模型辨识新方法及应用

史忠科

(西北工业大学自动控制系・西安, 710072)

摘要: 本文提出了一种有效的模型辨识新方法. 为了提高数值稳定性和计算效率, 本文给出了一种递阶最大信息量(AIC)新判据和参数估计新方法, 使单输出系统的极大似然方法及模型辨识的优选判据计算量成倍减少. 通过分析标量系统的 AIC 标准, 本文进一步导出了多输出情况下的 AIC 标准, 大大提高了计算效率.

关键词: 辨识判据; 系统辨识; 数值方法; 非线性系统建模

1 引 言

模型辨识是建模过程的关键技术[1,2]. 当模型结构的候选集确定之后, 模型辨识的优劣取决于建模准则和优选算法. 当模型结构比较准确时, 须采用 AIC 准则[1]. 此准则比较严格、准确, 但计算量很大, 按参数个数的阶乘来增长. 本文就是要解决它的计算效率和数值稳定性问题.

2 问题描述

设系统的状态方程和观测方程为

$$\begin{cases} \dot{x} = f(x,\Omega,t), \\ y(t) = g(x,\Omega,t), \\ z(k) = y(k) + V(k). \end{cases} \tag{1}$$

式中, $x \in \mathrm{R}^n$; $y \in \mathrm{R}^m$; $z(k)$ 为在 kT 时刻 y 的测量值; Ω 为未知参数向量, 维数未知; $V(k)$ 为量测噪声, 假定在 Ω 选定后为零均值高斯白噪声, 方差为 R. Ω 的维数(或 $f(x,\Omega,t)$, $g(x,\Omega,t)$ 的结构形式) 需要辨识, Ω 的值也需要辨识.

模型结构的确定应该依照一定的准则. 当对模型结构的准确度需要较高时, 人们常常使用最大信息量准则, 又称作赤池信息准则(AIC)[1]. 它是根据模型分布的密度函数导出的, 具体描述如下:

$$\mathrm{AIC} = -2\ln L + 2P. \tag{2}$$

式中, L 为极大似然函数, P 为模型中独立参数个数, 且

$$\ln L = -\frac{1}{2}\sum_{k=1}^{N} V^{\mathrm{T}}(k)R^{-1}V(k) - \frac{1}{2}N\ln|R| + \mathrm{const}. \tag{3}$$

该方法虽然比较严格, 但当 P 增大时, 计算复杂度呈指数增长.

---

* 霍英东基金资助项目.

本文于 1993 年 12 月 28 日收到, 1994 年 12 月 8 日收到修改稿.

3 递阶结构及新准则

为了减少 AIC 方法的计算量，采用了以下递阶结构.

设 $\Omega_j=[\theta_1,\theta_2,\cdots,\theta_j]^{\mathrm{T}}$ 已通过优选判据选入模型，现在考虑 θ_{j+1} 的选入或剔除问题.

求(3)式极大值，可得

$$\hat{\Omega}_j=A_j^{-1}b_j. \tag{4}$$

式中

$$A_j=\sum_{k=1}^{N}\left(\frac{\partial y(k)}{\partial \Omega_j^{\mathrm{T}}}\right)^{\mathrm{T}}R_j^{-1}\frac{\partial y(k)}{\partial \Omega_j^{\mathrm{T}}}=B_j^{\mathrm{T}}P_j^{-1}B_j,$$

$$b_j=\sum_{k=1}^{N}\left(\frac{\partial y(k)}{\partial \Omega_j^{\mathrm{T}}}\right)^{\mathrm{T}}R_j^{-1}[z(k)-y(k)],$$

$$P_j=\mathrm{diag}[R_j\quad R_j\quad\cdots\quad R_j],$$

$$B_j=\left[\left(\frac{\partial y(1)}{\partial \Omega_j^{\mathrm{T}}}\right)^{\mathrm{T}}\quad\left(\frac{\partial y(2)}{\partial \Omega_j^{\mathrm{T}}}\right)^{\mathrm{T}}\quad\cdots\quad\left(\frac{\partial y(N)}{\partial \Omega_j^{\mathrm{T}}}\right)^{\mathrm{T}}\right].$$

R_j 为当 $\Omega=\Omega_j$ 时对 R 的估计值.

