

新的递推有界 GM 回归估计算法

成立花<sup>1</sup>, 张俊敏<sup>2</sup>

(1. 西安工程大学 理学院, 陕西 西安 710048;  
2. 西安建筑科技大学 理学院, 陕西 西安 710055)

**摘要:** 提出一种新的递推有界广义极大似然类(GM)回归估计器,新估计器所用的风险函数基于更一般的框架,并采用有界 M-估计函数. 设计一个新的权函数拒绝或降低异常点对估计结果的影响,并增加一个增广项,提出一种具有较强自适应能力的面向自回归(AR)模型参数估计的算法. 仿真结果表明:提出的 GM 回归估计器及面向 AR 模型的算法对异常点不利影响(主要来自于回归变量中的加性异常点)的抑制效果均优于其他 GM 估计器;在参数不做任何调整的情况下,面向 AR 模型的算法对非平稳环境下的估计具有良好的估计精度和收敛性.

**关键词:** GM 估计器; 鲁棒估计; AR 模型; 加性异常点

**中图分类号:** TN 911.7; O 241.2; N 945.14 **文献标志码:** A

最小二乘法(LS)和递推的最小二乘法(RLS)是系统辨识中重要的算法. 然而,在实际应用中,由于各种干扰因素的存在,辨识所需的数据往往会被污染,包含数量和类型均未知的异常点. 这些异常点使估计精度降低,甚至完全失效. 解决这一问题通常有两种方式:从数据中找出异常点并剔除掉,用剩余的数据进行辨识;发展替代 LS 和 RLS 的鲁棒回归算法. 由于后者简便实用,备受学者关注,已发展出了许多算法<sup>[1-9]</sup>. 在自回归模型的参数估计中,数据中的异常点通常被分为两个基本类<sup>[10]</sup>:加性异常点(第一型的异常点)与革新异常点(第二型的异常点). 革新异常点在系统运行过程中遵循了真实系统的关系,因此,就系统的辨识而言,其不利影响较小. 然而,当数据中包含较大的加性异常点型杠杆点时,这些算法的估计性能会严重下降,甚至失效. 针对广义极大似然类(GM)估计器中存在的问题,本文提出新的递推有界 GM 回归估计算法.

1 问题描述

考虑模型  $y_i = \mathbf{x}_i \boldsymbol{\beta} + v_i, i = 1, 2, \cdots, n$ . 其中:  $\boldsymbol{\beta} = [b_1, b_2, \cdots, b_p]^T$ ;  $y_i$  是第  $i$  时刻响应变量观测值;  $\mathbf{x} = [x_{i,1}, x_{i,2}, \cdots, x_{i,p}]$ ;  $v_i$  是相互独立,且尺度参数  $\sigma$  未知的干扰项. 所用的数据记作  $(\mathbf{X}_n, \mathbf{Y}_n)$ . 其中:  $\mathbf{X}_n = [\mathbf{x}_1^T, \mathbf{x}_2^T, \cdots, \mathbf{x}_n^T]^T, \mathbf{Y} = [y_1, y_2, \cdots, y_n]^T$ . 获得第  $n$  时刻的参数估计  $\hat{\boldsymbol{\beta}}_n$  后,第  $i$  时刻相应的残差记作  $r_{n,i} = y_i - \mathbf{x}_i \hat{\boldsymbol{\beta}}_n$ . 而在迭代法中,相应于起始估计  $\hat{\boldsymbol{\beta}}_n^{(0)}$  在第  $i$  时刻的残差记作  $r_{n,i}^{(0)} = y_i - \mathbf{x}_i \hat{\boldsymbol{\beta}}_n^{(0)}, i = 1, 2, \cdots, n$ .

采用文献[9]中的一般框架,并加以改进的风险函数为

$$J_n(\boldsymbol{\beta}) = \sum_{i=1}^n \lambda^{n-i} \rho(\frac{r_{n,i}}{\sigma d(\mathbf{x}_i)}) d^2(\mathbf{x}_i) \xi(\mathbf{x}_i) \sigma^2. \tag{1}$$

式(1)中:  $\lambda$  为遗忘因子,  $0 < \lambda \leq 1$ ;  $d(\mathbf{x}_i), \xi(\mathbf{x}_i)$  为待定函数;  $\rho(\cdot)$  是一个改进于 Huber 函数的有界 M-估计函数,即

$$\rho(t)=\begin{cases}t^2/2, & |t|\leqslant M_1, \\ M_1|t|-M_1^2/2, & M_1<|t|\leqslant M_2, \\ M_2M_1-M_1^2/2, & M_2<|t|. \end{cases}$$

上式中: $M$  为常值参数. 新估计器是式(1)的解.

