응용통계연구 제 13권 2호, 2000년, pp. 371-381

# 편스플라인 추정량의 편의에 대한 점근 정규성

추인선[1] 최재룡[2]

요 약

비모수 회귀모형에 있어서 평활스플라인에 대하여 언급하고, 그 간단한 성질을 다룬다. 선형회귀나 다항식회귀에서는 적합하기 나쁜 데이터가 많이 존재한다. 설명변수가 여러 개인 경우에 준모수 회귀모형은 하나 혹은 그 이상의 변수에 대해서는 비모수 함수를 다른 변수에 대해서는 선형함수를 적합시켜 그들의 가법성을 가정한 것이다. 준모수 회귀모형에 있어서 선형부분의 회귀계수의 추정량에 편의가 발생하고, 여기서는 그 편의에 대한 점근 정규성을 다룬다.

주요용어: 평활화, 준모수회귀모형, 스플라인, 평활행렬.

## 1. 서 론

비모수 회귀는 데이터의 흩어진 정도에 근거하여 설명변수와 응답변수의 관계를 나타내는 함수, 즉 회귀곡선의 형태를 찾는 통계적 방법이다. 이것은 통상의 모수적 회귀에서는 적합하기 나쁜 데이터를 다룰 경우나 얻어진 데이터에 대하여 적합할 모형을 선택하기 위해 설명변수와 응답변수의 관계를 탐색적인 방법으로 요약할 경우에도 사용된다.

비모수 회귀에 있어서 곡선추정의 수순을 평활화(smoothing)라 한다. 평활화의 방법으로 핵함수를 이용한 평활화(kernel smoothing)와 평활스플라인(smoothing spline)이 잘 알려져 있다.

비모수 회귀를 중회귀로 발전시키는 것은 자연스러운 흐름이지만 완전한 일반화를 행하면 벌점(penalty)의 정식화가 어렵게 된다. 따라서 하나의 설명변수 $t$만을 비모수적으로 다룬 준모수(semiparametric) 회귀모형

$$y_i = \boldsymbol{x}_i^T \boldsymbol{\beta} + g(t_i) + \text{오차} \tag{1.1}$$

을 생각한다. 추정해야할 양은 $\boldsymbol{\beta}$ 및 $g(t)$이다.

실제로 준모수 회귀모형 (1.1)을 사용한 예는 Green *et al*(1985), Engle *et al*(1986) 등에서 찾아볼 수 있고, Rice(1986)는 모형 (1.1)에 있어서 $\boldsymbol{x}_i$와 $t_i$ 사이에 어떤 관계가 존재할 때, $\boldsymbol{\beta}$의 추정량이 일반적으로 편의(bias)를 갖는다는 것을 점근론적으로 나타내었다. 이 논문의 2절에서는 비모수 회귀모형에 있어서 평활스플라인을 언급하고, 3절과 4절에서는 편스플라인과 편스플라인 추정량의 점근적 성질을 다룬다.

1) (604-714) 부산광역시 사하구 하단동 840, 동아대학교 자연과학부 수학과, 강사
E-mail: i-s-chu@hanmail.net
2) (604-714) 부산광역시 사하구 하단동 840, 동아대학교 자연과학부 수학과, 교수
E-mail: jrchoi@mail.donga.ac.kr

## 2. 비모수 회귀모형에 있어서 평활스플라인

평활스플라인은 비모수 회귀모형에서 가장 많이 연구되고 있는 것 중에 하나이고 그 계산법, 소프트웨어도 잘 정비되어 있다. 단, 스플라인이라면 일반적으로 3차 스플라인을 가리키는 경우가 많고, 예를 들어, Green and Silverman(1994)도 3차 평활스플라인만을 논의하고 있다. 스플라인 함수는 다항식 모형이 갖는 결점을 보완하기 위해 개발되었다. 여기서는 다항식 모형 대신에 스플라인 함수를 회귀적으로 사용하는 회귀모형을 고려한다. 일반적으로 3차자연스플라인(natural cubic spline, NCS)을 사용하는 경우가 많다. 2차 평활스플라인에 관해서는 김(1998)을 참고한다.

비모수 회귀, 즉, 데이터 $(t_i, y_i)$ $(i = 1, \cdots, n)$가 주어졌을 때, 임의의 곡선 $g$에 대하여

$$y = g(t) + \text{오차}$$

라는 유연한 모형을 생각하여 데이터에의 적합을 시도한다. 최소제곱법과 함께 데이터에의 적합의 정도를 나타내는 잔차제곱합을 사용하거나, 곡선의 변동, 즉 **조잡함**(roughness)에 대한 벌점은 곡선의 2계미분 등을 이용하여 만든다. 이것에 의해 벌점붙은 제곱합(penalized sum of squares)은

$$S(g) = \sum_{i=1}^{n} \{y_i - g(t_i)\}^2 + \alpha \int_a^b \{g''(t)\}^2 dt$$

와 같은 형태가 된다.

