

비모수적 회귀함수 추정에서 평활량의 선택에 관한 연구¹⁾

석 경 하²⁾

요 약

비모수적 커널 회귀함수 추정법에서 평활량(bandwidth or smoothing parameter)의 선택은 아주 중요한 문제이다. 교차타당성(cross-validation)방법에 의한 평활량은 최적평활량으로의 상대적 수렴속도(relative convergence rate)가 $n^{-1/10}$ 로 상당히 느리다는 것을 알고 있다. 본 연구는 삽입방법(plug-in method)에 의해 선택된 평활량의 상대적 수렴속도가 교차타당성 방법보다 더 빠른 $n^{-2/7}$ 이 됨을 보였다. 그리고 모의실험을 통하여 소표본에서도 삽입방법이 교차타당성 방법보다 우수함을 입증하였다.

1. 서론

비모수적 커널 회귀함수 추정법은 여러 가지 평활법과 더불어 반응치 Y_1, \dots, Y_n 에 대한 회귀모형

$$Y_i = m(x_i) + \varepsilon_i, \quad i = 1, \dots, n \quad (1)$$

의 회귀함수(regression function) $m(x)$ 에 대한 추정법으로 많이 사용되고 있다. 여기에서 ε_i 는 서로 독립이고 평균이 0, 분산이 σ^2 인 확률변수이다. m 에 대한 커널추정법중 대표적인 것으로는 Nadaraya(1964), Watson(1964), Priestley and Chao(1972), Gasser and Müller(1979, 1984) 그리고 Fan(1992)의 연구 등이 있다. 이러한 여러 가지 방법에 대한 비교와 논의는 Chu and Marron(1991)과 Jones, Davis and Park(1994)에서 찾아 볼 수가 있다.

커널추정법에서 평활량의 선택은 아주 중요한 문제이다. 자료기저 평활량(data-driven bandwidth)의 선택에 관하여 Härdle, Hall and Marron(1988)에서 교차타당법에 의한 평활량 \hat{h}_{cv} 에 관하여 연구를 하였는데 이는 점근최적평활량(h_a)으로의 상대적 수렴속도($\hat{h}_{cv}/h_a - 1$)가 $n^{-1/10}$ 로 상당히 느리다는 것을 지적하고 있다.

본 연구에서는 삽입방법에 의한 평활량 \hat{h}_p 의 상대적 수렴속도($\hat{h}_p/h_a - 1$)는 차수가 작은 커널

1) 이 논문은 1994년도 한국학술진흥재단의 공모과제 연구비에 의하여 연구되었음

2) (621-749) 경남 김해시 어방동 607 인제대학교 통계학과 조교수.

함수를 사용했을 때에도 $n^{-2/7}$ 로 \hat{h}_{cv} 보다 훨씬 빠르다는 것을 2절에서 보이고 또한 소표본에서의 수행능력도 더 우수하다는 것을 3절에서 보일 것이다.

2. 삼입방법

평활량의 선택에 관한 연구가 본 연구의 초점이므로 계산의 간편성을 위하여 계획점(design points)이 단위구간에서 등간격고정(equally spaced on the unit interval and fixed, *i.e.*, $x_i = (i-0.5)/n, i = 1, \dots, n$)이라 가정하자. 그리고 회귀함수 m 에 관한 여러 가지 추정법 중에서 Pristley and Chao(1972)에 의한 방법,

$$\hat{m}_h(x) = \frac{1}{nh} \sum_{i=1}^n K\left(\frac{x-x_i}{h}\right) Y_i \quad (2)$$

을 이용하겠다. 여기에서 K 는 커널함수이고 h 는 평활량이다. 평활량 h 는 $\hat{m}_h(x)$ 에 결정적인 역할을 한다는 것은 많이 알려진 사실이다. 커널함수 K 는 $(-1, 1)$ 에서 정의된 연속이고 대칭인 함수이고 또한 차수(order)가 p , 즉

$$\int_{-1}^1 z^j K(z) dz = \begin{cases} 1, & j=0, \\ 0, & j=1, \dots, p-1, \\ \mu_{0,p} \neq 0, & j=p. \end{cases} \quad (3)$$

를 가정한다.

