第9卷第5期
1992年10月

控制理论与应用
CONTROL THEORY AND APPLICATIONS

Vol. 9, No. 5
Oct., 1992

线性多变量系统的联合辨识算法

丁 锋 谢新民
(清华大学自动化系·北京,100084)

摘要：本文提出了同时估计线性多变量系统所有参数的联合辨识算法(CIA)，并用随机过程理论分析了算法的收敛性. 与子系统辨识算法(SSIA)相比，CIA 的计算量要小得多. 各输出间存在相互作用噪声时的辨识问题通过使用递推广义增广最小二乘法(RGELS)[1]得到解决. 数字仿真结果表明了 CIA 的有效性.

关键词：多变量系统；参数估计；广义增广最小二乘；联合辨识算法

1 引 言

实际线性多变量过程(最小实现)都等价于下列 Guidorzi 辨识模型[2,3]

$$P(z^{-1})y(t)=Q(z^{-1})u(t)+e(t), \tag{1}$$

$$e(t)=T(z^{-1})v(t). \tag{2}$$

其中

$$P(z^{-1})=(p_{ij}(z^{-1}))=(\delta_{ij}+a_{ij}(1)z^{-1}+\cdots+a_{ij}(\nu_j)z^{-\nu_j})\in R^{m\times m},$$

$$Q(z^{-1})=(q_{ij}(z^{-1}))=(b_{ij}(1)z^{-1}+\cdots+b_{ij}(\nu)z^{-\nu})\in R^{m\times r},$$

$$\nu=\max(\nu_1,\cdots,\nu_m),\quad \delta_{ij}=\begin{cases}1, & i=j,\\ 0, & i\neq j.\end{cases}$$

噪声模型 $T(z^{-1})\in R^{m\times m}$ 为 z^{-1} 多项式阵或有理分式阵，$u(t)\in R^r$ 输入，$y(t)\in R^m$ 输出，$v(t)\in R^m$ 随机噪声，z^{-1} 为单位后移算子($z^{-1}y(t)=y(t-1)$).

$\{v(t)\}$ 是定义在概率空间 (Ω,F,P) 上的鞅差序列，且适应于递增 σ-代数 $(F_t,t\in N)$，其中 F_t 是由直到 t 时刻的观测数据生成的. $\{v(t)\}$ 满足假设

A1) $E[v(t)/F_{t-1}]=0$, a. s.

A2) $E[v(t)v^{\mathrm T}(t)/F_{t-1}]=\Gamma<\infty I$, a. s.

A3) $\sup_t \frac{1}{t}\sum_{s=1}^{t}\|v(s)\|^2<\infty$. a. s.

式中，上标 T 表转置，I 表单位阵，$\|X\|^2=\mathrm{tr}(X^{\mathrm T}X)$ (tr 表迹).

当 $T(z^{-1})$ 为下列三种特殊情况时，已有相应的算法估计模型(1)的参数：

a) $T(z^{-1})=0$ 或 I (Gauthier and Landau[4])；

b) $T(z^{-1})=P(z^{-1})$ (El-Sherief[5]的增广最小二乘法)；

c) $T(z^{-1})$ 为对角多项式阵(Sinha and Kwong[6]的广义最小二乘法 GLS).

以上诸方法一个共同特点，就是把模型(1)分解为 m 个子系统进行辨识(称之为子系统辨识算法 SSIA)，其计算量都很大. 而更一般的情形，$T(z^{-1})$ 为有理分式阵时，即当

本文于1991年2月9日收到. 1991年11月4日收到修改稿.

$$T(z^{-1}) = R^{-1}(z^{-1})D(z^{-1}),$$
$$R(z^{-1}) = r_{ij}(z^{-1}) \in R^{m\times m}, \quad D(z^{-1}) = (d_{ij}(z^{-1})) \in R^{m\times m}, \tag{3}$$
$$r_{ij}(z^{-1}) = \delta_{ij} + r_{ij}(1)z^{-1} + \cdots + r_{ij}(r_j)z^{-r_j}, \quad \text{etc.}$$

