

# 部分線形モデルの差分推定量の漸近理論

劉慶豊\*

Asymptotic Theory for Difference-based Estimator  
of Partially Linear Models

Qingfeng Liu\*

## 概要

本論文は部分線形モデルの重み付き差分法による推定量の漸近理論に関して論じる。一定の正則条件のもとで推定量の漸近正規性を確認する。さらに、漸近的にセミパラメトリックモデルの効率性を達成するための差分の次数のオーダーを明らかにし、重み付き差分の重みが満たすべき制約を明らかにする。

## 1 はじめに

部分線形モデルは線形モデルに非線形の部分を取り入れることにより、モデルの表現力が豊かになり、実世界にあるもっと複雑な現象を表現できるようになった。その一般的な表現は

$$Y_i = \alpha(\mathbf{U}_i) + \mathbf{X}_i' \boldsymbol{\beta} + \varepsilon_i, \quad i = 1, \dots, n. \quad (1)$$

となる。ただし、 $Y_i \in \mathbb{R}$ 、 $\mathbf{U}_i \in \mathbb{R}^r$ 、 $\mathbf{X}_i \in \mathbb{R}^l$ は観測できる確率変数と確率変数の列ベクトルである。 $\boldsymbol{\beta} \in \mathbb{R}^l$ は未知母数のベクトルで、 $\alpha(\mathbf{U}_i)$ は $\mathbb{R}^r \rightarrow \mathbb{R}$ の $\mathbf{U}_i$ により生成された $\sigma$ -fieldに関して可測な未知関数である。 $\varepsilon_i$ が平均0分散 $\sigma_\varepsilon^2$ のi.i.d.な確率変数である。本論文は線形部分の係数 $\boldsymbol{\beta}$ に着目し、 $\boldsymbol{\beta}$ の推定量の漸近理論に関して議論する。

部分線形モデルの推定法は様々存在するが、その中特にRobinson(1988)の方法は $\boldsymbol{\beta}$ の推定量に関してセミパラメトリックモデルの効率性を持つ、よく利用される方法の一つである。Robinson(1988)の方法では、まずモデルを

$$Y_i - E(Y_i | \mathbf{U}_i) = [\mathbf{X}_i' - E(\mathbf{X}_i' | \mathbf{U}_i)] \boldsymbol{\beta} + \varepsilon_i \quad (2)$$

\*小樽商科大学商学部、〒047-8501 小樽市緑3丁目5番21号; E-mail: qliu@res.otaru-uc.ac.jp

と書き換え，そして，モデル式の中の条件付期待値をノンパラメトリックな方法で推定し，得られた推定量をモデルに代入して，近似的な線形モデルを得る。最終的に最小二乗法で $\beta$ を推定する。もう一つ代替的な方法として差分法を利用した推定方法はPowell(1987)とAhn and Powell(1993)などにより提案された。差分法の基本的な発想は重み付き差分をとることによりモデルのノンパラメトリックな部分の影響を取り除くことである。このような差分法の発想はさらにFan and Wu(2008)とLiu et al. (2009)などにより変係数部分線形モデルに適用されている。差分法はRobinson(1988)の方法と比べて， $\beta$ を推定するためには条件付期待値のノンパラメトリックな推定が必要とされない。さらに，差分の次数を大きくすることでセミパラメトリックモデルの効率性に近づくことができる。Yatchew(1997)は差分の次数を増やすことで差分法の効率性が改善されることを示したが，漸近的にセミパラメトリックモデルの効率性を達成するための差分の次数のオーダーを明示的に提示していない。本論文の中心的な結果は $\beta$ の推定量の漸近的な性質に関するものである。特に，漸近的にセミパラメトリックモデルの効率性を達成するための差分の次数のオーダーを明らかにし，さらに重み付き差分の重みが満たすべき制約を明らかにして，Hall et al.(1990)により提案されたspike weightがこの制約を満たしていることを示す。

本論文の残りの部分の構成は次のようになる。第2章で差分推定量の定義を与える。第3章で論文の中心である差分推定量の漸近理論に関して議論していく。そして，第4章で有限標本の性質を確認するためのモンテカルロ実験の結果を示してから，最後の章で結論を述べる。定理の証明などは補章にある。

## 2 差分推定量

モデル(1)では $\alpha(\cdot)$ が未知関数であるため，普通の線形モデルのように推定できない。そこで，部分線形モデルの差分推定量はPowell(1987)とAhn and Powell(1993)などにより提案された。重み付き差分を取ることにより，モデル(1)の非線形でかつノンパラメトリックな部分， $\alpha(U_i)$ をモデルから取り除いて，そして残りの近似モデルのもとで線形部分の係数 $\beta$ を推定するというのはこの差分推定の基本的な発想である。以下からその定義を述べておこう。

モデル(1)により生成されたサイズが $n$ のサンプル $\{Y_i, U_i, X_i\}_{i=1}^n$ が観測されたとする。まず $U_i$ の大きさの順にgrid nearest neighbor algorithm<sup>1</sup>を利用してすべてのデータをソートする。一般性を失うことなく，ここでは $\{Y_i, U_i, X_i\}_{i=1}^n$

---

<sup>1</sup>補章を参照せよ。

を既にソートされたサンプルとする。モデル(1)に関して $q$ 次の重み付き差分を取ることで、新しいモデルとして

$$\sum_{j=1}^{q+1} \omega_j Y_{i+j-1} = \sum_{j=1}^{q+1} \omega_j \alpha(U_{i+j-1}) + \sum_{j=1}^{q+1} \omega_j \mathbf{X}'_{i+j-1} \boldsymbol{\beta} + \sum_{j=1}^{q+1} \omega_j \varepsilon_{i+j-1} \quad (3)$$

