第 41 卷 第 1 期 应 用 数 学 学 报 Vol. 41 No. 1
2018 年 1 月 ACTA MATHEMATICAE APPLICATAE SINICA Jan., 2018

部分线性函数多项式模型的联合探测 *

张 涛
(广西科技大学理学院，柳州 545006)
万艳玲
(广西科技大学社会科学学院，柳州 545006)
(E-mail: ylwanpiano@126.com)
王智文
(广西科技大学计算机科学与通讯工程学院，柳州 545006)

摘 要 本文提出了一个新的部分线性函数多项式回归模型，该模型中响应变量依赖于一个 p 阶函数多项式和一些非函数型数据的协变量. 函数多项式模型、函数线性模型和部分函数线性模型是该模型的特殊情形. 本文提出了一个模型探测方法，它能同时探测部分线性函数多项式回归模型中哪些阶是重要的以及哪些非函数型变量是重要的. 提出的方法能相合地识别真实的模型并有好的预测表现. 数值模拟能清晰地证实我们的理论结果.

关键词 函数型数据；函数型多项式模型；主成分；联合探测

MR(2000) 主题分类 62G05; 62N01

中图分类 O212.7

1 引言

随着社会的发展和科技的进步，数据收集的手段和技术越来越丰富，使得我们能收集和储存具有函数特征的数据. 于是，函数型数据成为了研究的热点和难点问题，近年来已经取得的一些进展可参见专著 Ramsay 和 Silverman[1], Ferraty 和 Vieu[2], 其中，函数型数据回归模型是函数型数据分析的一个重要方向，对这一领域，很多研究人员提出了不同的函数回归模型，其中，最广泛研究的是函数线性模型，包括对模型的估计，计算以及理论性质等，可参见 [3–6].

本文 2015 年 3 月 8 日收到. 2018 年 1 月 4 日收到修改稿.
* 国家自然科学基金 (11561006, 61462008), 广西高校科学技术研究项目 (KY2015YB171), 柳州市科学技术研究项目 (2016C050205), 以及广西科技大学创新团队项目 (gxkjdx201504) 资助.

因为函数线性回归模型在实际应用中可能存在模型的设定误差，当模型假定不成立时，基于模型假定所作的统计推断其表现可能很差，甚至导致错误的结论. 而函数非参数模型的收敛速度太慢，于是，很多作者考虑了函数线性模型和函数非参数模型的一种折中而使模型变得更加灵活同时拥有很好的理论性质. 如：2005 年 Müller 和 StadtMüller[7] 基于半参数拟似然方法研究了广义函数线性模型，2008 年 Ait-Saïdi[8] 等提出了单指标函数模型并利用交叉核实方法估计了系数函数和连接函数. 2011 年 Chen[9] 等也研究了单指标函数模型并将此模型扩展到多指标模型. 为了反映函数型协变量与其他的非函数型协变量的交互作用，2010 年 Li[10] 等建立了广义半参数单指标模型.

2010 年 Yao 和 Müller[11] 将函数线性模型推广到以下的函数多项式模型：

$$Y_i = \alpha + \int_T \gamma_1(s) X_i^c(s)\,\mathrm{d}s + \cdots + \int_{T^p} \gamma_p(t_1,\cdots,t_p) X_i^c(t_1)\cdots X_i^c(t_p)\,\mathrm{d}t_1\cdots\mathrm{d}t_p + \varepsilon_i, \quad (1.1)$$

其中：$E(\varepsilon_i|X_i)=0$, α 是截矩，γ_j, $1\le j\le p$ 为第 j 阶系数函数. $X_i^c(s)=X_i(s)-\mu_X(s)$ 为中心化的预测子过程，系数函数假定是光滑平方可积的. Yao 和 Müller[11] 主要研究了函数二次模型的估计问题.

2009 年 Shin[12] 考虑了部分函数线性模型

$$Y_i = Z_i^\top\theta + \int_0^1 \beta(t) X_i(t)\,\mathrm{d}t + \varepsilon_i, \quad (1.2)$$

其中 $\beta(\cdot)$ 是区间 $[0,1]$ 上的平方可积函数. Shin[13] 利用 Karhunen-Loève 展开研究了模型 (1.2) 的估计以及理论性质.

而在实际生活中，响应变量不仅线性地依赖于一些非函数型数据，而且非线性地依赖于一些函数型数据. 本文我们提出了下面的部分线性函数多项式回归模型

$$Y_i = Z_i^\top\theta + \int_T \gamma_1(t) X_i^c(t)\,\mathrm{d}t + \cdots + \int_{T^p} \gamma_p(t_1,\cdots,t_p) X_i^c(t_1)\cdots X_i^c(t_p)\,\mathrm{d}t_1\cdots dt_p + \varepsilon_i, \quad (1.3)$$

其中 Z_i 是 q- 维非函数型协变量，$\varepsilon_i\in R^1$ 是独立同分布的随机误差且满足 $E(\varepsilon_i|X_i,Z_i)=0$. 显然，模型 (1.1) 和 (1.2) 是模型 (1.3) 的特殊情形.

对模型 (1.3), 如果包含的非函数型协变量很多，那么，对这些协变量进行变量选择是必要的，同时，在函数多项式部分，我们也面临着需要选择多项式中重要的阶的问题. 例如，如果一个一阶多项式能够足够反映回归关系，我们希望发展一个方法能探测出二阶或更高的阶能够忽略不计. 为了同时实现这两种目标，我们提出了一个新的模型探测方法. 本文的主要贡献是：首先，我们提出了部分线性函数多项式模型，它包含了函数线性模型，函数多项式模型和部分函数线性模型作为特殊情形. 第二，我们提出了一个新的模型探测方法，它能同时探测出重要的非函数型协变量和函数多项式中重要的阶.

本文的结构如下：在第二部分，我们介绍了模型探测方法. 在第三部分，我们建立了提出的模型探测方法的理论结果. 在第四部分，我们报告了一些模拟结果. 在第五部分，我们用一个实例分析证实了我们提出的方法，所有的证明放在第六部分.

