带有 ARIMA(0,1,0)误差的非线性模型影响诊断

解锋昌[1],孙越泓[2],刘应安[3]

(1. 南京农业大学数学系,210095,江苏,南京)
(2. 南京师范大学数学与计算机科学学院,210097,江苏,南京)
(3. 南京林业大学信息学院,210037,江苏,南京)

[摘要] 在带有 ARIMA(0,1,0)误差的非线性回归模型中,利用梯度方向研究了观测数据对于其异方差检验的 Score 统计量的局部影响,分别得到了 因变量及自变量的微小扰动情况下度量最大局部影响的诊断统计量. 最后,给出了具体的数值实例,说明了本文结论的有效性.

[关键词] Score 统计量,梯度,局部影响,非线性回归模型,ARIMA(0,1,0)误差
[中图分类号]O212.1, [文献标识码]A, [文章编号]1001-4616(2005)04-0034-04

Influence Diagnosis for the Nonlinear Models with ARIMA(0,1,0) Errors

Xie Fengchang[1], Sun Yuehong[2], Liu Yingan[3]

(1. Department of Mathematics, Nanjing Agricultural University, 210095, Nanjing, China)
(2. School of Mathematics and Computer Science, Nanjing Normal University, 210097, Nanjing, China)
(3. College of Information Science and Technology, Nanjing Forestry University, 210037, Nanjing, China)

Abstract: In nonlinear regression models with ARIMA(0,1,0) errors, local influence of the observations on the Score statistic for heteroscedastic test is discussed based on gradient direction, and diagnostic statistics for measuring maximum local influence under perturbation in response to attributive variables and explanatory variables are obtained, respectively. In the end, a numerical example is given to illustrate our results.

Key words: Score statistic, gradient, local influence, nonlinear regression model, ARIMA(0,1,0) errors

0 引言

自回归单整滑动平均模型(ARIMA)是一种较为复杂的时间序列模型,在数量经济中有广泛的应用,相关研究已有不少,如[1—4]等. Bruce and Martin (1989)成功地将回归分析中单点剔除诊断方法应用到结构相当复杂的 ARIMA 模型. Tsai (1986),韦博成,胡跃清 (1994)分别研究了具有特殊的 ARIMA 误差的线性和非线性模型——具有 AR(1)误差的回归模型的异方差性和一阶自相关性检验问题,但是对误差为一般的 ARIMA 序列的回归模型的异方差性的诊断还比较困难. 文[4]讨论如下具有一类特殊的 ARIMA 误差的非线性模型——伴有 ARIMA(0,1,0)误差(或随机游动误差)的非线性回归模型的异方差的检验问题:

$$\begin{cases} y_i = f(x_i,\boldsymbol{\beta}) + \varepsilon_i, \\ \varepsilon_i = \sum_{t=1}^{i} e_t, \end{cases} \tag{1}$$

其中,$i = 1,2,\cdots,n$,x_i 是已知的协变量,$\boldsymbol{\beta}$ 是 p 维回归系数;f 为关于 $\boldsymbol{\beta}$ 的两阶可微函数;$e_1,e_2,\cdots,e_n$ 为白噪声序列,且 $e_i \sim N(0,\sigma_i^2)$. 假设,$\sigma_i^2 = \sigma^2 m_i = \sigma^2 m(z_i,\boldsymbol{\lambda})$,$z_i$ 是协变量,$\boldsymbol{\lambda}$ 为 $q(p+q+1<n)$ 维未知向

收稿日期:2005-05-28.
基金项目:国家自然科学基金资助项目(10371016)和南京农业大学青年科技创新基金资助项目(KJ04020).
作者简介:解锋昌,1969—,讲师,主要从事统计诊断和 EM 算法的教学与研究. E-mail: fcxie@ njau. edu. cn

