应用概率统计 第十一卷
第二期 1995年5月

Chinese Journal of Applied Probability
and Statistics Vol. 11, No. 2, May 1995

加权非线性回归的Score检验及其局部影响分析*

韦 博 成
(东南大学, 南京, 210018)

摘 要

在回归分析中, 随机误差是否存在方差齐性是理论与应用工作者都十分关心的问题. 对线性回归, 已有许多讨论 ([1], [2]). 本文则讨论加权非线性回归方差齐性的假设检验问题. 本文把Cook & Weiberg([3]) 关于线性回归方差齐性Score 检验的结果推广到非线性回归; 并利用Cox & Reid([4]) 的参数正交化方法得到修正的Score 检验统计量. 同时, 本文还讨论了微小扰动对于 Score 统计量的局部影响, 得到了度量最大局部影响的诊断统计量.

关键词: 方差齐性, 加权非线性回归, score 统计量, 参数正交性, 梯度, 局部影响.
学科分类号: 212. 1.

§1. 引 言

加权非线性回归可表示为

$$Y=\eta(\beta)+\varepsilon, \qquad \varepsilon\sim N(0,\sigma^2 M), \tag{1.1}$$

其中 $Y=(y_1,\ldots,y_n)^T$ 为因变量; $\eta(\beta)=(\eta(x_1,\beta),\ldots,\eta(x_n,\beta))^T=(\eta_1(\beta),\ldots,\eta_n(\beta))^T$, β 为 p 维未知参数向量, $\eta(\cdot,\cdot)$ 为已知函数, $x_1,\ldots,x_n$ 为自变量; $\varepsilon=(\varepsilon_1,\ldots,\varepsilon_n)^T$ 为随机误差向量, σ^2为尺度参数; $M=\mathrm{diag}\,(m_1,\ldots,m_n)$ 为加权矩阵, $\mathrm{Var}\,(\varepsilon_i)=\sigma^2 m_i$, m_i 为方差的权. 加权回归, 即异方差回归是一个老问题 ([1]). 近年来, 由于理论与实际问题的需要, 又重新受到人们的广泛重视. 1988 年, Carroll & Ruppert ([2]) 出版了专著, 系统地研究了加权回归的各种问题. 文 [3] 研究了线性回归方差齐性的 Score 检验统计量, 得到了很好的结果. 本文着重讨论加权非线性回归的情形, 并根据 Cox & Reid ([4]) 的参数正交化方法, 得到修正的 Score 检验统计量. §3 根据 Cook ([5]) 和 Thomas ([6]) 的方法讨论了微小扰动对于 Score 检验统计量的局部影响, 得到了度量最大局部影响的诊断统计量. 最后通过数值计算分析了所得到的结果.

以下介绍本文所讨论的模型、问题, 以及若干必要的记号和公式.

在模型 (1.1) 中, 通常假定方差 $\mathrm{Var}\,(\varepsilon_i)=\sigma^2 m_i$ 与某个协变量 Z_i 以及某个参数 δ 有关 ([2],[3]):

$$m_i=m(Z_i,\delta), \quad i=1,\ldots,n. \tag{1.2}$$

$m(\cdot,\cdot)$ 为已知函数, δ 的维数 $q\le n-p-1$, 并假设存在某个 δ_0, 对一切 i 有 $m(Z_i,\delta_0)\equiv 1$. 本

*国家自然科学基金资助项目.
本文1992 年 12 月 25 日收到. 1993 年 5 月 10 日收到修改稿.

文考虑以下假设检验问题:

$$H_0: \delta = \delta_0; \qquad H_1: \delta \neq \delta_0. \tag{1.3}$$

易见. 若 H_0 成立, 则 $M = I$ 为单位矩阵, 因而模型 (1.1) 具有方差齐性; 若 H_0 被否定, 则 (1.1) 具有异方差. 本文 §2 就是讨论模型 (1.1) 和 (1.2) 的假设检问题 H_0. 并求出 Score 统计量.

对于模型 (1.1) 和 (1.2), 参数为 $\theta = (\delta^T, \sigma^2, \beta^T)^T$, 相应的对数似然为:

$$L(\theta) = -\frac{n}{2}\log\sigma^2 - \frac{1}{2}\sum_{i=1}^{n}\log m_i - \frac{1}{2\sigma^2}S(\delta,\beta), \tag{1.4}$$

$$S(\delta,\beta) = \{Y - \eta(\beta)\}^T M^{-1}(\delta)\{Y - \eta(\beta)\}. \tag{1.5}$$

对于问题 (1.3), $\theta_1 = \delta$ 为有兴趣的参数, $\theta_2 = (\sigma^2, \beta^T)^T$ 为多余参数. 当 H_0 成立时, $M = I$, 相应的极大似然估计记为 $\widehat{\theta}_0 = (\delta_0^T, \widehat{\sigma}^2, \widehat{\beta}^T)^T$. 经直接计算可得 $\widehat{\sigma}^2 = n^{-1}S(\delta_0, \widehat{\beta})$, 且 $\widehat{\beta}$ 满足:

$$\{\partial S(\delta_0,\beta)/\partial\beta\}_{\beta=\widehat{\beta}} = 0 \tag{1.6}$$

对于模型 (1.1), 假设 $\eta(\beta)$ 关于 β 的前二阶导数存在, 并记为 $V(\beta) = \partial\eta/\partial\beta^T, W(\beta)= \partial^2\eta/\partial\beta\partial\beta^T$, 它们分别为 $n \times p$ 阶列满秩矩阵和 $n \times p \times p$ 阶立体阵. $V(\beta)$ 的上三角分解为 $V = (Q, N)(R^T, 0) = QR, (Q, N)$ 为正交阵, R 以及 $R^{-1} = K$ 为上三角矩阵. 又记 (1.1) 的固有曲率立体阵为 A^I (见 [7]). 对于 (1.2), 假设 $m(\delta)$ 关于 δ 的导数 $D = (\partial m_i/\partial\delta_a)$ 存在, 且为 $n \times q$ 阶列满秩矩阵.

