带有线性约束的指数族非线性回归模型

赵慧秀[1], 周秀轻[2]

(1. 南京理工大学理学院, 江苏 南京 210094)
(2. 南京师范大学数学科学学院, 江苏 南京 210046)

[摘要] 给出了带有线性约束的指数族非线性回归模型的最大似然估计量的随机展开, 所得到的随机展开式由相互独立的正态变量以及曲率立体阵表示, 使用简洁方便.

[关键词] 指数族非线性回归模型, 曲率, 随机展开, 立体阵
[中图分类号] O212.1 [文献标识码] A [文章编号] 1001-4616(2010)01-0001-05

Exponential Family Nonlinear Regression Models With Linear Restriction

Zhao Huixiu[1], Zhou Xiuqing[2]

(1. College of Science, Nanjing University of Science and Technology, Nanjing 210094, China)
(2. School of Mathematical Sciences, Nanjing Normal University, Nanjing 210046, China)

Abstract Exponential family nonlinear regression models with linear restriction are discussed in this paper. A set of stochastic expansions are given for the restricted maximum likelihood estimator. The stochastic expansions given here depend only on the independent normal random variables and curvature arrays, therefore it is convenient to use them.

Key words nonlinear exponential family regression models, curvature, stochastic expansions, arrays

1 模型

无约束的指数族非线性回归模型在文[1-7]中已有很多的论述, 本文研究的是带有线性约束的指数族非线性回归模型. 下面介绍一些有关概念.

假设 $\boldsymbol{y} = (\boldsymbol{y}_1, \boldsymbol{y}_2, \cdots, \boldsymbol{y}_n)^{\mathrm{T}}$ 的各分量相互独立, 且具有指数族的概率密度

$$p(\boldsymbol{y}; \theta, \phi) = \exp\left[\sum_{i=1}^{n}[\phi(y_i\theta_i - b(\theta_i) - c(y_i)) - \frac{1}{2}s(y_i, \phi)]\right], \tag{1}$$

其中 $\theta = (\theta_1, \cdots, \theta_n)^{\mathrm{T}}$ 为自然参数; ϕ 为已知的尺度参数; $b(\cdot)$, $c(\cdot)$, $s(\cdot, \cdot)$ 为已知的函数. 假定有兴趣的参数为 $\beta = (\beta_1, \cdots, \beta_p) \in \boldsymbol{B}$, $\boldsymbol{B}$ 为 p 维实数空间的子集, 其中 $p < n$, 且模型(1)满足条件:

$$\eta_i \triangleq g(\mu_i) = f(x_i; \beta) \triangleq f_i(\beta), \ i = 1, \cdots, n, \tag{2}$$

其中 $\mu_i = E(y_i)$, g 为已知的单调函数, f 为已知的回归函数, 而 $x_1, \cdots, x_n$ 为自变量, 则上述模型即为指数族非线性模型, 简记为模型 ENM(β, ϕ).

现记 $b(\theta) = \sum_{i=1}^{n} b(\theta_i)$, $\dot{\boldsymbol{b}}(\theta) = (\dot{\boldsymbol{b}}(\theta_1), \cdots, \dot{\boldsymbol{b}}(\theta_n))^{\mathrm{T}}$, $\ddot{b}(\theta) = \mathrm{diag}(\ddot{b}(\theta_1), \cdots, \ddot{b}(\theta_n))$, $\phi \triangleq \sigma^{-2}$, 其中 $\dot{b}(\theta_i)$ 和 $\ddot{b}(\theta_i)$ 分别为 $b(\theta_i)$ 的前二阶导数. 又记 $E(\boldsymbol{y}) = \mu(\theta)$, $\mathrm{var}(\boldsymbol{y}) = \sigma^2\boldsymbol{V}(\theta)$, 则由指数族的性质可得 $\mu(\theta) = \dot{\boldsymbol{b}}(\theta)$, $\boldsymbol{V}(\theta) = \ddot{\boldsymbol{b}}(\theta)$, 由(2)式可知, θ 为 β 的函数, 记为 $\theta = \theta(\beta)$. 由此又知 $\mu(\theta)$ 及 $\boldsymbol{V}(\theta)$ 也

收稿日期: 2009-09-28
基金项目: 国家社会科学基金(09BTJ004)、南京理工大学科研发展基金(XF09040).
通讯联系人: 赵慧秀, 博士, 讲师, 研究方向: 生物医学统计. E-mail: huixiu_zhao@yahoo.com

