文章编号: 1002-0411(2001)03-271-05

非线性系统的参数空间分割辨识方法

李 涌 韩崇昭
(西安交通大学电信学院综合所 710049)

摘 要: 本文提出了一种新的非线性系统 Volterra 级数模型辨识方法, 为非线性系统辨识中的"维数灾难"问题提供了一种满意的解决. 算法中参数空间分割和模型辨识同时完成, 降维依据采用输出拟合结果的均方误差, 最终得到输出拟合均方误差意义上的准最优解. 本算法也可以作为非线性系统模型的结构辨识算法, 并可以直接推广应用于其它很大一类非线性系统模型. 仿真试验结果表明, 算法计算量小, 精度高, 并具有较好的稳定性, 可以应用于在线实时辨识.

关键词: 非线性系统辨识 维数灾难 参数空间分割 结构辨识
中图分类号: TP271 文献标识码: B

NONLINEAR SYSTEM IDENTIFICATION WITH PARAMETER-SPACE SEGMENTING

LI Yong HAN Chong-zhao
(School of Electronic and Information Engineering, Xi'an Jiaotong University, Xi'an 710049)

Abstract: A new algorithm of nonlinear system identification for Volterra series model is presented in this paper. The algorithm provides a satisfactory solution to the dimension disaster of nonlinear system identification. Parameter-space segmenting and identification is completed synchronously in the algorithm. The target of P-Space segment is MSE of output approach and a near-optimal solution is achieved finally. The algorithm can be used as model structure identification and popularized to many other kinds of nonlinear system model. It is indicated by the result of simulation that the new method is of very limited calculation time, high accuracy and robustness. Therefore it can be used to identify the real plant on-line.

Keywords: nonlinear system identification, the dimension disaster, parameter-space segmenting, model structure identification

1 引言(Introduction)

Volterra 级数理论在非线性系统分析中占有十分重要的地位. 这一方面是因为 Volterra 级数理论有着十分坚实的数学理论基础, 另一方面是因为在大多数情况下, 一个一般的低阶非线性系统可以用截断 Volterra 级数描述, 并能取得满意的精度[1]. 基于 Volterra 级数的非线性系统辨识通过对系统输入输出数据的分析, 得到系统的物理意义明确的高阶脉冲响应函数, 或广义频域响应函数 GFRF(General Frequency Response Functions). 这一方法在科学和工程上成功地解决了非线性系统的一系列问题, 逐渐成为非线性研究的热点之一.

非线性系统 Volterra 级数辨识存在一个固有的问题: 维数灾难, 即 Volterra 级数的维数与系统模型阶次和记忆长度呈指数关系增长. 关于 Volterra 级数的非线性系统辨识, 虽然目前已存在一些较为成熟的算法[2,3], 但由于此普遍存在的问题, 一般的 Volterra 级数模型辨识算法不但计算量非常大, 而且需要的非线性系统的输入输出数据量也很大. 若采用如: 高斯信号[4]、多音正弦合成信号[5]等特性好的特殊输入信号作为非线性系统辨识的激励信号, 可以有效降低辨识算法的计算量和所需数据量. 但在工程应用场合, 对激励信号的过多限制却必定使算法的应用范围受到相应的约束. 一种经验化的两步式非线性系统模型预降维方法被用来解决维数灾难问题[6,7], 方法中利用一种经验性的指标对参数的重要性进行评价, 算法并不是任何辨识误差意义上的寻优[6].

本文提出一种新的非线性系统 Volterra 模型

收稿日期: 2000-07-14

辨识方法. 在这一递推算法中, 模型降维和模型辨识同时完成, 降维依据是输出拟合均方误差, 最终得到输出拟合均方误差意义上的准最优解. 对算法经过详细的理论推导后, 给出了仿真实验研究.

2 Volterra级数模型的维数灾难(The dimension disaster of volterra deries model)

考虑单输入单输出有限阶非线性离散系统的Volterra级数模型[8]

$$y(k) = \sum_{n=1}^{N}\sum_{m_1=0}^{M-1}\cdots\sum_{m_n=0}^{M-1}h_n(m_1, m_2, \cdots, m_n)\prod_{i=1}^{n}u(k-m_i) + e(k) \tag{1}$$

N、M、$e(k)$分别表示非线性系统的最高阶次、脉冲响应函数的长度(系统记忆长度)和模型的截尾误差.

