Journal of Statistical Theory & Methods
1998, Vol. 9, No. 1, pp. 19 ~ 27

커널 회귀함수 추정에서 점근최적인 평활량의 선택에 관한 연구 [1]

석 경 하 [2], 김 대 학 [3]

요약

본 논문에서는 커널 회귀함수의 추정방법에서 최적수렴율 $n^{-1/2}$을 가지는 평활량을 선택하는 방법에 대한 연구를 고려하였다. 이러한 평활량의 선택을 위하여 먼저 평활량의 수행측도인 기대평균제곱오차의 근사값을 4차항까지 테일러 급수전개를 하고 그 전개식을 최소화하는 평활량을 고려하였다. 이때 이 평활량이 포함하고 있는 미지의 범함수를 높은 차수의 커널함수를 이용하여 더욱 정확히 추정할수 있음을 제안한다. 또한 이렇게 구한 평활량과 최적 평활량과의 상대적 수렴율이 $n^{-1/2}$가 됨을 보였다.

주제어:커널회귀함수, 평활량, 최적수렴, 기대평균제곱오차, 테일러 전개, 범함수

1. 서론

여러 가지 비모수적 회귀함수의 추정방법중에서 커널회귀함수의 추정법이 최근에 많이 사용되고 있고 이 분야에 대한 연구도 많이 이루어지고 있다. 회귀함수의 커널추정방법중에서 대표적인 것으로는 Nadaraya(1964), Watson(1964), Priestley & Chao(1972) Gasser & Müller(1979,1984), Fan(1992), Fan & Gijbels(1995) 그리고 Park *et al.* (1997)등이 있는데 그 중에서 국소가중 회귀법(locally weighted regression)이 가장 우수하다는 평가를 받고 있다. Cleveland & Loader(1997) 에는 국소가중 회귀법에 관한 많은 내용이 설명되어있다.

회귀함수의 커널추정에서는 평활량(smoothing parameter, bandwidth)의 선택이 아주 중요한 문제이다. 그래서 평활량의 선택에 관해 연구가 많이 이루어지고 있다. 특히 Härdle *et al.*(1988)는 교차타당성 방법(cross-validation)으로 선택된 평활량($\hat{h}_{cv}$)는 최적평활량(h_{opt})으로의 상대적 수렴속도(relative convergence rate, $\hat{h}_{cv}/h_{opt} - 1$)가 $n^{-1/10}$로 상당히 느리다는 것을 보였다. Härdle *et al.*(1992)에서는 이른바 이중평활법(*double smoothing*)기법을 이용

---

[1]이 논문은 1996년 인제연구 장학재단의 지원을 받아 연구되었음
[2]621-749 경남 김해시 인제대학교 응용통계학과
[3]712-702 경북 경산시 하양읍 대구효성가톨릭대학교 정보통계학과

하여 평활량을 선택하는 방법에 대하여 연구를 하였다. 이 방법에 의한 평활량은 여러 가지의 적당한 조건하에서는 최적의 수렴율을 가질 수 있다는 것을 보였다.

본 논문에서는 Hall *et al*(1991)에 의한 결과로서 커널확률밀도함수 추정법에서 사용되는 여러 가지의 평활량 선택법 가운데 가장 이해가 쉽고 널리 쓰이는 소위 삽입방법(plug-in method)을 커널회귀함수의 추정에 응용하여 최적수렴율을 가지는 평활량의 선택방법을 제안하고자 한다. 이러한 평활량을 얻기 위하여 다음과 같은 절차를 고려하였다.

- 기대평균제곱오차 ($MASE$, Mean Averaged Squared Error)의 근사값을 테일러 급수 전개할때 통상적인 것 보다 2항 더 많이(4번째 항까지)하여 이 식을 최소화하는 평활량을 구하고
- 이 평활량이 포함하고 있는 미지의 범함수를 추정할 때 큰 차수의 커널함수를 이용하여 정확하게 추정한다.

그 결과 이렇게 얻어진 평활량이 최적의 수렴율을 가짐을 증명하였다.