量,存在 $\boldsymbol{\lambda}_0$,使得 $m(z_i,\boldsymbol{\lambda}_0) = 1$ 对所有 i 都成立,σ^2 为未知参数. 易见当 $\boldsymbol{\lambda} = \boldsymbol{\lambda}_0$ 时,白噪声序列 $\{\varepsilon_i\}$ 有共同方差. 则模型(1) 的异方差检验问题可表示为检验假设:

$$H_0: \boldsymbol{\lambda} = \boldsymbol{\lambda}_0,\ H_1: \boldsymbol{\lambda} \neq \boldsymbol{\lambda}_0. \tag{2}$$

文[4] 给出了模型(1) 在检验问题 H_0 下 Score 检验统计量:

$$SC = \{[\boldsymbol{u}^{\mathrm{T}} \bar{\dot{\boldsymbol{m}}}(\bar{\dot{\boldsymbol{m}}}^{\mathrm{T}} \bar{\dot{\boldsymbol{m}}})^{-1} \bar{\dot{\boldsymbol{m}}}^{\mathrm{T}}\boldsymbol{u}]/(2\sigma^4)\}_{\hat{\theta}}. \tag{3}$$

式中 $\hat{\boldsymbol{\theta}} = (\boldsymbol{\lambda}_0^{\mathrm{T}}, \hat{\sigma}^2, \hat{\boldsymbol{\beta}}^{\mathrm{T}})^{\mathrm{T}}$ 表示参数 $\boldsymbol{\theta}$ 在 H_0 成立时的极大似然估计;$\boldsymbol{u} = (e_1^2,e_2^2\cdots,e_n^2)^{\mathrm{T}}$, $\bar{\dot{\boldsymbol{m}}} = (\boldsymbol{I}_n - \boldsymbol{1}_n\boldsymbol{1}_n^{\mathrm{T}}/n)\dot{\boldsymbol{m}}$ 为 $\dot{\boldsymbol{m}}$ 的中心化,$\boldsymbol{I}_n$ 为 n 阶单位阵,$\boldsymbol{1}_n = (1,\cdots,1)^{\mathrm{T}}$,$\dot{\boldsymbol{m}} = (\partial m(z_i,\boldsymbol{\lambda})/\partial\lambda_j)_{n\times q}$.

1 局部影响分析

局部影响主要研究模型有微小扰动时,对于统计量的影响,由此导出某些特定的诊断统计量. Cook 于 1986 年从微分几何观点出发提出度量局部影响的基本方法[5],即假设 $\boldsymbol{\omega}$ 为刻画扰动的向量,$\boldsymbol{\omega}_0$ 对应于无扰动情形,由于受扰动以后的似然距离 $LD(\boldsymbol{\omega})$ 在 $\boldsymbol{\omega}_0$ 处的函数值与一阶导数值均为零,因此自然地用二阶导数即曲率来刻画扰动的影响. 但是,对于某些其它的统计量,它们在 $\boldsymbol{\omega}_0$ 处的一阶导数不为零,则可用方向导数达到最大的方向 $d_{\max}$,即梯度方向来刻画影响,如文献[6—8]. 在这些研究的基础上,文章用梯度方向来刻画扰动对于带有 ARIMA(0,1,0) 误差的非线性回归模型的异方差检验的 Score 统计量的局部影响,得到了度量影响的诊断统计量.

便于以下讨论,首先给出模型(1) 中参数 $\boldsymbol{\theta}$ 相应的对数似然函数:

$$l(\boldsymbol{\theta}) = C - \frac{n}{2}\log\sigma^2 - \frac{1}{2}\sum_{i=1}^{n}\log m_i - \frac{(\boldsymbol{y} - f(\boldsymbol{X},\boldsymbol{\beta}))^{\mathrm{T}}\boldsymbol{A}^{-T}\boldsymbol{M}^{-1}\boldsymbol{A}^{-1}(\boldsymbol{y} - f(\boldsymbol{X},\boldsymbol{\beta}))}{2\sigma^2}, \tag{4}$$