式(1)中非线性系统的输入输出测量序列$\{u(k)\}\{y(k)\}k=1, 2, \ldots$在最小二乘意义下有以下等式成立

$$Y = X\hat{\theta} \tag{2}$$

其中, X为输入矩阵, Y为输出向量, $\hat{\theta}$为非线性系统的截断Volterra核向量的估计值

$$\left|\begin{array}{l} X = [U(k), U(k+1), \cdots, U(k+L-1)]^{\mathrm{T}} \\ U(k) = [u(k), \cdots, u(k-M+1), u^2(k), \\ \quad u(k), u(k-1), \cdots, u^N(k-M-1)]^{\mathrm{T}} \end{array}\right. \tag{3}$$

$$Y = [y(k), y(k+1), \cdots, y(k+L-1)]^{\mathrm{T}} \tag{4}$$

$$\hat{\theta} = [\hat{h}_1(0), \cdots, \hat{h}_1(M-1), \hat{h}_2(0,0), \hat{h}_2(0,1), \cdots, \hat{h}_{N-1}(M-1, \cdots, M-1)]^{\mathrm{T}} \tag{5}$$

L是用于辨识的输出测量序列的数据长度, 即式(2)中方程的个数.

基于截断Volterra级数的非线性系统辨识, 就是在已知非线性系统的输入序列$u(k)$和输出序列$y(k)$的情况下, 利用式(2)求解Volterra核向量θ.

设使辨识结果满足预定精度要求的θ的长度为S(S由系统模型的阶数和滞后两个参数有决定). 并为叙述方便, 另记向量和矩阵X为:

$$\theta = [\theta_1, \theta_2, \cdots, \theta_s]^{\mathrm{T}}$$

$$X = \begin{vmatrix} x_1(k), & \cdots & x_s(k) \\ \vdots & \vdots & \vdots \\ x_1(k+L-1), & \cdots & x_s(k+L-1) \end{vmatrix} \tag{7}$$

其中, P和θ中的元素分别与(3), (5)式中的元素对应.

可以看出, S的大小随模型阶数和滞后成指数增长. 而由于要保证方程组的超定, 方程的个数也会按比例增加. 如, 对于一个3阶的Volterra级数模型, 当记忆长度为20时, 参数的个数达1770个, 矩阵X的规模将达到2000× 1770以上, 问题运算量十分巨大. 所以, 参数的数目应尽量在满足辨识精度的情况下尽量减少.

在系统分析中, 事先并不知道用来拟合数据所必需的模型参数的精确数目, 参数的数目在开始一般也不可能取很大, 所以不得不用不同的参数数目来试凑, 从而导致重解复杂的矩阵方程. 这无疑是一种很浪费的方法. 较理想的方法是利用已经完成的较少参数数目的运算结果完成进一步的逼近.

另外更重要的问题是, 如果当前的模型不能满足既定的精度要求, 一般是逐渐增加模型的阶数和滞后. 每当模型的阶数或滞后加1, 模型中将增加一组而不是一个参数, 这样直到达到精度. 一般的算法无法对这一组参数中各个参数的必要性分别讨论. 分析最后的模型会发现, 其实有大量的参数的值接近0. 也就是说, 为了达到要求的精度所需的参数远远小于沿用以上方法得到的模型所包含的参数数目. 在某种误差意义上, 哪些参数对估计重要, 哪些参数对估计不太重要, 这是解决维数灾难问题, 提高算法效率的关键所在.

3 参数空间分割及Volterra级数辨识(Volterra series identification with PSS)

根据(2)式构造出的Volterra级数模型中含有大量对系统贡献很小的元素, 为了简化计算, 可以不考虑它们对系统的影响, 认为它们的值为0. 如果在辨识前通过预处理, 找出对系统贡献大的元素, 对式(2)进行重构, 重构后的θ的维数必然下降, 式(2)中方程组的规模也将大大减小.