が得られる。ただし、 $\omega_j$ は重みを表し、 $\sum_{j=1}^{q+1} \omega_j = 0$ と $\sum_{j=1}^{q+1} \omega_j^2 = 1$ を満たす。重みの行列表現は

$$\mathbf{W} = \begin{bmatrix} \omega_1 & \omega_2 & \cdots & \omega_{q+1} & 0 & 0 & 0 & \cdots \\ 0 & \omega_1 & \omega_2 & \cdots & \omega_{q+1} & 0 & 0 & \cdots \\ 0 & 0 & \omega_1 & \omega_2 & \cdots & \omega_{q+1} & 0 & \cdots \\ & & & \cdots & & & & \cdots \end{bmatrix} \quad (4)$$

となる。 $\mathbf{W}$ は $(n-q) \times n$ の行列で、その第 $i$ 行の第 $i$ 番目の要素から第 $i+q$ 番目の要素は順番に $\omega_1, \omega_2, \dots, \omega_{q+1}$ となる。それ以外の要素はすべて0である。そして、(3)式の行列表現は

$$\mathbf{WY} = \mathbf{W}\alpha(\mathbf{U}) + \mathbf{WX}\boldsymbol{\beta} + \mathbf{W}\boldsymbol{\varepsilon} \quad (5)$$

となる。ただし、 $\mathbf{Y} = (Y_1, \dots, Y_n)'$ 、 $\mathbf{U} = (U_1, \dots, U_n)'$ 、 $\mathbf{X} = (\mathbf{X}_1, \dots, \mathbf{X}_n)'$ 、 $\boldsymbol{\varepsilon} = (\varepsilon_1, \dots, \varepsilon_n)'$ である。

ここで、 $\alpha(\cdot)$ に関して、ある有限な正の値 $k$ が存在してリプシッツ条件

$$\|\alpha(\mathbf{U}_i) - \alpha(\mathbf{U}_j)\| \leq \kappa \|\mathbf{U}_i - \mathbf{U}_j\| \quad (6)$$

が満たされると仮定する。この条件のもとでは、 $\mathbf{U}_i$ の値域が有界であると仮定すれば、データは $\mathbf{U}_i$ の大きさの順に沿って並べ替えたため、 $i = 1, 2, \dots, n-q$ 、 $j = i+1, i+2, \dots, i+q$ に関して $\alpha(\mathbf{U}_j) \approx \alpha(\mathbf{U}_i)$ となる。そして、重みに関する制約 $\sum_{j=1}^{q+1} \omega_j = 0$ を利用すれば、一定の条件のもとで

$$\mathbf{WY} = \mathbf{WX}\boldsymbol{\beta} + \mathbf{W}\boldsymbol{\varepsilon} + o_p(1) \quad (7)$$

$$\mathbf{WY} \approx \mathbf{WX}\boldsymbol{\beta} + \mathbf{W}\boldsymbol{\varepsilon} \quad (8)$$

となることが言える。そして、 $\boldsymbol{\beta}$ の差分推定量は

$$\hat{\boldsymbol{\beta}} = \{(\mathbf{WX})' \mathbf{WX}\}^{-1} (\mathbf{WX})' \mathbf{WY} \quad (9)$$

(8)式の最小二乗推定量として定義される。

### 3 差分推定量の漸近的性質

この小節では、差分の次数 $q$ が有限の値に固定された場合と $q$ が $n$ に伴って無限大になっていく場合の漸近理論を論じる。Yatchew(1997)は $q$ が固定された場合の漸近理論だけ与えた。本論文は $q$ が $n$ に伴って無限大になっていく場合の漸近理論に重点を置き、差分の次数 $q$ に関連する推定量 $\hat{\beta}$ の漸近効率性に関する理論を与える。一致性の証明に関しては漸近正規性の証明の過程の中でほぼ自明であるため省略する。

#### 3.1 $q$ が有限な値に固定された場合

まず、 $q$ が固定された場合の漸近理論のための正則条件を与える。

1.  $U_i$ がi.i.d.で、値域が有界である。その密度関数は至る所で0より大きい値をとる。
2.  $\varepsilon_i$ がi.i.d.で平均がゼロ、分散が $\sigma_\varepsilon^2$ で $U$ や $X$ と独立である。
3.  $\Gamma(U_i) \equiv E(X_i|U_i)$ ,  $V_i \equiv X_i - \Gamma(U_i)$ と定義する。ただし、 $\Gamma(U_i)$ は $U_i$ により生成された $\sigma$ -fieldに関して可測である。 $i \neq j$ に関して $E(V_i V_j') = \Omega_V$ 、それ以外は $E(V_i V_j') = 0$ 。 $V_i$ がi.i.d.で $U$ や $\varepsilon$ と独立である。
4.  $\|\alpha(U_i) - \alpha(U_j)\| \leq k_\alpha \|U_i - U_j\|$ と $\|\Gamma(U_i) - \Gamma(U_j)\| \leq k_\Gamma \|U_i - U_j\|$ を満たすような $\infty > k_\alpha > 0$ ,  $\infty > k_\Gamma > 0$ が存在する。
5.  $\sum_{j=1}^{q+1} \omega_j = 0$ ,  $\sum_{j=1}^{q+1} \omega_j^2 = 1$ 。

$q$ が有限の場合、以上の条件のもとで、Yatchew(1997)の結果と殆ど同様な次の定理が成り立つ。表記を簡単にするため、 $WY$ を $Y^*$ 、 $W\alpha(U)$ を $\zeta^*$ と記する、他の変数に関しても差分を取った変数を(\*)付の変数と記する。

**定理 1** 条件1 – 5が満たされるとする。 $q$ が有限で、 $q < n$ ,  $r \leq 3$ とする。差分推定量は以下の漸近的性質を持つ。

$$\sqrt{n-q} (\hat{\beta} - \beta) \xrightarrow{d} N(0, (1+\lambda) \sigma_\varepsilon^2 \Omega_V^{-1}) \quad (10)$$

