doi:10.16055/j.issn.1672-058X.2015.0007.002

函数带误差的部分线性模型约束下的统计推断

李梦含, 夏小超
(重庆大学 数学与统计学院,重庆 401331)

摘　要:考虑函数带误差的部分线性模型,研究约束条件下参数分量的统计推断.首先提出参数的约束估计并证明其渐进正态性;同时基于广义似然比统计量提出一种检验过程,并证明了即使在一般条件下函数中误差的 Wilks 现象仍然存在;最后,通过数值模拟检验约束估计值的一致性和检验统计量的有效性.

关键词:部分线性模型;函数带误差;统计推断;约束估计;Wilks 现象

中图分类号:O212　　**文献标识码**:A　　**文章编号**:1672-058X(2015)07-0005-11

0　引　言

近年来,半参技术发展迅速并广泛应用到经济、金融、政治、生态等科技领域.一方面,参数模型常因设定错误引起较大偏差,而半参技术可以减少设定错误的风险从而避免所谓的“维数灾难”;另一方面,半参技术还拥有非参模型的灵活性.而在半参模型中,部分线性模型发展尤为迅速,为研究温度和用电量的关系,Engle 等率先提出了以下形式的这种模型[1]

$$\boldsymbol{Y} = \boldsymbol{X}^{\mathrm{T}}\boldsymbol{\beta} + g(\boldsymbol{T}) + \varepsilon \tag{1}$$

这里 T 代表向量或矩阵的转置,$Y\in\mathbf{R}^1$ 为响应变量,$X\in\mathbf{R}^p$ 和 $T\in\mathbf{R}^1$ 是协变量,$g(\cdot)$是定义在[0,1]上的未知函数,$\boldsymbol{\beta}=(\beta_1,\cdots,\beta_p)^{\mathrm{T}}$ 是未知参数变量.在协变量给定时,误差项 ε 独立且条件均值为零.但在实际应用中,可能由于测量工具或环境因素的影响,使得协变量的测量存在误差,例如血清胆固醇水平、尿钠氯化物水平和接触污染物程度往往受测量误差影响[2].

当协变量的测量存在误差时,模型(1)被称为协变量误差模型或 EV 模型.一般有 3 种 EV 模型:

(i)只有 $\boldsymbol{X}$ 存在测量误差,即 $\boldsymbol{W}=\boldsymbol{X}+\xi$;

(ii)只有 $\boldsymbol{T}$ 存在测量误差,即 $\boldsymbol{U}=\boldsymbol{T}+\eta$;

(iii)$\boldsymbol{X}$ 和 $\boldsymbol{T}$ 都存在测量误差,即 $\boldsymbol{W}=\boldsymbol{X}+\xi,\boldsymbol{U}=\boldsymbol{T}+\eta$.

致力于研究 EV 模型参数估计和统计推断的文献也很多.为处理情形(i),Liang 等利用常用的衰减参数校正(parametric correction for attenuation)研究参数估计和非参估计的性质,并证明了估计值的渐近正态性和一致性[3];Cui 和 Li 利用最近邻广义二乘法(nearest neighbor-generalized least square method)得到了参数估计值、模型误差的方差和平滑函数[4],Cui 考虑了反复测量观察值时的参数估计问题[5];赵和周利用最小二乘和拉格朗日乘子检验进行了统计推断[6];You 等检验了统计推断的 3 个方面:带宽选择技术、拟合优度的检验、基于非凹惩罚似然法的变量选择[7];这些文献都是针对点估计进行的,当然也有很多基于经验似然构造参数置信区间的文献[8-10].总的来说,非参误差问题比参数误差更难处理,更涉及了非参回归模型中的反卷

收稿日期:2015-01-12;**修回日期**:2015-02-14.

作者简介:李梦含(1991-),女,山东济宁人,硕士研究生,从事模型变量选择和参数估计研究.

积技术.为研究参数估计的性质,Liang 首先将该方法推广到函数带误差的部分线性模型中[11].为了处理情形(ii),Huang 则采用经验似然法构造了参数的置信区间[12];此外 Zhu 和 Cui 也构造了参数估计值和非参核估计[13].

此处重点研究参数包含辅助信息的情形(ii).在统计应用中,样本外得到的辅助信息可提高参数估计的有效性,正如 Rao 等在线性模型中所述,当参数的先验信息表示成线性约束时,约束最小二乘估计比普通最小二乘更有效[14].而当线性部分的协变量存在测量误差时,Wei 对变系数部分线性模型做了统计推断[15].

受 Wei 的启发,对情形(ii)在如下约束条件下作统计推断:

$$\boldsymbol{A\beta} = \boldsymbol{b} \tag{2}$$

$\boldsymbol{A}$ 是 $k\times p$ 的已知矩阵,$\boldsymbol{b}$ 是 $k\times 1$ 的已知常数向量,并假定 $\mathrm{rank}(\boldsymbol{A})=k<p$.

第 2 节提出参数分量的约束估计和其主要性质;第 3 节对约束条件的合理性进行检验;第 4 节是数值模拟;主要结论的假设和证明则在第 5 节给出.

1 约束估计值的构造及其性质

为完成各种证明,需假设一些条件成立.

令 x_{ij} 为 $\boldsymbol{X}_i$ 的第 j 个分量,$h_j(t)=E(x_{ij}|\boldsymbol{T}_i=t)$,$\zeta_{ij}=x_{ij}-h_j(\boldsymbol{T}_i)$,$\boldsymbol{\zeta}_i=(\zeta_{i1},\cdots,\zeta_{ip})^{\mathrm{T}}$,$1\leqslant i\leqslant n$,$1\leqslant j\leqslant p$.首先提出平滑和超平滑的定义.

定义 1[16] u 的误差分布被称为 α 阶平滑的,如果它的特征函数 $\varphi_u(\cdot)$ 满足 $t\to\infty$ 时,

$$d_0|t|^{-\alpha} \leqslant |\varphi_u(t)| \leqslant d_1|t|^{-\alpha}$$

其中 d_0,d_1,α 均为正数.

定义 2[16] u 的误差分布被称为 α 阶超平滑的,如果它的特征函数 $\varphi_u(\cdot)$ 满足 $t\to\infty$ 时,

$$d_0|t|^{-\alpha_0}\exp(-|t|^{\alpha}/\gamma) \leqslant |\varphi_u(t)| \leqslant d_1|t|^{-\alpha_1}\exp(-|t|^{\alpha}/\gamma)$$

这里 d_0,d_1 同,α,γ 均为正数,α_0 和 α_1 为常数.

然后指出如下假设条件:

(C1) $g(\cdot)$ 和 $h_j(\cdot)(1\leqslant j\leqslant p)$ 一阶 Lipschitz 连续;

(C2) 不可观测协变量 $\boldsymbol{T}$ 的边缘密度在区间[0,1]上从零到无穷有界,且有有界的 m 阶导数,m 是正整数;误差 u 的分布是平滑或超平滑的,且其特征函数 $\varphi_u(\cdot)$ 不为 0;

(C3) 核函数 $K(\cdot)$ 对称,且为 m 阶对称,即满足 $K(-t)=K(t)$,$\int_{-\infty}^{\infty}K(t)\mathrm{d}t=1$,$\int_{-\infty}^{\infty}t^mK(t)\mathrm{d}t\neq 0$,$\int_{-\infty}^{\infty}t^jK(t)\mathrm{d}t=0$,其中 $j=1,2,\cdots,m-1$.

(C4) 误差分布满足下列两个条件之一:

(i) 误差分布是 α 阶平滑的,取平滑参数 $h=dn^{-1/(2m+2\alpha+1)}$,其中 $d>0$,$2m>2\alpha+1$,并假定对于常数 $c\neq 0$,当 $t\to\infty$ 时,$t^{\alpha}\varphi_u(t)\to c$,$t^{\alpha+1}\varphi'_u(t)=O(1)$,且有

$$\int_{-\infty}^{\infty}|t|^{\alpha+1}\{\varphi_K(t)+\varphi'_K(t)\}dt<\infty,\int_{-\infty}^{\infty}|t^{\alpha+1}\varphi_K(t)|^2\mathrm{d}t<\infty$$

(ii) 误差分布是 α 阶超平滑的,$\boldsymbol{X}$ 和 $\boldsymbol{T}$ 相互独立,$\varphi_K(t)$ 在 $|t|\leqslant M_0$ 上有有界的支撑,并取 $h=c(\log n)^{-1/\alpha}$,其中 $c>M_0(2/\gamma)^{1/\alpha}$.

