

具有结构变化的线性回归模型的局部影响分析

解锋昌¹ 李爱萍¹ 李 勇²

(1.南京农业大学数学系,江苏 南京 210095;2.东南大学数学系,江苏 南京 210096)

摘要 利用曲率方法研究具有结构变化的线性回归模型的局部影响,分别得到了在自变量、因变量和加权扰动下相应的诊断统计量,并通过具体的数值实例说明了本文结论的有效性.

关键词 局部影响;曲率;扰动;带有结构变化的线性回归模型

中图分类号:O212.2 文献标识码:A 文章编号:1000-1980(2005)06-0717-04

在计量经济学、生物学等学科中,人们经常需要研究具有结构变化的回归模型,并且已经做了大量的工作^[1~3].李勇^[4]系统地研究了具有结构变化的线性以及非线性回归模型的异方差检验问题.本文在此基础上,利用 Cook^[5]提出的曲率方法详细研究了具有结构变化的线性回归模型在各种扰动情况下的局部影响,并得到了相应的诊断统计量.该模型的一般形式可表示为

$$Y = \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} X_1\beta \\ X_2\beta \end{bmatrix} + \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \end{bmatrix} = X\beta + \varepsilon \quad (1)$$

式中: $\varepsilon = N_n(\mathbf{0}, Q)$, $Q = \text{diag}(\sigma_1^2 I_{n_1}, \sigma_2^2 I_{n_2})$, I_{n_1}, I_{n_2} 分别是 n_1, n_2 阶单位阵; $X = (X_1^T, X_2^T)^T$, X_k 是 $n_k \times p$ 矩阵, $k = 1, 2$; $\varepsilon = (\varepsilon_1^T, \varepsilon_2^T)^T$ 为误差向量, ε_k ($k = 1, 2$) 为 $n_k \times 1$ 阶误差向量; β 为 p 维未知参数.记模型(1)中参数为 $\theta = (\beta^T, \sigma_1^2, \sigma_2^2)^T$, 则参数 θ 相应于模型(1)的对数似然函数为

$$\ell(\theta) = -\frac{n}{2} \ln 2\pi - \frac{n_1}{2} \ln \sigma_1^2 - \frac{n_2}{2} \ln \sigma_2^2 - \frac{1}{2\sigma_1^2} (y_1 - X_1\beta)^T (y_1 - X_1\beta) - \frac{1}{2\sigma_2^2} (y_2 - X_2\beta)^T (y_2 - X_2\beta) \quad (2)$$

为方便起见,记

$$A = \begin{bmatrix} X_1^T \left(\frac{1}{\sigma_1^2} I_{n_1} - \frac{2}{n_1 \sigma_1^4} e_1 e_1^T \right) X_1 + X_2^T \left(\frac{1}{\sigma_2^2} I_{n_2} - \frac{2}{n_2 \sigma_2^4} e_2 e_2^T \right) X_2 \\ - 2AX_1^T e_1 / n_1 & - 2AX_2^T e_2 / n_2 \end{bmatrix}^{-1}$$
$$V = \begin{bmatrix} A & - 2AX_1^T e_1 / n_1 & - 2AX_2^T e_2 / n_2 \\ - 2e_1^T X_1 A^T / n_1 & 4e_1^T X_1 A X_1^T e_1 / n_1^2 & 4(n_1 n_2)^{-1} e_1^T X_1 A X_2^T e_2 \\ - 2e_2^T X_2 A^T / n_2 & 4(n_1 n_2)^{-1} e_2^T X_2 A X_1^T e_1 & 4e_2^T X_2 A X_2^T e_2 / n_2^2 \end{bmatrix}$$
$$e_k = (e_{k,1}, \dots, e_{k,n_k})^T \quad I_{n_k} = (1, \dots, 1)_{n_k \times 1}^T \quad s_k = (e_{k,1}^2, \dots, e_{k,n_k}^2)^T \quad e_{k,i} = y_{ki} - X_{ki}\beta \quad (k = 1, 2)$$
$$E_1 = (1, 0, \dots, 0)_{p \times 1}^T \quad G_1 = \text{diag}(e_{1,1}, \dots, e_{1,n_1}) \quad G_2 = \text{diag}(e_{2,1}, \dots, e_{2,n_2})$$

