线性回归模型的深度加权最小二乘估计和拟合检验

范允征[1]，林　路[2]

(1. 南通大学理学院, 江苏 南通 226007)
(2 山东大学数学与系统科学学院, 山东 济南 250100)

[摘要]　在线性回归模型中普通的最小二乘估计(LSE)许多情形下是不稳健的. 本文介绍了一种投影深度函数, 深度加权平均和深度加权 LSE, 这些估计量有符合需要的稳健性. 并讨论了在深度加权 LSE 情形下线性回归模型的拟合检验问题.

[关键词]　稳健性, 统计深度, 深度加权, LSE
[中图分类号] O212. 7　[文献标识码] A　[文章编号] 1001-4616(2008)03-0039-05

Depth-Weighted LSE for Linear Regression Model and the Fit-Test

Fan Yunzheng[1], Lin Lu[2]

(1. School of Science, Nantong University, Nantong 226007, China)
(2. School of Mathematics and System Science, Shandong University, Jinan 250100, China)

Abstract The ordinary Least Squares Estimate (LSE) in linear regression model is not robust for many cases. A class of projection-based depth functions, depth-weighted mean and depth-weighted LSE were studied. These estimates had desirable robustness. The problem of testing the fit of linear regression model under the depth-weighted LSE were also discussed.

Key words robustness, statistical depth, depth-weighted, LSE

为研究变量 Y如何随变量 X 变化, X, Y间的关系可由下式表示

$$Y_i = \theta_0 + \theta_1 x_i + \varepsilon_i, \quad i = 1, \cdots, n, \tag{1}$$

其中 ε_i是误差项, 且 $\varepsilon_1, \cdots, \varepsilon_n$互相独立.

令 $\boldsymbol{Y} = (Y_1, \cdots, Y_n)^{\mathrm{T}}$, $\boldsymbol{X} = \begin{pmatrix} 1 & \cdots & 1 \\ x_1 & \cdots & x_n \end{pmatrix} = (\boldsymbol{X}_1, \cdots, \boldsymbol{X}_n)$, $\theta = (\theta_0, \theta_1)^{\mathrm{T}}$, $\varepsilon = (\varepsilon_1, \cdots, \varepsilon_n)^{\mathrm{T}}$.

得(1)的矩阵形式

$$\boldsymbol{Y} = \boldsymbol{X}^{\mathrm{T}}\theta + \varepsilon \tag{2}$$

进一步假设 $\boldsymbol{Y} \in R$, $\boldsymbol{X} \in \aleph \subset R^p$, $\theta \in \Theta \subset R^p$即 θ是未知参数, ε满足 $E(\varepsilon) = \mathbf{0}$

通常选择 $\boldsymbol{q}(\theta, \boldsymbol{X}, \boldsymbol{Y}) = \boldsymbol{X}(\boldsymbol{Y} - \boldsymbol{X}^{\mathrm{T}}\theta)$[1] 作为 θ的拟得分函数, 由方程 $E(\boldsymbol{q}(\theta, \boldsymbol{X}, \boldsymbol{Y})) = \mathbf{0}$, 便有: 如果 $((Y_1, \boldsymbol{X}_1^{\mathrm{T}}), \cdots, (Y_n, \boldsymbol{X}_n^{\mathrm{T}}))$是$(\boldsymbol{Y}, \boldsymbol{X}^{\mathrm{T}})$的独立同分布的样本, 则方程的经验表达式为

$$\frac{1}{n}\sum \boldsymbol{X}_i(Y_i - \boldsymbol{X}_i^{\mathrm{T}}\theta) = 0, \tag{3}$$

由此得参数估计式

$$\hat{\theta} = \left(\frac{1}{n}\sum_{i=1}^{n}\boldsymbol{X}_i\boldsymbol{X}_i^{\mathrm{T}}\right)^{-1}\frac{1}{n}\sum_{i=1}^{n}\boldsymbol{X}_i Y_i. \tag{4}$$

(2)的最小二乘估计(LSE)是$(\boldsymbol{X}\boldsymbol{X}^{\mathrm{T}})^{-1}\boldsymbol{X}\boldsymbol{Y}$, 这和(4)一致. 经验方程(3)和参数估计(4)都是样本矩构造的, 故而不稳健, 因为样本矩对数据 $(Y_1, \boldsymbol{X}_1), \cdots, (Y_n, \boldsymbol{X}_n^{\mathrm{T}})$的扰动比较敏感.

