

非高斯系统的建模方法研究*

高 江

戴冠中

(西安交通大学系统工程研究所·西安,710049) (西北工业大学自动控制系·西安,710072)

摘要: 近年来,高阶统计信息理论被广泛应用于非高斯系统的建模问题中. 由于高阶统计信息方法特殊的信息结构,难以充分利用具体问题中的统计分布信息,从而对该方法的应用效果产生了影响. 针对这一问题,本文提出一种综合应用高阶统计信息理论与非高斯系统的似然函数指标的建模方法,为在非高斯系统建模问题中充分利用问题所提供的统计信息提供了新的途径. 仿真研究和实际应用取得了令人满意的效果.

关键词: 系统建模; 高阶统计信息; 似然函数; 非高斯系统

1 引 言

近年来,具有非高斯特性系统的建模问题受到人们的广泛重视. 传统的基于高斯统计分布的系统建模理论存在着许多缺陷,如不能辨识系统相位,难以应用于某些非线性系统中等^[1,2]; 当考虑了系统的非高斯特性后,这些问题中的许多可以得到完满的解决. 高阶统计信息(HOS)理论是近年来发展起来的用于研究非高斯系统的一种较完善的理论. 该理论应用非高斯随机变量的高阶统计特性对非高斯系统进行统一描述和处理,已得到了许多有效的估计算法,并被广泛应用于非高斯系统的建模问题中^[1~4].

通过对高阶统计信息理论的分析可以发现,该理论所研究的是非高斯随机变量的高阶统计相关特性,而并不涉及随机变量的具体统计分布. 这样做的优点是所得到的算法具有一定的普遍性,不受具体问题的限制;但也带来一些问题,如对算法性能的分析只能采用渐近指标,用渐近统计分布特性来研究^[3],特别是由于未充分考虑具体问题中的统计信息,必然会对其应用效果产生一定的影响.

基于上述的观点,本文提出一种综合应用高阶统计信息理论与问题本身统计分布信息的建模方法. 具体研究了一类具有 Bernoulli-Gaussian 输入的非高斯系统,推导了这类系统的似然函数,给出了综合应用 HOS 方法与系统似然函数指标的建模方法. 这种方法可以充分利用非高斯随机变量的统计分布信息,从而提高了系统建模的精度. 仿真与实际应用证实了这种方法的有效性.

2 系统描述

考虑如下具有 Bernoulli-Gaussian 输入的线性随机系统,

$$z(k) = y(k) + n(k) = h(k) * u(k) + n(k) = \sum_{i=1}^k h(k-i)u(i) + n(k). \quad (1)$$

其中 $h(k)$ 为系统的冲击响应,用 ARMA 模型表示为

$$H(z) = \frac{\sum_{i=0}^q b_i z^{-i}}{1 + \sum_{i=1}^p a_i z^{-i}}, \quad q \leq p. \quad (2)$$

* 中国博士后科学基金资助项目.
本文于 1995 年 11 月 27 日收到. 1996 年 10 月 14 日收到修改稿.

输入信号 $u(k)$ 是平稳的 Bernoulli-Gaussian 白噪声序列,

$$u(k) = q(k)r(k) + w(k). \quad (3)$$

其中 $q(k)$ 是 Bernoulli 0-1 序列, 由概率参数 λ 描述

$$Pr[q(k)] = \begin{cases} \lambda, & q(k) = 1, \\ 1 - \lambda, & q(k) = 0. \end{cases} \quad (4)$$

$r(k), w(k), n(k)$ 均为零均值 Gaussian 白噪声序列, 且

$$E\{r^2(k)\} = v_r, \quad E\{w^2(k)\} = v_w, \quad E\{n^2(k)\} = v_n. \quad (5)$$

$r(k), q(k), w(k), n(k)$ 相互独立. 系统的模型如图 1 所示.

