Journal of Statistical
Data & Information Science Society
1998, Vol. 9, No. 2, 179 ~ 188

변수평활량을 이용한 커널회귀함수 추정 [1]

석경하 [2] · 정성석 [3] · 김대학 [4]

요약

커널형 회귀함수의 추정법 중에서 국소 다항회귀 추정법이 가장 우수한 것으로 알려져 있다. 국소다항회귀 추정법에서도 다른 종류의 커널추정량과 마찬가지로 평활량이 중요한 역할을 한다. 특히 회귀함수가 복잡한 구조를 가질 때 변수평활량(variable bandwidth)을 사용하는 것이 타당할 것이다. 본 연구에서는 완전자료기저(fully automatic, fully data-driven) 변수평활량 선택법을 제안한다. 이 선택법은 편향과 분산의 예비추정에 필요한 평활량을 교차타당성 방법으로 선택하여 MSE를 추정하고 그 값을 최소화하는 평활량을 택하는 것이다. 제안된 방법의 우수성을 모의실험을 통하여 확인하였다. 그리고 제안된 방법은 자료점이 성긴(sparse)부분에서 생길 수 있는 문제점 즉 $X'X$의 비정칙성(non-singularity)을 해결할 수 있는 방법이라는 데에도 큰 의 미가 있다.

주제어: 커널회귀함수추정, 변수평활량, 완전자료기저, 삽입방법, 교차타당성, 모의실험

1. 서론

여러 가지 커널형 회귀함수 추정법 중에서 국소 다항회귀 추정법(LPE, Local Polynomial Estimation)은 가장 우수한 것으로 평가되고 있다. 이 추정법은 아주 직관적이어서 이해하기가 쉽고 또한 사용하기도 쉽다. 또한 이추정법에서는 커널추정법의 단점으로 지적되는 경계점의 편향문제가 자동으로 해결된다. 그리고 자료점의 분포가 어떠하더라도 좋은 결과를 얻을 수가 있고 도함수의 추정에도 쉽게 적용될수 있는 좋은 성질을 가지고 있다. 이 추정법에 관한 대표적인 참고서로는 Fan, Giblels(1996)와 Simonoff(1996)등을 들 수가 있다.

다항식의 차수와 평활량은 LPE의 수행에 결정적인 영향을 주는데 이 두 모수 사이에는 큰 상관관계가 존재한다. 그러므로 어떤 회귀함수를 추정할 때 평활량을 고정시키고 다항식의 차수를 조정하는 방법이 있고 반대로 다항식의 차수를 고정하고 평활량을 조정하

[1]이논문은 1997년도 한국학술진흥재단의 공모과제 연구비에 의해 연구되었음
[2]경남 김해시 어방동 인제대학교 응용통계학과 부교수
[3]전북 전주시 덕진동 전북대학교 통계학과 조교수
[4]경북 경산시 하양읍 대구효성가톨릭대학교 정보통계학과 부교수

는 방법이 있다. 혹은 이 두 모수를 동시에 적당히 조정하는 방법이 있을 것이다. 본 연구에서는 다항식의 차수가 1 인 국소 선형회귀 추정법(LLE, Local Linear Estimation)의 평활량 선택법에 대해서 다루기로 한다.

평활량은 추정하고자 하는 전 구간에서 같은 값을 가지는 고정평활량(global bandwidth)과 추정점에 따라 변하는 변수평활량(variable bandwidth, local bandwidth)의 두 종류가 있다. 어떤 평활량을 사용하는 것이 더 나은 추정을 할 수 있을 것인가 하는 것은 추정하고자 하는 함수의 형태가 어떠한가, 혹은 계획점의 분포가 어떠한가 등에 영향을 받을 것이나 일반적으로는 변수평활량을 사용하는 것이 더 나을 것이라는 것은 직관적으로도 자명하다. 즉 추정하고자 하는 함수의 평평한 부분에서는 큰 평활량을 사용하고 변동이 심한 부분에서는 작은 평활량을 사용하여야 할 것이며 자료점이 성긴(sparse)부분에서는 조밀한 부분에서보다 상대적으로 큰 평활량을 사용한 추정량이 더 우수할 것이다.

