

基于累积量极值的非线性系统盲反卷积

戴宪华

(汕头大学电子工程系, 汕头 515063)

摘 要: 提出一种新的基于累积量极值的非线性系统盲反卷积算法. 通过引入中间变量作为隐含观测量, 反卷积系统估计转化为两个子系统估计问题. 第一个子系统估计是一具有完备指导信号训练集的后非线性系统 (post nonlinear system) 辨识, 第二个子系统的估计则是一般基于累积量极值的线性系统盲反卷积. 与其它算法比较, 新算法具有明确的收敛性质和快速的收敛速度.

关键词: 非线性系统盲反卷积; 累积量极值; 隐含观测量

中图分类号: TP271⁺.61 **文献标识码:** A **文章编号:** 0372-2112 (2000) 09-0070-04

Nonlinear System Blind Deconvolution via Cumulants Extrema

DAI Xian-hua

(Dept. of Electronic Eng., Shantou University, Shantou 515063, China)

Abstract: A new cumulate extrema-based algorithm for blind deconvolution of the sandwich nonlinear system is developed. By introducing the intermediate variable as hidden variable, the learning of deconvolution system is fulfilled with an iteration learning process of two parts. The first part is a conventional learning of the post nonlinear system (PNS) with complete training sets, and the second part is in fact equivalent to blind deconvolution of a linear system. Compared with the other algorithm, the new algorithm has a faster convergence speed.

Key words: blind deconvolution of nonlinear system; cumulants extrema; hidden variables

1 引言

传统的盲反卷积^[1~3]大多基于线性系统的假设, 如线性有限长冲激响应 (FIR) 系统, 无限长冲激响应 (IIR) 系统等. 但是, 实际工程中遇到的许多系统或多或少存在非线性. 对于非线性系统, 线性盲反卷积仅是实际系统的一个线性近似解. 因此, 更为合理的反卷积系统模型应是非线性模型^[4~6, 10].

目前, 成熟的盲信号处理技术, 如基于累积量极值的盲反卷积算法, 大都是针对线性系统的盲处理^[1~3, 9]. 但是, 线性系统的盲处理技术不一定适合于非线性盲反卷积. 由于具有连续分布的随机变量将可能变换为统计独立分布的均匀或分段均匀的随机变量, 但该随机变量可能不是源信号, 即盲处理的解是非唯一的. 因此基于输出统计独立的盲反卷积也许根本不能获得原始源信号的估计^[4]. 另外, 非线性盲信号处理一般需要更精确的关于系统信号的统计分布信息, 如基于最小互信息, 最大熵等的非线性盲处理都需要系统信号的统计概率分布信息. 但实际的统计分布信息一般是未知的, 采用 Gram-Charlier 展开式来近似实际分布的算法^[4, 10]一般具有收敛速度慢、收敛结果有偏的缺点.

针对以上两个问题, 本文提出一新的盲反卷积算法. 首先, 假定实际系统可利用三明治模型描述, 从而避免了盲反卷

积解的非唯一性问题. 接着利用隐含观测量, 将三明治系统的盲学习分解为一具有完备训练集的后非线性系统 (PNS) 学习和一线性系统的盲学习. 其中 PNS 利用多局部线性化处理, 将非线性系统学习转化为由多个线性系统构成的混合统计系统的学习问题, 进而利用期望最大化 (EM) 算法获得系统估计. 线性系统的盲学习则采用一般线性系统的基于累积量极值的盲反卷积算法获得系统估计, 这是线性系统盲处理的直接推广. 由于采取以上处理方法, 新的盲反卷积具有更快的收敛速度.

2 三明治非线性反卷积模型

考虑如下非线性模型,

$$s(t) = B(z) [s(t)] = \sum_{m=0}^M b_m(t) s(t-m) \quad (1)$$

$$s(t) = N[A(z) [y(t)]] = N \left[\sum_{i=0}^I a_i(t) y(t-i) \right] \quad (2)$$

其中 $N[\cdot]$ 是一静态 (零记忆) 非线性算子, $A(z)$, $B(z)$ 为两个线性 FIR 系统的传输函数. \cdot 是一常数, 对零均值输入信号, $= 0$. $y(t)$ 是反卷积模型的输入信号, 它是激励信号 $x(t)$ 通过一线性系统产生的输出. 反卷积激励信号 $x(t)$ 被假设为独立同分布的随机序列. 反卷积模型 (1)、(2) 的所有参数记为 $\theta(t)$.