(C5) $\sup_{0\leqslant t\leqslant 1}E(\|\boldsymbol{X}_1\|^3|\boldsymbol{T}=t)<\infty$,且 $\boldsymbol{B}\hat{=}E(\boldsymbol{\zeta}_1\boldsymbol{\zeta}_1^{\mathrm{T}})$ 为正定阵;

(C6) $E(\varepsilon_i)=0$,$E(u_i)=0(1\leqslant i\leqslant n)$ 且 $\sup_{1\leqslant i\leqslant n}E(|\varepsilon_i|^3+|u_i|^3)<\infty$.

假设 $\{(\boldsymbol{X}_i,\boldsymbol{U}_i,\boldsymbol{Y}_i)\}_{i=1}^n$ 是模型(1)的情形(ii)生成的独立同分布的随机样本,即

$$\begin{cases} \boldsymbol{Y}_i = \boldsymbol{X}_i^{\mathrm{T}}\boldsymbol{\beta} + g(\boldsymbol{T}_i) + \varepsilon_i \\ \boldsymbol{U}_i = \boldsymbol{T}_i + u_i \end{cases} \tag{3}$$

这里,测量误差 u_i 均值为零,独立同分布,且独立于$(\mathbf{T}_i, \boldsymbol{X}_i, \varepsilon_i)$,$\boldsymbol{\beta} \in \Theta \subseteq \boldsymbol{R}^p$,$(\boldsymbol{X}, \boldsymbol{T})$给定时 ε_i 的条件均值为零,并假定 ε_i 同方差.另外为使模型(3)可识别,进一步要求 u 有已知分布的特征函数 $\varphi_u(\cdot)$.

记 $\boldsymbol{T}$ 和 $\boldsymbol{U}$ 的密度函数分别为$f_T(\cdot)$,$f_U(\cdot)$,定义$f_T(t)$的反卷积核估计为[16]

$$\hat{f}_n(t) = \frac{1}{nh}\sum_{j=1}^{n} K_n\left(\frac{t - \boldsymbol{U}_j}{h}\right) \tag{4}$$

这里 $h = h_n$ 是带宽,$K_n(t) = (2\pi)^{-1}\int_{-\infty}^{\infty} \exp(-its)\dfrac{\varphi_K(s)}{\varphi_u(s/h)}\mathrm{d}s$ 其中 $\varphi_K(\cdot)$是满足第 5 节一些条件的核函数 $K(\cdot)$的傅里叶变换.又 $g(t) = E(\boldsymbol{Y} - \boldsymbol{X}^{\mathrm{T}}\boldsymbol{\beta} \mid \boldsymbol{T} = t)$,对给定的$\boldsymbol{\beta}$,$g(t)$的估计值为[11]

$$g_n(t) = \sum_{j=1}^{n} W_{nj}(t)(\boldsymbol{Y}_j - \boldsymbol{X}_j^{\mathrm{T}}\boldsymbol{\beta}) \tag{5}$$

其中 $W_{nj}(t) = (nh)^{-1}K_n((t - \boldsymbol{U}_j)/h)/\hat{f}_n(t)$.用式(5)代替式(3)中的 $g(t)$,再利用最小二乘估计,得 $\boldsymbol{\beta}$ 的 profile 最小二乘估计(PLS)为

$$\hat{\boldsymbol{\beta}} = \arg\min_{\boldsymbol{\beta} \in \boldsymbol{\Theta}}((\tilde{\boldsymbol{Y}} - \tilde{\boldsymbol{X}}\boldsymbol{\beta})^{\mathrm{T}}(\tilde{\boldsymbol{Y}} - \tilde{\boldsymbol{X}}\boldsymbol{\beta})) = (\tilde{\boldsymbol{X}}^{\mathrm{T}}\tilde{\boldsymbol{X}})^{-1}\tilde{\boldsymbol{X}}^{\mathrm{T}}\tilde{\boldsymbol{Y}} \tag{6}$$

$\tilde{\boldsymbol{X}} = (\tilde{x}_1, \cdots, \tilde{x}_n)^{\mathrm{T}}$,$\tilde{\boldsymbol{Y}} = (\tilde{\boldsymbol{Y}}_1, \cdots, \tilde{\boldsymbol{Y}}_n)^{\mathrm{T}}$,$\tilde{\boldsymbol{X}}_i = (\tilde{x}_{i1}, \cdots, \tilde{x}_{ip})^{\mathrm{T}}$,$i = 1, 2, \cdots, n$,其中 $\tilde{x}_{is} = x_{is} - \sum_{j=1}^{n} W_{nj}(\boldsymbol{U}_i)x_{js}$,$\tilde{\boldsymbol{Y}}_i = \boldsymbol{Y}_i - \sum_{j=1}^{n} W_{nj}(\boldsymbol{U}_i)\boldsymbol{Y}_j$.因此,非参函数 $g(t)$的估计为

$$\hat{g}_n(t) = \sum_{j=1}^{n} W_{nj}(t)(\boldsymbol{Y}_j - \boldsymbol{X}_j^{\mathrm{T}}\hat{\boldsymbol{\beta}}) \tag{7}$$

而模型误差方差的估计值定义为

$$\hat{\sigma}_n^2 = \frac{1}{n}\sum_{i=1}^{n}(\boldsymbol{Y}_i - \boldsymbol{X}_i^{\mathrm{T}}\hat{\boldsymbol{\beta}} - \hat{g}_n(\boldsymbol{U}_i))^2 \tag{8}$$

下面的定理将说明式(7)和式(8)的一致性.

定理 1　假设条件(C1)-(C6)成立,有

$$\sup_{1 \leqslant i \leqslant n} | \hat{g}_n(\boldsymbol{U}_i) - g(\boldsymbol{U}_i) | = o_P(1) \tag{9}$$

且

$$\hat{\sigma}_n^2 = \sigma^2 + o_P(1) \tag{10}$$

下面将用两种方法构造参数的约束估计.

1.1　拉格朗日乘数法

Liang 证明了 PLS 估计值的一致性和渐进正态性[11],但并没有考虑约束条件的存在,而有效的约束可以减少估计偏差.本节考虑约束条件(2)并在第 3 节对约束条件的合理性进行检验.首先,应用拉格朗日乘数法构造惩罚函数

$$Q(\boldsymbol{\beta}, \lambda) = (\tilde{\boldsymbol{Y}} - \tilde{\boldsymbol{X}}\boldsymbol{\beta})^{\mathrm{T}}(\tilde{\boldsymbol{Y}} - \tilde{\boldsymbol{X}}\boldsymbol{\beta}) - 2\lambda^{\mathrm{T}}(\boldsymbol{b} - \boldsymbol{A}\boldsymbol{\beta}) \tag{11}$$

最小化式(11)得到参数估计值.通过求解最优化问题,即把 $Q(\beta, \lambda)$分别对 β 和 λ 求偏导令其为零,得到

$$\hat{\lambda} = [\boldsymbol{A}(\tilde{\boldsymbol{X}}^{\mathrm{T}}\tilde{\boldsymbol{X}})^{-1}\boldsymbol{A}^{\mathrm{T}}]^{-1}(\boldsymbol{A}\hat{\boldsymbol{\beta}} - \boldsymbol{b}) \tag{12}$$

和$\boldsymbol{\beta}$ 的约束 PLS 估计

$$\hat{\boldsymbol{\beta}}_r = \hat{\boldsymbol{\beta}} - (\tilde{\boldsymbol{X}}^{\mathrm{T}}\tilde{\boldsymbol{X}})^{-1}\boldsymbol{A}^{\mathrm{T}}[\boldsymbol{A}(\tilde{\boldsymbol{X}}^{\mathrm{T}}\tilde{\boldsymbol{X}})^{-1}\boldsymbol{A}^{\mathrm{T}}]^{-1}(\boldsymbol{A}\hat{\boldsymbol{\beta}} - \boldsymbol{b}) \tag{13}$$

这里 $\hat{\boldsymbol{\beta}}$ 由式(6)给定.