为 β的函数, 记为 $\mu(\beta)=\mu(\theta(\beta))$, $V(\beta)=V(\theta(\beta))$. 设 $D(\beta)$ 为 $\mu(\beta)$ 关于 β的一阶导数矩阵, $W(\beta)$ 为 $\mu(\beta)$ 关于 β的二阶导数立体阵. 最后记 $l(\beta)=\log p(y;\theta(\beta),\sigma)$, 其前二阶导数记为 $\dot{l}(\beta)$ 和 $\ddot{l}(\beta)$.

指数族非线性模型的改进的 BW (MBW)[2] 几何结构如下: 在 R^n 中取 $\mu=(\mu_1,\cdots,\mu_n)$ 为欧氏坐标, 定义解轨迹 π为 $\mu=\mu(\beta)$, 它是 R^n 中的 p 维曲面, 它在 β处的切空间 T_β 由 $D(\beta)$ 的列向量生成, 在 R^n 和 T_β 中定义向量 a, b 的内积为 $\langle a,b\rangle=a^{\mathrm T}V^{-1}b$. 在此内积下, 解轨迹 π在 β处的法空间记为 T'_β. 设 $D(\beta)$ 关于上述内积的 QR 分解为: $D(\beta)=(Q,N)\begin{pmatrix}R\\0\end{pmatrix}=QR$, 其中 R 为满秩上三角矩阵, Q 和 N 的列向量分别构成 T_β 和 T'_β 的标准正交基, 它们满足 $Q^{\mathrm T}V^{-1}Q=I_p$, $N^{\mathrm T}V^{-1}N=I_{n-p}$, $Q^{\mathrm T}V^{-1}N=0$ (I_p 和 I_{n-p} 为单位阵), 在上述内积下, 一个矩阵 D 的投影阵为 $P_D=D(D^{\mathrm T}V^{-1}D)^{-1}D^{\mathrm T}V^{-1}$, 特别地, T_β 和 T'_β 上的投影阵分别为 $P_T=QQ^{\mathrm T}V^{-1}$ 和 $P_N=NN^{\mathrm T}V^{-1}$. 根据上述标记, 可定义指数族非线性模型的固有曲率立体阵 A^I 和参数效应曲率立体阵 A^P 如下: $A^I=[N^{\mathrm T}V^{-1}][U]$, $A^P=[Q^{\mathrm T}V^{-1}][U]$, $U=L^{\mathrm T}WL$, $L=R^{-1}$, 其中[.][.]表示矩阵和立体阵的乘积[1]. 与此类似, 亦可基于自然参数 $\theta=\theta(\beta)$ 定义 D_θ 及曲率立体阵 A^I_θ 和 A^P_θ, 不同的是内积的形式为 $\langle a,b\rangle_\theta=a^{\mathrm T}Vb$, 并有以下重要关系:

$$D_\theta=V^{-1}D,\ Q_\theta=V^{-1}Q,\ N_\theta=V^{-1}N,\ L_\theta=R_\theta^{-1}=L,\ W_\theta=[V^{-1}][W-\Gamma],$$

$$A^I_\theta=[N^{\mathrm T}V^{-1}][L^{\mathrm T}(W-\Gamma)L]=A^I-\Gamma^I,\ A^P_\theta=[Q^{\mathrm T}V^{-1}][L^{\mathrm T}(W-\Gamma)L]=A^P-\Gamma^P,$$

其中 $\Gamma=D^{\mathrm T}V^{-1}SV^{-1}D$, $S=b^{(3)}(\theta)$ ($b(\theta)$ 关于 θ前三阶导数) 都是 $n\times n\times n$ 立体阵, $\Gamma^I=[N^{\mathrm T}V^{-1}][L^{\mathrm T}\Gamma L]$, $\Gamma^P=[Q^{\mathrm T}V^{-1}][L^{\mathrm T}\Gamma L]$.