收稿日期: 2007-06-12
基金项目: 江苏省自然科学基金指导性计划(06KJD110051)资助项目.
通讯联系人: 范允征, 副教授, 研究方向: 稳健统计. E-mail: fan. yzh@ ntu. edu. cn

统计深度是构造稳健估计的一个有力工具. Tukey[2], Liu[3], Zuo和 Serfling[4,5]等人介绍了一些统计深度函数及关于如何构造统计深度函数的原理. Zuo, Cui和 He[6]定义了深度加权平均, 深度加权方差和它们的估计量, 这些估计量有符合需要的稳健性, 它们的崩溃值将近 1/2.

本文提出了用投影深度估计线性回归模型(2)中的回归系数 θ, 从而把原来的 LSE推广到深度加权的 LSE, 并讨论了在此情形下如何进行拟合检验的问题.

1　深度加权 *LSE*

已有一些统计深度被研究, 例如, 半空间深度[2], 简单深度[3]和投影深度[4,5]. 下面仅用投影深度构造估计方程, 此法很容易推广到其它情形.

基于分布 F 的观察值 $\boldsymbol{Z}$ 的投影深度由下式定义:

$$D(\boldsymbol{Z}, F) = \frac{1}{1+O(\boldsymbol{Z}, F)},\tag{5}$$

其中 $O(\boldsymbol{Z}, F) = \sup\limits_{\|u\|=1}|\boldsymbol{u}^{\mathrm{T}}\boldsymbol{z} - \mathrm{med}(F_u)|\dfrac{1}{\mathrm{mad}(F_u)}$, F_u 是 $\boldsymbol{U}^{\mathrm{T}}\boldsymbol{Z}$ 的分布函数, $\mathrm{med}(F_u)$ 是 F_u 的中位数, $\mathrm{mad}(F_u)$ 是 $|\boldsymbol{u}^{\mathrm{T}}\boldsymbol{z} - \mathrm{med}(F_u)|$ 的分布的中位数, 若 $\boldsymbol{u}^{\mathrm{T}}\boldsymbol{z} - \mathrm{med}(F_u) = \mathrm{mad}(F_u) = 0$, 则定义 $|\boldsymbol{u}^{\mathrm{T}}\boldsymbol{z} - \mathrm{med}(F_u)|\dfrac{1}{\mathrm{mad}(F_u)} = 0$.

依据[6]中提出的随机变量的深度加权平均的概念, 对给定的分布函数 F 和深度函数 $D(\boldsymbol{Z}, F)$, 一个函数 $g(z)$ 的深度加权平均由下式定义:

$$DE(g(z)) = \frac{\int g(z)\omega(D(\boldsymbol{Z}, F))\mathrm{d}F(z)}{\int \omega(D(\boldsymbol{Z}, F))\mathrm{d}F(z)},\tag{6}$$

这里 $\omega(\cdot)$ 是一个给定的深度函数, 一般选择如下:

$$\omega(r) = \begin{cases} \dfrac{\exp\left[-k\left(1-\dfrac{r^s}{m^s}\right)^2\right] - \exp(-k)}{1-\exp(-k)}, & \text{如果}\quad r < m \\ 1, & \text{如果}\quad r \geqslant m \end{cases}\tag{7}$$