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.局部影响分析方法是 Cook^[5]从微分几何观点出发提出的.该方法的一般思想为:令 ρ 是未知的 q 维参数,其定义域为 R^q 中的某一开集; $\ell(\rho)$ 是原模型的似然函数; ω 是一个定义在 $\Omega \subset R^n$ 上的 n 维向量,表示对模型的扰动因素; $\ell(\rho|\omega)$ 是扰动模型的似然函数.用 $\hat{\rho}, \hat{\rho}(\omega)$ 分别表示 ρ 关于 $\ell(\rho), \ell(\rho|\omega)$ 的极大似然估计.假定存在一个 ω_0 使 $\ell(\rho|\omega_0) = \ell(\rho)$ 且 $\hat{\rho} = \hat{\rho}(\omega_0)$. Cook 定义的似然距离为: $LD(\omega) = 2[\ell(\hat{\rho}) - \ell(\hat{\rho}(\omega))]$.假定 $LD(\omega)$ 关于 ω 存在二阶以上连续导数.考虑影响图 $\eta(\omega) = (\omega^T, LD(\omega))^T$, 由文献[5, 6]可知,此影响图的

影响曲率可表示为 $c_d = 2 |d^T D^T \ddot{L} D d| = 2 |d^T \Delta^T \ddot{L}^{-1} \Delta d|$, 其中 d 是 n 维单位向量, $\ddot{L}$ 为 l 关于 ρ 的二阶导数, 在 $\hat{\rho}$ 处计值; $D = \partial \hat{\rho}(\omega) / \partial \omega$, $\Delta = \partial^2 \ell(\rho | \omega) / \partial \rho \partial \omega^T$, 它们都是 $q \times n$ 阶矩阵, 且在 $(\hat{\rho}, \omega_0)$ 处计值. 以上表明, 最大影响曲率可表示为 $c_{\max} = \lambda_1$, λ_1 为影响矩阵 $F = D^T \ddot{L} D = \Delta^T \ddot{L}^{-1} \Delta$ 的绝对值最大的特征值, 对应于 λ_1 的特征向量 $d_{\max}$ 为最大影响曲率方向. 通过最大影响曲率方向的指标图可观察到模型是怎样被扰动的.

另外, 人们经常考虑扰动 ω 对 ρ 中某一部分有兴趣参数估计的影响, 为此假定 $\rho = (\rho_1^T, \rho_2^T)^T$, ρ_1 是有兴趣的部分, ρ_2 为多余部分, 则子集参数 ρ_1 的影响曲率可表示为

$$c_s(\rho_1) = 2 |d^T \Delta^T (\ddot{L}^{-1} - B_{22}) \Delta d| \quad (3)$$

其中 $\hat{\rho}(\omega) = (\hat{\rho}_1^T(\omega), \hat{\rho}_2^T(\hat{\rho}_1(\omega)))^T$, $\hat{\rho}_2(\rho_1)$ 为 ρ_2 相应于 ρ_1 的极大似然估计. 同时 B_{22} 可由 $\ddot{L}$ 相应于 ρ_1, ρ_2 的分块矩阵得到:

$$B_{22} = \begin{bmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \ddot{L}_{22}^{-1} \end{bmatrix} \quad \ddot{L} = \begin{bmatrix} \ddot{L}_{11} & \ddot{L}_{12} \\ \ddot{L}_{21} & \ddot{L}_{22} \end{bmatrix}$$