커널함수의 선택에 관해서는 Gasser, Müller and Mammitzsch(1985)의 연구가 있다. 그런데 큰 차수의 커널함수를 사용함으로써 얻어지는 장점이 있는 반면 단점도 여러 가지 측면으로 지적됨으로 (Härdle, Hall and Marron (1992) 참조) 본 연구에서는 2차(second order, $p=2$) 커널함수를 사용하겠다.

$m(x)$ 의 추정량 $\hat{m}_h(x)$ 의 수행측도(performance measure)로써 Mean Averaged Squared Error,

$$MASE(h) = \frac{1}{n} \sum_i^* E\{\hat{m}_h(x_i) - m(x_i)\}^2 \quad (4)$$

를 고려해보자. 여기에서 첨자 i 의 범위는 $0 < c < d < 1$ 인 c 와 d 에 대하여 $c < x_i < d$ 를 만족하는 값들이다. (범위를 축소시킨 것은 경계점 근방에서 발생하는 큰 편의를 본 연구의 대상에서 제외시키고자 함이다.)

(4)식의 $MASE(h)$ 에 대한 근사치($AMASE$, Asymptotic MASE)는

$$AMASE(h) = \frac{(d-c)}{nh} \sigma^2 R(K) + h^4 \theta_{2,2} \mu_{0,2}^2 / 4 \quad (5)$$

이다. 여기에서 $R(K) = \int K^2(x) dx$, $\theta_{k,l} = \int_A m^{(k)}(x) m^{(l)}(x) dx$ 인데 $A = \{c < x < d\}$ 이다.

식(5)의 $AMASE(h)$ 를 최소화하는 점근최적평활량 h_a 는

$$h_a = \left\{ \frac{(d-c)\sigma^2 R(K)}{\mu_{0,2}^2 \theta_{2,2}} \right\}^{1/5} n^{-1/5} \quad (6)$$

이다. (6)식에서 모르는 부분은 σ^2 과 $\theta_{2,2}$ 인데 이것을 추정하여 대입하는 것이 삽입방법이다. σ^2 에 대한 비모수적 $n^{1/2}$ 일치추정량 $\hat{\sigma}^2$ 은 Gasser, Sroka and Jennen-Steinmetz(1986)에서 찾아볼 수가 있다. 그러므로 $\theta_{2,2}$ 에 대한 추정이 중요한 문제가 된다. 다음절은 바로 이 문제를 관심 있게 다룬다.

2.1 $\theta_{k,l}$ 의 추정

$\theta_{k,l}$ 에 대한 커널추정량으로 가장 자연스럽게 생각 할 수 있는 것은

$$\hat{\theta}_{k,l}(g) = \int_A \hat{m}_g^{(k)}(x) \hat{m}_g^{(l)}(x) dx \quad (7)$$

이다. 여기에서

$$\hat{m}_g^{(\nu)}(x) = \frac{1}{ng^{\nu+1}} \sum_{i=1}^n L_\nu\left(\frac{x-x_i}{g}\right) Y_i \quad (8)$$

이고, g 는 평활량, L_ν 는 커널함수인데 (2)식의 평활량 h 와 커널함수 K 와는 별개의 것이다. 커널함수 L_ν 는 $(-1, 1)$ 에서 정의된 연속이고 대칭인 함수이고 또한 차수(order)가 $q(\geq \nu+2)$, 즉,

$$\int_{-1}^1 z^j L_\nu(z) dz = \begin{cases} 0, & j=0, 1, \dots, \nu-1, \nu+1, \dots, q-1 \\ \nu!, & j=\nu, \\ \mu_{\nu,l} \neq 0, & j=q. \end{cases}$$

임을 가정하자. (2)식에서의 K 는 L_0 로 생각 할 수도 있겠다. 위에서 잠시 언급한바와 같은 이유로 본 연구에서는 차수가 최소($q = \nu+2$)인 커널함수 L_ν 를 사용하겠다. 회귀함수 m 을 추정할 때와 마찬가지로 $\theta_{k,l}$ 을 추정할 때도 평활량 g 가 중요한 역할을 한다. $\hat{\theta}_{k,l}(g)$ 의 평균제곱오차(MSE)를 최소화하는 평활량 g 는 아래의 정리로부터 쉽게 구해진다.