ただし、

$$\lambda = 2 \sum_{k=1}^q \left( \sum_{j=k+1}^{q+1} \omega_j \omega_k \right)^2 \quad (11)$$

である。さらに， $\sigma_\varepsilon^2$ と $\Omega_V$ の一致推定量はそれぞれ

$$\hat{\sigma}_\varepsilon^2 = \frac{1}{n - q - l} \sum_{i=1}^n \left( Y_i^* - \mathbf{X}_i^{*'} \hat{\beta} \right)^2 \quad (12)$$

$$\hat{\Omega}_V = \frac{1}{n - q} \mathbf{X}^{*'} \mathbf{X}^* \quad (13)$$

となる。

証明. 補章を参照せよ。□

標準的な統計学の理論に従えば，帰無仮説が $H_0 : \mathbf{R}\beta = 0$ で対立仮説が $H_1 : \mathbf{R}\beta \neq 0$ の $F$ 検定を行うことができる。 $\text{Rank}(\mathbf{R}) = c$ のとき，検定統計量

$$F = \frac{\hat{\beta}' \mathbf{R}' (\mathbf{R} \hat{\Omega}_V^{-1} \mathbf{R}')^{-1} \mathbf{R} \hat{\beta}}{c(1 + \lambda) \hat{\sigma}_\varepsilon^2} \quad (14)$$

は自由度が $(c, n - q - l)$ の $F$ 分布に従う。

Chamberlain(1992)に従って計算すれば，モデル(1)の $\beta$ の推定量のsemiparametric efficiency boundが $\sigma_\varepsilon^2 \Omega_V^{-1}$ であることが分かる。 $\lambda$ が非負であるため，定理1の結果から $q$ が有限な値に固定されると推定量 $\hat{\beta}$ が効率性を達成できないことが分かる。

### 3.2 $q$ が $n$ とともに無限大になっていく場合

これから， $q$ が $n$ とともに一定のオーダーで無限大になっていくとき，効率性が達成されることを示す。そのためには，一定の条件を満たす特殊な重みを使う必要がある。そのような重みは必ずしも一意に決まらない。ここでは，一つの例として，Hall et al. (1990)によって提案されたspike weightに少し調整を加えたものを利用する。混乱が生じないので，表記を変えずに $W$ でspike weightを表す。その要素 $\omega_j$ の定義は

$$\omega_j = \begin{cases} [q/(q+1)]^{1/2} & O(1) & j = 1 \\ -[q/(q+1)]^{-1/2} & O(1/q) & j = 2, 3, \dots, q+1 \\ 0 & & \text{else} \end{cases} \quad (15)$$

である。制約条件 $\sum_{j=1}^{q+1} \omega_j = 0$ と $\sum_{j=1}^{q+1} \omega_j^2 = 1$ が満たされていることは容易に確認できる。Spike weightを利用することで $k = 1, \dots, q$ に対して $\sum_{j=k+1}^{q+1} \omega_j \omega_k = O(1/q)$ となり， $\lambda = O(1/q)$ となる。一定のオーダーで $q \rightarrow \infty$ となっていくことにより，推定量 $\hat{\beta}$ の効率性が達成される。ここまでの結果を以下の定理にまとめる。

定理 2 正則条件1 – 5のもとで,  $q \ll n$ ,  $r \leq 3$ で,  $k = 1, \dots, q$ に対して  $\sum_{j=k+1}^{q+1} \omega_j \omega_k = O(1/q)$ すなわち  $\text{Tr}(\mathbf{W}'\mathbf{W}\mathbf{W}'\mathbf{W}) - 1 = O(1/q)$ を満たす重み行列 $\mathbf{W}$ が存在する。さらに,  $q = O(n^\eta)$ ,  $1/r - 1/4 > \eta > 0$ ,  $n$ に伴って無限大になっていくとき

$$\sqrt{n-q}(\hat{\beta} - \beta) \xrightarrow{d} N(\mathbf{0}, \sigma_\varepsilon^2 \Omega_{\mathbf{V}}^{-1}) \quad (16)$$

となる。 $\sigma_\varepsilon^2$ と $\Omega_{\mathbf{V}}$ の一致推定量はそれぞれ

$$\hat{\sigma}_\varepsilon^2 = \frac{1}{n-q-l} \sum_{i=1}^{n-q} (Y_i^* - \mathbf{X}_i^{*'} \hat{\beta})^2 \quad (17)$$

$$\hat{\Omega}_{\mathbf{V}} = \frac{1}{n-q} \sum_{i=1}^{n-q} \mathbf{X}_i^{*'} \mathbf{X}_i^* \quad (18)$$

で与える。

証明. 補章を参照せよ。□

以上の定理から $\hat{\beta}$ がsemiparametric efficiency bound,  $\sigma_\varepsilon^2 \Omega_{\mathbf{V}}^{-1}$ を達成していることが分かる。また,  $q$ が固定されたときと同様に係数 $\beta$ に関する $F$ 検定を行うことができる。

### 3.3 ノンパラメトリックの部分の推定

一旦 $\beta$ の一致推定量を得られたら, その推定量 $\hat{\beta}$ をモデル(1)に代入してから, 近似的に単純なノンパラメトリックモデル

$$\tilde{\mathbf{Y}} \approx \alpha(\mathbf{U}) + \varepsilon \quad (19)$$

を得ることができる。ただし,  $\tilde{\mathbf{Y}} = \mathbf{Y} - \mathbf{X}\hat{\beta}$ である。そして, モデル(19)に既存のノンパラメトリックな推定法を適用すれば良い。前述した定理によって, 差分推定量 $\hat{\beta}$ の $\beta$ に収束するオーダーが典型的なノンパラメトリック推定量の収束オーダーより速いため,  $\alpha(\cdot)$ の推定量の収束オーダーが利用したノンパラメトリックの推定法に対応する収束オーダーになる。例えば, local constant regression (Nadaraya-Watson estimator)を利用することで $\alpha(\cdot)$ を推定できる。