2. 삽입방법

반응치 $Y_1, \cdots, Y_n$ 에 대한 회귀모형

$$Y_i = m(x_i) + \epsilon_i, \qquad i = 1, \cdots, n$$

에 대한 커널회귀함수 $m(\cdot)$ 의 추정에 있어서 특히 평활량의 선택에 관한 연구가 본 연구의 주된 관심이므로 여러 가지 계산의 간편성을 위하여 계획점(design point)이 단위간격에서 등간격이고 고정(equally spaced on the unit interval and fixed *i.e.* $x_i = (i - 0.5)/n$, $i = 1, \cdots, n$)되어 있다고 가정을 한다. 그리고 여기에서 ϵ_i는 서로 독립이고 평균이 0 분산이 σ^2인 확률변수이다. 그리고 $m(\cdot)$에 관한 여러 가지 커널추정법중에서 Priest & Chao(1972)의 추정량

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

을 고려하였다. 여기에서 $K(\cdot)$는 커널함수이고 h는 평활량이다. 커널회귀함수의 추정량 $\hat{m}_h(x)$의 수행능력은 커널함수와 평활량을 어떻게 선택하느냐에 달려 있다. 특히 평활량 h가 결정적인 역할을 한다는 것은 잘 알려진 사실이다.

통상적으로 커널함수 K는 $(-1,1)$에서 정의된 연속이고 대칭인 함수이며 또한 차수가 p, 즉

$$\int_{-1}^{1} z^j K(z) = \begin{cases} 1, & j = 0, \\ 0, & j = 1, \cdots, p-1, \\ \neq 0, & j = p, \end{cases} \tag{2}$$

를 가정한다. 최적인 커널함수의 선택법에 대한 연구로는 Gasser *et al*(1985)이 있다. 물론 큰 차수의 커널함수를 이용하면 추정량의 편의를 줄일 수 있지만 이는 본 연구의 관심사가 아니므로 가장 이해가 쉬운 차수가 2(second order, $p=2$)인 커널함수를 사용하겠다.

$m(x)$의 커널추정량 $\hat{m}_h(x)$의 수행측도(performance measure)로써 기대평균제곱오차 ($MASE$), 즉

$$MASE(h) = \frac{1}{n}\sum_{i=1}^{n} E\{\hat{m}_h(x_i) - m(x_i)\}^2 \tag{3}$$

를 고려해보자. $MASE(h)$의 근사치 $AMASE(h)$ (Asymptotic MASE) 는

$$AMASE(h) = n^{-1}h^{-1}\sigma^2 R(K) + h^4\mu_2^2\theta_{2,2}/4 \tag{4}$$

가 된다. 여기에서 $R(f) = \int f^2(x)dx$, $\mu_p = \int x^p K(x)dx$ 이고 $\theta_{k,l} = \int m^{(k)}(x)m^{(l)}(x)dx$이다. 또 $AMASE(h)$를 최소화하는 평활량 h_a는

$$h_a = \left\{\frac{\sigma^2 R(K)}{\mu_2^2\theta_{2,2}}\right\}^{1/5} n^{-1/5} \tag{5}$$

가 됨은 Härdle *et al*(1992)에 증명되어 있다. 그리고 식(4)보다 2항 더 전개한 식 $AMASE^*(h)$는

$$AMASE^*(h) = n^{-1}h^{-1}\sigma^2 R(K) + h^4\mu_2^2\theta_{2,2}/4 + h^6\mu_2\mu_4\theta_{2,4}/24 - n^{-1}R(m) \tag{6}$$

가 됨을 알 수가 있다. 식(6)의 $AMASE^*(h)$를 최소화시키는 h는 다음의 관계식을 만족한다.

$$h = \left\{\frac{\sigma^2 R(K)}{\mu_2^2\theta_{2,2} + h^2\mu_2\mu_4\theta_{2,4}/4}\right\}^{1/5} n^{-1/5} \tag{7}$$

이 식을 테일러급수 전개한 후 다시 정리를 하면 식(6)을 최소화하는 평활량 h_b는

$$h_b = h_a - \frac{h_a^3\mu_4\theta_{2,4}}{20\mu_2\theta_{2,2}} \tag{8}$$

로 쓸 수가 있다. h_b와 $MASE(h)$를 최소화하는 최적평활량 h_{opt}와는 다음의 관계가 성립한다.

$$(h_b - h_{opt})/h_{opt} = O(n^{-1/2}) \tag{9}$$

식 (6) 에서 식 (9) 까지의 관계는 Hall *et al* (1991)에 잘 나타나 있다. 이러한 관계가 있기 때문에 $(\hat{h}_b - h_{opt})/h_{opt} = O(n^{-1/2})$ 를 만족하는 h_b의 추정량 $\hat{h}_b$를 구할 수가 있다. 이 추정량을 구하는 방법 중에서 h_b가 포함하는 모르는 부분, 즉, $\theta_{2,2}$, $\theta_{2,4}$, σ^2, 의 추정치를 대입하여 h_b의 추정량 $\hat{h}_b$을 구하는 것이 바로 삽입방법이다. 다음의 3절에서는 이러한 추정치를 구하는 방법에 대하여 알아보고 그 추정치의 상대적 수렴율이 $n^{-1/2}$가 됨을 증명한다.