在上述的指数族非线性模型中再对参数加个约束条件:

$$A'\beta=a,\tag{3}$$

其中 A 为 $p\times k$ 阶矩阵 ($k<p$), 则此模型就成了具有线性约束的指数族非线性模型. 今假定模型 (1) ~ (3) 的所有量均在下标加上 c. 不失一般性, 可假设 A 为列满秩矩阵. 所以约束条件 $A'\beta=a$ 可以反解为

$$\beta=A(A^{\mathrm T}A)^{-1}a+H\varphi,\tag{4}$$

其中 H 为 $p\times q$ 阶矩阵 ($q=p-k$), 满足 $A'H=0$, $H'A=0$, $H'H=E_q$, φ为 q 维实数空间的子集 Ψ上的未知参数.

今假定以下所讨论的问题都满足文献[2]中的正则条件 A、B、C, 且正则条件 A、B、C关于参数 β满足的条件关于参数 φ也满足, 以保证 β的估计 $\hat\beta_c$ 存在及其收敛性.

2　估计量的随机展开

定理 1　若模型 (1) ~ (3) 满足上述假定, 则 $\Delta\beta_c=\hat\beta_c-\beta_0$ 的一阶随机展开可表示为

$$\Delta\beta_c=L\tau_t+O_P(n^{-1}),\tag{5}$$

$$\sqrt{n}\,\Delta\beta_c\xrightarrow{L}N(0,\sigma^2K_c(\beta)),\tag{6}$$

其中 $K_c(\beta)=K^{-1}-K^{-1}A(A^{\mathrm T}K^{-1}A)^{-1}A^{\mathrm T}K^{-1}$, $\tau_t=P_t\tau$, $P_t=E_p-P_n$, $P_n=P_{L^{\mathrm T}A}$.

证明　由于 $\hat\beta_c$ 为 B 上的函数 $l(\beta)$ 在约束条件 $A^{\mathrm T}\beta=a$ 下的最大值点, 因此根据 Lagrange乘数法, 可令 $u(\beta,\rho)=l(\beta)+\rho^{\mathrm T}(A^{\mathrm T}\beta-a)$, 其中 ρ为 k 维参数向量, 对上式的 β和 ρ分别求导, 可知 $\hat\beta_c$ 和 $\hat\rho$ 应满足:

$$\begin{cases}\dot{l}(\hat\beta_c)+A\hat\rho=0,\\ A^{\mathrm T}\hat\beta_c=a.\end{cases}\tag{7}$$

将 $\dot{l}(\hat\beta_c)$ 在 β_0 点进行一阶展开得:

$$\dot{l}(\beta_0)+\ddot{l}(\beta_0)\Delta\beta_c+A\hat\rho+r\|\Delta\beta_c\|=0\ (\text{当 }\Delta\beta_c\to 0\text{ 时, }r\to 0).$$

由文献[2]的引理 2.3可得:

$$\phi D_\theta^{\mathrm T}(\beta)e+\phi[e^{\mathrm T}][W_\theta]\Delta\beta_c-\phi D_\theta^{\mathrm T}VD_\theta\Delta\beta_c+A\hat\rho+r\|\Delta\beta_c\|=0$$

由文献[2]的定理 4.1和引理 4.1可得 $\phi[e^{\mathrm{T}}][W_\theta]\Delta\beta_c = O_P(1)$, 所以由上式可解出

$$\Delta\beta_c = (\phi D_\theta^{\mathrm{T}} V D_\theta)^{-1}(\phi D_\theta^{\mathrm{T}} e + A\hat{\rho}) + a, \text{ 其中 } a = O_P(n^{-1}).$$

此等式两边同乘以 A^{T}, 整理得:

$$\hat{\rho} = -\phi\{A^{\mathrm{T}}(D_\theta^{\mathrm{T}} V D_\theta)^{-1}A\}^{-1}A^{\mathrm{T}}(\phi D_\theta^{\mathrm{T}} V D_\theta)^{-1}\phi D_\theta^{\mathrm{T}} e + O_P = -\phi(A^{\mathrm{T}} L L^{\mathrm{T}} A)^{-1}A^{\mathrm{T}} L\tau + O_P(1),$$

所以

$$\Delta\beta_c = (\phi D_\theta^{\mathrm{T}} V D_\theta)^{-1}[\phi D_\theta^{\mathrm{T}} e - \phi A(A^{\mathrm{T}} L L^{\mathrm{T}} A)^{-1}A^{\mathrm{T}} L\tau] + a = LP_t\tau + O_P(n^{-1}) = L\tau_t + O_P(n^{-1}),$$

从而(5)式得证.