2 联合探测

2.1 模型的化简

假定 $\{(X_i, Z_i, Y_i)_{i=1,\cdots,n}\}$ 是 $\{(X, Z, Y)_{i=1,\cdots,n}\}$ 中的独立同分布观测，其中 $Y_i \in R^1$ 为响应变量，$Z_i \in R^q$ 是 q 维协变量，X_i 是 $L_2(T)$ 中有光滑平均函数 $EX_i(s) = \mu_X(s)$ 和协方差函数 $cov(X_i(s_1), X_i(s_2)) = G_X(s_1, s_2)$ 的随机函数．根据 Karhunen-Loève 展开，X_i 能表示为

$$X_i(s) = \mu_X(s) + \sum_{j=1}^{\infty} \eta_{ij}\phi_j(s), \tag{2.1}$$

其中 η_{ij} 是满足 $E(\eta_{ij}) = 0$ 和 $\operatorname{var}(\eta_{ij}) = \nu_j$ 的不相关的随机变量， ϕ_j 是正交的特征函数序列， ν_j 是 $G_X(s_1, s_2)$ 的特征值．由于特征函数构成了一个完备的基函数空间，故模型中的系数函数也能表示成

$$\gamma_j(t_1, \cdots, t_j) = \sum_{k_1,\cdots,k_j=1}^{\infty} b_{k_1,\cdots,k_j}\phi_{k_1}(t_1)\cdots\phi_{k_j}(t_j), \tag{2.2}$$

其中 $(b_{j_1,\cdots,j_p})_{j_1,\cdots,j_p=1,2,\cdots}$ 满足 $\sum\limits_{j_1,\cdots,j_p} b_{j_1,\cdots,j_p} < \infty$.

将 (2.1) 和 (2.2) 带入 (1.3) 并利用特征函数的正交性，模型 (1.3) 能化简为

$$Y_i = Z_i^\top\theta + \sum_{j=1}^{p} U_{i,j}^\top\beta_j + \varepsilon_i, \tag{2.3}$$

其中：

$$\begin{aligned}
&\beta_1 = (b_1, \cdots, b_K)^\top,\\
&\beta_2 = (b_{1,1}, b_{2,1}, b_{2,2}\cdots, b_{K,K})^\top, \cdots, \beta_p = (b_{1,1,\cdots,1}, \cdots, b_{K,K,\cdots,K})^\top,\\
&U_{i,1} = (\eta_{i,1}, \cdots, \eta_{i,K})^\top,\\
&U_{i,2} = (\eta_{i,1}\eta_{i,1}, \eta_{i,2}\eta_{i,1}, \eta_{i,2}\eta_{i,2}, \cdots, \eta_{i,K}\eta_{i,K})^\top, \cdots,\\
&U_{i,p} = (\eta_{i,1}\eta_{i,1}\cdots\eta_{i,1}, \cdots, \eta_{i,K}\eta_{i,K}\cdots\eta_{i,K})^\top.
\end{aligned}$$

此处， K 是主成分的个数，对 K 的选取，我们可以选用交叉验证法或伪 -BIC 方法，我们在模拟中采用后者，在细节上参见 Yao[5] 等.

2.2 重要变量和重要的阶的联合探测

事实上, 我们还不能利用 (2.3) 作推断, 因为 (2.3) 中函数主成分得分 η_{ij} 是未知的. 我们必须首先估计函数主成分得分， η_{ij} 的估计可以利用函数主成分分析方法得到，对这一估计程序，参见 [5]. 基于估计的函数主成分得分，可得：

$$\begin{aligned}
&\widehat{U}_{i1} = (\widehat{\eta}_{i1}, \cdots, \widehat{\eta}_{iK})^\top,\\
&\widehat{U}_{i2} = (\widehat{\eta}_{i1}\widehat{\eta}_{i1}, \widehat{\eta}_{i2}\widehat{\eta}_{i1}, \widehat{\eta}_{i2}\widehat{\eta}_{i2}, \cdots, \widehat{\eta}_{iK}\widehat{\eta}_{iK})^\top,
\end{aligned}$$

$$\vdots$$

$$\widehat{U}_{ip} = (\widehat{\eta}_{i1}\widehat{\eta}_{i1}\cdots\widehat{\eta}_{i1},\cdots,\widehat{\eta}_{iK}\widehat{\eta}_{iK}\cdots\widehat{\eta}_{iK})^\top.$$

为了简单，我们定义：$\Omega_1=\{1,\cdots,p\}$ 和 $\Omega_2=\{1,\cdots,q\}$. 在这一部分，我们提出一个探测程序去识别以下两个集合： $A=\{j:\|\gamma_j\|\neq 0\}$ 和 $B=\{k:|\theta_k|\neq 0\}$. 注意到： $U_{i,j}^\top\beta_j$ 相应于函数多项式部分的第 j 项，则识别集合 A 等价于识别集合 $\{j:\|\beta_j\|\neq 0\}$. 为了同时识别 A 和 B, 我们考虑以下的目标函数

$$\begin{aligned} Q(\beta,\theta) =&\sum_{i=1}^{n}[Y_i - Z_i^\top\theta - \sum_{j=1}^{p}\widehat{U}_{ij}^\top\beta_j]^2 + \sum_{l=1}^{q}\mu_k|\theta_k| + \sum_{j=1}^{p}c_j\lambda_j\|\beta_j\| \\ \equiv&L(\beta,\theta) + \sum_{l=1}^{q}\mu_k|\theta_k| + \sum_{j=1}^{p}c_j\lambda_j\|\beta_j\|, \end{aligned} \tag{2.4}$$

其中 λ_j $(j=1,\cdots,p)$, μ_k $(k=1,\cdots,q)$ 是光滑参数， c_j $(j=1,\cdots,p)$ 是对群体大小的一个调节，一个简单的选择是 $c_j=|\beta_j|^{1/2}$ ($|\beta_j|$ 表示 β_j 中元素的个数). 最小化上面的目标函数 $Q(\beta,\theta)$, 我们可以获得 agLasso 估计 $\widehat{\beta}_j$ 和 aLasso 估计 $\widehat{\theta}_k$.

于是， $\gamma_l(t_1,\cdots,t_l)$, $1\le l\le p$ 的估计 $\widehat{\gamma}_l(t_1,\cdots,t_l)$, $1\le l\le p$ 能表示为：

$$\widehat{\gamma}_l(t_1,\cdots,t_l) = \sum_{j_1\le\cdots\le j_l}^{K}\widehat{b}_{j_1,\cdots,j_l}\widehat{\phi}_{j_1}(t_1)\cdots\widehat{\phi}_{j_l}(t_l), \qquad 1\le l\le p. \tag{2.5}$$

2.3 光滑参数选择

为了执行上述探测，我们必须选取恰当的光滑参数 λ_j $(1\le j\le p)$ 和 μ_k $(1\le k\le q)$. 类似如 [13–15] 的思想，我们采用下式来选择光滑参数

$$\lambda_j = \frac{\lambda_n}{\|\widetilde{\beta}_j\|}, \qquad \mu_k = \frac{\mu_n}{|\widetilde{\theta}_k|},$$

其中 $\widetilde{\beta}=(\widetilde{\beta}_1^\top,\cdots,\widetilde{\beta}_p^\top)^\top$ 和 $\widetilde{\theta}=(\widetilde{\theta}_1,\cdots,\widetilde{\theta}_q)$ 是未被惩罚的最小二乘估计，则原来的光滑参数的选取问题转化为一个二维参数 (λ_n,μ_n) 的选取问题.