3. 점근 최적 평활량의 선택법

먼저 식(5)와 (8)에서 모르는 부분 $\theta_{2,2}$,$\theta_{2,4}$, σ^2을 추정하여야 한다. 그런데 σ^2의 $n^{1/2}$일치 추정량 $\hat{\sigma}^2$은 Gasser *et al*(1986)에서 찾아 볼 수가 있다. 그러므로 $\theta_{2,2}$와 $\theta_{2,4}$의 추정이 중요한 문제가 된다. 지금부터는 이 문제를 해결하여 보도록 한다. $\theta_{k,l}$에 대한 커널추정량으로 가장 자연스럽게 생각해볼 수 있는 추정량은

$$\hat{\theta}_{k,l}(r,s,g) = \int \hat{m}_{g,r}^{(k)}(x)\hat{m}_{g,s}^{(l)}(x)dx, \tag{10}$$

이다. 여기에서

$$\hat{m}_{g,r}^{(\nu)}(x) = \frac{1}{ng^{\nu+1}}\sum_{i=1}^{n} L_{\nu,r}(\frac{x-x_i}{g})Y_i$$

이고, g는 평활량, $L_{\nu,r}$는 커널함수인데 (1)식의 평활량 h와 커널함수 K와는 별개의 것이다. 커널함수 $L_{\nu,r}$는 $(-1,1)$에서 정의된 연속이고 대칭인 함수이며 차수가 $\nu+r$, 즉,

$$\int_{-1}^{1} z^j L_{\nu,r}(z)dz = \begin{cases} 1, & j = 0,1,\cdots,\nu-1,\nu+1,\cdots,\nu+r-1, \\ \nu!, & j=\nu, \\ \mu_{\nu,r}\neq 0, & j=\nu+r, \end{cases} \tag{11}$$

임을 가정하자. 회귀함수 $m(\cdot)$를 추정할 때 평활량 h가 중요한 역할을 하는 것과 마찬가지로 $\theta_{k,l}$을 추정할 때도 평활량 g가 중요한 역할을 한다. 그러면 $\hat{\theta}_{k,l}(r,s;g)$의 평균제곱오차($MSE$)를 최소화하는 평활량 g를 구하여 보도록 하자. 식(10)의 $\hat{\theta}_{k,l}$은 간단한 계산에 의하여 다음과 같은 식으로 근사할 수 있다.

$$\begin{aligned} \hat{\theta}_{k,l} &\simeq n^{-2}g^{-k-l-2}\Big\{\sum_i Y_i^2 \int L_{k,r}(\frac{x-x_i}{g})L_{l,s}(\frac{x-x_i}{g})dx \\ &+ \sum_i\sum_{\neq j} Y_iY_j \int L_{k,r}(\frac{x-x_i}{g})L_{l,s}(\frac{x-x_i}{g})dx\Big\} \\ &\simeq n^{-2}g^{-k-l-1}\Big\{\sum_i Y_i^2 \int L_{k,r}(x)L_{l,s}(x)dx + \sum_i\sum_{\neq j} Y_iY_jL_{k,r} * L_{l,s}(\frac{x_i-x_j}{g})\Big\} \end{aligned}$$

여기에서 $*$는 함수의 합성(convolution)을 나타낸다. 이 근사식을 이용하면 $\hat{\theta}_{k,l}$의 평균제곱오차의 근사식을 쉽게 계산할 수 있다.