정리 1. “ m 은 $(0, 1)$ 에서 $\max\{\max(k, l)+2, k+l\}$ 번 미분가능하다” 라는 가정에서 $\hat{\theta}_{k,l}(g)$ 의 기대치와 분산의 근사치는 아래와 같다.

$$i) E\{\hat{\theta}_{k,l}(g)\} \simeq \theta_{k,l} + n^{-1}g^{-k-l-1}\{\theta_{0,0} + (d-c)\sigma^2\} \int L_k L_l \\ + g^2 \left\{ \frac{\mu_{k,k+2}}{(k+2)!} \theta_{k+2,l} + \frac{\mu_{l,l+2}}{(l+2)!} \theta_{k,l+2} \right\} + o(n^{-1}g^{-k-l-1} + g^2)$$

$$ii) VAR\{\hat{\theta}_{k,l}(g)\} \simeq n^{-2}g^{-2k-2l-1} [(d-c)\sigma^4 + 2\sigma^2 R_A(m)] \left[\int \left\{ \int L_k(y) L_l(y+z) dy \right\}^2 dz \right] \\ + 4n^{-1}\theta_{k+l,k+l} + o(n^{-2}g^{-2(k+l+1)} + n^{-1})$$

증명) $\hat{\theta}_{k,l}(g) \simeq n^{-2}g^{-k-l-2} \left\{ \int L_k L_l \sum_i^* Y_i^2 + \sum_{i \neq j}^* Y_i Y_j L_k * L_l \left(\frac{x_i - x_j}{g} \right) \right\}$ 임을 이용하면 i)은 쉽

게 증명이 된다. 여기에서 *는 합성(convolution)을 나타낸다. $VAR\{\hat{\theta}_{k,l}(g)\} = I_1 + I_2 + I_3 + I_4$ 로 쓰자. 여기에서

$$I_1 \simeq 2n^{-4}g^{-2(k+l+2)} \int L_k L_l \sum_{i \neq j}^* cov(Y_i^2, Y_i Y_j) L_k * L_l \left(\frac{x_i - x_j}{g} \right) \\ = 4n^{-2}g^{-k-l-1} \theta_{k,l} \int L_k L_l + o(n^{-2}g^{-k-l-1})$$

$$I_2 \simeq 4n^{-4}g^{-2(k+l+1)} \sigma^2 \sum_{\substack{i \neq j \\ i \neq i}}^* cov(Y_i Y_j, Y_i Y_j) L_k * L_l \left(\frac{x_i - x_j}{g} \right) L_k * L_l \left(\frac{x_i - x_j}{g} \right) \\ = 4n^{-1}\sigma^2 \theta_{k+l,k+l} + o(n^{-1})$$

$$I_3 \simeq n^{-4}g^{-2(k+l+1)} \left\{ \int L_k L_l \right\}^2 \sum_i^* var(Y_i^2) \\ = O(n^{-3}g^{-2(k+l+1)})$$

$$I_4 \simeq n^{-4}g^{-2(k+l+2)} \sum_{i \neq j}^* var(Y_i Y_j) \left\{ \int_A L_k \left(\frac{x-x_i}{g} \right) L_l \left(\frac{x-x_j}{g} \right) dx \right\}^2 \\ = n^{-2}g^{-2k-2l-1} [(d-c)\sigma^4 + 2\sigma^2 R_A(m)] \left[\int \left\{ \int L_k(y) L_l(y+z) dy \right\}^2 dz \right] \\ + o(n^{-2}g^{-2k-2l-1})$$

참고 1. 우리가 여기에서 크게 관심이 있는 경우는 $k=l=2$, 즉 $\theta_{2,2}$ 의 추정에 관심이 있다. 위의 정리로부터 MSE를 최소화하는 g 는 $O(n^{-1/7})$ 이고 이때 $\theta_{2,4}$ 가 양수이면 MSE는

$O(n^{-4/7})$ 이 되고 $\theta_{2,4}$ 가 음수이면 MSE 는 $O(n^{-5/7})$ 이 된다는 것을 쉽게 알 수가 있다.

2.2 삽입평활량 h_p 의 선택

(2) 식에서 σ^2 에 $\hat{\sigma}^2$ $\theta_{2,2}$ 에 $\hat{\theta}_{2,2}(g)$ (여기에서 g 는 $MSE\{\hat{\theta}_{2,2}(g)\}$ 를 최소화하는 값)을 각각 대입한 것이 바로 우리가 구하고자하는 삽입평활량 h_p 인데, 이는

$$h_p = \left\{ \frac{(d-c)\hat{\sigma}^2 R(K)}{\hat{\theta}_{2,2}(g)\mu_{0,2}^2} \right\}^{1/5} n^{-1/5} \quad (9)$$

이다. 아래의 정리는 h_p 의 상대적 수렴속도가 $n^{-2/7}$ 이 됨을 보여준다.

