第 9 卷第 1 期
1992 年 2 月

控制理论与应用
CONTROL THEORY AND APPLICATIONS

Vol. 9, No. 1
Feb., 1992

模型结构未知的多模型系统辨识的一种简化途径

韩志刚　王德进

(黑龙江大学应用数学研究所·哈尔滨,150080)

摘要：本文考虑多模型系统,包括模型结构时变系统的辨识问题.这里不假定模型结构已知,给出了多模型系统模型族的统一描述方法,包括模型结构与参数的辨识方法.并对所提出的方法进行了理论分析和计算机仿真.

关键词：多模型;时变参数;替代模型;节省能量准则;阶的判定

1　问题的提出

我们在文献[1]中讨论了一类结构时变系统的辨识问题.基本假设是系统的模型随着时间区段的不同而有所变化.但模型的结构与数目都必须假定是已知的.文献[2]在上述基本假设下,把文献[1]所考虑的辨识方法推广到结构随机变化的动态系统上去.

事实上,结构(随机)时变系统是一种多模型系统,系统模型结构的辨识问题是系统与模型的匹配问题.所以我们可以把文献[1]、[2]所考虑的问题统一为多模型系统的辨识问题.

多模型系统是有明显实际背景的.例如,商品的市场需求量预测系统,多级火箭系统等,都可以用多模型系统来描述.

但在实践中,一般情况下,解决系统辨识问题所能依赖的主要信息往往来自输入、输出观测数据.通常从实际背景,可以初步判断该系统是否为多模型系统.除此之外,关于系统的其它信息就很难获得了,在这种意义下,为了应用的一般性,我们应该把模型结构已知这一假定取消,即考虑模型结构未知的多模型系统的辨识问题.本文将以文献[3]所提出的多层递阶辨识方法为基础,给出解决上述问题的一种简化途径.

2　系统模型族的简化

文献[3]指出,对任何如下形式的模型:

$$Y_k = f[Y_{k-1}^{k-n}, U_{k-1}^{k-m}, \eta(k), k] + v_k. \tag{1}$$

其中 Y_k 是一维输出,U_k 是输入,而

$$Y_{k-1}^{k-n} = \{Y_{k-1}, Y_{k-2}, \cdots, Y_{k-n}\},$$
$$U_{k-1}^{k-m} = \{U_{k-1}, U_{k-2}, \cdots, U_{k-m}\}.$$

$\eta(k)$是时变参数,v_k 是模型的未知随机部分,n、m 是正整数,总可以找到一个如下形式的模型:

$$Y_k = \psi^\tau(k)\theta(k), \tag{2}$$

使得(1)与(2)是输入-输出等价的.此处

$$\psi(k) = (Y_{k-1}, Y_{k-2}, \cdots, Y_{k-n'}, U_{k-1}, U_{k-2}, \cdots, U_{k-m'})^\tau.$$

本文于1990年6月12日收到.1991年9月30日收到修改稿.

n'、m'是适当的正整数，一般有 $n'=n, m'=m$. 而 $\theta(k)$是一个随机时变参数向量.

在以下讨论中，不失一般性，假定 $m=n$. 并称 n 为模型(2)的阶数.

用 μ 表示多模型系统的模型集，即

$$\mu = \{M_i, \quad i = 1,2,\cdots,M\}.$$

M 是模型集 μ 中模型的个数. 假定模型 M_i 具有形式：

$$M_i: \quad Y_k = f_i[Y_{k-1}^{k-n'}, U_{k-1}^{k-m'}, \eta_i(k), k] + v_i(k).$$

其中 $\eta_i(k)$是未知时变参数，$v_i(k)$是未知的随机部分，$Y_{k-1}^{k-n'}$，$U_{k-1}^{k-m'}$意义同前.

对于上述模型，根据前面指出的结果(参看[3])，总可以在输入-输出等价的意义下，改写成如下形式：

$$Y_k = \varphi_i(k)^\tau \theta_i(k).$$

其中

$$\varphi_i(k) = (Y_{k-1},\cdots,Y_{k-h_i},U_{k-1},\cdots,U_{k-h_i})^\tau,$$

而 $\theta_i(k)$是 $2h_i$ 维的随机时变参数. 于是

$$\mu' = \{Y_k = \varphi_i(k)^\tau\theta_i(k), \quad i = 1,2,\cdots,M\}$$

就形成了在输入-输出等价意义下的“替代模型集”.