其中 $m = \mathrm{med}(D(\boldsymbol{Z}, F))$, $s \geqslant 1$ 且 k 是一个适当的正常数, 它可调节我们想要区分的具有不同投影深度的点的多少. 当 $s = 1$, $\omega(r)$ 是由[6]定义的权函数, 一般地, 权函数 $\omega(r)$ 的选择应满足如下条件:

$\omega(r)$ 连续可微且对某正常数 M 和 $s \geqslant 1$, 有 $0 \leqslant \omega(r) \leqslant Mr^s$. 假定为估计 P 维参数 θ 已定出一个合适的或最佳无偏估计函数 $\boldsymbol{q}(\theta, \boldsymbol{Z})$, 它是一个 P 维的深度加权平均意义下的无偏估计量, 即

$$DE(\boldsymbol{q}(\theta, \boldsymbol{Z})) = 0,\tag{8}$$

且深度加权平均 $DE(\cdot)$ 由(6)定义, 如果 $(Z_1, \dots, Z_n)$ 是 $\boldsymbol{Z}$ 的一个独立同分布的样本, 则由(6), 此方程的经验表达式有如下形式:

$$q_n^{(DW)}(\theta) \equiv \frac{1}{nW_n}\sum_{i=1}^{n}\omega(D(Z_i, F_n))\boldsymbol{q}(\theta, Z_i) = 0,\tag{9}$$

其中, $W_n = \dfrac{1}{n}\sum\limits_{i=1}^{n}\omega(D(Z_i, F_n))$, F_n 是由 $Z_1, \dots, Z_n$ 决定的经验分布函数. 记上述方程的根为 $\theta^{(DW)}$.

考虑线性回归模型(2), 它的拟得分函数选为 $\boldsymbol{q}(\theta, \boldsymbol{X}, \boldsymbol{Y}) = \boldsymbol{X}(\boldsymbol{Y} - \boldsymbol{X}^{\mathrm{T}}\theta)$, 由(9)得到它的经验表达式

$$\boldsymbol{q}_n^{(DW)}(\theta) \equiv \frac{1}{nW_n}\sum_{i=1}^{n}\boldsymbol{X}_i\omega(D((\boldsymbol{X}_i, Y_i), F_n))(Y_i - \boldsymbol{X}_i^{\mathrm{T}}\theta) = 0,\tag{10}$$

其中 $W_n = \dfrac{1}{n}\sum\limits_{i=1}^{n}\omega(D((\boldsymbol{X}_i, Y_i), F_n))$, 则方程(10)的根是

$$\theta^{(DW)} = \left(\frac{1}{nW_n}\sum_{i=1}^{n}\boldsymbol{X}_i\boldsymbol{X}_i^{\mathrm{T}}\omega(D((\boldsymbol{X}_i, Y_i), F_n))\right)^{-1}\cdot\frac{1}{nW_n}\sum_{i=1}^{n}\boldsymbol{X}_iY_i\omega(D((\boldsymbol{X}_i, Y_i), F_n)).\tag{11}$$

若 $\omega_i \triangleq \omega(D((\boldsymbol{X}_i, Y_i), F_n))$, 则 $W_n = \frac{1}{n}\sum_{i=1}^{n}\omega_i$, 且

$$\hat{\theta}^{(DW)} = \left(\frac{1}{nW_n}\sum_{i=1}^{n}\boldsymbol{X}_i\boldsymbol{X}_i^{\mathrm{T}}\omega_i\right)^{-1} \cdot \frac{1}{nW_n}\sum_{i=1}^{n}\boldsymbol{X}_iY_i\omega_i. \tag{12}$$