정리 2. 정리 1과 같은 가정에서 h_p 의 상대적 수렴속도는 아래와 같다.

$$(h_p/h_a - 1) = O(n^{-2/7}).$$

증명) $AMASE(h) = n^{-1}h^{-1}(d-c)\hat{\sigma}^2 R(K) + h^4\hat{\theta}_{2,2}(g)\mu_{0,2}^2/4$ 라면

$$\begin{aligned} AMASE(h_a) &= AMASE(h_a) - AMASE(h_a) \\ &= n^{-1}h_a^{-2}(d-c)R(K)(\sigma^2 - \hat{\sigma}^2) + h_a^3\mu_{0,2}^2(\hat{\theta}_{2,2}(g) - \theta_{2,2}) \\ &\quad + O(h_a^3(g^2 + n^{-1}g^{-5})) \end{aligned}$$

가 된다. 그러므로

$$\begin{aligned} AMASE(h_p) &= AMASE(h_a) + AMASE'(h_*)(h_a - h_p) \\ &= O(h_a^3(g^2 + n^{-1}g^{-5})) + O(h_a^2)(h_a - h_p) \end{aligned}$$

가 된다. 여기에서 h_* 는 h_a 와 h_p 사이의 값이다. 정리 1과 참고 1의 내용을 이용하면 증명은 쉽게 된다.

참고 2. (9)식의 h_p 에서 $\hat{\theta}_{2,2}(g)$ 를 역시 모르는 값이기 때문에 실제자료에서 바로 사용할 수가 없다. 그러므로 h_p 를 추정하는 것에 대해서 생각해 볼수가 있는데 그중에서 다음과 같은 \hat{h}_p 를 고려해보자.

$$\hat{h}_p = \left\{ \frac{(d-c)\hat{\sigma}^2 R(K)}{\hat{\theta}_{2,2}(\hat{g})\mu_{0,2}^2} \right\}^{1/5} n^{-1/5} \quad (10)$$

여기에서 \hat{g} 는 평활량인데

$$\widehat{MSE}\{\hat{\theta}_{2,2}(g)\} = \{n^{-1}g^{-5}R(L_2)\{\hat{\theta}_{0,0}(\hat{a}_1) + (d-c)\hat{\sigma}^2\} + 2g^2\mu_{2,4}\hat{\theta}_{2,4}(\hat{a}_2)/4!\}^2 + \hat{C}(g,2,2)$$

를 최소화하는 값이다. 여기에서

$$\hat{C}(g,k,l) = n^{-2}g^{-2k-2l-1}[(d-c)\hat{\sigma}^2 + 2\hat{\sigma}^2\hat{\theta}_{0,0}(\hat{b})] \left[\int \left\{ \int L_k(y)L_l(y+z)dy \right\}^2 dz \right]$$

이고 \hat{a}_1 과 \hat{a}_2 는 평활량인데

$$\widehat{MSE}\{\hat{\theta}_{0,0}(a_1)\} = \{n^{-1}a_1^{-1}R(L_0)\{\hat{\theta}_{0,0}(\hat{b}) + (d-c)\hat{\sigma}^2\} + a_1^2\mu_{0,2}\hat{\theta}_{0,2}(\hat{b})\}^2 + \hat{C}(a_1,0,0)$$

$$\begin{aligned} \widehat{MSE}\{\hat{\theta}_{2,4}(a_2)\} &= \{n^{-1}a_2^{-7}\{\hat{\theta}_{0,0}(\hat{b}) + (d-c)\hat{\sigma}^2\} \int L_2L_4 \\ &\quad + a_2^2\{\hat{\theta}_{2,6}(\hat{b})\mu_{4,6}/6! + \hat{\theta}_{4,4}(\hat{b})\mu_{2,4}/4!\}\}^2 + \hat{C}(a_2,2,4) \end{aligned}$$

을 각각 최소화하는 값이다. 그리고 \hat{b} 는 예비평활량(pilot bandwidth)이다.