令 df 是 $\widehat{\beta}$ 中的非零系数. 我们考虑了以下两个选择标准来选取 (λ_n,μ_n):

$$\mathrm{AIC} = \log\widehat{\sigma}_n^2 + \frac{2}{n_0}\mathrm{d}f, \qquad \mathrm{BIC} = \log\widehat{\sigma}_n^2 + \frac{\log(n_0)}{n_0}\mathrm{d}f,$$

其中

$$\widehat{\sigma}_n^2 = \frac{1}{n_0}\sum_{i=q+1}^{n}\left[Y_i - Z_i^\top\widehat{\theta} - \sum_{j=1}^{p}\widehat{U}_{ij}^\top\widehat{\beta}_j\right]^2,$$

则参数 (λ_n,μ_n) 能够最小化 AIC, BIC 进行选择.

3 理论性质

为了证明估计的相合性和模型探测的相合性，我们先介绍几个条件：

(a) X_i 有有界四阶矩，即： $\int E(X_i)^4 < \infty$.

(b) 存在正常数 C 和 $\alpha > 1$, 满足： $\nu_k - \nu_{k+1} \geq C^{-1}K^{-\alpha-1}$, 其中 $\{\nu_j\}_{j=1,2,\cdots}$ 是 X 的协方差函数的特征值.

(c) $n^{-1}K^{3\alpha+3} \to 0$.

(d) 定义： $U_i = (U_{i,1}^\top, \cdots, U_{i,p}^\top)^\top$, $W_i = (Z_i^\top, U_i^\top)^\top$, $W = (W_1, \cdots, W_n)$. 也定义 $\rho_{\min}(W^\top W)$ 为 $W^\top W$ 的最小特征值. 我们需要： $\rho_{\min}(W^\top W) = O_p(K^{-\alpha})$.

(e) 假设 $\sum\limits_{j\in A} c_j^2\lambda_j^2 = o(K^{5\alpha+2})$ 且 $\sum\limits_{k\in B} \mu_k = o(K^{5\alpha+2})$.

(f) $|b_{k_1,\cdots,k_l}| \leq C(k_1 \times \cdots \times k_l)^{-\alpha/l}$, $l = 1, \cdots, p$.

(g) $n^{1/2}K^{-2\alpha-2} \inf\limits_{j\in A^c} c_j\lambda_j \to \infty$ 且 $n^{1/2}K^{-2\alpha-2} \inf\limits_{k\in B^c} \gamma_k \to \infty$.

注 1 条件 (a) 是函数数据分析中的一个正则条件. 条件 (b) 需要相邻的特征值之差不是太小，也暗含着每个 ν_k 要大于一个常数乘以 $k^{-\alpha}$, 我们需要这个条件来得到 $\widehat{\phi}_j(t) - \phi_j(t)$ 和 $\widehat{\eta}_j - \eta_j$, $j = 1, \cdots, K$ 的界. 条件 (d) 类似于 [16]. 为了得到估计的相合性，我们需要条件 (e), 条件 (f) 暗含着系数 $b_{k_1,\cdots,k_l}$ 退化. 条件 (g) 是为了确保模型探测的相合性.

下面，我们给出模型探测方法的理论性质. 定理 1 建立了估计的相合性，定理 2 建立了模型探测的相合性，所有的证明放在第六部分.

定理 1 假设 (a)–(f) 成立, 我们有：$\|\widehat{\gamma}_j - \gamma_j\| = o_p(1)$, $j \in \Omega_1$, $\widehat{\theta}_k - \theta_k = o_p(1)$, $k \in \Omega_2$.

定义： $\widehat{A}_n = \{j : \|\widehat{\gamma}_j\| \neq 0\}$ 和 $\widehat{B}_n = \{j : |\widehat{\theta}_k| \neq 0\}$.

定理 2 假设 (a)–(g) 成立，我们有： $P(\widehat{A}_n = A) \longrightarrow 1$, $P(\widehat{B}_n = B) \longrightarrow 1$.

表 1 例 1 的模型探测结果

样本量	方法	变量选择			多项式阶的选择			模型选择的正确率
		O	C	U	O	C	U	C
I								
100	BIC	0.1200	0.8800	0.0000	0.3050	0.6950	0.0000	0.6250
	AIC	0.2525	0.7475	0.0000	0.3100	0.6900	0.0000	0.5525
200	BIC	0.1300	0.8700	0.000	0.2450	0.7550	0.0000	0.6650
	AIC	0.2800	0.7200	0.000	0.2650	0.7350	0.0000	0.5575
400	BIC	0.0625	0.9375	0.0000	0.2525	0.7675	0.0000	0.7475
	AIC	0.2100	0.7900	0.0000	0.02725	0.7225	0.0000	0.6025
II								
100	BIC	0.4425	0.5575	0.0000	0.2825	0.7175	0.0000	0.4325
	AIC	0.5925	0.4075	0.0000	0.3050	0.6950	0.0000	0.4125
200	BIC	0.4225	0.5775	0.000	0.2425	0.7575	0.0000	0.4425
	AIC	0.5175	0.4825	0.000	0.2475	0.7525	0.0000	0.4125
400	BIC	0.4125	0.5825	0.0000	0.2475	0.7825	0.0000	0.4800
	AIC	0.4950	0.5050	0.0000	0.2625	0.7375	0.0000	0.4000

4 模拟研究

为了评价提出的方法的有限样本表现，我们做了一些模拟研究. 为了评估模型探测的表现，我们计算了在 200 次模拟中正确拟合 (即探测出真正的模型，记为 ‘C’), 欠拟合 (即至少一个重要的多项式的阶被忽略记为 ‘U’) 和过分拟合 (至少一个多余的多项式的阶被包含进模型中来，记为 ‘O’) 的百分比. 为了评估预测的表现，我们取样本量 n 分别为 100, 200 和 400 的样本为训练样本，分别计算采用 BIC 标准和 AIC 标准选择光滑参数时得到的估计和 oracle 估计 (即基于真实的模型，记为 ‘OE’), 未被惩罚的估计 (即基于全模型，记为 ‘UE’). 然后，取样本量为 100 的样本作为测试样本，分别计算了上叙 4 个方法的相对预测误差 (RPE)

$$\mathrm{RPE}_i = (Y_i - \widehat{Y}_i)^2 / Y_i^2,$$

的 $25\%, 50\%$ 和 75% 分位点，其中： Y_i 和 $\widehat{Y}_i$ 分别为测试样本的真实值和预测值.

