系 统 科 学 与 数 学
J. Sys. Sci. & Math. Scis.
28(10) (2008, 10), 1297–1307

响应变量随机缺失下的广义变系数模型的估计*

李 志 强

(北京化工大学理学院，北京 100029; 北京工业大学应用数理学院，北京 100022)

薛 留 根

(北京工业大学应用数理学院，北京 100022)

摘要 在响应变量随机缺失时，利用拟似然方法给出了广义变系数模型中非参数函数系数的估计. 研究了所得到的估计的渐近性质，求出了估计的渐近偏差与渐近方差，并进行模拟比较.

关键词 随机缺失，广义变系数模型，拟似然估计，渐近正态性.

MR(2000) 主题分类号 62G08

1 引言和估计方法

变系数模型是由 Hastie 和 Tibshirani[1] 于 1993 年提出的，其通过允许回归系数依赖于某些协变量，使应用范围变广. 现已有许多文献研究变系数模型，如 Cai 等人 [2] 利用基于似然的方法研究变系数模型中函数系数的非参数估计的渐近性质. 当响应变量随机缺失时，本文将 Cai 等人 [2] 的模型利用拟似然方法推广到缺失数据下，即通过回归方法填补缺失的响应变量值，再利用完全数据方法估计函数系数. 本文讨论了估计的渐近分布，并对完整个体方法和本文方法进行模拟比较. 由模拟可以看出，本文方法要比仅用未缺失的完整个体的方法稳定.

设 (Y_i, X_i, U_i), $i=1,2,\cdots,n$ 为来自 (Y,X,U) 的独立同分布样本，这里 $Y,U\in R$, $X=(X_1,X_2,\cdots,X_p)^{\mathrm T}\in R^p$, 其中响应变量 Y_i 随机缺失 [3]. 当 Y_i 缺失时，令 $\delta_i=0$, 否则令 $\delta_i=1$. 在响应变量随机缺失的假定下，选取概率满足

$$P\{\delta_i=1|X_i=x,Y_i=y,U_i=u\}=P\{\delta_i=1|X_i=x,U_i=u\}\equiv\pi(x,u). \tag{1.1}$$

* 国家自然科学基金 (10571008), 北京市自然科学基金 (1072004) 和河南省基础与前沿技术研究计划 (072300410090) 资助课题.

收稿日期：2006-02-09, 收到修改稿日期：2007-04-25.

假定响应变量 Y 关于 X,U 的条件期望函数为 $\mu(x,u)=E[Y|X=x,U=u]$, 条件方差有形式 $\mathrm{Var}(Y|X=x,U=u)=\sigma^2V\big(\mu(x,u)\big)$, 则相应的拟似然函数 $Q(\omega,y)$ 满足

$$\frac{\partial Q(\omega,y)}{\partial\omega}=\frac{y-\omega}{V(\omega)}. \tag{1.2}$$

假设对给定的已知联系函数 $g(\cdot)$, 满足

$$\eta(x,u)=g\{\mu(x,u)\}=\sum_{j=1}^{p}a_j(u)x_j,$$

其中 $x=(x_1,x_1,\cdots,x_p)^{\mathrm{T}}$, 而 $a_j(u),\ j=1,2,\cdots,p$ 为未知的函数系数.

下面给出诸函数系数 $a_j(u)$ 的估计. 设 $K(u)$ 为对称概率密度, 记 $K_h(t)=\frac{K(\frac{t}{h})}{h}$, 其中 h 为窗宽. 对给定点 u 及 u 邻域中点 v, 用线性函数逼近 $a_j(v)$, 即 $a_j(v)\approx a_j+b_j(v-u)$. 易见 $a_j=a_j(u),b_j=a'_j(u)$. 记 $a=(a_1,a_2,\cdots,a_p)^{\mathrm{T}},\ b=(b_1,b_2,\cdots,b_p)^{\mathrm{T}}$.

设 h_1,h_2 为窗宽. 记 $(\widehat{a},\widehat{b})$ 极大化局部拟似然函数

$$l_c(a,b)=\sum_{i=1}^{n}Q\Big(g^{-1}\Big\{\sum_{j=1}^{p}\big(a_j+b_j(U_i-u)\big)X_{ij}\Big\},Y_i\Big)K_{h_1}(U_i-u)\delta_i\ . \tag{1.3}$$

由此可得 $a(u)$ 的完整个体估计 $\widehat{a}_c(u)=\widehat{a}$, 则 $\widehat{\eta}(X_i,U_i)=\sum\limits_{k=1}^{p}\widehat{a}_k(U_i)X_{ik}=X_i^{\mathrm{T}}\widehat{a}_c(U_i)$. 令

$$\widehat{Y}_i^*=\delta_iY_i+(1-\delta_i)g^{-1}\big(\widehat{\eta}(X_i,U_i)\big). \tag{1.4}$$

然后利用 $\widehat{Y}_i^*$ 填补缺失的 Y_i 值，再对全部数据利用局部线性拟似然求 $(\widehat{a},\widehat{b})$ 以极大化

$$l_I(a,b)=\sum_{i=1}^{n}Q\Big(g^{-1}\Big\{\sum_{j=1}^{p}\big(a_j+b_j(U_i-u)\big)X_{ij}\Big\},\widehat{Y}_i^*\Big)K_{h_2}(U_i-u)\ . \tag{1.5}$$

由此得到 $a(u)$ 的借补估计 $\widehat{a}_I(u)=\widehat{a}$.