由式(5),定义 $g(t)$ 的非参约束估计为

$$\hat{g}_{nr}(t)=\sum_{j=1}^{n}W_{nj}(t)(\boldsymbol{Y}_j-\boldsymbol{X}_j^{\mathrm{T}}\hat{\boldsymbol{\beta}}_r) \tag{14}$$

此时 σ^2 的约束估计为

$$\hat{\sigma}_{nr}^2=\frac{1}{n}\sum_{i=1}^{n}(\boldsymbol{Y}_i-\boldsymbol{X}_i^{\mathrm{T}}\hat{\boldsymbol{\beta}}_r-\hat{g}_{nr}(\boldsymbol{U}_i))^2 \tag{15}$$

下面的定理 2 也将给出 $\hat{\boldsymbol{\beta}}_r$ 的渐进正态性和 $\hat{g}_{nr}(t)$ 的一致性.记 $\zeta_1=\boldsymbol{X}_1-E(\boldsymbol{X}_1|\boldsymbol{T}_1)$,$\boldsymbol{I}_k$ 是 $k\times k$ 的单位阵.

定理 2 (i) 假设(C1)-(C5)成立,则有

$$\sqrt{n}(\hat{\boldsymbol{\beta}}_r-\boldsymbol{\beta})\to_D N(0,\sum)$$

这里 $\sum=\sigma^2\boldsymbol{D}\boldsymbol{B}^{-1}\boldsymbol{D}^{\mathrm{T}}$,$\boldsymbol{D}=I_p-\boldsymbol{B}^{-1}\boldsymbol{A}^{\mathrm{T}}[\boldsymbol{A}\boldsymbol{B}^{-1}\boldsymbol{A}^{\mathrm{T}}]^{-1}\boldsymbol{A}$,其中 $\boldsymbol{B}=E(\boldsymbol{\zeta}_1\boldsymbol{\zeta}_1^{\mathrm{T}})$,$\to_D$ 为依分布收敛.

(ii) 假设(C1-C5)成立,则有

$$\sup_{t\in[0,1]}|\hat{g}_{nr}(t)-g(t)|=o_P(1),\hat{\sigma}_{nr}^2=\sigma^2+o_P(1)$$

注:容易验证约束估计 $\hat{\boldsymbol{\beta}}_r$ 的协方差矩阵 $\sum$ 不大于无约束估计 $\hat{\boldsymbol{\beta}}$ 的协方差矩阵 $\sum_0=\sigma^2\boldsymbol{B}^{-1}$,即 $\sum_0-\sum$ 非负定.

根据定理 2(i),只有估计出 $\hat{\boldsymbol{\beta}}_r$ 的渐进协方差才能进行统计推断,从而构造置信区间.由文献[11]和定理 1,知 $n^{-1}\sum_{i=1}^{n}\tilde{\boldsymbol{X}}_i\tilde{\boldsymbol{X}}_i^{\mathrm{T}}$ 和 $\hat{\sigma}_n^2$ 分别是 $\boldsymbol{B}$ 和 σ^2 的一致估计,则 $\sum$ 的一致估计为 $\hat{\sum}=\hat{\sigma}_n^2\hat{\boldsymbol{D}}\hat{\boldsymbol{B}}^{-1}\hat{\boldsymbol{D}}^{\mathrm{T}}$,这里 $\hat{\boldsymbol{D}}$ 为 $\boldsymbol{D}$ 的估计值,而 $\boldsymbol{B}$ 用 $\hat{\boldsymbol{B}}=n^{-1}\tilde{\boldsymbol{X}}^{\mathrm{T}}\tilde{\boldsymbol{X}}$ 代替.此时有如下推论:

推论 1 在定理 2 的条件下,若 $\boldsymbol{\beta}$ 接近参数的真值,则有

$$\kappa_n\to_D\chi_p^2$$

这里 $\kappa_n=n(\hat{\boldsymbol{\beta}}_r-\boldsymbol{\beta})^{\mathrm{T}}(\hat{\sigma}_n^2\hat{\boldsymbol{D}}\hat{\boldsymbol{B}}^{-1}\hat{\boldsymbol{D}}^{\mathrm{T}})^{-1}(\hat{\boldsymbol{\beta}}_r-\boldsymbol{\beta})$,$\chi_p^2$ 是自由度为 p 的卡方分布.

接下来介绍另一种构造 $\boldsymbol{\beta}$ 约束估计值的方法.

1.2 方法 2

将 Wei 在部分线性 EV 模型中得到的参数约束估计方法应用到本文的模型中[15],过程如下.

定义 $p\times(p-k)$ 矩阵 $\boldsymbol{R}$ 使得 $\boldsymbol{Q}^{\mathrm{T}}\hat{=}(\boldsymbol{A}^{\mathrm{T}},\boldsymbol{R})$ 满秩且 $\boldsymbol{A}\boldsymbol{R}=0$,此时 $\boldsymbol{R}$ 存在但不唯一[17].记 $\boldsymbol{Q}^{-1}=[\boldsymbol{A}^{\mathrm{T}}(\boldsymbol{A}\boldsymbol{A}^{\mathrm{T}})^{-1},\boldsymbol{R}(\boldsymbol{R}^{\mathrm{T}}\boldsymbol{R})^{-1}]$,再令 $\boldsymbol{\theta}=\boldsymbol{Q}\boldsymbol{\beta}$,则有 $\boldsymbol{\theta}=(\boldsymbol{\theta}_1^{\mathrm{T}},\boldsymbol{\theta}_2^{\mathrm{T}})^{\mathrm{T}}$,其中 $\boldsymbol{\theta}_1=\boldsymbol{A}\beta$,$\boldsymbol{\theta}_2=\boldsymbol{R}^{\mathrm{T}}\boldsymbol{\beta}$.

令 $\boldsymbol{G}=(g(\boldsymbol{T}_1),\cdots,g(\boldsymbol{T}_n))^{\mathrm{T}}$,$\varepsilon=(\varepsilon_1,\cdots,\varepsilon_n)^{\mathrm{T}}$,知模型(1)的矢量形式为 $\boldsymbol{Y}=\boldsymbol{X}\boldsymbol{\beta}+\boldsymbol{G}+\varepsilon$.再由 $\boldsymbol{A}\boldsymbol{\beta}=\boldsymbol{b}$,则模型可改写为

$$\boldsymbol{Y}=\boldsymbol{X}\boldsymbol{A}^{\mathrm{T}}(\boldsymbol{A}\boldsymbol{A}^{\mathrm{T}})^{-1}\boldsymbol{b}+\boldsymbol{X}\boldsymbol{R}(\boldsymbol{R}^{\mathrm{T}}\boldsymbol{R})^{-1}\boldsymbol{\theta}_2+\boldsymbol{G}+\varepsilon \tag{16}$$

记 $\boldsymbol{Y}^*=\boldsymbol{Y}-\boldsymbol{X}\boldsymbol{A}^{\mathrm{T}}(\boldsymbol{A}\boldsymbol{A}^{\mathrm{T}})^{-1}\boldsymbol{b}$,$\boldsymbol{X}^*=\boldsymbol{X}\boldsymbol{R}(\boldsymbol{R}^{\mathrm{T}}\boldsymbol{R})^{-1}$,模型(16)转化为部分线性模型 $\boldsymbol{Y}^*=\boldsymbol{X}^*\theta_2+\boldsymbol{G}+\varepsilon$,即

$$\boldsymbol{Y}_i^*=\boldsymbol{X}_i^{*T}\boldsymbol{\theta}_2+g(\boldsymbol{T}_i)+\varepsilon_i \tag{17}$$

这里 $\boldsymbol{Y}^*=(\boldsymbol{Y}_1^*,\cdots,\boldsymbol{Y}_n^*)^{\mathrm{T}}$,$\boldsymbol{X}^*=(\boldsymbol{X}_1^*,\cdots,\boldsymbol{X}_n^*)^{\mathrm{T}}$.当 T_i 可观测时,应用 profile 最小二乘有 $\boldsymbol{\theta}_2$ 的 PLS 估计:

$$\bar{\boldsymbol{\theta}}_2=(\bar{\boldsymbol{X}}^{*\mathrm{T}}\bar{\boldsymbol{X}}^*)^{-1}\bar{\boldsymbol{X}}^{*\mathrm{T}}\bar{\boldsymbol{Y}}^* \tag{18}$$