§2. Score 检验统计量

对于问题 (1.3), 其 Score 检验统计量为

$$SC = \left\{\left(\frac{\partial L}{\partial\delta}\right)^T J^{\delta\delta}\left(\frac{\partial L}{\partial\delta}\right)\right\}_{\theta=\widehat{\theta}_0}, \tag{2.1}$$

其中 $J^{\delta\delta}$ 为 J^{-1} 中对应于参数 δ 的分块矩阵, J 为 (1.4) 式参数 θ 的 Fisher 信息阵, 且 SC 的渐近分布为 $\chi^2(q)$, 由 (1.4) 式经直接计算可得, J 按 δ, σ^2 和 β 的分块矩阵为:

$$\begin{pmatrix} \frac{1}{2}D^T M^{-2} D & \frac{1}{2\sigma^2}D^T M^{-1}\mathbf{1} & 0 \\ \frac{1}{2\sigma^2}\mathbf{1}^T M^{-1} D & \frac{n}{2\sigma^4} & 0 \\ 0 & 0 & \frac{1}{\sigma^2}V^T M^{-1} V \end{pmatrix}, \tag{2.2}$$

其中 $\mathbf{1} = (1, \ldots, 1)^T$ 为 n 维向量. 由 (2.1) 和 (2.2) 可得:

定理 1 对于模型 (1.1) 和 (1.2), 假设检验问题 (1.3) 的 Score 检验统计量可表示为

$$SC = u^T P_{\overline{D}} u/(2\widehat{\sigma}^4), \tag{2.3}$$

其中 $u = (\widehat{e}_1^2, \ldots, \widehat{e}_n^2)^T, \widehat{e}_i = y_i - \eta_i(\widehat{\beta}), \overline{D} = (I - \mathbf{1}\mathbf{1}^T/n)D, P_{\overline{D}} = \overline{D}(\overline{D}^T\overline{D})^{-1}\overline{D}^T$ 且在 $\widehat{\theta}_0$ 处计值.

证明 以下均在 $\widehat{\theta}_0$ 处计值, 故略去. 由 (1.4) 式经直接计算可得

$$\frac{\partial L}{\partial\delta} = -\frac{1}{2}D^T(\mathbf{1} - u/\widehat{\sigma}^2).$$

由于 $\mathbf{1}^T u = n\widehat{\sigma}^2$, 代入上式可得

$$\frac{\partial L}{\partial \delta} = -\frac{1}{2}D^T\left(\frac{\mathbf{1}\mathbf{1}^T u}{n\widehat{\sigma}^2} - \frac{u}{\widehat{\sigma}^2}\right) = \frac{1}{2}D^T(I - \mathbf{1}\mathbf{1}^T/n)u\widehat{\sigma}^{-2} = \overline{D}^T u/(2\widehat{\sigma}^2). \tag{2.4}$$

又在 (2.2) 式中, $\delta = \delta_0$ 时 $M = I$, 因而由分块矩阵求逆公式可得

$$\begin{aligned} J^{\delta\delta} &= \left\{\frac{1}{2}D^T D - \left(\frac{1}{2\widehat{\sigma}^2}D^T\mathbf{1}\right)\left(\frac{n}{2\widehat{\sigma}^4}\right)^{-1}\left(\frac{1}{2\widehat{\sigma}^4}\mathbf{1}^T D\right)\right\}^{-1} \\ &= 2\{D^T(I - \mathbf{1}\mathbf{1}^T/n)D\}^{-1} = 2(\overline{D}^T\overline{D})^{-1}. \end{aligned} \tag{2.5}$$

(2.4) 式、(2.5) 式代入 (2.1), 即可得到 (2.3) 式. □

如果 (1.1) 式为线性模型, 即 $\eta(\beta) = X\beta$, X 为 $n \times p$ 阶矩阵, 则 $\widehat{\beta} = (X^T X)^{-1}X^T Y$, $\widehat{e} = Y - X\widehat{\beta}$, 因而 (2.3) 式的结果与文 [3] 的公式完全一致, 这正是我们所预期的.