为便于处理, 将上述系统用向量形式表示为

$$z = WQr + Ww + n. \quad (6)$$

式中

$$W = \begin{bmatrix} h(0) & 0 & \cdots & 0 \\ h(1) & h(0) & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ h(N-1) & h(N-2) & \cdots & h(0) \end{bmatrix},$$

$$r = \text{col}[r(1), r(2), \dots, r(N)], \quad w = \text{col}[w(1), w(2), \dots, w(N)],$$

$$Q = \text{diag}[q(1), q(2), \dots, q(N)], \quad n = \text{col}[n(1), n(2), \dots, n(N)].$$

N 为观测数据的长度.

3 系统的似然函数及优化方法

3.1 似然函数的推导

因为 $r(k), w(k), n(k)$ 均服从高斯分布, 且相互独立, 用[5]的方法可推得以 ARMA 参数 θ 为参量的似然函数为

$$L[\theta|z] = p[z|\theta] = (2\pi)^{-\frac{N}{2}} |\Omega|^{-\frac{1}{2}} \exp\left\{-\frac{1}{2} z' \Omega^{-1} z\right\}. \quad (7)$$

式中

$$\Omega = v_r W Q W' + v_w W W' + v_n I. \quad (8)$$

对数似然函数为

$$l_1\{\theta|z\} = -\frac{1}{2} \ln|\Omega| - \frac{1}{2} z' \Omega^{-1} z - \frac{N}{2} \ln(2\pi). \quad (9)$$

上述似然函数的最大值对应于 θ 的极大似然 (ML) 估计. 因此, 可将 (9) 式做为优化指标, 对 ARMA 参数进行优化.

注意到上述 (8) 式中 $q(k)$ 仍未知. 在实际算法的实现中需要附加对 $q(k)$ 的估计环节, $q(k)$ 可由二值检测的方法得到[7].

从 (9) 式中略去与 θ 参数无关的量, 得到 ARMA 参数辨识的如下优化指标:

$$l\{\theta|z\} = -\frac{1}{2} \ln|\Omega| - \frac{1}{2} z' \Omega^{-1} z. \quad (10)$$

3.2 优化方法的实现及性能

由于 Ω 矩阵的维数为 N , 直接基于 (10) 式的优化算法运算量很大. 设系统 (1) 的状态空间实现为 $\{A, B, C\}$, 由极大似然估计和 Kalman 滤波理论[6], 上述指标函数可以等价地由系统

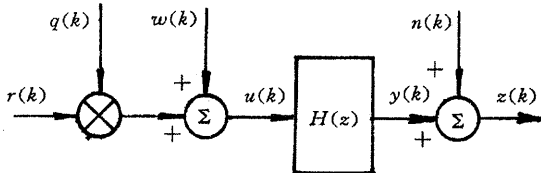


图 1 系统模型

(1)的 Kalman 滤波新息序列 $\tilde{z}(k|k-1)$ 表示如下:

$$l\{\theta|z\} = -\frac{1}{2} \sum_{k=1}^N \left[\frac{\tilde{z}^2(k|k-1)}{\eta(k)} + \ln(\eta(k)) \right], \quad (11)$$

其中 $\eta(k)$ 为新息 $\tilde{z}(k|k-1)$ 的方差. 使 $l\{\theta|z\}$ 达到极大的 θ 的估计就是使下述指标

$$J\{\theta|z\} = \sum_{k=1}^N \left[\frac{\tilde{z}^2(k|k-1)}{\eta(k)} + \ln(\eta(k)) \right] \quad (12)$$

达极小. $J(\theta|z)$ 对于 ARMA 参数 θ 的梯度向量及海赛阵的计算可以通过 Kalman 灵敏度滤波完成^[6]. 这样, (10)式指标的集总参数优化问题便转化为由 Kalman 灵敏度滤波实现的状态空间中的递推优化, 大大减小了运算量.