在这一部分，我们考虑函数型数据由下列过程所生成

$$X_i(s) = \sum_{j=1}^{2} \eta_{ij}\phi_j(s),$$

其中： $\phi_1(s) = -\sqrt{2}\cos(2\pi s),\ s \in [0,1],\ \phi_2(s) = \sqrt{2}\sin(2\pi s),\ s \in [0,1].$ η_{ij} 从正态分布 $N(0, \lambda_j)$ 生成，其中： $\lambda_1 = 1,\ \lambda_2 = 1/2$. 对每个过程的观测，我们在 [0, 1] 区间中等距离的选取 50 个点， $(Z_{i1}, \cdots, Z_{i7})^\top$ 由均值为 0 协方差为 $\mathrm{cov}\,(Z_{ij_1}, Z_{ij_2}) = 0.5^{|j_1 - j_2|}$, $1 \leq j_1,\ j_2 \leq 7$ 的多元正态分布生成.

表 2 例 2 的模型探测结果

样本量	方法	变量选择			多项式阶的选择			模型选择的正确率
		O	C	U	O	C	U	C
III								
100	BIC	0.1475	0.8525	0.0000	0.2850	0.7150	0.0000	0.6150
	AIC	0.2650	0.7350	0.0000	0.3000	0.7000	0.0000	0.5250
200	BIC	0.0875	0.9125	0.000	0.1900	0.8100	0.0000	0.7475
	AIC	0.2075	0.7925	0.000	0.2025	0.7975	0.0000	0.6375
400	BIC	0.1200	0.9200	0.0000	0.2025	0.8275	0.0000	0.7500
	AIC	0.2300	0.7700	0.0000	0.02175	0.7825	0.0000	0.5975
IV								
100	BIC	0.1225	0.8775	0.0000	0.4425	0.5575	0.0000	0.4900
	AIC	0.2775	0.7225	0.0000	0.4650	0.5350	0.0000	0.4050
200	BIC	0.0800	0.9200	0.000	0.3275	0.6725	0.0000	0.6350
	AIC	0.1925	0.8075	0.000	0.3400	0.6600	0.0000	0.5575
400	BIC	0.1250	0.9350	0.0000	0.2875	0.7125	0.0000	0.6550
	AIC	0.2575	0.7425	0.0000	0.3100	0.6900	0.0000	0.5275

例 1 部分线性函数二次模型 在这个例子中 Y_i 由下列模型生成：

$$\mathrm{I}: Y_i = Z_i^\top\theta + 2\eta_{i1} + 1.5\eta_{i2} + e_i,$$
$$\mathrm{II}: Y_i = Z_i^\top\theta + \frac{1}{2}\eta_{i1}^2 + \eta_{i2}^2 + e_i,$$

其中： e_i 由 $N(0,0.5)$ 生成且 $\theta=(2,1.5,1,0,0,0,0)$. 我们容易看到：模型 I 是没有函数二次项的部分线性函数二次模型，模型 II 是没有函数线性项的部分线性函数二次模型.

表 3 例 1 的预测结果

样本量	方法	25		50		75	
				I			
100	BIC	0.0279	(0.0607)	0.1445	(0.2978)	0.7342	(1.4842)
	AIC	0.0279	(0.0605)	0.1443	(0.2973)	0.7368	(1.4888)
	OE	0.0290	(0.0634)	0.1486	(0.3066)	0.7692	(1.5574)
	UE	0.0289	(0.0619)	0.1470	(0.2976)	0.7571	(1.5024)
200	BIC	0.0202	(0.0506)	0.1016	(0.2374)	0.5212	(1.1755)
	AIC	0.0202	(0.0504)	0.1019	(0.2381)	0.5215	(1.1749)
	OE	0.0211	(0.0526)	0.1043	(0.2442)	0.5473	(1.2406)
	UE	0.0205	(0.0504)	0.1027	(0.2359)	0.5298	(1.1815)
400	BIC	0.0161	(0.0413)	0.0863	(0.2063)	0.4419	(1.0150)
	AIC	0.0161	(0.0412)	0.0865	(0.2065)	0.4417	(1.0130)
	OE	0.0169	(0.0435)	0.0889	(0.2128)	0.4609	(1.0557)
	UE	0.0164	(0.0425)	0.0872	(0.2069)	0.4448	(1.0067)
				II			
100	BIC	0.0071	(0.0028)	0.0399	(0.0134)	0.2250	(0.0897)
	AIC	0.0072	(0.0029)	0.0402	(0.0134)	0.2281	(0.0926)
	OE	0.0072	(0.0028)	0.0396	(0.0131)	0.2315	(0.0961)
	UE	0.0075	(0.0030)	0.0415	(0.0131)	0.2421	(0.0984)
200	BIC	0.0066	(0.0025)	0.0382	(0.0118)	0.2197	(0.0908)
	AIC	0.0067	(0.0025)	0.0382	(0.0118)	0.2208	(0.0935)
	OE	0.0069	(0.0026)	0.0392	(0.0122)	0.2276	(0.0936)
	UE	0.0066	(0.0025)	0.0386	(0.0120)	0.2252	(0.0924)
400	BIC	0.0063	(0.0023)	0.0359	(0.0108)	0.2094	(0.0825)
	AIC	0.0063	(0.0023)	0.0359	(0.0108)	0.2091	(0.0817)
	OE	0.0066	(0.0024)	0.0370	(0.0114)	0.2175	(0.0860)
	UE	0.0064	(0.0023)	0.0363	(0.0107)	0.2093	(0.0850)

例 2 部分线性函数三次模型 在这个例子中 Y_i 由下列模型生成：

$$\mathrm{III}: Y_i = Z_i^\top\theta + 2\eta_{i1} + \eta_{i2} + \eta_{i1}^2 + \eta_{i2}^2 + e_i,$$
$$\mathrm{IV}: Y_i = Z_i^\top\theta + 2\eta_{i1} + \eta_{i2} + e_i,$$

其中： e_i 由 $N(0,0.5)$ 生成且 $\theta=(2,1.5,1,0,0,0,0)$. 模型 III 是没有函数三次项的部分线性函数三次模型，模型 IV 是没有函数二次项和三次项的部分线性函数三次模型.

