第 13 卷第 4 期
1996 年 8 月

控制理论与应用
CONTROL THEORY AND APPLICATIONS

Vol. 13, No. 4
Aug., 1996

含 ARMA 噪声系统模型的参数辨识方法*

王正明　易东云
(国防科技大学系统工程与数学系・长沙, 410073)

摘要: 实际问题中, 大量的动态系统控制问题可归结为含 MA, ARMA 噪声系统模型的参数辨识问题. 本文提出 RMA, RARMA 两种系统模型参数辨识的一种新方法, 主要手段是构造和研究特殊的辅助线性模型. 理论分析和实际计算表明, 本文方法较传统方法精度有明显提高.

关键词: 系统辨识; RARMA 模型

1 引　言

系统模型

$$y_i = \sum_{j=1}^{N} \beta_j \psi_j(t_i) + e_i, i = 1,2,\cdots,m; \quad \{e_i\} \text{为平稳零均值 ARMA 序列.} \tag{1}$$

在气象预报、动态系统数据分析、机械振动、航天测量等一系列实际问题中有广泛应用[1~5], 这里 $\sum_{j=1}^{N} \beta_j \psi_j(t)$ 是系统中所含确定性趋势部分的函数描述, 其中 $\psi_j(t)(j = 1,2,\cdots,N)$, 是一组基函数, e_i 为系统噪声. 模型(1)的参数辨识包括两个问题: 参数 $\beta = [\beta_1,\beta_2,\cdots,\beta_N]^{\mathrm{T}}$ 的辨识和平稳白色噪声的时序参数辨识. 由于系统噪声 e_i 的不可观测性, 该问题到目前为止尚未得到有效的解决.

记 $Y = [y_1,y_2,\cdots,y_m]^{\mathrm{T}}, e = [e_1,e_2,\cdots,e_m]^{\mathrm{T}}, x_{ij} = \psi_j(t_i), (i = 1,2,\cdots,m; j = 1,2,\cdots,N), X = (x_{ij})_{m\times N}$; 则模型(1) 可改写为

$$Y = X\beta + e, \quad \{e_i\} \text{为平稳的零均值 ARMA}(p,q) \text{ 序列.} \tag{2}$$

本文首先讨论了 $\{e_i\}$ 为零均值 MA(q) 序列的情况, 利用 MA(q) 序列自相关函数的特点, 我们构造一种特殊的辅助线性回归模型, 给出了时间序列参数及 β 的估计. 对于 $\{e_i\}$ 为 ARMA(p,q) 序列的情况, 一方面沿用上述思想, 另一方面利用回归分析中自变量选择的思想和非线性最小二乘估计的方法, 给出了系统参数的辨识. 为了实用方便, 我们还给出了配套的算法. 理论分析和仿真计算表明, 本文方法精度较传统方法有明显提高. 将本文方法应用到航空客票数据的分析和预测中, 取得令人满意的结果.

2 RMA 模型的参数辨识

考虑模型(2)的特殊情况

$$Y = X\beta + e, \quad \{e_i\} \text{为零均值 MA}(q) \text{ 序列.} \tag{3}$$

此时,

$$Ee = 0, \quad \mathrm{cov}(e) \equiv K = (\sigma_{ij})_{m\times m},$$

* 国防预研基金资助及湖南省自然基金资助项目.
本文于 1995 年 5 月 17 日收到. 1995 年 12 月 18 日收到修改稿.

$$\sigma_{ij}=\begin{cases}0, & |i-j|>q,\\ r_{|i-j|}, & |i-j|\leqslant q.\end{cases}\tag{4}$$

对模型(3)，若K已知，则$\beta_{\rm LS}=(X^{\rm T}K^{-1}X)^{-1}X^{\rm T}K^{-1}Y$为$\beta$的方差一致最小的线性无偏估计[6]. 因而，若能给出K较精确的估计$\hat{K}$，则可用$\hat{\beta}=(X^{\rm T}\hat{K}^{-1}X)^{-1}X^{\rm T}\hat{K}^{-1}Y$作为$\beta$的估计. 由(4)式知，只要能给出$r=[r_0,r_1,\cdots,r_q]^{\rm T}$的估计，就可给出$K$的估计. 另外，当$\{e_i\}$为可逆的MA($q$)序列时，由$r$的估计，也可唯一地给出MA($q$)序列的$q+1$个参数的估计值[3].