我们使用 Marquart-Levinberg 算法^[8]对 ARMA 参数进行寻优, 算法的实现过程是, 用 Kalman 灵敏度滤波器计算出上述优化指标对 ARMA 参数的梯度向量 $g(\theta_i)$ 和海赛阵 $H(\theta_i)$, 按以下方式对参数进行迭代优化:

$$\theta_{i+1} = \theta_i - [H(\theta_i) + D_i]^{-1}g(\theta_i).$$

其中, D_i 为一合适的对角阵, 用以保证 $H(\theta_i) + D_i$ 正定.

根据极大似然估计理论^[7], 在上述算法的性能方面有如下结论:

1) 令指标函数 $J(\theta|z)$ 的期望值为

$$E[J(\theta|z)] = E \left\{ \sum_{k=1}^N \left[\frac{\tilde{z}^2(k|k-1)}{\eta(k)} + \ln(\eta(k)) \right] \right\}, \quad (13)$$

$$C(N) = \frac{1}{N^2} E[J(\theta|z) - E[J(\theta|z)]]^2, \quad (14)$$

则有
$$C(N) \leq \frac{2b^2(1+c^2)}{Nv_n^2(1-c^2)}. \quad (15)$$

其中, b, c 为常数, $0 \leq c < 1, N$ 为观测数据的长度.

由 (15) 式可见, 指标函数 $J(\theta|z)$ 均方收敛于 $E[J(\theta|z)]$, 且方差与 N 成反比. 这保证了随着观测数据量的增加, 似然函数较快地收敛于其期望值.

2) 以 Bernoulli-Gaussian 序列为输入的系统, 其 ARMA 参数对应于似然优化指标期望值 $E[J(\theta|z)]$ 的唯一最小点.

由上述结论可知, 随着 N 的增大, $J(\theta|z)$ 较快地收敛于 $E[J(\theta|z)]$, 系统参数是可辨识的.

4 高阶统计信息建模方法

4.1 AR 参数辨识

一般地, 用 $C_{k,x}(\tau_1, \tau_2, \dots, \tau_{k-1})$ 表示随机过程 $x(t)$ 的 k 阶矩, $k > 2$. 上述 Bernoulli-Gaussian 序列 $u(k)$ 的 k 阶矩为

$$C_{k,u}(\tau_1, \tau_2, \dots, \tau_{k-1}) = \begin{cases} \gamma_{k,u}, & \tau_1 = \tau_2 = \dots = \tau_{k-1} = 0, \\ 0, & \text{else.} \end{cases} \quad (16)$$

令
$$C_{k,x}(m, n) = C_{k,x}(m, n, 0, \dots, 0), \quad (17)$$

根据高阶统计信息理论及 ARMA 模型的性质, 系统的 AR 参数满足如下正则方程:

$$C_{k,y}(m, n) + \sum_{i=1}^p a_i C_{k,y}(m-i, n) = \gamma_{k,u} \sum_{i=0}^{\infty} h^{k-2}(i) h(i+n) b_{i+m} = 0, \quad m > q. \quad (18)$$

其中 $C_{k,y}(m, n)$ 为由 (17) 式定义的 $y(k)$ 的 k 阶矩. 为保证算法的可辨识性引入以下结论.

引理 1^[4] 若在 (18) 式中选取 $m = q+1, \dots, q+p+M, n = q-p, \dots, q, M \geq 0$, 则可

以保证算法对任意相位系统的可辨识性.

将 $m = q + 1, \dots, q + p + M, n = q - p, \dots, q, M \geq 0$ 代入(18)式并展开得线性方程组

$$\bar{C}_{k,y}(p, q, M)\bar{a} = 0. \quad (19)$$

其中 $\bar{a} = \text{col}(a_p, a_{p-1}, \dots, a_1, 1)$. 由上述可辨识性引理知上式左端矩阵 $\bar{C}_{k,y}(p, q, M)$ 的秩为 p . 由于系统的阶未知, 在算法实施中可用足够大的 \hat{p}, \hat{q} 替代上面的 p, q , 得到 $\bar{C}_{k,y}$, 再由矩阵 $\bar{C}_{k,y}$ 的有效秩确定模型 AR 部分的阶 p .