참고 3. 교차타당성 평활량 \hat{h}_{cv} 의 상대적 수렴속도 ($\hat{h}_{cv}/h_a - 1$)는 $O(n^{-1/10})$ (Härdle, Hall and Marron(1988))로 삽입평활량의 수렴속도보다 상당히 느리다.

참고 2에서 예비평활량 \hat{b} 에 교차타당성 평활량 \hat{h}_{cv} 을 사용했을 때 \hat{h}_p 의 상대적 수렴속도는 $O(n^{-2/7})$ 로 변화가 없다는 것을 정리 1과 정리 2의 내용으로부터 쉽게 알 수가 있다.

3. 소표본 모의실험

이 절에서는 모의실험을 통하여 삽입방법이 교차타당성 방법보다 우수함을 보이려고 한다. 이 실험을 위하여 3개의 회귀함수, $m(x) = \sin^3(2\pi x^3)$, $2 - 5x + 5\exp(-(x-0.5)^2/0.01)$ 그리고 $2\sin(4\pi x)$ 를 선택하였다. 그리고 $x_i = (i-0.5)/n$ 에서 Y_i 를 관측하였다. 여기에서 사용되는 $m(x)$ 중 $\sin^3(2\pi x^3)$ 은 Härdle, Hall and Marron(1992), $2 - 5x + 5\exp(-(x-0.5)^2/0.01)$ 은 Gasser, Kneip and Köhler(1991)에서 인용하였고 변동이 큰 회귀함수의 비교를 위하여 $2\sin(4\pi x)$ 를 사용하였다. Gasser, Müller and Mammitsch(1985)의 연구 결과를 이용하여 다음과 같은 최적커널함수를 사용하였다.

$$K(x) = \frac{3}{4}(1-x^2)I_{(-1 < x < 1)},$$

$$L_2(x) = \frac{105}{16}(-5x^4 + 6x^2 - 1)I_{(-1 < x < 1)},$$

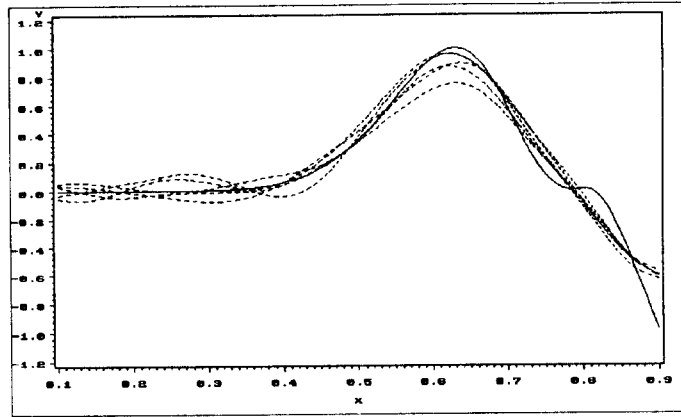


그림 1. $m(x) = \sin^3(2\pi x^3)$, $\sigma = 0.05$, $n = 50$ 일 때 \hat{h}_ω 를 이용한 추정량.
 $m(x)$ 는 실선으로 $\hat{m}_h(x)$ 는 점선으로 표시되어 있음.

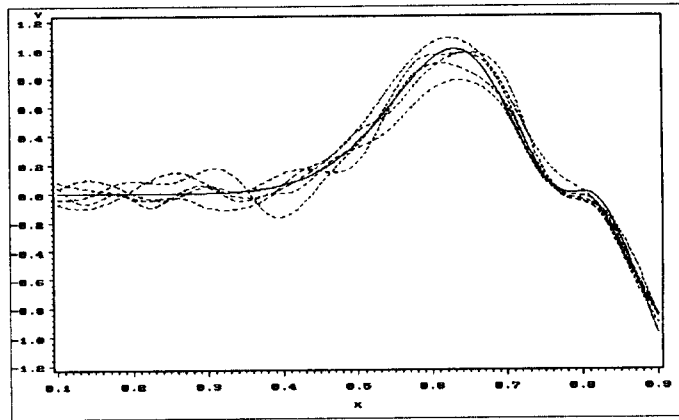


그림 2. $m(x) = \sin^3(2\pi x^3)$, $\sigma = 0.05$, $n = 50$ 일 때 \hat{h}_p 를 이용한 추정량.
 $m(x)$ 는 실선으로 $\hat{m}_h(x)$ 는 점선으로 표시되어 있음.

$$L_4(x) = \frac{10395}{32} (-21x^6 + 35x^4 - 15x^2 + 1) I_{(-1 < x < 1)}$$

그리고

$$L_6(x) = 39590.33 (-85.8x^8 + 184.8x^6 - 126x^4 + 28x^2 - 1) I_{(-1 < x < 1)}$$

이다.