时，本文用 RGELS 方法解决了模型(1)的参数辨识问题．本文建立的联合辨识算法，可以同时估计模型(1)的所有参数，并运用随机过程理论分析了 CIA 估值收敛性．

2 联合辨识算法

利用(3)式，(2)式可以写成

$$R(z^{-1})e(t) = D(z^{-1})v(t). \tag{4}$$

置

$$\theta^{\mathrm T} = [\theta_1^{\mathrm T} \quad \theta_2^{\mathrm T}],$$
$$\theta_1^{\mathrm T} = \begin{bmatrix} a_{11} & \cdots & a_{1m} & b_{11} & \cdots & b_{1r} \\ & \vdots & & & \vdots & \\ a_{m1} & \cdots & a_{mm} & b_{m1} & \cdots & b_{mr} \end{bmatrix}, \quad \theta_2^{\mathrm T} = \begin{bmatrix} r_{11} & \cdots & r_{1m} & d_{11} & \cdots & d_{1m} \\ & \vdots & & & \vdots & \\ r_{m1} & \cdots & r_{mm} & d_{m1} & \cdots & d_{mm} \end{bmatrix}, \tag{5}$$
$$a_{ij} = [a_{ij}(1) \quad \cdots \quad a_{ij}(\nu_j)], \quad b_{ij} = [b_{ij}(1) \quad \cdots \quad b_{ij}(\nu)],$$
$$r_{ij} = [r_{ij}(1) \quad \cdots \quad r_{ij}(r_j)], \quad d_{ij} = [d_{ij}(1) \quad \cdots \quad d_{ij}(d_j)],$$
$$\varphi^{\mathrm T}(t) = [\varphi_1^{\mathrm T}(t) \quad \varphi_2^{\mathrm T}(t)],$$
$$\varphi_1^{\mathrm T}(t) = [-\bar{y}_1(t) \quad \cdots \quad -\bar{y}_m(t) \quad \bar{u}_1(t) \quad \cdots \quad \bar{u}_r(t)], \tag{6}$$
$$\varphi_2^{\mathrm T}(t) = [-\bar{e}_1(t) \quad \cdots \quad -\bar{e}_m(t) \quad \bar{v}_1(t) \quad \cdots \quad \bar{v}_m(t)],$$
$$\bar{u}_i(t) = [u_i(t-1) \quad \cdots \quad u_i(t-\nu)],$$
$$\bar{y}_i(t) = [y_i(t-1) \quad \cdots \quad y_i(t-\nu_i)]. \quad \text{etc.}$$

则(1)，(4)可以写为

$$y(t) = \theta_1^{\mathrm T}\varphi_1(t) + e(t), \tag{7}$$
$$y(t) = \theta^{\mathrm T}\varphi(t) + v(t). \tag{8}$$

仿照文[1]的 RGELS 方法，不难推出多变量系统的递推联合辨识算法(RCIA)为

$$\hat\theta(t) = \hat\theta(t-1) + P(t)\psi(t)[y^{\mathrm T}(t) - \psi(t)\hat\theta(t-1)], \tag{9}$$
$$P^{-1}(t) = P^{-1}(t-1) + \psi(t)\psi^{\mathrm T}(t), \tag{10a}$$

或

$$P(t) = P(t-1) - \frac{P(t-1)\psi(t)\psi^{\mathrm T}(t)P(t-1)}{1 + \psi^{\mathrm T}(t)P(t-1)\psi(t)}, \tag{10b}$$

$\hat\theta(0) =$ 很小正矩阵，

$P(0) = \mathrm{diag}[P_1(0) \quad P_2(0)], \quad P_1(0) = \alpha I, \quad P_2(0) = \beta I, \quad \alpha \gg 1, 0 < \beta \ll 1.$

