doi:10.3969/j.issn.1001-4616.2024.01.002

多元广义线性模型经验似然方法分析

朱春华[1],单苗慧[1],高启兵[2]

(1.南京审计大学统计与数据科学学院,江苏 南京 211815)
(2.南京师范大学数学科学学院,江苏 南京 210046)

[摘要]　针对多元广义线性模型,基于估计相关阵、广义估计方程和经验似然方法,本文构造出经验似然比统计量,此统计量能克服"工作相关阵"方法的误设定问题. 在一定的条件下,本文也获得了经验似然比统计量渐近 Wilks 性质,该结果可用作未知参数向量置信域的构造. 最后,通过数值模拟对所提方法的有效性进行验证.

[关键词]　多元广义线性模型,广义估计方程,经验似然,置信域

[中图分类号]O212.4　[文献标志码]A　[文章编号]1001-4616(2024)01-0007-07

Analysis of Empirical Likelihood Methods for Multivariate Generalized Linear Models

Zhu Chunhua[1],Shan Miaohui[1],Gao Qibing[2]

(1.School of Statistics and Data Science,Nanjing Audit University,Nanjing,211815,China)
(2.School of Mathematical Sciences,Nanjing Normal University,Nanjing,210046,China)

Abstract:For the generalized linear models with multivariate responses,based on the estimating correlation matrix,the generalized estimating equations and empirical likelihood methods,this paper constructs the empirical likelihood ratio statistics which can overcome the mistakenly specification caused by the method of "working correlation matrix". Under certain assumptions,this paper also obtains the asymptotic Wilks property of the empirical likelihood ratio statistics, which can be used to construct the confidence region of the unknown parameter. Last,the validity of the proposed method is verified through the numerical simulations.

Key words: multivariate generalized linear models, generalized estimating equations, empirical likelihood, confidence region

考虑多元响应变量广义线性模型,即设 $\boldsymbol{y}_i=(y_{i1},\cdots,y_{im})^{\mathrm{T}}$ 为独立的 m 元响应变量观测数据,$\boldsymbol{X}_i=(\boldsymbol{x}_{i1},\cdots,\boldsymbol{x}_{im})^{\mathrm{T}}$ 为已知的 $m\times p$ 阶设计矩阵,其中 $i=1,\cdots,n$ 和 $j=1,\cdots,m$,此处 n 和 m 分别表示观测个体数和个体重复观测的次数,且对每个 i 和 j 有

$$E(y_{ij})=\boldsymbol{h}(\boldsymbol{x}_{ij}^{\mathrm{T}}\boldsymbol{\beta}),\mathrm{Var}(\boldsymbol{x}_{ij}^{\mathrm{T}}\boldsymbol{\beta})=\sigma^2(\boldsymbol{x}_{ij}^{\mathrm{T}}\boldsymbol{\beta}),$$

式中,y_{ij} 和 $\boldsymbol{x}_{ij}$ 分别是一元响应变量和 p 元自变量,$\boldsymbol{\beta}$ 是 p 元未知参数向量且具有真值 $\boldsymbol{\beta}_0$,$\boldsymbol{h}:R\to R$ 和 $\sigma^2:R\to R$ 是充分光滑的已知均值和方差函数. 为获得未知参数估计,Liang 等[1]提出如下广义估计方程:

$$\sum_{i=1}^{n}\boldsymbol{X}_i^{\mathrm{T}}\boldsymbol{A}_i^{1/2}(\boldsymbol{\beta})\boldsymbol{R}_w^{-1}\boldsymbol{A}_i^{-1/2}(\boldsymbol{\beta})[\boldsymbol{y}_i-\boldsymbol{h}_i(\boldsymbol{\beta})]=0,$$

式中,$\boldsymbol{A}_i(\boldsymbol{\beta})=\mathrm{diag}[\sigma^2(\boldsymbol{x}_{i1}^{\mathrm{T}}\boldsymbol{\beta}),\cdots,\sigma^2(\boldsymbol{x}_{im}^{\mathrm{T}}\boldsymbol{\beta})]$,$\boldsymbol{h}_i(\boldsymbol{\beta})=[\boldsymbol{h}(\boldsymbol{x}_{i1}^{\mathrm{T}}\boldsymbol{\beta}),\cdots,\boldsymbol{h}(\boldsymbol{x}_{im}^{\mathrm{T}}\boldsymbol{\beta})]^{\mathrm{T}}$,$\boldsymbol{R}_w$ 表示多维响应变量 $\boldsymbol{y}_i$ 的工作相关阵,其可以根据经验或先验信息选取. 该广义估计方程可用于对均衡重复观测的纵向数据进行统计推断. 进一步研究发现,在一定条件下,基于上述广义估计方程的统计推断依然有效. 其他一些相关工作参见文献[2-5],更多有关广义线性模型的内容,建议参考文献[6]. 以上讨论都是基于事先选取好的工作相关阵进行讨论,Liang 等[1]指出工作相关阵的选取好坏会一定程度影响推断的效率,因此在推

收稿日期:2022-07-07.
基金项目:国家社科基金项目(21BTJ030).
通讯作者:朱春华,副教授,研究方向:近代回归分析. E-mail:zhuchunhua@ nau.edu.cn

断中如何利用响应向量重复测量数据内部相关信息进行有效推断是个值得进一步探讨的问题,本文将结合经验似然方法对其进行研究.

经验似然方法[7]是流行的非参数统计工具,已被广泛且成功应用于各种统计分析. 如 Xue 等[8]将经验似然应用于部分线性模型;Chen 等[9]将经验似然方法应用到一元广义线性模型;Yan 等[10]考虑独立相关矩阵下的广义线性模型广义估计方程经验似然方法,但在其方法中没有考虑响应向量内部相关信息,等等. 经验似然方法之所以被广大统计研究者和实际工作者青睐,主要是其能无缝地利用辅助信息、具有 Bartlett 可校正性、置信域的方向和形状完全由数据决定等优点. 本文中,我们将在经验似然方法中融入响应向量重复测量数据的内部相关信息,以期能提高推断效率,并讨论其理论性质. 另外,需要指出的是我们的方法对响应向量的相关阵结构不做任何参数化和依据先验信息或经验的设定,这样可以避免因相关阵的错误设定带来的推断效率的降低.