其中,C 为常数,$\boldsymbol{A}^{-\mathrm{T}} = (\boldsymbol{A}^{-1})^{\mathrm{T}}$. 为方便计,后文记 $\boldsymbol{e} = \boldsymbol{A}^{-1}(\boldsymbol{y} - f(\boldsymbol{X},\boldsymbol{\beta}))$. $\boldsymbol{\theta} = (\boldsymbol{\lambda}^{\mathrm{T}},\boldsymbol{\beta}^{\mathrm{T}},\sigma^2)^{\mathrm{T}}$,$\boldsymbol{y} = (y_1,y_2,\cdots,y_n)^{\mathrm{T}}$, $f(\boldsymbol{X},\boldsymbol{\beta}) = (f(x_1,\boldsymbol{\beta}), f(x_2,\boldsymbol{\beta}),\cdots,f(x_n,\boldsymbol{\beta}))^{\mathrm{T}} = (f_1(\boldsymbol{\beta}),\cdots,f_n(\boldsymbol{\beta}))^{\mathrm{T}} = f(\boldsymbol{\beta})$,$\boldsymbol{M} = \mathrm{diag}(m_1,m_2,\cdots,m_n)$ 和 $\boldsymbol{A} = \begin{bmatrix} 1 & 0 & 0 & \cdots & 0 \\ 1 & 1 & 0 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 1 & 1 & 1 & \cdots & 1 \end{bmatrix}$.

设在扰动 $\boldsymbol{\omega}$ 下上面的 Score 统计量 SC 表示为 $SC(\boldsymbol{\omega})$,$SC(\boldsymbol{\omega}_0) = SC$ 表示模型无扰动情形. 根据以上分析,则扰动 $\boldsymbol{\omega}$ 对于 $SC(\boldsymbol{\omega})$ 局部影响最大的方向应满足:

$$\boldsymbol{d}_{\max} \propto \{\partial SC(\boldsymbol{\omega})/\partial\boldsymbol{\omega}\}_{(\omega_0,\hat{\theta})}. \tag{5}$$

本文分别考虑:(a) 因变量的扰动,即 $\boldsymbol{y} \to \boldsymbol{y} + \boldsymbol{\omega}$,$\boldsymbol{\omega}_0 = \boldsymbol{0}$. (b) 自变量的扰动,即 $f(\boldsymbol{\beta}) \to f(\boldsymbol{\beta} \mid \boldsymbol{\omega})$.

设模型有扰动 $\boldsymbol{\omega}$ 时,(4) 式记为 $l(\boldsymbol{\theta} \mid \boldsymbol{\omega})$,相应的极大似然估计为 $\hat{\boldsymbol{\theta}}(\boldsymbol{\omega}) = (\boldsymbol{\lambda}_0^{\mathrm{T}},\sigma^2(\boldsymbol{\omega}),\hat{\boldsymbol{\beta}}^{\mathrm{T}}(\boldsymbol{\omega}))^{\mathrm{T}}$. 当 $\boldsymbol{\omega} = \boldsymbol{\omega}_0$ 时,则以上统计量还原到原来无扰动的情形. 以下讨论均在 $(\boldsymbol{\omega}_0,\hat{\boldsymbol{\theta}})$ 计值,大多将其略去. 为得出 $SC(\boldsymbol{\omega})$ 的局部影响最大方向,先给出两个引理.

引理 1 对于扰动(a),$\hat{\boldsymbol{\beta}}(\boldsymbol{\omega})$,$\sigma^2(\boldsymbol{\omega})$ 在 $\boldsymbol{\omega}_0$ 处关于 $\boldsymbol{\omega}$ 的导数分别为

$$\frac{\partial\hat{\boldsymbol{\beta}}(\boldsymbol{\omega})}{\partial\boldsymbol{\omega}^{\mathrm{T}}} = \boldsymbol{D}^{-1}\dot{\boldsymbol{f}}^{\mathrm{T}}\boldsymbol{A}^{-\mathrm{T}}\boldsymbol{A}^{-1};\ \frac{\partial\hat{\sigma}^2(\boldsymbol{\omega})}{\partial\boldsymbol{\omega}} = 2(\boldsymbol{I}_n - \dot{\boldsymbol{f}}\boldsymbol{D}^{-1}\dot{\boldsymbol{f}}^{\mathrm{T}}\boldsymbol{A}^{-\mathrm{T}}\boldsymbol{A}^{-1})^{\mathrm{T}}\boldsymbol{A}^{-\mathrm{T}}\hat{\boldsymbol{e}}/n, \tag{6}$$