정리1. "m은 $(0,1)$에서 $max\{max(k+r,l+s),k+l\}$번 미분가능하다" 라는 가정 하에서 $\hat{\theta}_{k,l}(r,s;g)$의 기대치와 분산의 근사치는 아래와 같다.

$$E\{\hat{\theta}_{k,l}(r,s;g)\} \simeq \theta_{k,l} + c_1 n^{-1} g^{-k-l-1} + c_2 g^r + c_3 g^s$$
$$var\{\hat{\theta}_{k,l}(r,s;g)\} \simeq c_4 n^{-1} + c_5 n^{-2} g^{-2k-2l-1}$$

여기에서

$$c_1 = (\theta_{0,0} + \sigma^2) \int L_{k,r}(x) L_{l,s}(x) dx$$
$$c_2 = \theta_{k+r,l}\mu_{k,r,k+r}/(k+r)!$$
$$c_3 = \theta_{l+s,k}\mu_{l,s,l+s}/(l+s)!$$
$$c_4 = 4\theta_{k+l,k+l}\sigma^2$$
$$c_5 = (\sigma^4 + 2\sigma^2\theta_{0,0})\{\int [\int L_{k,r}(y) L_{l,s}(y+z) dy]^2 dz\}$$

이다. 그리고 $\mu_{a,b,c} = \int x^c L_{a,b}(x) dx$를 나타낸다.

(증명) 먼저 기대치에 대한 증명을 해본다.

$$E\{\hat{\theta}_{k,l}(r,s;g)\} \simeq n^{-1} g^{-k-l-1} (\theta_{0,0} + \sigma^2) \int L_{k,r} L_{l,s}$$
$$+ g^{-k-l-2} \int\int\int m(x) m(y) L_{k,r}(\frac{t-y}{g}) L_{l,s}(\frac{t-x}{g}) dx dy dt$$
$$\simeq c_1 n^{-1} g^{-k-l-1} + g^{-k-l} \int\int\int \{\frac{g^k t_1^k}{k!} m^{(k)}(t) + \frac{g^{k+r} t_1^{k+r}}{(k+r)!} m^{(k+r)}(t)\} L_{k,r}(t_1)$$
$$\{\frac{g^l t_2^l}{l!} m^{(l)}(t) + \frac{g^{l+s} t_2^{l+s}}{(l+s)!} m^{(l+s)}(t)\} L_{l,s}(t_2) dt dt_1 dt_2$$

이제 분산에 관한 것을 증명하도록 한다. $var\{\hat{\theta}_{k,l}(r,s;g)\} \simeq A_1 + A_2 + A_3 + A_4$ 로 쓰자. $B_{ij} = \int L_{k,r}(\frac{x-x_i}{g}) L_{l,s}(\frac{x-x_j}{g}) dx$ 라고 하면

$$A_1 \simeq 2n^{-4} g^{-2(k+l+2)} \sum_i \sum_{\neq j} cov(Y_i^2, Y_i Y_j) B_{ii} B_{ij}$$
$$A_2 \simeq 4n^{-4} g^{-2(k+l+2)} \sum_i \sum_{\neq j} \sum_{\neq i'} cov(Y_i Y_j, Y_{i'} Y_j) B_{ij} B_{i'j}$$
$$A_3 \simeq n^{-4} g^{-2(k+l+2)} \sum_i var(Y_i^2) B_{ii}^2$$
$$A_4 \simeq n^{-4} g^{-2(k+l+2)} \sum_i \sum_{\neq j} var(Y_i Y_j) B_{ij}^2$$

로 쓸 수 있다. 그리고 각각은 다음과 같은 근사식으로 표현할 수 있다.

$$A_1 = 2n^{-4} g^{-2k-2l-3} \int L_{k,r}(x) L_{l,s}(x) dx \sum_i \sum_{\neq j} 2\{m(x_i) m(x_j)\} \sigma^2 B_{ij}$$

$$
\begin{aligned}
&= 4n^{-2}g^{-2k-2l-1}\sigma^2 \int L_{k,r}(x)L_{l,s}(x)dx \int\int\int m(x-gy)m(x-gz) \\
&\quad L_{k,r}(y)L_{l,s}(z)dxdydz \\
&= 4n^{-2}g^{-k-l-1}\sigma^2\theta_{k,l}\int L_{k,r}L_{l,s} + o(n^{-2}g^{-k-l-1})
\end{aligned}
$$