1　方法和主要结果

为引入本文经验似然方法,定义辅助随机变量

$$\boldsymbol{\eta}_{ni}(\boldsymbol{\beta})=\boldsymbol{X}_i^{\mathrm{T}}\boldsymbol{A}_i^{1/2}(\boldsymbol{\beta})\boldsymbol{R}_n^{-1}(\boldsymbol{\beta})\boldsymbol{A}_i^{-1/2}(\boldsymbol{\beta})[\boldsymbol{y}_i-\boldsymbol{h}_i(\boldsymbol{\beta})],$$

式中,$\boldsymbol{R}_n(\boldsymbol{\beta})=1/n\sum_{i=1}^{n}\boldsymbol{A}_i^{-1/2}(\boldsymbol{\beta})[\boldsymbol{y}_i-\boldsymbol{h}_i(\boldsymbol{\beta})][\boldsymbol{y}_i-\boldsymbol{h}_i(\boldsymbol{\beta})]^{\mathrm{T}}\boldsymbol{A}_i^{-1/2}(\boldsymbol{\beta})$,$\boldsymbol{R}_n=\boldsymbol{R}_n(\boldsymbol{\beta}_0)$. 据此建立对数经验似然比函数

$$LR(\boldsymbol{\beta})=\max\left\{\prod_{i=1}^{n}np_i \mid p_i\geqslant 0,\sum_{i=1}^{n}p_i=1,\sum_{i=1}^{n}p_i\boldsymbol{\eta}_{ni}(\boldsymbol{\beta})=0\right\}. \tag{1}$$

假设零向量在$\{\boldsymbol{\eta}_{ni},i=1,\cdots,n\}$凸包内,则基于拉格朗日乘子法对数经验似然比函数可表示为 $LR(\boldsymbol{\beta})=-\sum_{i=1}^{n}\log[1+\boldsymbol{t}^{\mathrm{T}}\boldsymbol{\eta}_{ni}(\boldsymbol{\beta})]$,其中 $\boldsymbol{t}$ 满足下列方程

$$\sum_{i=1}^{n}\frac{\boldsymbol{\eta}_{ni}(\boldsymbol{\beta})}{1+\boldsymbol{t}^{\mathrm{T}}\boldsymbol{\eta}_{ni}(\boldsymbol{\beta})}=0. \tag{2}$$

为叙述方便,以下我们做一下记号约定:设 $c,c_1,c_2,\cdots$ 等表示一般正的常数;$\|\cdot\|$ 表示向量的 Euclidean 模或矩阵的 Frobenius 模;记 $\boldsymbol{Q}_n=\sum_{i=1}^{n}\boldsymbol{X}_i^{\mathrm{T}}\boldsymbol{A}_i^{1/2}\boldsymbol{R}_0^{-1}\boldsymbol{A}_i^{1/2}\boldsymbol{X}_i$,$\boldsymbol{A}_i=\boldsymbol{A}_i(\boldsymbol{\beta}_0)$,$\boldsymbol{h}_i=\boldsymbol{h}_i(\boldsymbol{\beta}_0)$,$\boldsymbol{e}_i=\boldsymbol{y}_i-\boldsymbol{h}_i$ 表示零均值误差向量,$\boldsymbol{\eta}_{ni}=\boldsymbol{\eta}_{ni}(\boldsymbol{\beta}_0)$ 和 $\boldsymbol{\eta}_i=\boldsymbol{X}_i^{\mathrm{T}}\boldsymbol{A}_i^{1/2}\boldsymbol{R}_0^{-1}\boldsymbol{A}_i^{-1/2}\boldsymbol{e}_i$,$\lambda_n$ 表示矩阵 $\boldsymbol{S}_n=\sum_{i=1}^{n}\boldsymbol{X}_i^{\mathrm{T}}\boldsymbol{X}_i$ 的最小特征根.

为获得本文主要结果,我们需要以下假设:

A1. $\sup_{i\geqslant 1}\|\boldsymbol{X}_i\|<\infty$;

A2. 当 $n\to\infty$ 时,$\lambda_n\to\infty$;

A3. 对 $\boldsymbol{t}\in R$,$\sigma^2(\boldsymbol{t})>0$ 和 $\boldsymbol{h}(\boldsymbol{t})$ 连续可微,$\sup_{i\geqslant 1}E\|\boldsymbol{e}_i\|^{2+\delta}<\infty$,其中 $\delta>0$,且存在正常数 c_1 和 c_2 使得 $c_1\boldsymbol{I}_m\leqslant\boldsymbol{R}_0\leqslant c_2\boldsymbol{I}_m$. 其中 $\boldsymbol{R}_0$ 表示 $\boldsymbol{y}_i$ 未知的真实相关阵.

本文主要结果如下:

定理 1　若假设 A1~A3 成立,则有 $-2LR(\boldsymbol{\beta}_0)\to\chi^2(p)$,其中 $\chi^2(p)$ 表示自由度为 p 的卡方分布.

注 1　假设 A1 是广义线性模型渐近性质的必要条件;假设 A2 等同于对费希尔信息阵的最小特征根的假设,该假设已是广义线性模型渐近性质的研究中的最弱假设;A3 则是常用的一些光滑性和正则性假设. 总的来说以上假设是广义线性模型渐近性质研究中最弱和常用的假设条件;具体细节可参见文[2-5]中的讨论.

注 2　利用定理 1 的结果,我们可以构造未知参数 $\boldsymbol{\beta}$ 的置信域如下:

$$\{\boldsymbol{\beta}\mid -2LR(\boldsymbol{\beta})<\chi^2_{1-\alpha}(p)\},$$

式中,$\chi^2_{1-\alpha}(p)$ 表示自由度为 p 的卡方分布左侧 $1-\alpha$ 分位数,该置信域可用于未知参数向量的假设检验.