根据 Cox & Reid ([4]) 的论证, 当有兴趣的参数 (本文为 δ) 与多余参数 (本文为 σ^2 和 β) 不正交时 (本文 $J_{\delta\sigma} = D^T M^{-1}\mathbf{1}/(2\sigma^2) \neq 0$), 似然比统计量以及相应的 Score 统计量需要加以调整. 为此, 首先作参数变换, 使得在新参数下, 有兴趣的参数与多余参数正交, 然后再加上适当的调整项. 这种参数变换不一定存在, 但是本文情况, 这种变换恰好存在. 我们有

引理 1 若取新参数 $\widetilde{\theta} = (\delta^T, \gamma, \beta^T)^T$:,

$$\gamma = \sigma^2 m_g(\delta), \qquad m_g(\delta) = \left[\prod_{i=1}^{n} m_i(\delta)\right]^{1/n}, \tag{2.6}$$

则在新参数下, δ 与 γ, β 正交. 这时 (1.1) 式中, 变量 Y 关于 $\widetilde{\theta}$ 的 Fisher 信息阵 $\widetilde{J}$ 为分块对角阵

$$\widetilde{J} = \mathrm{diag}(\widetilde{J}_{\delta\delta}, \widetilde{J}_2). \tag{2.7}$$

$\widetilde{J}_{\delta\delta} = 2^{-1}D^T M^{-1}(I - \mathbf{1}\mathbf{1}^T/n)M^{-1}D$, $\widetilde{J}_2 = \mathrm{diag}(n/2\gamma, m_g^T V^T M^{-1}V/\gamma)$.

证明 根据文 [4] 的论证, 可令 $\sigma^2 = \sigma^2(\delta, \gamma)$ 为 γ, δ 的函数, 它应满足以下微分方程:

$$\frac{\partial\sigma^2(\delta, \gamma)}{\partial\delta} = -J_{\sigma\sigma}^{-1}J_{\sigma\delta}.$$

上式右端由 (2.2) 式确定, $J_{\sigma\sigma} = n/2\sigma^4$, $J_{\sigma\delta} = D^T M^{-1}\mathbf{1}/2\sigma^2$, 因此上述微分方程化为

$$\frac{\partial\sigma^2}{\partial\delta_a} = -\frac{\sigma^2}{n}\sum_{i=1}^{n} m_i^{-1}\frac{\partial m_i}{\partial\delta_a}, \qquad a = 1, \ldots, q.$$

该式可进一步表示为

$$\frac{\partial\log\sigma^2}{\partial\delta_a} = -\frac{1}{n}\sum_{i=1}^{n}\frac{\partial\log m_i}{\partial\delta_a} = \frac{\partial\log(m_g^{-1})}{\partial\delta_a}.$$

由此即可解出 $\sigma^2 = \gamma m_g^{-1}$, 此即 (2.6) 式, $\sigma^2 = \gamma m_g^{-1}(\delta)$ 代入 (1.4) 式, 经直接计算可得到 (2.7). □

根据 Cox & Reid ([4]) 的公式, 修正的 Score 统计量为 $SC_m = SC + \Delta SC$,

$$\Delta SC = -\sum_{a=1}^{q} tr\left\{\widetilde{J}_2^{-1}(\delta)\partial[\widetilde{J}_2(\delta)]/\partial\delta_a\right\}(\widetilde{\delta}_a - \delta_{0a}), \tag{2.8}$$

其中 $\widetilde{\delta}=(\widetilde{\delta}_1,\cdots,\widetilde{\delta}_q)^T$ 为 δ 的 MLE, $\delta_0=(\delta_{01},\cdots,\delta_{0q})^T, \widetilde{\delta}-\delta_0$ 可近似表示为

$$\widetilde{\delta}-\delta_0\approx\left[\widetilde{J}_{\delta\delta}^{-1}\frac{\partial L}{\partial\delta}\right]. \tag{2.9}$$

以上 (2.8) 式和 (2.9) 式均在 $\widehat{\theta}_0=(\delta_0^T,\widehat{\sigma}^2,\widehat{\beta}^T)^T$ 处计值, 以下同此, 不再赘述.

定理 2 对于模型 (1.1) 和 (1.2), 假设检验问题 (1.3) 的 Score 统计量可表示为

$$SC_m=(u+2\widehat{\sigma}^2h)^TP_{\overline{D}}u/(2\widehat{\sigma}^4), \tag{2.10}$$

其中 $h=(h_{11},\ldots,h_{nn})^T, h_{11},\ldots,h_{nn}$ 为 $H=V(V^TV)^{-1}V^T$ 的对角元素.

证明 由 (2.7) 式可知

$$\frac{\partial\widetilde{J}_2(\delta)}{\partial\delta_a}=\operatorname{diag}\left(0,\quad\gamma^{-1}V^T\frac{\partial(m_gM^{-1})}{\partial\delta_a}V\right),$$

其中 m_gM^{-1} 为对角阵, 其第 i 个对角元为 $a_i=m_gm_i^{-1}$. 因此由 (2.6) 式可知

$$\begin{aligned}\left(\frac{\partial a_i}{\partial\delta_a}\right)_{\widehat{\theta}_0}&=\left[a_i\frac{\partial\log\, a_i}{\partial\delta_a}\right]_{\widehat{\theta}_0}=\left[a_i(-m_i^{-1}D_{ia}+n^{-1}\sum_{j=1}^n m_j^{-1}D_{ja})\right]_{\widehat{\theta}_0}\\&=-D_{ia}+n^{-1}\sum_{j=1}^n D_{ja}=-\overline{D}_{ia},\end{aligned}$$

其中 $D=(D_{ia}),\overline{D}=(\overline{D}_{ia})$. 因此在 $\widehat{\theta}_0$ 处有

$$\frac{\partial(m_gM^{-1})}{\partial\delta_a}=\operatorname{diag}(\overline{D}_{ia}),\qquad i=1,\ldots,n.$$