## 4 モンテカルロ実験

有限標本の性質を確かめるため，モンテカルロ実験を行う。DGPとして， $\mathbf{U} = (U_1, U_2)' \in \mathbb{R}^2$ ， $\mathbf{X} = (X_1, X_2)' \in \mathbb{R}^2$ ， $U_1, U_2$ は独立なUniform(0, 1)分布に従い， $\varepsilon \sim N(0, 1)$ ， $V_1$ と $V_2$ が互いに独立な標準正規分布に従うとする。また， $\mathbf{U}$ ， $\varepsilon$ ， $V_1$ ， $V_2$ はすべて互いに独立であるとする。モデルは

$$Y = \alpha(\mathbf{U}) + b_1 X_1 + b_2 X_2 + \varepsilon$$

$$X_1 = \sin(U_1 + U_2) + V_1 \quad (20)$$

$$X_2 = \cos(U_1 - U_2) + V_2$$

$$\alpha(\mathbf{U}) = 5 \sin(2\pi(U_1 + U_2)) \quad (21)$$

とする。パラメーターの真の値は $b_1 = 0.3$ ， $b_2 = 0.6$ とする。サンプル数は300, 500と1000の三種類で， $q$ も1, 2と4の三種類にする。以上の設定で実験を繰り返し1000回行う。

$b_1$ と $b_2$ の推定結果は表1と表2にまとめた。差分推定がうまく機能していることが分かる。さらに， $b_1$ の場合サンプル数が300のとき， $q = 1$ のケースのMSEが最も小さい，そして，サンプル数が500のとき， $q = 2$ のケース，1000になると， $q = 4$ のケースのMSEが最も小さい。サンプル数に従って差分の次数を大きくすることで，推定量の効率性が高まっていくことが分かる。 $b_2$ の場合も同じような傾向が見られる。

表1.  $b_1$ の推定結果

$n$	300			500			1000		
$q$	Bias	SD	MSE	Bias	SD	MSE	Bias	SD	MSE
1	0.0018	0.0934	0.0087	0.0007	0.0674	0.0045	0.0012	0.0440	0.0019
2	0.0041	0.0995	0.0099	0.0018	0.0655	0.0043	0.0013	0.0432	0.0019
4	0.0071	0.1359	0.0185	0.0054	0.0738	0.0055	0.0015	0.0423	0.0018

表2.  $b_2$ の推定結果

$n$	300			500			1000		
$q$	Bias	SD	MSE	Bias	SD	MSE	Bias	SD	MSE
1	-0.0007	0.0941	0.0089	0.0030	0.0686	0.0047	0.0004	0.0422	0.0018
2	-0.0004	0.0989	0.0098	0.0035	0.0668	0.0045	-0.0012	0.0417	0.0017
4	-0.0006	0.1373	0.0188	0.0058	0.0747	0.0056	-0.0015	0.0425	0.0018

ノンパラメトリックの部分 $\alpha(\cdot)$ に関しては、 $b_1$ と $b_2$ の推定量をモデル式に代入して、そして、Nadaraya-Watson法で推定を行った。図1に推定結果の一つの例を表示した。

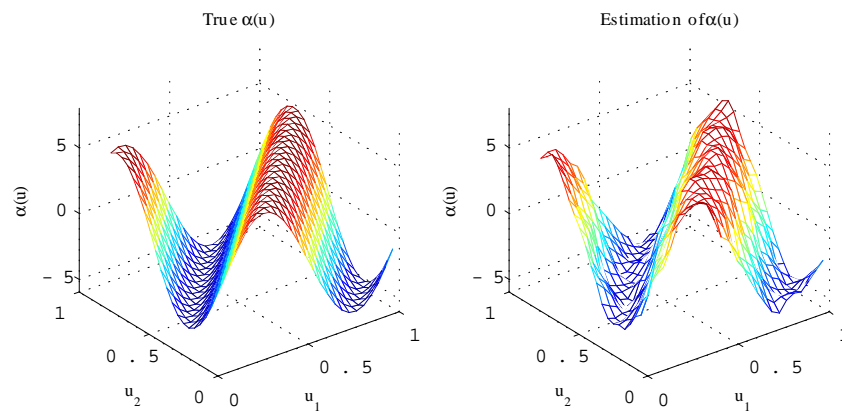


図1 .  $a(\cdot)$ の推定結果

## 5 結論

本論文はYatchew(1997)の結果を拡張して、差分法による部分線形モデルの推定法の漸近理論に関して論じた。特に、差分の次数の増加するオーダーに関する制約を明らかにした。そして、効率性を達成するために、重み行列が満たすべき条件を定理の中で与えた。差分法に関する漸近理論を既存の研究結果よりさらに厳密なものにした。また、論文の中でモンテカルロ実験を行い、推定量の有限標本の性質を確認した。査読者に指摘されたように、データが得られたときにそのデータから差分の最適な次数 $q$ を得るための方法が実用的には関心の対象となる。一つの方法としてMSEの推定量を最小にする $q$ を最適な $q$ とすることが考えられる。しかし、そのために推定量の有限標本の性質を導出し、それを利用してMSEを推定することが必要となる。推定量の有限標本の性質がまだ解明されていない。今後の研究課題にしたい。