例如在(2)中, 深度加权的 LSE 是

$$\hat{\theta}^{(DW)} = \begin{pmatrix}\hat{\theta}_0^{(DW)}\\ \hat{\theta}_1^{(DW)}\end{pmatrix} = \left(\frac{1}{nW_n}\sum_{i=1}^{n}\boldsymbol{X}_i\boldsymbol{X}_i^{\mathrm{T}}\omega_i\right)^{-1} \cdot \frac{1}{nW_n}\sum_{i=1}^{n}\boldsymbol{X}_iY_i\omega_i = \left(\frac{1}{nW_n}\sum_{i=1}^{n}\begin{pmatrix}1 & x_i\\ x_i & x_i^2\end{pmatrix}\omega_i\right)^{-1} \cdot \frac{1}{nW_n}\sum_{i=1}^{n}\begin{pmatrix}Y_i\\ x_iY_i\end{pmatrix}\omega_i =$$

$$\frac{nW_n}{nW_n\sum x_i^2\omega_i - \left(\sum x_i\omega_i\right)^2}\begin{pmatrix}\sum x_i^2\omega_i & -\sum x_i\omega_i\\ -\sum x_i\omega_i & nW_n\end{pmatrix}\begin{pmatrix}\bar{Y}\\ \frac{1}{nW_n}\sum x_iY_i\omega_i\end{pmatrix} =$$

$$\frac{1}{\sum x_i^2\omega_i - nW_n\bar{X}^2}\begin{pmatrix}\bar{Y}\sum x_i^2\omega_i - \bar{X}\sum x_iY_i\omega_i\\ \sum(x_i-\bar{X})(Y_i-\bar{Y})\omega_i\end{pmatrix},$$

其中 $\bar{X} = \frac{1}{nW_n}\sum x_i\omega_i$, $\bar{Y} = \frac{1}{nW_n}\sum Y_i\omega_i$, 显然 $\sum(x_i-\bar{X})\omega_i = 0$, $\sum(Y_i-\bar{Y})\omega_i = 0$, 因此

$$\hat{\theta}_1^{(DW)} = \frac{\sum(x_i-\bar{X})(Y_i-\bar{Y})\omega_i}{\sum(x_i-\bar{X})^2\omega_i} \triangleq \frac{L_{XY}}{L_{XX}},$$

$$\hat{\theta}_0^{(DW)} = \frac{1}{L_{XX}}\left[\bar{Y}\sum(x_i-\bar{X})x_i\omega_i + \bar{Y}\bar{X}\sum x_i\omega_i - \bar{X}\sum(x_i-\bar{X})Y_i\omega_i - \bar{X}^2\sum Y_i\omega_i\right] = \bar{Y} - \frac{L_{XY}}{L_{XX}}\bar{X} = \bar{Y} - \hat{\theta}_1^{(DW)}\bar{X}.$$

上面的式子表明, 如果在原来的 LSE 框架下的样本矩用深度加权样本矩代替, 得到的估计量正是深度加权估计量 $\hat{\theta}^{(DW)}$. 由于深度加权样本矩有符合需要的稳健性[6], 可以期望, 深度加权估计量 $\hat{\theta}^{(DW)}$ 也有符合需要的稳健性. 事实上, [7] 中的定理 2 表明, $\hat{\theta}^{(DW)}$ 有非常高的崩溃值接近 1/2, 远大于 1/3, [8] 中给出了回归深度估计的近似崩溃值.

2 线性回归模型的拟合检验

现在将文章开始部分的模型(1)推广到一般的具有 P 元回归参数的情形:

$$Y_i = \theta_0 + \theta_1x_{1i} + \theta_2x_{2i} + \cdots + \theta_{p-1}x_{(p-1)i} + \varepsilon_i,\ i = 1, \cdots, n. \tag{13}$$

它的矩阵形式与(2)相同, 即 $\boldsymbol{Y} = \boldsymbol{X}^{\mathrm{T}}\theta + \varepsilon$, 其中 $E(\varepsilon) = \boldsymbol{0}$, 且

$$\mathrm{cov}(\varepsilon) = \sigma^2\boldsymbol{I}, \tag{14}$$

这里 $\boldsymbol{X} = \begin{pmatrix}1 & 1 & \cdots & 1\\ x_{11} & x_{12} & \cdots & x_{1n}\\ \vdots & \vdots & \vdots & \vdots\\ x_{(p-1)1} & x_{(p-1)2} & \vdots & x_{(p-1)n}\end{pmatrix}$, $\theta = \begin{pmatrix}\theta_0\\ \theta_1\\ \vdots\\ \theta_{p-1}\end{pmatrix}$, $\boldsymbol{Y} = \begin{bmatrix}Y_1\\ Y_2\\ \vdots\\ Y_n\end{bmatrix}$.