2　一些引理和主要结果证明

为证明定理 1,我们需要以下一系列引理.

引理 1 若假设 A1 ~ A3 成立,则有: $\| \boldsymbol{R}_n^{-1}-\boldsymbol{R}_0^{-1} \| =o_p(1)$.

证明 注意到 $\boldsymbol{R}_n=1/n \sum\limits_{i=1}^{n} \boldsymbol{A}_i^{-1/2}(\boldsymbol{\beta}_0)[\boldsymbol{y}_i-\boldsymbol{h}_i(\boldsymbol{\beta}_0)][\boldsymbol{y}_i-\boldsymbol{h}_i(\boldsymbol{\beta}_0)]^{\mathrm{T}}\boldsymbol{A}_i^{-1/2}(\boldsymbol{\beta}_0)=(r_{njk})_{j,k=1,\cdots,m}$,其中 $r_{njk}=1/n \sum\limits_{i=1}^{n} \tilde{e}_{ij}\tilde{e}_{ik}$, $\tilde{e}_{ij}=e_{ij}/\sigma_{ij}$, $\tilde{e}_{ik}=e_{ik}/\sigma_{ik}$, $j,k=1,\cdots,m$. 由假设 A1 和 A3 知 $E(\tilde{e}_{ij})=0$, $\sup\limits_{i\geqslant 1,j=1,\cdots,m} E|\tilde{e}_{ij}|^{2+\delta}<\infty$ 且 $\{\tilde{e}_{ij},j=1,\cdots,m\}$ 相互独立. 因此由柯西−施瓦兹不等式知:对每个固定的 $j,k=1,\cdots,m$ 有

$$\sum_{i=1}^{n} E|\tilde{e}_{ij}\tilde{e}_{ik}|^{1+\delta/2} i^{-(1+\delta/2)} \leqslant \sum_{i=1}^{n} \sqrt{E|\tilde{e}_{ij}|^{2+\delta}E|\tilde{e}_{ik}|^{2+\delta}}\, i^{-(1+\delta/2)}<\infty .$$

所以由文[11]中推论 3.1 可知 $r_{nij}\to 0$ 几乎处处成立. 由此可得引理 1 的结论.

引理 2 若假设 A1 ~ A3 成立,则有以下结果成立:

(i). $\left\| \boldsymbol{Q}_n^{-1/2} \sum\limits_{i=1}^{n} (\boldsymbol{\eta}_{ni}-\boldsymbol{\eta}_i) \right\| =o_p(1)$;

(ii). $\boldsymbol{Q}_n^{-1/2} \sum\limits_{i=1}^{n} \boldsymbol{\eta}_{ni}\to N(\boldsymbol{0},\boldsymbol{I}_p)$. 其中 $\boldsymbol{I}_p$ 为 p 阶单位矩阵.

证明 在假设 A1 ~ A3 下,由中心极限定理知 $\boldsymbol{Q}_n^{-1/2} \sum\limits_{i=1}^{n} \boldsymbol{\eta}_i\to N(\boldsymbol{0},\boldsymbol{I}_p)$. 因此若结论(i)成立,则结论(ii)也成立. 从而为证引理 2,只需证结论(i)成立即可.

在假设 A1 ~ A3 下,注意到 $E\boldsymbol{e}_i=0$,易见随机矩阵 $\sum\limits_{i=1}^{n} \boldsymbol{A}_i^{-1/2}\boldsymbol{e}_i\boldsymbol{u}^{\mathrm{T}}\boldsymbol{Q}_n^{-1/2}\boldsymbol{X}_i^{\mathrm{T}}\boldsymbol{A}_i^{1/2}$ 的每个元素方差有限. 另外

$$\operatorname{tr}\left[\boldsymbol{u}^{\mathrm{T}}\boldsymbol{Q}_n^{-1/2}\sum_{i=1}^{n}(\boldsymbol{\eta}_{ni}-\boldsymbol{\eta}_i)\right]=\operatorname{tr}\left[(\boldsymbol{R}_n^{-1}-\boldsymbol{R}_0)\sum_{i=1}^{n}(\boldsymbol{A}_i^{-1/2}\boldsymbol{e}_i\boldsymbol{u}^{\mathrm{T}}\boldsymbol{Q}_n^{-1/2}\boldsymbol{X}_i^{\mathrm{T}}\boldsymbol{A}_i^{1/2})\right] ,$$

其中 $\boldsymbol{u}$ 为任意单位向量. 此式结合引理 1,(i)可得证. 从而完成引理 2 的证明.

引理 3 若假设 A1 ~ A3 成立,则有以下结果成立:

$$\left\|\boldsymbol{Q}_n^{-1/2}\left(\sum_{i=1}^{n}\boldsymbol{\eta}_{ni}\boldsymbol{\eta}_{ni}^{\mathrm{T}}\right)\boldsymbol{Q}_n^{-1/2}-\boldsymbol{I}_p\right\|=o_p(1),$$

式中,$\boldsymbol{I}_p$ 为 p 阶单位矩阵.

证明 注意到有下列分解表达式

$$\boldsymbol{Q}_n^{-1/2}\left(\sum_{i=1}^{n}\boldsymbol{\eta}_{ni}\boldsymbol{\eta}_{ni}^{\mathrm{T}}\right)\boldsymbol{Q}_n^{-1/2}-\boldsymbol{I}_p=\boldsymbol{J}_{n1}-\boldsymbol{I}_p+\boldsymbol{J}_{n2}+\boldsymbol{J}_{n3}+\boldsymbol{J}_{n4},$$