2 估计量的渐近分布

记 $q_l(x,y)=(\frac{\partial^l}{\partial x^l})Q\{g^{-1}(x),y\}$, $l=1,2,3$. 由拟似然函数的定义可知

$$q_1(x,y)=\{y-g^{-1}(x)\}\rho_1(x),\ q_2(x,y)=\{y-g^{-1}(x)\}\rho'_1(x)-\rho_2(x). \tag{2.1}$$

其中

$$\rho_l(t)=\frac{\Big(\frac{\mathrm{d}g^{-1}(t)}{\mathrm{d}t}\Big)^l}{V(g^{-1}(t))}.$$

易知 $q_l(x,y)$ 为 y 的线性函数. 记 $X^{\otimes2}=XX^{\mathrm{T}},X^{\otimes3}=XX^{\mathrm{T}}X,X^{\otimes4}=XX^{\mathrm{T}}XX^{\mathrm{T}}$. 下面给出一些证明本文主要定理所用的条件.

条件 1 $\forall x\in R$ 及在响应变量 Y 的取值范围内的 y, 函数 $q_2(x,y)<0$.

条件 2 U 具有紧支撑，且存在某个正数 $\delta > 0$, 使得 U 的密度函数 $f(u) \geq \delta$, $f(u)$ 具有连续的一阶导数； (X,U) 具有密度函数 $f(x,u) > 0$, 且 $f(x,u)$ 具有连续的一阶偏导数；$\pi(x,u) > 0$, 关于 x,u 的一阶偏导数存在.

条件 3 $a''(u), g'''(u), V'\big(\mu(x,u)\big)$ 和 $E[|X_k X_l X_m||U=u]$ $(1 \leq k,l,m \leq p)$ 在点 u 连续，$E[Y^4|X=x,U=u]$ 在 u 的邻域内有界.

条件 4 核函数 $K(\cdot)$ 为有紧支撑的对称概率密度函数，且 2 阶矩存在.

条件 5 对任意 $1 \leq k,l,m \leq p$, 条件期望 $E\big[|[g^{-1}(\eta(X,U))]'\rho_1'(\eta(X,U))X_k X_l X_m|\big|U=u\big]$, $E[|q_3(\eta(X,U),Y)X_k X_l X_m||U=u]$, $E[\rho_2^2(\eta(X,U))X^{\otimes 4}|U=u]$ 存在.

条件 6 作为 u 的函数，$E[\pi(X,U)q_2(\eta(X,U),Y)|U=u]$, $E[\pi(X,U)q_2(\eta(X,U),Y)X^{\mathrm{T}}|U=u]$, $E[\pi(X,U)q_2(\eta(X,U),Y)XX^{\mathrm{T}}|U=u]$ 和 $E[q_2^2(\eta(X,U),Y)|U=u]$ 在点 u 连续，并且对某个 $\delta > 2$, 成立 $E[q_1^{2+\delta}(\eta(X,U),Y)] < \infty$.

定义

$$\kappa_i = \int t^i K(t)\mathrm{d}t, \quad \upsilon_i = \int t^i K(t)^2 \mathrm{d}t, \quad i=0,1,2, \quad \varLambda = \frac{\sigma^2 \upsilon_0}{f(u)}\varSigma^{-1}(u),$$

$$\varSigma(u) = E[\pi(X,U)\rho_2\big(\eta(X,U)\big)XX^{\mathrm{T}}|U=u], \quad \varSigma_1(u) = E[\rho_2\big(\eta(X,U)\big)XX^{\mathrm{T}}|U=u]. \tag{2.2}$$

下面的定理建立了 $a(u)$ 的估计的渐近正态性.

定理 2.1 在条件 1–6 下，当 $h_1 \to 0, nh_1 \to \infty$ 时，由 (1.3) 式求解得到的 $\widehat{a}_c(u)$ 满足

$$\sqrt{nh_1}\Big(\widehat{a}_c(u) - a(u) - \frac{\kappa_2}{2}a''(u)h_1^2\Big) \xrightarrow{\mathcal{L}} N(0,\varLambda). \tag{2.3}$$

定理 2.2 在条件 1–6 下，当 $h_1 \to 0, \frac{nh_1^2}{\log(\frac{1}{h_1})} \to \infty$, $h_2 \to 0$, $nh_2 \to \infty$ 时，

$$\sqrt{nh_2}(\widehat{a}_I(u) - a(u) - \lambda(u)) \xrightarrow{\mathcal{L}} N(0,\varLambda). \tag{2.4}$$

这里 $\lambda(u) = \frac{\kappa_2}{2}\varSigma_1^{-1}(u)[\varSigma_1(u)a''(u)h_2^2 + \gamma_1(u)a''(u)h_1^2]$, $\gamma_1(u) = \varSigma_1(u) - \varSigma(u)$.

3 模拟研究

我们采用下面的例子来描述本文给出的方法. 设不完全随机样本为 $(Y_i, X_i, U_i, \delta_i)$, $i = 1,2,\cdots,n$, 其中 $X_i = (X_{i1}, X_{i2})^{\mathrm{T}}$, 假定 Y 关于 X, U 的条件分布为两点分布，即 $P\{Y = 1|X_1 = x_1, X_2 = x_2, U = u\} \equiv \pi(x_1,x_2,u)$. 考虑下列广义变系数模型

$$\eta(X_i,U_i) = \log\left\{\frac{\pi(X_{i1},X_{i2},U_i)}{1+\pi(X_{i1},X_{i2},U_i)}\right\} = \sin(4U_i)X_{i1} + 4U_i(1-U_i)X_{i2}, \quad i=1,2,\cdots,n.$$