벌점붙은 제곱합 $S(g)$를 최소화하여 얻어지는 곡선이 스플라인, 즉 구분적인 다항식으로 표현되는 매끄러운 곡선이다. 이것을 **평활 스플라인**이라 부른다. Green and Silverman(1994)은 이와 같은 사고방식을 "roughness penalty approach"라 부르고 상세하게 소개하고 있다. 곡선을 적합하는 것은 단지 데이터에의 적합성을 좋게하는 것뿐만아니라, 그다지 변동이 심하지 않는 매끄러운 곡선을 추정하는 것이다. $g$를 2계미분가능이라 하고, 간편한 계산을 위해 $g''(t)$의 제곱적분 $\int \{g''(t)\}^2 dt$를 roughness의 측도라 하자(상세한 것은 Green and Silverman(1994)을 참조).

각 점 $t_1, \cdots, t_n$에 있어서 $z_1, \cdots, z_n$이라는 값이 주어져 있다고 할 때, 점 $(t_i, z_i)$를 보간하는 매끄러운 곡선 $g(t)$를 구하는 문제를 생각한다. 즉, $g(t_i) = z_i$가 성립하면 된다. 그러나 그와 같은 곡선은 무수히 존재하고, 어느 정도 미분가능성을 보장한 채로 비교적 변동이 작은 곡선을 선택한다.

먼저 $(2m-1)$차 스플라인은

$$W_2^m = \{g \mid g\text{는 } [a, b]\text{에서 } (m-1)\text{번 절대연속미분가능이고, } g^{(m)}(t) \in L_2[a, b]\}$$

라는 매끄러운 함수족에서 $\int \{g''(t)\}^2 dt$를 최소로 하는 것을 보일 수 있다.

**정리** 2.1 점 $(t_i, z_i)$를 보간하면서 $W_2[a, b]$에 속한 $g(t)$ 중에서 $\int \{g''(t)\}^2 dt$를 최소로 하는 것은 NCS이다.

주어진 모든 점을 보간하는 NCS는 유일하게 존재한다. Green and Silverman(1994)의 논의대로 서술하면 다음과 같다.

$t_1, \cdots, t_n$을 절점으로 갖는 NCS를 $g$라 하면 각 절점에 있어서 $g$ 자신 및 2차 도함수의 값을 부여함으로써 $g$가 결정된다. 즉,

$$g_i = g(t_i), \quad \gamma_i = g''(t_i), \quad (i = 1, \cdots, n)$$

을 준다. 자연스플라인의 정의에 의해 $\gamma_1 = \gamma_n = 0$이다. 벡터 $\boldsymbol{g} = (g_1, \cdots, g_n)^T$, $\gamma = (\gamma_2, \cdots, \gamma_{n-1})^T$가 3차자연 스플라인 $g$를 결정하기 위한 필요충분조건은

$$Q^T \boldsymbol{g} = R\gamma$$

이다. 여기서 $h_i = t_{i+1} - t_i \, (i = 1, \cdots, n-1)$이라 두고,

$$Q = \begin{bmatrix} h_1^{-1} & 0 & \cdots & 0 \\ -h_1^{-1} - h_2^{-1} & h_2^{-1} & & \vdots \\ h_2^{-1} & -h_2^{-1} - h_3^{-1} & \ddots & 0 \\ \vdots & & & -h_{n-2}^{-1} - h_{n-1}^{-1} \\ 0 & \cdots & 0 & h_{n-1}^{-1} \end{bmatrix},$$

$$R = \frac{1}{3} \begin{bmatrix} h_1 + h_2 & h_2/2 & 0 & \cdots & 0 \\ h_2/2 & h_2 + h_3 & h_3/2 & & \vdots \\ 0 & & \cdots & h_{n-3}/2 & 0 \\ \vdots & & \ddots & h_{n-3} + h_{n-2} & h_{n-2}/2 \\ 0 & \cdots & 0 & h_{n-2}/2 & h_{n-2} + h_{n-1} \end{bmatrix}.$$

$R$은 정부호행렬이므로 $\boldsymbol{g} = (z_1, \cdots, z_n)^T$가 주어지면 $\gamma$는 $\gamma = R^{-1}Q^T\boldsymbol{g}$로 유일하게 결정된다. 따라서 다음 정리를 얻는다.

**정리** 2.2 절점 $t_1, \cdots, t_n$에 대하여 $g(t_i) = z_i \, (i = 1, \cdots, n)$를 만족하는 NCS는 유일하게 존재한다.

NCS에 대해서는 roughness penalty가 $\boldsymbol{g}$ 혹은 $\gamma$의 2차형식 $\int_a^b \{g''(t)\}^2 dt = \boldsymbol{g}^T K \boldsymbol{g}$로 표현할 수 있다. 단, $K = QR^{-1}Q^T$이고, 이것은 양의 정부호대칭행렬이다. $m \geq 3$인 경우에도 $(2m-1)$차 스플라인의 최적성, 특징, 유일성을 논의할 수 있다.