式中,$\boldsymbol{J}_{n1}=\boldsymbol{Q}_n^{-1/2}\left(\sum\limits_{i=1}^{n}\boldsymbol{\eta}_i\boldsymbol{\eta}_i^{\mathrm{T}}\right)\boldsymbol{Q}_n^{-1/2}$, $\boldsymbol{J}_{n2}=\boldsymbol{Q}_n^{-1/2}\left(\sum\limits_{i=1}^{n}(\boldsymbol{\eta}_{ni}-\boldsymbol{\eta}_i)\boldsymbol{\eta}_i^{\mathrm{T}}\right)\boldsymbol{Q}_n^{-1/2}$,

$$\boldsymbol{J}_{n3}=\boldsymbol{Q}_n^{-1/2}\left(\sum_{i=1}^{n}\boldsymbol{\eta}_i(\boldsymbol{\eta}_{ni}-\boldsymbol{\eta}_i)^{\mathrm{T}}\right)\boldsymbol{Q}_n^{-1/2},\boldsymbol{J}_{n4}=\boldsymbol{Q}_n^{-1/2}\left(\sum_{i=1}^{n}(\boldsymbol{\eta}_{ni}-\boldsymbol{\eta}_i)(\boldsymbol{\eta}_{ni}-\boldsymbol{\eta}_i)^{\mathrm{T}}\right)\boldsymbol{Q}_n^{-1/2}.$$

首先证明 $\|\boldsymbol{J}_{n1}-\boldsymbol{I}_p\|=o_p(1)$. 由 $\boldsymbol{J}_{n1}$ 的对称性知此等价于需要证明依概率 1 有

$$\boldsymbol{u}^{\mathrm{T}}\boldsymbol{J}_{n1}\boldsymbol{u}=\sum_{i=1}^{n}\boldsymbol{d}_{ni}^{\mathrm{T}}\boldsymbol{e}_i^{\mathrm{T}}\boldsymbol{e}_i\boldsymbol{d}_{ni}\to 1,$$

式中,$\boldsymbol{u}$ 为任意单位向量及 $\boldsymbol{d}_{ni}=\boldsymbol{u}^{\mathrm{T}}\boldsymbol{Q}_n^{-1/2}\boldsymbol{X}_i^{\mathrm{T}}\boldsymbol{A}_i^{1/2}\boldsymbol{R}_0^{-1}\boldsymbol{A}_i^{-1/2}$. 记 $d_n=\max\limits_{1\leqslant i\leqslant n}|\boldsymbol{d}_{ni}|$,由 A1 ~ A3 知 $d_n\leqslant c\|\boldsymbol{Q}_n^{-1/2}\|=o(1)$ 和 $\sum\limits_{i=1}^{n}\|\boldsymbol{d}_{ni}\|^2\leqslant c$. 另易见 $E(\boldsymbol{u}^{\mathrm{T}}\boldsymbol{J}_{n1}\boldsymbol{u})=1$,且对任意 $\varepsilon>0$,由马尔可夫不等式有

$$\sum_{i=1}^{n}E[\boldsymbol{d}_{ni}^{\mathrm{T}}\boldsymbol{e}_i^{\mathrm{T}}\boldsymbol{e}_i\boldsymbol{d}_{ni}I(\boldsymbol{d}_{ni}^{\mathrm{T}}\boldsymbol{d}_{ni}>\varepsilon)]\leqslant\sum_{i=1}^{n}\|d_{ni}\|^2E[\|\boldsymbol{e}_i\|^2I(\|\boldsymbol{e}_i\|^2>\varepsilon/\|d_{ni}\|^2)]\leqslant$$

$$\sup_{i\geqslant 1}E\|\boldsymbol{e}_i\|^2\varepsilon^{-\delta/2}d_n^{\delta/2}\sum_{i=1}^{n}\|d_{ni}\|^2\leqslant cd_n^{\delta/2}\to 0.$$

因此由文[12]中第 358 页推论可得 $\|\boldsymbol{J}_{n1}-\boldsymbol{I}_p\|=o_p(1)$.

下面证明 $\|\boldsymbol{J}_{nk}\|=o_p(1)$, $k=2,3,4$. 首先证明 $\|\boldsymbol{J}_{n2}\|=o_p(1)$.

由假设 A1 和 A3,我们可知 $\sup\limits_{i\geqslant 1}\|\boldsymbol{A}_i^{1/2}\|^2\ \|\boldsymbol{R}_0^{-1}\|<\infty$ 以及

$$E\left(\sum_{i=1}^{n}\|\boldsymbol{Q}_n^{-1/2}\boldsymbol{X}_i^{\mathrm{T}}\|^2\|\boldsymbol{A}_i^{-1/2}\boldsymbol{e}_i\|^2\right)\leqslant c_1\sum_{i=1}^{n}\|\boldsymbol{Q}_n^{-1/2}\boldsymbol{X}_i^{\mathrm{T}}\|^2=c_1\mathrm{tr}\left[\boldsymbol{Q}_n^{-1/2}\left(\sum_{i=1}^{n}\boldsymbol{X}_i^{\mathrm{T}}\boldsymbol{X}_i\right)\boldsymbol{Q}_n^{-1/2}\right]\leqslant c_2.$$

由此结合引理 1 可得

$$\|\boldsymbol{J}_{n2}\|=\left\|\boldsymbol{Q}_n^{-1/2}\left(\sum_{i=1}^{n}(\boldsymbol{\eta}_{ni}-\boldsymbol{\eta}_i)\boldsymbol{\eta}_i^{\mathrm{T}}\right)\boldsymbol{Q}_n^{-1/2}\right\|=\left\|\boldsymbol{Q}_n^{-1/2}\left(\sum_{i=1}^{n}\boldsymbol{X}_i^{\mathrm{T}}\boldsymbol{A}_i^{1/2}(\boldsymbol{R}_n^{-1}-\boldsymbol{R}_0^{-1})\boldsymbol{A}_i^{-1/2}\boldsymbol{e}_i\boldsymbol{e}_i^{\mathrm{T}}\boldsymbol{A}_i^{-1/2}\boldsymbol{R}_0^{-1}\boldsymbol{A}_i^{1/2}\boldsymbol{X}_i\right)\boldsymbol{Q}_n^{-1/2}\right\|\leqslant$$

$$\|(\boldsymbol{R}_n^{-1}-\boldsymbol{R}_0^{-1})\|\left(\sum_{i=1}^{n}\|\boldsymbol{Q}_n^{-1/2}\boldsymbol{X}_i^{\mathrm{T}}\|^2\|\boldsymbol{A}_i^{-1/2}\boldsymbol{e}_i\|^2\right)\sup_{i\geqslant 1}\|\boldsymbol{A}_i^{1/2}\|^2\|\boldsymbol{R}_0^{-1}\|=O_p(n^{-1/2})=o_p(1).$$

同理可得 $\|\boldsymbol{J}_{nk}\|=o_p(1)$, $k=3,4$. 由以上已证的结果立即可得定理结论成立.