更一般地, (13) 可表为 $Y_i = r(x_{1i}, x_{2i}, \cdots, x_{(p-1)i}) + \varepsilon_i$, 其中 $E(\varepsilon_i) = 0$ 且

$$\mathrm{var}(\varepsilon_i) = \sigma^2 < \infty,\ i = 1, \cdots, n. \tag{15}$$

下面将要检验假设 H_0:

$$r(x_1, x_2, \cdots, x_{p-1}) = \theta_0 + \theta_1x_1 + \theta_2x_2 + \cdots + \theta_{p-1}x_{p-1}. \tag{16}$$

假定 $\varepsilon_1, \cdots, \varepsilon_n$ 是独立同分布的, 将讨论处理非高斯数据的方法: 大样本检验, 且研究基于如下统计量 V_n 的拟合检验.

假定(14)中的 $\boldsymbol{X}$ 是行满秩的, 此条件确保系数 $\theta_0, \cdots, \theta_{p-1}$ 的最小二乘估计的惟一性. 回归模型(14)的深度加权 LSE 用 $\hat{\theta}_0^{(DW)}, \hat{\theta}_1^{(DW)}, \cdots, \hat{\theta}_{p-1}^{(DW)}$ 表示, 由(12)有

$$\hat{\theta}^{(DW)} = (\hat{\theta}_0^{(DW)}, \hat{\theta}_1^{(DW)}, \cdots, \hat{\theta}_{p-1}^{(DW)})^{\mathrm{T}} = \left(\frac{1}{nW_n}\sum_{i=1}^{n}\boldsymbol{X}_i\boldsymbol{X}_i^{\mathrm{T}}\omega_i\right)^{-1} \cdot \frac{1}{nW_n}\sum_{i=1}^{n}\boldsymbol{X}_iY_i\omega_i,$$

它的矩阵形式是:

$$\hat{\theta}^{(DW)} = (X\omega X^{\mathrm{T}})X\omega Y, \tag{17}$$

其中, $\omega = \begin{pmatrix} \omega_1 & & \\ & \ddots & \\ & & \omega_n \end{pmatrix}$, 且 $\omega_i \triangleq \omega(D((X_i, Y_i), F_n))$.

定义残差向量 $\boldsymbol{e}$的第 i个分量 $e_i = Y_i - \hat{\theta}_0^{(DW)} - \hat{\theta}_1^{(DW)}x_{1i} - \cdots - \hat{\theta}_{p-1}^{(DW)}x_{(p-1)i}$, $i = 1, \cdots, n$.

显然 $\boldsymbol{e} = \boldsymbol{Y} - \boldsymbol{X}^{\mathrm{T}}\hat{\theta}^{(DW)}$, 由(17)有: $\boldsymbol{e} = \boldsymbol{Y} - \boldsymbol{X}^{\mathrm{T}}(\boldsymbol{X}\omega\boldsymbol{X}^{\mathrm{T}})^{-1}\boldsymbol{X}\omega\boldsymbol{Y} = (\boldsymbol{I} - \boldsymbol{F})\boldsymbol{Y}$, 其中 $\boldsymbol{F} = \boldsymbol{X}^{\mathrm{T}}(\boldsymbol{X}\omega\boldsymbol{X}^{\mathrm{T}})^{-1}\boldsymbol{X}\omega$, $\boldsymbol{I}$是一个 $n\times n$单位阵, 显然 $\boldsymbol{F}\boldsymbol{X}^{\mathrm{T}} = \boldsymbol{X}^{\mathrm{T}}$, $\boldsymbol{X}\omega\boldsymbol{F} = \boldsymbol{X}\omega$. 一个方差的基本估计是 $\sigma_m^2 = \dfrac{1}{n-p}\sum_{i=1}^{n}e_i^2$.