其中,$\dot{\boldsymbol{f}} = \partial f(\boldsymbol{\beta})/\partial\boldsymbol{\beta}^{\mathrm{T}}$ 为 $n \times p$ 阶矩阵,$\ddot{\boldsymbol{f}} = \partial^2 f(\boldsymbol{\beta})/\partial\boldsymbol{\beta}\partial\boldsymbol{\beta}^{\mathrm{T}}$ 是一个 $n \times p \times p$ 阶立体阵,$\boldsymbol{D} = -\boldsymbol{e}^{\mathrm{T}}\boldsymbol{A}^{-1}[\ddot{\boldsymbol{f}}] + \dot{\boldsymbol{f}}^{\mathrm{T}}\boldsymbol{A}^{-\mathrm{T}}\boldsymbol{A}^{-1}\dot{\boldsymbol{f}}$.

证明 对于扰动(a),由 $l(\boldsymbol{\theta} \mid \boldsymbol{\omega})$ 易得在 $\boldsymbol{\lambda}_0$ 处有

$$\hat{\sigma}^2(\boldsymbol{\omega}) = (\boldsymbol{y} + \boldsymbol{\omega} - f(\hat{\boldsymbol{\beta}}(\boldsymbol{\omega})))^{\mathrm{T}}\boldsymbol{A}^{-\mathrm{T}}\boldsymbol{A}^{-1}(\boldsymbol{y} + \boldsymbol{\omega} - f(\hat{\boldsymbol{\beta}}(\boldsymbol{\omega})))/n.$$

同时有

$$\dot{\boldsymbol{f}}^{\mathrm{T}}\boldsymbol{A}^{-\mathrm{T}}\boldsymbol{A}^{-1}(\boldsymbol{y} + \boldsymbol{\omega} - f(\hat{\boldsymbol{\beta}}(\boldsymbol{\omega}))) = 0. \tag{7}$$

现在对(7) 式两边关于 $\boldsymbol{\omega}$ 求导并整理得:

$$([\dot{\boldsymbol{f}}]^{\mathrm{T}}\boldsymbol{A}^{-\mathrm{T}}\boldsymbol{A}^{-1}\dot{\boldsymbol{f}} - e^{\mathrm{T}}\boldsymbol{A}^{-1}\ddot{\boldsymbol{f}})[\partial\hat{\boldsymbol{\beta}}(\boldsymbol{\omega})/\partial\boldsymbol{\omega}^{\mathrm{T}}] = [\dot{\boldsymbol{f}}]^{\mathrm{T}}\boldsymbol{A}^{-\mathrm{T}}\boldsymbol{A}^{-1}.$$

利用前面定义的记号得

$$D[\partial\hat{\boldsymbol\beta}(\boldsymbol\omega)/\partial\boldsymbol\omega^{\mathrm T}] = [\dot{\boldsymbol f}]^{\mathrm T}\boldsymbol A^{-\mathrm T}\boldsymbol A^{-1}, \tag{8}$$

同时可得 $\hat\sigma_k^2(\omega)$ 关于 ω 的导数为

$$\frac{\partial\hat\sigma^2(\boldsymbol\omega)}{\partial\boldsymbol\omega} = \frac{2}{n}(\boldsymbol I_n - \dot{\boldsymbol f}\frac{\partial\hat{\boldsymbol\beta}(\boldsymbol\omega)}{\partial\boldsymbol\omega^{\mathrm T}})^{\mathrm T}\boldsymbol A^{-\mathrm T}\hat{\boldsymbol e}, \tag{9}$$

所以由(8,9)式即得(6)式成立.