引理 4　若假设 A1～A3 成立,则有以下结果成立:

$$\max_{i\geqslant 1}\|\boldsymbol{Q}_n^{-1/2}\boldsymbol{\eta}_{ni}\|=o_p(1).$$

证明　易见 $\max_{i\geqslant 1}\|\boldsymbol{Q}_n^{-1/2}\boldsymbol{\eta}_{ni}\|\leqslant\max_{i\geqslant 1}\|\boldsymbol{Q}_n^{-1/2}\boldsymbol{\eta}_i\|+\max_{i\geqslant 1}\|\boldsymbol{Q}_n^{-1/2}(\boldsymbol{\eta}_{ni}-\boldsymbol{\eta}_i)\|$. 首先由假设 A1～A3 知有 $\omega_n=\sum_{i=1}^{n}\|\boldsymbol{Q}_n^{-1/2}\boldsymbol{X}_i^{\mathrm{T}}\|^2=O(1)$,而

$$P(\max_{i\geqslant 1}\|\boldsymbol{Q}_n^{-1/2}\boldsymbol{\eta}_i\|>\varepsilon)\leqslant\sum_{i=1}^{n}P(\|\boldsymbol{Q}_n^{-1/2}\boldsymbol{\eta}_i\|>\varepsilon)\leqslant\varepsilon^{-(2+\delta)}\sup_{i\geqslant 1}E\|\boldsymbol{e}_i\|^{2+\delta}\|\boldsymbol{Q}_n^{-1/2}\|^{\delta}\omega_n=o(1),$$

结合假设 A1～A3 即得 $\max_{i\geqslant 1}\|\boldsymbol{Q}_n^{-1/2}\boldsymbol{\eta}_i\|=o_p(1)$. 类似处理方法并结合引理 1 的结论可得 $\max_{i\geqslant 1}\|\boldsymbol{Q}_n^{-1/2}(\boldsymbol{\eta}_{ni}-\boldsymbol{\eta}_i)\|=o_p(1)$. 从而引理 4 结论得证.

引理 5　若假设 A1～A3 成立,则有以下结果成立:

$$\|\boldsymbol{Q}_n^{1/2}\boldsymbol{t}\|=O_p(1),$$

式中,$\boldsymbol{t}$ 满足 $\sum_{i=1}^{n}\frac{\boldsymbol{\eta}_{ni}}{1+\boldsymbol{t}^{\mathrm{T}}\boldsymbol{\eta}_{ni}}=0$.

证明　令 $\boldsymbol{t}=\|\boldsymbol{t}\|\boldsymbol{u}$,其中 $\|\boldsymbol{u}\|=1$,利用 $|\boldsymbol{t}^{\mathrm{T}}\boldsymbol{\eta}_{ni}|\leqslant\|\boldsymbol{t}\|\ \|\boldsymbol{Q}_n^{1/2}\boldsymbol{u}\|\ \|\boldsymbol{Q}_n^{-1/2}\boldsymbol{\eta}_{ni}\|$,对式(2)进行放大并整理得

$$\|\boldsymbol{t}\|\sum_{i=1}^{n}(\boldsymbol{u}^{\mathrm{T}}\boldsymbol{\eta}_{ni})^2\leqslant\|\boldsymbol{Q}_n^{1/2}\boldsymbol{u}\|\ \|\boldsymbol{Q}_n^{-1/2}\sum_{i=1}^{n}\boldsymbol{\eta}_{ni}\|+\|\boldsymbol{t}\|\ \|\boldsymbol{Q}_n^{1/2}\boldsymbol{u}\|^2\|\boldsymbol{Q}_n^{-1/2}\sum_{i=1}^{n}\boldsymbol{\eta}_{ni}\|\max_{1\leqslant i\leqslant n}\|\boldsymbol{Q}_n^{-1/2}\boldsymbol{\eta}_{ni}\|.$$

另由引理 3 知:对任意 $0<\varepsilon<1$,依概率 1 有 $\|\boldsymbol{t}\|\sum_{i=1}^{n}(\boldsymbol{u}^{\mathrm{T}}\boldsymbol{\eta}_{ni})^2\geqslant(1-\varepsilon)\|\boldsymbol{t}\|\ \|\boldsymbol{Q}_n^{1/2}\boldsymbol{u}\|^2$.

因此,我们有 $(1-\varepsilon)\|\boldsymbol{t}\|\ \|\boldsymbol{Q}_n^{1/2}\boldsymbol{u}\|\leqslant\|\boldsymbol{t}\|\ \|\boldsymbol{Q}_n^{1/2}\boldsymbol{u}\|\ \left\|\boldsymbol{Q}_n^{-1/2}\sum_{i=1}^{n}\boldsymbol{\eta}_{ni}\right\|\max_{1\leqslant i\leqslant n}\|\boldsymbol{Q}_n^{-1/2}\boldsymbol{\eta}_{ni}\|+\left\|\boldsymbol{Q}_n^{-1/2}\sum_{i=1}^{n}\boldsymbol{\eta}_{ni}\right\|$.

此式与引理 2 的(ii)和引理 4 结合,我们有

$$(1-\varepsilon)\|\boldsymbol{t}\|\ \|\boldsymbol{Q}_n^{1/2}\boldsymbol{u}\|\leqslant\|\boldsymbol{t}\|\ \|\boldsymbol{Q}_n^{1/2}\boldsymbol{u}\|O_p(1)o_p(1)+O_p(1).$$

这可推得引理 5 结论.