收稿日期: 1999-02-04; 修回日期: 2000-01-29

基金项目: 国家自然科学基金 (No. 69872021)

上述非线性实际上是一三明治模型,其中(2)是一 PNS 系统.实际工程中的许多系统都可用该模型描述.如:数字卫星通信的信道可等效于一输入复用线性滤波,静态非线性畸变及一输出复用线性滤波级联而成的三明治非线性系统^[5,6].

反卷积的目的就是根据信号 $y(t)$ 恢复激励 $x(t)$.与经典的反卷积不同,盲反卷积是在系统未知的情况下恢复激励 $x(t)$.因此,盲反卷积需要估计非线性系统(1)、(2).由于盲反卷积的期望输出是未知的,非线性系统(1)、(2)的估计实际上是一仅利用系统输出测量的系统估计,即盲辨识^[2].目前,大部分盲反卷积的研究都是基于线性系统模型的假设^[1~3,9].但是,线性盲处理算法不一定适合非线性系统的估计.另一方面,累积量是系统输出的非线性函数,而系统输出又是系统参数的非线性函数,因此基于累积量极值的算法是一复杂的非线性优化问题.其它的非线性盲处理一般都需要精确的统计分布信息,如基于最小互信息,最大熵等的非线性盲处理都需要系统信号的统计分布信息.但实际的统计分布信息一般是未知的,通常采用 Gann-Charier 展开式来近似实际的分布^[4,10].这类处理方法带来的缺点一是算法收敛较慢,二是算法采用实际分布的近似解可能导致收敛结果有偏.

为解决以上问题,本文提出一新的算法,将复杂系统估计转化为两个简单子系统的估计问题.具体做法是引入非线性系统(2)的期望信号作为隐含观测量,记为: $d_h(t)$.基于该隐含观测量,非线性系统(1)、(2)的估计转化为 PNS 系统(2)的估计及线性系统(1)的估计问题.其中, PNS 系统(2)的估计是一具有完备输入、输出训练信号的非线性系统参数估计问题.线性系统(1)的估计则是一般的线性系统盲反卷积问题,即以 $d_h(t)$ 为输入信号的盲反卷积问题,它可采用一般的累积量极值算法获得参数估计^[1~3,9].

几乎可以肯定,上述分解系统的估计要比原始非线性系统估计更加简单有效.但分解的关键是估计隐含观测量 $d_h(t)$.为了同时估计反卷积系统参数和隐含观测量,本文提出一新的迭代估计算法.

3 新的迭代反卷积算法

新的算法由隐含观测量估计和系统估计两部分构成.首先讨论隐含观测量的估计问题.

3.1 隐含观测量估计

在隐含观测量估计时,反卷积系统的参数 $\vartheta(t)$ 假定为已知.隐含观测量估计就是在给定系统参数的条件下,寻找最优信号估计使得反卷积系统(1)、(2)最佳地满足:由输入信号 $y(t)$ 和输出统计独立条件构成的整体系统输入、输出关系.

首先隐含观测量 $d_h(t)$ 是线性系统(1)的期望输入信号,即它作为输入信号通过线性系统(1)产生激励信号 $x(t)$ 的最优估计.那么,根据累积量极值原则,隐含观测量估计应使得^[1~3,9],

$$J_1 = K_s(4, 2, t) = \frac{s^4(t) - 3s^2(t)}{|s^2(t)|^2} = \frac{[B(z)[d_h(t)]^4 - 3[B(z)[d_h(t)]^2]^2}{|[B(z)[d_h(t)]^2|^2} \quad (3)$$

在激励源具有负归一化累积量时最小化;在激励源具有正归一化累积量时最大化.这里的优化是对隐含变量求极值.实际上,式(3)分子的统计期望即为四阶累积量,也称为 kurtosis^[1~3].当算法处于稳态,式(3)极值将等价于归一化 kurtosis 极值.式(3)中多项式 $B(z)$ 假定为已知.