对模型探测，我们分别使用 BIC 和 AIC 选择光滑参数，例 1 和例 2 中正确拟合、过分拟合和欠拟合的百分比总结在表 1 和表 2，从表 1 和表 2 可以看出，由于使用 BIC 标准选择光滑参数时，能够更加准确的选择重要的非函数协变量和重要的函数多项式的阶，于是，得到的模型探测结果好于 AIC 标准. AIC 标准经常选择更多的多项式的阶和更多的非函数协变量. 而且，用 BIC 标准选择光滑参数时，正确选择的百分比随着样本容量的增加而增大.

例 1 和 2 的预测结果分别总结在表 3 和表 4. 表中报告了基于 200 次模拟时 RPE 的 25%, 50% 和 75% 分位点的平均值（括号中为相应的标准差）. 从表 3 和表 4 可以得出以下结论：所有估计的相对误差的均值和标准差都随样本量的增加而减小. 基于 BIC 标准和 AIC 标准得到的估计比未被惩罚得估计有更好的预测表现，这也证实了降低模型的复杂度能提高预测的精度.

表 4 例 2 的预测结果

样本量	方法	25		50		75	
				III			
100	BIC	0.0294	(0.0639)	0.1376	(0.2732)	0.7380	(1.4931)
	AIC	0.0295	(0.0644)	0.1379	(0.2735)	0.7390	(1.4946)
	OE	0.0306	(0.0665)	0.1411	(0.2801)	0.7803	(1.5811)
	UE	0.0301	(0.0645)	0.1415	(0.2757)	0.7623	(1.5086)
200	BIC	0.0234	(0.0552)	0.1027	(0.2189)	0.5381	(1.2401)
	AIC	0.0234	(0.0551)	0.1022	(0.2182)	0.5378	(1.2399)
	OE	0.0244	(0.0574)	0.1059	(0.2268)	0.5378	(1.3063)
	UE	0.0235	(0.0548)	0.1040	(0.2183)	0.5489	(1.2664)
400	BIC	0.0201	(0.0515)	0.0907	(0.2102)	0.4643	(1.0276)
	AIC	0.0201	(0.0515)	0.0909	(0.2107)	0.4641	(1.0275)
	OE	0.0209	(0.0538)	0.0936	(0.2160)	0.4931	(1.1049)
	UE	0.0202	(0.0517)	0.0908	(0.2088)	0.4749	(1.0527)
				IV			
100	BIC	0.0300	(0.0658)	0.1582	(0.3365)	0.8095	(1.7065)
	AIC	0.0301	(0.0661)	0.1578	(0.3347)	0.8153	(1.7227)
	OE	0.0313	(0.0687)	0.1635	(0.3495)	0.8153	(1.7987)
	UE	0.0320	(0.0690)	0.1647	(0.3391)	0.8564	(1.7611)
200	BIC	0.0247	(0.0610)	0.1127	(0.2498)	0.5810	(1.2947)
	AIC	0.0246	(0.0608)	0.1128	(0.2496)	0.5794	(1.2924)
	OE	0.0258	(0.0640)	0.1168	(0.2604)	0.6106	(1.3653)
	UE	0.0254	(0.0611)	0.1151	(0.2469)	0.5882	(1.2756)
400	BIC	0.0211	(0.0550)	0.1006	(0.2414)	0.5177	(1.1823)
	AIC	0.0211	(0.0549)	0.1006	(0.2414)	0.5183	(1.1848)
	OE	0.0211	(0.0575)	0.1037	(0.2483)	0.5468	(1.2487)
	UE	0.0212	(0.0545)	0.1027	(0.2447)	0.5243	(1.1919)

5 实例分析

在这一节，我们将提出的方法分析食品工业中的谱数据．谱数据是通过光谱仪对猪肉样本进行测量，得到波段在 850–1050 之间的谱曲线，每两个波段进行一次测量，从而测得 100 个光谱数据 (记为 $\{X_1,\cdots,X_{100}\}$)．同时，利用分析化学，也测得了样本的水分百分比，脂肪百分比和蛋白质百分比．样本量为 215, 其中，前面 160 个用来估计，后面 55 个用作测试样本．我们的目的是利用光谱数据，水分百分比和蛋白质百分比来预测脂肪的百分比．为了去掉漂移影响，我们用原始谱曲线的二阶导数来作为函数型数据．根据 K 的选取方法，我们选取前面的 3 个主成分，这三个主成分的贡献率超过了 97%.

我们首先考虑了以下模型：

$$Y_i = Z_i^{\top}\theta + \int_T \gamma_1(s)X_i^c(s)\,\mathrm{d}s + \cdots + \int_{T^p}\gamma_p(t_1,\cdots,t_p)X_i^c(t_1)\cdots X_i^c(t_p)\,\mathrm{d}t_1\cdots\mathrm{d}t_p + \varepsilon_i, \quad (5.1)$$

其中： Z_i 是 4 维非函数协变量： (水分百分比，蛋白质百分比， X_{49}, X_{100}).

首先，我们分别使用 BIC 和 AIC 选择光滑参数对模型进行联合探测，在函数多项式部分，模型选择的结果为二次多项式．这说明一次多项式不足以反映函数型数据和响应变量的关系，这一结论也与 Ferraty[6] 等的结论一致．在非函数型部分，得到的结论是 X_{100} 为不重要的协变量，这一结论与 Li[10] 等和 Amato[17] 等一致，在他们对这一实例的分析中，仅仅考虑了波段在 902 到 1028 这一段，因为这一段包含了最多的信息，而剩下的几乎可以忽略不计．另外， X_{49} 是重要的协变量，这一结论也与 Ferraty[18] 等得到的结果一致．

对预测，我们先比较了提出的估计与未被惩罚的估计 (即基于全模型，记为 ‘UE’) 的预测表现．我们分别利用训练样本计算了采用 BIC 标准和 AIC 标准选择光滑参数时得到的估计和基于全模型得到的估计．然后，分别计算了上述 3 种方法在测试样本中得到的预测值的相对预测误差 (RPE), 并计算了 RPE 的分位点，计算结果显示在表 5. 从表 5 可以看出，惩罚方法明显好于未被惩罚方法．

表 5 实例分析中的预测结果

模型	方法	25	50	75
(5.1)	BIC	0.0028	0.0078	0.0438
	AIC	0.0117	0.0306	0.1354
	UE	0.0116	0.0362	0.1670
(5.2)	BIC	0.0101	0.0501	0.1649
	AIC	0.0101	0.0501	0.1649
	UE	0.0203	0.0619	0.1596
(5.3)	UE	0.0131	0.0561	0.1128