其中 U_i 是服从 [0,1] 区间上均匀分布的随机变量， X_{i1} 和 X_{i2} 为服从标准正态分布的随机变量. 分别给出两个选取概率，即 $\pi_1(X_1 = x_1, X_2 = x_2, U = u) \equiv 0.8$,

$$\pi_2(x_1,x_2,u) = \left\{\begin{array}{cc} 0.7 + 0.3(|x_1| + |x_2| + |u-0.5|), & |x_1| + |x_2| + |u-0.5| \leq 1, \\ 0.7, & \text{其它}. \end{array}\right.$$

在具体估计 $a_j(u)$, $j = 1,2$ 时，核函数取为 Epanechnikov 核函数 $K(t) = 0.75(1-t^2)_+$. 窗宽 h_1 和 h_2 可分别用交叉核实法 [4] 得到. 但窗宽 h_1 和 h_2 的选取范围不同，为了保证借补

估计达到非参数估计的最优收敛速度，由定理 2 的条件知窗宽 h_2 可取到最优窗宽 $O(n^{-\frac{1}{5}})$，而 h_1 无法达到．记 $\widehat{h}_{opt}$ 为利用完整个体和交叉核实方法得到的最优窗宽，由文献 [4] 可知 $\widehat{h}_{opt}$ 的阶数为 $n^{-\frac{1}{5}}$．由于定理 2 中窗宽条件的限制，为了保证借补估计达到非参数估计的最优收敛速度，选取窗宽 h_1 为 $\widehat{h_1}=\widehat{h}_{opt}\times n^{\frac{1}{5}}\times n^{-\frac{7}{24}}=\widehat{h}_{opt}\times n^{-\frac{11}{120}}$．窗宽 h_2 的选取利用借补数据和交叉核实方法得到，由文献 [4] 可知 h_2 的阶数为 $n^{-\frac{1}{5}}$．由定理 1 和 2 可知，对给定的窗宽，函数系数的两个估计具有相同的渐近方差．但由文献 [5] 知，对于非参数回归函数的估计，数据驱动窗宽与样本容量有关．当利用借补数据时，有较大的样本容量，所选取的窗宽也更稳定，由此可导致较稳定的非参数估计及较小的方差．从而使其有限抽样性质得到提高．这也是本文在估计函数系数时需要两个窗宽的原因．

对于两种估计，我们利用渐近均方误差 (AMSE) 来评价它们的表现，定义

$$\mathrm{AMSE}=\frac{1}{n}\sum_{i=1}^{n}\sum_{r=1}^{2}[\widehat{a}_r(U_i)-a_r(U_i)]^2.$$

下表分别给出样本容量为 100, 200, 300 的模拟，在上述缺失概率下利用本文方法重复做 100 次得到的 $\widehat{a}$ 的模拟结果．其中 mean 表示当样本容量为 n 时，由 100 次模拟得到的 AMSE 的平均值，而 std 表示 100 次模拟中 AMSE 的标准差．由表 1 和表 2 中的取值情况可知，在样本容量固定时，借补估计 $\widehat{a}_I(u)$ 对应的 AMSE 要小些，且对应的标准差也较小．这说明当利用交叉核实方法分别选取窗宽时，借补估计 $\widehat{a}_I(u)$ 更稳定，效率也更高．故本文提出的填补缺失值的估计方法要好些．

表 1　π_1 时估计的 AMSE 的均值和标准差

	n	100	200	300
$\widehat{a}_c$	mean	0.7452	0.4421	0.2262
	std	0.9901	0.3763	0.1666
$\widehat{a}_I$	mean	0.6275	0.2586	0.1975
	std	0.5137	0.2633	0.1585

表 2　π_2 时估计的 AMSE 的均值和标准差

	n	100	200	300
$\widehat{a}_c$	mean	0.7523	0.4350	0.2527
	std	0.9473	0.3418	0.1679
$\widehat{a}_I$	mean	0.6342	0.2437	0.2498
	std	0.6293	0.2716	0.1525

表 3 和表 4 在 π_2 下分别给出 $a_1(0.3)=\sin(4\times0.3)=0.932$ 和 $a_2(0.6)=4\times0.6(1-0.6)=0.96$ 的估计在 100 次模拟中的均值和标准差．由表 3 和表 4 中的取值可以看出，样本容量越大估计越精确．当样本容量较小时，$\widehat{a}_I$ 的标准差要比 $\widehat{a}_c$ 的标准差小得多．而当样本容量较大时，两个估计的标准差较接近，这与定理 1,2 的结论相符合．

表 3　π_2 时 $a_1(0.3)$ 的估计的均值和标准差

	n	100	200	300
$\widehat{a}_{1c}(0.3)$	mean	0.7943	0.9470	0.9289
	std	0.4073	0.2929	0.1676
$\widehat{a}_{1I}(0.3)$	mean	0.8559	0.9190	0.9386
	std	0.3302	0.2396	0.1648

表 4　π_2 时 $a_2(0.6)$ 的估计的均值和标准差

	n	100	200	300
$\widehat{a}_{2c}(0.6)$	mean	0.9124	0.9986	0.9671
	std	0.3915	0.2816	0.1672
$\widehat{a}_{2I}(0.6)$	mean	0.9391	1.0082	0.9547
	std	0.3282	0.2382	0.1671