另一方面,隐含观测量又是 PNS 系统(2)的期望信号,它应满足式(2)的输入、输出关系约束.根据系统(2),假定满足,

$$p(d_h(t) | y(t), \vartheta(t)) = \exp \left\{ -\frac{1}{2} [d_h(t) - N[A(z)[y(t)]]^T \cdot \frac{1}{1} [d_h(t) - N[A(z)[y(t)]] \right\} \quad (4)$$

其中 $\frac{1}{1}$ 为协方差矩阵.

根据最大似然算法,隐含观测量可通过如下似然函数最大化获得,

$$J_2 = -\frac{1}{2} [d_h(t) - N[A(z)[y(t)]]^T \frac{1}{1} [d_h(t) - N[A(z)[y(t)]] \quad (5)$$

基于反卷积系统输入、输出关系,导出了两个估计隐含观测量的准则.实际上,隐含观测量的估计应同时满足以上两个准则,使得整体非线性系统满足输入、输出关系.结合准则函数(4)、(5),隐含观测量 $d_h(t)$ 可利用以下优化算法估计:当激励源具有负归一化累积量时, $d_h(t)$ 可通过 $J_3 = J_1 + J_2$ 最小化估计;在激励源具有正归一化累积量时, $d_h(t)$ 可通过 $J_4 = J_1 - J_2$ 最大化估计.其中 $\frac{1}{1}$ 为常数,用以调整以上两个优化准则对隐含变量估计的贡献.隐含观测量估计记为: $\bar{d}_h(t)$.

3.2 后非线性系统估计

基于隐含观测量估计 $\bar{d}_h(t)$, PNS 系统(2)的估计实际上是一给定输入、输出完备训练集的非线性系统参数估计问题,训练集为: $y(t)$, $\bar{d}_h(t)$.系统估计包括:静态非线性算子 $N[\cdot]$ 和 $a_i(t)$ 估计.

对 PNS 系统(2),已有文献^[5]讨论它的估计.这里,提出一新的具有更快收敛速度的估计算法.

3.2.1 新的统计模型

首先,当静态非线性算子 $N[\cdot]$ 取为连续算子时,在定义域的一小邻域,非线性算子可利用一线性函数近似.而在算子的不同小邻域,算子则可利用多个线性函数近似.因此, PNS 系统(2)可近似为

$$s(t) = \sum_{k=1}^K s_k(t) \quad (6)$$

$$s_k(t) = N(s_k^0) + \frac{\partial N(u(t))}{\partial u(t)} \bigg|_{u(t)=s_k^0} [u(t) - s_k^0] = N(s_k^0) + l(s_k^0) [u(t) - s_k^0] \quad (7)$$

$$u(t) = \sum_{i=0}^{\infty} a_i(t) y(t-i) \quad (8)$$

其中 $k = \begin{cases} 1, & u(t) \in A_k \\ 0, & \text{Otherwise} \end{cases}$, A_k 是一中心点在 s_k^0 处的小邻域,满足 $\bigcup_{k=1}^K A_k = A$, A 表示 $N[\cdot]$ 的定义域.参数 $N(s_k^0)$, $l(s_k^0)$ 表示 $N[\cdot]$ 在 s_k^0 处的函数值和一阶导数.

非线性系统经如上的近似后,非线性算子可利用参数 $N(s_k^0)$, $l(s_k^0)$ 采用多项式曲线拟合方法估计.因此, PNS 系统估

计转化为参数 $N(s_k^0)$, $l(s_k^0)$ 和 $a_i(t)$ 的估计问题. 分析上线性近似, PNS 系统在其定义域中为多个线性系统近似. 该近似可认为是多个线性系统的并行组合, 每个线性系统定义在对应的区域解决一函数逼近问题. 某一时刻, 每个线性系统都可单独作为非线性系统的近似. 这一多系统近似实质上构成一混合专家网络 (mixture of experts, ME)^[7,8]. 因此, 可以采用统计混合专家网络来描述上系统近似.