$\operatorname{diag}(\overline{D}_{ia})$ 为一对角阵 (a 固定). 又由于在 $\widehat{\theta}_0$ 处有 $\widetilde{J}_2^{-1}=\operatorname{diag}(2\widehat{\sigma}^2/n,\widehat{\sigma}^2(V^TV)^{-1})$, 因此有

$$\begin{aligned}\operatorname{tr}\left\{\widetilde{J}_2^{-1}(\delta)\partial[\widetilde{J}_2(\delta)/\partial\delta_a]\right\}&=\operatorname{tr}\left\{(V^TV)^{-1}V^T\operatorname{diag}(-\overline{D}_{ia})V\right\}\\&=-\operatorname{tr}\left\{H\operatorname{diag}(\overline{D}_{ia})\right\}=-\sum_{i=1}^n hii\overline{D}_{ia}.\end{aligned} \tag{2 .11}$$

又 (2.4)、(2.7) 的结果代入 (2.9) 式可得

$$\widetilde{\delta}-\delta_0\approx\sigma^{-2}(\overline{D}^T\overline{D})^{-1}\overline{D}^Tu. \tag{2.12}$$

(2.11) 和 (2.12) 式代入 (2.8) 式可得

$$\Delta SC=\widehat{\sigma}^{-2}h^T\overline{D}(\overline{D}^T\overline{D})^{-1}\overline{D}^Tu.$$

由此即得到 (2.10) 式. □

§3. Score统计量的局部影响分析

局部影响主要研究模型有微小扰动时, 对于统计量的影响, 由此可导出某些特定的诊断统计量. 局部影响首先由 Cook ([5]) 于 1986 年提出. 假设 ω 为刻划扰动的向量, ω_0 对应于无扰动情形. 在文 [5] 的讨论中, 由于受扰动以后的似然距离 $LD(\omega)$ 在 ω_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划扰动的影响. 但是, 对于某些其他的统计量, 它们在 ω_0 处的一阶导数不为零,则可用方向导数达到最大的方向, 即梯度方向来刻化影响, 即扰动的局部影响最大的方向 d_{max} 为 ω_0 处的梯度方向. 文 [6] 曾经用梯度方向来刻化扰动对于置信域体积的影响; 文 [10] 曾经用梯度方向来刻化扰动对于数据变换的影响. 本文则用梯度方向来研究扰动对于 Score 检验统计量的局部影响. 设在扰动 ω 下 (2.4) 式表示为 $SC(\omega), SC(\omega_0) = SC$ 表示无扰动情形. 根据以上分析, 扰动 ω 对于 $SC(\omega)$ 局部影响最大的方向应满足

$$d_{mas} \propto \left\{\frac{\partial SC(\omega)}{\partial \omega}\right\}_{(\omega_0, \widehat{\theta})}. \tag{3.1}$$

本文分别考虑因变量的扰动和自变量的扰动:

(a) 因变量的扰动, 即 $\quad Y \to Y + \omega, \quad \omega_0 = 0.$

(b) 自变量的扰动, 即 $\quad \eta(\beta) \to \eta(\beta|\omega)$. 例如, 若 (1.1) 为线性模型, 常见的扰动为 $\eta(\beta) \to \eta(\beta|\omega) = X(\omega)\beta, X(\omega)$ 表示 X 的某几行或某几列受到扰动.

设模型 (1.1) 有扰动 ω 时 (1.4) 式和 (1.5) 式分别记为 $L(\theta|\omega)$ 和 $S(\delta, \beta|\omega)$, 相应的 MLE 为 $\widehat{\theta}(\omega) = (\delta_0^T, \widehat{\sigma}^2(\omega))^T$. 当 $\omega = \omega_0$ 时, 则以上统计量还原到原来无扰动的情形. 本节的目的就是计算 $SC(\omega)$ 在 $(\omega_0, \widehat{\theta})$ 处的梯度方向. 以下讨论中, 有关的量均在 $(\omega_0, \widehat{\theta})$ 处计值, 大多略去. 根据定理 1, $SC(\omega)$ 可表示为

$$SC(\omega) = \frac{n^2}{2} S^{-2}(\delta_0, \widehat{\beta}(\omega)|\omega) u^T(\omega) P_{\overline{D}} u(\omega), \tag{3.2}$$

式中 $P_{\overline{D}}$ 与 ω 无关, 我们需要求出 $S(\delta_0, \widehat{\beta}(\omega)|\omega)$ 与 $u(\omega)$ 在 ω_0 处的导数. 为此先证明两个引理.

引理 2 对于扰动 (a) 和 (b), 在 $(\omega_0, \widehat{\theta})$ 分别有

$$(a) \qquad \frac{\partial S(\delta_0, \widehat{\beta}(\omega)|\omega)}{\partial \omega} = 2\widehat{e}, \tag{3.3}$$

$$(b) \qquad \frac{\partial S(\delta_0, \widehat{\beta}(\omega)|\omega)}{\partial \omega} = -2\Delta_1^T \widehat{e}, \tag{3.4}$$

其中 $\Delta_1 = \partial \eta(\beta|\omega)/\partial \omega^T$, 且在 $(\omega_0, \widehat{\theta})$ 处计值.