## 6 補章

Grid Nearest Neighbor Algorithm



$U_i$ の次元 $r = 1$ のとき,  $u_1 \leq u_2 \leq \dots \leq u_n$ となるように $u_i$ を並べ替えるだけで良い。 $r = 2, 3$ のとき, Yatchew (1997)のGrid Nearest Neighbor Algorithmを用いてデータをソートする。Yatchew (1997)はレンマの中でソートしたデータに関しては, きわめて0に近い $0 < \delta < 1$ があって,

$$\max_i \|U_i - U_{i+1}\|^2 = O_p(n^{-2(1-\delta)/r}) \cong O_p(n^{-2/r}) \quad (22)$$

となる、さらに

$$\frac{1}{n} \sum_{i=1}^n \|U_i - U_{i-1}\|^2 = O_p(n^{-2(1-\delta)/r}) \cong O_p(n^{-2/r}) \quad (23)$$

となることを示した。

以下から定理1と2を証明する。証明の方針はYatchew (1997)の定理の証明をベースにしている。

定理1の証明. 定理1の証明は定理2の証明から殆ど自明であるため証明を省略する。□

定理2の証明. 表記を簡潔にするため, ここから $\tilde{n}$ で $n - q$ を表す。Spike weighを利用すれば

$$\sum_{j=k+1}^{q+1} \omega_j \omega_k = O(1/q) \quad (24)$$

すなわち

$$E(\text{Tr}(\mathbf{W}'\mathbf{W}\mathbf{W}'\mathbf{W})) - 1 = O_p(1/q) \quad (25)$$

が満たされる。さらにYatchew(1997)のレンマにより, grid nearest neighbor algorithmを利用してデータをソートしたので,  $q = O_p(n^\eta)$ であれば $\max_{1 \leq j \leq q} \|U_i - U_{i+j}\|^2 = O_p(n^{2\eta-2(1-\delta)/r}) \cong O_p(n^{2\eta-2/r})$ となる。ただし,  $1 > \delta > 0$ で任意に0に近い小さい定数である。そして,  $\infty > \kappa > 0$ に対して

$\|f(\mathbf{U}_i) - f(\mathbf{U}_j)\| \leq \kappa \|\mathbf{U}_i - \mathbf{U}_j\|$ を満たす可測関数 $f(\cdot)$ は

$$\begin{aligned}
& \left( \sum_{j=1}^{q+1} \omega_j f(U_{i+j-1}) \right)^2 \\
&= (\omega_2 (f(U_i) - f(U_{i+1})) + \cdots + \omega_{q+1} (f(U_i) - f(U_{i+q})))^2 \\
&\leq \sum_{j=1}^q \sum_{j'=1}^q |\omega_{j+1} \omega_{j'+1}| \|f(U_i) - f(U_{i+j})\| \|f(U_i) - f(U_{i+j'})\| \\
&\leq \max_{1 \leq h \leq q} \|f(U_i) - f(U_{i+h})\|^2 \sum_{j=1}^q \sum_{j'=1}^q |\omega_{j+1} \omega_{j'+1}| \\
&\cong O_p(n^{2\eta-2/r}) O_p(1) = O_p(n^{2\eta-2/r})
\end{aligned} \tag{26}$$

を満たすことが分かる。上式の導出の中で, spike weightの性質から導出できる一つの結果 $\sum_{j=1}^q \sum_{j'=1}^q |\omega_{j+1} \omega_{j'+1}| = O_p(1)$ であることを利用した。

これから

$$\begin{aligned}
& \sqrt{\tilde{n}} (\hat{\beta} - \beta) \\
&= (\tilde{n}^{-1} \mathbf{X}^* \mathbf{X}^*)^{-1} (\tilde{n}^{1/2} \tilde{n}^{-1} \mathbf{X}' (\boldsymbol{\varepsilon}^* + \boldsymbol{\zeta}^*)) \\
&= \underbrace{\left( \tilde{n}^{-1} \sum_{i=1}^{\tilde{n}} \mathbf{X}_i^* \mathbf{X}_i^* \right)^{-1}}_{\Delta} \underbrace{\tilde{n}^{1/2} \tilde{n}^{-1} \sum_{i=1}^{\tilde{n}} \mathbf{X}_i^* (\varepsilon_i^* + \zeta_i^*)}_{\nabla}.
\end{aligned} \tag{27}$$

に関して考えていく。 $\Delta$ の $(m, s)$ 要素,  $\Delta_{ms}$ は

$$\begin{aligned}
\Delta_{ms} &= \tilde{n}^{-1} \sum_{i=1}^{\tilde{n}} X_{im}^* X_{is}^* \\
&= \tilde{n}^{-1} \sum_{i=1}^{\tilde{n}} \left\{ \left( \sum_{j=1}^{q+1} \omega_j \Gamma(u_{i+j-1})_m + \sum_{j=1}^{q+1} \omega_j v_{i+j-1, m} \right) \right. \\
&\quad \left. \times \left( \sum_{j=1}^{q+1} \omega_j \Gamma(u_{i+j-1})_s + \sum_{j=1}^{q+1} \omega_j v_{i+j-1, s} \right) \right\}
\end{aligned} \tag{28}$$

と展開できる。さらに，コーシー・シュワルツ不等式を利用すれば

$$\begin{aligned}
& \tilde{n}^{-1} \left| \sum_{i=1}^{\tilde{n}} \left( \sum_{j=1}^{q+1} \omega_j \Gamma(u_{i+j-1})_m \sum_{j=1}^{q+1} \omega_j \Gamma(u_{i+j-1})_s \right) \right| \\
& \leq \tilde{n}^{-1} \left[ \sum_{i=1}^{\tilde{n}} \left( \sum_{j=1}^{q+1} \omega_j \Gamma(u_{i+j-1})_m \right)^2 \right]^{1/2} \left[ \sum_{i=1}^{\tilde{n}} \left( \sum_{j=1}^{q+1} \omega_j \Gamma(u_{i+j-1})_s \right)^2 \right]^{1/2} \\
& \cong \tilde{n}^{-1} O_p(n^{1/2+\eta-1/r}) O_p(n^{1/2+\eta-1/r}) \\
& = O_p(n^{2\eta-2/r}) = o_p(1)
\end{aligned} \tag{29}$$