其中

$$\psi^{\mathrm T}(t) = [\varphi_1^{\mathrm T}(t) \quad \psi_2(t)], \tag{11}$$
$$\psi_2^{\mathrm T}(t) = [-\hat{\bar{e}}_1(t) \quad \cdots \quad -\hat{\bar{e}}_m(t) \quad \hat{\bar{v}}_1(t) \quad \cdots \quad \hat{\bar{v}}_m(t)],$$
$$\hat{\bar{e}}_i(t) = [\hat e_i(t-1) \quad \cdots \quad \hat e_i(t-r_i)], \quad \hat{\bar{v}}_i(t) = [\hat v_i(t-1) \quad \cdots \quad \hat v_m(t-1)],$$
$$\hat e(t) = y(t) - \hat\theta_1^{\mathrm T}(t-1)\varphi_1(t), \tag{12}$$
$$\hat v(t) = y(t) - \hat\theta^{\mathrm T}(t-1)\psi(t), \tag{13}$$

或

$$\hat{e}(t)=y(t)-\hat{\theta}_1^{\mathrm{T}}(t-1)\varphi_1(t),\tag{14}$$

$$\hat{v}(t)=y(t)-\hat{\theta}^{\mathrm{T}}(t-1)\psi(t).\tag{15}$$

3 RCIA收敛性分析

在没有附加条件下，RCIA与RGELS[1]和RGLS[7]一样，可能产生局部收敛点. 本文对$\hat{\theta}(t)$作某种限制(同样，也可按照文[8]的方法对矩阵作某种修正)，运用随机过程理论分析算法(9)～(13)的收敛性，同时给出参数收敛的条件.

假设

i) (A1)～(A3)成立；

ii) $P(z^{-1})$稳定，$T(z^{-1})$稳定且逆稳定；

iii) $v(t)$为独立同分析的随机向量，且各态遍历；

iv) 设D_s是一个闭集，且$\theta\in D_s$. 将(9)修改为

$$\hat{\theta}'(t)=\hat{\theta}(t-1)+P(t)\psi(t)[y^{\mathrm{T}}(t)-\psi^{\mathrm{T}}(t)\hat{\theta}(t-1)],$$

如果 $$\hat{\theta}'(t)\in D_s,\ \text{则}\ \hat{\theta}(t)=\hat{\theta}'(t).$$

否则 $$\hat{\theta}(t)=\hat{\theta}(t-1),\quad P(t)=P(t-1).$$

v) 修正持续激励条件成立[8]

$$\lim_{t\to\infty}\frac{1}{t}\sum_{s=1}^{t}\psi(s)\psi^{\mathrm{T}}(s)=E[\psi_t\psi_t^{\mathrm{T}}]=G^{-1}>0.\tag{16}$$

定义参数估计误差$\tilde{\theta}(t)=\hat{\theta}(t)-\theta(0)$，且总认为$E\|\theta(0)\|^2\leqslant M_0<\infty$，由(9)得

$$\tilde{\theta}(t)=P(t)P^{-1}(0)\tilde{\theta}(0)+\sum_{s=1}^{t}P(t)\psi(s)[x(s)+v(s)]^{\mathrm{T}}$$
$$\triangleq A+B.\tag{17}$$

式中

$$x(t)=\theta_2^{\mathrm{T}}[\psi_2(t)-\varphi_2(t)],\quad \|x(t)\|^2\leqslant M<\infty.$$

利用公式 $2E[\mathrm{tr}(A^{\mathrm{T}}B)]\leqslant 2\sqrt{E[\mathrm{tr}(A^{\mathrm{T}}A)]E[\mathrm{tr}(B^{\mathrm{T}}B)]}$，有

$$E\|\tilde{\theta}(t)\|^2=E\|A+B\|^2=E[\mathrm{tr}(A^{\mathrm{T}}A)]+2\mathrm{tr}(A^{\mathrm{T}}B)+\mathrm{tr}(B^{\mathrm{T}}B)$$
$$\leqslant E\|A\|^2+2\sqrt{E\|A\|^2E\|B\|^2}+E\|B\|^2.\tag{18}$$

由(10a)和(16)可导出

$$\lim_{t\to\infty}\mathrm{tr}[P(t)]=\lim_{t\to\infty}\frac{\mathrm{tr}(G)}{t}=0.\tag{19}$$