若 θ_{j+1} 选入模型，则令

$$\Omega_{j+1}^{\mathrm{T}}=[\theta_{j+1}\quad\Omega_j^{\mathrm{T}}], \tag{5}$$

有

$$A_{j+1}=\begin{bmatrix}C_{j+1}^{\mathrm{T}}\\ B_j^{\mathrm{T}}\end{bmatrix}P_{j+1}^{-1}[C_{j+1}\quad B_j]=\begin{bmatrix}C_{j+1}^{\mathrm{T}}P_{j+1}^{-1}C_{j+1} & C_{j+1}^{\mathrm{T}}P_{j+1}^{-1}B_j\\ B_j^{\mathrm{T}}P_{j+1}^{-1}C_{j+1} & B_j^{\mathrm{T}}P_{j+1}^{-1}B_j\end{bmatrix}. \tag{6}$$

式中

$$C_{j+1}^{\mathrm{T}}=\left[\left(\frac{\partial y(1)}{\partial \theta_{j+1}}\right)^{\mathrm{T}}\quad\left(\frac{\partial y(2)}{\partial \theta_{j+1}}\right)^{\mathrm{T}}\quad\cdots\quad\left(\frac{\partial y(N)}{\partial \theta_{j+1}}\right)^{\mathrm{T}}\right].$$

令

$$\begin{cases}V_j(k)=z(k)-g(\hat{x}_j,\hat{\Omega}_j,k),\\ V_{j+1}(k)=z(k)-g(\hat{x}_{j+1},\hat{\Omega}_{j+1},k).\end{cases} \tag{7}$$

有

$$\begin{cases}\hat{R}_j=\dfrac{1}{N}\displaystyle\sum_{k=1}^{N}V_j(k)V_j^{\mathrm{T}}(k),\\[2ex] \hat{R}_{j+1}=\dfrac{1}{N}\displaystyle\sum_{k=1}^{N}V_{j+1}(k)V_{j+1}^{\mathrm{T}}(k).\end{cases}$$

则

$$\mathrm{AIC}(j+1)-\mathrm{AIC}(j)=2+N\ln\frac{\displaystyle\sum_{K=1}^{N}V_{j+1}^2(k)}{\displaystyle\sum_{k=1}^{N}V_j^2(k)},\quad m=1, \tag{9}$$

及

$$\begin{aligned}&\mathrm{AIC}(j+1)-\mathrm{AIC}(j)\\ &=2+N\ln\frac{\det\left[\displaystyle\sum_{k=1}^{N}V_{j+1}(k)V_{j+1}^{\mathrm{T}}(k)\right]}{\det\left[\displaystyle\sum_{k=1}^{N}V_j(k)V_j^{\mathrm{T}}(k)\right]}\\ &\quad+\sum_{k=1}^{N}[V_{j+1}^{\mathrm{T}}(k)R_{j+1}^{-1}(k)V_{j+1}(k)-V_j^{\mathrm{T}}(k)R_j^{-1}V_j(k)],\quad m>1.\end{aligned} \tag{10}$$

若 $\mathrm{AIC}(j+1)-\mathrm{AIC}(j)<0$，则 $\mathrm{AIC}(j+1)$ 趋于极小值，可将 θ_{j+1} 选入模型，否则剔除.

当 $m=1$ 时，我们可得

$$2 + N\ln\frac{R_{i+1}}{R_j} < 0, \tag{11}$$

即
$$G_j = \frac{R_{j+1}}{R_j} > e^{2/N}. \tag{12}$$

(12)式就是当 $z(k)$ 为标量时的 AIC 新准则.

当 $m > 1$ 时,(10)式的计算比较复杂,特别是当 m 较大时,右端第二项的计算量大得惊人!为了减少计算量,可对 R_j, R_{j+1} 采用 U-D 分解.