1.1 自变量扰动的局部影响分析

考虑单个自变量扰动模型, 不妨设第一个自变量上有扰动, 令扰动向量为 $\omega = (\omega_1^T, \omega_2^T)^T$, $\omega_k = (\omega_{k,1}, \dots, \omega_{k,n_k})_{n_k \times 1}^T, k=1, 2$, $\omega^0 = \mathbf{0}$ 表示模型没有扰动. 则扰动后模型(1)的对数似然函数为

$$\begin{aligned} \ell(\theta | \omega) = & -\frac{n}{2} \ln 2\pi - \frac{n_1}{2} \ln \sigma_1^2 - \frac{1}{2\sigma_1^2} (y_1 - (X_1 + \omega_1 E_1^T) \beta)^T (y_1 - (X_1 + \omega_1 E_1^T) \beta) - \\ & \frac{n_2}{2} \ln \sigma_2^2 - \frac{1}{2\sigma_2^2} (y_2 - (X_2 + \omega_2 E_2^T) \beta)^T (y_2 - (X_2 + \omega_2 E_2^T) \beta) \end{aligned} \quad (4)$$

定理1 当 β 为有兴趣参数, σ_1^2, σ_2^2 为多余参数时, $\ell(\omega)$ 在 ω^0 处的正则曲率为

$$c_s(\beta) = 2 |d^T (\Delta_1^T V \Delta_1) d|$$

其中

$$\Delta_1^T = \begin{bmatrix} (e_1 E_1^T - \beta_1 X_1) / \sigma_1^2 & -\beta_1 e_1 / \sigma_1^4 & \mathbf{0} \\ (e_2 E_1^T - \beta_1 X_2) / \sigma_2^2 & \mathbf{0} & -\beta_1 e_2 / \sigma_2^4 \end{bmatrix}$$

证明 由模型(1)的对数似然函数(2)可得 $\ell(\theta)$ 在 $\hat{\theta}$ 处关于 θ 的二阶导数为

$$\begin{aligned} \frac{\partial^2 \ell(\theta)}{\partial \beta \partial \beta^T} \Big|_{\theta=\hat{\theta}} &= -\frac{1}{\hat{\sigma}_1^2} X_1^T X_1 - \frac{1}{\hat{\sigma}_2^2} X_2^T X_2 & \frac{\partial^2 \ell(\theta)}{\partial \beta \partial \sigma_1^2} \Big|_{\theta=\hat{\theta}} &= -\frac{1}{\hat{\sigma}_1^4} X_1^T (y_1 - X_1 \hat{\beta}) \\ \frac{\partial^2 \ell(\theta)}{\partial \beta \partial \sigma_2^2} \Big|_{\theta=\hat{\theta}} &= -\frac{1}{\hat{\sigma}_2^4} X_2^T (y_2 - X_2 \hat{\beta}) & \frac{\partial^2 \ell(\theta)}{(\partial \sigma_1^2)^2} \Big|_{\theta=\hat{\theta}} &= -\frac{n_1}{2\hat{\sigma}_1^4} \\ \frac{\partial^2 \ell(\theta)}{\partial \sigma_1^2 \partial \sigma_2^2} \Big|_{\theta=\hat{\theta}} &= 0 & \frac{\partial^2 \ell(\theta)}{(\partial \sigma_2^2)^2} \Big|_{\theta=\hat{\theta}} &= -\frac{n_2}{2\hat{\sigma}_2^4} \end{aligned}$$

于是得观测信息矩阵 $-\ddot{L}$ 为

$$-\ddot{L} = \begin{bmatrix} X_1^T X_1 / \hat{\sigma}_1^2 + X_2^T X_2 / \hat{\sigma}_2^2 & X_1^T (y_1 - X_1 \hat{\beta}) / \hat{\sigma}_1^4 & X_1^T (y_2 - X_2 \hat{\beta}) / \hat{\sigma}_2^4 \\ (y_1 - X_1 \hat{\beta})^T X_1 / \hat{\sigma}_1^4 & n_1 / (2\hat{\sigma}_1^4) & 0 \\ (y_2 - X_2 \hat{\beta})^T X_2 / \hat{\sigma}_2^4 & 0 & n_2 / (2\hat{\sigma}_2^4) \end{bmatrix} \quad (5)$$