## 3. 편스플라인(PARTIAL SPLINE)

비모수 회귀모형에 있어서 평활화와 같은 방법으로 벌점붙은 제곱합을 정식화한다. 단, 여기서는 보다 일반적인 가중치 $w_1, \cdots, w_n$을 붙인 형태로 표현한다. $g$가 $W_2^m[a, b]$에 속한

임의의 함수라 하면 벌점붙은 제곱합은

$$S_W(\boldsymbol{\beta}, g) = \sum_{i=1}^{n} w_i\{y_i - \boldsymbol{x}_i^T\boldsymbol{\beta} - g(t_i)\}^2 + \alpha\int_a^b \{g^{(m)}(t)\}^2 dt \tag{3.1}$$

와 같이 정의된다.

$\boldsymbol{\beta}$ 및 $g$의 추정량은 이 $S_W(\boldsymbol{\beta}, g)$를 최소로 하는 것이다. 그 결과 $g$의 추정량은 $(2m-1)$차 자연스플라인일 필요가 있고, $S_W(\boldsymbol{\beta}, g)$의 최소화를 선형대수문제로 귀착시킬 수 있다.

그런데 특히 중회귀의 경우에는 비모수적 성분의 설명변수 $t_i$에 대하여 순서를 붙이거나 동점을 다루기 위해 비모수 회귀문제에서 다른 방법보다도 형식적인 접근을 생각하는 것이 바람직하다.

$t_1, \cdots, t_n$은 순서와 동점을 무시한다. 이들에 순서를 붙여 서로 다른 값만을 나열한 것을 $s_1 < \cdots < s_q$ $(q \geq 3)$라 한다. 여기서 $t_1, \cdots, t_n$과 $s_1, \cdots, s_q$와의 관계를 파악하기 위해 결합행렬(incidence matrix) $N$을 사용한다. $N$은 $n \times q$행렬이고, 그 각 원소는

$$N_{ij} = \begin{cases} 1 & \text{if } t_i = s_j, \\ 0 & \text{otherwise} \end{cases}$$

로 정의된다. $\boldsymbol{g} = (g(s_1), \cdots, g(s_q))^T$라 두면

$$(g(t_1), \cdots, g(t_n))^T = N\boldsymbol{g}$$

가 된다.

$g$를 $s_1, \cdots, s_q$를 절점으로 갖는 $(2m-1)$차 자연스플라인이라 할 때, 결합행렬 $N$을 도입함으로써 $S_W(\boldsymbol{\beta}, g)$의 각 항은 $\boldsymbol{\beta}$ 및 $\boldsymbol{g}$의 함수인 2차형식으로 표현된다. 관측벡터를 $\boldsymbol{y} = (y_1, \cdots, y_n)^T$, 모수적 성분의 계획행렬(design matrix)을 $X = (\boldsymbol{x}_1, \cdots, \boldsymbol{x}_n)^T = (x_{ij})$, 가중행렬을 $W = \mathrm{diag}(w_1, \cdots, w_n)$이라 하면

$$S_W(\boldsymbol{\beta}, g) = (\boldsymbol{y} - X\boldsymbol{\beta} - N\boldsymbol{g})^T W(\boldsymbol{y} - X\boldsymbol{\beta} - N\boldsymbol{g}) + \alpha\boldsymbol{g}^T K\boldsymbol{g} \tag{3.2}$$

가 된다. 여기서 $K = QR^{-1}Q^T$는 $q \times q$행렬이고, $Q$, $R$의 정의는 정리 2.2에서 $t_i$ 대신에 $s_j$ $(j = 1, \cdots, q)$를 사용한 것이다.

(3.2)를 최소로하는 $\boldsymbol{\beta}$, $\boldsymbol{g}$는

$$X^T WX\boldsymbol{\beta} = X^T W(\boldsymbol{y} - N\boldsymbol{g}), \qquad (\boldsymbol{g}: \text{ known}) \tag{3.3}$$

$$(N^T WN + \alpha K)\boldsymbol{g} = N^T W(\boldsymbol{y} - X\boldsymbol{\beta}), \qquad (\boldsymbol{\beta}: \text{ known}) \tag{3.4}$$

의 해로써 주어진다. 즉, 구하려는 $\boldsymbol{\beta}$, $\boldsymbol{g}$의 추정량은 행렬방정식

$$\begin{bmatrix} X^T WX & X^T WN \\ N^T WX & N^T WN + \alpha K \end{bmatrix} \begin{bmatrix} \boldsymbol{\beta} \\ \boldsymbol{g} \end{bmatrix} = \begin{bmatrix} X^T \\ N^T \end{bmatrix} W\boldsymbol{g} \tag{3.5}$$

의 해이다.

행렬방정식 (3.5)가 유일한 해를 갖기 위해서는 임의의 $\alpha$에 대하여 (3.5)의 좌변의 행렬이 양의 정부호 대칭행렬이면 된다. 자연스플라인에 의한 보간의 유일성에 의해 (3.1)의

$S_W(\boldsymbol{\beta},\, g)$를 최소로하는 곡선 $g$도 유일하게 존재한다. 이것을 **편스플라인**(partial spline)이라 부른다.