$$
\begin{aligned}
A_2 &= 4n^{-1}g^{-2(k+l+1)}\sigma^2 \int\int\int m(x)m(z)L_{k,r} * L_{l,s}(\frac{x-y}{g}) \\
&\quad L_{k,r} * L_{l,s}(\frac{y-z}{g})dxdydz \\
&= 4n^{-1}g^{-2(k+l)} \int\int\int\int\int m(y+g(x+u_1))m(y-g(z+u_2)) \\
&\quad L_{k,r}(x)L_{l,s}(u_1)L_{k,r}(z)L_{l,s}(u_2)dxdydzdu_1du_2 \\
&= 4n^{-1}\theta_{k+l,k+l}\sigma^2 + o(n^{-1}) \\
A_3 &= n^{-3}g^{-2(k+l+1)}var(Y_i^2)\{\int L_{k,r}L_{l,s}\}^2 \\
A_4 &= n^{-4}g^{-2k-2l-4}\sum_i\sum_{\neq j}\{\sigma^4+\sigma^2m^2(x_i)+\sigma^2m^2(x_j)\} \\
&\quad \{g\int L_{k,r}(y)L_{l.s}(y+\frac{x_i-x_j}{g})dy\}^2 \\
&= n^{-2}g^{-2k-2l-1}\{\sigma^4+2\sigma^2\theta_{0,0}\}\{\int[\int L_{k,r}(y)L_{l,s}(y+z)dy]^2dz\}
\end{aligned}
$$

그러므로 이상의 내용을 종합하면 증명은 쉽게 이루어진다.

참고 1. 우리는 위의 정리로부터 $MSE = c_1^2n^{-2}g^{-2k-2l-2} + c_2^2g^{2r} + c_3^2g^{2s} + 2c_1c_2n^{-1}g^{-k-l+r-1} + 2c_1c_3n^{-1}g^{-k-l+s-1} + 2c_2c_3g^{r+s} + c_4n^{-1} + c_5n^{-2}g^{-2k-2l-1}$ 가 됨을 알 수 있다. 여기에서 c_1, c_4 그리고 c_5는 항상 양수이지만 c_2 와 c_3의 부호는 알 수가 없다.

참고 2. $k = l = 2$,$r = s = 2$ 그리고 $c_2 > 0, c_3 > 0$ 인 경우에 MSE를 최소화하는 $\hat{g} = (n^{-1/7})$이고 이때 $MSE = O(n^{-4/7})$이 된다. 그리고 $c_2 > 0, c_3 > 0$가 아닌 경우에 MSE를 최소화하는 $\hat{g} = (n^{-1/7})$로 같지만 이때는 $MSE = O(n^{-5/7})$ 혹은 $O(n^{-4/7})$가 된다. 그리고 $k = l = 2$,$r = s = 6$ 인 경우에는 c_2 와 c_3의 부호와 관계없이 MSE를 최소화하는 $\hat{g} = (n^{-1/11})$이고 이때 $MSE = O(n^{-1})$이 된다.

위의 참고 1,2를 이용하여 다음과 같은 정리를 얻을 수 있다.

정리 2 정리1과 같은 가정을 하고 h_b의 모르는 부분,$\theta_{2,2}$와 $\theta_{2,4}$를 $\hat{\theta}_{2,2}(6,6;\hat{g}_1)$ 와 $\hat{\theta}_{2,4}(2,2;\hat{g}_2)$로 로 추정한 $\hat{h}_b$는

$$\frac{\hat{h}_b}{h_{opt}} - 1 = O(n^{-1/2})$$

를 만족한다. 여기에서 $\hat{g}_1$ 과 $\hat{g}_2$는 $MSE\{\hat{\theta}_{2,2}(6,6;g_1)\}$ 과 $MSE\{\hat{\theta}_{2,4}(2,2;g_2)\}$를 각각 최소화하는 g값들이다.

증명 $AMASE^*(h)$를 최소화하는 h_b가 $(h_b - h_{opt})/h_{opt} = O(n^{-1/2})$을 만족하며 $MSE\{\hat{\theta}_{2,2}(6,6;\hat{g}_1)\}$ $=O(n^{-1})$ 이고 $MSE\{\hat{\theta}_{2,4}(2,2;\hat{g}_2)\}$ $=O(n^{-4/9})$ 그리고 $\hat{g}_1 = O(n^{-1/11})$, $\hat{g}_2 = O(n^{-1/9})$이라는 사실을 이용하면 증명은 쉽게된다.

참고 3 위의 정리에서 $\hat{g}_1$과 $\hat{g}_2$를 계산하기 위해서는 또 다른 평활량이 필요하다. 그러나 이 단계에서 요구되는 평활량은 어떠한 것을 쓰더라도 정리 2의 결과에는 영향을 미치지 못한다. 그러므로 예비평활량(pilot bandwidth)으로서는 가장 쉽게 구할 수 있는 교차타당성 평활량(cross-validated bandwidth)을 사용하여도 무난할 것으로 생각된다.