2 新的 GM 估计器

相应于式(1)的“正规方程”为

$$\sum_{i=1}^n\lambda^{n-i}\psi(\frac{r_{n,i}}{\sigma d(\mathbf{x}_i)})d(\mathbf{x}_i)\xi(\mathbf{x}_i)\alpha\mathbf{x}_i^T=0. \tag{2}$$

式(2)中: $\psi=\rho'$ .

定义权函数为

$$w(t)=\begin{cases}1, & |t|\leqslant M_1, \\ M_1\text{sign}(t)/t, & M_1<|t|\leqslant M_2, \\ 0, & M_2<|t|. \end{cases}$$

上式中: $\text{sign}(\cdot)$ 是符号函数. 因此,式(2)可写为

$$\sum_{i=1}^n\lambda^{n-i}w(\frac{r_{n,i}}{\sigma d(\mathbf{x}_i)})\xi(\mathbf{x}_i)\mathbf{x}_i^Tr_{n,i}=0. \tag{3}$$

用矩阵形式描述,有

$$\mathbf{X}_n^T\mathbf{A}_n\mathbf{W}_n\mathbf{Y}_n=\mathbf{X}_n^T\mathbf{A}_n\mathbf{W}_n\mathbf{X}_n\boldsymbol{\beta}_n. \tag{4}$$

式(4)中: $\mathbf{A}_n=\text{diag}(\lambda^{n-1},\lambda^{n-2},\cdots,\lambda^{n-i},\cdots,1)$ ;  $\mathbf{W}_n=\text{diag}(\omega_1,\omega_2,\cdots,\omega_i,\cdots,\omega_n)$ ,  $\omega_i=w(\frac{r_{n,i}}{\sigma d(\mathbf{x}_i)})\xi(\mathbf{x}_i)$ .

定义

$$\mathbf{P}_n=(\mathbf{X}_n^T\mathbf{A}_n\mathbf{W}_n\mathbf{X}_n)^{-1}, \tag{5}$$

由式(4)可得

$$\boldsymbol{\beta}_n=\mathbf{P}_n\mathbf{X}_n^T\mathbf{A}_n\mathbf{W}_n\mathbf{Y}_n. \tag{6}$$

式(5),(6)是新算法的基础公式.

为确定  $d(\mathbf{x}_i)$  和  $\xi(\mathbf{x}_i)$  的具体形式,定义 3 个矩阵,有

$$\begin{aligned}\widetilde{\mathbf{H}}_n&\triangleq\mathbf{X}_n\mathbf{P}_n\mathbf{X}_n^T, \\ \widetilde{\mathbf{H}\mathbf{W}}_n&\triangleq\widetilde{\mathbf{H}}_n\mathbf{A}_n\mathbf{W}_n, \\ \mathbf{H}\mathbf{W}_n&\triangleq(\mathbf{W}_n)^{1/2}\mathbf{A}_n^{1/2}\widetilde{\mathbf{H}}_n\mathbf{A}_n^{1/2}(\mathbf{W}_n)^{1/2}.\end{aligned}$$

上式中: $(\mathbf{W}_n)^{1/2}=\text{diag}(\sqrt{\omega_1},\sqrt{\omega_2},\cdots,\sqrt{\omega_n})$ ;  $(\mathbf{A})^{1/2}=\text{diag}(\sqrt{\lambda^{n-1}},\sqrt{\lambda^{n-2}},\cdots,1)$ . 一般地,这 3 个矩阵具有如下 4 个性质:1)  $\widetilde{\mathbf{H}}_n$  是对称,但非幂等的;2)  $\widetilde{\mathbf{H}\mathbf{W}}_n$  是幂等,但非对称的;3)  $\mathbf{H}\mathbf{W}_n$  是幂等对称的,且其对角元素  $h\omega_i$  的范围为  $0\leqslant h\omega_i\leqslant 1$ ;4)  $\widetilde{\mathbf{H}\mathbf{W}}_n$  的对角元素  $\widetilde{h\omega_i}$  的范围为  $0\leqslant \widetilde{h\omega_i}\leqslant 1$ .

**引理<sup>[11]</sup>** 设  $\mathbf{X}_n^T\mathbf{X}_n$  可逆,则矩阵  $\mathbf{H}_n=\mathbf{X}_n(\mathbf{X}_n^T\mathbf{X}_n)^{-1}\mathbf{X}_n^T$  为幂等对称矩阵,且其对角元素  $h_i$  的范围为  $0\leqslant h_i\leqslant 1$ . 矩阵  $\mathbf{H}_n$  的对角元素通常被用来检测杠杆点.

证明 性质 1),2)的结论很容易得到,所以只需证明性质 3)即可.