定理 1 的证明　由引理 4 和引理 5,我们有 $\max_{1\leqslant i\leqslant n}|\boldsymbol{t}^{\mathrm{T}}\boldsymbol{\eta}_{ni}|=o_p(1)$. 因此对对数经验似然比统计量 $LR(\boldsymbol{\beta}_0)=-\sum_{i=1}^{n}\log[1+\boldsymbol{t}^{\mathrm{T}}\boldsymbol{\eta}_{ni}]$ 关于 $\boldsymbol{t}^{\mathrm{T}}\boldsymbol{\eta}_{ni}$ 进行泰勒公式展开得

$$-2LR(\boldsymbol{\beta}_0)=2\sum_{i=1}^{n}[\boldsymbol{t}^{\mathrm{T}}\boldsymbol{\eta}_{ni}-(\boldsymbol{t}^{\mathrm{T}}\boldsymbol{\eta}_{ni})^2/2]+o_p(1).\tag{3}$$

由式(2)可得

$$0=\sum_{i=1}^{n}\frac{\boldsymbol{t}^{\mathrm{T}}\boldsymbol{\eta}_{ni}}{1+\boldsymbol{t}^{\mathrm{T}}\boldsymbol{\eta}_{ni}}=\sum_{i=1}^{n}\boldsymbol{t}^{\mathrm{T}}\boldsymbol{\eta}_{ni}-\boldsymbol{t}^{\mathrm{T}}\left(\sum_{i=1}^{n}\boldsymbol{\eta}_{ni}\boldsymbol{\eta}_{ni}^{\mathrm{T}}\right)\boldsymbol{t}+\sum_{i=1}^{n}\frac{(\boldsymbol{t}^{\mathrm{T}}\boldsymbol{\eta}_{ni})^3}{1+\boldsymbol{t}^{\mathrm{T}}\boldsymbol{\eta}_{ni}}\tag{4}$$

和

$$0=\sum_{i=1}^{n}\frac{\boldsymbol{\eta}_{ni}}{1+\boldsymbol{t}^{\mathrm{T}}\boldsymbol{\eta}_{ni}}=\sum_{i=1}^{n}\boldsymbol{\eta}_{ni}-\left(\sum_{i=1}^{n}\boldsymbol{\eta}_{ni}\boldsymbol{\eta}_{ni}^{\mathrm{T}}\right)\boldsymbol{t}+\sum_{i=1}^{n}\frac{\boldsymbol{\eta}_{ni}(\boldsymbol{t}^{\mathrm{T}}\boldsymbol{\eta}_{ni})^2}{1+\boldsymbol{t}^{\mathrm{T}}\boldsymbol{\eta}_{ni}}.\tag{5}$$

由(4)及引理 3、引理 4 和引理 5 得

$$\sum_{i=1}^{n}(\boldsymbol{t}^{\mathrm{T}}\boldsymbol{\eta}_{ni})^2=\sum_{i=1}^{n}\boldsymbol{t}^{\mathrm{T}}\boldsymbol{\eta}_{ni}+\sum_{i=1}^{n}\frac{(\boldsymbol{t}^{\mathrm{T}}\boldsymbol{\eta}_{ni})^3}{1+\boldsymbol{t}^{\mathrm{T}}\boldsymbol{\eta}_{ni}}=\sum_{i=1}^{n}\boldsymbol{t}^{\mathrm{T}}\boldsymbol{\eta}_{ni}+o_p(1). \tag{6}$$

将(6)代入(3)得

$$-2LR(\boldsymbol{\beta}_0)=\sum_{i=1}^{n}(\boldsymbol{t}^{\mathrm{T}}\boldsymbol{\eta}_{ni})^2+o_p(1). \tag{7}$$

另由(5)可得

$$\boldsymbol{t}=\left(\sum_{i=1}^{n}\boldsymbol{\eta}_{ni}\boldsymbol{\eta}_{ni}^{\mathrm{T}}\right)^{-1}\sum_{i=1}^{n}\boldsymbol{\eta}_{ni}+\left(\sum_{i=1}^{n}\boldsymbol{\eta}_{ni}\boldsymbol{\eta}_{ni}^{\mathrm{T}}\right)^{-1}\sum_{i=1}^{n}\frac{\boldsymbol{\eta}_{ni}(\boldsymbol{t}^{\mathrm{T}}\boldsymbol{\eta}_{ni})^2}{1+\boldsymbol{t}^{\mathrm{T}}\boldsymbol{\eta}_{ni}},$$

将上式代入(7)并结合引理3、引理4和引理5得

$$-2LR(\boldsymbol{\beta}_0)=\left(\boldsymbol{Q}_n^{-1/2}\sum_{i=1}^{n}\boldsymbol{\eta}_{ni}\right)^{\mathrm{T}}\left(\boldsymbol{Q}_n^{-1/2}\sum_{i=1}^{n}\boldsymbol{\eta}_{ni}\boldsymbol{\eta}_{ni}^{\mathrm{T}}\boldsymbol{Q}_n^{-1/2}\right)\left(\boldsymbol{Q}_n^{-1/2}\sum_{i=1}^{n}\boldsymbol{\eta}_{ni}\right)+o_p(1).$$

此式结合引理2(ii)和引理4,我们有$-2LR(\boldsymbol{\beta}_0)\xrightarrow{d}\chi^2(p)$.