接下来，我们考虑了以下模型：

$$Y_i = \alpha + \int_T \gamma_1(s) X_i^c(s)\,\mathrm{d}s + \int_{T^2} \gamma_2(t_1, t_2) X_i^c(t_1) X_i^c(t_2)\,\mathrm{d}t_1\,\mathrm{d}t_p + \varepsilon_i \tag{5.2}$$

和

$$Y_i = \alpha + \int_T \gamma_1(s) X_i^c(s)\,\mathrm{d}s + \varepsilon_i. \tag{5.3}$$

对模型 (5.2) 和 (5.3), 我们类似的分别使用全模型和惩罚方法计算了 RPE 的分位点，所得结果也显示在表 5. 从表 5 我们也能够看出，提出的惩罚方法好于未被惩罚方法.

6 证明

为了证明估计的相合性和模型选择的相合性，我们先介绍一个引理.

引理 1 假设条件 (a) 和 (b) 成立，则：

$$\begin{aligned}
&|\widehat{\eta}_{ij} - \eta_{ij}| = O_p(K^{\alpha+1}/\sqrt{n}),\\
&\Big|\sum_{i=1}^{n}(\widehat{\eta}_{ij_1}\widehat{\eta}_{ij_2})/n - E(\eta_{ij_1}\eta_{ij_2})\Big| = O_p(K^{\alpha+1}/\sqrt{n}),\\
&\quad \vdots,\\
&\Big|\sum_{i=1}^{n}(\widehat{\eta}_{ij_1}\cdots\widehat{\eta}_{ij_p})/n - E(\eta_{ij_1}\cdots\eta_{ij_p})\Big| = O_p(K^{\alpha+1}/\sqrt{n}),\\
&\|\widehat{\phi}_j(s) - \phi_j(s)\| = O_p\Big(\frac{K^{2\alpha+2}}{n}\Big).
\end{aligned}$$

引理 1 的证明 利用条件 (c) 和 [4] 中的 (5.2), 可得 $|\widehat{\eta}_{ij} - \eta_{ij}| = O_p(K^{\alpha+1}/\sqrt{n})$ 和 $\|\widehat{\phi}_j(s) - \phi_j(s)\| = O_p(\frac{K^{2\alpha+2}}{n})$. 通过直接计算可得

$$\begin{aligned}
&\sum_{i=1}^{n}(\widehat{\eta}_{ij_1}\widehat{\eta}_{ij_2})/n - E(\eta_{ij_1}\eta_{ij_2})\\
=&\Big(\sum_{i=1}^{n}(\widehat{\eta}_{ij_1}\widehat{\eta}_{ij_2})/n - \sum_{i=1}^{n}(\eta_{ij_1}\eta_{ij_2})/n\Big) + \Big(\sum_{i=1}^{n}(\eta_{ij_1}\eta_{ij_2})/n - E(\eta_{ij_1}\eta_{ij_2})\Big)\\
=&B_1 + B_2.
\end{aligned}$$

利用 $|\widehat{\eta}_{ij} - \eta_{ij}| = O_p(K^{\alpha+1}/\sqrt{n})$, 可得：

$$\begin{aligned}
B_1 =&\frac{1}{n}\sum_{i=1}^{n}(\widehat{\eta}_{ij_1}\widehat{\eta}_{ij_2} - \eta_{ij_1}\eta_{ij_2})\\
=&\frac{1}{n}\sum_{i=1}^{n}\big[(\widehat{\eta}_{ij_1} - \eta_{ij_1})\widehat{\eta}_{ij_2} + \eta_{ij_1}(\widehat{\eta}_{ij_2} - \eta_{ij_2})\big]\\
=&O_p(K^{\alpha+1}/\sqrt{n}).
\end{aligned}$$

显然 $B_2=O_p(n^{-1/2})$. 通过与前面相似的道理可以证明引理 1 中其他的结论. 证毕.

引理 2 定义

$$\widehat{U}_i=(\widehat{U}_{i,1}^\top,\cdots,\widehat{U}_{i,p}^\top)^\top,\qquad \widehat{W}_i=(Z_i^\top,\widehat{U}_i^\top)^\top,\qquad \widehat{W}=(\widehat{W}_1,\cdots,\widehat{W}_n).$$

定义 ρ^* 是 $\widehat{W}^\top\widehat{W}$ 的最小特征值，假设条件 (a)–(e) 成立，则： $\rho^*=O_p(nK^{-\alpha})$.

引理 2 是 [16] 的一个直接结论.

定理 1 的证明 定义：$\beta=(\beta_1^\top,\cdots,\beta_p^\top)^\top$, $\widehat{\beta}=(\widehat{\beta}_1^\top,\cdots,\widehat{\beta}_p^\top)^\top$, $\theta=(\theta_1^\top,\cdots,\theta_q^\top)^\top$, $\widehat{\theta}=(\widehat{\theta}_1^\top,\cdots,\widehat{\theta}_q^\top)^\top$, $\delta=(\theta^\top,\beta^\top)^\top$, $\widehat{\delta}=(\widehat{\theta}^\top,\widehat{\beta}^\top)^\top$.

首先，我们证明：对 $j\in\Omega_1$, 有 $\|\widehat{\beta}_j-\beta_j\|=o_p(1)$；对 $k\in\Omega_2$, 有 $|\widehat{\theta}_k-\theta_k|=o_p(1)$.

注意到：

$$\begin{aligned}
&L(\widehat{\beta},\widehat{\theta})-L(\beta,\theta)\\
=&\sum_{i=1}^n[Y_i-Z_i^\top\widehat{\theta}-\widehat{U}_i^\top\widehat{\beta}]^2-\sum_{i=1}^n[Y_i-Z_i^\top\theta-\widehat{U}_i\beta]^2\\
=&\sum_{i=1}^n[(Y_i-Z_i^\top\theta-\widehat{U}_i\beta)+(Z_i^\top\theta-Z_i^\top\widehat{\theta})+(\widehat{U}_i\beta-\widehat{U}_i\widehat{\beta})]^2-\sum_{i=1}^n[Y_i-Z_i^\top\theta-\widehat{U}_i\beta]^2\\
=&2\sum_{i=1}^n[(Y_i-Z_i^\top\theta-\widehat{U}_i\beta)[\widehat{W}_i^\top(\delta-\widehat{\delta})]+\sum_{i=1}^n[\widehat{W}_i^\top(\delta-\widehat{\delta})]^2\\
\equiv&A_1+A_2.
\end{aligned}\tag{6.1}$$