由于系统(1)中 $n(k)$ 为零均值 Gaussian 白噪, 由高阶矩的性质可知下式成立,

$$C_{k,z}(\tau_1, \tau_2, \dots, \tau_{k-1}) = C_{k,y}(\tau_1, \tau_2, \dots, \tau_{k-1}). \quad (20)$$

由(19)与(20)式就可以得到 AR 参数的估计. 由于 Bernoulli-Gaussian 序列的对称分布特性, 我们用观测数据的四阶矩构造上述方程组. 为保证上述算法的精度与数值稳定性, 使用了文[10]的奇异值分解方法求解线性方程组(19).

4.2 MA 参数辨识

利用所得到的 AR 参数, 求出残差序列:

$$\tilde{z}(k) = z(k) - \sum_{i=1}^p a_i z(k-i). \quad (21)$$

MA 参数用优化方法^[1]得到. 优化指标为残差序列的四阶矩估计与按模型参数计算的四阶矩的平方误差

$$N\{b_i | \hat{C}_{4,\tilde{z}}\} = \|\hat{C}_{4,\tilde{z}} - C_{4,\tilde{z}}\|_2. \quad (22)$$

具体表示为

$$N\{b_i | \hat{C}_{4,\tilde{z}}\} = \sum_R \left[\sum_{k=0}^q b_k b_{k+i} b_{k+m} b_{k+n} - \hat{C}_{4,\tilde{z}}(l, m, n) \right]^2. \quad (23)$$

其中 $b_i, i = 0, \dots, q$ 为系统的 MA 参数, R 表示所有非冗余的延迟 l, m 与 n . 系统 MA 部分的阶用阈值检测的方法^[9]确定.

这样, 我们就用高阶统计信息方法得到 ARMA 参数的估计. 该参数用做以 $J(\theta|z)$ 为优化指标的辨识算法的初始模型参数.

5 仿真实例

仿真对象取自具有任意相位的地震子波辨识问题. 系统模型为

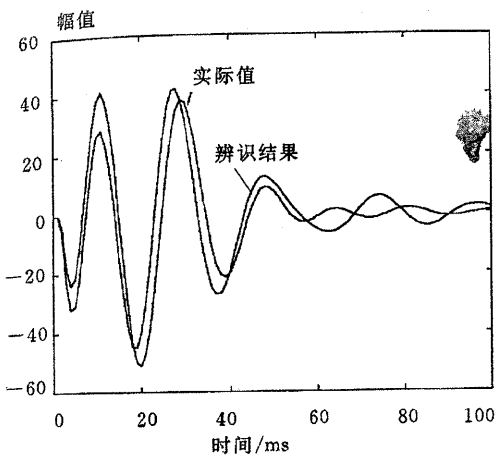
$$H(z) = \frac{-4.0907z^{-1} - 3.6666z^{-2} + 21.5072z^{-3} - 14.0035z^{-4}}{1 - 3.5887z^{-1} + 4.9960z^{-2} - 3.1890z^{-3} + 0.7895z^{-4}}.$$

输入为 Bernoulli-Gaussian 序列, 其中 $v_r = 1, v_w = 0.02, \lambda = 0.15$. 观测噪声为高斯白噪声, $v_n = 10$, 取 800 个观测信号对系统建模. 首先应用上述的高阶矩方法得到系统 ARMA 参数的初始估计, 结果如图 2(a) 所示. 再以 $J(\theta|z)$ 为优化指标对参数进行优化, 得到最终的模型估计. 由似然函数优化得到的最终系统模型估计与真实系统十分接近. 而若仅用高阶矩方法, 要达到同样的精度则需要长得多的观测信号序列. ARMA 参数的收敛过程如图 3 所示.