LLE에서 고정평활량을 선택하는 방법 중에서 최근의 대표적인 방법으로는 삽입(Plug-in)방법을 이용한 Ruppert 등(1995)과 교차타당성 방법을 응용한 Hart 와 Yi (1998)의 연구 그리고 Fan 등(1996) 등이 있다. 그리고 변수평활량의 선택법에 관한 연구로는 Fan 과 Gijbels(1995) 와 Fan 등(1996)이 있다. 본 연구에서는 완전자료기저 변수평활량 선택법을 소개한다. 완전자료기저 변수평활량 선택법이란, 먼저 추정하고자 하는 점에서의 MSE의 추정량을 교차타당성 방법에 의한 평활량을 사용하여 구한다. 그 다음 이 추정량을 최소화하는 평활량을 구하는데 이것이 바로 본 연구에서 제안하는 방법이다. 이런 방법에 의해 구해진 변수평활량은 완전자료기저(fully automatic)이고 또한 자료점의 분포가 어떠하더라도 수정 없이 사용될 수 있다는 큰 장점을 가지고 있다. 그리고 이러한 방법이 Ruppert(1995)에 의한 방법보다 더 우수함을 모의실험을 통하여 알 수가 있었다. 그러나 제안된 평활량 선택법은 계산에 시간이 많이 걸린다는 단점을 가지고 있다.

제 2절에서는 LLE에 대해서 간단히 소개를 하고 제 3절에서는 하나의 변수 평활량 선택법을 소개를 한다. 그리고 제 4절에서는 간단한 모의실험을 통하여 제안된 방법의 우수성을 입증하였다.

2. 국소 선형회귀 추정법

어떤 모집단으로부터 이변량 랜덤표본 $(X_1, Y_1), \cdots, (X_n, Y_n)$이 주어졌을 때 두 변수 사이의 회귀관계식을 아래와 같이 쓸 수 있다.

$$Y_i = m(X_i) + \sigma(X_i)\epsilon_i,\ i = 1, \cdots, n. \tag{2.1}$$

여기에서 ϵ_i 는 $E(\epsilon_i) = 0,\ var(\epsilon_i) = 1$인 iid 확률변수이다. m에 대한 비모수적 추정법중에서 LPE가 가장 우수한 것으로 평가되고 있다. 만약 $m(x)$가 x_0점에서 $(p+1)$번째 도함수를 가진다면, 다음과 같이 p차 다항식으로 $m(x)$를 근사 시킬 수 있다.

$$m(x) = m(x_0) + m'(x_0)(x - x_0) + \cdots + m^{(p)}(x - x_0)^p/p! \tag{2.2}$$

이 사실을 이용하여 다음 식과 같은 국소 다항회귀 모형을 적합시킬 수 있다.

$$\min_{\beta}\sum_{i=1}^{n}\{Y_i-\sum_{j=1}^{p}\beta_j(X_i-x)^j\}K(\frac{X_i-x}{h}) \tag{2.3}$$

여기에서 $\beta=(\beta_0,\cdots,\beta_p)^T$, K는 음이 아닌 가중치 함수이고 h는 평활량이다. 이 추정법에 영향을 주는 것은 가중치 함수 K와 다항식의 차수 p 그리고 평활량 h인데 평활량이 가장 많은 영향을 끼친다. $\hat{\beta}=(\hat{\beta}_0,\cdots,\hat{\beta}_p)$ 가 (2.2)식을 만족하는 해라고 한다면 Taylor 전개식 (2.1)로 부터 $\nu!\hat{\beta}_\nu(x_0)$가 $m^{(\nu)}(x_0)$ 의 추정량이 된다는 것을 알 수가 있다. 문제를 좀 더 쉽게 이해하기 위해 그리고 표현을 간편하게 하기 위하여 위의 (2.3)식을 행렬형태로 표현하면

$$\min_{\beta}(\mathbf{y}-X\beta)^TW(h)(\mathbf{y}-X\beta) \tag{2.4}$$

이 된다. 여기에서 $\mathbf{y}=(Y_1,\cdots,Y_n)^T$이고

$$X=\begin{pmatrix}1 & (X_1-x_0) & \cdots & (X_1-x_0)^p\\ 1 & (X_2-x_0) & \cdots & (X_2-x_0)^p\\ \vdots & \vdots & \vdots & \vdots\\ 1 & (X_n-x_0) & \cdots & (X_n-x_0)^p\end{pmatrix}$$

$$W(h)=\mathrm{diag}\{K(\frac{X_i-x_0}{h})\}_{n\times n} \tag{2.5}$$

이다. (2.4)식을 만족하는 최소 제곱추정량은 아래와 같다.