크기가 50이고 오차항 e_i 의 분포가 $N(0, 0.05^2)$ 인 5개의 유사표본으로부터 $m(x)$ 를 추정하여 그림으로 나타내었다. 그림 1과 2는 $m(x) = \sin^3(2\pi x^3)$ 일 때 그림 3과 4는 $m(x) = 2 - 5x + 5\exp(-(x-0.5)^2/0.01)$ 일 때 \hat{h}_ω 와 \hat{h}_p 로 추정된 $\hat{m}_h(x)$ 를 그린 것이다. 그림 1과 2

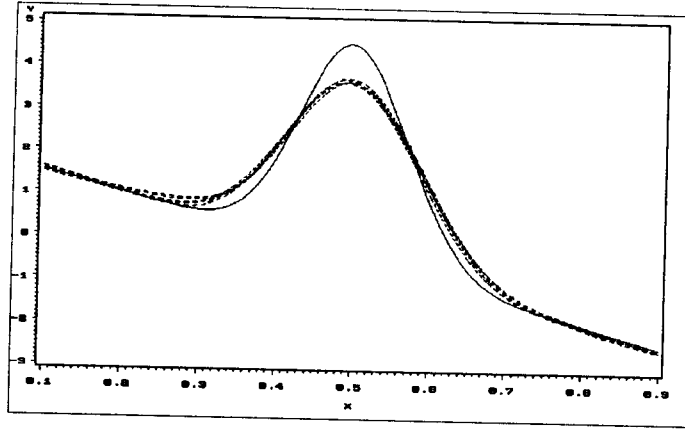


그림 3. $m(x) = 2 - 5x + 5\exp(-(x-0.5)^2/0.01)$, $\sigma = 0.05$, $n = 50$ 일 때 h_ω 를 이용한 추정량. $m(x)$ 는 실선으로 $\hat{m}_{\hat{h}_\omega}(x)$ 는 점선으로 표시되어 있음.

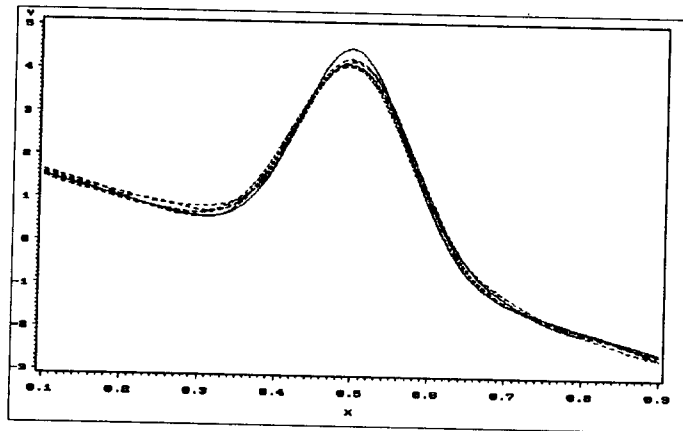


그림 4. $m(x) = 2 - 5x + 5\exp(-(x-0.5)^2/0.01)$, $\sigma = 0.05$, $n = 50$ 일 때 \hat{h}_p 를 이용한 추정량. $m(x)$ 는 실선으로 $\hat{m}_{\hat{h}_p}(x)$ 는 점선으로 표시되어 있음.

를 보면 전체적으로 $\hat{m}_{\hat{h}_p}(x)$ 는 $\hat{m}_{\hat{h}_\omega}(x)$ 보다 변동이 크다. 그러나 $\hat{m}_{\hat{h}_\omega}(x)$ 는 $x=0.8$ 근방의 어깨부분을 거의 추정을 못할 뿐 아니라 0.9근방에서도 아주 큰 편차를 가진다. 반면 $\hat{m}_{\hat{h}_p}(x)$ 는 0.8 근방에서의 추정이 좋을 뿐 아니라 0.9근방에서의 편차도 아주 작다.