4 定理的证明

定理 2.1 的证明　定义

$$c_n=(nh_1)^{-\frac{1}{2}},\quad X_i^*=\left(X_i^{\mathrm T},\left(\tfrac{U_i-u}{h_1}\right)X_i^{\mathrm T}\right)^{\mathrm T},\quad D=\begin{pmatrix}XX^{\mathrm T} & 0\\ 0 & \kappa_2XX^{\mathrm T}\end{pmatrix},\tag{4.1}$$

$$\beta^*=(c_n^{-1}\big(a^{\mathrm T}-a^{\mathrm T}(u)\big),h_1c_n^{-1}\big(b^{\mathrm T}-a'^{\mathrm T}(u)\big))^{\mathrm T},$$

$$\overline{\eta}_i(u)=\sum_{j=1}^p\{a_j(u)+a_j'(u)(U_i-u)\}X_{ij}.$$

则由 (1.3) 知 $\widehat{\beta}^*$ 为下式的极大化解

$$l_n(\beta^*)=h_1\sum_{i=1}^n\Big[Q\{g^{-1}\big(c_n\beta^{*T}X_i^*+\overline{\eta}_i(u)\big),Y_i\}-Q\{g^{-1}\big(\overline{\eta}_i(u)\big),Y_i\}\Big]K_{h_1}(U_i-u)\delta_i.\tag{4.2}$$

由 (2.2) 式中 κ_i 和 υ_i 的定义及条件 4 知, $\kappa_0=1$, $\kappa_1=0$, $\upsilon_1=0$. 记

$$W_n=h_1c_n\sum_{i=1}^nq_1\{\overline{\eta}_i(u),Y_i\}X_i^*K_{h_1}(U_i-u)\delta_i,$$

$$A=f(u)E[\pi(X,U)\rho_2\big(\eta(X,U)\big)D|U=u].\tag{4.3}$$

通过与文献 [2] 定理 1 证明中类似的计算可证

$$l_n(\beta^*)=W_n^{\mathrm T}\beta^*-\frac{1}{2}\beta^{*T}A\beta^*+o_p(1)\ .\tag{4.4}$$

应用二次逼近引理 [5], 可得

$$\widehat{\beta}^*=A^{-1}W_n+o_p(1)\ .\tag{4.5}$$

进一步，通过考虑 (4.5) 式中的前 p 个分量，有

$$\widehat{a}_c(u)-a(u)=\frac{1}{nf(u)}\sum_{i=1}^nW_iK_{h_1}(U_i-u)\delta_i+o_p(c_n),\tag{4.6}$$

其中 $W_i=q_1(\overline{\eta}_i(u),Y_i)\Sigma^{-1}(u)X_i$, $\Sigma(u)$ 由 (2.2) 式给出.

对 (4.6) 中的 $q_1(\overline{\eta}_i(u),Y_i)$ 在 $\eta(X_i,U_i)$ 处利用 Taylor 展开，再通过计算一、二阶矩，可得

$$\begin{aligned}&\widehat{a}_c(u)-a(u)\\
=&\frac{1}{nf(u)}\sum_{i=1}^nq_1(\eta(X_i,U_i),Y_i)\Sigma^{-1}(u)X_iK_{h_1}(U_i-u)\delta_i-\frac{1}{nf(u)}\sum_{i=1}^nq_2(\eta(X_i,U_i),Y_i)\\
&\times\frac{X_i^{\mathrm T}a''(U_i)}{2}(U_i-u)^2\Sigma^{-1}(u)X_iK_{h_1}(U_i-u)\delta_i+o_p(c_n)\\
=&\frac{1}{nf(u)}\sum_{i=1}^nq_1(\eta(X_i,U_i),Y_i)\Sigma^{-1}(u)X_iK_{h_1}(U_i-u)\delta_i-\frac{\Sigma^{-1}(u)}{2f(u)}\\
&\times E\big\{E\big[\delta_1q_2(\eta(X_1,U_1),Y_1)X_1X_1^{\mathrm T}a''(U_1)K_{h_1}(U_1-u)(U_1-u)^2|U_1\big]\big\}+o_p(c_n)\\
=&\frac{1}{nf(u)}\sum_{i=1}^nq_1(\eta(X_i,U_i),Y_i)\Sigma^{-1}(u)X_iK_{h_1}(U_i-u)\delta_i+\frac{a''(u){h_1}^2}{2}+o_p(c_n).\end{aligned}\tag{4.7}$$

此时通过直接计算条件期望，易证

$$E\Big[\frac{1}{nf(u)}\sum_{i=1}^{n}q_1(\eta(X_i,U_i),Y_i)\Sigma^{-1}(u)X_iK_{h_1}(U_i-u)\delta_i\Big]=0\ ,$$

$$\begin{aligned}&\mathrm{Var}\Big(\frac{1}{nf(u)}\sum_{i=1}^{n}q_1(\eta(X_i,U_i),Y_i)\Sigma^{-1}(u)X_iK_{h_1}(U_i-u)\delta_i\Big)\\&=(nh_1)^{-1}\frac{\sigma^2\upsilon_0}{f(u)}\Sigma^{-1}(u)+o_p((nh_1)^{-1})\ .\end{aligned}$$

再根据条件，可证 Lyapounov 条件成立，从而即证 (2.3) 式. 进一步，根据文献 [5] 可得比 (4.6) 式更强的结果.