가중행렬 $W$, 결합행렬 $N$에 대한 스플라인 **평활행렬**(smoothing matrix)을

$$S_\alpha = N(N^TWN + \alpha K)^{-1}N^TW$$

라 하면 (3.4)에 의해 각 절점에 있어서 $\boldsymbol{g}$의 값은

$$N\boldsymbol{g} = (g(t_1),\, \cdots,\, g(t_n))^T = S_\alpha(\boldsymbol{y} - X\boldsymbol{\beta}) \tag{3.6}$$

와 같이 쓸 수 있다. 이 방법의 결점은 $\alpha$가 매우 클(혹은 작을) 때에 수렴속도가 느리다는 것이다(실제로는 큰 문제가 되지 않는다).

연립방정식 (3.3), (3.4)는 반복법을 사용하지 않고도 직접 풀 수 있다. (3.4)는 (3.6)과 같이 쓸 수 있으므로, 이것을 (3.3)에 대입하여 $V\boldsymbol{g}$를 제거하면

$$X^TWX\boldsymbol{\beta} = X^TW\boldsymbol{y} - X^TW\,S_\alpha(\boldsymbol{y} - X\boldsymbol{\beta}),$$
$$X^TW(I - S_\alpha)X\boldsymbol{\beta} = X^TW(I - S_\alpha)\boldsymbol{y}$$

가 된다. (3.5)의 행렬이 정부호 대칭이라면 $X^TW(I - S_\alpha)X$도 정부호 대칭이므로 유일해가 존재한다. 이상으로부터 $\boldsymbol{\beta}$ 및 $\boldsymbol{g}$의 추정량은 다음과 같이 쓸 수 있다.

$$\begin{aligned}\hat{\boldsymbol{\beta}} &= \{X^T(I - S_\alpha)X\}^{-1}X^T(I - S_\alpha)\boldsymbol{y},\\ N\hat{\boldsymbol{g}} &= S_\alpha(\boldsymbol{y} - X\hat{\boldsymbol{\beta}}).\end{aligned} \tag{3.7}$$

이들 계산은 그다지 어렵지 않고, 비모수 회귀모형에 대한 평활화 알고리즘을 이용하면 $O(n)$의 계산시간으로 추정량을 구할 수 있다.

## 4. 편스플라인 추정량의 점근적 성질

Rice(1986)는 준모수회귀모형 (1.1)에 있어서 $\boldsymbol{x}_i$와 $t_i$ 사이에 어떤 관계가 존재할 때, $\boldsymbol{\beta}$의 추정량이 일반적으로 편의를 갖는다는 것을 점근론적으로 증명하였다. 이 절에서는 $\hat{\boldsymbol{\beta}}$의 점근 정규성을 증명한다.

간편한 계산을 위해 $[a,\, b] = [0,\, 1]$이라 하자. $0 \le t_{1n} < t_{2n} < \cdots < t_{nn} \le 1$에 대하여 모형을

$$y_{in} = \boldsymbol{x}_{in}^T\boldsymbol{\beta} + g(t_{in}) + \varepsilon_{in}, \qquad (i = 1,\, 2,\, \cdots,\, n)$$

과 같이 쓸 수 있다. 여기서 $\{\varepsilon_{in}\}$은 평균 0, 분산 $\sigma^2$인 독립이고 동일 분포에 따르는 확률변수라 하자. 또한 벌금붙은 제곱합을

$$S(\boldsymbol{\beta},\, g) = \frac{1}{n}\sum_{i=1}^{n}\{y_{in} - \boldsymbol{x}_{in}^T\boldsymbol{\beta} - g(t_{in})\}^2 + \lambda\int_0^1\{g^{(m)}(t)\}^2dt$$

라 정의하고($m \ge 2$), 이것을 최소로하는 $\boldsymbol{\beta} \in R^p$ 및 $g \in W_2^m$을 추정량이라 하자((3.1)에서 $w_i = 1$, 평활모수 $\alpha$를 $n\lambda$로 바꾼 것).

$\boldsymbol{y} = (y_{1n}, \cdots, y_{nn})^T$, $X = (x_{1n}, \cdots, x_{nn})^T$라 둔다. 비모수 단순회귀에 있어서 평활 행렬은

$$S_\lambda = (I + n\lambda K)^{-1}, \quad (K\text{는 } m \text{ 및 } t_{1n}, \cdots, t_{nn}\text{에 의존하는 양의 정부호대칭행렬})$$

이 된다. 따라서 (3.7)에 의해 $\boldsymbol{\beta}$ 및 $\boldsymbol{g} = (g(t_{1n}), \cdots, g(t_{nn}))^T$의 추정량 즉, 편스플라인 추정량은

$$\begin{aligned} \hat{\boldsymbol{\beta}} &= \{X^T(I - S_\lambda)X\}^{-1}X^T(I - S_\lambda)\boldsymbol{y}, \\ \hat{\boldsymbol{g}} &= S_\lambda(\boldsymbol{y} - X\hat{\boldsymbol{\beta}}) \end{aligned} \tag{4.1}$$

와 같이 쓸 수 있다.