首次分析目标系统的时候, 需要完成模型降维和系统辨识两项任务. 本文采用参数空间分割(PSS Parameter-Space Segmenting)完成模型降维. 整个算法是一个逐步求精的过程, 在不断选取参数的同时进行辨识, 并利用当前的辨识结果进一步完成参数空间分割, 算法两部分尽量相互利用计算结果, 减少整个算法的运算量.

在不断选取参数的同时进行辨识, 是一种参数数目不断增加的系统辨识, 使用已经完成了的p个参数问题的计算来进行$p+1$个参数问题的估计, 不但将大大改善算法的效率, 而且可以对所增加的

参数的重要性进行评估, 进而优先选择最重要的参数, 最终以最少的参数达到满意的精度.

设模型(2)中当前已经有 p 个参数时, $\{\theta_{i_2}, \theta_{i_2}, \cdots, \theta_{i_p}\}$, $1 \leqslant i_1, i_2, \cdots, i_p \leqslant S$ 时, 辨识结果没有达到要求的精度. 现在考察在模型中加入第 $p+1$ 个参数, 即:

$$\theta = [\theta_{i_1}, \theta_{i_2}, \cdots, \theta_{i_p} \vdots \theta_{i_{p+1}}]^{\mathrm{T}} = [\Theta_1 \vdots \Theta_2]^{\mathrm{T}} \quad (8)$$

时问题解答的改进情况. 此时, 相应的输入矩阵 X 为:

$$X = \begin{vmatrix} x_{i_1}(k), \cdots, x_{i_p}(k) & x_{i_{p+1}}(k) \\ \vdots & \vdots \\ x_{i_1}(k+L-1), \cdots, x_{i_p}(k+L-1) & x_{i_{p+1}}(k+L-1) \end{vmatrix} = [X_1 \vdots X_2] \quad (9)$$

根据(2)式, 我们有:

$$\begin{vmatrix} X_1^{\mathrm{T}}X_1 & X_1^{\mathrm{T}}X_2 \\ X_2^{\mathrm{T}}X_1 & X_2^{\mathrm{T}}X_2 \end{vmatrix} \begin{vmatrix} \Theta_1 \\ \Theta_2 \end{vmatrix} = \begin{vmatrix} X_1^{\mathrm{T}}Y \\ X_2^{\mathrm{T}}Y \end{vmatrix} \quad (10)$$

于是可得:

$$\left| \begin{array}{l} X_1^{\mathrm{T}}X_1\hat{\Theta}_1 + X_1^{\mathrm{T}}X_2\hat{\Theta}_2 = X_1^{\mathrm{T}}Y \\ X_2^{\mathrm{T}}X_1\hat{\Theta}_1 + X_2^{\mathrm{T}}X_2\hat{\Theta}_2 = X_2^{\mathrm{T}}Y \end{array} \right. \quad (11)$$

解这个方程组就可以得到 $\hat{Q}_1, \hat{Q}_2$:

$$\left| \begin{array}{l} \hat{\Theta}_1 = \hat{\hat{Q}}_1 - AX_2^{\mathrm{T}}(Y - X_1\hat{\hat{Q}}_1) \\ \hat{Q}_2 = BX_2^{\mathrm{T}}(Y - X^1\hat{\hat{Q}}_1) \end{array} \right. \quad (12)$$

其中

$$\left| \begin{array}{l} A = (X_1^{\mathrm{T}}X_1)^{-1}X_1^{\mathrm{T}}X_2B \\ B = [X_2^{\mathrm{T}}X_2 - X_2^{\mathrm{T}}X_1(X_1^{\mathrm{T}}X_1)^{-1}X_1^{\mathrm{T}}X_2]^{-1} \end{array} \right. \quad (13)$$

$\hat{\hat{Q}}_1 = (X_1^{\mathrm{T}}X_1)^{-1}X_1^{\mathrm{T}}Y$ 是核向量取 p 个参数时的最小二乘估计.