置

$$n = \max_i h_i$$

及

$$\varphi(k) = (Y_{k-1},\cdots,Y_{k-n},U_{k-1},\cdots,U_{k-n})^\tau,$$

$$\theta(k) = (\alpha_1(k),\cdots,\alpha_n(k),\beta_1(k),\cdots,\beta_n(k))^\tau.$$

那么，模型

$$Y_k = \varphi(k)^\tau\theta(k) \tag{3}$$

是“替代模型集”μ'中阶数最大的模型.

为讨论方便，我们说模型

$$Y_k = \psi(k)^\tau\theta(k)$$

与

$$Y_k = \varphi(k)^\tau\eta(k)$$

是结构相同的，如果 $\psi(k)=\varphi(k), \quad \forall\ k$.

我们说模型

$$Y_k = \varphi(k)^\tau\theta(k)$$

是模型

$$Y_k = \psi(k)^\tau\eta(k)$$

的边际模型，如果

$$\psi(k) = [\varphi(k)^\tau \mid 0,\cdots,0]^\tau,$$

其中 $\varphi(k)$是 n 维向量，$\psi(k)$是 m 维向量，而 $n\leqslant m$.

应用上述术语，显然，“替代模型集”μ'中，每个模型都是模型(3)的边际模型. 由模型(3)得出相应的边际模型，只需令参数 $\theta(k)$的相应分量等于零就行了. 就这种意义而言，只要知道了模型(3)，那么一切边际模型都可以被确定. 在实际应用中，如果系统的真实模型不是模型(3)，而是某个边际模型，那么从模型(3)出发，依据观测数据对其参数进行估计，其中某些参数的估值将会接近于零. 所以，未知模型结构的模型的确定问题，就变成了模型(3)的阶的确定问题.

3 节省“能量”准则

由于在模型(3)

$$Y_k = \varphi(k)^\tau \theta(k)$$

中，$\theta(k)$是随机时变参数，那么，$\theta(k)-\theta(k-1)$就是$\theta(k)$在单位时间内的变化值．可以认为它表示了一种变化速率．如果$\hat{\theta}(k)$是$\theta(k)$的估值，那么$\hat{\theta}(k)-\hat{\theta}(k-1)$就是这种变化速率的估值，而$\|\hat{\theta}(k)-\hat{\theta}(k-1)\|^2$是这种速率估值的平方，所以可以看作为是一种“能量”估值的表示．

置

$$J_N(n) = \sum_{k=1}^{N} \|\hat{\theta}(k) - \hat{\theta}(k-1)\|^2.$$

此处N是观测的终止时刻，n是模型的阶数．于是，$J_N(n)$可以看成是总“能量”的估值．这种能量表示使得参数$\theta(k)$发生变化的总的能量消耗，它事实上是参数变化的“尺度”的一种度量．一个自然的想法是：对于描述一个动态系统的数学模型，如果它使得参数$\theta(k)$变化所消耗的总能量最小，则这个模型可以认为是适宜的．在这种意义下，我们引入如下的模型阶的判别准则，称之为节省能量准则，或称为参数变化尺度最小准则．

阶的判别准则：

如果n_0使$J_N(n)$对一切$n \geqslant 1$及一切可能的参数估值算法，皆能达到最小值：

$$J_N(n_0) = \text{最小}.$$

就说n_0是模型

$$Y_k = \varphi(k)^\tau \theta(k) \tag{4}$$

在节省能量意义下的最佳阶数．

文献[3]、[4]指出估计模型(4)中随机时变参数的一种适宜算法是

$$\hat{\theta}(k) = \hat{\theta}(k-1) + \frac{1}{\|\varphi(k)\|^2}\varphi(k)\{Y_k - \varphi(k)^\tau\hat{\theta}(k-1)\}. \tag{5}$$

由这一算法得出的估值$\hat{\theta}(k)$满足

$$Y_k = \varphi(k)^\tau\hat{\theta}(k). \tag{6}$$

事实上，我们有下述引理(参看[4]或[5])：

引理 在Y_k、$\varphi(k)$和$\hat{\theta}(k-1)$已给定的条件下，由算法(5)确定的$\hat{\theta}(k)$必满足条件：

$$Y_k = \varphi(k)^\tau\hat{\theta}(k),$$

并极小化指标函数：

$$J = \|\hat{\theta}(k) - \hat{\theta}(k-1)\|^2.$$

由此不难得出如下的推论：

推理 在观测数据及初值$\hat{\theta}(0)$均已给定的条件下，由递推算法(5)所确定的估值序列$\hat{\theta}(1),\hat{\theta}(2),\cdots,\hat{\theta}(N)$必满足约束条件：