引理[5] 设 D 为 R 上的紧子集，令 $||a||$ 表示向量 a 的模. 在定理 1 条件下，当 $h_1\to 0$, $\frac{nh_1^2}{\log(\frac{1}{h_1})}\to\infty$ 时，有

$$\sup_{u\in D}\Big\|\widehat{a}_c(u)-a(u)-\frac{1}{nf(u)}\sum_{i=1}^{n}W_iK_{h_1}(U_i-u)\delta_i\Big\|=O_p\Big(h_1^2c_n+c_n^2\log^{\frac{1}{2}}\Big(\frac{1}{h_1}\Big)\Big).\qquad(4.8)$$

定理 2.2 的证明 令 $\overline{\eta}_i(u)$ 如定理 1 中所设， $d_n=(nh_2)^{-\frac{1}{2}}$, $Y_i^*=\delta_iY_i+(1-\delta_i)g^{-1}\big(\eta(X_i,U_i)\big)$,

$$X_i^*=\left(X_i^{\mathrm{T}},\ \left(\frac{U_i-u}{h_2}\right)X_i^{\mathrm{T}}\right)^{\mathrm{T}},\quad \beta_I^*=\Big(d_n^{-1}\big(a^{\mathrm{T}}-a^{\mathrm{T}}(u)\big),h_2d_n^{-1}\big(b^{\mathrm{T}}-a'^{\mathrm{T}}(u)\big)\Big)^{\mathrm{T}}.$$

则由 (1.5) 可知 $\widehat{\beta}_I^*$ 为下式的极大化解

$$l_n(\beta_I^*)=h_2\sum_{i=1}^{n}\big[Q\{g^{-1}\big(d_n\beta_I^{*\mathrm{T}}X_i^*+\overline{\eta}_i(u)\big),\widehat{Y}_i^*\}-Q\{g^{-1}\big(\overline{\eta}_i(u)\big),\widehat{Y}_i^*\}\big]K_{h_2}(U_i-u).\qquad(4.9)$$

利用 Taylor 展开，可得

$$l_n(\beta_I^*)=V_n^{\mathrm{T}}\beta_I^*+\frac{1}{2}\beta_I^{*\mathrm{T}}B_n\beta_I^*+o_p(1),\qquad(4.10)$$

其中

$$V_n=h_2d_n\sum_{i=1}^{n}q_1\{\overline{\eta}_i(u),\widehat{Y}_i^*\}X_i^*K_{h_2}(U_i-u),$$
$$B_n=h_2d_n^2\sum_{i=1}^{n}q_2\{\overline{\eta}_i(u),\widehat{Y}_i^*\}X_i^*X_i^{*\mathrm{T}}K_{h_2}(U_i-u).$$

先讨论 B_n, 定义 $\|\ \widehat{a}(u)-a(u)\ \|_\infty=\sup_{u\in D}||\widehat{a}(u)-a(u)||$. 应用引理的结论知， $\|\ \widehat{a}(u)-a(u)\ \|_\infty=o_p(1)$. 由 B_n 的定义知， $2p\times 2p$ 矩阵 B_n 的每个元具有形式

$$B_{n,j,k,l}=\frac{1}{n}\sum_{i=1}^{n}q_2\{\overline{\eta}_i(u),\widehat{Y}_i^*\}X_{ik}X_{il}\Big(\frac{U_i-u}{h_2}\Big)^jK_{h_2}(U_i-u),$$

其中 $k,l=1,2,\cdots,p,\ j=0,1,2$. 由 (2.1) 式中 $q_2(x,y)$ 及 $\widehat{\eta}(X_i,U_i)$ 的定义知

$$\begin{aligned}&q_2\{\overline{\eta}_i(u),\widehat{Y}_i^*\}-q_2\{\overline{\eta}_i(u),Y_i^*\}=\{\widehat{Y}_i^*-Y_i^*\}\rho_1'(\overline{\eta}_i(u))\\=&(1-\delta_i)[g^{-1}(\widehat{\eta}(X_i,U_i))-g^{-1}(\eta(X_i,U_i))]\rho_1'(\overline{\eta}_i(u))\\=&(1-\delta_i)[g^{-1}(\eta(X_i,U_i))]'X_i^{\mathrm{T}}(\widehat{a}_c(U_i)-a(U_i))\rho_1'(\overline{\eta}_i(u))(1+o_p(1)).\end{aligned}$$

因此利用假定条件可得

$$\begin{aligned}&\left|B_{n,k,l,j}-\frac{1}{n}\sum_{i=1}^n q_2\big(\overline{\eta}_i(u),Y^*\big)X_{ik}X_{il}\Big(\frac{U_i-u}{h_2}\Big)^jK_{h_2}(U_i-u)\right|\\=&\left|\frac{1}{n}\sum_{i=1}^n[q_2\{\overline{\eta}_i(u),\widehat{Y}_i^*\}-q_2\{\overline{\eta}_i(u),Y_i^*\}]X_{ik}X_{il}\Big(\frac{U_i-u}{h_2}\Big)^jK_{h_2}(U_i-u)\right|\\\le&\left|\frac{1}{n}\sum_{i=1}^n(1-\delta_i)\left[g^{-1}(\eta(X_i,U_i))\right]'\rho_1'(\overline{\eta}_i(u))X_{ik}X_{il}\Big(\frac{U_i-u}{h_2}\Big)^jK_{h_2}(U_i-u)\sum_{s=1}^pX_{is}\right|\\&\times O_p(\|\,\widehat{a}(u)-a(u)\,\|_\infty)=o_p(1).\end{aligned}$$

利用大数定律和条件期望公式，通过直接计算可得

$$\begin{aligned}B_n&=\frac{1}{n}\sum_{i=1}^n q_2\big(\overline{\eta}_i(u),Y_i^*\big)X_i^*X_i^{*T}K_{h_2}(U_i-u)+o_p(1)\\&=E[q_2\big(\overline{\eta}_1(u),Y_1^*\big)X_1^*X_1^{*T}K_{h_2}(U_1-u)]+o_p(1)\\&=-f(u)E\big[\rho_2\big(\eta(X,U)\big)D\big|U=u\big]+o_p(1)=-B+o_p(1),\end{aligned}\tag{4.11}$$