应该特别注意的是, B 是标量, 式(13)中的矩阵求逆, 实际上是简单的求倒.

为了使算法能计算下一个更高阶的模型的参数, 必须计算 A B 中都用到的逆矩阵 $(X^{\mathrm{T}}X)^{-1}$. 利用分块矩阵求逆特性可得:

$$(X^{\mathrm{T}}X)^{-1} = \begin{vmatrix} X_1^{\mathrm{T}}X_1 & X_1^{\mathrm{T}}X_2 \\ X_2^{\mathrm{T}}X_1 & X_2^{\mathrm{T}}X_2 \end{vmatrix}^{-1}$$

$$= \begin{vmatrix} C + AX_2^{\mathrm{T}}X_1C & -A \\ -A^{\mathrm{T}} & B \end{vmatrix} \quad (14)$$

其中 $C = (X_1^{\mathrm{T}}X_1)^{-1}$

记第 p 次迭代和 $p+1$ 次迭代的输出拟合误差向量分别为:

$$E^{(p)} = Y - X_1\hat{\hat{Q}}_1 \quad (15\mathrm{a})$$

$$E^{(p+1)} = Y - X\theta \quad (15\mathrm{b})$$

在很多应用场合, 系统辨识的目的都是为了实现对系统行为的某种预测或逼近. 以辨识模型和实际系统的在同一激励下输出的均方误差作为衡量辨识效果的标准有普遍的意义, 使用这一误差作为评价参数重要性的指标.

第 p 次迭代和 $p+1$ 次迭代的输出拟合均方误差分别为

$$e^{(p)} = \frac{1}{p}(E^{(p)})^{\mathrm{T}}E^{(p)} \quad (16\mathrm{a})$$

$$e^{(p+1)} = \frac{1}{p+1}(E^{(p+1)})^{\mathrm{T}}E^{(p+1)} \quad (16\mathrm{b})$$

引进第 $p+1$ 个参数后辨识模型的输出拟合均方误差的改进为

$$\Delta e^{(p+1)} = e^{(p)} - e^{(p+1)} \quad (17)$$

$\Delta e^{(p)}$, $p = 1, 2, 3, \ldots$ 反应了最后引进的这个参数对辨识效果的贡献, 它将作为评价模型待定参数重要性的评价准则.

在辨识之前首先要根据经验选定一个最大系统阶数和记忆长度, 由此确定问题的整个参数空间, 以后的参数评价和空间分割都基于这一空间. 设当前模型中已经包含 p 个参数, 令 $\Delta e_j^{(p+1)}$ ($j \in [1, S]$, $j \notin \{i_1, i_2, , \cdots, i_p\}$) 是引进核向量的第 j 个元素作为模型的第 $p+1$ 个参数后, 辨识模型的输出拟合均方误差的改善, 记 $k = \arg\min_j\{\Delta e_j^{(p+1)}\}$, 则选择核向量的第 k 个元素作为模型的第 $p+1$ 个参数. 当 $p = 1$ 时, $\Delta e_j^{(1)}$ 是模型中只含有核向量的第 j 个元素时, 含一个参数的模型对应的均方误差. 记 $l = \arg\min_j\{\Delta e_j^{(p)}\}$, 则选择核向量的第 l 个元素为模型的第一个参数.

4　仿真研究(Simulation)

考虑如下三阶非线性模型:

$$\begin{aligned} y(n) = &-0.64u(n) + 0.89u(n-1) \\ &-0.95u(n-2) - 10u^2(n-1) \\ &+10.6u(n-1)u(n-2) - 0.5u^3(n) \\ &+u^3(n-1) - 1.4u^3(n-2) \end{aligned}$$

仿真时, 非线性系统的激励信号为限带白噪声, 在非线性系统的输出端加均匀分布的白噪声, 信噪比(SNR)为 30db. 共进行 3 次仿真试验, 每次采样 500 对输入输出数据.

参考文献[7]中使用同样的对象, 其中取记忆长度为 17, 参考其取值, 这里模型的记忆长度取为 12.