$$Y_k = \varphi(k)^\tau\hat{\theta}(k), \quad k = 1,2,\cdots,N,$$

并极小化指标函数：

$$J_N(n) = \sum_{k=1}^{N} \|\hat{\theta}(k) - \hat{\theta}(k-1)\|^2.$$

其中N是观测的终止时刻，n是模型的阶数．

由上述推论不难看出，对阶的判别准则(节省能量准则)而言，在一切可能的参数估值

算法中，极小化 $J_N(n)$ 的是递推算法(5). 于是当把参数估值算法固定为算法(5)时，节省能量准则中的极小化手续，就简化成了仅对估值算法的初值 $\hat{\theta}(0)$ 及 n.

4　模型阶的确定准则

为方便计算，我们把回归向量

$$(Y_{k-1},\cdots,Y_{k-i},U_{k-1},\cdots,U_{k-i})^\tau$$

表示成为 $\varphi_i(k)$，相应的参数向量表示为 $\theta_i(k)$. 并把目标函数记为

$$J_N(i)=\sum_{k=1}^{N}\|\hat{\theta}_i(k)-\hat{\theta}_i(k-1)\|^2.$$

进一步引入记号

$$G_i=\{(\hat{\theta}_i(1),\cdots,\hat{\theta}_i(N))^\tau|Y_k=\varphi_i(k)^\tau\hat{\theta}_i(k),\quad k=1,2,\cdots,N\},$$

$$\theta'_N(I)=(\hat{\theta}_i(1),\cdots,\hat{\theta}_i(N))^\tau$$

置

$$J_N^*(i)=\min_{\hat{\theta}_i(0)}\ \min_{\theta'_N(i)\in G_i}J_N(i).$$

根据前节的推论，当把参数估值算法固定为算法(5)时，我们有

$$J_N^*(i)=\min_{\hat{\theta}_i(0)}J_N(i).$$

此时，节省能量准则变为选取 n_0，使其满足

$$J_N^*(n_0)=\min_i J_N^*(i).$$

下面的定理说明了这种手续只能在近似的意义下完成.

定理　在观测数据已给定的条件下，如果参数估值序列是由算法(5)给定的，那么必有

$$J_N^*(i)\geqslant J_N^*(i-1),\quad i=1,2,\cdots,N-1.$$

证　设 $\{\hat{\theta}_i(0),\hat{\theta}_i(1),\cdots,\hat{\theta}_i(N)\}$ 是由算法(5)所产生的估值序列，则必有

$$Y_k=\varphi_i(k)^\tau\hat{\theta}_i(k),\quad k=1,2,\cdots,N.$$

置

$$\hat{\theta}_{i+1}^*(k)=[\hat{\theta}_i(k)\ \vdots\ 0]^\tau,\quad k=0,1,\cdots,N$$

因为

$$\varphi_{i+1}(k)^\tau\hat{\theta}_{i+1}^*(k)=\varphi_i(k)^\tau\hat{\theta}_i(k)+0\cdot Y_{k-i-1}=\varphi_i(k)^\tau\hat{\theta}_i(k)=Y_k,$$

所以

$$(\hat{\theta}_{i+1}^*(1),\hat{\theta}_{i+1}^*(2),\cdots,\hat{\theta}_{i+1}^*(N))^\tau\in G_{i+1}.$$

不难看出

$$\sum_{k=1}^{N}\|\hat{\theta}_i(k)-\hat{\theta}_i(k-1)\|^2=\sum_{k=1}^{N}\|\hat{\theta}_{i+1}^*(k)-\hat{\theta}_{i+1}^*(k-1)\|^2\geqslant J_N^*(i+1),$$