这里 $\bar{\boldsymbol{X}}^*$ 和 $\bar{\boldsymbol{Y}}^*$ 如式(6)中所定义,但权重 $W_{nj}(\cdot)$ 的 $K_n(\cdot)$ 却是一般核函数的重新排列.当替代变量 U 可观测时,类似于式(6)有 $\boldsymbol{\theta}_2$ 的估计值

$$\hat{\boldsymbol{\theta}}_2=(\tilde{\boldsymbol{X}}^{*\mathrm{T}}\tilde{\boldsymbol{X}}^*)^{-1}\tilde{\boldsymbol{X}}^{*\mathrm{T}}\tilde{\boldsymbol{Y}}^* \tag{19}$$

其中 $\tilde{\boldsymbol{X}}^*$ 和 $\tilde{\boldsymbol{Y}}^*$ 和式(6)的 $\tilde{\boldsymbol{X}}$ 和 $\tilde{\boldsymbol{Y}}$ 类似,是 $\boldsymbol{X}_i,\boldsymbol{Y}_i$ 的代替值.既然 $\tilde{\boldsymbol{X}}^*=\tilde{\boldsymbol{X}}\boldsymbol{R}(\boldsymbol{R}^{\mathrm{T}}\boldsymbol{R})^{-1}$, $\tilde{\boldsymbol{Y}}^*=\tilde{\boldsymbol{Y}}-\tilde{\boldsymbol{X}}\boldsymbol{A}^{\mathrm{T}}(\boldsymbol{A}\boldsymbol{A}^{\mathrm{T}})^{-1}\boldsymbol{b}$,则有

$$\hat{\boldsymbol{\theta}}_2=\boldsymbol{R}^{\mathrm{T}}\boldsymbol{R}(\boldsymbol{R}^{\mathrm{T}}\tilde{\boldsymbol{X}}^{\mathrm{T}}\tilde{\boldsymbol{X}}\boldsymbol{R})^{-1}\boldsymbol{R}^{\mathrm{T}}\tilde{\boldsymbol{X}}^{\mathrm{T}}\tilde{\boldsymbol{Y}}-\boldsymbol{R}^{\mathrm{T}}\boldsymbol{R}(\boldsymbol{R}^{\mathrm{T}}\tilde{\boldsymbol{X}}^{\mathrm{T}}\tilde{\boldsymbol{X}}\boldsymbol{R})^{-1}\boldsymbol{R}^{\mathrm{T}}\tilde{\boldsymbol{X}}^{\mathrm{T}}\tilde{\boldsymbol{X}}\boldsymbol{A}^{\mathrm{T}}(\boldsymbol{A}\boldsymbol{A}^{\mathrm{T}})^{-1}\boldsymbol{b} \tag{20}$$

又知 $\boldsymbol{\beta}=\boldsymbol{Q}^{-1}\boldsymbol{\theta}$,简单计算可得 $\boldsymbol{\beta}$ 的约束估计

$$\tilde{\boldsymbol{\beta}}_r=\boldsymbol{R}(\boldsymbol{R}^{\mathrm{T}}\tilde{\boldsymbol{X}}^{\mathrm{T}}\tilde{\boldsymbol{X}}\boldsymbol{R})^{-1}\boldsymbol{R}^{\mathrm{T}}\tilde{\boldsymbol{X}}^{\mathrm{T}}\tilde{\boldsymbol{Y}}+[\boldsymbol{I}_p-\boldsymbol{R}(\boldsymbol{R}^{\mathrm{T}}\tilde{\boldsymbol{X}}^{\mathrm{T}}\tilde{\boldsymbol{X}}\boldsymbol{R})^{-1}\boldsymbol{R}^{\mathrm{T}}\tilde{\boldsymbol{X}}^{\mathrm{T}}\tilde{\boldsymbol{X}}]\boldsymbol{A}^{\mathrm{T}}(\boldsymbol{A}\boldsymbol{A}^{\mathrm{T}})^{-1}\boldsymbol{b} \tag{21}$$

比较式(21)和(13),则有如下结论.

定理 3　如果矩阵 $\tilde{\boldsymbol{X}}^{\mathrm{T}}\tilde{\boldsymbol{X}}$ 非奇异,则有

$$\tilde{\boldsymbol{\beta}}_r=\hat{\boldsymbol{\beta}}_r$$

其中 $\hat{\boldsymbol{\beta}}_r,\tilde{\boldsymbol{\beta}}_r$ 见式(13)(21).

2　参数的约束条件检验

考虑模型(3)线性部分参数带有约束条件的情形,对约束条件的合理性进行检验.不失一般性的考虑如下带有线性假设的检验:

$$\mathrm{H}_0:\boldsymbol{A\beta}=\boldsymbol{b}\ \text{vs}\ \mathrm{H}_1:\boldsymbol{A\beta}\neq\boldsymbol{b} \tag{22}$$

Fan 和 Huang 提出部分变系数模型参数的 profile 广义极大似然比检验[18],并证明了 Wilks 现象的存在,即原假设成立时该统计量近似服从与 σ^2 无关的卡方分布.应用该方法检验模型(3)的式(22),发现 Wilks 现象仍然存在,但当线性部分存在测量误差时却不存在 Wilks 现象[15].本节检验过程如下:

原假设成立,即 $\boldsymbol{A\beta}=\boldsymbol{b}$ 时,参数的约束估计 $\hat{\boldsymbol{\beta}}_r$ 和非参估计 $\hat{g}_{nr}(t)$ 分别由式(13)和式(14)给出.相应的残差平方和为

$$\mathrm{RSS}_0=\sum_{i=1}^{n}(\boldsymbol{Y}_i-\boldsymbol{X}_i^{\mathrm{T}}\hat{\boldsymbol{\beta}}_r-\hat{g}_{nr}(\boldsymbol{U}_i))^2 \tag{23}$$

而 H_1 成立时的残差平方和为

$$\mathrm{RSS}_1=\sum_{i=1}^{n}(\boldsymbol{Y}_i-\boldsymbol{X}_i^{\mathrm{T}}\hat{\boldsymbol{\beta}}-\hat{g}_n(\boldsymbol{U}_i))^2 \tag{24}$$

此时定义 GLR 检验统计量为[18]

$$\boldsymbol{T}_n=n\log\frac{\mathrm{RSS}_0}{\mathrm{RSS}_1}\approx n\frac{\mathrm{RSS}_0-\mathrm{RSS}_1}{\mathrm{RSS}_1} \tag{25}$$

如果 H_0 为真,直观上 RSS_0 和 RSS_1 不应相差过大.所以当 GLR 统计量较大时,应拒绝原假设.理论说明由定理 4 给出.

定理 4　若检验式(22)的原假设和(C1)-(C6)成立,则有, $\boldsymbol{T}_n\to_D\chi_k^2$,这里 χ_k^2 是自由度为 k 的卡方分布.

定理 5　若检验式(22)的备则假设和(C1)-(C6)成立,则有, $\boldsymbol{T}_n\to_D\chi_k^2(\delta)$,这里 $\chi_k^2(\delta)$ 代表自由度为 k,非中心化的卡方随机变量,其中非中心参数为

$$\delta=\lim_{n\to\infty}n(\boldsymbol{A\beta}-\boldsymbol{b})^{\mathrm{T}}(\sigma^2\boldsymbol{A}\boldsymbol{B}^{-1}\boldsymbol{A}^{\mathrm{T}})^{-1}(\boldsymbol{A\beta}-\boldsymbol{b})$$

注:定理 4 说明原假设成立时,$\boldsymbol{T}_n$ 与 σ^2,$\boldsymbol{\beta}$ 和 $g(\cdot)$ 无关,近似服从自由度为 k 的卡方分布.这个定理既提供函数带误差的部分线性模型参数分量检验的方法,也说明了 Wilks 现象依然存在.虽然只考虑了约束参数分量的检验,但也可用类似的方法进行非参函数的检验.

3 数值模拟

为对约束估计值和统计量 $\boldsymbol{T}_n$ 进行检验,本节在有限样本下作数值模拟,数据由下产生

$$\begin{cases}\boldsymbol{Y}_i=\boldsymbol{x}_{i1}\beta_1+x_{i2}\beta_2+g(\boldsymbol{T}_i)+\varepsilon_i\\ \boldsymbol{U}_i=\boldsymbol{T}_i+u_i\end{cases}\tag{26}$$

这里(x_{i1},x_{i2})由相关系数为 0.4 的二维标准正态分布产生,$\boldsymbol{T}_i\sim N(0.5,0.25^2)$,$g(t)=t_+^3(1-t)_+^3$.为研究误差分布对参数估计值的影响,检验如下两种情形:u_i 由双指数分布(平滑情况)产生;u_i 由正态分布(超平滑情况)产生.假设 σ_0^2 为误差 u_i 的方差,并取 $\sigma_0^2=\frac{3}{7}\mathrm{Var}(\boldsymbol{T})=\frac{3}{10}\mathrm{Var}(\boldsymbol{U})$,则该信噪比可达 0.7[16].