假设  $\overline{\mathbf{W}}_n$  是  $\mathbf{W}_n$  中非零的相应部分, $\overline{\mathbf{A}}_n,\overline{\mathbf{X}}_n$  分别是  $\mathbf{A}_n,\mathbf{X}_n$  中相应于  $\overline{\mathbf{W}}_n$  的部分,且类似于普通最小二乘法中  $\mathbf{X}_n^T\mathbf{X}_n$  是可逆的假设. 新估计器中假定  $\overline{\mathbf{X}}_n^T\overline{\mathbf{X}}_n$  是可逆的,式(5)可简化为  $\mathbf{W}\mathbf{P}_n=(\overline{\mathbf{X}}_n^T\overline{\mathbf{A}}_n\overline{\mathbf{W}}_n\overline{\mathbf{X}}_n)^{-1}$ . 而  $\overline{\mathbf{A}}_n^{1/2},\overline{\mathbf{W}}_n^{1/2}$  的乘积是非奇异的,令  $\mathbf{D}=(\overline{\mathbf{W}}_n)^{1/2}\overline{\mathbf{A}}_n^{1/2}\mathbf{X}_n$ ,在  $\mathbf{H}\mathbf{W}_n$  中相应于  $\overline{\mathbf{W}}_n$  的部分  $\overline{\mathbf{H}\mathbf{W}}_n$  可以表示为

$$\overline{\mathbf{H}\mathbf{W}}_n=(\overline{\mathbf{W}}_n)^{1/2}\overline{\mathbf{A}}_n^{1/2}(\overline{\mathbf{X}}_n^T\overline{\mathbf{A}}_n\overline{\mathbf{W}}_n\overline{\mathbf{X}}_n)^{-1}\overline{\mathbf{X}}_n^T\overline{\mathbf{A}}_n^{1/2}(\overline{\mathbf{W}}_n)^{1/2}=\mathbf{D}_n(\mathbf{D}_n^T\mathbf{D}_n)^{-1}\mathbf{D}_n^T.$$

根据性质 1), $\mathbf{H}\mathbf{W}_n$  是幂等对称的,且其对角元素  $\overline{h\omega_j}$  的范围为  $0\leqslant \overline{h\omega_j}\leqslant 1$ . 又因为  $\mathbf{H}\mathbf{W}_n$  中相应于权重为 0 的对角元素也是 0,所以对于  $\mathbf{H}\mathbf{W}_n$  的所有对角元素,有  $0\leqslant h\omega_j\leqslant 1$ .

下面分析基于式(5),(6)的残差特性.

**定理** 残差  $r_{n,i}$  的方差  $(\sigma_{\lambda}^{(i)})^2$  与  $\widetilde{h}\omega_i, \sigma^2$  之间满足  $(\sigma_{\lambda}^{(i)})^2 \leq (1 - \widetilde{h}\omega_i)\sigma^2$ .

**证明** 根据式(6),残差向量的形式有

$$\mathbf{r}_n = \mathbf{Y}_n - \mathbf{X}_n \mathbf{P}_n \mathbf{X}_n^T \mathbf{A}_n = (\mathbf{I}_n - \mathbf{X}_n \mathbf{P}_n \mathbf{X}_n^T \mathbf{A}_n \mathbf{W}_n) \mathbf{Y}_n = (\mathbf{I}_n - \widetilde{\mathbf{H}} \mathbf{W}_n) \mathbf{Y}_n. \quad (7)$$

式(7)中:  $\mathbf{I}_n$  为  $n$  阶单位矩阵.

在观测噪声为独立同分布的假设之下,残差向量  $\mathbf{r}_n$  的协方差矩阵为

$$\text{cov}(\mathbf{r}_n) = (\mathbf{I}_n - \widetilde{\mathbf{H}} \mathbf{W}_n) \sigma^2 (\mathbf{I}_n - (\widetilde{\mathbf{H}} \mathbf{W}_n)^T) = (\mathbf{I}_n - \widetilde{\mathbf{H}} \mathbf{W}_n - (\widetilde{\mathbf{H}} \mathbf{W}_n)^T + \widetilde{\mathbf{H}} \mathbf{W}_n (\widetilde{\mathbf{H}} \mathbf{W}_n)^T) \sigma^2. \quad (8)$$

于是残差  $r_{n,i}$  的方差为

$$(\sigma_{\lambda}^{(i)})^2 = (1 - 2\widetilde{h}_i \lambda^{n-i} \omega_i + \sum_{j=1}^n (\lambda^{n-j})^2 \widetilde{h}_{i,j}^2 \omega_j^2) \sigma^2. \quad (9)$$