应用前段的推论，即可得出

$$J_N^*(i)\geqslant J_N^*(i+1).\qquad \text{证毕.}$$

由于恒有 $J_N^*(i)\geqslant 0$，所以当 N 足够大时，我们可以得出模型阶的判定法则如下：

方法 1) 在容许的条件下，取 n_0，使其足够的大；

方法 2) 对于事先给定的小正数 ε，取 n_0 使得下述不等式成立：

$$J_N^*(n_0-1)-J_N^*(n_0)>\varepsilon,$$

$$J_N^*(j)-J_N^*(j+1)\leqslant\varepsilon,\quad 对\ j\geqslant n_0.$$

5 仿真运算结果

为了充分体现系统模型结构时变的特点，我们采用如下两个具有随机时变参数的模型：

$$Y_k = (1 + 0.9^k)e^{-\frac{1}{k^2+1}Y_{k-1}} + (1 - 0.95^k)Y_{k-2}U_{k-1},$$

$$Y_k = \ln(1 + \sqrt[3]{K}Y_{k-1}) + (1 - 0.01k)e^{0.9kU_{k-1}},$$

每隔 25～30 步，交替产生 200 个数据. 其中 U_k 是(0,2)区间上均匀分布的随机数. 然后，用"替代模型集"

$$Y_k = \varphi_i(k)^\tau\theta_i(k), \quad i = 1,2,\cdots,9$$

中的模型分别对数据进行拟合. 其中

$$\varphi_i(k) = (Y_{k-1},\cdots,Y_{k-i},U_{k-1},\cdots,U_{k-i})^\tau.$$

$\theta_i(k)$为相应的参数向量. 采用随机梯度算法进行参数估计，各参数初值均取为零. 对各拟合模型分别计算准则函数

$$J_N^*(i) = \sum_{k=1}^{N} \| \hat{\theta}_i(k) - \hat{\theta}_i(k-1) \|^2.$$

或写成递推形式

$$J_N^*(i) = J_{N-1}^*(i) + \| \hat{\theta}_i(N) - \hat{\theta}_i(N-1) \|^2.$$

结果如表 1 所示.

表 1

阶 i	$J_N^*(i)(N=50)$	$J_N^*(i)(N=100)$	$J_N^*(i)(N=150)$	$J_N^*(i)(N=200)$
1	21.9255	2828.97	3506.07	3692.89
2	4.12854	9.90008	19.4271	28.5612
3	3.11810	8.82524	17.0690	24.6207
4	2.33777	5.45439	10.2755	14.1189
5	1.91839	4.84129	9.96414	12.7685
6	1.64961	4.07408	6.94360	10.1142
7	1.50589	3.80193	6.07894	8.55003
8	1.33710	3.60744	5.70516	7.87273
9	1.26581	3.44391	5.50231	7.33325

从上表我们看出，仿真结果满足上段定理的结论：

$$J_N^*(i) \geqslant J_N^*(i+1).$$

以上结果说明，尽管我们没有对算法的初值 $\hat{\theta}(0)$极小化 J，所得结论仍正确.

6 结束语

本文针对模型结构未知的多模型系统，提出的结构辨识的一种简化方法，在实际应用中只能作为一种近似方法得以实现. 从仿真结果中，也可以看出这一点. 它实际上体现了一种逼近的思想.

参 考 文 献

[1] 韩志刚. 一类结构时变系统的辨识. 控制理论与应用, 1988, 5(4): 86—91

[2] 韩志刚. 结构随机变化系统的多层递阶预报. 自动化学报, 1990, 16(5): 423—428

[3] Han Zhigang. Multi-Level recursive identification method. Chinese Jaurnal of Automation, 1989, 1(1): 101—106

[4] 韩志刚. 多层递阶方法及其应用. 北京: 科学出版社, 1989

[5] Goodwin, G. C. and Sin, K. S.. Adaptive Filtering. Prodiction and Control. Prentice-Hall, INC Englewood Cliffs. New Jersey, 1984

A Kind of Reduced Approach to Multi-Model System Identification of Unknown Model Structure

HAN Zhigang and WANG Dejin

(Institute of Applied Mathematics, Heilongjiang University · Haerbin, 150080, PRC)

Abstract: In this paper, the identification problem of multi-model system including time-varying model structure system is considered. Here, we suppose that the model structure is unknown. A unified description method of model set of multi-model system, including the identification method of model structure and parameter, is given. Theoretical analysis and computer simulation to the proposed method are made.

Key words: multi-model; time-varying parameter; substitute model; spare energy criterion; determination of order

本文作者简介

韩志刚　1958 年毕业于吉林大学. 现任黑龙江大学教授. 主要从事系统辨识, 自适应预报与控制的研究, 曾提出多层递阶方法. 目前的研究领域是结构与参数皆时变的系统的辨识及自适应控制.

王德进　1982 年毕业于黑龙江大学物理系无线电专业. 现任黑龙江大学讲师. 主要从事系统辨识的研究. 目前的研究领域是系统模型结构的辨识.