引理2 对于扰动(b),$\hat{\boldsymbol\beta}(\boldsymbol\omega)$,$\sigma_1^2(\omega)$,$\sigma_2^2(\omega)$ 在 ω_0 处关于 ω 的导数分别为

$$\partial\hat{\boldsymbol\beta}(\boldsymbol\omega)/\partial\boldsymbol\omega^{\mathrm T} = \boldsymbol D^{-1}\boldsymbol H, \partial\hat{\boldsymbol\sigma}^2(\boldsymbol\omega)/\partial\boldsymbol\omega = -2(\dot{\boldsymbol f}\boldsymbol D^{-1}\boldsymbol H + \dot{\boldsymbol f}_\omega)^{\mathrm T}\boldsymbol A^{-\mathrm T}\hat{\boldsymbol e}/n, \tag{10}$$

其中 $\dot{\boldsymbol f}_\omega = \partial f(\hat{\boldsymbol\beta}(\boldsymbol\omega)\mid\boldsymbol\omega)/\partial\boldsymbol\omega^{\mathrm T}$ 是一个 $n\times n$ 阶矩阵,$\ddot{\boldsymbol f}_\omega = \partial^2 f(\hat{\boldsymbol\beta}(\boldsymbol\omega)\mid\boldsymbol\omega)/\partial\boldsymbol\beta\partial\boldsymbol\omega^{\mathrm T}$ 是一个 $n\times p\times n$ 阶立体阵,$\boldsymbol H = \hat{\boldsymbol e}^{\mathrm T}\boldsymbol A^{-1}\ddot{\boldsymbol f}_\omega - \dot{\boldsymbol f}^{\mathrm T}\boldsymbol A^{-\mathrm T}\boldsymbol A^{-1}\dot{\boldsymbol f}_\omega$.

证明 对于扰动(b),由 $l(\boldsymbol\theta\mid\boldsymbol\omega)$ 易得在 $\boldsymbol\lambda_0$ 处有

$$\hat\sigma^2(\boldsymbol\omega) = (y - f(\hat{\boldsymbol\beta}(\boldsymbol\omega)\mid\boldsymbol\omega))^{\mathrm T}\boldsymbol A^{-\mathrm T}\boldsymbol A^{-1}(y - f(\hat{\boldsymbol\beta}(\boldsymbol\omega)\mid\boldsymbol\omega))/n,$$

同时有

$$\dot{\boldsymbol f}^{\mathrm T}\boldsymbol A^{-\mathrm T}\boldsymbol A^{-1}(y - f(\hat{\boldsymbol\beta}(\boldsymbol\omega)\mid\boldsymbol\omega)) = 0. \tag{11}$$

现在对(11)式两边关于 ω 求导并整理得:

$$(\dot{\boldsymbol f}^{\mathrm T}\boldsymbol A^{-\mathrm T}\boldsymbol A^{-1}\dot{\boldsymbol f} - \hat{\boldsymbol e}^{\mathrm T}\boldsymbol A^{-1}\ddot{\boldsymbol f})[\partial\hat{\boldsymbol\beta}(\boldsymbol\omega)/\partial\boldsymbol\omega^{\mathrm T}] = \hat{\boldsymbol e}^{\mathrm T}\boldsymbol A^{-1}\ddot{\boldsymbol f}_\omega - \dot{\boldsymbol f}^{\mathrm T}\boldsymbol A^{-\mathrm T}\boldsymbol A^{-1}\dot{\boldsymbol f}_\omega,$$

利用前面的记号可得

$$D[\partial\hat{\boldsymbol\beta}(\boldsymbol\omega)/\partial\boldsymbol\omega^{\mathrm T}] = \boldsymbol H, \tag{12}$$

同时可得 $\hat\sigma_k^2(\omega)$ 关于 ω 的导数为

$$\frac{\partial\hat\sigma^2(\boldsymbol\omega)}{\partial\boldsymbol\omega} = -\frac{2}{n}(\dot f\frac{\partial\hat{\boldsymbol\beta}(\boldsymbol\omega)}{\partial\boldsymbol\omega^{\mathrm T}} + \dot{\boldsymbol f}_\omega)^{\mathrm T}\boldsymbol A^{-\mathrm T}\hat{\boldsymbol e}. \tag{13}$$

所以由(12,13)式即得到(10)式成立.