式(9)中:  $\widetilde{h}_{i,j}$  是  $\widetilde{\mathbf{H}}_n$  的第  $i$  行第  $j$  列的元素;  $\widetilde{h}_i$  表示其对角线上的相应元素. 注意到  $(\lambda^{n-j}) \leq \lambda^{n-j}, j=1, 2, \dots, n$ , 且  $0 \leq \omega_j \leq 1$ , 易得

$$(\sigma_{\lambda}^{(i)})^2 \leq (1 - 2\widetilde{h}_i \lambda^{n-i} \omega_i + \sum_{\substack{j=1 \\ i \neq j}}^n \lambda^{n-j} \widetilde{h}_{i,j}^2 \omega_j + (\lambda^{n-i})^2 \widetilde{h}_i^2 \omega_i^2) \sigma^2. \quad (10)$$

又根据  $\widetilde{\mathbf{H}} \mathbf{W}_n$  的幂等性, 可知其对角元素与其他元素的关系, 有

$$\widetilde{h}\omega_i = \widetilde{h}_i \lambda^{n-i} \omega_i = \sum_{\substack{j=1 \\ i \neq j}}^n \lambda^{n-j} \widetilde{h}_{i,j}^2 \omega_j + (\lambda^{n-i})^2 \widetilde{h}_i^2 \omega_i^2. \quad (11)$$

因而, 把式(11)代入式(10), 可得

$$(\sigma_{\lambda}^{(i)})^2 \leq (1 - \widetilde{h}\omega_i) \sigma^2 = (1 - \widetilde{h}\omega_i) \sigma^2. \quad (12)$$

由性质4), 又因为式(12)右边是非负的, 考虑到在权重矩阵和遗忘因子确定的情况下,  $\widetilde{h}\omega_i$  是关于  $\mathbf{x}_i$  的函数, 取  $d\mathbf{x}_i = \sqrt{1 - \widetilde{h}\omega_i}$ ,  $0 \leq \widetilde{h}\omega_i < 1$ , 并定义一个统计量  $\tilde{t}_i = r_{i,n} / (\hat{\sigma}_n \sqrt{1 - \widetilde{h}\omega_i})$ . 考虑到  $\mathbf{H}_n$  的对角元素  $h_i$  是  $\widetilde{h}\omega_i$  的特殊形式, 且对普通的最小二乘法而言,  $h_i$  有检测回归变量中异常点的功能. 为了避免  $\widetilde{h}\omega_i$  的值接近 1 时, 造成计算上的误差, 定义  $\xi(\cdot)$  函数为

$$\xi(\mathbf{x}_i) = \begin{cases} 1, & \widetilde{h}\omega_i < K, \\ 0, & \text{其他.} \end{cases}$$

上式中:  $K$  为可调的参数. 该参数确定数据  $\mathbf{x}_i$  是否为异常点. 如果  $\widetilde{h}\omega_i < K$ , 认为  $\mathbf{x}_i$  是正常的; 否则, 认为  $\mathbf{x}_i$  是异常的. 为方便, 将  $\xi(\mathbf{x}_i), d(\mathbf{x}_i)$  分别记作  $\tilde{\xi}(\widetilde{h}\omega_i), \tilde{d}(\widetilde{h}\omega_i)$ .

新递推估计器采用一步迭代法推导. 假定在  $n-1$  时刻的权矩阵  $\mathbf{W}_{n-1}$  和相关的估计  $\hat{\mathbf{p}}_{n-1}, \mathbf{P}_{n-1}$  已经得到, 那么相应的式(5), (6)可以表示为

$$\mathbf{P}_{n-1} = (\mathbf{X}_{n-1}^T \mathbf{A}_{n-1} \mathbf{W}_{n-1} \mathbf{X}_{n-1})^{-1}, \quad (13)$$

$$\hat{\mathbf{p}}_{n-1} = \mathbf{P}_{n-1} \mathbf{X}_{n-1}^T \mathbf{A}_{n-1} \mathbf{W}_{n-1} \mathbf{Y}_{n-1}. \quad (14)$$

此时, 初始权重矩阵取  $\mathbf{W}_n^{(0)} = \text{diag}(\text{diag}(\mathbf{W}_{n-1}), 1)$ , 而  $\hat{\mathbf{p}}_n, \mathbf{P}_n$  的初始估计为

$$\mathbf{P}_n^{(0)} = (\mathbf{X}_n^T \mathbf{A}_n \mathbf{W}_n^{(0)} \mathbf{X}_n)^{-1} = (\lambda \mathbf{X}_{n-1}^T \mathbf{A}_{n-1} \mathbf{W}_{n-1} \mathbf{X}_{n-1} + \mathbf{x}_n^T \mathbf{x}_n)^{-1} = (\lambda \mathbf{P}_{n-1}^{-1} + \mathbf{x}_n^T \mathbf{x}_n)^{-1}, \quad (15)$$