现在希望用一方差的估计量来推断线性模型(16)是否成立. 考虑 $\hat{\sigma}^2 = \dfrac{1}{a_n}\sum_{i=2}^{n}(e_i - e_{i-1})^2 = \dfrac{1}{a_n}\boldsymbol{e}^{\mathrm{T}}\boldsymbol{H}\boldsymbol{e}$, 其中 $a_n = 2(n-1) - \mathrm{tr}(\boldsymbol{H}(\boldsymbol{F} + \boldsymbol{F}^{\mathrm{T}} - \boldsymbol{F}\boldsymbol{F}^{\mathrm{T}}))$ 且 $\boldsymbol{H}$ 是一个 $n\times n$ 对角阵

$$\boldsymbol{H} = \begin{pmatrix} 1 & -1 & 0 & 0 & 0 & \cdots & 0 & 0 & 0 \\ -1 & 2 & -1 & 0 & 0 & \cdots & 0 & 0 & 0 \\ 0 & -1 & 2 & -1 & 0 & \cdots & 0 & 0 & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \ddots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & 0 & 0 & \cdots & -1 & 2 & -1 \\ 0 & 0 & 0 & 0 & 0 & \cdots & 0 & -1 & 1 \end{pmatrix}.$$

当线性模型成立时, 此方差的估计量是 σ^2的无偏估计, 这是因为,

$$\begin{aligned} E(\boldsymbol{e}^{\mathrm{T}}\boldsymbol{H}\boldsymbol{e}) = E[\boldsymbol{Y}^{\mathrm{T}}(\boldsymbol{I} - \boldsymbol{F}^{\mathrm{T}})\boldsymbol{H}(\boldsymbol{I} - \boldsymbol{F})\boldsymbol{Y}] \triangleq E(\boldsymbol{Y}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{Y}) = \\ \mathrm{tr}(\boldsymbol{A}\sigma^2\boldsymbol{I}) + (\boldsymbol{X}^{\mathrm{T}}\theta)^{\mathrm{T}}\boldsymbol{A}(\boldsymbol{X}^{\mathrm{T}}\theta) = \\ \sigma^2[\mathrm{tr}(\boldsymbol{H}(\boldsymbol{I} - \boldsymbol{F})(\boldsymbol{I} - \boldsymbol{F}^{\mathrm{T}}))] + 0\ (\text{因为 } \boldsymbol{F}\boldsymbol{X}^{\mathrm{T}} = \boldsymbol{X}^{\mathrm{T}}) = \\ \sigma^2[2(n-1) - \mathrm{tr}(\boldsymbol{H}(\boldsymbol{F} + \boldsymbol{F}^{\mathrm{T}} - \boldsymbol{F}\boldsymbol{F}^{\mathrm{T}}))] = a_n\sigma^2, \end{aligned}$$

其中 $\boldsymbol{A} = (\boldsymbol{I} - \boldsymbol{F}^{\mathrm{T}})\boldsymbol{H}(\boldsymbol{I} - \boldsymbol{F})$.

现在用方差比 $V_n = \dfrac{\sigma_m^2}{\hat{\sigma}^2}$作为统计量, 下面的定理提出了几个条件, 这些条件下, V_n有渐近正态分布[9].