그림 3과 4로부터 $\hat{m}_{\hat{h}_p}(x)$ 와 $\hat{m}_{\hat{h}_\omega}(x)$ 는 모두 다 $m(x)$ 를 잘 추정하고 있다는 것을 알 수가 있다. 그러나 삽입방법이 봉우리 부분($x=0.5$ 근방)과 어깨부분($x=0.35, 0.65$ 근방)을 더욱더 정확하게 추정함을 알 수가 있다. 이상의 비교를 통하여 교차타당성방법은 과평활(oversmoothing)

하는 경향이 있어 삽입방법보다 수행능력이 뒤떨어진다는 것을 알 수가 있다.

좀 더 나은 비교를 위하여 $MASE(h)$ 의 추정값, 즉

$$M\widehat{ASE}(h) = \frac{1}{rep} \sum \frac{1}{n} \sum_i \{\widehat{m}_h(x_i) - m(x_i)\}^2$$

를 비교하여 본다. 여기에서 rep 는 유사표본의 개수 즉 반복수를 나타내는데 본 실험에서는 200을 사용하였다. 이 실험을 위하여 표본의 크기 (n)는 25, 50, 그리고 오차항 e_i 의 분포는 $N(0, \sigma^2)$, $\sigma^2 = 0.05^2, 0.1^2$ 을 사용하였다. 여기에 대한 결과가 표 1에서 표 3까지에 나와 있다. 이들 표에는 $M\widehat{ASE}(h)$ 뿐만 아니라 ASE 값의 표준편차 ($S.D(ASE)$), 추정된 평활량의 평균치 ($AVE(\widehat{h})$)와 표준편차 ($S.D(\widehat{h})$) 그리고 (4)식의 $MASE(h)$ 를 최소화하는 h_{MASE} 의 값이 나타나 있다.

일반적으로 삽입방법이 교차타당성방법보다 더 우수함을 표로부터 알 수가 있다. 특히 $m(x) = 2 - 5x + 5\exp(-(x-0.5)^2/0.01)$ 에서 삽입방법의 우수성이 돋보인다. 그리고 변동이 큰 회귀함수의 추정에서도 좋은 결과를 보인다. 그렇지만 $M\widehat{ASE}(h)$ 의 값은 $\sin^3(2\pi x^3)$ 에서 가장 작게 나타난다.

전체적으로 볼 때 \widehat{h}_{cv} 값은 최적평활량 h_{MASE} 보다 아주 큰 값을 나타낸다. 이러한 결과는 위의 그림을 통해서 혹은 밀도함수의 추정에서 예상할 수가 있는 결과이다. 이러한 과평활로 인해서 $M\widehat{ASE}(h)$ 의 값이 크게 나온다. 그러나 추정된 평활량의 표본변동(sample variability, $S.D(\widehat{h})$)은 예상치 못할 정도로 작게 나온다. 이는 우리가 일반적으로 알고 있는 바와는 상반된 결과인데 이것에 대한 좀더 깊은 연구가 필요한 것으로 보인다. 또한 이 실험의 결과로 알 수 있는 것은 \widehat{h}_{cv} 값은 σ 의 값에 크게 영향을 받지 않는다는 것이다. \widehat{h}_p 값도 대체적으로 h_{MASE} 보다 큰 값을 가지지만 \widehat{h}_{cv} 보다 편의가 훨씬 적다. 그런데 ASE 의 변동은 교차타당성 방법과 비슷하게 나타난다.

이상의 결과를 종합할 때 삽입방법이 교차타당성 방법보다 상당히 만족스럽다고 결론을 내릴 수가 있다. 그러나 삽입방법의 편의도 상당히 큰 것으로 나타난다. 이를 줄이는 방법에 관한 연구가 계속 되었으면 좋겠다.