4. 결론 및 향후과제들

최적수렴율을 가지는 평활량의 선택법을 제안하였다. 제안된 평활량이 근사적으로는 최적이지만 유한인 실제 자료에서 좋은 수행능력을 가지기 위해서는 해결해야될 과제가 있다.

앞의 절에서 볼수 있듯이 $\hat{h}_b$의 정확도는 $\theta_{2,2}$와 $\theta_{2,4}$를 얼마나 잘 추정하느냐에 달려있다. 그런데 밀도함수에서와는 다르게 $\theta_{2,2}$ 와 $\theta_{2,4}$가 아주 큰 값을 가지기 때문에 이를 추정할 때 아주 신중해야 하겠다. 하나의 예로 변동이 그렇게 크지 않는 함수 $m(x) = sin^3(2\pi x)$의 $\theta_{2,2} \sim 15,000$로 아주 큰 값은 가진다. 그러므로 h_b값을 결정하는데 $\theta_{2,2}$와 $\theta_{2,4}$의 값이 크게 영향을 미친다. (참고로 $\int \phi^{(2)}(x)dx \sim 0.211$, ϕ는 $N(0,1)$의 $p.d.f$). 그러므로 $\theta_{k,l}$형태의 모수를 포함하지 않는 방법, 예를 들면 교차타당성 방법 혹은 wavelets을 이용할수 이용할 수 있을 것으로 생각된다.

참고문헌

1. Cleveland, W. S. & Loader, C. (1997). Smoothing by local regression; principles and methods (with discussion), *Computational Statistics*. (to appear)

2. Fan, j. (1992) Design-adaptive nonparasitic regression, *Journal of the American Statistical Association,* Vol. 87, 998-1004

3. Fan, J. & Gijbels, I. (1995). Adaptive order polynomial fitting; bandwidth robustification and bias reduction. *Journal of Computation & Graphical Statistics.* Vol. 4, 213-227

4. Hărdle, W. & Hall, P. & 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.

5. Hărdle, W. & Hall, P. & Marron, J. S. (1992). Regression smoothing parameters that are not far from their optimum, *Journal of the American Statistical Association,* Vol. 87, 227-233.

6. Gasser ,T. & Műller, H. G. (1979). "Kernel estimation of regression function" in *Smoothing Technique for Curve Estimation,* eds. Gasser & Rosenblatt, M., Berlin: Springer-Verlag, pp23-68.

7. Gasser ,T. & Műller, H. G. (1984). Estimating regression function and their derivatives by the kernel method, *Scandinavian Journal of Statistics,* Vol. 11, 171-185.

8. Hall, P. & Sheather, S. J. & Park, B. U. (1991). On optimal data based bandwidth selection in kernel density estimation, *Biometrika,* 78, 263-269

9. Nadaraya, E. A. (1964). On estimating regression, *Theory of Probability and Its Application,* Vol. 9, 141-142.

10. Park, B. U. & Kim, W. C. & Ruppert, D. & Jones, M. C. & Signorini, D. F. Kohn, R. (1997). Simple transformation technique for improved nonparametric regression, *Scandinavian Journal of Statistics,* Vol. 24, 145-163.

11. Priestley, M. & Chao, M. T. (1972). Non-parametric function fitting, *Journal of the Royal Statistical Society,* Ser. B, Vol. 34, 385-392.

12. Watson, G. S, (1964). Smooth regression analysis, *Sankhyā,* Ser. A. Vol. 26, 359-372

Asymptotic optimal bandwidth selection in kernel regression function estimation

Kyoungha Seog [4] · **Daehak Kim** [5]

Abstract

We considered the bandwidth selection method which has asymptotic optimal convergence rate $n^{-1/2}$ in kernel regression function estimation. For the proposed bandwidth selection, we considered Mean Averaged Squared Error as a performance criterion and its Taylor expansion to the fourth order. Then we estimate the bandwidth which minimizes the estimated approximate value of MASE. Finally we show the relative convergence rate between optimal bandwidth and proposed bandwidth.

[4]Department of applied statistics, Inje University

[5]Department of Statistical infomation, Catholic University of Taegu-Hyosung