令 $P_{\widehat{W}}=\widehat{W}(\widehat{W}^\top\widehat{W})\widehat{W}^\top$. 因为： $Y_i-Z_i^\top\theta-\widehat{U}_i\beta=\varepsilon_i+(U_i\beta-\widehat{U}_i\beta)+V_i$, 其中：

$$V_i=\sum_{j=K+1}^\infty\gamma_j\eta_{i,j}+\cdots+\sum_{j_1,\cdots,j_p=K+1}^\infty\gamma_{j_1,\cdots,j_p}\eta_{i,j_1}\cdots\eta_{i,j_p}.$$

利用与 [16] 中引理 3 相似的方法，可得：

$$\|P_{\widehat{W}}\varepsilon\|^2=O_p(K),\qquad \|P_{\widehat{W}}(U\beta-\widehat{U}\beta)\|^2=O_p(K^{2\alpha+2}),\qquad \|P_{\widehat{W}}V\|^2=O_p(K^{2\alpha+2}).\tag{6.2}$$

根据 (6.2), 可得：

$$A_1\geq-O_p(K^{2\alpha+2})^{1/2}\|\widehat{W}^\top(\delta-\widehat{\delta})\|\geq-O_p(K^{2\alpha+2})-\frac{1}{2}\|\widehat{W}^\top(\delta-\widehat{\delta})\|^2.\tag{6.3}$$

根据 (6.1) 和 (6.3), 可得：

$$0\geq L(\widehat{\beta},\widehat{\theta})-L(\beta,\theta)\geq-O_p(K^{2\alpha+2})+\frac{1}{2}\|\widehat{W}^\top(\delta-\widehat{\delta})\|^2.\tag{6.4}$$

利用 Cauchy-Schwartz 不等式，可得：

$$n\sum_{j=0}^p c_j\lambda_j\|\widehat{\beta}_j\|+n\sum_{k=0}^q\mu_k|\widehat{\theta}_k|-n\sum_{j=0}^p c_j\lambda_j\|\beta_j\|+n\sum_{k=0}^q\mu_k|\theta_k|$$

$$
\begin{aligned}
&\geq n\sum_{j\in A}c_j\lambda_j(\|\widehat{\beta}_j\|-\|\beta_j\|)+n\sum_{j\in B}\mu_k(|\widehat{\theta}_k|-|\theta_k|)\\
&\geq-(\sqrt{n\rho^*/2}\|\widehat{\beta}-\beta\|)\Big(\frac{2n}{\rho^*}\sum_{j\in A}c_j^2\lambda_j^2\Big)^{1/2}-(\sqrt{n\rho^*/2}\|\widehat{\theta}-\theta\|)\Big(\frac{2n}{\rho^*}\sum_{j\in B}\mu_k^2\Big)^{1/2}\\
&\geq-\frac{n}{\rho^*}\sum_{j\in A}c_j^2\lambda_j^2-\frac{n\rho^*}{4}\|\widehat{\beta}-\beta\|^2-\frac{n}{\rho^*}\sum_{j\in B}\mu_k^2-\frac{n\rho^*}{4}\|\widehat{\theta}-\theta\|^2.
\end{aligned}
\tag{6.5}
$$

根据 (6.4), (6.5) 以及引理 2, 可得：

$$
\begin{aligned}
0\geq&Q(\widehat{\beta},\widehat{\theta})-Q(\beta,\theta)\\
\geq&-O_p(K^{2\alpha+2})+\frac{1}{2}\|\widehat{W}^\top(\delta-\widehat{\delta})\|^2\\
&-\frac{n}{\rho^*}\sum_{j\in A}c_j^2\lambda_j^2-\frac{n\rho^*}{4}\|\widehat{\beta}-\beta\|^2-\frac{n}{\rho^*}\sum_{j\in B}\mu_k^2-\frac{n\rho^*}{4}\|\widehat{\theta}-\theta\|^2\\
\geq&-O_p(K^{2\alpha+2})+\frac{n\rho^*}{4}(\|\widehat{\beta}-\beta\|^2+\|\widehat{\theta}-\theta\|^2)-\frac{n}{\rho^*}\sum_{j\in A}c_j^2\lambda_j^2-\frac{n}{\rho^*}\sum_{k\in B}\mu_k^2.
\end{aligned}
$$

利用假设 (c) 和 (e), 可得：

$$
\|\widehat{\beta}-\beta\|^2+\|\widehat{\theta}-\theta\|^2=O_p\Big(\frac{K^{3\alpha+2}}{n^2}\Big).
$$

于是，我们证得对 $j\in\Omega_1$ 有： $\|\widehat{\beta}_j-\beta_j\|=o_p(1)$, 对 $k\in\Omega_2$ 有 $|\widehat{\theta}_k-\theta_k|=o_p(1)$.

接下来，我们证明 $\|\widehat{\gamma}_l(t_1,\cdots,t_l)-\gamma_l(t_1,\cdots,t_l)\|=o_p(1)$, $1\leq l\leq p$. 对 $l=1$, 注意到：

$$
\begin{aligned}
&\|\widehat{\gamma}_1(t)-\gamma_1(t)\|\leq\sum_{j=1}^K\|\widehat{b}_j\widehat{\phi}_j(t)-b_j\phi_j(t)\|+\|\sum_{j>K}b_j\phi_j(t)\|\\
\leq&\sum_{j=1}^K[|\widehat{b}_j-b_j|\|\phi_j(s)\|+|\widehat{b}_j-b_j|\|\widehat{\phi}_j(s)-\phi_j(s)\|+|b_j|\|\widehat{\phi}_j(s)-\phi_j(s)\|]\\
&+\Big\|\sum_{j>K}b_j\phi_j(s)\Big\|\\
\equiv&F_1+F_2.
\end{aligned}
\tag{6.6}
$$

根据 (6.1) 和引理 1, 可得：

$$
F_1=O_p\Big(\frac{K^{5\alpha/2+1}}{n}+\frac{K^{2\alpha+3}}{n}\Big).\tag{6.7}
$$

利用假设 (f), 可得：

$$
F_2=O_p\Big(\frac{K^{2\alpha+2}}{n}\Big).\tag{6.8}
$$

由 (6.6), (6.7), (6.8) 和假设 (c), 可得： $\|\widehat{\gamma}_1(t)-\gamma_1(t)\|=o_p(1)$.