자연스플라인 공간의 기저의 하나로, Demmler and Reinsch(1975)에 의해 도입된 것으로 다음과 같이 정의되는 쌍직교성(biorthogonality)을 갖는 $(2m-1)$차 자연스플라인의 집합 $\{\phi_{jn}(t)\}_{j=1}^n$이 있다.

$$\frac{1}{n}\sum_{i=1}^n \phi_{jn}(t_{in})\phi_{kn}(t_{in}) = \delta_{jk}, \qquad \int_0^1 \phi_{jn}^{(m)}(t)\phi_{kn}^{(m)}(t)\,dt = \lambda_{kn}\delta_{jk},$$

여기서 $\delta_{jk}$는 Kronecker Delta이다.

LEMMA 4.1 $\phi_{jn} = (\phi_{jn}(t_{1n}), \cdots, \phi_{jn}(t_{nn}))^T$는 행렬 $K = QR^{-1}Q^T$의 고유벡터이고, 대칭인 $K$의 고유값을 $\omega_{jn}$이라 하면, $\lambda_{jn} = \omega_{jn}$이다. 따라서 $j = 1, \cdots, m$에 대하여 $\phi_{jn}(t)$는 각각 $(j-1)$차 다항식이고, 이 때, $\lambda_{jn} = 0$이다.

이 Lemma로부터 $S_\lambda = (I + n\lambda K)^{-1}$의 고유값은 $(1 + \lambda_{kn}K)^{-1}$ $(k = 1, \cdots, n)$이 된다. 또한 $\boldsymbol{y} = (y_{1n}, \cdots, y_{nn})^T$에 $S_\lambda$를 곱한다는 것(즉, spline smoothing)을 기저 $\{\phi_{kn}(t)\}$를 이용하여 표현할 수 있다. $\gamma_{kn} = n^{-1/2}\sum_{i=1}^n y_{jn}\phi_{kn}(t_{in})$이라 두면 $\boldsymbol{y}$는 $\{\phi_{kn}(t)\}$에 의해 직교분해할 수 있고,

$$y_{jn} = \frac{1}{\sqrt{n}}\sum_{k=1}^n \gamma_{kn}\phi_{kn}(t_{in})$$

이 된다. 이 사실로터 $S_\lambda\boldsymbol{y}$의 $i$번째 성분은

$$\{S_\lambda\boldsymbol{y}\}_i = \frac{1}{\sqrt{n}}\sum_{k=1}^n \frac{\gamma_{kn}}{1 + \lambda_{kn}\lambda}\phi_{kn}(t_{in})$$

과 같이 쓸 수 있다.

Rice(1986), Chen and Shiau(1991) 등이 설정한 가정은 $\boldsymbol{x}_{in} = (x_{i1n}, \cdots, x_{ipn})^T$와 $t_{in}$ 사이에 매끄러운 함수 $h_j$를 사용하여

$$x_{ijn} = h_j(t_{in}) + z_{ijn}, \qquad (j = 1, 2, \cdots, p)$$

라는 관계가 성립한다고 하자. 다음은 $k = 1, \cdots, n$에 대하여

$$\xi_{kjn} = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} z_{ijn} \phi_{kn}(t_{in}), \qquad h_{kjn} = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} h_j(t_{in}) \phi_{kn}(t_{in})$$

와 같이 직교변환하고, $i, j = 1, 2, \cdots, p$에 대하여 다음과 같이 가정한다.

[가정 1] $\lim_{n \to \infty} n^{-1} \sum_{i=1}^{n} \xi_{kjn} = 0$,

[가정 2] $p \times p$ 정부호행렬 $\Sigma = (\delta_{ij})$가 존재하고, $\lim_{n \to \infty} n^{-1} \sum_{i=1}^{n} \xi_{kin} \xi_{kjn} = \delta_{ij}$,

[가정 3] $\sup_{k=1,\ldots,n} |\xi_{kjn}| = O(\log n)$.

예를 들어, $z_{in} = (z_{i1n}, \cdots, z_{ipn})^T \sim$ i.i.d $N_p(0, \Sigma)$일 때, [가정 1], [가정 2]는 확률수렴의 의미로, [가정 3]은 $O_p(\log n)$의 의미로 각각 성립한다. 또한 $t_{in}$ $(i = 1, \cdots, n)$에 대해서는 다음을 가정한다.

[가정 4] 점 $t_{in}$은 $[0, 1]$에서 연속인 밀도함수 $f(t)$에 대하여 $\dfrac{2n-1}{2n} = \displaystyle\int_0^{t_{in}} f(t)dt$와 같이 쓸 수 있다는 의미로 정칙이다.

Rice(1986), Chen and Shiau(1991) 등은 다음의 두 Lemma가 Speckmann(1982)의 미제출논문에서 증명되어져 있다고 언급하고 있다. 여기서 $\dfrac{a(n)}{b(n)}$이 $n \to \infty$일 때 0에 수렴하지 않고, 또한 무한대에 발산하지 않을 때, $a(n) \asymp b(n)$이라 쓴다.