定理　假设模型(15)成立, r有线性形式(16), 且 $\varepsilon_1, \cdots, \varepsilon_n$是互相独立的随机变量, 具有相同的方差 σ^2, 对所有的 i及正常数 δ成立 $E|\varepsilon_i|^{2+\delta} < M$, 如果 $\hat{\theta}_0, \cdots, \hat{\theta}_{p-1}$满足 $E(\hat{\theta}_j - \theta_j)^2 = O\left(\dfrac{1}{n}\right)$, $j = 0, 1, \cdots, p-1$, 且 $|x_j|$对每一个 j都是有界的, 则统计量 V_n是渐近正态的, 即 $\sqrt{n}(V_n - 1) \xrightarrow{D} N(0, 1)$ 当 $n \to \infty$.

证明　由于 $\dfrac{2(n-p)}{a_n}(V_n - 1) = 2\dfrac{n-p}{a_n}\cdot\dfrac{\sigma_m^2}{\hat{\sigma}^2} - 1 + \left(1 - 2\dfrac{n-p}{a_n}\right)$,

即

$$\frac{2(n-p)}{a_n}(V_n - 1) = \frac{2\sum_{i=1}^{n}e_i^2 - \sum_{i=2}^{n}(e_i - e_{i-1})^2}{a_n\hat{\sigma}^2} + \left(1 - 2\frac{n-p}{a_n}\right), \tag{18}$$

且

$$e_1 = Y_1 - \sum_{j=0}^{p-1}\hat{\theta}_j x_{ji} = \varepsilon_1 + \sum_{j=0}^{p-1}(\theta_j - \hat{\theta}_j)x_{ji} \leqslant \varepsilon_1 + c\sum_{j=0}^{p-1}(\theta_j - \hat{\theta}_j),\quad (c\text{是常数})$$

$$\begin{aligned} e_1^2 \leqslant \varepsilon_1^2 + c^2\Big[\sum_{j=0}^{p-1}(\theta_j - \hat{\theta}_j)\Big]^2 + 2\varepsilon_1 c\sum_{j=0}^{p-1}(\theta_j - \hat{\theta}_j) \leqslant \\ \varepsilon_1^2 + c^2(p-1)\sum_{j=0}^{p-1}(\theta_j - \hat{\theta}_j)^2 + 2\varepsilon_1 c\Big[(p-1)\sum_{j=0}^{p-1}(\theta_j - \hat{\theta}_j)^2\Big]^{\frac{1}{2}}, \end{aligned}$$

因此 $Ee_1^2 \leqslant \sigma^2 + O\left(\dfrac{1}{n}\right)$. 类似有 $E(e_n^2) \leqslant \sigma^2 + O\left(\dfrac{1}{n}\right)$, 从而 $e_1^2 + e_n^2 < O_p(1)$.

因此, (18)的右边第一项中的分子为 $2\sum_{i=2}^{n}e_ie_{i-1} + e_1^2 + e_n^2 = 2\sum_{i=2}^{n}e_ie_{i-1} + O_p(1)$.

而 $\sum_{i=2}^{n} e_i e_{i-1} = \sum_{i=2}^{n}\left(Y_i - \sum_{j=0}^{p-1}\hat{\theta}_j x_{ji}\right)\left(Y_{i-1} - \sum_{j=0}^{p-1}\hat{\theta}_j x_{j(i-1)}\right) =$

$$\sum_{i=2}^{n}\left[\varepsilon_i + \sum_{j=0}^{p-1}(\theta_j - \hat{\theta}_j)x_{ji}\right]\left[\varepsilon_{i-1} + \sum_{j=0}^{p-1}(\theta_j - \hat{\theta}_j)x_{j(i-1)}\right] =$$

$$\sum_{i=2}^{n}\varepsilon_i\varepsilon_{i-1} + \sum_{i=2}^{n}\varepsilon_{i-1}\left[\sum_{j=0}^{p-1}(\theta_j - \hat{\theta}_j)x_{ji}\right] + \sum_{i=2}^{n}\varepsilon_i\left[\sum_{j=0}^{p-1}(\theta_j - \hat{\theta}_j)x_{j(i-1)}\right] +$$

$$\sum_{i=2}^{n}\left[\sum_{j=0}^{p-1}(\theta_j - \hat{\theta}_j)x_{j(i-1)}\right]\left[\sum_{j=0}^{p-1}(\theta_j - \hat{\theta}_j)x_{ji}\right],$$