令
$$R_j = U_{Rj}D_{Rj}U_{Rj}^{\mathrm{T}}, \quad R_{j+1} = U_{R(j+1)}D_{R(j+1)}U_{R(i+1)}^{\mathrm{T}},$$

且 $U_{Rj}, U_{R(j+1)}$ 为单位上三角阵,$D_{Rj}, D_{R(j+1)}$ 均为对角阵,
$$D_{Rj} = \mathrm{diag}[D_{Rj}(1),\ D_{Rj}(2),\ \cdots,\ D_{Rj}(m)],$$
$$D_{R(j+1)} = \mathrm{diag}[D_{R(j+1)}(1),\ D_{R(j+1)}(2),\ \cdots,\ D_{R(j+1)}(m)].$$

代入(10)式中得
$$\mathrm{AIC}(j+1) - \mathrm{AIC}(j) = 2 + Q_j + N\ln\frac{\prod_{i=1}^{m} D_{R(j+1)}(i)}{\prod_{i=1}^{m} D_{Rj}(i)}. \tag{13}$$

式中
$$Q_j = \sum_{k=1}^{N}[V_{j+1}^{\mathrm{T}}(k)R_{j+1}^{-1}V_{j+1}(k) - V_j^{\mathrm{T}}(k)R_j^{-1}V_j(k)].$$

(13)式的判据还可写成
$$\prod_{i=1}^{m}\frac{D_{R(j+1)}(i)}{D_{Rj}(i)} > e^{\frac{2-Q_j}{N}}. \tag{14}$$

在(12),(14)式中,建立了判断准则与数据长度的关系,解决了行列式的计算量问题.

4 参数递阶估计新方法

在考虑 θ_{j+1} 的选入或剔除时,都必须涉及 Ω 的参数估计计算,为了减少计算量,本文采用如下的递阶估计方法.

令
$$A_j = U_jD_jU_j^{\mathrm{T}}, \quad A_{j+1} = U_{j+1}D_{j+1}U_{j+1}^{\mathrm{T}}. \tag{15}$$

式中,U_j, U_{j+1} 均为单位上三角阵;D_j, D_{j+1} 为对角阵,将(15)代入(6)式中,得
$$U_{j+1}D_{j+1}U_{j+1}^{\mathrm{T}} = \begin{bmatrix} C_{j+1}^{\mathrm{T}}P_{j+1}^{-1}C_{j+1} & C_{j+1}^{\mathrm{T}}P_{j+1}^{-1}B_j \\ B_j^{\mathrm{T}}P_{j+1}^{-1}C_{j+1} & B_j^{\mathrm{T}}P_{j+1}^{-1}B_j. \end{bmatrix}. \tag{16}$$

当 $m = 1$ 时,(16)式可简化为
$$U_{j+1}D_{j+1}U_{j+1}^{\mathrm{T}} = \begin{bmatrix} 1 & U_{12} \\ 0 & U_j \end{bmatrix}\begin{bmatrix} d_{j+1} & 0 \\ 0 & D_j\Lambda_j \end{bmatrix}\begin{bmatrix} 1 & 0 \\ U_{12}^{\mathrm{T}} & U_j^{\mathrm{T}} \end{bmatrix}. \tag{17}$$

根据(16),(17)式可得
$$\begin{aligned} U_{12} &= C_{j+1}^{\mathrm{T}}P_{j+1}^{-1}B_jU_j^{-\mathrm{T}}D_j^{-1}\Lambda_j^{-1}, \\ d_{j+1} &= C_{j+1}^{\mathrm{T}}P_{j+1}^{-1}(P_{j+1} - B_jU_j^{-\mathrm{T}}D_{j1}^{-1}U_j^{-1}B_j^{\mathrm{T}})P_{j+1}^{-1}C_{j+1}, \\ \theta_{j+1} &= d_{j+1}^{-1}C_{j+1}^{\mathrm{T}}P_{j+1}^{-1}(E_j - B_j\hat{\Omega}_j), \\ \Omega_j &= \hat{\Omega}_j - U_j^{-\mathrm{T}}U_{12}^{\mathrm{T}}\theta_{j+1}. \end{aligned} \tag{18}$$

式中,$D_{j1} = D_j\Lambda_j$,$\hat{\Omega}_j$ 为未考虑 θ_{j+1} 时对 Ω_j 的估计.

当 $m>1$ 时，假定 θ_{j+1} 的初值为零，算法与 $m=1$ 时完全相同.