そして

$$\begin{aligned}
& \tilde{n}^{-1} \left| \sum_{i=1}^{\tilde{n}} \left( \sum_{j=1}^{q+1} \omega_j \Gamma(u_{i+j-1})_m \sum_{j=1}^{q+1} \omega_j v_{i+j-1,s} \right) \right| \\
& \leq \tilde{n}^{-1} \left[ \sum_{i=1}^{\tilde{n}} \left( \sum_{j=1}^{q+1} \omega_j \Gamma(u_{i+j-1})_m \right)^2 \right]^{1/2} \left[ \sum_{i=1}^{\tilde{n}} \left( \sum_{j=1}^{q+1} \omega_j v_{i+j-1,s} \right)^2 \right]^{1/2} \\
& \cong \tilde{n}^{-1/2} O_p(n^{1/2+\eta-1/r}) \left\{ 1/\tilde{n} \sum_{i=1}^{\tilde{n}} \left( \sum_{j=1}^{q+1} \omega_j v_{i+j-1,s} \right)^2 \right\}^{1/2} \\
& = O_p(n^{\eta-1/r}) = o_p(1)
\end{aligned} \tag{30}$$

が分かる。ただし，

$$1/\tilde{n} \sum_{i=1}^{\tilde{n}} \left( \sum_{j=1}^{q+1} \omega_j V_{i+j-1} \right)^2 \rightarrow \Omega_{V,s,s}, \tag{31}$$

である。従って， $n \rightarrow \infty$ のとき

$$\begin{aligned}
\Delta_{ms} & \cong \tilde{n}^{-1} \sum_{i=1}^{\tilde{n}} \left( \sum_{j=1}^{q+1} \omega_j V_{i+j-1} \sum_{j=1}^{q+1} \omega_j V_{i+j-1} \right) \\
& = \Omega_{V,m,s}
\end{aligned} \tag{32}$$

となり，すなわち

$$\Delta \cong \Omega_V \tag{33}$$

となる。

以下から、添え字( $i \cdot$ )で行列の第 $i$ 行を表す。続いて $\nabla = \tilde{n}^{1/2} \tilde{n}^{-1} \sum_{i=1}^{\tilde{n}} (\Gamma_{i \cdot}^{*'} + V_{i \cdot}^{*'}) (\varepsilon_i^* + \zeta_i^*)$ の収束先を導出する。

$$\tilde{n}^{-1/2} \sum_{i=1}^{\tilde{n}} \Gamma_{i \cdot}^{*'} \zeta_i^* \quad (34)$$

の第 $m$ 要素

$$\begin{aligned} & \tilde{n}^{1/2} \tilde{n}^{-1} \left| \sum_{i=1}^{\tilde{n}} \left[ \sum_{j=1}^{q+1} \Gamma(U_{i+j-1})_m \sum_{j=1}^{q+1} \alpha(U_{i+j-1}) \right] \right| \\ & = O_p(n^{1/2}) O_p(n^{2\eta-2/r}) = O_p(n^{1/2+2\eta-2/r}) = o_p(1) \end{aligned} \quad (35)$$

0に確率収束する。さらに、

$$\begin{aligned} & E \left\{ \tilde{n}^{-1} \sum_{i=1}^{\tilde{n}} \Gamma_{im}^* \varepsilon_i^* \sum_{i=1}^{\tilde{n}} \Gamma_{is}^* \varepsilon_i^* \right\} \quad (36) \\ & = \tilde{n}^{-1} E \left\{ \sum_{i=1}^{\tilde{n}} \left[ \sum_{j=1}^{q+1} \omega_j \Gamma(u_{i+j-1})_m \sum_{j=1}^{q+1} \omega_j \varepsilon_{i+j-1} \right] \right. \\ & \quad \left. \sum_{i=1}^n \left[ \sum_{j=1}^{q+1} \omega_j \Gamma(u_{i+j-1})_s \sum_{j=1}^{q+1} \omega_j \varepsilon_{i+j-1} \right] \right\} \\ & = \tilde{n}^{-1} E \sum_{i=1}^{\tilde{n}} \left[ \sum_{j=1}^{q+1} \omega_j \Gamma(u_{i+j-1})_m \sum_{j=1}^{q+1} \omega_j \Gamma(u_{i+j-1})_s \sum_{j=1}^{q+1} \omega_j^2 \varepsilon_{i+j-1}^2 \right] \\ & + \tilde{n}^{-1} E \sum_{i=1}^{\tilde{n}} \sum_{i' \neq i} \left[ \sum_{j=1}^{q+1} \omega_j \Gamma(u_{i+j-1})_m \sum_{j=1}^{q+1} \omega_j \Gamma(u_{i'+j-1})_s \right. \\ & \quad \left. \times \sum_{j=1}^{q+1} \omega_j \varepsilon_{i+j-1} \sum_{j=1}^{q+1} \omega_j \varepsilon_{i'+j-1} \right] \\ & = \tilde{n}^{-1} E \sum_{i=1}^{\tilde{n}} \left[ \sum_{j=1}^{q+1} \omega_j \Gamma(u_{i+j-1})_m \sum_{j=1}^{q+1} \omega_j \Gamma(u_{i+j-1})_s \right] \sigma_\varepsilon^2 \\ & + \tilde{n}^{-1} E \sum_{i=1}^{\tilde{n}} 2 \left\{ \sum_{k=1}^q \left( \sum_{j=k+1}^{q+1} \omega_j \Gamma(u_{i+j-1})_m \sum_{j=k+1}^{q+1} \omega_j \Gamma(u_{i+j-1})_s \sum_{j=k+1}^{q+1} \omega_j \omega_k \right) \right\} \sigma_\varepsilon^2 \end{aligned}$$