定理 对于模型(1),假设检验问题 H_0 的 Score 统计量在扰动(a),(b)下,其局部影响最大的方向 $\boldsymbol d_{\max}$ 应分别满足:

$$(a)\ \boldsymbol d_{\max} \propto -\frac{2}{n\sigma^2}\boldsymbol G^{\mathrm T}\boldsymbol A^{-\mathrm T}\hat{\boldsymbol e}\boldsymbol u^{\mathrm T}\boldsymbol P_{\tilde m}\boldsymbol u + 2(\boldsymbol I_n - \boldsymbol E)^{\mathrm T}\boldsymbol G^{\mathrm T}\mathrm{diag}(\hat{\boldsymbol e})\boldsymbol P_{\tilde m}\boldsymbol u, \tag{14}$$

$$(b)\ \boldsymbol d_{\max} \propto \frac{2}{n\hat\sigma^2}\boldsymbol Q^{\mathrm T}\boldsymbol A^{-\mathrm T}\hat{\boldsymbol e}\boldsymbol u^{\mathrm T}\boldsymbol P_{\tilde m}\boldsymbol u - 2(\boldsymbol I_n - \boldsymbol E)^{\mathrm T}\boldsymbol Q^{\mathrm T}\mathrm{diag}(\hat{\boldsymbol e})\boldsymbol P_{\tilde m}\boldsymbol u, \tag{15}$$

其中 $\boldsymbol G = \boldsymbol I_n - \dot{\boldsymbol f}\boldsymbol D^{-1}\dot{\boldsymbol f}^{\mathrm T}\boldsymbol A^{-\mathrm T}\boldsymbol A^{-1}$,$\boldsymbol Q = \dot{\boldsymbol f}\boldsymbol D^{-1}\boldsymbol H + \dot{\boldsymbol f}_\omega$,$\boldsymbol P_{\tilde m} = \tilde{\boldsymbol m}^{\mathrm T}\tilde{\boldsymbol m}^{-1}\tilde{\boldsymbol m}^{\mathrm T}$ 以及

$$\boldsymbol E = \begin{bmatrix} 0 & 0 & 0 & \cdots & 0 & 0 \\ 1 & 0 & 0 & \cdots & 0 & 0 \\ 0 & 1 & 0 & \cdots & 0 & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & \cdots & 1 & 0 \end{bmatrix}_{n\times n}.$$

证明 由(3)式可得 $SC(\omega)$ 关于 ω 的导数为

$$\frac{\partial SC(\boldsymbol\omega)}{\partial\boldsymbol\omega} = -\frac{1}{\hat\sigma^6(\boldsymbol\omega)}\frac{\partial\hat\sigma^2(\boldsymbol\omega)}{\partial\boldsymbol\omega}\boldsymbol u^{\mathrm T}(\boldsymbol\omega)\boldsymbol P_{\tilde m}\boldsymbol u(\boldsymbol\omega) + \frac{1}{\hat\sigma^4(\boldsymbol\omega)}[\frac{\partial\boldsymbol u(\boldsymbol\omega)}{\partial\boldsymbol\omega^{\mathrm T}}]^{\mathrm T}\boldsymbol P_{\tilde m}\boldsymbol u(\boldsymbol\omega). \tag{16}$$

下面计算 $[\partial\boldsymbol u(\boldsymbol\omega)/\partial\boldsymbol\omega^{\mathrm T}]_{\omega_0}$,