所以

$$0\leqslant\lim_{t\to\infty}E\|A\|^2=\lim_{t\to\infty}E[\mathrm{tr}(\tilde{\theta}^{\mathrm{T}}(0)P^{-1}(0)P^2(t)P^{-1}(0)\tilde{\theta}(0))]$$
$$\leqslant\lim_{t\to\infty}E[(\mathrm{tr}(P(t)))^2\mathrm{tr}(P^{-2}(0))\|\tilde{\theta}(0)\|^2]$$
$$\leqslant\lim_{t\to\infty}\left[\frac{\mathrm{tr}(G)}{t}\right]\mathrm{tr}[P^{-2}(0)]M_0=0,\tag{20}$$

$$0\leqslant\lim_{t\to\infty}E\|B\|^2=\lim_{t\to\infty}E[\sum_{s,i}^{t}\psi^{\mathrm{T}}(s)P^2(t)\psi(i)(x(s)+v(s))^{\mathrm{T}}(x(i)+v(i))]$$

$$=\lim_{t\to\infty}\sum_{|s-i|\leqslant N}^{t}E[\psi^{\mathrm{T}}(s)P^2(t)\psi(i)(x(s)+v(s))^{\mathrm{T}}(x(i)+v(i))]\quad (\text{某 } N>\sum_{i=1}^{m}r_i)$$

$$=\lim_{t\to\infty}[(2N+1)t-N(N+1)]E|\psi^{\mathrm{T}}(s)P^2(t)\psi(i)(x(s)+v(s))^{\mathrm{T}}(x(i)+v(i))|,$$

$$(|s-i|\leqslant N).$$

利用公式 $2\psi^{\mathrm{T}}(s)Q\psi(i)\leqslant\psi^{\mathrm{T}}(s)Q\psi(s)+\psi^{\mathrm{T}}(i)Q\psi(i)$($Q$ 正定)和同分布假设以及 Frechet 定理可得

$$0\leqslant\lim_{t\to\infty}E\|B\|^2\leqslant\lim_{t\to\infty}[(2N+1)t-N(N+1)]E[\psi^{\mathrm{T}}(s)P^2(t)\psi(s)\|x(s)+v(s)\|^2]$$

$$\leqslant\lim_{t\to\infty}[(2N+1)t-N(N+1)]2E[\psi^{\mathrm{T}}(s)P^2(t)\psi(s)(\|x(s)\|^2+\|v(s)\|^2]$$

$$\leqslant\lim_{t\to\infty}2[(2N+1)t-N(N+1)]\left[\frac{\mathrm{tr}(G)}{t}\right]^2\mathrm{tr}(G^{-1})(M+\mathrm{tr}(\Gamma))]=0, \tag{21}$$

由此可得

$$\lim_{t\to\infty}\|\tilde{\theta}(t)\|^2=0, \tag{22}$$

即参数估计误差均方收敛于零.

同理,我们也可以证明算法(9)~(11),(14),(15)的收敛性.

4 数字仿真

例 1 考虑下列仿真对象

$$P(z^{-1})=\begin{bmatrix}1-0.3z^{-1} & -0.1z^{-1}\\ -0.1z^{-1} & 1-0.7z^{-1}\end{bmatrix},\quad Q(z^{-1})=\begin{bmatrix}1+0.3z^{-1} & 0.4z^{-1}\\ 0.7z^{-1} & 1+0.5z^{-1}\end{bmatrix},$$

$$R(z^{-1})=\begin{bmatrix}1-0.1z^{-1} & -0.1z^{-1}\\ -0.1z^{-1} & 1+0.2z^{-1}\end{bmatrix},\quad D(z^{-1})=\begin{bmatrix}1+0.1z^{-1} & 0.1z^{-1}\\ -0.2z^{-1} & 1+0.3z^{-1}\end{bmatrix}.$$

