第14卷 第5期 自动化学报 Vol. 14, No. 5
1988年9月 ACTA AUTOMATICA SINICA Sept., 1988

多层递阶辨识方法

韩志刚
(黑龙江大学)

摘要

本文提出了作为多层递阶预报方法的基础的多层建模的方法,称为多层递阶辨识方法,并指出这种方法可能是解决非线性系统辨识问题的一种有效的途径.

一、引言

笔者在文献[1]中提出了多层递阶预报方法,在文献[2]中讨论了预报模型的时变参数的辨识问题.事实上在这两篇文章中,已经提出了多层递阶辨识方法的基本思想,但还不够明确.目前多层递阶预报方法已逐渐的被人们所接受[3,4,5].它的理论基础是多层建模,所以对多层辨识方法的理论分析已属十分必要了.文献[6]也说明了,通过模型多层化的途径,可能会使某些问题的处理得到简化.

本文的目的是企图对多层辨识方法作一个简明的介绍,并说明,这种用多层线性模型来描述动态系统的方法,可能对某些非线性系统的建模带来方便.

二、确定模型结构的基本定理

模型结构的确定,特别是非线性的情形,是系统辨识中的一个困难问题.然而,多层递阶辨识方法可以借助于层数的增加,用多层的线性模型来描述所考虑的系统,从而减轻或避免了上述的困难.

这里将说明,下述形式的模型

$$y(k)=\boldsymbol{\varphi}(k)^{\tau}\boldsymbol{\theta}(k) \tag{1}$$

在多层递阶辨识方法中是有基本意义的.

定义:设 S_1 和 S_2 是两个动态系统,它们具有相同的输入集和输出集.如果在过去的输入和输出皆相同的条件下,现在的同一个输入,两个系统皆对应着同样的输出,则说这两个系统是输入-输出等价的,或简称等价的.

由于有输出可分离性定理[2],所以以下仅考虑(广义)单输出模型.

简单结论 1:设系统 S_1 具有形式

$y(k)=\boldsymbol{\varphi}(k)^{\tau}\boldsymbol{\eta}(k)+v(k)$,

本文于1986年2月28日收到.

其中 $\boldsymbol{\eta}(k)$ 是时变参数向量，$v(k)$ 是模型的未知随机部分，则必存在一组时变随机参数 $\boldsymbol{\theta}(k)$，使得系统 S_2

$$y(k)=\boldsymbol{\varphi}(k)^{\tau}\boldsymbol{\theta}(k)$$

与系统 S_1 等价.

证明：只须证明这样的随机时变参数 $\boldsymbol{\theta}(k)$ 的存在性. 事实上，只要取

$$\boldsymbol{\theta}(k)=\boldsymbol{\eta}(k)+\frac{1}{\|\boldsymbol{\varphi}(k)\|^2}\boldsymbol{\varphi}(k)v(k)$$

就够了. 因为此时有：

$$\boldsymbol{\varphi}(k)^{\tau}\boldsymbol{\theta}(k)=\boldsymbol{\varphi}(k)^{\tau}\boldsymbol{\eta}(k)+\boldsymbol{\varphi}(k)^{\tau}\frac{\boldsymbol{\varphi}(k)}{\|\boldsymbol{\varphi}(k)\|^2}v(k)$$
$$=\boldsymbol{\varphi}(k)^{\tau}\boldsymbol{\eta}(k)+v(k)=y(k).$$

简单结论 2：如果非线性系统的模型具有形式

$$y(k)=f[\boldsymbol{Y}_{k-1},\ \boldsymbol{U}(k),\ \boldsymbol{\xi}(k),\ k]+e(k),$$

其中 $y(k)$ 是一维的输出，$\boldsymbol{U}(k)$ 是输入，$\boldsymbol{\xi}(k)$ 是时变参数向量，$e(k)$ 是随机噪声，而

$$\boldsymbol{Y}_{k-1}=\{y(0),\ \cdots,\ y(k-1)\},$$
$$\boldsymbol{U}(k)=\{u(0),\ \cdots,\ u(k)\},$$

则在输入-输出等价的意义下，可以把上述系统的模型写成下述的形式：

$$y(k)=\boldsymbol{\varphi}(k)^{\tau}\boldsymbol{\theta}(k),$$

其中 $\boldsymbol{\varphi}(k)$ 是由 $k-1$ 及其以前时刻的某些输出，和 k 及其以前时刻的某些输入组成的向量. $\boldsymbol{\theta}(k)$ 是随机时变参数.