这样参数空间分割前θ共有 454 个元素, 取允许的均方误差为$(a\times 0.01)^2$, 其中 a 为该次辨识所用输出采样的最大幅值. 表 1 给出了试验的有关数据, 表 2 给出了最终模型中包含的参数.

表 1 模型辨识结果

Table. 1 Result of identification

序号	输出最大幅值	允许均方误差	模型参数个数	输出拟合最大误差	输出拟合均方误差
1	3.7689	0.001420	12	0.1715	0.001192
2	6.7039	0.004494	12	0.2252	0.003954
3	5.4323	0.002833	10	0.2485	0.002203

表 2 辨识模型中包含的参数

Table. 2 Parameter in each model

序号	参数个数	模型包含的参数在θ中的编号(依选择顺序排列)											
1	12	24	13	3	21	108	30	14	16	1	2	22	107
2	12	113	24	13	2	21	36	110	14	20	15	123	17
3	10	106	29	13	18	2	15	17	5	1	3		

可以看到, 在进行参数空间分割后模型仅包含 10~12 个参数, 核向量的维数和算法的复杂程度明显降低, 并远少于文献 7 的最终辨识模型包含的 131 个元素之数, 参数选择的效率和合理性明显提高.

通过将辨识出的模型结果应用于不同的采样数据, 比较实际输出和辨识模型在同一激励下的预测输出, 可检验辨识模型的有效性和普适性. 表 3 是将辨识出的 3 个模型交叉应用于 3 次采样数据的输出拟合均方误差.

纵向比较表 3 中的各列数据会发现, 对采样数据应用由其它数据得到的辨识模型并没有造成输出拟合误差的增大; 横向比较表 3 中的各行数据会发现, 同一个辨识模型对不同的输出输入采样数据组的拟合情况大致相同.

表 3 模型普适性检验

Tabel. 3 Test for universality of model

模型序号 \ 数据序号	1	2	3
1		0.001641	0.003449
2	0.004330		0.007506
3	0.000992	0.001261	

应用本算法得到对象的模型之后, 也就得到了非线性系统模型的结构(参数空间分割结果). 以后的辨识可以直接利用这一结果, 通过各种简单的一次性辨识方法方便地完成, 这对于工业现场对象的长期监视十分有意义. 这种情况下实际是将本文的辨识方法作为模型结构辨识, 经过本算法, 随后的辨识面对的是一个规模很小的拟合问题, 运算量很小. 辨识的效果检验了参数空间分割的普适性(表 4).

表 4 参数空间分割普适性检验

Tabel. 4 Test for universality of PSS

模型序号 \ 数据序号	1	2	3
1		0.001261	0.002797
2	0.003332		0.005833
3	0.000897	0.001087	

可以看出, 表 4 中的各数据均比表 3 中的相应误差小 9.6%~23% 左右.

仿真试验结果表明, 最终模型只使用了原参数空间的 2.20%~2.64% 就达到了预期的精度, 算法的效率令人满意. 另外可以看出, 并不是参数越多的模型拟合精度越好, 精度与参数的选择有很大的关系, 这也是本算法的基础和意义所在. 而且, 从不同输入输出采样数据得到的模型中包含的参数并不相同, 也就是说, 可以使用不同的 Volterra 模型逼近同一个非线性系统的输入输出行为, 并均达到相当的精度.

最终模型中使用的参数的序号最大只有 123, 这就意味着开始的记忆长度取 8(核长度 164) 就已经足够. 当然这是无法事先准确预知的, 只能通过积累每次应用时的经验逐步提高选取初值的准确性. 本算法即使在记忆长度取 8 得情况下, 使用的参数也只占参数空间的 60.97%~73.17%, 效果还是十分明显的, 而且这一效果在对记忆长度大的系统进行辨识时更加明显.

模型普适性检验结果表明, 决定输出拟合效果的主要是使用的辨识模型, 而不是输入输出采样数据组. 这说明, 算法得到的模型确实反应的是系统本身的本质特性. 对采样数据的不敏感表明算法的稳定性良好.