例 1 双指数误差

假设误差 u 有如下双指数密度函数

$$f_u(u)=\frac{1}{\sqrt{2}\sigma_0}\exp\left(-\frac{\sqrt{2}\,|u|}{\sigma_0}\right)\tag{27}$$

核函数 $K(\cdot)$是高斯核,即标准正态密度.简单计算可知式(4)中的核 $K_n(\cdot)$可由如下定义

$$K_n(x)=\left(\sqrt{2\pi}\right)^{-1}\exp(-x^2/2)\{1-(x^2-1)\sigma_0^2/2h^2\}\tag{28}$$

根据条件(C4)(i)选取 $h=1.16\cdot\mathrm{sd}(\boldsymbol{T})\cdot n^{-1/9}$[19].

例 2 正态误差

假设 $u\sim N(0,\sigma_0^2)$,核函数 $K(\cdot)$的傅里叶变换为 $\varphi_K(t)=(1-t^2)_+^3$,根据式(4)有

$$K_n(x)=\frac{1}{\pi}\int_0^1\cos(tx)(1-t^2)^3\exp\left(\frac{\sigma_0^2t^2}{2h^2}\right)\mathrm{d}t\tag{29}$$

根据(C4)(ii)选取 $h=1.1\sigma_0(\log n)^{-1/2}$.

3.1 约束估计的一致性检验

在模型(26)中,令 $\beta_1=1$,$\beta_2=3$.考虑约束 $3\beta_1+\beta_2=6$ 和模型误差 ε 分别是均匀分布、正态分布、学生 t-分布、卡方分布的情形,分别给出约束估计 $\hat{\boldsymbol{\beta}}_r$ 和 $\hat{\sigma}_{nr}^2$的样本均方误差(MSE)和样本标准差(SD),其中

$$\mathrm{MSE}(\hat{\boldsymbol{\beta}}_r)=N^{-1}\sum_{k=1}^{N}|\hat{\boldsymbol{\beta}}_r^{(k)}-\hat{\boldsymbol{\beta}}|^2,\mathrm{SD}(\hat{\boldsymbol{\beta}}_r)=\sqrt{N^{-1}\sum_{k=1}^{N}|\boldsymbol{\beta}_r^{(k)}-\bar{\boldsymbol{\beta}}_r|^2}$$

这里 $\hat{\boldsymbol{\beta}}_r^{(k)}$ 是运行 k 次得到的 $\boldsymbol{\beta}$ 的约束估计,$\bar{\boldsymbol{\beta}}_r$ 是所有 $\hat{\boldsymbol{\beta}}_r^{(k)}$ 的平均值,$\|\cdot\|$表示欧几里得范数.进行 $N=1\ 000$ 次模拟检验其在不同样本量下的表现,结果见表 1.

表 1 $\hat{\boldsymbol{\beta}}_r$ 和 $\hat{\sigma}_{nr}$的均方误差和标准差

		正态误差 u				双指数误差 u			
		$\hat{\boldsymbol{\beta}}_r$		$\hat{\sigma}_{nr}$		$\hat{\boldsymbol{\beta}}_r$		$\hat{\sigma}_{nr}$	
ε	n	MSE	SD	MSE	SD	MSE	SD	MSE	SD
	50	0.105 955	0.087 871	0.001 288	0.035 835	0.103 654	0.086 240	0.001 899	0.029 353
(a)	100	0.073 343	0.062 464	0.001 040	0.032 081	0.072 241	0.061 064	0.000 880	0.023 058
	150	0.060 941	0.065 653	0.002 695	0.051 703	0.056 521	0.048 600	0.000 568	0.019 855

续表 1

		正态误差 u				双指数误差 u			
		$\hat{\boldsymbol{\beta}}_r$		$\hat{\sigma}_{nr}$		$\hat{\boldsymbol{\beta}}_r$		$\hat{\sigma}_{nr}$	
ε	n	MSE	SD	MSE	SD	MSE	SD	MSE	SD
(b)	50	0.100 718	0.085 269	0.003 011	0.054 864	0.104 343	0.088 478	0.003 934	0.054 700
	100	0.073 391	0.066 161	0.001 589	0.039 789	0.070 152	0.069 898	0.002 714	0.049 206
	150	0.060 612	0.058 282	0.004 041	0.063 356	0.060 243	0.059 272	0.001 459	0.036 298
(c)	50	0.100 094	0.084 631	0.004 471	0.066 862	0.100 504	0.084 050	0.004 413	0.058 514
	100	0.072 925	0.064 478	0.004 812	0.069 370	0.072 572	0.063 513	0.003 375	0.056 329
	150	0.060 612	0.058 282	0.003 635	0.060 112	0.058 087	0.054 198	0.001 972	0.042 785
(d)	50	0.101 710	0.086 436	0.003 963	0.062 700	0.104 075	0.087 832	0.004 695	0.060 802
	100	0.072 958	0.071 923	0.030 116	0.173 474	0.072 664	0.065 861	0.002 649	0.047 467
	150	0.060 129	0.062 533	0.002 747	0.052 306	0.059 061	0.049 065	0.001 464	0.035 682

(a):$U(-\frac{\sqrt{3}}{2},\frac{\sqrt{3}}{2})$;(b):$N(0,0.5^2)$;(c):$\frac{\sqrt{3}}{4}t(8)$;(d):$\frac{1}{8}\chi_8^2-1$.

由表 1 知,当样本量增加时,均方误差和标准差在递减.说明随着样本量增多,约束估计逐渐接近真实的参数,与结论一致.

3.2　检验统计量有效性检验

对模型式(26),考虑如下检验:

$$H_0:\beta_1=\beta_2 \text{ vs } H_1:\beta_1\neq\beta_2 \tag{30}$$

关于 $\beta_1=2,\beta_2=2-c,c=0$ 表示原假设,否则就是备则假设.

原假设成立时,对样本量为 $n=100$ 的情形运行 1 000 次来检验统计量 $\boldsymbol{T}_n$ 是否服从定理 4 的 $\chi_k^2(k=1)$. 图 1,2 分别描绘了均匀误差下例 1,例 2 的误差 Q-Q 图,也揭示了 1 000 个 GLR 统计量的四分位数和 χ_1^2 分布四分位数的关系,可以看出 GLR 统计量可以很好的拟合期望的卡方分布,也与之前结果一致.

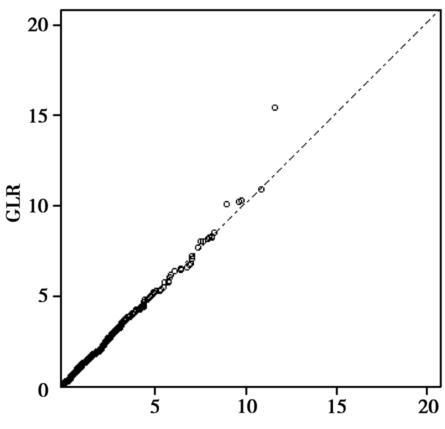

图 1　例 1 的 Q-Q 图

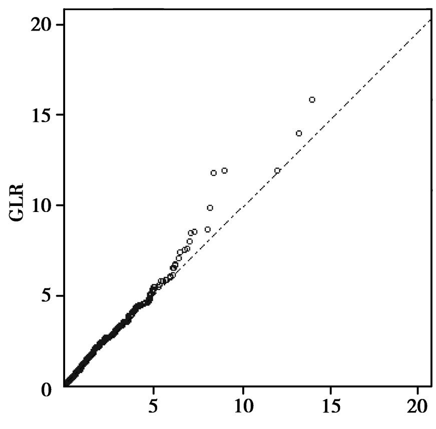

图 2　例 2 的 Q-Q 图

为评估第 3 节提出检验过程的有效性,重复 1 000 次得到检验统计量的功效曲线.图 3 描绘了 GLR 检验的功效曲线,拒绝率是根据显著水平 $\alpha=0.05$ 在不同的样本量下计算的,从图 3 可以看出当样本量增大时检验效果变好,这也说明了检验过程是有效性.图 4 描绘了固定样本量 $n=100$ 时施加不同模型误差的情形,如图所示,模型误差时正态分布、卡方分布、学生 t-分布的情形相似,但当 c 离 0 较近时,均匀分布下的情形有所不同.例 2 的结论类似可得(图 5,6).