となり，その最初の項は

$$\begin{aligned}
& \tilde{n}^{-1} E \sum_{i=1}^{\tilde{n}} \left| \sum_{j=1}^{q+1} \omega_j \Gamma(u_{i+j-1})_m \sum_{j=1}^{q+1} \omega_j \Gamma(u_{i+j-1})_s \right| \sigma_\varepsilon^2 \\
& \leq \tilde{n}^{-1} E [O(n) O_p(n^{\eta-2/r})] \sigma_\varepsilon^2 \\
& = O_p(n^{2\eta-2/r})
\end{aligned} \tag{37}$$

であるため，0に確率収束する。二番目の項も

$$\begin{aligned}
& \tilde{n}^{-1} E \sum_{i=1}^{\tilde{n}} 2 \left| \sum_{k=1}^q \left( \sum_{j=k+1}^{q+1} \omega_j \Gamma(u_{i+j-1})_m \sum_{j=k+1}^{q+1} \omega_j \Gamma(u_{i+j-1})_s \sum_{j=k+1}^{q+1} \omega_j \omega_k \right) \right| \sigma_\varepsilon^2 \\
& = \sigma_\varepsilon^2 O(n^{-1}) O(n) E \{ O_p(n^{\eta-1/r}) O_p(n^{\eta-1/r}) o_p(1) \} \\
& = o_p(n^{2\eta-2/r}).
\end{aligned} \tag{38}$$

であるため，0に確率収束する。従って，

$$E \left\{ \tilde{n}^{-1} \sum_{i=1}^{\tilde{n}} \Gamma_{i \cdot}^* \varepsilon_i^* \sum_{i=1}^{\tilde{n}} \Gamma_{i \cdot}^* \varepsilon_i^* \right\} \rightarrow \mathbf{0} \tag{39}$$

$$E \left\{ \tilde{n}^{-1} \sum_{i=1}^{\tilde{n}} V_{i \cdot}^* \zeta_i^* \sum_{i=1}^{\tilde{n}} V_{i \cdot}^* \zeta_i^* \right\} \rightarrow \mathbf{0} \tag{40}$$

となることが分かる。さらに，

$$E \left( \tilde{n}^{-1/2} \sum_{i=1}^{\tilde{n}} \Gamma_{i \cdot}^* \varepsilon_i^* \right) = E \left( \tilde{n}^{-1/2} \sum_{i=1}^{\tilde{n}} V_{i \cdot}^* \zeta_i^* \right) = \mathbf{0} \tag{41}$$

より，マルコフの不等式を利用すれば

$$\tilde{n}^{-1/2} \sum_{i=1}^{\tilde{n}} \Gamma_{i \cdot}^* \varepsilon_i^* = o_p(1) \tag{42}$$

$$\tilde{n}^{-1/2} \sum_{i=1}^{\tilde{n}} V_{i \cdot}^* \zeta_i^* = o_p(1) \tag{43}$$

が証明される。

$\tilde{n}^{-1/2}\mathbf{V}'\boldsymbol{\varepsilon}^*$ の分散共分散行列を計算する。定理の中の条件を利用すれば

$$\begin{aligned}
& E(\tilde{n}^{-1}\mathbf{V}'\mathbf{W}'\mathbf{W}\boldsymbol{\varepsilon}\boldsymbol{\varepsilon}'\mathbf{W}'\mathbf{W}\mathbf{V}) \\
&= \tilde{n}^{-1}E[\text{Tr}(\mathbf{W}'\mathbf{W}\mathbf{W}'\mathbf{W})]\sigma_\varepsilon^2\boldsymbol{\Omega}_V \\
&= \left\{ 1 + 2 \sum_{k=1}^q \left( \sum_{j=k+1}^{q+1} \omega_j \omega_k \right)^2 \right\} \sigma_\varepsilon^2 \boldsymbol{\Omega}_V \\
&= (1 + o_p(1)) \sigma_\varepsilon^2 \boldsymbol{\Omega}_V
\end{aligned} \tag{44}$$

が得られる。そして、補題3、中心極限定理(Scott 1973 Theorem 3)とCramer-Word Theoremを利用すれば、

$$\frac{1}{\sqrt{\tilde{n}}}\mathbf{V}'\boldsymbol{\varepsilon}^* \xrightarrow{d} N(\mathbf{0}, \sigma_\varepsilon^2 \boldsymbol{\Omega}_V) \tag{45}$$

となることが分かる。さらに、(35), (42), (43)と(45)式により

$$\nabla = \frac{1}{\sqrt{\tilde{n}}}\mathbf{X}'(\boldsymbol{\varepsilon}^* + \boldsymbol{\zeta}^*) \cong \frac{1}{\sqrt{\tilde{n}}}\mathbf{V}'\boldsymbol{\varepsilon}^* \xrightarrow{d} N(\mathbf{0}, \sigma_\varepsilon^2 \boldsymbol{\Omega}_V) \tag{46}$$

となる。最後にCramér-word theoremと(33), (46)式より

$$\sqrt{\tilde{n}}(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta}) \xrightarrow{d} N(\mathbf{0}, \sigma_\varepsilon^2 \boldsymbol{\Omega}_V^{-1}) \tag{47}$$