$$\hat{\mathbf{p}}_n^{(0)} = \mathbf{P}_n^{(0)} \mathbf{X}_n^T \mathbf{A}_n \mathbf{W}_n^{(0)} \mathbf{Y}_n = \mathbf{P}_n^{(0)} (\lambda \mathbf{X}_{n-1}^T \mathbf{A}_{n-1} \mathbf{W}_{n-1} \mathbf{Y}_{n-1} + \mathbf{x}_n^T \mathbf{y}_n) = \mathbf{P}_n^{(0)} (\lambda \mathbf{P}_{n-1}^{-1} \hat{\mathbf{p}}_{n-1} + \mathbf{x}_n^T \mathbf{y}_n) = \quad (16)$$

$$\mathbf{P}_n^{(0)} ((\mathbf{P}_{n-1}^{(-)} - \mathbf{x}_n^T \mathbf{x}_n) \hat{\mathbf{p}}_{n-1} + \mathbf{x}_n^T \mathbf{y}_n) = \hat{\mathbf{p}}_{n-1} + \mathbf{P}_n^{(0)} \mathbf{x}_n^T (\mathbf{y}_n - \mathbf{x}_n \hat{\mathbf{p}}_{n-1}).$$

式(15), (16)中:  $\mathbf{P}_n^{(-)}$  为  $\mathbf{P}_n^{(0)}$  的逆矩阵.

令  $\mathbf{A} = \lambda \mathbf{P}_{n-1}^{(-)}, \mathbf{C} = \mathbf{I}_p, \mathbf{B} = \mathbf{D}^T = \mathbf{x}_n^T$ , 并将矩阵的逆的公式  $(\mathbf{A} + \mathbf{BCD})^{-1} = \mathbf{A}^{-1} - \mathbf{A}^{-1} \mathbf{B} (\mathbf{DA}^{-1} \mathbf{B} + \mathbf{C}^{-1})^{-1} \mathbf{DA}^{-1}$  应用于式(15), 则  $\hat{\mathbf{p}}_n^{(0)}$  的递推计算公式变为

$$\mathbf{P}_n^{(0)} = \frac{1}{\lambda} \mathbf{P}_{n-1} (\mathbf{I}_p - \frac{\mathbf{x}_n^T \mathbf{x}_n \mathbf{P}_{n-1}}{\lambda + \mathbf{x}_n^T \mathbf{P}_{n-1} \mathbf{x}_n}). \quad (17)$$

迭代开始后, 需要计算再加权矩阵  $\widetilde{\mathbf{W}}_n$ . 此时, 只对  $\omega_n$  进行更新, 保持  $\mathbf{W}_n^{(0)}$  中的相应元素不变, 即取  $\widetilde{\mathbf{W}}_n = \text{diag}(\text{diag}(\mathbf{W}_{n-1}), \omega_n)$ . 由于  $\lambda^0 = 1, \omega_n^{(0)} = 1$ , 易知  $\widetilde{h}_n^{(0)} = \widetilde{h}\omega_n^{(0)} = \mathbf{x}_n^T \mathbf{P}_n^{(0)} \mathbf{x}_n, \tilde{h}_n^{(0)} = \frac{r_{n,n}^{(0)}}{\hat{\sigma} \sqrt{1 - \tilde{h}_n^{(0)}}}$ . 因此,

$\omega_n = \omega(\tilde{t}_n^{(0)})\tilde{\xi}(\tilde{h}_n^{(0)})$ . 一个更好的做法是将  $W_n^{(0)}$  的对角元素都进行更新. 然而, 为了得到递推的计算公式, 这种做法不得不放弃. 但是, 引入的遗忘因子和所用 M-估计函数的有界性可以降低这种影响.

在推导  $P_n, \hat{\beta}_n$  的递推计算公式时, 除  $D = \omega_n x_n$  不同外, 其他做法类似于式(16), (17)的推导, 可得

$$P_n = \frac{1}{\lambda} P_{n-1} (I_p - \frac{x_n^T \omega_n x_n P_{n-1}}{\lambda + \omega_n x_n P_{n-1} x_n^T}), \tag{18}$$

$$\hat{\beta}_n = P_n X_n^T \Lambda_n \tilde{W}_n Y_n = \hat{\beta}_{n-1} + P_n \omega_n x_n^T (y_n - x_n \hat{\beta}_{n-1}). \tag{19}$$

式(16)~(19)形成了递推有界广义极大似然类(RBGM)回归估计器. 为了降低计算量, 新估计器还可以描述为

输入:  $\hat{\beta}_{n-1}, P_{n-1}, (x_n, y_n)$ ;

输出:  $\hat{\beta}_n, P_n$ ;

- 1) 令  $P_\lambda = \frac{1}{\lambda} P_{n-1}, q_n = x_n P_{n-1} x_n^T, Q_n = x_n^T x_n P_{n-1}, X_y = x_n^T (y_n - x_n \hat{\beta}_{n-1})$ ;
- 2)  $P_n^{(0)} = P_\lambda (I_p - \frac{Q_n}{\lambda + q_n}), \beta_n^{(0)} = \hat{\beta}_{n-1} + P_n^{(0)} X_y$ ;
- 3)  $\omega_n = \omega(\tilde{t}_n^{(0)})\tilde{\xi}(\tilde{h}_n^{(0)})$ ;
- 4)  $P_n = P_\lambda (I_p - \frac{\omega_n Q_n}{\lambda + \omega_n q_n}), \hat{\beta}_n = \hat{\beta}_{n-1} + \omega_n P_n X_y$ .