证明 易见 (1.6) 式对任意的 $\beta = \widehat{\beta}(\omega)$ 成立, 因此

$$\left[\frac{\partial S(\delta_0, \beta|\omega)}{\partial \beta}\right]_{\beta = \widehat{\beta}(\omega)} = 0, \tag{3.5}$$

$$\frac{\partial S(\delta_0, \widehat{\beta}(\omega)|\omega)}{\partial \omega} = \frac{\partial S(\delta_0, \widehat{\beta}|\omega)}{\partial \beta} \frac{\widehat{\beta}(\omega)}{\partial \omega^T} + \frac{\partial S(\delta_0|\widehat{\beta}|\omega)}{\partial \omega} = \frac{\partial S(\delta_0|\widehat{\beta}|\omega)}{\partial \omega}.$$

以上最后表示 $\widehat{\beta}$ 固定, 仅对 S 中最后一项 ω 求导数.

根据 (1.5) 式, 对扰动 (a) 有

$$S(\delta_0, \widehat{\beta}(\omega)|\omega) = \left\{Y + \omega - \eta(\widehat{\beta}(\omega))\right\}^T \left\{Y + \omega - \eta(\widehat{\beta}(\omega))\right\}, \tag{3.6}$$

$$\left[\frac{\partial S(\delta_0, \widehat{\beta}(\omega))}{\partial \omega}\right]_{\omega_0, \widehat{\theta}} = 2\widehat{e}.$$

对于扰动 (b) 有

$$S(\delta_0 \widehat{\beta}(\omega)|\omega) = \left\{Y - \eta(\widehat{\beta}(\omega)|\omega)\right\}^T \left\{Y - \eta(\widehat{\beta}(\omega)|\omega)\right\}, \tag{3.7}$$

$$\left[\frac{\partial S(\delta_0, \widehat{\beta}|\omega)}{\partial \omega}\right]_{(\omega_0, \widehat{\theta})} = -2\left[\frac{\partial \eta(\widehat{\beta}|\omega)}{\partial \omega^T}\right]^T_{(\omega_0, \widehat{\theta})} \widehat{e} = -2\Delta_1^T \widehat{e}. \quad \square$$

引理 3 对于扰动 (a) 和 (b), $\widehat{\beta}(\omega)$ 在 ω_0 处关于 ω 的导数可分别表示为

$$(a) \qquad \left[\frac{\partial \widehat{\beta}(\omega)}{\partial \omega^T}\right]_{\omega_0} = K(I - B)^{-1} Q^T, \tag{3.8}$$

$$(b) \qquad \left[\frac{\partial \widehat{\beta}(\omega)}{\partial \omega^T}\right]_{\omega_0} = K(I - B)^{-1} K^T \left\{[\widehat{e}^T][\Delta_2] - V^T \Delta_1\right\}, \tag{3.9}$$

其中 $B = [\widehat{e}^T][K^T W K]$ (见 [11]), $\Delta_2 = \partial^2 \eta(\beta|\omega)/\partial\beta\partial\omega^T$. 以上所有矩阵或立体阵均在 $(\omega_0, \widehat{\theta})$处计值.

证明 对于扰动 (a), 根据 (3.5) 和 (3.6) 式, $\widehat{\beta}(\omega)$ 应满足

$$\left\{\frac{\partial \eta(\beta)}{\partial \beta^T}\right\}^T \{Y + \omega - \eta(\beta)\} = 0, \qquad \beta = \widehat{\beta}(\omega).$$

该式对 ω 求导, 并在 $\omega_0 = 0$ 处计值可得

$$[\widehat{e}^T]\left[\frac{\partial^2 \eta}{\partial\beta\partial\beta^T}\frac{\partial \widehat{\beta}(\omega)}{\partial \omega^T}\right] + \left(\frac{\partial \eta}{\partial \beta^T}\right)^T \left\{I - \frac{\partial \eta}{\partial \beta^T}\frac{\partial \widehat{\beta}(\omega)}{\partial \omega^T}\right\} = 0,$$

$$\left([\widehat{e}^T][W] - V^T V\right)\left[\frac{\partial \widehat{\beta}(\omega)}{\partial \omega^T}\right]_{\omega_0} + V^T = 0.$$

把关系式 $V = QR, V^T V = R^T R, R^{-1} = K, B = [\widehat{e}^T][K^T W K]$ 代人上式, 即可得到 (3.8) 式.

对于扰动 (b), 由 (3.5) 和 (3.7) 式可知, $\widehat{\beta}(\omega)$ 满足

$$\left\{\frac{\partial \eta(\beta|\omega)}{\partial \beta^T}\right\}^T \{Y - \eta(\beta|\omega)\} = 0, \qquad \beta = \widehat{\beta}(\omega).$$

该式对 ω 求导, 并在 ω_0 处取值可得

$$[\widehat{e}^T]\left[\frac{\partial^2 \eta(\beta|\omega)}{\partial\beta\partial\beta^T}\frac{\partial \widehat{\beta}(\omega)}{\partial \omega^T} + \frac{\partial^2 \eta(\beta|\omega)}{\partial\beta\partial\omega^T}\right]_{\omega_0} + V^T\left[-\frac{\partial \eta(\beta|\omega)}{\partial \beta^T}\frac{\partial \widehat{\beta}(\omega)}{\partial \omega^T} - \frac{\partial \eta(\beta|\omega)}{\partial \omega^T}\right]_{\omega_0} = 0.$$