$$\Delta\beta_c = (D^{\mathrm{T}} V^{-1} D)^{-1} D^{\mathrm{T}} V^{-1} e - (D^{\mathrm{T}} V^{-1} D)^{-1} A(A^{\mathrm{T}} L L^{\mathrm{T}} A)^{-1} A^{\mathrm{T}} L\tau + a =$$
$$(D^{\mathrm{T}} V^{-1} D)^{-1} D^{\mathrm{T}} V^{-1} e - (D^{\mathrm{T}} V^{-1} D)^{-1} A(A^{\mathrm{T}}(D^{\mathrm{T}} V^{-1} D)^{-1} A)^{-1} A^{\mathrm{T}}(D^{\mathrm{T}} V^{-1} D)^{-1} D^{\mathrm{T}} V^{-1} e + a.$$

由文献[2]的引理 4.1, 可得: $\dfrac{D^{\mathrm{T}} V^{-1} e}{\sqrt{n}} \xrightarrow{L} N(0, \sigma^2 K(\beta))$.

$$\sqrt{n}\Delta\beta_c \xrightarrow{L} N(0, \sigma^2 K_c(\beta)), \quad K_c(\beta) = K^{-1} - K^{-1}A(A^{\mathrm{T}} K^{-1} A)^{-1} A^{\mathrm{T}} K^{-1},$$

从而(6)式得证.

定理 2 若模型(1)、(2)、(3)满足定理 1的条件, 且满足条件 D: 假定 θ_t 在 B 存在三阶连续导数, 并且下列有关导数的极限存在

$$\overline{\lim_{n\to\infty}} \frac{1}{n}\sum_{t=1}^{n}\left(\frac{\partial^3\theta_t}{\partial\beta_i\partial\beta_j\partial\beta_k}\right)^2 = G_{ijk}(\beta) < +\infty,$$

则 $\Delta\beta_c = \hat{\beta}_c - \beta_0$ 的二阶随机展开可表示为:

$$\Delta\beta_c = LP_t\left\{\tau_t + [\lambda^{\mathrm{T}}][A_\theta^I]\tau_t + [\tau_n^{\mathrm{T}}][A_\theta^P]\tau_t - \frac{1}{2}\tau_t^{\mathrm{T}} A^P \tau_t\right\} + O_P(n^{-1}), \tag{8}$$

其中 $\tau_n = P_n\tau = \tau - \tau_t$, $\lambda = N^{\mathrm{T}} V^{-1} e$.

证明

$$\frac{\partial l}{\partial\beta_i} = \phi\sum_{t=1}^{n}\left[y_t\frac{\partial\theta_t}{\partial\beta_i} - \dot{b}(\theta_t)\frac{\partial\theta_t}{\partial\beta_i}\right] = \phi\sum_{t=1}^{n}(y_t - \mu_t(\beta))\frac{\partial\theta_t}{\partial\beta_i},$$

$$\frac{\partial^2 l}{\partial\beta_i\partial\beta_j} = \phi\sum_{t=1}^{n}\left(e_t\frac{\partial^2\theta_t}{\partial\beta_i\partial\beta_j} - \frac{\partial\mu_t(\beta)}{\partial\beta_j}\frac{\partial\theta_t}{\partial\beta_i}\right), \text{其中 } e_t = y_t - \mu_t(\beta),$$

$$\frac{\partial^3 l}{\partial\beta_i\partial\beta_j\partial\beta_k} = \phi\sum_{t=1}^{n}\left(e_t\frac{\partial^3\theta_t}{\partial\beta_i\partial\beta_j\partial\beta_k} - \frac{\partial\mu_t(\beta)}{\partial\beta_k}\frac{\partial^2\theta_t}{\partial\beta_i\partial\beta_j} - \frac{\partial^2\mu_t(\beta)}{\partial\beta_j\partial\beta_k}\frac{\partial\theta_t}{\partial\beta_i} - \frac{\partial\mu_t(\beta)}{\partial\beta_j}\frac{\partial^2\theta_t}{\partial\beta_i\partial\beta_k}\right) =$$
$$\phi\sum_{t=1}^{n}\left(e_t\frac{\partial^3\theta_t}{\partial\beta_i\partial\beta_j\partial\beta_k} - D_{tk}W_{\theta tij} - W_{tjk}D_{\theta ti} - D_{tj}W_{\theta tik}\right).$$