が証明される。他の推定量の一致性はLLNに従う。□

次の補題を証明するため、便利上データを逆順に並べ替える。 $\lambda$ を $\lambda'\lambda = 1$ を満たす任意の $l \times 1$ 実数値ベクトルとする。新しい表記として $\xi_i = \lambda'V_i^*\varepsilon_i^*$ とする。さらに、 $M_i$ を $\{\xi_j\}_{j=-\infty}^i$ により生成された $\sigma$ -fieldとし、 $M_{-\infty} = \bigcap_{k=-\infty}^i M_k$ ,  $M_\infty = \bigcup_{k=-\infty}^i M_k$ とする。 $V_i^*$ の第 $s$ 要素を $V_{is}^*$ ,  $s = 1, 2, \dots, l$ , と表す。

補題 3 定理2の条件のもとで、

$$\sum_{m=0}^{\infty} (\|E\{\xi_i | M_{i-m}\}\| + \|\xi_i - E\{\xi_i | M_{i+m}\}\|) < \infty \tag{48}$$

が満たされる。ただし、ここに限ってノルムの意味合いとしては $\|X\|$ が $(EX^2)^{1/2}$ を意味する。

証明.  $m = 0$ のとき、

$$\|E\{V_{is}^*\varepsilon_i^* | M_i\}\| = \left\{ E \left( \sum_{j=1}^{q+1} \omega_j V_{i-j+1,s} \sum_{j=1}^{q+1} \omega_j \varepsilon_{i-j+1} \right)^2 \right\}^{1/2} < \infty \tag{49}$$

となることが自明である。  $m = 1$  の場合 ,

$$\begin{aligned}
& \|E \{V_{is}^* \varepsilon_i^* | M_{i-1}\}\| \tag{50} \\
&= \left\{ E \left[ \sum_{j=2}^{q+1} \omega_j V_{i-j+1,s} \sum_{j=2}^{q+1} \omega_j \varepsilon_{i-j+1} \right. \right. \\
&\quad \left. \left. + E \left( \sum_{j=2}^{q+1} \omega_j V_{i-j+1,s} \omega_i \varepsilon_i + \omega_i V_{i,s} \sum_{j=2}^{q+1} \omega_j \varepsilon_{i-j+1} + \omega_i V_{i,s} \omega_i \varepsilon_i \middle| M_{i-1} \right) \right]^2 \right\}^{1/2} \\
&= \left\{ E \left( \sum_{j=2}^{q+1} \omega_j V_{i-j+1,s} \sum_{j=2}^{q+1} \omega_j \varepsilon_{i-j+1} \right)^2 \right\}^{1/2} \\
&= \left\{ E \left( \sum_{j=2}^{q+1} \omega_j^2 \right)^2 \right\}^{1/2} \sigma_\varepsilon \sigma_{v_s}
\end{aligned}$$

となる。ただし,  $\sigma_{v_s}$  は  $\Omega_V$  の  $s$  番目の対角要素の平方根を表す。  $j \neq 1$  のとき  $\omega_j = O_p(1/q)$  であるため,  $\|E \{V_{is}^* \varepsilon_i^* | M_{i-1}\}\| = O(1/q)$  が分かる。さらに,  $m = 2, \dots, q$  に関しても同様な結果が得られる。一方では,  $m > q$  のとき  $\|E \{V_{is}^* \varepsilon_i^* | M_{i-m}\}\| = 0$  となる。従って  $\sum_{m=0}^{\infty} \|E \{V_{is}^* \varepsilon_i^* | M_{i-m}\}\| < \infty$ 。さらに,  $m = 0, 1, \dots, \infty$  に関して  $\|V_{is}^* \varepsilon_i^* - E \{V_{is}^* \varepsilon_i^* | M_{i+m}\}\| = 0$  であるため,

$$\sum_{m=0}^{\infty} (\|E \{V_{is}^* \varepsilon_i^* | M_{i-m}\}\| + \|V_{is}^* \varepsilon_i^* - E \{V_{is}^* \varepsilon_i^* | M_{i+m}\}\|) < \infty \tag{51}$$

となる。最後に  $\lambda' \lambda = 1$  は  $\lambda$  の要素  $\lambda_s$  が有限であることを意味するので,  $\xi_i = \sum_{s=1}^l \lambda_s V_{is}^* \varepsilon_i^*$  であることから, 補題が証明される。  $\square$

## 参考文献

- [1] Ahmad, I., Leelahanon, S. and Li, Q. (2005), Efficient Estimation of a Semiparametric Partially Linear Varying Coefficient Model, *Annals of Statistics*, 33, 258-283.
- [2] Ahn, H., Powell, J. (1993), Semiparametric estimation of censored selection models with a nonparametric selection mechanism, *Journal of Econometrics*, 58, 2-29.

- [3] Chamberlain, G. (1992), Efficiency Bounds for Semiparametric Regression, *Econometrica*, 60, (3), 567-596.
- [4] Fan, J. and Wu, Y. (2008). Semiparametric estimation of covariance matrices for longitudinal data, *Journal of American Statistical Association*, to appear.
- [5] Hall, P., Kay, J. and Titterton, D. M. (1990), Asymptotically optimal difference-based estimation of variance in nonparametric regression. *Biometrika*, 77, 521-528.
- [6] Liu, Q., Wu, Y. and Fan, J. (2009), Difference-based Estimation for Semiparametric Varying-Coefficient Partially Linear Models, Working paper.
- [7] Powell, J. (1987), Semiparametric Estimation of Bivariate Latent Variable Models, *Working paper 8704*, Social Systems Research Institute of Wisconsin, University of Wisconsin, Madison.
- [8] Robinson, P.M. (1988), Root-N consistent semiparametric regression, *Econometrica*, 56, 931-954.
- [9] Scott, D.J. (1973), Central limit theorems for martingales and for processes with stationary increments using a skorokhod representation approach, *Advances in Applied Probability*, 5, (1), 119-137.
- [10] Yatchew, A. (1997), An elementary estimator for the partially linear model, *Economics Letters*, 57, 135-143.