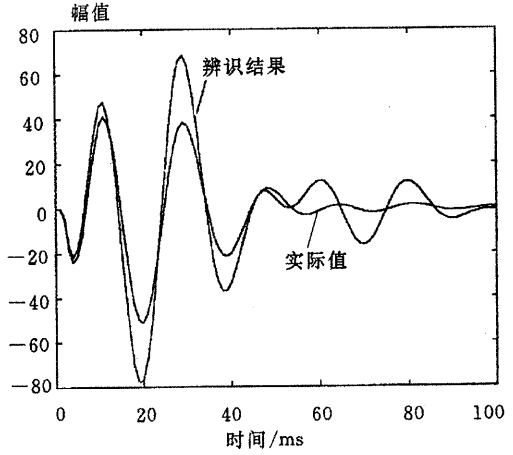
6 结 语

针对一类非高斯系统, 给出了一种基于高阶统计信息理论和问题本身特定的似然指标函数的优化建模方法, 并分析了其性能. 这为在实际问题中如何有效利用高阶统计信息理论和问题本身的信息以达到最好的辨识效果提供了一个思路.

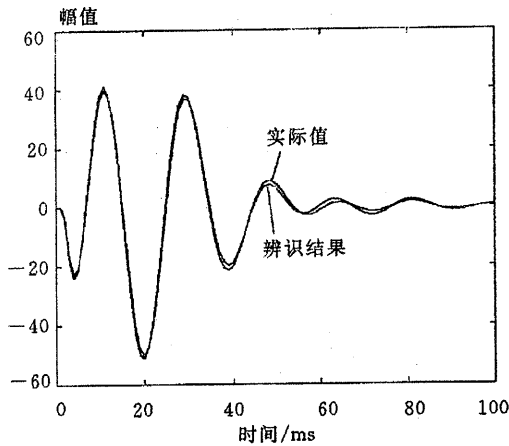
应当指出的是, 这里所给出的似然指标函数仅是综合问题中统计信息的一种方法, 其它的统计工具也可以担当这样的角色. 随着高阶统计信息理论在实际问题中应用的不断深入, 必将