3 数值模拟

在这部分,我们提供一些数值模拟研究本文提出的方法在有限样本下的表现,主要从覆盖概率及置信椭球方面与三类广义估计方程的经验似然结果进行比较研究,考虑下列非线性边际模型:

$$h_{ij}=E(y_{ij})=\frac{\exp(\boldsymbol{x}_{ij}^{\mathrm{T}}\boldsymbol{\beta}_0)}{1+\exp(\boldsymbol{x}_{ij}^{\mathrm{T}}\boldsymbol{\beta}_0)},\mathrm{Var}(y_{ij})=h_{ij}(1-h_{ij}),j=1,2,3,i=1,\cdots,n,$$

式中,参数$\boldsymbol{\beta}_0=(-1,0.5,0.2)^{\mathrm{T}}$,$\boldsymbol{x}_{ij}=(x_{ij1},x_{ij2},x_{ij3})^{\mathrm{T}}$来自3维正态分布,边际均值和方差分别为0和0.16,且分量间两两相关系数为0.5. 另外响应变量来自3维正态分布,分量间两两相关系数也取为0.5. 我们考虑以下4种广义估计方程下的经验似然方法(分别记为GEE1-EL,GEE2-EL,GEE3-EL和GEE4-EL)和一种经典广义估计方程方法(记为GEE5):

GEE1:$\sum_{i=1}^{n}\boldsymbol{X}_i^{\mathrm{T}}[\boldsymbol{y}_i-\boldsymbol{h}_i(\boldsymbol{\beta})]=0$;

GEE2:$\sum_{i=1}^{n}\boldsymbol{X}_i^{\mathrm{T}}\boldsymbol{A}_i(\boldsymbol{\beta})[\boldsymbol{y}_i-\boldsymbol{h}_i(\boldsymbol{\beta})]=0$;

GEE3:$\sum_{i=1}^{n}\boldsymbol{X}_i^{\mathrm{T}}\boldsymbol{A}_i^{1/2}(\boldsymbol{\beta})\boldsymbol{R}_0\boldsymbol{A}_i^{-1/2}(\boldsymbol{\beta})[\boldsymbol{y}_i-\boldsymbol{h}_i(\boldsymbol{\beta})]=0$;

GEE4:$\sum_{i=1}^{n}\boldsymbol{X}_i^{\mathrm{T}}\boldsymbol{A}_i^{1/2}(\boldsymbol{\beta})\boldsymbol{R}_n^{-1}(\boldsymbol{\beta})\boldsymbol{A}_i^{-1/2}(\boldsymbol{\beta})[\boldsymbol{y}_i-\boldsymbol{h}_i(\boldsymbol{\beta})]=0$;

GEE5:$\sum_{i=1}^{n}\boldsymbol{X}_i^{\mathrm{T}}\boldsymbol{A}_i^{1/2}(\boldsymbol{\beta})\hat{\boldsymbol{R}}_n^{-1}\boldsymbol{A}_i^{-1/2}(\boldsymbol{\beta})[\boldsymbol{y}_i-\boldsymbol{h}_i(\boldsymbol{\beta})]=0$,

式中,$\hat{\boldsymbol{R}}_n=1/n\sum_{i=1}^{n}\boldsymbol{A}_i^{-1/2}(\tilde{\boldsymbol{\beta}}_n)[\boldsymbol{y}_i-\boldsymbol{h}_i(\tilde{\boldsymbol{\beta}}_n)][\boldsymbol{y}_i-\boldsymbol{h}_i(\tilde{\boldsymbol{\beta}}_n)]^{\mathrm{T}}\boldsymbol{A}_i^{-1/2}(\tilde{\boldsymbol{\beta}}_n)$,此处$\tilde{\boldsymbol{\beta}}_n$是GEE1的解.

注 GEE1相当于工作相关阵取为单位阵的广义估计方程;GEE2相当于基于最小二乘方法所得的广义估计方程;GEE3相当于用真实相关阵建立的广义估计方程,实际中由于真实相关阵未知而不可用,此处是为了模拟比较而采用;GEE4是本文方法所基于的广义估计方程,其中$\boldsymbol{R}_n(\boldsymbol{\beta})$充当GEE3中真实相关阵的角色,据我们所知该方程求解极其复杂,但经验似然方法置信域的构造并不需要求解未知参数估计(这也是经验似然方法的一大优势),所以不影响本文方法的使用;GEE5是一种两步方法的广义估计方程,即先基于GEE1参数的初始估计构造相关阵的估计,然后基于此估计再构造广义估计方程,Wang[13]在发散维协变情形系统地讨论估计相关阵和广义估计方程的渐近理论.

模拟中样本容量n分别取为150,300,450,模拟重复5 000次,我们计算5 000次模拟中置信域覆盖参数真值$\boldsymbol{\beta}_0$比例. 另从5 000次模拟中选取一个样本,画出GEE1-EL,GEE2-EL,GEE3-EL,GEE4-EL和GEE5(基于Wald型检验统计量构造置信椭球)方法的参数95%的置信椭球,为能在平面上展示,我们固定第3个参数值在真值处并画出另两个参数的置信椭圆,同时我们计算该样本下5种方法下置信椭圆精

度(置信椭圆的面积占比分别在横轴[-3,1]和纵轴[-1,3]上等间隔取 30 个点,逐点计算 5 种方法的枢轴量的值,统计落在置信椭圆内点在全部 900 个点中的占比).

5 种情形参数的 95% 置信域的覆盖概率结果见表 1.

表 1　参数 95%覆盖概率模拟

Table 1　Simulations for 95% coverage probability of parameter

n	GEE1-EL	GEE2-EL	GEE3-EL	GEE4-EL	GEE5
150	0.940 4	0.940 0	0.944 6	0.944 6	0.935 8
300	0.944 6	0.947 2	0.949 8	0.948 4	0.947 0
450	0.944 6	0.943 0	0.949 8	0.949 8	0.943 8

从表 1 可以看出,随样本量增加,5 种情形下覆盖概率趋近于 95%. 此外,GEE3-EL 的结果即真实的相关阵模拟效果最好,GEE4-EL 的结果即本文提出的方法的模拟效果仅次于 GEE3-EL 的结果,且随着样本量的增加,GEE4-EL 与 GEE3-EL 的模拟效果几乎接近,这一点也说明在真实相关阵未知情况下,本文方法具有较好的可靠性,另外与 GEE5 方法下 Wald 型统计量构造的传统置信域比较,GEE4-EL 方法在各种样本量下也有较好表现. 针对选取的一个样本,参数的 95%置信椭圆如图 1 所示,其中“+”符号点标记表示真参数值点,“ * ”符号点标记表示基于 GEE5 所得参数的估计值点.