5　计算效率分析比较

为了说明本文方法的计算效率，我们将新方法的计算量与常规方法进行了比较. 表 1 给出了 AIC 准则的计算量统计，表 2 给出了本文方法与普通递推方法的比较.

表 1　AIC 准则计算量的比较

算　法	加　　法	乘　　法	除法	对数
新方法	$Nm(m+1)+1$	$Nm(m+1)$	m	1
AIC 方法	$m!+mN(m+1)+m(m-1)^2$	$m!+m(m+1)N+m(m+1)^2$	$m(m-1)+1$	1

表 2　参数估计计算量的比较

算　法	加　　法	乘　　法	除　法
新 方 法	$(N-1)m(j+2)+j^2+3j-1$	$Nm(j+3)+j^2+3j+2$	$jm+1$
递推算法	$1.5[(j+1)^2+(j+1)]Nm$	$[1.5(j+1)^2+5.5(j+1)]NM$	Njm

表 1，表 2 中计算量均指一次性能指标计算及一次参数估计的计算量，j 为已经确认(或已处理)选入模型中的参数个数. 若 Ω 有 q 个候选参数，则 AIC 计算与参数辨识计算量分别如表 3，表 4 所示. 以乘法为例，说明计算量的统计方法. 新 AIC 准则的乘法总量计算式为

$$\sum_{l=1}^{q} Nm(m+1)=qNm(m+1).$$

表 3　AIC 准则总计算量比较

算　法	加　　法	乘　　法	除　法	对数
新方法	$qNm(m+1)+q$	$qNm(m+1)$	mq	q
AIC 方法	$qm!+qNm(m+1)+qm(m-1)^2$	$qm!+qNm(m+1)+qm(m+1)^2$	$qm(m-1)$	q

表 4　参数估计总计算量的比较

算　法	加　　法	乘　　法	除　　法
新方法	$(N-1)m\frac{1}{2}q(q+5)-q+\frac{3}{2}q(q+1)+\frac{1}{6}q(q+1)(2q+1)$	$Nm\left[3q+\frac{1}{2}q(q+1)\right]+2q+\frac{3}{2}q(q+1)+\frac{1}{6}q(q+1)(2q+1)$	$\frac{1}{2}mq(q+1)+q$
一般递推算法	$1.5Nm\left[\frac{1}{2}(q+1)(q+2)+\frac{1}{6}(q+1)(q+2)(2q+3)-2\right]$	$1.5Nm\left[\frac{1}{6}(q+1)(q+2)(2q+3)-1\right]+5.5Nm\left[\frac{1}{2}(q+1)(q+2)-1\right]$	$\frac{1}{2}Nmq(q+1)$

由于本文方法利用了原有估计的结果，由表 3，表 4 可知，本文提出的新方法有较高

的计算效率.

对于飞机短周期运动，若有10个输出，100个待选参数，数据长度为200时，新的参数估计方法的计算效率提高90倍！AIC指标计算效率提高100倍！

参考文献

[1] Akaike, H. A.. A New-Look at the Statistical Model Indentification. IEEE. Trans. Automat. Contr. , 1974, AC-19 (6): 716—722

[2] Bilings, S. A.. Indentification of Nonlinear Systems-A Survey. IEE Proc. Pt. D. , 1983, 127(6): 272—285

A New Approach for Model Identification and Its Application

SHI Zhongke

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

Abstract: A new efficient method for nonlinear model identification is presented in this paper. To get high numerical stability and computational efficiency, a new AIC criterion and recursive parameter estimation methods are developed which make both the maximum likelihood method and the optimal criterion for model choice of SISO system very simple. For MIMO system, a new AIC criterion is given and the computation formulas of the cost function is simplified. The results of aerodynamic model identification show that the new method can give accurate model structures of unknown system.

Key words: criterion for model identification; system identification; numeric method

本文作者简介

史忠科 1956年生. 西北工业大学自控系教授. 曾在南昌航空工业学院和试飞研究院从事过教学和科研工作. 先后在飞行试验动力学、最优估计等方面做了系统的研究工作. 目前主要从事随机系统、大系统鲁棒控制、智能控制、飞行力学等方面的研究工作.