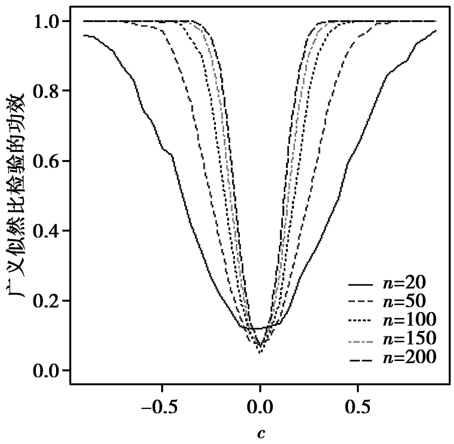

图 3 不同样本量下例 1 的功效曲线

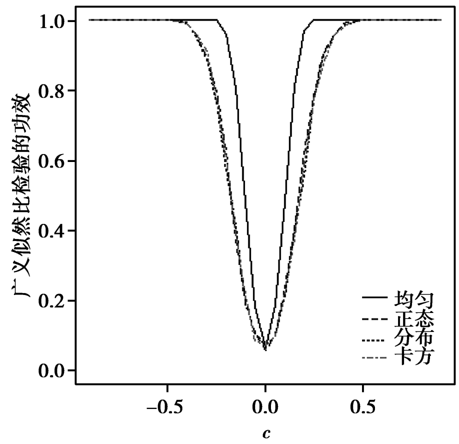

图 4 不同模型误差下例 1 的功效曲线

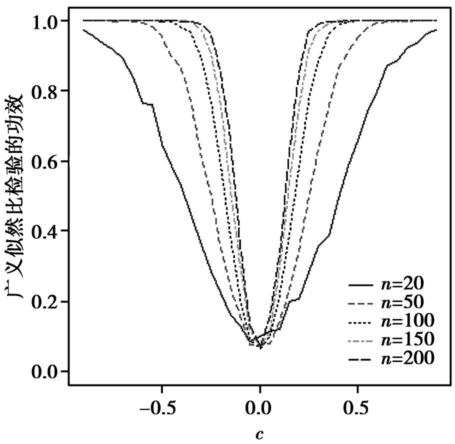

图 5 不同样本量下例 2 的功效曲线

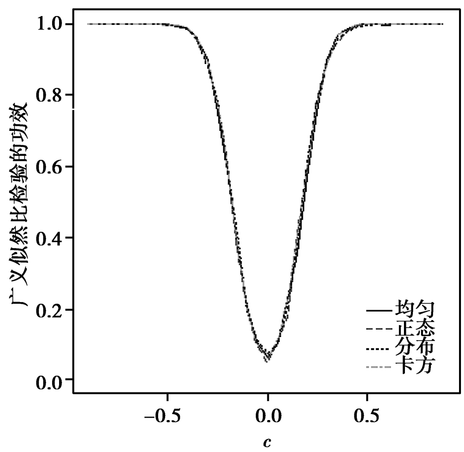

图 6 不同模型误差下例 2 的功效曲线

4 主要结论的假设和证明

最后在证明结论前,先介绍如下引理.

引理 1[11] 令 $\xi_1,\xi_2,\cdots,\xi_n$ 为零均值的独立随机变量,$\sup_{1\leqslant i\leqslant n}E|\xi_i|^r\leqslant C<\infty\ (r\geqslant 2)$.假设$\{a_{ki},1\leqslant k,i\leqslant n\}$为正数数列,其对于 $0<p_1<1$ 满足 $\sup_{1\leqslant k,i\leqslant n}|a_{ki}|\leqslant n^{-p_1}$,对于 $p_2\geqslant\max(0,2/r-p_1)$满足 $\max_{1\leqslant i\leqslant n}\sum_{k=1}^{n}a_{ki}\leqslant n^{p_2}$,则有 $\max_{1\leqslant i\leqslant n}\left|\sum_{k=1}^{n}a_{ki}\xi_k\right|=O(n^{-s}\log n)s=(p_1-p_2)/2$ a.s.

引理 2[11] 假设条件(C1)-(C5)成立,则有当 $n\to\infty$ 时,$n^{-1}\tilde{\boldsymbol{X}}^{\mathrm{T}}\tilde{\boldsymbol{X}}\to_p\boldsymbol{B}$

引理 3[11] 假设条件(C1)-(C6)成立,则有,$\sqrt{n}(\hat{\beta}-\boldsymbol{\beta})\to_D N(0,\sigma^2\boldsymbol{B}^{-1})$

引理 4[11] 假设条件(C3)和(C5)成立,

(i) 如果 U 是平滑误差,取 $2m>2\alpha+1$,则有 $\max_{1\leqslant j\leqslant n}\left|r_j(\boldsymbol{T}_i)-\sum_{k=1}^{n}W_{nk}(\boldsymbol{U}_i)r_j(\boldsymbol{T}_k)\right|=o(n^{-1/4})$,其中 $j=0,1,\cdots,p$, $r_0(\cdot)=g(\cdot)$, $r_j(\cdot)=h_j(\cdot)(1\leqslant j\leqslant p)$.

(ii) 如果 $\boldsymbol{U}$ 是超平滑误差, $\boldsymbol{X}$ 和 $\boldsymbol{T}$ 独立,(i)的结论对于 $j=1,\cdots,p$ 仍然成立,但

是$\max_{1\le i\le n}\left|g(\boldsymbol{T}_i)-\sum_{k=1}^{n}W_{nk}(\boldsymbol{U}_i)g(\boldsymbol{T}_k)\right|=o(1)$.

定理 1 的证明　首先证明式(9).由

$$\sup_{1\le i\le n}|g(\boldsymbol{U}_i)-\hat{g}_n(\boldsymbol{U}_i)|\le\sup_{1\le i\le n}\left|g(\boldsymbol{U}_i)-\sum_{j=1}^{n}W_{nj}(U_i)g(\boldsymbol{T}_j)\right|+$$
$$\sup_{1\le i\le n}\left|\sum_{j=1}^{n}W_{nj}(\boldsymbol{U}_i)\boldsymbol{X}_j^{\mathrm{T}}(\hat{\beta}-\boldsymbol{\beta})\right|+\sup_{1\le i\le n}\left|\sum_{k=j}^{n}W_{nj}(\boldsymbol{U}_i)\varepsilon_j\right| \tag{31}$$

回顾 $W_{nj}(\cdot)$的定义,取 $h=n^{-1/3}$[11]可以验证 $\max_{1\le i,j\le n}|W_{nj}(\boldsymbol{U}_i)|=O(n^{-2/3})$.由条件(C3)和引理 1,取 $r=3$,$\xi_k=\varepsilon_k$,$a_{ki}=W_{nk}(\boldsymbol{U}_i)$,$p_1=2/3$ 和 $p_2=0$ 有

$$\max_{1\le i\le n}\left|\sum_{k=1}^{n}W_{nk}(\boldsymbol{U}_i)\varepsilon_k\right|=O(n^{-1/3}\log n)\ \text{a.s} \tag{32}$$

利用引理 4 可知式(31)右边第一项是 $o(1)$,中间项由条件(C5),引理 3 和上面的讨论知为 $o(1)$.再由式(32)即得式(9).