由式(4)得 $\ell(\theta | \omega)$ 在 $(\hat{\theta}, \omega^0)$ 处关于 θ, ω 的二阶导数

$$\begin{aligned} \frac{\partial^2 \ell(\theta | \omega)}{\partial \beta_1 \partial \omega_{1i}} \Big|_{(\hat{\theta}, \omega^0)} &= \frac{1}{\hat{\sigma}_1^2} (y_{1i} - \sum_{c=2}^p X_{1,ic} \hat{\beta}_c - X_{1,i1} \hat{\beta}_1) - \frac{1}{\hat{\sigma}_1^4} \hat{\beta}_1 X_{1,i1} \\ \frac{\partial^2 \ell(\theta | \omega)}{\partial \beta_1 \partial \omega_{2j}} \Big|_{(\hat{\theta}, \omega^0)} &= \frac{1}{\hat{\sigma}_2^2} (y_{2j} - \sum_{c=2}^p X_{2,jc} \hat{\beta}_c - X_{2,j1} \hat{\beta}_1) - \frac{1}{\hat{\sigma}_2^4} \hat{\beta}_1 X_{2,j1} \\ \frac{\partial^2 \ell(\theta | \omega)}{\partial \beta_k \partial \omega_{1i}} \Big|_{(\hat{\theta}, \omega^0)} &= -\frac{1}{\hat{\sigma}_1^2} \hat{\beta}_1 X_{1,ik} & \frac{\partial^2 \ell(\theta | \omega)}{\partial \beta_k \partial \omega_{2j}} \Big|_{(\hat{\theta}, \omega^0)} &= -\frac{1}{\hat{\sigma}_2^2} \hat{\beta}_1 X_{2,jk} \quad (k=2, \dots, p) \\ \frac{\partial^2 \ell(\theta | \omega)}{\partial \sigma_1^2 \partial \omega_{1i}} \Big|_{(\hat{\theta}, \omega^0)} &= -\frac{\hat{\beta}_1}{\hat{\sigma}_1^4} (y_{1i} - \sum_{c=1}^p X_{1,ic} \hat{\beta}_c) & \frac{\partial^2 \ell(\theta | \omega)}{\partial \sigma_1^2 \partial \omega_{2j}} \Big|_{(\hat{\theta}, \omega^0)} &= 0 \end{aligned}$$

$$\frac{\partial^2 \ell(\boldsymbol{\theta}|\boldsymbol{\omega})}{\partial \sigma_2^2 \partial \omega_{1i}} \Big|_{(\hat{\boldsymbol{\theta}}, \boldsymbol{\omega}^0)} = 0 \qquad \frac{\partial^2 \ell(\boldsymbol{\theta}|\boldsymbol{\omega})}{\partial \sigma_2^2 \partial \omega_{2j}} \Big|_{(\hat{\boldsymbol{\theta}}, \boldsymbol{\omega}^0)} = -\frac{\hat{\beta}_1}{\hat{\sigma}_2^4} \Big(y_{2j} - \sum_{c=1}^p X_{2j} \hat{\beta}_c \Big)$$

于是在 $\boldsymbol{\omega}^0$ 处可得 Δ_1 , 将 $\ddot{L}, \Delta_1$ 代入子集参数曲率公式 (3), 并结合前面记号得定理成立.

1.2 响应变量扰动模型

考虑响应变量扰动模型, 即 $(y_1^T, y_2^T)^T$ 上有扰动变为 $(y_1^T, y_2^T)^T + \boldsymbol{\omega}$, $\boldsymbol{\omega} = (\omega_1^T, \omega_2^T)^T$, $\omega_k = (\omega_{k,1}, \dots, \omega_{k,n_k})^T$ ($k = 1, 2$), $\boldsymbol{\omega}^0 = \mathbf{0}$ 表示模型没有扰动. 则扰动后模型 (1) 的对数似然函数为

$$\begin{aligned} \ell(\boldsymbol{\theta}|\boldsymbol{\omega}) = & -\frac{n}{2} \ln 2\pi - \frac{n_1}{2} \ln \sigma_1^2 - \frac{1}{2\sigma_1^2} (y_1 + \omega_1 - X_1 \boldsymbol{\beta})^T (y_1 + \omega_1 - X_1 \boldsymbol{\beta}) - \\ & \frac{n_2}{2} \ln \sigma_2^2 - \frac{1}{2\sigma_2^2} (y_2 + \omega_2 - X_2 \boldsymbol{\beta})^T (y_2 + \omega_2 - X_2 \boldsymbol{\beta}) \end{aligned} \tag{6}$$