因此 $\sum_{i=2}^{n} e_i e_{i-1} = \sum_{i=2}^{n}\varepsilon_i\varepsilon_{i-1} + R_n$, 其中

$$R_n = R_{n1} + R_{n2} + R_{n3},$$

$$R_{n1} = \sum_{i=2}^{n}\varepsilon_{i-1}\left[\sum_{j=0}^{p-1}(\theta_j - \hat{\theta}_j)X_{ji}\right]\ (\text{注意 } X_{oi} = 1) = \sum_{j=0}^{p-1}(\theta_j - \hat{\theta}_j)\rho_j,$$

而 $\rho_j = \sum_{i=2}^{n}\varepsilon_{i-1}X_{ji},\quad j = 0, 1, \cdots, p-1.$

由于 $E\rho_j^2 = O(n)$, $E(\theta_j - \hat{\theta}_j)^2 = O(n^{-1})$, 而 p 是有限数, 于是有 $R_{n1} = O_p(1)$, 类似地, R_n 中的另外两项也有同样结果, 故有 $R_n = O_p(1)$. 结合前面的结论及 $1 - 2\frac{n-p}{a_n} = O(n^{-1})$, 因此有:

$$2\frac{n-p}{a_n}\sqrt{n}(V_n - 1) = \frac{1}{\sqrt{n}}\frac{\sum_{i=2}^{n}\varepsilon_i\varepsilon_{i-1}}{\sigma^2} + O_p\left(\frac{1}{\sqrt{n}}\right).$$

于是使用 Hoeffding 和 Robbins 在 [10] 中提出的具有 m - dependent 的随机变量的中心极限定理, 立即得本定理之结果.

由 [7] 中的引理 1, $\sqrt{n}(\hat{\theta}^{(DW)} - \theta) = O_p(1)$, 则 $\hat{\theta}^{(DW)}$ 满足本定理之条件, 于是 (16) 的水平为 α 的假设检验之零假设的拒绝式为 $V_n \geqslant 1 + \frac{Z_\alpha}{\sqrt{n}}$ 如果 V_n 的观察值是 u, 则有近似的 P - 值 $1 - \Phi[\sqrt{n}(u-1)]$.

[参考文献]

[1] Godambe V P. Estimating Functions[M]. Oxford: Clarendon Press, 1991.

[2] Tukey J W. Mathematics and picturing data[J]. Proc Intern Congr Math Vancouver, 1975, 2(2): 523-531.

[3] Liu R Y. On a notion of data depth based on random simplices[J]. Ann Statist, 1990, 18(2): 405-414.

[4] Zuo Y J, Serfling R. General notions of statistical depth function[J]. Ann Statist, 2000a, 28(2): 461-482.

[5] Zuo Y J, Serfling R. Structural properties and convergence results for contours of sample statistical depth functions[J]. Ann Statist, 2000b, 28(2): 483-499.

[6] Zuo Y J, Cui H J, He X M. On the stahel-Donoho estimator and depth-weighted means of multivariate data[J]. Ann Statist, 2004, 32(1): 167-188.

[7] Lin L, Chen M H. Robust estimating equation based on statistical depth[J]. Statistical Papers, 2006, 47(2): 263-278

[8] Rousseeuw P J, Hubert M. Regression depth[J]. JASA, 1999, 94: 389-433.

[9] Jeffrey D Hart. Non Parametric Smoothing and Lack-of-Fit Tests[M]. New York, Berlin, Heidelberg: Springer-Verlag, 1977.

[10] Hoeffding W, Robbins H. The central limit theorem for dependent random variables[J]. Duke Math J, 1948, 15(3): 773-780.

[责任编辑: 陆炳新]