令 $\tilde{g}(\cdot)=g(\cdot)-\sum_{j=1}^{n}W_{nj}(\cdot)g(\boldsymbol{T}_j)$,$\bar{\varepsilon}_i=\sum_{j=1}^{n}W_{nj}(\boldsymbol{U}_i)\varepsilon_j$,为证明式(10),将式(8)中的$\hat{\boldsymbol{\sigma}}_n^2$ 作如下分解

$$\hat{\boldsymbol{\sigma}}_n^2=n^{-1}\sum_{i=1}^{n}(\tilde{\boldsymbol{Y}}_i-\tilde{\boldsymbol{X}}_i^{\mathrm{T}}\hat{\boldsymbol{\beta}})^2=$$
$$n^{-1}\sum_{i=1}^{n}\varepsilon_i^2+n^{-1}\sum_{i=1}^{n}(\hat{\boldsymbol{\beta}}-\boldsymbol{\beta})^{\mathrm{T}}\tilde{\boldsymbol{X}}_i\tilde{\boldsymbol{X}}_i^{\mathrm{T}}(\hat{\boldsymbol{\beta}}-\boldsymbol{\beta})+n^{-1}\sum_{i=1}^{n}\tilde{g}^2(\boldsymbol{T}_i)+n^{-1}\sum_{i=1}^{n}\bar{\varepsilon}_i^2-2n^{-1}\sum_{i=1}^{n}\tilde{g}^2(\boldsymbol{T}_i)\tilde{X}_i^{\mathrm{T}}(\hat{\boldsymbol{\beta}}-\boldsymbol{\beta})-$$
$$2n^{-1}\sum_{i=1}^{n}(\hat{\boldsymbol{\beta}}-\boldsymbol{\beta})^{\mathrm{T}}\tilde{\boldsymbol{X}}_i\varepsilon_i+2n^{-1}\sum_{i=1}^{n}(\hat{\boldsymbol{\beta}}-\boldsymbol{\beta})^{\mathrm{T}}\tilde{\boldsymbol{X}}_i\bar{\varepsilon}_i+2n^{-1}\sum_{i=1}^{n}\tilde{g}(\boldsymbol{T}_i)\varepsilon_i-2n^{-1}\sum_{i=1}^{n}\tilde{g}(\boldsymbol{T}_i)\bar{\varepsilon}_i-2n^{-1}\sum_{i=1}^{n}\varepsilon_i\bar{\varepsilon}_i\hat{=}$$
$$M_{n1}+M_{n2}+M_{n3}+M_{n4}-2M_{n5}-2M_{n6}+2M_{n7}+2M_{n8}-2M_{n9}-2M_{n10}$$

则只需说明 M_{n1}依概率收敛到 σ^2,$2\le k\le 10$ 时 M_{nk}依概率收敛到 0.前者可直接应用大数定律得到,后者证明如下.

引理 3 表明$\hat{\boldsymbol{\beta}}-\boldsymbol{\beta}=O_P(n^{-1/2})$,再由引理 2 可得 $M_{n2}=o_P(1)$.由引理 4 知 $M_{n3}=o_P(1)$.显然应用式(32)有 $M_{n4}=o_P(1)$.既然由引理 4 可以得到 $\max_{1\le i\le n}\|\tilde{\boldsymbol{X}}_i\|=o_P(n^{1/2})$,再由引理 2,3 易知 $M_{n5}=o_P(1)$.由 Cauchy-Schwarz 不等式和以上所述,有

$$|M_{n6}|\le O_P(n^{-1/2})n^{-1}\sum_{i=1}^{n}\|\tilde{\boldsymbol{X}}_i\varepsilon_i\|\le O_P(n^{-1/2})\left(n^{-1}\sum_{i=1}^{n}\|\tilde{\boldsymbol{X}}_i\|^2\right)^{1/2}\left(n^{-1}\sum_{i=1}^{n}|\varepsilon_i|^2\right)^{1/2}=o_P(1)$$

再由 $\max_{1\le i\le n}|\bar{\varepsilon}_i|=O_P(n^{-1/3}\log n)$知

$$|M_{n7}|\le O_P(n^{-1/2})\max_{1\le i\le n}\|\tilde{X}_i\|\max_{1\le i\le n}|\bar{\varepsilon}_i|=o_P(1)$$

相似地可以得到 $M_{n8}=M_{n9}=M_{n10}=o_P(1)$,证毕.

定理 2 的证明　直接应用引理 2 和引理 3 的证明即得(i),所以只需证明(ii)即可.一方面,

$$\sup_{0\le t\le 1}|\hat{g}_{nr}(t)-g(t)|\le\sup_{0\le t\le 1}|\hat{g}_n(t)-g(t)|+\sup_{0\le t\le 1}\left|\sum_{j=1}^{n}W_{nj}(t)\tilde{\boldsymbol{X}}_j^{\mathrm{T}}(\hat{\boldsymbol{\beta}}_r-\boldsymbol{\beta})\right|$$

由 Cauchy-Schwarz 得

$$\sup_{0\le t\le 1}\left|\sum_{j=1}^{n}W_{nj}(t)\tilde{X}_j^{\mathrm{T}}(\hat{\boldsymbol{\beta}}_r-\boldsymbol{\beta})\right|\le\sup_{0\le t\le 1}\left|\sum_{j=1}^{n}W_{nj}^2(t)\right|^{1/2}\cdot\left|(\hat{\boldsymbol{\beta}}_r-\boldsymbol{\beta})^{\mathrm{T}}\sum_{j=1}^{n}\tilde{\boldsymbol{X}}_j\tilde{\boldsymbol{X}}_j^{\mathrm{T}}(\hat{\boldsymbol{\beta}}_r-\boldsymbol{\beta})\right|^{1/2}\le$$
$$\sup_{0\le t\le 1}|W_{nj}(t)|^{1/2}\sup_{0\le t\le 1}\left|\sum_{j=1}^{n}W_{nj}(t)\right|^{1/2}\cdot O_P(1)=O_P(n^{-1/3})=o_P(1)$$

再由式(9)就得到了(ii)的第一个结论.

另一方面,对 $\hat{\sigma}_{nr}^2$ 作如下分解:

$$\hat{\sigma}_{nr}^2 = n^{-1}\sum_{i=1}^{n}(\tilde{\boldsymbol{Y}}_i - \tilde{\boldsymbol{X}}_i^{\mathrm{T}}\hat{\boldsymbol{\beta}})^2 + 2n^{-1}\sum_{i=1}^{n}(\tilde{\boldsymbol{Y}}_i - \tilde{\boldsymbol{X}}_i^{\mathrm{T}}\hat{\boldsymbol{\beta}})\tilde{\boldsymbol{X}}_i^{\mathrm{T}}(\hat{\boldsymbol{\beta}}_r - \hat{\boldsymbol{\beta}}) + n^{-1}\sum_{i=1}^{n}(\hat{\boldsymbol{\beta}}_r - \hat{\boldsymbol{\beta}})^{\mathrm{T}}\tilde{\boldsymbol{X}}_i\tilde{\boldsymbol{X}}_i^{\mathrm{T}}(\hat{\boldsymbol{\beta}}_r - \hat{\boldsymbol{\beta}}) =$$

$$\hat{\sigma}_n^2 + \Delta_{n1} + \Delta_{n2}$$

由(i),引理 2,$\hat{\sigma}_n^2$ 依概率收敛到 σ^2(定理 1)知 $\Delta_{n2}=O_P(n^{-1})$,再由 Cauchy−Schwarz 不等式得

$$|\Delta_{n1}| \leqslant \hat{\sigma}_n |\Delta_{n2}|^{1/2} = O_P(n^{-1/2}) = o_P(1)$$

证毕.

定理 3 的证明 简单计算易知

$$\begin{bmatrix} \boldsymbol{R}^{\mathrm{T}}\tilde{\boldsymbol{X}}^{\mathrm{T}}\tilde{\boldsymbol{X}} \\ \boldsymbol{A} \end{bmatrix}(\hat{\boldsymbol{\beta}}_r - \tilde{\boldsymbol{\beta}}_r) = 0 \tag{33}$$

又 $\tilde{\boldsymbol{\beta}}_r=\hat{\boldsymbol{\beta}}_r$ 等价于上述等式(33)的系数矩阵非奇异,类似文献[15]中定理 2.2 的证明即得.