定理2 当 $\boldsymbol{\beta}$ 为有兴趣参数, σ_1^2, σ_2^2 为多余参数时, $\eta(\boldsymbol{\omega})$ 在 $\boldsymbol{\omega}^0$ 处的正则曲率为

$$c_s(\boldsymbol{\beta}) = 2 |d^T (\Delta_2^T V \Delta_2) d|$$

其中
$$\Delta_2^T = \begin{bmatrix} X_1/\sigma_1^2 & e_1/\sigma_1^4 & \mathbf{0} \\ X_2/\sigma_2^2 & \mathbf{0} & e_2/\sigma_2^4 \end{bmatrix}$$

证明 由式 (6) 得 $\ell(\boldsymbol{\theta}|\boldsymbol{\omega})$ 在 $\boldsymbol{\omega}^0$ 处关于 $\boldsymbol{\theta}, \boldsymbol{\omega}$ 的二阶导数为

$$\begin{aligned} \frac{\partial^2 \ell(\boldsymbol{\theta}|\boldsymbol{\omega})}{\partial \boldsymbol{\beta} \partial \omega_1^T} \Big|_{\boldsymbol{\omega}^0} &= \frac{1}{\sigma_1^2} X_1^T & \frac{\partial^2 \ell(\boldsymbol{\theta}|\boldsymbol{\omega})}{\partial \boldsymbol{\beta} \partial \omega_2^T} \Big|_{\boldsymbol{\omega}^0} &= \frac{1}{\sigma_2^2} X_2^T & \frac{\partial^2 \ell(\boldsymbol{\theta}|\boldsymbol{\omega})}{\partial \sigma_1^2 \partial \omega_1^T} \Big|_{\boldsymbol{\omega}^0} &= \frac{1}{\sigma_1^4} e_1^T \\ \frac{\partial^2 \ell(\boldsymbol{\theta}|\boldsymbol{\omega})}{\partial \sigma_1^2 \partial \omega_2^T} \Big|_{\boldsymbol{\omega}^0} &= \mathbf{0} & \frac{\partial^2 \ell(\boldsymbol{\theta}|\boldsymbol{\omega})}{\partial \sigma_2^2 \partial \omega_1^T} \Big|_{\boldsymbol{\omega}^0} &= \mathbf{0} & \frac{\partial^2 \ell(\boldsymbol{\theta}|\boldsymbol{\omega})}{\partial \sigma_2^2 \partial \omega_2^T} \Big|_{\boldsymbol{\omega}^0} &= \frac{1}{\sigma_2^4} e_2^T \end{aligned}$$

于是在 $\boldsymbol{\omega}^0$ 处可得 Δ_2 , 将 $\ddot{L}, \Delta_2$ 代入子集参数曲率公式 (3), 并结合前面记号得定理成立.

1.3 加权扰动模型

首先赋予每个数据点的对数似然函数贡献不同的权, 从而得到扰动模型. 令 $\boldsymbol{\omega} = (\omega_1^T, \omega_2^T)^T$, $\omega_k = (\omega_{k,1}, \dots, \omega_{k,n_k})^T$ ($k = 1, 2$) 为加权扰动向量, $\boldsymbol{\omega}^0 = (1, \dots, 1)^T$ 表示模型没有扰动. 则扰动后模型 (1) 的对数似然函数为

$$\begin{aligned} \ell(\boldsymbol{\theta}|\boldsymbol{\omega}) = & \sum_{i=1}^{n_1} \omega_{1i} \left\{ -\frac{1}{2} \ln 2\pi - \frac{1}{2} \ln \sigma_1^2 - \frac{1}{2\sigma_1^2} (y_{1i} - X_{1i}^T \boldsymbol{\beta})^2 \right\} + \\ & \sum_{j=1}^{n_2} \omega_{2j} \left\{ -\frac{1}{2} \ln 2\pi - \frac{1}{2} \ln \sigma_2^2 - \frac{1}{2\sigma_2^2} (y_{2j} - X_{2j}^T \boldsymbol{\beta})^2 \right\} \end{aligned} \tag{7}$$