其中 D 由 (4.1) 式给出. 应用二次逼近引理可得 $\widehat{\beta}_I^*(u)=B^{-1}V_n+o_p(1)$. 记

$$\begin{aligned}B_1&=f(u)E\big[\rho_2\big(\eta(X,U)\big)XX^{\mathrm{T}}\big|U=u\big]=f(u)\Sigma_1(u),\\V_{n1}&=h_2d_n\sum_{i=1}^n q_1\{\overline{\eta}_i(u),\widehat{Y}_i^*\}X_iK_{h_2}(U_i-u).\end{aligned}$$

由 $\widehat{\beta}_I^*(u)$ 及其渐近表达式的前 p 个分量可知

$$d_n^{-1}\big(\widehat{a}_I(u)-a(u)\big)=B_1^{-1}V_{n1}+o_p(1).\tag{4.12}$$

利用 V_{n1} 可证 $\widehat{a}_I(u)$ 的渐近正态性. 下面讨论 V_{n1}.

$$\begin{aligned}V_{n1}=&\,h_2d_n\sum_{i=1}^n q_1\{\overline{\eta}_i(u),Y_i^*\}X_iK_{h_2}(U_i-u)\\&+h_2d_n\sum_{i=1}^n[q_1\{\overline{\eta}_i(u),\widehat{Y}_i^*\}-q_1\{\overline{\eta}_i(u),Y_i^*\}]X_iK_{h_2}(U_i-u)\\=&\,V_{n2}+V_{n3}\ .\end{aligned}$$

由 (2.1) 中 $q_1(x,y)$ 的表达式及引理中 (4.8) 式，利用 Taylor 展开可得

$$\begin{aligned}
V_{n3} &= h_2 d_n \sum_{i=1}^{n} (\widehat{Y}_i^* - Y_i^*)\rho_1\big(\overline{\eta}_i(u)\big) X_i K_{h_2}(U_i - u) \\
&= h_2 d_n \sum_{i=1}^{n} (1-\delta_i)[g^{-1}\big(\eta(X_i,U_i)\big)]'[\widehat{\eta}(X_i,U_i) - \eta(X_i,U_i)] \\
&\qquad \times \rho_1\big(\overline{\eta}_i(u)\big) X_i K_{h_2}(U_i - u)(1+o_p(1)) \\
&= h_2 d_n \sum_{i=1}^{n} (1-\delta_i)[g^{-1}\big(\eta(X_i,U_i)\big)]'\rho_1\big(\overline{\eta}_i(u)\big) K_{h_2}(U_i - u) \\
&\qquad \times X_i X_i^{\mathrm{T}}\big(\widehat{a}_c(U_i) - a(U_i)\big) + o_p(1) \\
&= h_2 d_n \sum_{i=1}^{n} (1-\delta_i)[g^{-1}\big(\eta(X_i,U_i)\big)]'\rho_1\big(\overline{\eta}_i(u)\big) K_{h_2}(U_i - u) X_i X_i^{\mathrm{T}} \\
&\qquad \times \frac{1}{nf(U_i)} \sum_{j=1}^{n} q_1\big(\eta(X_j,U_j),Y_j\big)\Sigma^{-1}(U_i) X_j K_{h_1}(U_j - U_i)\delta_j \\
&\quad - h_2 d_n \sum_{i=1}^{n} (1-\delta_i)[g^{-1}\big(\eta(X_i,U_i)\big)]'\rho_1\big(\overline{\eta}_i(u)\big) K_{h_2}(U_i - u) X_i X_i^{\mathrm{T}} \frac{1}{nf(U_i)} \\
&\qquad \times \sum_{j=1}^{n} q_2\big(\eta(X_j,U_j),Y_j\big) X_j^{\mathrm{T}} \frac{a''(U_i)}{2}(U_j - U_i)^2 \Sigma^{-1}(U_i) X_j K_{h_1}(U_j - U_i)\delta_j \\
&\quad + o_p(1) \\
&= T_{n1} + T_{n2} + o_p(1)\ .
\end{aligned}$$

下面通过计算一、二阶矩，要证以下结论成立

$$T_{n1} - T'_{n1} \xrightarrow{\mathcal{P}} 0, \quad T_{n2} - T'_{n2} \xrightarrow{\mathcal{P}} 0, \tag{4.13}$$

其中

$$T'_{n1} = h_2 d_n \sum_{j=1}^{n} \delta_j q_1\big(\eta(X_j,U_j),Y_j\big)\gamma_1(U_j)\Sigma^{-1}(U_j) X_j K_{h_2}(U_j - u)\ ,$$

$$\begin{aligned}
T'_{n2} = -\frac{\kappa_2}{2} h_1^2 h_2 d_n &\sum_{j=1}^{n} \delta_j q_2\big(\eta(X_j,U_j),Y_j\big)\gamma_1(U_j)\Sigma^{-1}(U_j) X_j \\
&\cdot \sum_{k=1}^{p} a''_k(U_j) X_{jk} K_{h_2}(U_j - u),
\end{aligned} \tag{4.14}$$

这里 $\gamma_1(u) = \Sigma_1(u) - \Sigma(u)$.

下面只证 (4.13) 的的第一个式子. 事实上由 $E[q_1(\eta(X_j,U_j),Y_j)|X_j,U_j] = 0$, 可证 $E(T_{n,1} - T'_{n,1}) = 0$. 因此只需证明 $\mathrm{Var}(T_{n,1} - T'_{n,1}) \to 0$, 再由车贝雪夫不等式可证 $T_{n,1} - T'_{n,1} \xrightarrow{\mathcal{P}} 0$.