### 3 AR 模型参数估计算法

在观测数据中含有较大量加性异常点的情况下进行 AR 模型参数估计时, RBGM 也是有偏的. 为此, 需要对其进行必要的改造, 把面向 AR 模型参数估计的算法记作 AR-RBGM.

首先, 用  $\tilde{x}_n, \tilde{\beta}_n$  替换 RBGM 中的  $x_n, \beta_n, \tilde{\beta}_n = [b_1, b_2, \cdots, b_p, \eta_1], \tilde{x}_n = [y_{n-1}, y_{n-2}, \cdots, y_{n-p}, \epsilon_{n-1}]$  ( $\eta_1$  为增广参数,  $\epsilon_{n-1}$  为增广变量). 在初始化时,  $\epsilon_{n-1}$  和  $\eta_1$  都赋 0 值. 在运行期间,  $\eta_1$  随着算法自动更新.

其次, 增加一个加性异常点的检测过程. 在此过程中涉及两种残差的计算, 一是不包括增广项的  $r_{n,n}$ , 二是包括增广项的残差  $\tilde{r}_{n,n} = y_n - \tilde{x}_n \tilde{\beta}_n$ . 在当前  $n$  时刻的估计已经得到后, 假设以前的加性异常点的影响已经体现在增广项里, 且当前的观测数据不是加性异常点, 那么相应的接受域为  $|\tilde{r}_{n,n}| \leq \gamma \hat{\sigma}_n$ , 而相应的拒绝域为  $|\tilde{r}_{n,n}| > \gamma \hat{\sigma}_n$ . 其中:  $\gamma$  是个可调常数, 一般取 1.5 左右即可. 据检验结果可知: 如果拒绝了假设, 则认为  $y_n$  是加性异常点, 其影响应该体现在增广项中, 相应地取  $\epsilon_n = r_{n,n}$ ; 否则, 取  $\epsilon_n = 0$ . 此外,  $\hat{\sigma}_n$  是用  $\tilde{r}_{n,i}$  替换  $r_{n,i}$  后, 通过式(20)计算得到的, 而在用 RLS 进行启动的阶段, 可直接用 med 估计, 也可以借用先验估计结果.

在 AR 模型参数估计时, 无论数据中包含或不包含异常点, 异常点是革新或加性, AR-RBGM 中的参数都无须重置. 另外, RBGM 和 AR-RBGM 追踪系统突变的能力主要依赖于遗忘因子的大小, 当需要追踪突变时, 遗忘因子相对取小一点, 如 0.99 等; 不需要追踪突变时, 遗忘因子则取得相对大一点, 如 0.999 或 0.999 9 等. 在估计精度和对突变的追踪能力方面需要折中取舍.

### 4 仿真实验

在实验中, 假定需要辨识的真实系统为

$$z_i - 1.2z_{i-1} + 0.52z_{i-2} = e_i. \tag{20}$$

式(20)中:  $e_i$  为相互独立同分布的随机变量, 且  $e_i \sim N(0, \sigma_1^2)$ . 该模型已经在文献[5, 10]中被用来检验 M 和 GM 估计器的鲁棒性.

在观测值的仿真实验中, 如果第一型异常点出现, 相应的观测值  $y_i$  用  $y_i = z_i + v_i$  进行模拟. 其中:  $v_i \sim (1 - \kappa_1) \times "0" + \kappa_1 \times N(0, \sigma_2^2)$ , 如果第二型异常点出现, 相应的观测值  $y_i$  通过把式(21)中的  $e_i$  替换为另一个随机变量  $n_i$ , 并令  $y_i = z_i$  进行模拟. 其中:  $n_i \sim N(0, \sigma_1^2) + \kappa_2 \times N(0, \sigma_3^2)$ . 在仿真实验中, 设计了 4 种更加复杂的情形, 观测数据如表 1 所示. 每一种情形都将整个过程分成 3 个阶段, 并且假定第一型异常点或者第二型异常点分别出现在某一个阶段.