类似的，可证： $\|\widehat{\gamma}_l(t_1,\cdots,t_l)-\gamma_l(t_1,\cdots,t_l)\|=o_p(1)$, $2\leq l\leq p$. 证毕.

定理 2 的证明 假设存在 $\delta^*=(\widehat{\beta}^{*\top},\widehat{\theta}^{*\top})^\top$ 满足：对 $j\in A$, 有 $\widehat{\beta}^*_j=\widehat{\beta}_j$, 且对 $k\in B$, 有 $\widehat{\theta}^*_k=\widehat{\theta}_k$, 对 $j^{'}\in A^c$, 有 $\widehat{\beta}^*_{j^{'}}=0$ 且对 $k^{'}\in B^c$ 有 $\widehat{\theta}^*_{k^{'}}\neq 0$. 令

$$\varpi^*=\sum_{i=1}^n \widehat{W}_i(\widehat{W}_i^\top\widehat{W}_i)^{-1}\widehat{W}_i^\top(Y_i-Z_i^\top\widehat{\theta}^*-\widehat{U}_i^\top\widehat{\beta}^*),$$

则有： $\|\varpi^*\|^2\leq 2\|\varpi\|^2+2\|\widehat{U}^\top(\widehat{\beta}^*-\beta)\|^2+2\|Z^\top(\widehat{\theta}^*-\theta)\|^2=O_p(K^{2\alpha+2})$.

通过与定理 1 类似的方法，可得：

$$\begin{aligned}
&Q(\widehat{\beta},\widehat{\theta})-Q(\widehat{\beta}^*,\widehat{\theta}^*)\\
=&2\sum_{i=1}^n\left[(Y_i-Z_i^\top\widehat{\theta}^*-\widehat{U}_i\widehat{\beta}^*)\right]\left[\widehat{W}_i^\top(\widehat{\delta}-\delta^*)\right]+\sum_{i=1}^n\left[\widehat{W}_i^\top(\widehat{\delta}-\delta^*))\right]^2\\
&+n\sum_{j=0}^p c_j\lambda_j\|\widehat{\beta}_j\|+n\sum_{k=0}^q\mu_k|\widehat{\theta}_k|-n\sum_{j=0}^p c_j\lambda_j\|\widehat{\beta}^*_j\|-n\sum_{k=0}^q\mu_k|\widehat{\theta}^*_k|\\
\geq&-O_p(K^{2\alpha+2})\sqrt{n}\sum_{j\in A^c\cup B^c}\|\widehat{\delta}_j\|+n\sum_{j\in A^c}c_j\lambda_j\|\widehat{\beta}_j\|+n\sum_{k\in B^c}\mu_k|\widehat{\theta}_k|.
\end{aligned}$$

根据假设 (g), 可得： $Q(\widehat{\beta},\widehat{\theta})-Q(\widehat{\beta}^*,\widehat{\theta}^*)>0$ 以概率趋向于 1 成立，这将与 $Q(\widehat{\beta},\widehat{\theta})$ 最小相矛盾，于是，证毕.

参 考 文 献

[1] Ramsay J, Silverman B, Functional Data Analysis, Second ed. Berlin: Springer, 2005

[2] Ferraty F, Vieu P. Nonparametric Functional Data Analysis. New York: Springer, 2006

[3] Cai T T, Hall P. Prediction in functional linear regression. *Ann. Statist.*, 2006, 34: 2159–2179

[4] Hall P, Horowitz J L. Methodology and convergence rates for functional linear regression. *Ann.Statist*, 2007, 35: 70–91

[5] Yao F, Müller H G, Wang J L. Functional data analysis for sparse longitudinal data. *J. Amer. Statist. Assoc.*, 2005, 100: 577–590

[6] Ferraty F, Keilegom I, Vieu P. Regression when both response and predictor are functions. *Journal of Multivariate Analysis*, 2012, 109: 10–28

[7] Müller H G, Stadtmüller U. Generalized functional linear models. *Ann. Statist.*, 2005, 33: 774–805

[8] Ait-Saïdi A, Ferraty F, Kassa R, Vieu P. Cross-validated estimations in the single-functional index model. *Statistics*, 2008, 42: 475–494

[9] Chen D, Hall P, Müller H G. Single and multiple index functional regression models with nonparametric link. *Ann. Statist.*, 2011, 39: 1720–1747

[10] Li Y H, Wang Naisyin, Carroll R J. Generalized functional linear models with semiparametric single-index interactions. *J. Amer. Statist. Assoc.*, 2010, 105: 621–633

[11] Yao F, Müller H G. Functional quadratic regression. *Biometrika*, 2010, 97: 49–64

[12] Shin H. Partial functional linear regression. *Journal of Statistical Planning and Inference*, 2009, 139; 3405–3418

[13] Zou H. The Adaptive LASSO and Its Oracle Properties. *Journal of the Ameriean Statistical Association*, 2006, 101: 1418–1429

[14] Zhang H H, Lu W. Adaptive lasso for Cox's proportional hazard model. *Biometrika*, 2007, 94: 691–703

[15] Wang H, Xia Y. Shrinkage Estimation of the Varying Coefficient Model. *Journal of the Ameriean Statistical Association*, 2009, 104: 747–757

[16] Lian H. Shrinkage estimation and selection for multiple functional regression. *Statistica Sinica*, 2013, 23: 51–74

[17] Amatoa U, Antoniadisb A, Feisa I D. Dimension reduction in functional regression with applications. *Computational Statistics and Data Analysis*, 2006, 50: 2422–2446

[18] Ferraty F, Hall P, Vieu P. Most predictive design points for functional data predictor. *Biometrika*, 2010, 97: 807–824

Joint Detection for Partial Linear Functional Polynomial Model

ZHANG Tao

(*School of Science, Guangxi University of Science and Technology, Liuzhou* 545006, *China*)

WAN Yanling

(*School of Social Sciences, Guangxi University of Science and Technology, Liuzhou* 545006, *China*)

(*E-mail: ylwanpiano@126.com*)

WANG Zhiwen

(*School of Computer Science and Communication Engineering,*

Guangxi University of Science and Technology, Liuzhou 545006, *China*)

Abstract In this article, we propose a general class of semiparametric rates models for recurrent event data, which includes the proportional rates model and a semiparametric additive rates model as special cases. For the inference on the model parameters, estimating equation approaches are developed. The consistency and asymptotic normality properties of the proposed estimators are established.

Key words functional data; functional polynomial regression model; principal component; joint detection

MR(2000) Subject Classification 62G05; 62N015

Chinese Library Classification O212.7