以下，我们主要讨论r的估计. 记$H=(h_{ij})_{m\times m}=X(X^{\rm T}X)^{-1}X^{\rm T}$，$A=(a_{ij})_{m\times m}=I-H$，$\xi=AY$，事实上

$$\xi=(I-H)Y=(I-H)(X\beta+e)=(I-H)e.\tag{5}$$

从(5)式可以看到，尽管e是不可观测的，但我们可由Y的数据及模型(3)得到$\xi=(I-H)e$的数据，而ξ与β无关.

以下，我们利用ξ的数据给出r的估计.

引理 1 设K由(4)式给出，$u=[u_1,u_2,\cdots,u_m]^{\rm T}$，$v=[v_1,v_2,\cdots,v_m]^{\rm T}$；则

$$u^{\rm T}Kv=\sum_{j=1}^{q+1}d_jr_{j-1}.$$

其中 $$d_j=\sum_{i=1}^{m-j+1}(u_iv_{i+j-1}+u_{i+j-1}v_i),\quad(1\leqslant j\leqslant q+1).$$

记

$$s=(m+1)(k-1)-\frac{k}{2}(k-1)+(l-k+1),\ g_s=\xi_k\xi_l,1\leqslant k\leqslant l\leqslant m;$$

$$z_{sj}=\sum_{i=1}^{m-j+1}(a_{ki}a_{li+j-1}+a_{ki+j-1}a_{li}).\tag{6}$$

由(5)式可知，${\rm E}\xi\xi^{\rm T}=(I-H)K(I-H)$. 综合以上讨论，有

$$Eg_s=\sum_{j=1}^{q+1}z_{sj}r_{j-1},\quad s=1,2,\cdots,\frac{m(m+1)}{2}\tag{7}$$

记$g=[g_1,g_2,\cdots,g_{m(m+1)/2}]^{\rm T}$，$Z=(z_{sj})$为$\dfrac{m(m+1)}{2}\times(q+1)$矩阵；由(7)式可知

$$g=Zr+\eta,\quad {\rm E}\eta=0.\tag{8}$$

下面，我们利用模型(8)给出r的线性无偏估计.

引理 2 设$X=X_{m\times N}$，${\rm rank}(X)=N$，$H=X(X^{\rm T}X)^{-1}X^{\rm T}$，$A=(a_{ij})_{m\times m}=I-H$，$Z$的元素由(6)式给出，$m\geqslant N+q+1$；则$Z^{\rm T}Z$为正定矩阵.

证 注意到$Z^{\rm T}Z$为常值矩阵，要证明$Z^{\rm T}Z$为正定矩阵，只须证明：当$r\neq0$时，$\|Zr\|\neq0$，由(6)式可知，Zr的$\dfrac{m(m+1)}{2}$个元素正好对应于对称矩阵$(I-H)K(I-H)$的下三角(包括对角线)部分的全部元素，因此$\|Zr\|=0\Leftrightarrow(I-H)K(I-H)=0$. 这样，引理的证明转化为证明：

$$r\neq0\Rightarrow{\rm rank}[(I-H)K(I-H)]>0.\tag{9}$$

注意到$r\neq0$时${\rm rank}(K)\geqslant m-q$，又由引理条件知$m-q\geqslant N+1$，故，${\rm rank}(K)\geqslant N+1$. 由矩阵分析的知识可知，存在正交阵$P$使$I-H=PDP^{\rm T}$，$D={\rm diag}(0,\cdots,0,1,\cdots,1)$($N$

个0,$m-N$个1).记$B=P^{\mathrm{T}}KP$,那么,$\mathrm{rank}(B)=\mathrm{rank}(K)\geqslant N+1$,于是$r\neq 0$时,$\mathrm{rank}[(I-H)K(I-H)]=\mathrm{rank}(DP^{\mathrm{T}}KPD)=\mathrm{rank}(DBD)\geqslant 1$. 由此即知(9)式.　　证毕.

综合上述讨论,我们有

定理 1　设$\mathrm{rank}(X)=N, m\geqslant N+q+1$;则$\hat{r}=(Z^{\mathrm{T}}Z)^{-1}Z^{\mathrm{T}}g$为$r$的线性无偏估计.

很明显,MA(q)模型为(3)的特例. 此时,$A=I$,Z的无素只取0和1两个值,并且每行只有一个元素,取值为1,第j列有$m-j+1$个元素为1而其余的元素为0,此时Z为正交矩阵,并且,由定理1给出的$\hat{r}$与通常的矩估计是一样的.