定理3 当 $\boldsymbol{\beta}$ 为有兴趣参数, σ_1^2, σ_2^2 为多余参数时, $\eta(\boldsymbol{\omega})$ 在 $\boldsymbol{\omega}^0$ 处的正则曲率为

$$c_s(\boldsymbol{\beta}) = 2 |d^T (\Delta_3^T V \Delta_3) d|$$

其中
$$\Delta_3^T = \begin{bmatrix} G_1^T X_1/\sigma_1^2 & s_1/(2\sigma_1^4) - I_{n_1}/(2\sigma_1^2) & \mathbf{0} \\ G_2^T X_2/\sigma_2^2 & \mathbf{0} & s_2/(2\sigma_2^4) - I_{n_2}/(2\sigma_2^2) \end{bmatrix}$$

证明 由式 (7) 得 $\ell(\boldsymbol{\theta}|\boldsymbol{\omega})$ 在 $\boldsymbol{\omega}^0$ 处关于 $\boldsymbol{\theta}, \boldsymbol{\omega}$ 的二阶导数为

$$\begin{aligned} \frac{\partial^2 \ell(\boldsymbol{\theta}|\boldsymbol{\omega})}{\partial \boldsymbol{\beta} \partial \omega_1^T} \Big|_{\boldsymbol{\omega}^0} &= \frac{1}{\sigma_1^2} X_1^T G_1 & \frac{\partial^2 \ell(\boldsymbol{\theta}|\boldsymbol{\omega})}{\partial \boldsymbol{\beta} \partial \omega_2^T} \Big|_{\boldsymbol{\omega}^0} &= \frac{1}{\sigma_2^2} X_2^T G_2 & \frac{\partial^2 \ell(\boldsymbol{\theta}|\boldsymbol{\omega})}{\partial \sigma_1^2 \partial \omega_1^T} \Big|_{\boldsymbol{\omega}^0} &= \frac{s_1^T}{2\sigma_1^4} - \frac{I_{n_1}^T}{2\sigma_1^2} \\ \frac{\partial^2 \ell(\boldsymbol{\theta}|\boldsymbol{\omega})}{\partial \sigma_1^2 \partial \omega_2^T} \Big|_{\boldsymbol{\omega}^0} &= \mathbf{0} & \frac{\partial^2 \ell(\boldsymbol{\theta}|\boldsymbol{\omega})}{\partial \sigma_2^2 \partial \omega_1^T} \Big|_{\boldsymbol{\omega}^0} &= \mathbf{0} & \frac{\partial^2 \ell(\boldsymbol{\theta}|\boldsymbol{\omega})}{\partial \sigma_2^2 \partial \omega_2^T} \Big|_{\boldsymbol{\omega}^0} &= \frac{s_2^T}{2\sigma_2^4} - \frac{I_{n_2}^T}{2\sigma_2^2} \end{aligned}$$

于是在 $\boldsymbol{\omega}^0$ 处可得 Δ_3 , 将 $\ddot{L}, \Delta_3$ 代入子集参数曲率公式 (3), 并结合前面记号得定理成立.

2 南方澳大利亚地区洋葱产量和密度数据

南方澳大利亚农业局的 Ian Rogers 提供了棕色西班牙良种洋葱在 Mt. Gambier (MG) 和 Uraidla (U) 两个南方澳大利亚地区栽培得到的产量和密度数据^[7]. 其中用 $x_{1,i}, y_{1,i}$ ($i = 1, \dots, 42$) 分别表示在 MG 地区栽培的产

量和密度 $x_{2j}, y_{2j} (j=1, \dots, 42)$ 分别表示在 U 地区栽培的产量和密度. 其中, 有一个明显的异常点, 在计算中略去不计. 根据文献 [7] 的讨论, 利用下面具有结构变化的线性回归模型度量洋葱的产量的倒数与种植密度的关系.