为了利用 ME 描述上系统, 引入一个新的隐含变量 $\vartheta(t)$, 它是一取值于 $\{1, \dots, K\}$ 的离散随机变量, 以概率 $p_k = \text{prob}\{\vartheta(t) = k\}$ 对应于 K 个线性近似系统 (7)、(8). 利用新隐含变量, 近似系统 (6) ~ (8) 可利用下统计模型表示^[7,8],

$$p(\bar{d}_h(t), \vartheta(t) | y(t), \vartheta(t)) = \prod_{k=1}^K p(\bar{d}_h(t) | s_k(t), \vartheta(t)) \prod_{k=1}^K p_k(t) \quad (9)$$

其中 $p(\bar{d}_h(t) | s_k(t), \vartheta(t)) = B \exp(-(\bar{d}_h(t) - s_k(t))^2)$, B 为常数. $p_k(t) = \begin{cases} 1, & \vartheta(t) = k \\ 0, & \text{其他} \end{cases}$ 及 $\sum_{k=1}^K p_k(t) = 1$. 条件 $p_k(t) = 1$ 意味着 $\bar{d}_h(t)$ 可作为第 k 个线性近似系统的期望信号.

对应式 (9) 的对数似然函数为,

$$Q(\bar{d}(t), \vartheta(t) | y(t), \vartheta(t)) = - \sum_{k=1}^K \ln p_k(t) - \sum_{k=1}^K (\bar{d}_h(t) - s_k(t))^2 + C \quad (10)$$

其中 C 是一常数.

3.2.2 基于 EM 算法的参数估计

通过以上分析, PNS 系统估计转化为对数似然函数描述的参数估计问题. 基于似然函数最大化算法 (MLA), 可以获得非线性系统的参数估计. 但是, 由于上面的对数似然函数中包含隐含观测量, 因此不能直接用于参数估计. 上述统计模型实际上是一不完全观测的统计学习问题. 在不完全观测的学习中, 比较有效的参数估计算法是期望最大化算法 (EM)^[7,8].

EM 算法包含两步, 第一, 基于系统估计, 获得隐含观测量估计, 进而获得似然函数的统计期望估计; 第二, 利用隐含观测量估计, 采用 MLA 估计系统参数.

在上统计模型中, 隐含变量 $\vartheta(t)$ 是一离散随机变量. 为计算对数似然函数的期望, 需要计算 $\vartheta(t)$ 对应的概率. 根据统计模型 (9), 概率估计为,

$$h_k(t) = p(\vartheta(t) = k) = \exp(-(\bar{d}_h(t) - s_k(t))^2) / \sum_{n=1}^K \exp(-(\bar{d}_h(t) - s_n(t))^2) \quad (11)$$

基于概率估计, 对数似然函数的统计期望为

$$Q(\vartheta | \vartheta(t)) = - \sum_{k=1}^K h_k(t) (\bar{d}_h(t) - s_k(t))^2 + C \quad (12)$$

上述对数似然函数中不再包含隐含变量, 成为一完全观测的似然函数. 它与一般线性系统的高斯模型十分相似. 因此, 可采用最大化期望的参数估计算法估计系统参数.

令: $y(t) = [y_0(t), \dots, y_1(t)]^T$, $s_k(t) = [N(s_k^0), \dots, N(s_k^1)]^T$, $l(s_k^0) = [l(s_k^0), \dots, l(s_k^1)]^T$. 基于式 (12) 最大化, 获得参数估计为,

$$y(t+1) = y(t) - \mu Y(t) \sum_{k=1}^K h_k(t) (\bar{d}_h(t) - s_k(t) - l(s_k^0)) \quad (13)$$

其中 $Y(t) = [y(t), \dots, y(t-1)]^T$. 这是一线性系统的参数估计公式, 即 PNS 系统的参数估计转化为 EM 算法后, 已是一线性系统的参数估计.

与传统的非线性系统局部线性化参数估计方法比较, 新的参数估计算法以统计概率求和的形式利用了静态非线性算子的多点局部线性化信息. 多点局部线性化信息以竞争学习的方式综合估计系统参数. 只有当 $h_k(t) = 1$, $h_j(t) = 0$; $j \neq k$ 时, 新的 EM 参数估计算法才退化为传统的非线性系统局部线性化参数估计方法. 因此可以断言, 新算法将在收敛速度、多局部解等问题上优于传统的非线性系统局部线性化参数估计方法. 由于篇幅限制, 不再讨论.