$(\Delta\beta_c)^{\mathrm{T}}\dddot{l}\,\Delta\beta_c$ 的第 i 个分量为:

$$\phi\sum_{j=1}^{p}\sum_{k=1}^{p}\Delta\beta_{cj}\Delta\beta_{ck}\frac{\partial^3 l}{\partial\beta_i\partial\beta_j\partial\beta_k} =$$
$$\phi\sum_{t=1}^{n}\sum_{j=1}^{p}\sum_{k=1}^{p}\left(e_t\frac{\partial^3\theta_t}{\partial\beta_i\partial\beta_j\partial\beta_k} - \Delta\beta_{cj}\Delta\beta_{ck}D_{tk}W_{\theta tij} - \Delta\beta_{cj}\Delta\beta_{ck}D_{tj}W_{\theta tik} - \Delta\beta_j\Delta\beta_k W_{tjk}D_{\theta ti}\right),$$

所以

$$(\Delta\beta_c)^{\mathrm{T}}\dddot{l}\,\Delta\beta_c = \phi(\Delta\beta_c)^{\mathrm{T}}\left\{e^{\mathrm{T}}\left(\frac{\partial^3\theta}{\partial\beta^3}\right)^{\mathrm{T}}\right\}\Delta\beta_c - 2\phi[(D\Delta\beta_c)^{\mathrm{T}}][W_\theta\Delta\beta] - \phi(\Delta\beta_c)^{\mathrm{T}}[D_\theta^{\mathrm{T}}][W](\Delta\beta_c), \tag{9}$$

其中上式右端第一项仅仅是一个符号, 它表示一个向量. 其第 i 个分量为

$$\sum_{j=1}^{p}\sum_{k=1}^{p}\sum_{t=1}^{n}\phi e_t\frac{\partial^3\theta_t}{\partial\beta_i\partial\beta_j\partial\beta_k}\Delta\beta_{ck}\Delta\beta_{cj}.$$

由 Lagrange乘数法将(7)式把 $\dot{l}(\hat{\beta}_c)$ 在 β_0 点二阶展开得:

$$\begin{cases}\dot{l}(\beta_0) + \ddot{l}(\beta_0)\Delta\beta_c + A\hat{\rho} + \dfrac{1}{2}(\Delta\beta_c)^{\mathrm{T}}\dddot{l}\,\Delta\beta_c + r\|\Delta\beta_c\|^2 = 0,\\ A^{\mathrm{T}}\hat{\beta}_c = \alpha\end{cases}$$

由文献[2]的引理2.3及(9)式可得:

$$\phi D_\theta^{\mathrm{T}}(\beta)e+\phi[e^{\mathrm{T}}][W_\theta]\Delta\beta_c-\phi D_\theta^{\mathrm{T}}VD_\theta\Delta\beta_c+\frac{1}{2}\phi(\Delta\beta_c)^{\mathrm{T}}\left\{e^{\mathrm{T}}\left(\frac{\partial^3\theta}{\partial\beta^3}\right)^{\mathrm{T}}\right\}\Delta\beta_c-$$

$$\phi[(D\Delta\beta_c)^{\mathrm{T}}][W_\theta\Delta\beta]-\frac{1}{2}\phi(\Delta\beta_c)^{\mathrm{T}}[D_\theta^{\mathrm{T}}][W](\Delta\beta_c)+A\rho+r\|\Delta\beta_c\|^2=0$$

解得:

$$\Delta\beta_c=(\phi D_\theta^{\mathrm{T}}VD_\theta)^{-1}\phi D_\theta^{\mathrm{T}}(\beta)e+(\phi D_\theta^{\mathrm{T}}VD_\theta)^{-1}\phi[e^{\mathrm{T}}][W_\theta]\Delta\beta_c+$$

$$\frac{1}{2}(\phi D_\theta^{\mathrm{T}}VD_\theta)^{-1}\phi(\Delta\beta_c)^{\mathrm{T}}\left\{e^{\mathrm{T}}\left(\frac{\partial^3\theta}{\partial\beta^3}\right)^{\mathrm{T}}\right\}\Delta\beta_c-(\phi D_\theta^{\mathrm{T}}VD_\theta)^{-1}\phi[(D\Delta\beta_c)^{\mathrm{T}}][W_\theta\Delta\beta]$$

$$-\frac{1}{2}(\phi D_\theta^{\mathrm{T}}VD_\theta)^{-1}\phi(\Delta\beta_c)^{\mathrm{T}}[D_\theta^{\mathrm{T}}][W](\Delta\beta_c)+(\phi D_\theta^{\mathrm{T}}VD_\theta)^{-1}A\rho+r\|\Delta\beta_c\|^2. \tag{10}$$