在所有情形中,取  $\sigma_1^2=25, \kappa_1=0.05, \kappa_2=0.05, \sigma_2^2=\sigma_3^2=400$ . 采用对数化平均相对误差 (LMRE) 来刻画估计精度及收敛性,即

$$\text{LMRE}(n)=20\log_{10}((\sum_{m=1}^M\|\hat{\boldsymbol{\beta}}_m^{(n)}-\boldsymbol{\beta}\|_2/\|\boldsymbol{\beta}\|_2)/M).$$

(21)

式(21)中: $M$  为程序运行的次数; $\boldsymbol{\beta}$  为系统的真实参数; $\hat{\boldsymbol{\beta}}_m^{(n)}$  为在第  $m$  次运行时, $n$  时刻  $\boldsymbol{\beta}$  的估计.

实验中用经典的 RLS 和另外两种递推的 GM 估计器 RGM<sup>[8]</sup> 和 RKW<sup>[4]</sup> 做了比较. 在没有异常点的情况下这几种估计器的表现,如图 1 所示. 图 1 中:LMRE 为对数化平均相对误差; $n$  为时序. 在比较时,每种估计器都进行了参数调整以使其表现尽可能达到最好. 具体地,对 RLS, $\lambda=0.9999$ ;对 RGM, $\lambda=0.9999, c=2.8, P(0)=100I_2$ ;对 RKW, $P(0)=A^{-1}(0)=100I_2, \lambda=0.9999, c=2, a=5$ ;对 RBGM 和 AR-RBGM, $\lambda=0.9999, N_{\text{Initial}}=10, \lambda_{\sigma}=0.98, \gamma=1.5, L=20, K=0.85, M_1=1.88, M_2=2.41$ . 为了方便比较,所有的初始尺度估计中, $\sigma_0=5$ . 实验结果表明:在此情形下,除了 AR-RBGM 的表现稍差以外,其他估计器表现相当. 因为 AR-RBGM 中的增广项改变了原始模型的结构,对其性能产生了有限的影响.

非平稳环境(情形 2~4)下估计性能比较,如图 2~4 所示. 图 2~4 中:LMRE 为对数化平均相对误差; $n$  为时序. 由图 2,3 可知:即使当加性异常点出现在算法已达到或者接近稳定状态时,RLS, RGM, RKW 完全失效,而 RBGM 对这些影响抑制效果是明显的,但精度仍然不高;AR-RBGM 能在各种情形下都保持较高的精度和收敛性,其精度与没有异常点时的精度差别很小. 由图 4 可知:对加性异常点出现在初始阶段的情形,仍然是 AR-RBGM 保持较高精度和良好收敛性,其他估计器全部失;加性异常点出现在起始阶段时,AR-RBGM 的相应的指标会变大一点,这主要是受到非鲁棒启动算法 RLS 的影响.

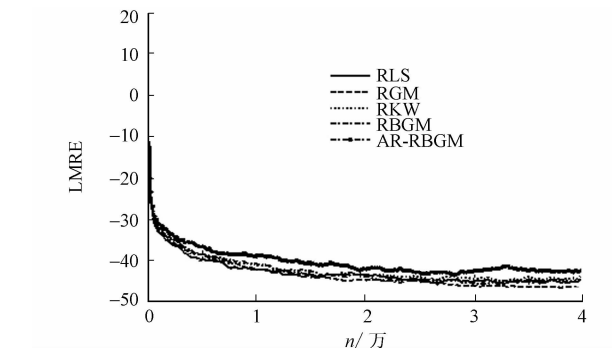


图 1 平稳环境(情形 1)下估计性能比较

Fig. 1 Performance comparison of the algorithm for the stability environment (case 1)

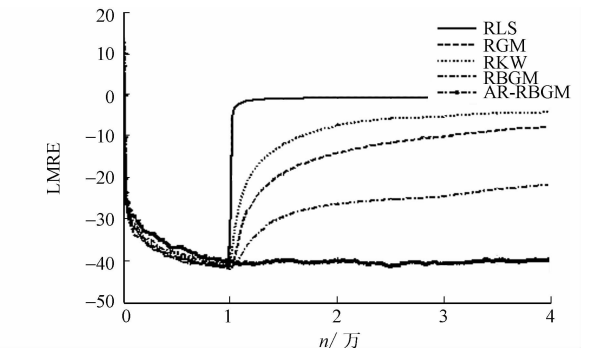


图 2 非平稳环境(情形 2)下估计性能比较

Fig. 2 Performance comparison of the algorithm for the instability environment (case 2)

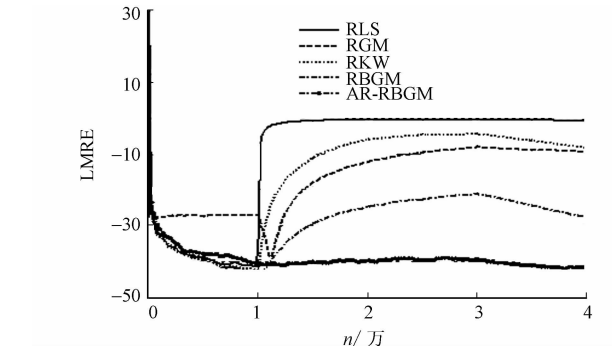


图 3 非平稳环境(情形 3)下估计性能比较

Fig. 3 Performance comparison of the algorithm for the instability environment (case 3)