例 2 考虑下列仿真对象[6]

$$P(z^{-1})=\begin{bmatrix}1-0.7z^{-1}+0.125z^{-2} & 0.5z^{-1}-0.25z^{-2}\\ 0.5z^{-1}-0.6z^{-2} & 1-z^{-1}+0.24z^{-2}\end{bmatrix},$$

$$Q(z^{-1})=\begin{bmatrix}1-0.2z^{-1} & 0.1z^{-1}\\ 0.3z^{-1} & 1-0.5z^{-1}\end{bmatrix},$$

$$R(z^{-1})=\begin{bmatrix}1-0.7z^{-1} & 0\\ 0 & 1-0.8z^{-1}\end{bmatrix},$$

$$D(z^{-1})=I.$$

仿真时输入采用零均值单位方差不相关随机噪声向量序列,$\{v(t)\}$采用零均值白噪声向量序列,改变其方差可以控制噪信比 $N/S1$ 和 $N/S2$. 将 RCIA 用于辨识这两个系统的参数. 表 1 和表 2 示出了不同数据长度 L 和噪信比 F 的仿真结果,其中参数估值相对误差范数定义为 PEN$=\|\hat{\theta}_1-\theta_1\|/\|\theta_1\|$,$\theta_1$ 为系统模型参数真值,$\hat{\theta}_1$ 为 θ_1 的估计. 例 2 引自文[6],表 2 示出了 RCIA 与 RGLS 辨识结果的比较,RCIA 的估值精度高于 RGLS 的估值精度. 从表 1 中可以看出,随着 L 的增加,PEN 不断下降,这表明算法收敛的速度. 数字仿真表明 RCIA 可以给出系统模型参数的良好估计.

表 1 例 1 模型参数估计值

	$N/S1=11.61\%$ $N/S2=9.83\%$			
$L=$	300	500	1000	2000
$a_{11}(1)=-0.300$	−0.36567	−0.34127	−0.32964	−0.31863
$a_{12}(1)=-0.100$	−0.08189	−0.09399	−0.09263	−0.09636
$b_{11}(1)=\ 0.300$	0.30242	0.30077	0.29980	0.30041
$b_{12}(1)=\ 0.400$	0.40526	0.40566	0.39973	0.39945
$a_{21}(1)=-0.100$	−0.09667	−0.09560	−0.09307	−0.09822
$a_{22}(1)=-0.700$	−0.70451	−0.70253	−0.70191	−0.70050
$b_{21}(1)=\ 0.700$	0.69740	0.69722	0.69832	0.70086
$b_{22}(1)=\ 0.500$	0.50133	0.50268	0.49962	0.50501
PEN %	5.445	3.377	2.492	1.566
	$N/S1=37.17\%$ $N/S2=30.47\%$			
$L=$	300	500	1000	2000
$a_{11}(1)=-0.300$	−0.53701	−0.48160	−0.43001	−0.38060
$a_{12}(1)=-0.100$	−0.02717	−0.05394	−0.06131	−0.07973
$b_{11}(1)=\ 0.300$	0.30562	0.30762	0.30114	0.30224
$b_{12}(1)=\ 0.400$	0.40637	0.40950	0.39376	0.39689
$a_{21}(1)=-0.100$	−0.12430	−0.07974	−0.08682	−0.08414
$a_{22}(1)=-0.700$	−0.70000	−0.71325	−0.70665	−0.70861
$b_{21}(1)=\ 0.700$	0.69258	0.69177	0.69501	0.70292
$b_{22}(1)=\ 0.500$	0.50275	0.50774	0.49792	0.51533
PEN %	19.779	15.039	10.841	6.864