根据本文所定义的符号, 上式可化为

$$\left(V^T V - [\widehat{e}^T][W]\right)\left[\frac{\partial \widehat{\beta}(\omega)}{\partial \omega^T}\right]_{\omega_0} = [\widehat{e}^T][\Delta_2] - V^T \Delta_1.$$

由于 $V^T V - [\widehat{e}^T][W] = R^T(I - B)R$, 该式代人上式, 即可得到 (3.9) 式. $\square$

根据以上引理可得本节主要定理如下

定理 3 对于模型 (1.1) 和 (1.2), 假设检验问题 (1.3) 的 Score 统计量在扰动 (a), (b) 下, 其局部影响最大的方向 d_{max} 应满足

$$(a) \qquad d_{max} \propto a = \alpha_1 - \alpha_2 - \alpha_3, \tag{3.10}$$

$$(b) \qquad d_{max} \propto b = \Delta_1^T a + \left([\widehat{\epsilon}^T][\Delta_2]\right)^T (V^T V)^{-1} V^T \alpha_2, \tag{3.11}$$

其中 $\alpha_1 = \mathrm{diag}(\widehat{e}) P_{\overline{D}} u, \alpha_2 = Q(I-B)^{-1} Q^T \alpha_1, \alpha_3 = (\widehat{e}\widehat{e}^T / n\widehat{\sigma}^2)\alpha_1$.

证明 记 $S = S(\delta_0, \widehat{\beta}(\omega)|\omega)$, 由 (3.2) 式可知

$$n^{-2}\frac{\partial SC(\omega)}{\partial \omega} = -S^{-3}\frac{\partial S}{\partial \omega} u^T P_{\overline{D}} u + S^{-2}\left[\frac{\partial u(\omega)}{\partial \omega^T}\right]_{\omega_0}^T P_{\overline{D}} u(\omega). \tag{3.12}$$

对于扰动 (a), 由 (3.1), (3.3) 式可知

$$d_{max} \propto -2S^{-1}\widehat{e} u^T P_{\overline{D}} u + \left[\frac{\partial u(\omega)}{\partial \omega^T}\right]_{\omega_0}^T P_{\overline{D}} u. \tag{3.13}$$

由于 $u = \mathrm{diag}(\widehat{e})\widehat{e}$, 因此上式第一项可表示为 $-2S^{-1}\widehat{e} u^T P_{\overline{D}} u = -2(\widehat{e}\widehat{e}^T / n\widehat{\sigma}^2)\alpha_1$. 今计算 $[\partial u(\omega)/\partial \omega^T]_{\omega_0}$, 由于 $u_i(\omega) = \widehat{e}_i^2(\omega) = \left\{y_i + \omega_i - \eta_i(\widehat{\beta}(\omega))\right\}^2$, 因此在 ω_0 处有

$$\left[\frac{\partial u_i(\omega)}{\partial \omega_k}\right]_{\omega_0} = 2\widehat{e}_i\left[\delta_{ik} - \sum_{a=1}^{p}\frac{\partial \eta_i}{\partial \beta_a}\frac{\partial \widehat{\beta}_a(\omega)}{\partial \omega_k}\right]_{\omega_0},$$

其中 $\delta_{ik} = 1(i=k), \delta_{ik} = 0(i \neq k)$, β_a 为 β 的第 a 个分量. 因此由引理 3 的 (3.8) 式有

$$\left[\frac{\partial u(\omega)}{\partial \omega^T}\right]_{\omega_0} = 2\mathrm{diag}(\widehat{e})\left\{I - VK(I-B)^{-1}Q^T\right\}.$$

以上各项结果代入 (3.13) 式即得 $d_{max} \propto a$.

对于扰动 (b), 由 (3.4), (3.12) 式可得

$$d_{max} \propto 2S^{-1}\Delta_1^T \widehat{e} u^T P_{\overline{D}} u + \left[\frac{\partial u(\omega)}{\partial \omega^T}\right]_{\omega_0}^T P_{\overline{D}} u. \tag{3.14}$$

上式第一项可表示为 $2S^{-1}\Delta_1^T \widehat{e} u^T P_{\overline{D}} u = 2\Delta_1^T(\widehat{e}\widehat{e}^T / n\widehat{\sigma}^2)\alpha_1$, 今计算 $[\partial u(\omega)/\partial \omega^T](\omega_0)$, 由于 $u_i(\omega) = \left\{y_i - \eta_i(\widehat{\beta}(\omega)|\omega)\right\}^2$, 因此在 ω_0 处有

$$\left[\frac{\partial u_i(\omega)}{\partial \omega^T}\right]_{\omega_0} = -2\widehat{e}_i\left[\frac{\partial \eta_i(\widehat{\beta}|\omega)}{\partial \omega_k} + \sum_{a=1}^{p}\frac{\partial \eta_i}{\partial \beta_a}\frac{\partial \widehat{\beta}_a(\omega)}{\partial \omega_k}\right]_{\omega_0},$$

$$\begin{aligned}\left[\frac{\partial u}{\partial \omega^T}\right]_{\omega_0} &= -2\mathrm{diag}(\widehat{e})\left[\Delta_1 + V\frac{\partial \widehat{\beta}}{\partial \omega^T}\right]_{\omega_0} \\ &= -2\mathrm{diag}(\widehat{e})\left\{\Delta_1 + Q(I-B)^{-1}K^T([\widehat{e}^T][\Delta_2] - V^T\Delta_1)\right\}.\end{aligned}$$