基于同样的原理, 可以获得参数 $s_2(t)$, $s_3(t)$ 的估计. 利用 $s_2(t)$, $s_3(t)$ 可以重建算子 $N[\cdot]$. 设 $N(s) = \sum_{p=1}^P p s^p$, 那么其系数可由下方程获得估计,

$$\begin{cases} \sum_{p=1}^P p (s_k^0)^p = N(s_k^0), \\ \left. \frac{\partial N(s)}{\partial s} \right|_{s=s_k^0} = l(s_k^0), \end{cases} \quad k=1, \dots, K \quad (14)$$

3.3 线性系统盲估计

在获得隐含观测量估计后, 线性系统 (1) 的估计实际上就是以 $\bar{d}_h(t)$ 为输入的线性系统盲反卷积. 利用输出累积量极值算法可以获得线性系统 (1) 的估计^[1~3].

3.4 算法综合

基于以上讨论, 新的迭代算法可以总结为表 1.

表 1 非线性系统盲反卷积

给定系统参数的初始猜测 $\vartheta(0)$, 令 $t=0$ 开始如下迭代.
步 1: 基于 $\vartheta(t)$, 估计隐含变量 $d_h(t)$
步 2: 利用 $\bar{d}_h(t)$, 估计 PNS 系统 (2) 和线性系统 (1), 进而获得 $\vartheta(t+1)$
最后, 基于非线性系统估计, 获得反卷积输出.

4 模拟实验

在实验中, 假设产生观测变量 $y(t)$ 的非线性系统为三明治系统, 其中输入、输出线性系统取为低通滤波器, 静态非线性算子取为 $\tanh(\cdot)$. 系统的激励 $x(t)$ 取为 4 电平脉冲幅度调制 (4-PAM) 信号, 即以等概率取值于集合 $\{0, 1, 2, 3\}$. 实验反卷积模型取形如式 (1)、(2) 的三明治系统. 其中, 多项式 $A(z)$, $B(z)$ 的阶次分别取为 12 和 14. 反卷积模型的静态非线性算子利用多项式拟合方法估计.

实验结果图 1 为系统的激励 $x(t)$ 、系统输出 $y(t)$ 和反卷积系统输出信号. 图 2 为静态非线性算子及其估计结果. 图 3 为输入、输出线性低通滤波器幅度和相位估计. 图 4 为反卷积系统估计的学习曲线, 它们是 40 次独立的计算机模拟实验结果的平均. 从实验结果图可看出, 新的反卷积算法可以有效地恢复激励信号. 与最小互信息反卷积算法比较, 新的算法具有更快的收敛速度.