定理 4 的证明 方便起见,记 $\boldsymbol{P}=(\tilde{\boldsymbol{X}}^{\mathrm{T}}\tilde{\boldsymbol{X}})^{-1}\boldsymbol{A}^{\mathrm{T}}[\boldsymbol{A}(\tilde{\boldsymbol{X}}^{\mathrm{T}}\tilde{\boldsymbol{X}})^{-1}\boldsymbol{A}^{\mathrm{T}}]^{-1}$,由式(13)中 $\hat{\boldsymbol{\beta}}_r$ 的定义,有 $\hat{\boldsymbol{\beta}}_r=\hat{\boldsymbol{\beta}}-\boldsymbol{P}(\boldsymbol{A}\hat{\boldsymbol{\beta}}-\boldsymbol{b})$.由式(23)将 RSS_0 作如下分解

$$\begin{aligned}\mathrm{RSS}_0 &= (\tilde{\boldsymbol{Y}} - \tilde{\boldsymbol{X}}\hat{\boldsymbol{\beta}}_r)^{\mathrm{T}}(\tilde{\boldsymbol{Y}} - \tilde{\boldsymbol{X}}\hat{\boldsymbol{\beta}}_r) = \\ &\mathrm{RSS}_1 + 2(\tilde{\boldsymbol{Y}} - \tilde{\boldsymbol{X}}\hat{\boldsymbol{\beta}})^{\mathrm{T}}\tilde{\boldsymbol{X}}\boldsymbol{P}(\boldsymbol{A}\hat{\boldsymbol{\beta}} - \boldsymbol{b}) + (\boldsymbol{A}\hat{\boldsymbol{\beta}} - \boldsymbol{b})^{\mathrm{T}}\boldsymbol{P}^{\mathrm{T}}\tilde{\boldsymbol{X}}^{\mathrm{T}}\tilde{\boldsymbol{X}}\boldsymbol{P}(\boldsymbol{A}\hat{\boldsymbol{\beta}} - \boldsymbol{b})\end{aligned} \tag{34}$$

其中 $\mathrm{RSS}_1=(\tilde{\boldsymbol{Y}}-\tilde{\boldsymbol{X}}\hat{\boldsymbol{\beta}})^{\mathrm{T}}(\tilde{\boldsymbol{Y}}-\tilde{\boldsymbol{X}}\hat{\boldsymbol{\beta}})$.又式(34)右边中间项为 0,即有

$$\mathrm{RSS}_0 = \mathrm{RSS}_1 + (\boldsymbol{A}\hat{\boldsymbol{\beta}} - \boldsymbol{b})^{\mathrm{T}}\boldsymbol{P}^{\mathrm{T}}\tilde{\boldsymbol{X}}^{\mathrm{T}}\tilde{\boldsymbol{X}}\boldsymbol{P}(\boldsymbol{A}\hat{\boldsymbol{\beta}} - \boldsymbol{b}) \tag{35}$$

又知 $\hat{\sigma}_n^2=n^{-1}\mathrm{RSS}_1$,则由式(10)和(25)得

$$\begin{aligned}T_n &= \sigma^{-2}(\boldsymbol{A}\hat{\boldsymbol{\beta}} - \boldsymbol{b})^{\mathrm{T}}\boldsymbol{P}^{\mathrm{T}}\tilde{\boldsymbol{X}}^{\mathrm{T}}\tilde{\boldsymbol{X}}\boldsymbol{P}(\boldsymbol{A}\hat{\boldsymbol{\beta}} - \boldsymbol{b}) + o_{\boldsymbol{P}}(1) = \\ &\sigma^{-2}(\boldsymbol{A}\hat{\boldsymbol{\beta}} - \boldsymbol{b})^{\mathrm{T}}[\boldsymbol{A}(\tilde{\boldsymbol{X}}^{\mathrm{T}}\tilde{\boldsymbol{X}})^{-1}\boldsymbol{A}^{\mathrm{T}}]^{-1}(\boldsymbol{A}\hat{\boldsymbol{\beta}} - \boldsymbol{b}) + o_{\boldsymbol{P}}(1)\end{aligned} \tag{36}$$

由引理 2 和引理 3 知

$$\sqrt{n}\boldsymbol{A}(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta}) \to_D N(0,\sigma^2\boldsymbol{A}\boldsymbol{B}^{-1}\boldsymbol{A}^{\mathrm{T}}) \tag{37}$$

且

$$n\boldsymbol{A}(\tilde{X}^{\mathrm{T}}\ \ \tilde{X})^{-1}\boldsymbol{A}^{\mathrm{T}} \to_P \boldsymbol{A}\boldsymbol{B}^{-1}\boldsymbol{A}^{\mathrm{T}} \tag{38}$$

再由式(37),即得

$$\sqrt{n}(\boldsymbol{A}\hat{\boldsymbol{\beta}} - \boldsymbol{b} + \boldsymbol{A}\boldsymbol{\beta} - \boldsymbol{b}) \to_D N(0,\sigma^2\boldsymbol{A}\boldsymbol{B}^{-1}\boldsymbol{A}^{\mathrm{T}})$$

所以,在原假设成立即 H_0 :$\boldsymbol{A}\boldsymbol{\beta}-\boldsymbol{b}=0$ 时,可得 $T_n \to_D \chi_k^2$.证毕.

定理 5 的证明 证明方法和定理 4 的证明相同,此处省略.

参考文献:

[1] ENGLE R,GRANGER C,RICE J,et al.Nonparametric Estimates of the Relation Between Weather and Electricity Sales[J].Journal of American Statistical Association,1986(81):310-320

[2] YOU J H,ZHOU Y,CHEN G M.Corrected Local Polynomial Estimation in Varying-coefficient Models with Measurement Errors

[J].The Canadian Journal of Statistics,2006(34):391-410

[3] LIAN H,HÄRDLE W,CARROLL R J.Estimation in a Semiparametric Partially Linear Errors-in-variables Model[J].The Annals of Statistics,1999(27):1519-1535

[4] CUI H J,LI R C.On Parameter Estimation for Semi-linear Errors-in-variable Models[J].Journal of Multivariable Analysis,1998(64):1-24

[5] CUI H J.Estimation in Partial Linear EV Models with Replicated Observations[J].Science China Mathematics,Series A.2004(34):467-482

[6] 赵培信,周小双.线性误差协变量下部分线性模型的约束统计推断[J].山东大学学报:理学版,2014,49(7):69-74

[7] YOU J H,XU Q F,ZHOU B.Statistical Inference for Partially Linear Regression Models with Measurement Errors[J].Chinese Annals of Mathematics,Series B,2008(29):207-222

[8] LI G R,XUE L G.Empirical Likelihood Confidence Region for the Parameter in Partially Linear Errors-in-variables Model[J].Communications in Statistics - Theory and Methods,2008(37):1552-1564

[9] WONG H,LIU F,CHEN M,et al.Empirical Likelihood Based Diagnostics for Heteroscedasticity in Partially Linear Errors-in-variables Models[J].Journal of Statistical Planning and Inference,2009(139):916 - 929

[10] LIU Q,XUE L G.Empirical Likelihood Confidence Regions of Parameters in Nonlinear EV Models under Missing Data[J].Acta Mathematica Scientia,Chinese Series A,2012(32):233-245

[11] LIANG H.Asymptotic Normality of Parametric Part in Partially Linear Models with Measurement Errors in the Nonparametric Part[J].Journal of Statistical Planning and Inference,2000(86):51-62

[12] HUANG Z S. Empirical Likelihood for the Parametric Part in Partially Linear Errors-in-function Models[J]. Statistics and Probability Letters,2012(82):63-66

[13] ZHU L X,CUI H J.A Semi-parametric Regression Model with Errors in Variables[J].Board of the Foundation of the Scandinavian Journal of Statistics,2003(30):429-442

[14] RAO C R,TOUTENBURG H,SHALABH,et al.Linear Models and Generalizations:Least Squares and Alternatives[M].Berlin:Springer,2008

[15] WEI C H. Statistical Inference for Restricted Partially Linear Varying Coeficient Errors-in-variables Models[J]. Journal of Statistical Planning and Inference,2012(142):2464-2472

[16] FAN J Q,TRUONG Y.Nonparametric Regression with Errors in Variables[J].The Annals of Statistics,1993(21):1900-1925

[17] AMEMIYA T.Advanced Econometrics[M].Boston:Harvard University Press,1985

[18] FAN J Q,HUANG T.Profile Likelihood Inferences on Semiparametric Varying-coefficient Partially Linear Models[J].Bernoulli,2005(11):1031-1057

[19] CHEN X,CUI H J.Empirical Likelihood Inference for Parameters in a Partially Linear Errors-in-variables Model[J].Statistics,2011(46):745-757

Statistical Inference for Partially Linear Models of Errors-in-function Under Restricted Conditions

LI Meng-han, XIA Xiao-chao

(School of Mathematics and Statistics,Chongqing University,Chongqing 401331,China)

Abstract: This paper considers a partially linear model of errors-in-function. Statistical inference for the parametric component under restrictions is investigated.Firstly,a restricted estimator for the parameter is derived and is demonstrated to be asymptotic normal.Then,a test procedure based on the generalized likelihood ratio statistic is proposed.It is proven that the Wilks phenomenon for errors in function still holds under some mild conditions. Finally,a simulation study is carried out to examine the consistency of the proposed restricted estimators and the validity of the test statistic.

Key words: partially linear model; error in function; statistical inference; restricted estimator; Wilks phenomenon