首先由定义知 $T_{n,1}-T'_{n,1}=h_2d_n\sum\limits_{j=1}^{n}M_j$, 其中

$$\begin{aligned}M_j=\delta_jq_1(\eta(X_j,U_j),Y_j)\Big\{&\frac{1}{n}\sum_{i=1}^{n}\frac{(1-\delta_i)}{f(U_i)}[g^{-1}(\eta(X_i,U_i))]'X_iX_i^{\mathrm{T}}\rho_1(\overline{\eta}_i(u))\\&\times\Sigma^{-1}(U_i)K_{h_1}(U_j-U_i)K_{h_2}(U_i-u)\\&-\gamma_1(U_j)\Sigma^{-1}(U_j)K_{h_2}(U_j-u)\Big\}X_j.\end{aligned}$$

通过直接计算可得 $\mathrm{Var}(T_{n,1}-T'_{n,1})=(h_2d_n)^2\sum\limits_{j=1}^{n}\mathrm{Var}(M_j)=h_2\mathrm{Var}(M_1)$. 记

$$M_{n,i}=\frac{(1-\delta_i)}{f(U_i)}[g^{-1}(\eta(X_i,U_i))]'X_iX_i^{\mathrm{T}}\rho_1(\overline{\eta}_i(u))\Sigma^{-1}(U_i)X_1K_{h_1}(U_1-U_i)K_{h_2}(U_i-u).$$

由 M_1 的定义，可知

$$M_1=\delta_1q_1(\eta(X_1,U_1),Y_1)\Big[\frac{1}{n}\sum_{i=1}^{n}M_{n,i}-\gamma_1(U_1)\Sigma^{-1}(U_1)K_{h_2}(U_1-u)X_1\Big].$$

由 M_1 的定义易证 $E(M_1)=0$. 因此由 $q_1(x,y)$ 的定义，通过计算条件期望可得

$$\begin{aligned}&\mathrm{Var}(M_1)\\&=E\Big\{\delta_1q_1^2(\eta(X_1,U_1),Y_1)\Big[\frac{1}{n}\sum_{i=1}^{n}M_{n,i}-\gamma_1(U_1)\Sigma^{-1}(U_1)K_{h_2}(U_1-u)X_1\Big]^{\otimes2}\Big\}\\&=\sigma^2E\Big\{\pi(X_1,U_1)\rho_2(\eta(X_1,U_1))\Big[\frac{1}{n}\sum_{i=1}^{n}M_{n,i}-\gamma_1(U_1)\Sigma^{-1}(U_1)K_{h_2}(U_1-u)X_1\Big]^{\otimes2}\Big\}\ .\end{aligned}$$

通过类似的计算可得

$$\begin{aligned}&\frac{1}{n}\sum_{i=1}^{n}M_{n,i}-\gamma_1(U_1)\Sigma^{-1}(U_1)K_{h_2}(U_1-u)X_1\\&=\frac{1}{n}\sum_{i=1}^{n}(M_{n,i}-E[M_{n,i}|X_1,U_1])\\&\quad+\Big(\frac{1}{n}\sum_{i=1}^{n}E[M_{n,i}|X_1,U_1]-\gamma_1(U_1)\Sigma^{-1}(U_1)K_{h_2}(U_1-u)X_1\Big)\\&=\frac{1}{n}\sum_{i=1}^{n}(M_{n,i}-E[M_{n,i}|X_1,U_1])+o_p(1).\end{aligned}$$

因此利用 $M_{n,i}$ 的定义，通过计算条件期望可得

$$\begin{aligned}&\mathrm{Var}(T_{n,1}-T'_{n,1})=h_2\mathrm{Var}(M_1)\\&=h_2\sigma^2E\Big\{\pi(X_1,U_1)\rho_2(\eta(X_1,U_1))\Big[\frac{1}{n}\sum_{i=1}^{n}(M_{n,i}-E[M_{n,i}|X_1,U_1])\Big]^{\otimes2}\Big\}+o(1)\\&=h_2\sigma^2E\Big\{\frac{1}{n^2}\sum_{i=1}^{n}\pi(X_1,U_1)\rho_2(\eta(X_1,U_1))E\big[[(M_{n,i}-E[M_{n,i}|X_1,U_1])]^{\otimes2}|X_1,U_1\big]\Big\}+o(1)\end{aligned}$$

$$
\begin{aligned}
&= \frac{h_2}{n}\sigma^2 E\left\{\pi(X_1,U_1)\rho_2(\eta(X_1,U_1))E\left[[(M_{n,2}-E[M_{n,2}|X_1,U_1])]^{\otimes 2}\,|X_1,U_1\right]\right\}+o(1)\\
&= O\Big(\frac{1}{nh_1}\Big)+o(1)=o(1).
\end{aligned}
$$

由此可证 (4.13) 的第一个式子成立．通过类似的计算可证 (4.13) 的第二个式子亦成立．

利用与以上类似步骤，可得

$$
\begin{aligned}
T'_{n2} &= -\frac{\kappa_2}{2}h_1^2 d_n^{-1}E\big[\delta_1 q_2\big(\eta(X_1,U_1),Y_1\big)\gamma_1(U_1)\Sigma^{-1}(U_1)X_1X_1^{\mathrm{T}}a''(U_1)K_{h_2}(U_1-u)\big]\\
&\quad\cdot(1+o_p(1))\\
&= \frac{\kappa_2}{2}f(u)\gamma_1(u)a''(u)d_n^{-1}h_1^2+o_p(d_n^{-1}h_1^2).
\end{aligned}
\tag{4.15}
$$