LEMMA 4.2 (Speckmann) [가정 4]에 의해

(1) $\lambda_{kn} = ck^{2m}(1 + o(1))$이다. 여기서 $c$는 밀도 $f$에 의존하는 상수이고, $o(1)$은 $k_{1n} \to \infty$, $k_{2n} = o(n^{2/(2m+1)})$인 임의의 수열 $k_{1n}$, $k_{2n}$에 대하여 $n \to \infty$일 때, $k_{1n} \le k \le k_{2n}$이 되는 $k$에 관하여 평등하게 0에 수렴하는 항이다.

(2) $\lambda \asymp n^{-\delta}$ $(0 < \delta < 1)$일 때, $\delta$만에 의존하는 양의 상수 $\tau$에 대하여

$$\sum_k \frac{1}{1 + \lambda_{kn}\lambda} = O(n^{-1/2-\tau}), \qquad \sum_k \frac{1}{(1 + \lambda_{kn}\lambda)^2} = O(n^{-1/2m})$$

이다.

다음은 $\boldsymbol{g} = (g(t_{1n}), \cdots, g(t_{nn}))^T$, $\boldsymbol{h}_j = (h(t_{1n}), \cdots, h(t_{nn}))^T$ $(j = 1, \cdots, p)$라 두고, $B_{1,\lambda}^2 = n^{-1} \boldsymbol{g}^T (I - S_\lambda)^2 \boldsymbol{g}$, $B_{2j,\lambda}^2 = n^{-1} \boldsymbol{h}_j^T (I - S_\lambda)^2 \boldsymbol{h}_j$라 하자. $B_{1,\lambda}^2$은 $\boldsymbol{\beta} = \mathbf{0}$임을 알 수 있을 때, 편스플라인에 있어서 $\boldsymbol{g}$의 추정량의 Average Squared Bias를 나타낸다. 즉, $\hat{\boldsymbol{g}} = S_\lambda \boldsymbol{y}$에 대하여 $B_{1,\lambda}^2 = n^{-1} \sum_{i=1}^{n} \{g(t_{in}) - E\hat{g}(t_{in})\}^2$이다. 이에 대하여 다음 Lemma가 성립한다.

LEMMA 4.3 (Speckmann) 가정 4에 의해 $g \in W_2^m$이라면 $B_{1,\lambda}^2 = O(\lambda)$이다.

또한 $B_{2j,\lambda}^2$에 대해서도 $h \in W_2^m$이면 $B_{2j,\lambda}^2 = O(\lambda)$가 성립한다. 다음은 Rice(1986)의 핵심이 되는 Lemma를 다룬다. 여기서 $\boldsymbol{x}_j = (x_{1jn}, \cdots, x_{njn})^T$라 두면 $X = [\boldsymbol{x}_1, \cdots, \boldsymbol{x}_p]$이다.

LEMMA 4.4 [가정 1]-[가정 4]가 성립하고, $g, h_j \in W_2^m$ $(j = 1, \cdots, p)$라 하자. $\lambda \asymp n^{-\delta}$ $(0 < \delta < 1)$일 때, $i, j = 1, \cdots, p$에 대하여

(a) $n^{-1}\boldsymbol{x}_i^T(I - S_\lambda)\boldsymbol{x}_j = \sigma_{ij} + o(1)$,

(b) $n^{-1}\boldsymbol{x}_i^T(I - S_\lambda)^2\boldsymbol{x}_j = \sigma_{ij} + o(1)$,

(c) $n^{-1}\boldsymbol{x}_i^T(I - S_\lambda)\boldsymbol{g} = o(n^{-1/2}) + O(\lambda)$,
단, $h_i(t)$가 기껏해야 $(m-1)$차 다항식일 때, $O(\lambda)$ 항은 0이다.

(d) $n^{-1}\boldsymbol{x}_i^T S_\lambda^2 \boldsymbol{x}_j = O(1)$,

(e) trace$S_\lambda^2 = O(\lambda^{-1/2m})$.

오차 $\boldsymbol{\varepsilon} = (\varepsilon_{1n}, \cdots, \varepsilon_{nn})^T$에 분포를 가정하면 $E(\boldsymbol{y}) = X\boldsymbol{\beta} + \boldsymbol{g}$, $\mathrm{Var}(\boldsymbol{y}) = \sigma^2 I$이다. 따라서 $\boldsymbol{\beta}$의 추정량 (4.1)의 평균과 분산은

$$E(\hat{\boldsymbol{\beta}}) = \boldsymbol{\beta} + \{X^T(I - S_\lambda)X\}^{-1}X^T(I - S_\lambda)\boldsymbol{g},$$
$$\mathrm{Var}(\hat{\boldsymbol{\beta}}) = \sigma^2\{X^T(I - S_\lambda)X\}^{-1}X^T(I - S_\lambda)^2X\{X^T(I - S_\lambda)X\}^{-1}$$

와 같이 표현된다. 이들 점근적 성질을 다음 정리에서 다룬다.