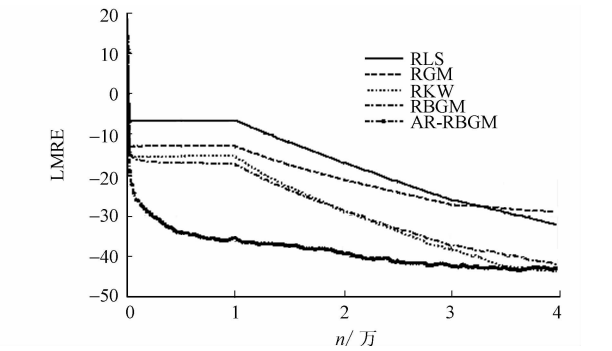


图 4 非平稳环境(情形 4)下估计性能比较

Fig. 4 Performance comparison of the algorithm for the instability environment (case 4)

5 结 束 语

改进 GM 估计器的一般框架,提出一种新的递推 GM 回归估计器(RBGM),并针对 AR 模型参数估计提出了 AR-RBGM. RBGM 和 AR-RBGM 均能对回归变量中的加性异常点的影响起到抑制作用,特别是 AR-RBGM 能在非平稳环境下实现自适应的估计,并保持良好的精度和收敛性. 新估计器还可以进行改进,一方面,鲁棒的启动算法可以提高性能;另一方面,可进一步推广到 ARMA 模型参数的估计中,从而获得相应的鲁棒算法. 然而,由于 AR-RBGM 引入了增广变量,增加了额外的计算量.

参考文献:

[1] HUBER P J. Robust regression: Asymptotics, conjectures and monte carlo[J]. Annals of Statistics, 1973, 1(5): 799-821.

[2] CAMPBELL K. Recursive computation of M-estimates for the parameters of a finite autoregressive process[J]. The Annals of Stat,1982,10(2):442-453.

[3] ANTOCH J, EKBLÖM H. Recursive robust regression computational aspects and comparison[J]. Computational Statistics and Data Analysis,1995,19(2):115-128.

[4] SEJLING K, et al. Methods for recursive robust estimation of AR parameters[J]. Computational Statistics and Data Analysis,1994,17(5):509-536.

[5] PHAM D S, ZOUBIR A M. A sequential algorithm for robust parameter estimation[J]. IEEE Signal Processing Lett,2005,12(1):21-24.

[6] VEGA L R, REY H, BENESTY J, et al. A robust recursive least squares algorithm[J]. IEEE Trans Signal Process, 2009,57(3):1209-1216.

[7] KRASKER W S, WELSCH R E. Efficient bounded-influence regression estimation[J]. Journal of the American Statistical Association,1982,77(379):595-604.

[8] GRILLENZONI C. Recursive generalized M-estimators of system parameters[J]. Technometrics,1997,39(2):211-224.

[9] ENGIUND J E. Recursive versions of the algorithm by Krasker and Welsch[J]. Sequential Analysis,1991,10(3/4): 211-234.

[10] MARONNA R A, MARTIN R D, YOHAI V J. Robust statistics: Theory and methods[M]. West Sussex: John Wiley & Sons,2006:888-889.

[11] ROUSSEEUW P J, LEROY A M. Robust regression and outlier detection[M]. New York: Wiley,1987:340-347.

A New Recursive Bounded GM Estimator for Regression

CHENG Li-hua<sup>1</sup>, ZHANG Jun-min<sup>2</sup>

(1. College of Science, Xi'an Polytechnic University, Xi'an 710048, China;  
2. College of Science, Xi'an University of Architecture and Technology, Xi'an 710055, China)

**Abstract:** A new recursive bounded GM estimator for regression is proposed. Unlike other GM estimators, the new estimator is based on one more general framework and uses a cost function with bounded M-estimate function. The new estimator, in effect, is a recursive one-step iteration solution of the "normal equations" corresponding to the cost function. In the new estimator, a weight function is designed to reject or to reduce the influence of the outliers. Furthermore, by introducing an augment variable, the proposed estimator is modified to a very adaptive version for the estimation of autoregressive parameters. The simulation results show that both the proposed estimator and its modification are more effective than other related estimators in suppressing the adverse influence of outliers; the proposed estimator, with the same settings, can keep a high accuracy and stable convergence performance in a variety of non-stationary environments.

**Keywords:** generalized maximum likelihood type estimator; robust estimation; autoregressive parameters; additive outliers