表 2 例 2 模型参数估计值

	$L=300$		$L=400$	
	$N/S1=9.8\%$	$N/S2=10.2\%$	$N/S1=29.4\%$	$N/S2=30.6\%$
	RCIA	RGLS*	RCIA	RGLS*
$a_{11}(1)=-0.750$	−0.7918	−0.7257	−0.8700	−0.7004
$a_{11}(2)=\ 0.125$	0.1428	0.1088	0.1700	0.0946
$a_{12}(1)=\ 0.500$	0.4924	0.4704	0.4831	0.3492
$a_{12}(2)=-0.250$	−0.2636	−0.2177	−0.2911	−0.1310
$b_{11}(1)=\ 1.000$	1.0007	0.9962	1.0067	0.9751
$b_{11}(2)=-0.200$	−0.2435	−0.1790	−0.3203	−0.1556
$b_{12}(1)=\ 0.000$	0.0000	0.0009	−0.0022	−0.0091
$b_{12}(2)=\ 0.100$	0.0935	0.0593	0.0844	−0.0643
$a_{21}(1)=\ 0.500$	0.4981	0.5000	0.6356	0.3916
$a_{21}(2)=-0.600$	−0.5915	−0.5394	−0.6357	−0.4035
$a_{22}(1)=-1.000$	−1.0191	−1.0220	−1.0659	−1.0200
$a_{22}(2)=\ 0.240$	0.2472	0.2278	0.3160	0.1736
$b_{21}(1)=\ 0.000$	−0.0053	0.0009	0.0069	0.0096
$b_{21}(2)=\ 0.300$	0.2846	0.1983	0.3988	0.1800
$b_{22}(1)=\ 10.000$	0.9963	0.9961	0.9950	0.9851
$b_{22}(2)=-0.500$	−0.5254	−0.5159	−0.5787	−0.4987
PEN %	3.382	6.326	12.634	16.811

* RGLS 参数估计值引自文[6].

5 结　语

多变量系统现有参数估计算法计算量大一直是辨识领域中有待解决的问题之一.本文提出的 RCIA,利用 RGELS 方法大大地减小了多变量系统辨识算法的计算量.并用随机过程理论分析了算法的收敛性,给出了参数估计收敛条件.该算法本质上是递推的,可用于在线辨识.

参考文献

[1] 丁　锋.辨识 Box-Jenkins 模型参数的递推广义增广最小二乘法.控制与决策,1990,5(6):53—56

[2] Guidorzi, R.. Canonical Structures in the Identification of Multivariable Systems. Automatica, 1975, 11:361—374

[3] El-Sherif, H. and Sinha, N. K.. Choices of Models for the Identification of Linear Multivariable Discrete-Time Systems. IEE Proc. PT. D, 1979, 126(12):1326—1330

[4] Gauthier, A. and Laudau, I. D.. On the Recursive Identification of Multi-Input, Multi-Output Systems. Automatica, 1978, 14:609—614

[5] El-Sherief, H.. Parametric Identification of a State-Space Model of Multivariable Systems Using the Extnded Least-Squares Method. IEEE Trans. Syst., Man, Cybern, 1981, 11(3):223—227

[6] Sinha, N. K. and Kwong, Y. H.. Recursive Estimation of the Parameters of Linear Multivariable Systems. Automatica, 1979, 15:471—475

[7] Soderstrom, T.. Convergence Properties of the Generalized Least Squares Identification Method. Automatica, 1974, 10:617—626

[8] Goodwin, G. C. and Sin, K. S. Adaptive Filtering Prediction and Control. Prentice-Hall, INC, Englewood Cliffs, New Jersey, 329—333

Combined Identification Algorithm for Linear Multivariable Systems

DING Feng and XIE Xinmin

(Department of Automation, Tsinghua University • Beijing, 10084, PRC)

Abstract: In this paper, the combined identification algorithm (CIA) for linear multivariable systems is presented which can simultaneously estimate all parameters of the whole system. The convergence of CIA is analyzed by means of stochastic process theory. The problem of identification in the presence of interactive noises at different outputs is solved by using the recursive generalized extended least-squares method. Compared with subsystem identification algorithm, CIA has much less amount of the calculation. The simulation examples are included.

Key words: multivariable system; parameter estimation; generalized extended least squares; combined identification algorithm

本文作者简介

丁　锋　见本刊 1992 年第 4 期第 409 页.

谢新民　见本刊 1992 年第 4 期第 409 页.