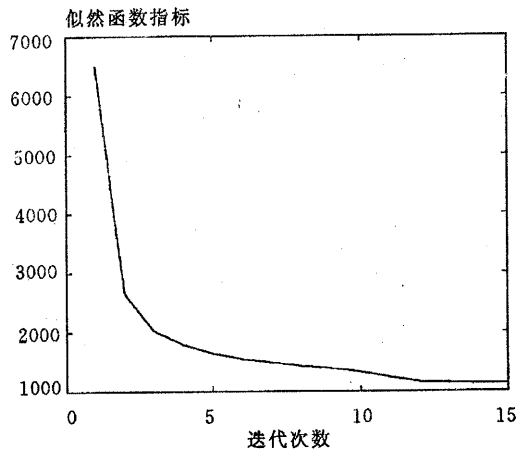
(a) 由 HOS 方法得到的初始模型的脉冲响应



(b) 第 11 次迭代后的优化结果

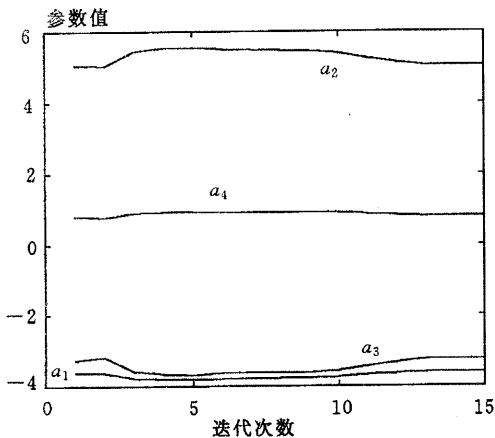


(c) 第 15 次迭代后的最终优化结果

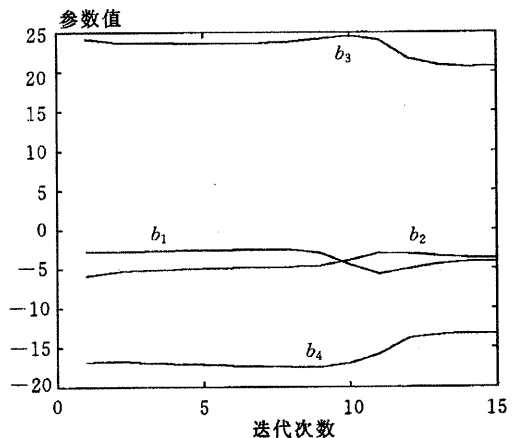


(d) 似然指标函数 $J(\theta|z)$ 的收敛过程

图 2 辨识算法的迭代收敛过程



(a) AR 参数的收敛过程



(b) MA 参数的收敛过程

图 3 ARMA 模型参数的收敛过程

得到更多有效而实用的方法.

本文的算法已应用于“具有任意相位的地震子波辨识”地震信号处理问题中,并取得了令人满意的效果.

参 考 文 献

- 1 Mendel, J. M.. Tutorial on higher-order statistics (spectra) in signal processing and system theory: theoretical results and some applications. Proc. IEEE, 1991,79(3):278-305
- 2 高江,戴冠中. 高阶矩法辨识非最小相位系统. 控制理论与应用, 1994,11(1):1-7
- 3 Friedlander, B. and Porat, B.. Asymptotically optimal estimation of MA and ARMA parameters of non-Gaussian processes from high-order moments. IEEE Trans. Automat. Contr., 1990, AC-35(1):27-35
- 4 Giannakis, G. B.. On the identifiability of non-Gaussian ARMA models using cumulants. IEEE Trans. Automat. Contr., 1990, AC-35(1):18-26
- 5 Edwards, A. W. F.. Likelihood. London: Cambridge University Press, 1972
- 6 Mehra, R. K.. Identification of stochastic linear dynamic systems using Kalman filtering representation. AIAA Journal, 1971,9(1):28-31
- 7 Mendel, J. M.. Maximum-Likelihood Deconvolution: A Journey into Model-based Signal Processing. New York: Springer Verlag, 1990
- 8 Himmelblau, D. M.. Applied Nonlinear Programming. New York: McGraw-Hill, 1972
- 9 Giannakis, G. B. and Mendel, J. M.. Cumulant-based order determination of non-Gaussian ARMA models. IEEE Trans. Acoust., Speech, Signal Processing, 1989, ASSP-37(8):1411-1423
- 10 Cadzow, J. A.. Spectral estimation: an overdetermined rational model equation approach. Proc. IEEE, 1982,75:907-938

On the Modeling of Non-Gaussian Systems

GAO Jiang

(Institute of Systems Engineering, Xi'an Jiaotong University, Xi'an, 710049, PRC)

DAI Guanzhong

(Department of Automatic Control, Northwestern Polytechnical University • Xi'an, 710072, PRC)

Abstract: Higher-order statistics methods are widely used in the modelling of non-Gaussian systems in recent years. Due to the specific information structure in the HOS methods, the statistical distribution information of the specific problems in applications can't be fully used. This problem has limited the application of HOS theory. To overcome this difficulty, a new method based on HOS and the maximum-likelihood index of the interested system is proposed in this paper. This provides a method to utilize the specific information in applications. Satisfactory results are obtained in simulations and applications.

Key words: system identification; higher-order statistics; likelihood function; non-Gaussian system

本文作者简介

高江 1969年生, 1990年毕业于西安交通大学信息与控制工程系, 获学士学位, 1996年于西北工业大学自动控制系获博士学位, 现在西安交通大学做博士后研究工作. 研究方向为系统辨识与最优估计, 数字信号处理及其并行化, 计算机视觉, 工业生产过程的调度、监控与仿真.

戴冠中 1937年生, 西北工业大学教授、校长, 自动控制理论及应用学科博士生导师. 主要的研究方向为大系统的控制与估计理论, 智能控制理论及应用, 面向控制的并行处理方法与并行计算机以及控制理论在石油勘探信号处理中的应用等.