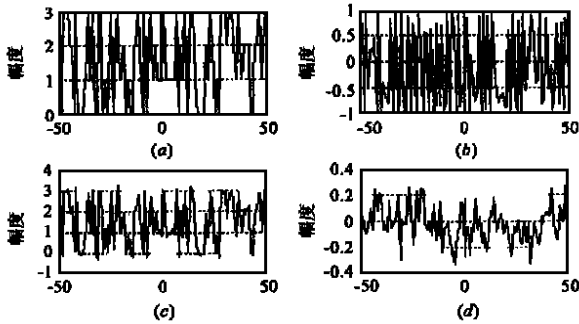


图1 反卷积信号:(a)激励信号(b)非线性系统输出信号
(c)反卷积输出;(d)误差信号

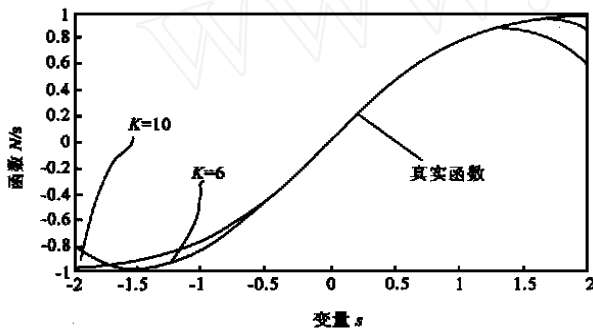


图2 静态(零记忆)非线性算子及其估计

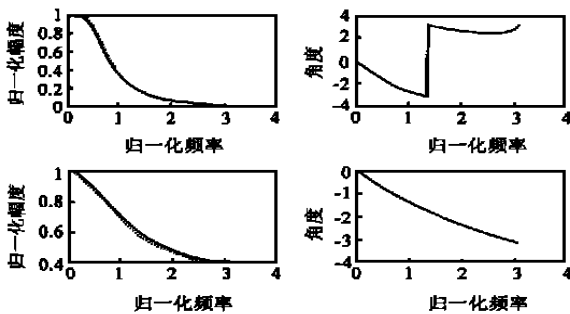


图3 三名治非线性系统输入、输出线性系统的幅频、相频响应
其中实线为真实系统响应,虚线为反卷积系统逆系统的
频响

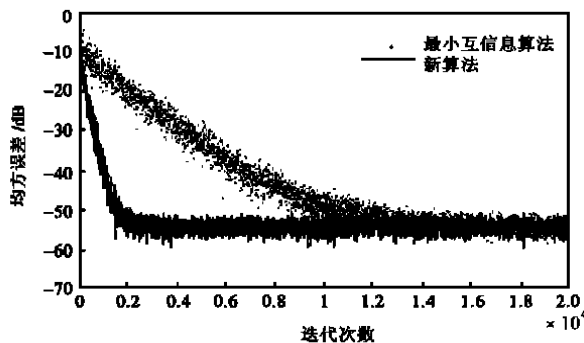


图4 非线性系统学习曲线,它们是40次独立实验均方误差平均
的结果.其中,最小互信息算法参考文献[4]

5 结论

论文提出了一种新的基于累积量极值的反卷积算法.由于新算法将盲反卷积系统估计分解一具有完备训练集的非线性估计和一线性系统的盲反卷积,以及非线性系统估计采用线性化处理,进而转化为统计学习问题.因此新的算法具有更快的收敛速度.新的盲反卷积算法也可直接推广应用于解决其它非线性系统的盲反卷积问题.

参考文献:

- [1] James A Cadzow. Blind deconvolution via cumulants [J]. IEEE Signal Processing Magazine, May, 1996.
- [2] J. M. Mendel. Tutorial on higher-order statistics (spectra) in signal processing and system theory: theoretical results and some applications [J]. Proc. of the IEEE, 1991, 79 (3).
- [3] J. Karhunen, Erkki Oja. A class of neural networks for independent component analysis [J]. IEEE Trans. on Neural Networks, May 1997, 8.
- [4] Howard Hua Yang, Shun-ichi Amari. Information-theoretic approach to blind separation of sources in non-linear mixture [J]. Signal Processing, 1998, 64.
- [5] H. W. Chen. Modeling and identification of parallel nonlinear systems: structural classification and parameter estimation methods [J]. Proc. of the IEEE, 1995, 83 (1).
- [6] Mohamed Ibnkahia. Neural network modeling and identification of non-linear channels with memory: algorithms, applications, and analytic models [J]. IEEE Trans. on Signal Processing, 1998, 46 (5).
- [7] Michael I Jordan. Convergence results for the EM approach to mixtures of experts architectures [J]. Neural Networks, 1995, 8 (9).
- [8] Shun-ichi Amari. Information geometry of the EM and em algorithms for neural networks [J]. Neural Networks, 1995, 8 (9).
- [9] Aapo Hyvarinen Erkki Oja. Independent component analysis by general nonlinear hebbian-like learning rules [J]. Signal Processing, 1998, 64.
- [10] Aniss Taleb. Source separation in post nonlinear mixtures: an entropy-based algorithm [J]. Proc. of ICASSP 1998: 2089.

作者简介:



戴宪华 1992 年于东南大学获博士学位, 1993 年 - 1994 年为华南理工大学博士后, 1995 年调入汕头大学工作, 现为汕头大学电子工程系教授. 主要研究兴趣: 非线性自适应信号处理、盲信号处理、通信信道均衡等.