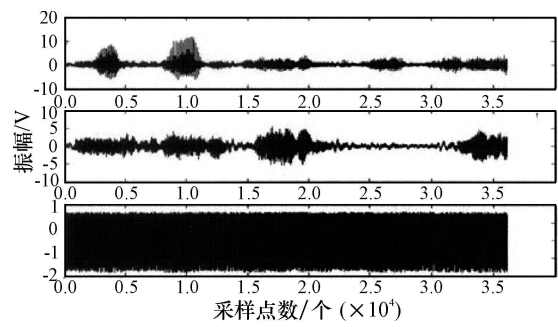


图 7 还原声音信号图

5 结语

基于峭度的独立分量分析采用非高斯极大化准则分离源信号, 能够成功的进行一大类非高斯源信号的分离, 算法收敛速度快、计算量小, 并且不论初始条件如何总能够成功的从混合信号中分离出源信号. 实际的分离过程中, 每次得到的恢复信号的比例和排列都是不同的. 算法的不足之处在于消减过程中施密特正交化带来的误差量的累积, 使得算法在有大量源信号存在的情况下会导致后续分离信号质量的迅速下降^[9]. 换句话说, 峭度并不是一种非高斯性的稳健测度^[8], 因此进一步研究求取一种稳健的基于峭度的独立分量分析算法是一项有意义的工作.

参考文献:

[1] JUTTEN C. Blind separation of sources[J]. Signal Processing, 1991, 24(1): 1-10

[2] COMON P. Independent component analysis: a new concept[J]. Signal Processing, 1994, 36(3): 287-314

[3] LEE T W. Independent component analysis using an extended infomax algorithm for mixed subgaussian and supergaussian source [J]. Neural Computation, 1999(11): 417-441.

[4] 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

[5] DOUGLAS S C, KUNG S Y. Kuicnet algorithms for blind deconvolution[C]// IEEE Workshop Neural Networks Signal Process Cambridge, U. K: IEEE, 1998 3-12

[6] HYVARNEN A, OJA E. Independent Component Analysis: Algorithms and Application[J]. Neural Networks Centre, 2000 13(4/5): 411-430

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

[8] GEORGIE P, CICHOCKI A, AMARI S. On some extensions of the natural gradient algorithm[C]// Third International Conference on Independent Component Analysis and Blind Signal Separation (ICA 2001). USA: San Diego, 2001: 581-585

[9] HYVARNEN A, KARHUNEN J OJA E. Independent component analysis [M]. New York: John Wiley and Sons, 2001

A New ICA Algorithm Based on Normalized Kurtosis

CUI Zhi-tao, WEN Jie-chang, HU Nan-hui

(Faculty of Applied Mathematics, Guangdong University of Technology, Guangzhou 510006, China)

Abstract The traditional approach to gradient to search for the cost function's extreme in the independent component analysis was based on kurtosis. The path we are looking for is crucial to the convergence of the algorithm. In the Kuicnet algorithm, after replacing the gradient of the cost function with its scaling we obtained a new ICA algorithm. Adjusting the path of the gradient timely in the iterative process, it accelerates the convergence. The simulation illustrates the effectiveness.

Key words independent component analysis; kurtosis; fast-ICA algorithm