**정리** 4.1 (Rice(1986)). Lemma 4.4와 같은 조건에서

(a) $E(\hat{\boldsymbol{\beta}}) = \boldsymbol{\beta} + o(n^{-1/2}) + O(\lambda)$, 단, $h_j$ $(j = 1, \cdots, p)$가 기껏해야 $(m-1)$차 다항식일 때, $O(\lambda)$ 항은 0이다.

(b) $\mathrm{Var}(\hat{\boldsymbol{\beta}}) = n^{-1}\sigma^2\Sigma^{-1} + o(n^{-1})$이다.

**증명**: (a)는 Lemma 4.4의 (a)와 (c), (b)는 Lemma 4.4의 (a)와 (b)로부터 증명된다. □

이와 같이 설명변수 $X$, $t$ 사이에 비선형적인 상관관계가 존재할 때, $\hat{\boldsymbol{\beta}}$에는 참인 곡선 $g$와 평활모수 $\lambda$에 의존하는 편의가 생긴다는 것을 알 수 있다. 따라서 $\boldsymbol{\beta}$를 모수적 수렴속도 $n^{-1/2}$로 추정하기 위해서는 $O(\lambda)$ 항이 $o(n^{-1/2})$ 항보다도 빠르게 수렴하도록 $\lambda = o(n^{-1/2})$이라 하자. 다시말하면 $\lambda$의 값을 작게할 필요가 있다. 그러면 추정곡선 $\hat{\boldsymbol{g}}$의 매끄러움이 억제되어(undersmoothing)버린다. 결국, 여기서는 다루지 않지만, 평활모수 $\lambda$를 cross-validation 등의 방법으로 자동적으로 선택할 경우에는 주의하여야 한다.

모수적 성분의 수렴속도 $n^{-1/2}$로 $\boldsymbol{\beta}$를 추정할 수 있을 때, 즉, $\lambda = o(n^{-1/2})$인 경우에는 다음의 가정을 추가하면 $\boldsymbol{\beta}$의 점근 정규성을 보일 수 있다.

[**가정** 5] $\boldsymbol{z}_{in} = (z_{i1n}, \cdots, z_{ipn})^T$는 평균 0, 공분산행렬 $\Sigma$이고, 유한인 절대3차적률을 갖는 확률벡터이다($\{\boldsymbol{z}_{in}\} \sim$ i.i.d.).

[**가정** 6] $\{\varepsilon_{in}\}$은 평등하게 유계인 3차적률을 갖는 독립확률변수이고, $\varepsilon_{in}$은 $\boldsymbol{z}_{in}$과 독립이다.

[가정 5]에서는 Lemma 4.4의 (a)-(d)는 확률수렴의 의미를 갖는다.

다음 정리는 Chen and Shiau(1991)의 결과를 편스플라인의 경우로 수정한 것이다.

**정리** 4.2 Lemma 4.4의 조건과 [가정 5], [가정 6]에서 $\lambda = o(n^{-1/2})$이면

$$\sqrt{n}(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta}) \to N(\mathbf{0}, \sigma^2\Sigma^{-1}) \qquad \text{(in distribution)}$$

이 성립한다.

**증명**: $\hat{\boldsymbol{\beta}} - \boldsymbol{\beta} = \{X^T(I - S_\lambda)X\}^{-1}X^T(I - S_\lambda)(\boldsymbol{g} + \boldsymbol{\varepsilon})$이고, 정리 4.1 (a)의 증명으로부터

$$\sqrt{n}\{X^T(I - S_\lambda)X\}^{-1}X^T(I - S_\lambda)\boldsymbol{g} = \sqrt{n}\{o_p(n^{-1/2}) + O_p(\lambda)\} = o_p(1)$$

이다. $x_{ijn} = h_j(t_{in}) + z_{ijn}$에 대응시켜 $\boldsymbol{x}_j = \boldsymbol{h}_j + \boldsymbol{z}_j$, $X = H + Z$라 쓰고,

$$X^T(I - S_\lambda)\boldsymbol{\varepsilon} = H^T(I - S_\lambda)\boldsymbol{\varepsilon} + Z^T\boldsymbol{\varepsilon} - Z^TS_\lambda\boldsymbol{\varepsilon}$$

과 같이 분해한다. 여기서 $E\{\boldsymbol{h}_j^T(I - S_\lambda)\boldsymbol{\varepsilon}\} = 0$이고, Lemma 4.3에 의해

$$\begin{aligned}\mathrm{Var}\{n^{-1/2}\boldsymbol{h}_j^T(I - S_\lambda)^2\boldsymbol{\varepsilon}\} &= n^{-1}\sigma^2\boldsymbol{h}_j^T(I - S_\lambda)\boldsymbol{h}_j \\ &= \sigma^2B_{2j,\lambda}^2 = O(\lambda) = o(1)\end{aligned}$$