표 1. $m(x) = \sin^3(2\pi x)$ 일 때 모의실험 결과

σ	0.05				0.1			
	25		50		25		50	
n								
방법	CV	PI	CV	PI	CV	PI	CV	PI
$M\widehat{ASE}(\times 10^2)$	1.048	0.941	0.269	0.168	1.191	1.100	0.386	0.304
$S.D(ASE)(\times 10^3)$	1.005	1.095	0.426	0.422	2.185	2.354	1.185	0.979
$AVE(\hat{h})(\times 10)$	1.002	0.954	0.730	0.593	1.000	0.956	0.750	0.630
$S.D(\hat{h})(\times 10^3)$	1.218	2.403	1.658	1.567	2.622	5.385	5.792	4.105
h_{MASE}	0.03046		0.03473		0.04932		0.04721	

표 2. $m(x) = 2 - 5x + 5\exp(-(x-0.5)^2/0.01)$ 일 때 모의실험 결과

σ	0.05				0.1			
	25		50		25		50	
n								
방법	CV	PI	CV	PI	CV	PI	CV	PI
$M\widehat{ASE}(\times 10^2)$	7.873	5.012	4.739	0.659	8.074	5.298	4.837	0.884
$S.D(ASE)(\times 10^3)$	3.477	4.799	2.084	0.966	7.057	10.64	4.216	2.227
$AVE(\hat{h})(\times 10)$	1.045	0.901	0.902	0.537	1.045	0.903	0.903	0.547
$S.D(\hat{h})(\times 10^3)$	0.077	1.397	0.095	0.141	0.163	3.301	0.163	0.407
h_{MASE}	0.03015		0.024636		0.03019		0.03465	

표 3. $m(x) = 2\sin(4\pi x)$ 일 때 모의실험 결과

σ	0.05				0.1			
	25		50		25		50	
n								
방법	CV	PI	CV	PI	CV	PI	CV	PI
$M\widehat{ASE}(\times 10^2)$	4.155	2.777	1.057	0.593	4.320	2.962	1.149	0.813
$S.D(ASE)(\times 10^3)$	3.335	2.821	1.321	1.291	6.804	5.747	2.665	2.830
$AVE(\hat{h})(\times 10)$	1.024	0.865	0.731	0.572	1.024	0.870	0.732	0.586
$S.D(\hat{h})(\times 10^3)$	0.026	0.270	0.077	0.609	0.129	0.819	0.056	3.137
h_{MASE}	0.05583		0.03391		0.05726		0.04834	

참고문헌

- [1] Chu, C. K., and Marron, J. S. (1991), Choosing a kernel regression estimator (with comment), *Statistical Science*, Vol. 6, 404-436.
- [2] Fan, J. (1992), Design-adaptive nonparametric regression, *Journal of the American Statistical Association*, Vol. 87, 998-1004.
- [3] Gasser, T., Kneip, A., and Köhler, W. (1991), A flexible and fast method for automatic smoothing, *Journal of the American Statistical Association*, Vol. 86, 643-652.
- [4] Gasser, T., and Müller, H. G. (1979), "Kernel estimation of regression function" in *Smoothing Techniques for Curve Estimation*, eds. T. Gasser and M. Rosenblatt, Berlin:Springer-Verlag, pp. 23-68.
- [5] Gasser, T., and Müller, H. G. (1984), Estimating regression function and their derivatives by the kernel method, *Scandinavian Journal of Statistics*, Vol. 11, 171-185.
- [6] Gasser, T., Müller, H. G., and Mammitzsch, V. (1985), Kernels for nonparametric curve estimation, *Journal of the Royal Statistical Society, Ser. B*, Vol. 47, 238-252.
- [7] Gasser, T., Sroka, L., and Jennen-Steinmetz, C. (1986), Residual variance and residual pattern in nonlinear regression, *Biometrika*, Vol. 73, 625-633.
- [8] Härdle, W., Hall, P., and Marron, J. S. (1988), How far are automatically chosen regression parameters from their optimum? (with discussion), *Journal of the American Statistical Association*, Vol. 83, 86-99.
- [9] Härdle, W., Hall, P., and Marron, J. S. (1992), Regression smoothing parameters that are not far from their optimum, *Journal of the American Statistical Association*, Vol. 87, 227-233.
- [10] Jones, M. C., Davies, S. J., and Park, B.U. (1994), Versions of kernel-type regression estimators, *Journal of the American Statistical Association*, Vol. 89, 825-832.
- [11] Nadaraya, E. A. (1964), On estimating regression, *Theory of probability and Its Application*, Vol. 9, 141-142.
- [12] Priestley, M. B., and Chao, M. T. (1972), Non-parametric function fitting, *Journal of the Royal Statistical Society, Ser. B*, Vol. 34, 385-392.
- [13] Watson, G. S. (1964), Smooth regression analysis, *Sankhya*, Ser. A. Vol. 26, 359-372.