证明：我们有

$$y(k)=f[\boldsymbol{Y}_{k-1},\boldsymbol{U}_k,\boldsymbol{\xi}(k),k]+e(k)$$
$$=\boldsymbol{\varphi}(k)^{\tau}\boldsymbol{\eta}(k)+f[\boldsymbol{Y}_{k-1},\boldsymbol{U}_k,\boldsymbol{\xi}(k),\ k]-\boldsymbol{\varphi}(k)^{\tau}\boldsymbol{\eta}(k)+e(k),$$

只要置：

$$v(k)=f[\boldsymbol{Y}_{k-1},\boldsymbol{U}_k,\boldsymbol{\xi}(k),k]-\boldsymbol{\varphi}(k)^{\tau}\boldsymbol{\eta}(k)+e(k),$$

就有：

$$y(k)=\boldsymbol{\varphi}(k)^{\tau}\boldsymbol{\eta}(k)+v(k).$$

然而 $v(k)$ 是随机部分，于是由结论 1 即得出所需的结论.

注：上述结论中涉及到的模型

$$y(k)=\boldsymbol{\varphi}(k)^{\tau}\boldsymbol{\theta}(k)$$

中的 $\boldsymbol{\varphi}(k)$ 的确定具有一定的任意性，在实际应用中将借助于定阶的方法来确定.

上述的两条结论说明了具有形式

$$y(k)=\boldsymbol{\varphi}(k)^{\tau}\boldsymbol{\theta}(k)$$

的模型是具有一定的普遍意义的. 此处 $\boldsymbol{\theta}(k)$ 是随机时变参数.

三、模型的多层结构与阶的确定

由于模型

$$y(k)=\boldsymbol{\varphi}(k)^{\tau}\boldsymbol{\theta}(k) \tag{1}$$

中存在有随机时变参数，所以用常规的方法来分析和处理它，是存在着困难的，然而它却适宜于模型的多层化，其步骤如下．

显然 $\{\boldsymbol{\theta}(k)\}$ 是一个多维的随机序列，自然可以考虑对这个序列的建模问题，不妨设 $\boldsymbol{\theta}(k)$ 的维数是 n_1．

如果存在常数 n_2 及非时变的参量矩阵 $A_1^{(1)}, A_2^{(1)}, \cdots, A_{n_2}^{(1)}$，使得有：

$$\boldsymbol{\theta}(k)=A_1^{(1)}\boldsymbol{\theta}(k-1)+\cdots+A_{n_2}^{(1)}\boldsymbol{\theta}(k-n_2)+e_2(k), \tag{2}$$

其中 $\{e_2(k)\}$ 为不相关零均值的噪声．我们就说原系统是两层，n_1、n_2 阶的，记作 Lev $(2, n_1, n_2)$．(1)和(2)是这个系统的两层模型，其参数用文献[2]中的方法确定．

如果对于任何的 n_2，模型

$$\boldsymbol{\theta}(k)=A_1^{(1)}(k)\boldsymbol{\theta}(k-1)+\cdots+A_{n_2}^{(1)}(k)\boldsymbol{\theta}(k-n_2)+e_2(k)$$

当 $\{e_2(k)\}$ 为零均值的不相关噪声时，参量矩阵 $A_1^{(1)}(k), \cdots, A_{n_2}^{(1)}(k)$ 皆为时变的，则取 $n_2=n_1$，建立模型

$$\boldsymbol{\theta}(k)=B_1^{(1)}(k)\boldsymbol{\theta}(k-1)+\cdots+B_{n_1}^{(1)}(k)\boldsymbol{\theta}(k-n_1), \tag{2°}$$

其中 $B_1^{(1)}(k), \cdots, B_{n_1}^{(1)}(k)$ 是随机时变参数矩阵．对 $B_1^{(1)}(k), \cdots, B_{n_1}^{(1)}(k)$ 的所有元素进行重新排列(次序无关重要)，组成向量 $\boldsymbol{\theta}_2(k)$，于是得到了一个新的随机序列 $\{\boldsymbol{\theta}_2(k)\}$．考虑这个序列的建模问题．

如果存在 n_3 及非时变参数矩阵 $A_1^{(2)}, \cdots, A_{n_3}^{(2)}$，使得：

$$\boldsymbol{\theta}_2(k)=A_1^{(2)}\boldsymbol{\theta}_2(k-1)+\cdots+A_{n_3}^{(2)}\boldsymbol{\theta}_2(k-n_3)+e_3(k), \tag{3}$$

其中 $\{e_3(k)\}$ 是零均值不相关噪声，我们就说原系统是三层，n_1、n_3 阶的，记作 Lev $(3, n_1, n_3)$，(1)、(2°)、(3) 就是这个系统的三层模型．参数估计算法同前．

如果这样的 n_3 不存在，则取 $n_3=n_1$，建立模型

$$\boldsymbol{\theta}_2(k)=B_1^{(2)}(k)\boldsymbol{\theta}_2(k-1)+\cdots+B_{n_1}^{(2)}(k)\boldsymbol{\theta}_2(k-n_1), \tag{3°}$$