이므로 Chebyshev 부등식을 이용하면 $n^{-1/2}H^T(I - S_\lambda)\boldsymbol{\varepsilon} = o_p(1)$임을 알 수 있다. 또한 $E(\boldsymbol{z}_j^TS_\lambda\boldsymbol{\varepsilon}) = 0$이고, Lemma 4.4의 (e), [가정 5]와 [가정 6]으로부터

$$\begin{aligned}\mathrm{Var}(\boldsymbol{z}_j^TS_\lambda\boldsymbol{\varepsilon}) &= E_z[E\{(\boldsymbol{z}_j^TS_\lambda\boldsymbol{\varepsilon})^2|\boldsymbol{z}_j\}] \\ &= E_z[\mathrm{Var}(\boldsymbol{z}_j^TS_\lambda\boldsymbol{\varepsilon}|\boldsymbol{z}_j) + \{E(\boldsymbol{z}_j^TS_\lambda\boldsymbol{\varepsilon}|\boldsymbol{z}_j)\}^2] \\ &= \sigma^2E(\boldsymbol{z}_j^TS_\lambda\boldsymbol{z}_j) \\ &= \sigma^2\sigma_{jj}\ \mathrm{trace}\{S_\lambda^2\} = O(\lambda^{-1/2m}) = o(n^{1/2})\end{aligned}$$

이므로 Chebyshev 부등식에 의해 $n^{-1/2}Z^TS_\lambda\boldsymbol{\varepsilon} = o_p(1)$임을 알 수 있고, 마지막으로

$$n^{-1/2}Z^T\boldsymbol{\varepsilon} \to N(\mathbf{0}, \sigma^2\Sigma) \quad \text{(in distribution)} \tag{4.2}$$

임을 보이면 된다. 임의의 $p$차 열벡터 $\boldsymbol{v}$에 대하여 $\boldsymbol{a} = (a_1, \cdots, a_n)^T = Z\boldsymbol{v}$라 두면 $\boldsymbol{v}^TZ^T\boldsymbol{\varepsilon} = \sum_{i=1}^n a_i\varepsilon_{in}$이다. [가정 5]와 [가정 6]으로부터 $a_1, \cdots, a_n$도 독립인 확률변수이고,

$$n^{-3/2}\sum_{i=1}^n E|a_i\varepsilon_{in}|^3 \le n^{-1/2}\sup_i E|\varepsilon_{in}|^3E|a_1|^3 = o(1)$$

이므로 $\boldsymbol{v}^TZ^T\boldsymbol{\varepsilon}$는 Lyapounov 조건을 만족한다. 따라서 중심극한정리에 의해

$$n^{-1/2}\boldsymbol{v}^TZ^T\boldsymbol{\varepsilon} \to N(\mathbf{0}, \sigma^2\boldsymbol{v}^T\Sigma) \quad \text{(in distribution)}$$

임을 알 수 있다. 따라서 Cramer-Wold device에 의해 (4.2)가 증명되고, 이 사실과 Lemma 4.4 (a)에 의해 정리가 증명되었다. □

# 참고문헌


[1] 김 종태 (1998). 이차 평활스플라인. < 응용통계연구 >. 제11권. 363-376

[2] Chen, H. and Shiau, J. (1991). A two-stage spline smoothing method for partially linear models. *Journal of Statistical Planning and Inference*, 27, 187-201.

[3] Demmler, A. and Reinsch, C. (1975). Oscillation Matrices with Spline Smoothing. *Numer. Math.*, 24, 375-382.

[4] Green, P. and Silverman, B.W. (1994). *Nonparametric Regression and Generalized Linear Models.* Chapman and Hall.

[5] Green, P., Jennison, C. and Seheult, A. (1985). Analysis of field experiments by least squares smoothing. *Journal of the Royal Statistical Society.* Series B, 47, 299-315.

[6] Engle, R.F., Granger, C.W.J., Rice, J. and Weiss, A. (1986). Semiparametric estimates of the relation between weather and electricity sales. *Journal of American Statistical Association.* 81, 310-320.

[7] Rice, J. (1986). Convergence Rates for Partially Splined Models. *Statistics and Probability Letters*, 4. 203-208.

[8] Speckman, P. (1982). *Efficient nonparametric regression with cross-validated smoothing splines.* manuscript.

[9] Speckman, P. (1988). Kernel Smoothing in Partial Linear Models. *Journal of the Royal Statistical Society.* Series B, 50, 413-436.


[ 1999년 3월 접수, 2000년 4월 채택 ]

# Asymptotics Normality for Bias of Partial Spline Estimator

In-Sun Chu[1) ] Jae-Ryong Choi[2)]

ABSTRACT

In this paper, we describes smoothing spline in nonparametric regression and some asymptotic results for estimates of the regression cofficients in the parametric part were biased on semiparametric regression estimator.

*Keywords:* Smoothing; Semiparametric regression model; Spline; Smoothing matrix.

1) Lecturer, Department of Mathematics, Dong-A University. E-mail: i-s-chu@hanmail.net
2) Professor, Department of Mathematics, Dong-A University. E-mail: jrchoi@mail.donga.ac.kr