由文献[2]中的引理4.1及条件D可得

$$(\Delta\beta_c)^{\mathrm{T}}\left\{e^{\mathrm{T}}\left(\frac{\partial^3\theta}{\partial\beta^3}\right)^{\mathrm{T}}\right\}\Delta\beta_c=O_P\left(n^{-\frac{1}{2}}\right),$$

所以(10)式可变为:

$$\Delta\beta_c=(\phi D_\theta^{\mathrm{T}}VD_\theta)^{-1}(\phi D_\theta^{\mathrm{T}}e+A\rho)+(\phi D_\theta^{\mathrm{T}}VD_\theta)^{-1}\left\{[\phi e-\phi D\Delta\beta_c]^{\mathrm{T}}[W_\theta]\Delta\beta_c-\right.$$

$$\left.\frac{1}{2}\phi(\Delta\beta_c)^{\mathrm{T}}[D_\theta^{\mathrm{T}}][W](\Delta\beta_c)\right\}+O_P(n^{-\frac{3}{2}}), \tag{11}$$

上式两边同乘以A^{T}得:

$$A^{\mathrm{T}}(\phi D_\theta^{\mathrm{T}}VD_\theta)^{-1}A\rho=-A^{\mathrm{T}}(\phi D_\theta^{\mathrm{T}}VD_\theta)^{-1}\phi D_\theta^{\mathrm{T}}e-$$

$$A^{\mathrm{T}}(\phi D_\theta^{\mathrm{T}}VD_\theta)^{-1}\left\{[\phi e-\phi D\Delta\beta_c]^{\mathrm{T}}[W_\theta]\Delta\beta_c-\frac{1}{2}\phi(\Delta\beta_c)^{\mathrm{T}}[D_\theta^{\mathrm{T}}][W](\Delta\beta_c)\right\}+O_P(n^{-\frac{3}{2}})=$$

$$-A^{\mathrm{T}}L\tau-A^{\mathrm{T}}Lw+O_P(n^{-\frac{3}{2}}), \tag{12}$$

其中 $w=L^{\mathrm{T}}\left\{[(e-D\Delta\beta_c)^{\mathrm{T}}][W_\theta]\Delta\beta_c-\frac{1}{2}(\Delta\beta_c)^{\mathrm{T}}[D_\theta^{\mathrm{T}}][W](\Delta\beta_c)\right\}.$　(13)

由(11)、(12)式,整理得:

$$\Delta\beta_c=L\tau+\phi^{-1}LL^{\mathrm{T}}A(A^{\mathrm{T}}LL^{\mathrm{T}}A)^{-1}\{-\phi A^{\mathrm{T}}L\tau-\phi A^{\mathrm{T}}Lw\}+Lw+O_P(n^{-\frac{3}{2}}),$$

$$\Delta\beta_c=L\tau-LL^{\mathrm{T}}A(A^{\mathrm{T}}LL^{\mathrm{T}}A)^{-1}A^{\mathrm{T}}L(\tau+w)+Lw+O_P(n^{-\frac{3}{2}}). \tag{14}$$

把(5)式和 $e=(QQ^{\mathrm{T}}V^{-1}+NN^{\mathrm{T}}V^{-1})e=Q\tau+N\lambda$带入(13)式,整理得:

$$w=L^{\mathrm{T}}\left\{[(e-DL\tau_t)^{\mathrm{T}}][W_\theta]L\tau_t-\frac{1}{2}\tau_t^{\mathrm{T}}[D_\theta^{\mathrm{T}}][L^{\mathrm{T}}WL]\tau_t\right\}+O_P(n^{-1})=$$

$$[\lambda^{\mathrm{T}}][A_\theta^I]\tau_t+[\tau_n^{\mathrm{T}}][A_\theta^P]\tau_t-\frac{1}{2}\tau_t^{\mathrm{T}}A^P\tau_t+O_P(n^{-1}). \tag{15}$$