3　RARMA 模型的辨识

考虑模型

$$Y=X\beta+e,\quad \{e_i\}\text{ 为平稳零均值 ARMA}(p,q)\text{ 序列}. \tag{10}$$

记$r=(r_0,r_1,\cdots,r_{m-1})^{\mathrm{T}}, K=(r_{|i-j|})_{m\times n}=\mathrm{cov}(e)$. 与模型(8)的推导相同,我们有

$$g=Zr+\eta,\quad E\eta=0. \tag{11}$$

模型(11)与模型(8)的区别在于:模型(11)中,Z的列数和待估参数的个数都为m(待估参数比模型(8)多),而且模型(11)中$Z^{\mathrm{T}}Z$为奇异矩阵. 直接利用模型(11)估计r是困难的,而且精度不高.

注意到$\{e_i\}$为平稳的零均值ARMA(p,q)序列,因而r有如下特点[3]:

$$|r_j|\leqslant T_1e^{-T_2j},\quad T_1,T_2\text{ 为正值常数}, \tag{12}$$

$$-r_j=a_1r_{j-1}+a_2r_{j-2}+\cdots+a_pr_{j-p},\quad (j>p). \tag{13}$$

这里,$a=[a_1,a_2,\cdots,a_p]^{\mathrm{T}}$为ARMA($p,q$)中AR部分的参数. 以下的分析表明,这些特点对估计r是有利的. 下面,我们分析Z矩阵的特点.

引理 3　设Z_j为Z的第j列,则

$$\|Z_j\|^2\leqslant m-j+1,\quad j=1,2,\cdots,m. \tag{14}$$

证　以$\|U\|_F^2$记方阵U的全体元素的平方之和,令

$$S=(s_{kl})_{m\times m},\quad s_{kl}=\begin{cases}0, & |k-l|\neq j,\\ r_j, & |k-l|=j;\end{cases} \tag{15}$$

那么,由Z矩阵的构造可知

$$\|Z_1r_0\|^2=\|(I-H)(r_0I)(I-H)\|_F^2=r_0^2\|(I-H)\|_F^2\leqslant mr_0^2,$$

$$\|Z_jr_{j-1}\|^2=\frac{1}{2}\|(I-H)S(I-H)\|_F^2,\quad (j\geqslant 2).$$

由于对任何实值向量a,$\|(I-H)a\|^2=\|a\|^2-\|Ha\|^2$. 故当$j\geqslant 2$时

$$\|Z_jr_{j-1}\|^2\leqslant\frac{1}{2}\|S(I-H)\|_F^2\leqslant\frac{1}{2}\|S\|_F^2=(m-j+1)r_j^2.$$

由于$r_j(j=0,1,\cdots,m-1)$可取任意实数,综合以上各式得(14)式.　　证毕.

从(12)式和(14)式可以看到,随着j的增长,$\|Z_j\|^2$和r_j^2分别呈等差级数和等比级数递减,因此,当j充分大时,可认为$\|Z_jr_{j-1}\|^2\approx 0$. 进一步地,把(15)式改为

$$S=(s_{kl})_{m\times m},\quad s_{kl}=\begin{cases}0, & |k-l|\leqslant n-1;\\ r_j, & |k-l|\geqslant n.\end{cases}$$

沿用引理3方法,可以证明

$$\| \sum_{j=n+1}^{m} Z_j r_{j-1} \|^2 \leqslant \sum_{j=n+1}^{m} (m-j+1) r_{j-1}^2. \tag{16}$$

综合上述论可知，当 n 充分大时

$$\| \sum_{j=n+1}^{m} Z_j r_{j-1} \|^2 \approx 0. \tag{17}$$

在(17)式中，n 究竟取多大合适，要参考$[a_1,a_2,\cdots,a_n]^{\mathrm{T}}$ 的值，一般取

$$n = \min\{m-N-p-q-1, N+q+1+2p\}.$$

根据(17)式，$Z_p=[Z_1,Z_2,\cdots,Z_n]$，$r_p=[r_0,r_1,\cdots,r_{n-1}]^{\mathrm{T}}$；模型(11)可改为

$$g = Z_p r_p + \eta. \tag{18}$$

根据选择最优回归模型的思想[7,8]，模型(18)更有利估计 r_p 容易证明

引理 4 设 $m > n+N$，则 $Z_p^{\mathrm{T}} Z_p$ 为正定矩阵.

由引理 4 及模型(18)，我们可得 r_p 的 LS 的估计

$$\hat{r}_p = (Z_p^{\mathrm{T}} Z_p)^{-1} Z_p^{\mathrm{T}} g. \tag{19}$$

由 $\mathrm{E}\eta = 0$ 及(17)式，我们可把 $\hat{r}_p$ 视为 r_p 的无偏估计.