由 $q_1(x,y)$ 和 Y_i^* 的定义可知 $q_1\big(\eta(X_i,U_i),Y_i^*\big)=\delta_i q_1\big(\eta(X_i,U_i),Y_i\big)$. 利用 Taylor 展开，通过类似上面的讨论可得

$$
\begin{aligned}
V_{n2} &= h_2 d_n\sum_{i=1}^n q_1\big(\overline{\eta}_i(u),Y_i^*\big)X_iK_{h_2}(U_i-u)\\
&= h_2 d_n\sum_{i=1}^n q_1\big(\eta(X_i,U_i),Y_i^*\big)X_iK_{h_2}(U_i-u)\\
&\quad -h_2 d_n\sum_{i=1}^n q_2\big(\eta(X_i,U_i),Y_i^*\big)X_iX_i^{\mathrm{T}}\frac{a''(u)}{2}(U_i-u)^2K_{h_2}(U_i-u)\big(1+o_P(1)\big)\\
&= h_2 d_n\sum_{i=1}^n \delta_i q_1\big(\eta(X_i,U_i),Y_i\big)X_iK_{h_2}(U_i-u)+\frac{\kappa_2}{2}h_2^2d_n^{-1}f(u)\Sigma_1(u)a''(u)\\
&\quad +o_p(d_n^{-1}h_2^2).
\end{aligned}
\tag{4.16}
$$

设 E_p 为 p 阶单位阵，则由 $\gamma_1(u)$ 的定义有 $E_p+\gamma_1(u)\Sigma^{-1}(u)=\Sigma_1(u)\Sigma^{-1}(u)$. 由 (4.12)–(4.16) 知

$$
\begin{aligned}
&d_n^{-1}\Big\{\widehat{a}_I(u)-a(u)-\frac{\kappa_2}{2}\Sigma_1^{-1}(u)[\Sigma_1(u)a''(u)h_2^2+\gamma_1(u)a''(u)h_1^2]\Big\}\\
&= \frac{h_2d_n}{f(u)}\sum_{i=1}^n \delta_i q_1\big(\eta(X_i,U_i),Y_i\big)\Sigma_1^{-1}(u)\big[E_p+\gamma_1(U_i)\Sigma^{-1}(U_i)\big]X_iK_{h_2}(U_i-u)+o_p(1)\\
&= \frac{h_2d_n}{f(u)}\sum_{i=1}^n \delta_i q_1\big(\eta(X_i,U_i),Y_i\big)\Sigma_1^{-1}(u)\Sigma_1(U_i)\Sigma^{-1}(U_i)X_iK_{h_2}(U_i-u)+o_p(1).
\end{aligned}
\tag{4.17}
$$

通过直接计算易证

$$
E\Big[\frac{h_2d_n}{f(u)}\sum_{i=1}^n \delta_i q_1\big(\eta(X_i,U_i),Y_i\big)\Sigma_1^{-1}(u)\Sigma_1(U_i)\Sigma^{-1}(U_i)X_iK_{h_2}(U_i-u)\Big]=0\ ,
$$

$$
\begin{aligned}
&\operatorname{Var}\Big(\frac{h_2 d_n}{f(u)}\sum_{i=1}^{n}\delta_i q_1\big(\eta(X_i,U_i),Y_i\big)\varSigma_1^{-1}(u)\varSigma_1(U_i)\varSigma^{-1}(U_i)X_iK_{h_2}(U_i-u)\Big)\\
=&\frac{h_2}{f^2(u)}E\Big[\delta_1 q_1^2\big(\eta(X_1,U_1),Y_1\big)\varSigma_1^{-1}(u)\varSigma_1(U_1)\varSigma^{-1}(U_1)X_1X_1^{\mathrm{T}}\varSigma^{-1}(U_1)\\
&\qquad\times\varSigma_1(U_1)\varSigma_1^{-1}(u)K_{h_2}^2(U_1-u)\Big]\\
=&\frac{\sigma^2\upsilon_0}{f(u)}\varSigma^{-1}(u)+o_p(1).
\end{aligned}
$$

再根据条件，可证 Lyapounov 条件成立，从而可证 (2.4) 成立.

参　考　文　献

[1] Hastie R J, Tibshirani T J. Varying-coefficient models. *J. Roy. Statist. Soc. Ser. B*, 1993, **55**: 757–796.

[2] Cai Z W, Fan J and Li R Z. Efficient estimation and inference for varying-coefficient models. *J. Amer. Statist. Assoc.*, 2000, **95**: 888–902.

[3] Little R J A and Rubin D B. Statistical Analysis with Missing Data. John Wiley, New York, 1987.

[4] 唐庆国，王金德. 变系数模型中的一步估计法. 中国科学 A 辑， 2005, **35**(1): 23–38.

[5] Chen J W, Fan J, Li K H and Zhou H. Local quasi-likelihood estimation with data missing at random. *Statistica Sinica*, 2006, **16**: 1071–1100.

THE ESTIMATION OF GENERALIZED VARYING-COEFFICIENT MODEL WITH RESPONSE VARIABLES MISSING AT RANDOM

LI Zhiqiang

(*College of Sciences, Beijing University of Chemical Technology, Beijing* 100029;

College of Applied Sciences, Beijing University of Technology, Beijing 100022)

XUE Liugen

(*College of Applied Sciences, Beijing University of Technology, Beijing* 100022)

Abstract　A quasi-likelihood estimation is considered for nonparameter coefficient function in generalized varying-coefficient model with response variables missing at random. It is shown that the proposed estimation is asymptotically normal, moreover, the asymptotical bias and the variance of the estimation are given.

Key words　Missing at random, generalized varying-coefficient model, quasi-likelihood estimation, asymptotical normality.