把(15)式带入(14)式,整理得:

$$\Delta\beta_c=L\tau-LP_n(\tau+w)+Lw+O_P(n^{-\frac{3}{2}})=L(\tau-\tau_n)-LP_nw+Lw+O_P(n^{-\frac{3}{2}})=$$

$$L\tau_t+LP_tw+O_P(n^{-\frac{3}{2}})=LP_t(\tau_t+w)+O_P(n^{-\frac{3}{2}})=$$

$$LP_t\left\{\tau_t+[\lambda^{\mathrm{T}}][A_\theta^I]\tau_t+[\tau_n^{\mathrm{T}}][A_\theta^P]\tau_t-\frac{1}{2}\tau_t^{\mathrm{T}}A^P\tau_t\right\}+O_P(n^{-\frac{3}{2}}),$$

所以(8)式得证.

特殊情形:当模型为带线性约束的正态非线性回归模型时,

$$V=E_n,A_\theta^I=[N^{\mathrm{T}}][U]\triangleq I,A_\theta^P=[Q^{\mathrm{T}}][U]\triangleq P,A^P=P,$$

所以

$$\Delta\beta_c=LP_t\left\{\tau_t+[\lambda^{\mathrm{T}}][I]\tau_t+[\tau_n^{\mathrm{T}}][P]\tau_t-\frac{1}{2}\tau_t^{\mathrm{T}}P\tau_t\right\}+O_P(n^{-\frac{3}{2}}). \tag{16}$$

这与文献[1]中的结论是一致的.

由定理 1、2可得以下推论:

推论 1 若模型(1)、(2)、(3)满足定理 2的条件, 则

$$\Delta\mu_c = \mu(\beta_c) - \mu(\beta_0) = \boldsymbol{QP}_t\{\tau_t + [\lambda^{\mathrm{T}}][\boldsymbol{A}_\theta^I]\tau_t + [\tau_n^{\mathrm{T}}][\boldsymbol{A}_\theta^P]\tau_t\} + \frac{1}{2}\boldsymbol{N}(\tau_t^{\mathrm{T}}\boldsymbol{A}'\tau_t) + \frac{1}{2}\boldsymbol{QP}_n\tau_t^{\mathrm{T}}\boldsymbol{A}^P\tau_t + \boldsymbol{O}_P(\boldsymbol{n}^{-1}). \tag{17}$$

推论 2 若模型(1)、(2)、(3)满足定理 2的条件, 则

$$\hat{\boldsymbol{e}}_c = \boldsymbol{y} - \mu(\hat{\beta}_c) = \boldsymbol{Q}\tau_n + \boldsymbol{N}\lambda - \boldsymbol{QP}_t[\lambda^{\mathrm{T}}][\boldsymbol{A}_\theta^I]\tau_t - \boldsymbol{QP}_t[\tau_n^{\mathrm{T}}][\boldsymbol{A}_\theta^P]\tau_t - \frac{1}{2}\boldsymbol{N}(\tau_t^{\mathrm{T}}\boldsymbol{A}'\tau_t) - \frac{1}{2}\boldsymbol{QP}_n\tau_t^{\mathrm{T}}\boldsymbol{A}^P\tau_t + \boldsymbol{O}_P(\boldsymbol{n}^{-1}). \tag{18}$$

[参考文献]

[1] 韦博成. 近代非线性回归分析[M]. 北京: 清华大学出版社, 1989: 88-210.

[2] Wei Bocheng. Exponential Family Nonlinear Models[M]. Singapore: Springer, 1998: 1-71.

[3] Bates D M, Watts D G. Relative curvature measures of nonlinearity[J]. J Roy Statist Soc Ser B, 1980, 42(1): 1-25.

[4] Bates D M, Watts D G. Parameter transformations for improved approximate confidence region in nonlinear least squares[J]. Ann Statist, 1981, 9(6): 1152-1167.

[5] Bates D M, Watts D G. Nonlinear Regression Analysis and Its Application[M]. New York: Wiley, 1988: 1-216.

[6] Beale E M L. Confidence regression in non-linear estimation[J]. J Roy Statist Soc Ser B, 1960, 22(1): 41-88.

[7] Belsley D A, Kuhe E, Welsch R E. Regression Diagnostics[M]. New York: Wiley, 1980: 1-105.

[责任编辑: 丁 蓉]