$$Y = \begin{bmatrix} Y_1 \\ Y_2 \end{bmatrix} = \begin{bmatrix} \beta_1 I_{n_1} + \beta_2 x_1 \\ \beta_1 I_{n_2} + \beta_2 x_2 \end{bmatrix} + \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \end{bmatrix} = \begin{bmatrix} X_1 \\ X_2 \end{bmatrix} \beta + \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \end{bmatrix}$$

式中: $Y_1 = (y_{11}, \dots, y_{142})^T, x_1 = (x_{11}, \dots, x_{142})^T; Y_2 = (y_{21}, \dots, y_{241})^T, x_2 = (x_{21}, \dots, x_{241})^T; X_1, X_2$ —— 42×2 和 41×2 矩阵; β —— 2×1 未知参数; ε_k ——误差 ($k=1, 2$). 下面考虑自变量、因变量扰动时的局部影响, 结果见图 1 (其中 m 表示数据点). 由图 1 可知, MG 地区第 33、41、42 号点是强影响点, U 地区第 79、80、83 号点是强影响点.

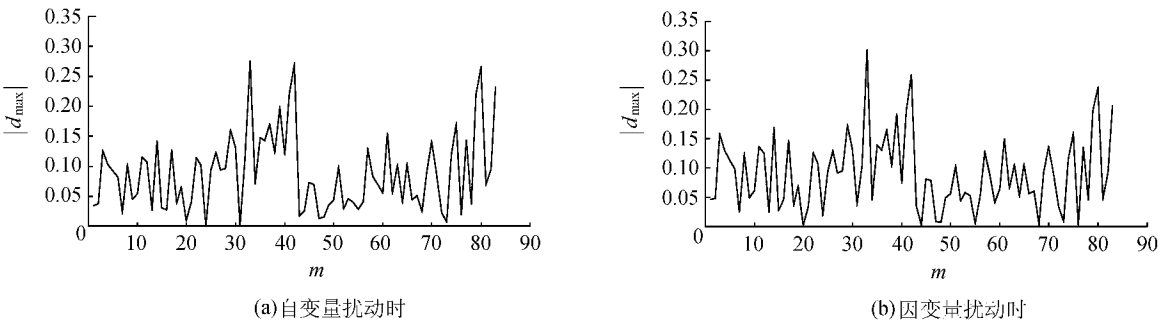


图 1 变量扰动时 $|d_{\max}|$ 的散点图
Fig.1 Index plot of $|d_{\max}|$ under perturbation of explanatory and response variables

参考文献:

[1] JIRO H. Finite-sample properties of single-equation estimators under-structural change[J]. Journal of Econometrics, 1992, 53(1—3): 189—209.
[2] DONALD W K, ANDREWS R C F. Inference in nonlinear econometrics models with structural change[J]. The Review of Economic Studies, 1988, 55(4): 615—639.
[3] 吴启光, 徐兴忠. 带有结构变化的线性模型中参数估计的一些结果[J]. 数学年刊, 2001, A-22: 607—616.
[4] 李勇. 具有结构变化的非线性回归模型的阶段异方差和相关性检验[D]. 南京: 东南大学, 2003.
[5] COOK R D. Assessment of local influence[J]. J R Statist Soc, 1986, B-48: 133—169.
[6] 韦博成, 鲁国斌, 史建清. 统计诊断引论[M]. 南京: 东南大学出版社, 1991. 302—345.
[7] RATKOWSK Y D A. 非线性回归模型——统一的实用方法[M]. 洪再吉, 韦博成, 吴诚鸥, 等, 译. 南京: 南京大学出版社, 1986. 49—59.

Local influence of linear regression model with structural change

XIE Feng-chang¹, LI Ai-ping¹, LI Yong²

(1. Department of Mathematics, Nanjing Agricultural University, Nanjing 210095, China;
2. Department of Mathematics, Southeast University, Nanjing 210096, China)

Abstract The study was performed by use of the curvature method. The diagnostic statistics under perturbation of explanatory variables, response variable and case weights were obtained respectively, and the effectiveness of the conclusion was verified by a numerical example.

Key words local influence; curvature; perturbation; linear regression model with structural change