对于扰动(a),由于 $e_1^2(\boldsymbol\omega) = \{y_1 + \omega_1 - f_1(\hat{\boldsymbol\beta}(\boldsymbol\omega))\}^2$,$e_i^2(\boldsymbol\omega) = \{[y_i + \omega_i - f_i(\hat{\boldsymbol\beta}(\boldsymbol\omega))] - [y_{i-1} + \omega_{i-1} - f_{i-1}(\hat{\boldsymbol\beta}(\boldsymbol\omega))]\}^2$,$i = 2,\cdots,n$. 因此在 ω_0 处有

$$[\frac{\partial e_1^2(\boldsymbol\omega)}{\partial\omega_k}]_{\omega_0} = 2\hat e_1[\delta_{1k} - \sum_{a=1}^{p}\frac{\partial f_1}{\partial\beta_a}\frac{\partial\hat\beta_a(\boldsymbol\omega)}{\partial\omega_k}]_{\omega_0},$$

$$[\frac{\partial e_i^2(\boldsymbol\omega)}{\partial\omega_k}]_{\omega_0} = 2\hat e_i\{[\delta_{ik} - \sum_{a=1}^{p}\frac{\partial f_i}{\partial\beta_a}\frac{\partial\hat\beta_a(\boldsymbol\omega)}{\partial\omega_k}] - [\delta_{i-1,k} - \sum_{a=1}^{p}\frac{\partial f_{i-1}}{\partial\beta_a}\frac{\partial\hat\beta_a(\boldsymbol\omega)}{\partial\omega_k}]\}_{\omega_0}.$$

其中 $\delta_{ik} = 1(i = k)$,$\delta_{ik} = 0(i \neq k)$,$\beta_a$ 为 $\boldsymbol\beta$ 的第 a 个分量. 则由引理1可得:

$$[\partial\boldsymbol u(\boldsymbol\omega)/\partial\boldsymbol\omega^{\mathrm T}]_{\omega_0} = 2\mathrm{diag}(\hat{\boldsymbol e})(\boldsymbol I_n - \dot{\boldsymbol f}\boldsymbol D^{-1}\dot{\boldsymbol f}^{\mathrm T}\boldsymbol A^{-\mathrm T}\boldsymbol A^{-1})(\boldsymbol I_n - \boldsymbol E).$$

从而结合引理 1 和(16) 式可得(14) 成立.

对于扰动(b),由于 $e_1^2(\omega) = \{y_1 - f_1(\hat{\beta}(\omega) \mid \omega)\}^2$,$e_i^2(\omega) = \{[y_i - f_i(\hat{\beta}(\omega) \mid \omega)] - [y_{i-1} - f_{i-1}(\hat{\beta}(\omega) \mid \omega)]\}^2$,$i = 2,\cdots,n$. 因此在 ω_0 处有

$$\left[\frac{\partial e_1^2(\omega)}{\partial \omega_k}\right]_{\omega_0} = -2\hat{e}_1\left[\frac{\partial f_1(\hat{\beta} \mid \omega)}{\partial \omega_k} + \sum_{a=1}^{p}\frac{\partial f_1}{\partial \beta_a}\frac{\partial \hat{\beta}_a(\omega)}{\partial \omega_k}\right]_{\omega_0},$$

$$\left[\frac{\partial e_i^2(\omega)}{\partial \omega_k}\right]_{\omega_0} = -2\hat{e}_i\left\{\left[\frac{\partial f_i(\hat{\beta} \mid \omega)}{\partial \omega_k} + \sum_{a=1}^{p}\frac{\partial f_i}{\partial \beta_a}\frac{\partial \hat{\beta}_a(\omega)}{\partial \omega_k}\right] - \left[\frac{\partial f_{i-1}(\hat{\beta} \mid \omega)}{\partial \omega_k} + \sum_{a=1}^{p}\frac{\partial f_{i-1}}{\partial \beta_a}\frac{\partial \hat{\beta}_a(\omega)}{\partial \omega_k}\right]\right\}_{\omega_0},$$

其中 $\partial f_i(\hat{\beta} \mid \omega)/\partial \omega_k$ 表示 β 固定仅对 ω 求导. 则由引理 2 可得:

$$[\partial u(\omega)/\partial \omega^{T}]_{\omega_0} = -2\mathrm{diag}(\hat{e})(\dot{f}_{\omega} + \dot{f}D^{-1}H)(I_n - E).$$

从而结合引理 2 和(16) 式可得(15) 成立.