至此，我们利用了 r 的特点(12)式及 Z 矩阵的特点，给出了 r_p 的估计；但这个估计仍未用上(13)式，下面，我们再进一步改进 r_p 的估计 $\hat{r}_p$.

事实上 $\hat{r}_p$ 是极值问题

$$\| g - Z_p r_p \|^2 = \min \tag{20}$$

的解，若利用约束条件(14)式，则极值问题变为：

$$\begin{cases} r_j = -a_1 r_{j-1} - \cdots - a_p r_{j-p}, \quad (j>q), \\ F(r_0,r_1,\cdots,r_p,a_1,a_2,\cdots,a_p) = \| g - Z_p r_p \|^2 = \min. \end{cases} \tag{21}$$

注意到 $Z_p^{\mathrm{T}}(g - Z_p \hat{r}_p) = 0$，于是 $\| g - Z_p r_p \|^2 \equiv \| g - Z_p \hat{r}_p \|^2 + \| Z_p(\hat{r}_p - r_p) \|^2$.

设 $B=(b_{ij})_{n\times n} = Z_p^{\mathrm{T}} Z_p$，极值问题(21)又可改写为如下的等价形式：

$$\begin{cases} r_j = -a_1 r_{j-1} - a_2 r_{j-1} \cdots - a_p r_{j-p}, \quad (j>q), \\ G(r_0,r_1,\cdots,r_q,a_1,a_2,\cdots,a_q) \equiv \sum_{i,j=1}^{m} b_{ij}(r_{i-1}-\hat{r}_{i-1})(r_{j-1}-\hat{r}_{j-1}) = \min. \end{cases} \tag{22}$$

RARMA 模型参数辨识，按如下的步骤进行

Setp 1 求 $H, A = I - H, \xi = AY$；

Setp 2 求 $B = Z_p^{\mathrm{T}} Z_p,\quad f = Z_p^{\mathrm{T}} P$；

Setp 3 求 $\hat{r} = B^{-1} f$；

Setp 4 求解极值问题(22)，得到 ARMA(p,q) 序列的待估参数的估计值 $\bar{r}_0,\bar{r}_1,\cdots,\bar{r}_q,\bar{a}_1,\bar{a}_2,\cdots,\bar{a}_p$；

Setp 5 求 $\bar{r}_j = -\bar{a}_1 r_{j-1} - \bar{a}_2 r_{j-2} - \cdots - \bar{r}_p \bar{r}_{j-p}, (j>q)$；$\hat{K} = (\hat{\sigma}_{ij})_{m\times m}, \hat{\sigma}_{ij} = \bar{r}_{|i-j|}$；

Setp 6 $\hat{\beta} = (X^{\mathrm{T}} \hat{K}^{-1} X)^{-1} X^{\mathrm{T}} \hat{K}^{-1} Y$.

4 仿真计算与实例分析

例 1 设模型(1)中

$$\begin{cases} y_i = \beta_1 \sin t_i + \beta_2 e^{-t_i} + \beta_3 \ln t_i + \beta_4 t_i^2 + e_i, \quad t_i = 0.005i, \\ e_i + 0.2e_{i-1} + 0.1e_{i-2} = \varepsilon_i + 0.5\varepsilon_{i-1} + 0.1\varepsilon_{i-2}, \quad i = 1,2,\cdots,200. \end{cases} \tag{23}$$

即系统噪声 e 为 ARMA(2.2) 序列. 表 1 给出了应用本文方法及通常的近似方法[3] 对模型(23) 中参数估值的对比. 结果表明本文方法的参数估计精度有较明显的提高.

表 1　例 1 的参数辨识

	β_1	β_2	β_3	β_4	r_0	r_1	r_2	a_1	a_2
真　值	9.0	8.0	7.0	6.0	1.094	0.283	−0.066	0.200	0.100
本文方法	9.03	8.67	7.14	6.01	1.058	0.261	−0.054	0.230	0.092
传统方法	8.14	8.97	6.30	6.85	0.894	0.213	−0.049	0.258	0.131