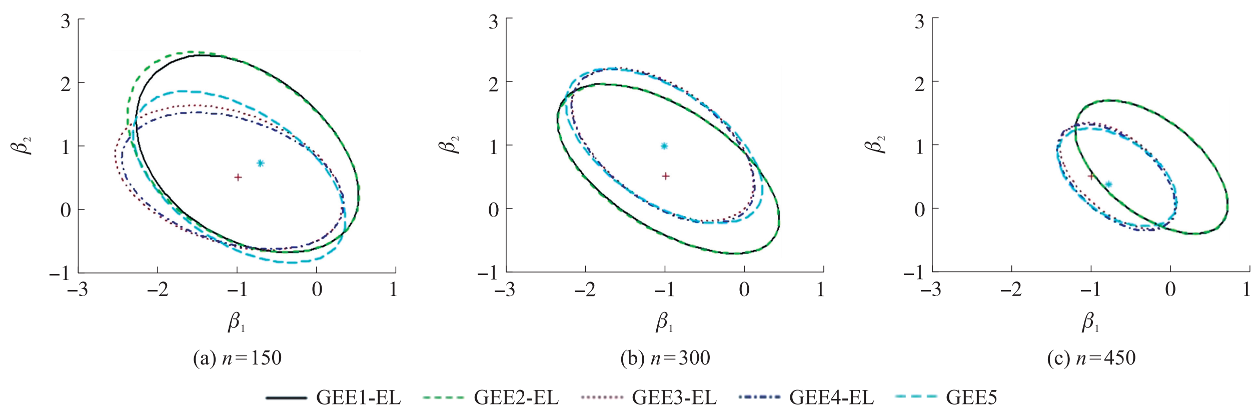

图 1　参数置信椭圆

Fig. 1　Confidence ellipse of parameter

可以看出,随着样本量增加,5 种情形对应的置信椭圆都有变小的趋势,另 GEE3-EL,GEE4-EL 和 GEE5 方法置信椭圆较小且较接近,特别是随着样本量增大,三者椭圆几乎重合,而 GEE1-EL 和 GEE2-EL 方法的置信椭圆相对较大,这说明考虑正确(GEE3-EL,GEE4-EL 和 GEE5 方法的意义下)相关信息有助于改善统计推断效果,不过随着样本量增加,GEE1-EL 和 GEE2-EL 方法的结果与 GEE3-EL,GEE4-EL 和 GEE5 方法的结果差异也呈现出缩小趋势. 同样,针对上述选取的样本,表 2 中参数的 95%覆盖精度即 5 种方法下置信椭圆面积占比也支持类似的结论. 总的来说,本文所提方法在置信域的可靠性和精度方面都有不错的表现.

表 2　参数 95%覆盖精度模拟

Table 2　Simulations for 95% coverage accuracy of parameter

n	GEE1-EL	GEE2-EL	GEE3-EL	GEE4-EL	GEE5
150	0.368 9	0.386 7	0.281 1	0.260 0	0.285 6
300	0.276 7	0.277 8	0.217 8	0.215 6	0.230 0
450	0.148 9	0.147 8	0.094 4	0.097 8	0.094 4

4　结论

本文考虑多维响应变量广义线性模型经验似然方法. 我们的主要贡献体现在以下两个方面:第一、在本文的方法中,我们将响应向量的相关阵信息融入到经验似然比统计量的构造中,用以构造未知参数的置信域,以期提高未知参数置信域的可靠性和精度;这里,我们需要指出的是在本文的方法中对响应向量的相关信息无需做任何参数化假定,这样可以避免参数化相关阵错误设定的问题出现. 第二、在费希尔信息

阵最弱的假定和一些正则性假设下,本文证明了经验似然比统计量的渐近 Wilks 性质. 此外,数值模拟结果显示本文所提出方法的结果在有限样本量下置信域的可靠性和精度方面的表现好于工作相关阵选取与真实相关阵不一致的情形,与真实相关阵情形表现极为相似,说明本文方法的有效性.

[参考文献]

[1] LIANG K Y,ZEGER S L. Longitudinal data analysis using generalized linear models[J]. Biometrika,1986,73(1):13-22.

[2] CHEN K,HU I,YING Z. Strong consistency of maximum quasi-likelihood estimators in generalized linear models with fixed and adaptive design[J]. The annals of statistics,1999,27:1155-1163.

[3] CHEN X,CHEN X R. Adaptive quasi-likelihood estimates in generalized linear models[J]. Science in China seriers A mathematics,2005,48(6):829-846.

[4] YIN C M,ZHAO L C. Asymptotic normality and strong consistency of maximum quasi-likelihhod in generalized linear models[J]. Science in China series A mathematics,2006,49:145-157.

[5] GAO Q B,LIN J G,WU Y H,et al. Asymptotic normality of maximum quasi-likelihood estimators in generalized linear models with adaptive design[J]. Statistics,2012,46(6):833-846.

[6] MCCULLAGH P,NELDER J A. Generalized linear models[M]. 2nd ed. New York:Chapman & Hall,1989.

[7] OWEN A B. Empirical likelihood[M]. New York:Chapman and Hall-CRC,2001.

[8] XUE L G,ZHU L X. Empirical likelihood for a varying coefficient model with longitudinal data[J]. Journal of American Statistical Association,2007,102(478):642-654.

[9] CHEN X,CUI H J. Empirical likelihood inference for parameters in a partially linear errors-in-variables model[J]. Statistics,2012,46(6):745-757.

[10] YAN L,CHEN X. Empirical likelihood for generalized linear models with fixed and adaptive designs[J]. Statistics,2015,49(5):978-988.

[11] 林正炎,陆传荣,苏中根. 概率极限理论基础[M]. 北京:高等教育出版社,1999.

[12] CHOW Y S,TEICHER H. Probability theory:independence,interchangeability,martigales[M]. New York:Springer,1988.

[13] WANG L. GEE analysis of clustered binarudata with diverging number of covariates[J]. The annals of statistics,2011,39(1):389-417.

[责任编辑:陆炳新]