2 应用

我们以氯化物数据[9] 来说明本文讨论的方法的有效性. 氯化物数据分析了氯离子通过血球壁迁移的浓度与时间的关系. 根据离子迁移理论,氯化物浓度的观测值 y_i(%) 与时间 x_i(单位:分) 的关系式是:

$$y_i = \theta_1(1 - \theta_2\exp(-\theta_3 x_i)) + \varepsilon_i,$$

其中,θ_1 为氯化物的最终百分率浓度,θ_3 是比例常数,而 θ_2 说明该氯化物未知的始末浓度以及未知的反应速度;ε_i 表示随机误差. 假设随机误差 ε_i 是 ARIMA(0,1,0) 误差序列,即:$\varepsilon_i = \sum_{t=1}^{i} e_t$,且 $e_1,e_2,\cdots,e_n$ 相互独立,$e_i \sim N(0,\sigma^2 m_i)$,$i = 1,2,\cdots,n$,同时我们取:$m_i = \exp(\lambda_1 + \lambda_2 x_i)$,记 $\lambda = (\lambda_1,\lambda_2)^{T}$. 今根据定理计算异方差检验统计量 $SC(\omega)$ 对 ω 的局部影响,考虑如下氯化物浓度受到扰动:$y_i \to y_i + \omega_i(\omega_0 = 0)$. 计算结果见表 1,由此可以看出第 1 号点是强影响点.

表 1 氯化物浓度受到扰动时 Score 统计量的局部影响分析

i	$d_{\max i}$	i	$d_{\max i}$	i	$d_{\max i}$	i	$d_{\max i}$	i	$d_{\max i}$	i	$d_{\max i}$
1	-0.5655	10	0.0798	19	-0.0653	28	-0.0482	37	0.0557	46	-0.1579
2	-0.0054	11	0.0028	20	0.0042	29	-0.1219	38	-0.0639	47	0.0529
3	0.0273	12	-0.1127	21	0.0040	30	0.1150	39	0.0058	48	-0.1608
4	-0.0446	13	0.2271	22	0.0037	31	-0.1398	40	0.0187	49	0.1496
5	0.0432	14	-0.0557	23	0.0596	32	0.1320	41	-0.0058	50	0.0913
6	-0.0507	15	0.0377	24	0.0005	33	-0.1466	42	-0.0482	51	0.0032
7	0.0344	16	-0.2263	25	-0.0503	34	0.1123	43	0.0230	52	0.1203
8	0.0820	17	0.1530	26	0.0506	35	-0.0707	44	0.1002	53	-0.0887
9	-0.1516	18	0.0020	27	0.0899	36	0.0265	45	0.0160	54	-0.0809

[参考文献]

[1] Bruce C, Martin R D. Leave-k-out diagnostics for time series [J]. J Roy Statist Soc, 1989, 51B:363—424.
[2] Tsai C L. Score test for the first-order autoregressive model with heteroscedasticity [J]. Biometrika, 1986, 73:455—460.
[3] 韦博成, 胡跃清. 非线性回归模型相关性和异方差性的检验 [J]. 工程数学学报, 1994, 11(4):1—12.
[4] 林金官. 非线性模型的异方差和变离差检验 [D]. 南京:东南大学数学系, 2002.
[5] Cook R D. Assessment of local influence [J]. J Roy Statist Soc, 1986, 48B:133—169.
[6] Thomas W. Influence on confidence regions for regression coefficients in generalized linear models [J]. Journal of the American Statistical Association, 1990, 85:393—397.
[7] Lawrance A J. Regression transformation diagnostics using local influence [J]. Journal of the American Statistical Association, 1988, 83:1067—1072.
[8] 韦博成. 加权非线性回归的 Score 检验及其局部影响分析 [J]. 应用概率统计, 1995, 11(2):147—156.
[9] Bates D M, Watts D G. Nonlinear Regression Analysis and Its Applications [M]. New York:John Wiley Sons, 1988.

[责任编辑:陆炳新]