以上最后式用到了 (3.9) 式. 以上各式代入 (3.14) 式即得到 (3.11)式. □

在定理 3 中, 矩阵 B 称为有效残差曲率阵 (见 [9], [11]), 是非线性回归中十分重要的统计量它亦可表示为 $B = [\widehat{\epsilon}^T N][A^I]$. 它仅与模型的固有曲率有关. 定理说明了, 对 $SC(\omega)$ 局部影响最大的方向与残差以及模型的固有曲率有关. 这显然是合理的.

若 (1.1) 为线性模型, 设 $\eta(\beta) = X\beta, \eta(\beta|\omega) = X(\omega)\beta$, $X(\omega)$ 可按列扰动与按行扰动分为: (b'): $X(\omega) = (X_1 + \omega, X_2, \ldots, X_p)$, (b''): $X^T(\omega) = (x_1 + \omega, x_2, \ldots, x_n)$; 其中 $X_1, \ldots, X_p$ 和 $x_1, \ldots x_n$ 分别为 X 的列向量和行向量.

推论 若 (1.1) 为线性模型, 则对于扰动 $(a),(b'),(b'')$ 分别有

$$(a) \qquad d_{max} \propto a, \qquad a = (I-H)\alpha_1 - \alpha_3, \tag{3.15}$$

$$(b') \qquad d_{max} \propto b', \qquad b' = \widehat{\beta}_1 a + \xi_1 \widehat{\epsilon}, \tag{3.16}$$

$$(b'') \qquad d_{max} \propto b'', \qquad b'' = a_1 \widehat{\beta} + \widehat{e}_1 \xi, \tag{3.17}$$

其中 $\xi = (X^TX)^{-1}X^T\alpha_1; \widehat{\beta}_1, a_1, \widehat{e}_1$ 分别为 $\widehat{\beta}, a, \widehat{e}$ 的第一分量, $H = X(X^TX)^{-1}$.

证明 若 (1.1) 为线性模型, 则 $B=0, V=X, Q^TQ=H$. 因此由 (3.10) 式可直接得到 (3.15) 式. 今证 (3.16) 式. 对于扰动 $(b'), \eta(\beta|\omega) = X\beta + \beta_1\omega, \Delta_1 = \widehat{\beta}_1 I$. 为求 Δ_2, 今记 $X=(X_{ia})$, 则有 $\partial\eta_i(\beta|\omega)/\partial\beta_a = X_{ia} + \delta_{1a}\omega_i, \partial^2\eta_i(\beta|\omega)/\partial\beta_a\partial\omega_k = (\Delta_2)_{iak} = \delta_{1a}\delta_{ik}, [\widehat{e}^T][\Delta_2] = \sum_i \widehat{e}_i E_{1i} = d_1\widehat{e}^T$ (其中 E_{1i} 为 (1,i) 处为 1 其它为 0 的 $p\times n$ 阶矩阵, d_1 为第一元素为 1 其它元素为 0 的 p 维向量). 以上有关 Δ_1, Δ_2 的公式代入 (3.11) 式可得 $b' = \widehat{\beta}_1 a + \widehat{e}d_1^T(X^TX)^{-1}X^TH\alpha_1 = \widehat{\beta}_1 + \xi_1\widehat{e}$. 此即 (3.16) 式. 类似地, 可得到 (3.17) 式. □

例 鱼场数据. 该数据的详细讨论可参见文献 [2]. 该文献指出, 这组数据可用以下非线性模型进行似合:

$$X_i = y_i^{0.3} = [\beta_1 x_i exp(-\beta_i x_i)]^{0.3} + \varepsilon_i, \quad i = 1, \ldots, 28.$$

根据文献 [2] 的详细分析, 4, 12, 16 等数据点都是影响比较强的点, 而 12 号点的影响最强. 为检验方差的影响, 今假定 $\mathrm{Var}(\varepsilon_i) = \sigma^2 m_i, m_i = x_i^\delta, \delta$ 为一维参数, δ_0 时 $m_i = 1$. 经直接计算可知, $SC = 2.6434$, 接近临界值. 今根据公式 (3.10), (3.11) 计算异方差检验统计量 $SC(\omega)$ 对 ω 的局部影响. 考虑两种扰动: $(a) \quad y_i \to y_i + \omega_i, (b) \quad x_i \to x_i + \omega_i$ (这时都有 $\omega_0 = 0$). 计算结果见表 1, 其中 x_i, y_i 为原始数据, a_i, b_i 分别为 (3.10) 和 (3.11) 的相应分量. 该表说明, 第 12 号和第 16 号点都是强影响点. 这些结果与文 [2] 的分析一致. (i, a_i) 和 (i, b_i) 的散点图见图 1 和图 2.

史建清同志协助进行了很多计算, 作者深表感谢.