其中 $B_1^{(2)}(k), \cdots, B_{n_1}^{(2)}(k)$ 是随机时变参数矩阵．对它们的所有元素进行重排，组成向量 $\boldsymbol{\theta}_3(k)$，于是又得到了一个新的随机序列 $\{\boldsymbol{\theta}_3(k)\}$，如此继续下去．

对于每个形如(1)的模型，可能有两种情形发生：

(1) 存在某个层数 I，正整数 n_2 和非时变参数矩阵 $A_1^{(I-1)}, \cdots, A_{n_I}^{(I-1)}$，使得

$$\theta_{I-1}(k)=A_1^{(I-1)}\boldsymbol{\theta}_{I-1}(k-1)+\cdots+A_{n_I}^{(I-1)}\boldsymbol{\theta}_{I-1}(k-n_I)+e_I(k), \tag{I}$$

其中 $\{e_I(k)\}$ 是零均值不相关噪声．此时可以说原系统是 I 层，n_1、n_I 阶的，记作 Lev (I, n_1, n_I)．模型组 (1), (2), ⋯ $(I-1)$, (I) 就是这个系统的 I 层模型．

(2) (1)中所述的情形不发生．于是对每个层数 I，皆有模型

$$\boldsymbol{\theta}_{I-1}(k)=B_1^{(I-1)}(k)\boldsymbol{\theta}_{I-1}(k-1)+\cdots+B_{n_I}^{(I-1)}(k)\boldsymbol{\theta}_{I-1}(k-n_I), \tag{I°}$$

我们就说原系统是无限层的．此时可以考虑选取适当的层数，对原系统进行近似描述．

从上述讨论中可以看出，模型的阶数 n_1 和 n_I 的确定，是多层模型建模的重要环节．二层以上的模型虽然都是多维的，但由于有输出可分离性定理[2]，所以只考虑一维的情形就够了．

以模型

$$y(k)=\boldsymbol{\varphi}(k)^{\tau}\boldsymbol{\theta}(k)$$

为例. 如果 $\boldsymbol{\varphi}(k)$ 的维数是 n, 显然,当观测数据 $(\boldsymbol{\varphi}(k), y(k))$ 给定的条件下,准测函数

$$J=\sum_{k=1}^{n} w(k)(y(k)-\boldsymbol{\varphi}(k)^{\tau}\hat{\boldsymbol{\theta}}(k-1))^{2}$$

必然依赖于 n 和 $\hat{\boldsymbol{\theta}}(0)$, 即 $J=J(n, \hat{\boldsymbol{\theta}}(0))$ 置

$$J^{*}(n)=\min_{\hat{\boldsymbol{\theta}}(0)} J(n, \hat{\boldsymbol{\theta}}(0)).$$

于是选取 n_1, 使得下式满足就行了:

$$J^{*}(n_1)=\min_{n} J^{*}(n).$$

对于经分离后的高层广义单输出模型,关于它的阶的确定,当然可以应用上述的方法.

四、应 用

多层递阶辨识方法实质上已有了一些较好的应用, 在预报方面的应用可参看文献[6]、[7]等等,在自适应控制方面的应用可参看文献[8].

参 考 文 献

[1] 韩志刚,动态系统预报的一种新方法,自动化学报, Vol. **9**, No. 3, 1983, 161—168.
[2] 韩志刚,动态系统时变参数的辨识,自动化学报, Vol. **10**, No. 4, 1984, 330—337.
[3] 蔡季冰,蔡治平. 自适应预报方法在铁路客流量预测上的应用,信息与控制, Vol. **14**, No. 1, 1985, 33—35.
[4] 罗新红,刘樵良,多层递阶预报方法的改进及其应用,信息与控制, Vol. **15**, No. 1, 1986, 1—5.
[5] 张恩恕等,西太平洋夏季副高特征量多层递阶长期预报模型,高原气象, Vol. **3**, No. 2, 1984, 51—56.
[6] Han, Z. G., Tang, B. Y., Multistratum Recursive Linearization Prediction Method and Its Application in the oil Field, IFAC The Conference on Digital Computer Application to Process Control, Vienne, September, 1985.
[7] 韩志刚,多层递阶预报方法及其应用,自然杂志, Vol. **8**, No. 1, 1985, 39—42.
[8] 韩志刚等,参数预报自适应控制方法及其在农业中的应用,控制理论与应用, Vol. **2**, No. 4, 1986, 44—52.

MULTI-LEVEL RECURSIVE IDENTIFICATION METHOD

Han Zhigang

(Institute of Applied Mathematics, Heilongjiang University)

Abstract

In this paper, the multi-level modelling method called the multi-level recursive identification method is introduced. It is indicated that this method may be a kind of effective approach for solving non-linear system identification problems.