一次结构辨识多次参数辨识应用结果(表 4)的精度明显高于模型普适性检验得到的结果, 原因是表 4 采用的模型本身就是利用新的采样值计算的, 包含了更多的当前信息. 而精度提高的代价是额外的小规模最小二乘问题直接求解的运算量开销.

5　结论(Conclusion)

非线性系统辨识的维数灾难问题是一个普遍存在的难题, 本文利用参数空间分割的方法, 形成了一种具有良好降维能力的非线性系统辨识算法. 本文重点讨论的是非线性系统 Volterra 级数模型的辨识, 但该方法可以很方便地推广到相当多种类的非线性系统模型. 作为一种结构辨识方法, 本算法的结果有良好的稳定性, 降维效果十分显著.

参　考　文　献(References)

1　J F Barret. The Use of Functions in The Analysis of Nonlinear System, Journal of Electronics Controls, 1963, **15**(6): 567~ 615

2　Kim K I, Powers E j. A Digital Method of Modeling Quadratically Nonlinear System with a General Random Input, IEEE Trans on ASSP, 1988, **36**(2): 1758~ 1769

3　Tseng Ching-Hsiang. A Mixed Domain Method for Identification of Quadratically Nonlinear System. IEEE Trans on SP, 1997, **45**(4): 1013~ 1024

4　Tick L J. The Estimation of Transfer Function of Quadratic Systems, Technometrics, 1961, **3**(11): 563~ 567

5　韩崇昭, 王立琦, 马　川. 非线性动态系统的非参数辨识算法. 第二界全球华人智能控制与智能自动化大会论文集, 西安, 1997: 1765~ 1769

6　Billings S A, Tsang K M. Spectral Analysis for Non-linear Systems. Part I: Parametric Non-linear Spectral Analysis

7　唐晓泉, 王文正, 方洋旺, 李　涌, 韩崇昭等, 子集优化在非线性系统辨识中的应用. 西安交通大学学报, 1999, **33**(3): 19~ 22

8　焦李成. 非线性传递函数理论与应用. 西安: 西安电子科技大学出版社, 1992, 77~ 128

作者简介

李　湧(1972-), 西安交通大学电子与信息工程学院综合自动化所博士研究生. 研究领域为非线性系统的频域分析、非线性系统辨识、非线性系统故障诊断等方面的的研究工作, 同时进行汽车悬挂部件的在线检测和故障诊断方面的应用研究.

韩崇昭(1943-), 1968 年毕业于西安交通大学电机工程系, 1981 年于中科院研究生院获硕士学位, 现为西安交大电子与信息工程学院副院长、教授、博士生导师. 研究领域为随机与自适应控制、工业过程控制与稳态优化、非线性系统的频域分析和多目标信息融合等.

(上接第 270 页)

参　考　文　献(References)

1　宋天虎. 积极发展适合我国国情的虚拟制造技术. 中国机械工程, 1998, **9**(11)

2　牛军钰, 赵大哲, 赵　宏. 一个开放集成的工作流管理系统. 第十届中国计算机网络与数据通信学术会议, 南京, 1998, 287~ 290

3　B R Gaines, K H Norrie, A Z Lapsley. An Intelligent Information System Supporting the Virtual Manufacturing Enterprise, http://ksi.cpsc.ucalgary.ca80/articles

4　S Joosten. Workflow Management Research Area Overwiew, Proceedings of the 2nd Americas Conference on Information Systems, Phenix, Arizona, August 16- 18, 1996, 914~ 916

5　张维明, 邓苏等. 信息系统建模技术与应用. 电子工业出版社, 1997

6　J Billington, A Tokmakoff. Petri Nets and Traders: Enabling Technologies for Virtual Enterprises. In IEEE Sixth Workshop on Enabling Technologies, Infrastructure for Collaborative Enterprises (WETICE'97), Cambridge, MA, 18- 20 June 1997. IEEE Press

作者简介

李业丽(19-), 博士生. 研究领域为系统工程、数据挖掘、多媒体信息处理.

常桂然(19-), 教授. 研究领域为计算机网络、多媒体技术、数据挖掘等.

潘德惠(19-), 教授. 研究领域为分布参数控制、管理科学、控制理论.