图1. 因变量扰动对 Score 统计量的局部影响

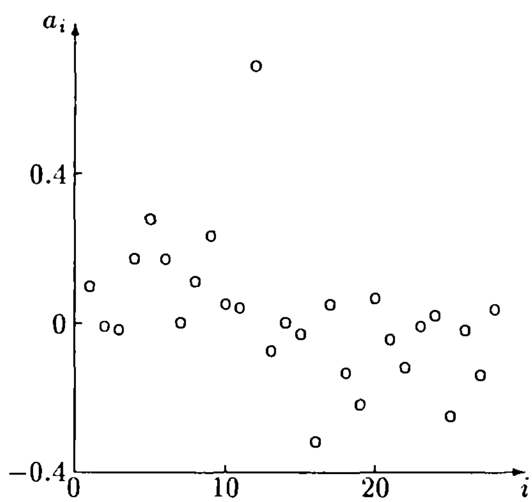
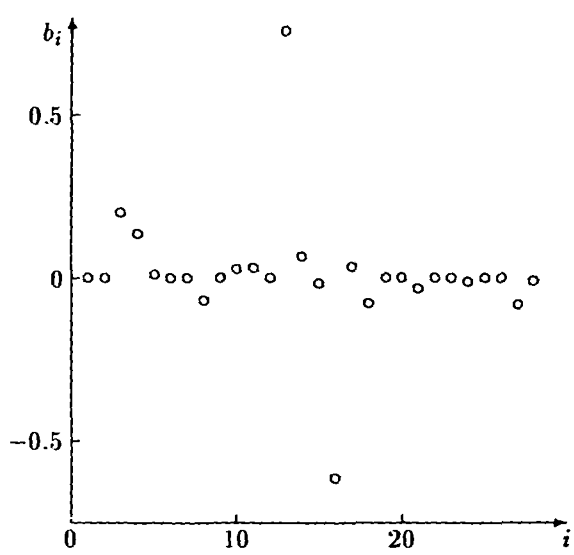


表1　　鱼场数据 *Score* 统计量的局部分析

No.	y_i	x_i	a_i	b_i
1	2215	963	0.1943	0.0053
2	1334	572	−0.0115	−0.0056
3	800	305	−0.0373	−0.0241
4	438	272	0.1728	0.1353
5	3071	824	0.2767	0.0091
6	957	940	−0.1704	0.0025
7	934	486	−0.0068	0.0038
8	971	307	−0.01114	−0.0762
9	2257	1066	0.2322	0.0024
10	1451	480	−0.0502	−0.0281
11	686	393	0.0452	0.0319
12	127	176	0.6867	0.7585
13	700	237	−0.0776	−0.0658
14	1381	700	−0.0002	0.0008
15	1393	511	−0.0301	−0.0168
16	363	87	−0.3181	−0.6145
17	668	370	0.0516	0.0359
18	2067	448	−0.01339	−0.0794
19	644	819	−0.02210	−0.0021
20	1747	799	0.0641	0.0041
21	744	273	−0.1205	−0.0388
22	1087	936	−0.1205	0.0028
23	1335	558	−0.0141	−0.0071
24	1981	597	−0.0192	−0.0120
25	627	848	−0.2500	−0.0019
26	1099	619	−0.0236	0.0006
27	1532	397	−0.1402	−0.0820
28	2086	616	0.0363	−0.0094

图2.　自变量扰动对Score统计量的局部影响



参 考 文 献

[1] Seber, G. A. F., *Linear Regression Analysis*. Wiley, 1977.
[2] Carroll, R. J. and Ruppert, D., *Transformation and Weighting in Regression*. Chapmam and Hall, London, 1988.
[3] Cook, D. R. and Weisberg, S., Diagnostic for heteroscedasticity in regression. Biomertrika **70**(1983), 1-10.
[4] Cox, D. R. and Reid , N., Parameter ortrhogonality and approximate conditional inference. J. R. Statist. Soc. B **49**(1987), 1-39.
[5] Cook, R. D., Assessment of local influence. J. R. Statist. Soc. B **48**(1986), 133-169.
[6] Thomas, W., Influence regions for regression coefficients in generalized linear models. JASA **85**(1990), 393-397.
[7] Cox, D. R. and Hinkley, D. V., *Theorctical Statistics*. Chapman and Hall, London, 1974.
[8] 韦博成、鲁国斌、史建清, 统计诊断引论, 东南大学出版社, 1991.
[9] 韦博成, 近代非线性回归分析, 东南大学出版社, 1989.
[10] Lawrance, A. J., Regression transformation diagnostics using local influence. JASA **83**(1988), 1067-1072.
[11] Hamilton, D. C., Watts, D. C. and Bates, D. M., Accounting for intrinsic nonlinearity in nonlinear regression paramater inference regions. Ann. Statist. **10**(1982), 386-393.

Score Statistic and Its Local Influence Analysis in Weighted Nonlinear Regression

WEI BOCHENG

(*Southeast University, Nanjing*)

The assumption of homoscedasticity is a commonly concerned problem in regression analysis. Cook and Weisberg (1983) provided a diagnostic test for heteroscedasticity based on the score statistic in linear regression. In this paper, we extend their result to nonlinear regression . We also obtain a modified score statistic based on the method of Cox and Reid (1987) in which the orthogonality between parameters of interest and nuisance parameters is used. Furthermore,we study the local influence of small perturbations on the score statistic and corresponding diagnostic measures are presented. A numerical example is given to illustrate our results.