应用本文方法对[3] 中某国际航空公司客票数的数据集 $\{y_i, i = 1,2,\cdots,144\}$ 进行分析. 由于该数据序列具有指数函数趋势, 取 $x_i = \ln y_i, i = 1,2,\cdots,144$, 用 $\beta_0 + \beta_1 \frac{t}{12} + \beta_2 (\frac{t}{12})^2 + \sum_{k=1}^{5} (\beta_{2k+1} \cos\omega_k t + \beta_{2k+2} \sin\omega_k t), \omega_k = \frac{2\pi}{12}$, 拟合 $\{x_i\}$ 得 $\{\hat{x}_i\}$, 残差 $\mathrm{RSS} = \sum_{i=1}^{144} (y_i - e^{\hat{x}_i})^2 = 24936$, 小于[3] 中用模型 $R_1 e^{r_1 t} + \sum_{k=1}^{5} B_k e^{b_k t} (C_k \sin\omega_k t + (\sqrt{-C_k^2}) \cos\omega_k t)$ 非线性拟合 $\{y_i\}$ 所得残差平方和 28908. 进一步对残差数据 $\{x_i - \hat{x}_i\}$ 的自相关及偏相关分析表明该数据系统噪声可用 ARMA(1,3) 描述. 由此, 可对 $\{x_i\}$ 建立如下 RARMA 模型:

$$\begin{cases} x_i = \beta_0 + \beta_1 \frac{t}{12} + \beta_2 (\frac{t}{12})^2 + \sum_{k=1}^{5} (\beta_{2k+1} \cos\omega_k t + \beta_{2k-2} \sin\omega_k t + e_i, \\ e_i \text{ 为 ARMA(1,3) 序列}, e_i - \varphi_1 e_{i-1} = \varepsilon_i - \theta_1 \varepsilon_{i-1} - \theta_2 \varepsilon_{i-2} - \theta_3 \varepsilon_{i-3}. \end{cases} \tag{24}$$

应用数据集中前 11 年的 132 个数据对模型(24)进行参数辨识结果列于表 2

表 2　航空客票数据集模型(24)的参数辨识

β_0	β_1	β_2	β_3	β_4	β_5	β_6	β_7	β_8
4.736	0.158	0.003	−0.049	−0.142	0.079	−0.023	−0.009	0.027
β_9	β_{10}	β_{11}	β_{12}	φ_1	θ_1	θ_2	θ_3	
0.026	0.022	0.021	0.006	0.459	−0.158	−0.265	−0.097	

将表 2 中辨识所得参数代入(24), 对数据集中第 12 年的客票数进行一步预报, 实际值和所得预得值列于表 3.

表 3　模型(24)的预报

实际值	417	391	419	461	472	535	622	606	508	461	390	432
一步预报值	418	396	432	452	475	536	606	598	509	451	392	434

由表 3 的数据结果可知, 预报的根方差为 7.7, 效果令人满意.

参 考 文 献

[1] Pandit, S. M., Wu, S. N.. Time Series and System Analysis with Applications. John Wily and Sons, 1983
[2] 徐南荣等. 系统辨识. 南京: 南京大学出版社, 1991
[3] 杨位饮, 顾岚. 时间序列分析与动态数据建模. 北京: 北京理工大学出版社, 1988

[4] 杨叔子. 时间序列分析的工程应用. 武汉：华中理工大学出版社，1992
[5] 吴澄，彭允许. 陀螺随机漂移率建模研究. 宇航学报，1993，(1)：21－28
[6] 陈希孺，王松桂. 近代回归分析——原理、方法及应用. 合肥：安徽教育出版社，1987
[7] 沙钰，王正明. 建立最优线性回归模型的一种新方法和算法. 数理统计与应用概率，1991，(4)：325－332
[8] 王正明，回归系数的改进主成分估计. 数学的实践与认识，1990，(1)：59－64

A New method to Parameters Identification for System Models with ARMA Noise

WANG Zhengming　and　YI Dongyun

(Department of System Engineering and Mathematics, National University of Defense Thchnology • Changsha, 410073, PRC)

Abstract: A lot of actual questions in system model analysis and control can be summed up in parameters identification for system models with ARMA noise. In this paper, we present a new method to parameters identification for system models with ARMA noise by use of constructing the special assistant linear model. Theory analysis and simulations show that the new method has much better precision than usual methods. Applying this method to an actual system model analysis and prediction, we obtain satisfying results.

Key word: system identification; RARMA model

本文作者简介

王正明　1962年生. 1986年毕业于华中理工大学应用数学系，获硕士学位. 现为国防科技大学系统工程与数学系教授. 已获部委级二等奖三项，发表学术论文35篇，合作出版著作两本. 目前主要从事应用数学教学和动态系统分析与测量数据处理方面的研究工作.

易东云　1965年生. 1992年毕业于国防科技大学系统工程与数学系概率统计专业，获硕士学位，讲师. 合作出版专著一本，已生发表学术论文16篇. 目前主要从事动态系统分析、最优估计以及测量数据处理等方面的研究工作.