$$\hat{\beta}=(X^TW(h)X)^{-1}X^TW(h)\mathbf{y} \tag{2.6}$$

이 추정량의 조건부 편의와 분산은 다음과 같이 주어진다.

$$\mathrm{bias}(\hat{\beta})=\mathrm{bias}(\hat{\beta}|X_1,\cdots,X_n)=(X^TW(h)X)^{-1}X^TW(h)\mathbf{r}$$

$$\mathrm{var}(\hat{\beta})=\mathrm{var}(\hat{\beta}|X_1,\cdots,X_n)=(X^TW(h)X)^{-1}(X^T\Sigma X)(X^TW(h)X)^{-1}, \tag{2.7}$$

여기에서

$$\mathbf{r}=\mathbf{m}-X\beta,$$

$$\Sigma=\mathrm{diag}\{K^2(\frac{X_i-x_0}{h})\sigma^2(X_i)\}_{n\times n},$$

그리고

$$\mathbf{m}=\{m(X_1),\cdots,m(X_n)\}^T$$

이다. 한편, 다항식의 차수 p가 1 인 경우에도 $\hat{m}$의 수행능력이 크게 떨어지지 않으므로(Fan 과 Gijbels (1995)) 문제의 간편성을 위하여 본 연구에서는 p 가 1인 국소 선형 회귀추정량(LLE)을 고려하도록 한다. 이 경우에 $\beta_0(x_0) = m(x_0)$의 추정량은

$$\hat{m}_l(x_0) = \frac{\sum_{i=1}^{n}\{S_{n,2} - (X_i - x_0)S_{n,1}\}K((X_i - x_0)/h)Y_i}{\sum_{i=1}^{n}\{S_{n,2} - (X_i - x_0)S_{n,1}\}K((X_i - x_0)/h)} \tag{2.8}$$

여기에서

$$S_{n,j} = \sum_{i=1}^{n}(X_i - x_0)^j K((X_i - x_0)/h), j = 0, 1, 2, \cdots$$

이다. 앞 절에서 잠깐 언급한 것처럼 (2.7)식의 추정량이 포함하는 평활량 h가 $\hat{m}_l$ 의 수행능력에 지대한 영향을 미친다. 다음절에서는 변수평활량을 선택하는 방법을 제안 한다.

3. 변수평활량 선택법

고정평활량을 선택하는 대표적인 방법은 Ruppert 등(1995)을 들 수가 있다. 이 방법은 평균 누적오차제곱($MISE$, Mean Integrated Squared Error)의 근사식을 최소화하는 h_{MISE}를 구하고 이 식이 포함하는 모르는 부분을 추정하여 삽입하는 삽입(plug-in)추정량 $\hat{h}_R$을 찾는 방법이다. 본 연구에서 다루고자 하는 변수평활량 선택법에 대한 대표적인 연구는 Fan 과 Gijbels(1995)을 들 수가 있다. 이 연구는 단순히 변수평활량의 선택방법에 관해서만 소개를 하였고, 이 방법의 이론적인 규명은 Fan 과 Huang(1998)에서 이루어졌다. 일반적으로 변수평활량을 선택하는 방법은 점근 평균오차제곱(asymptotic MSE)을 최소화하는 h_{opt}를 구하고 이 식이 포함하는 모르는 부분을 추정하여 대입하는 소위 점근 후 대입법으로만 생각을 하기가 쉬웠다. 그러나 Fan 과 Gijbels(1995)에서는 점근전 대치법(pre-asymptotic substitution)을 소개하였는데 이 방법은 편이(bias)와 분산(var)의 추정치를 구한 다음 MSE의 추정치를 최소화하는 평활량을 찾는 방법이다. 이 방법을 사용하는 이유는 점근 후 대치법에서는 h_{opt}가 f_X(계획점의 밀도함수)에 의존하는 문제를 해결하기 위하여 f_X를 추정하여 대입할 수도 있고 계획변환(design transformation)을 사용하여 해결할 수도 있겠지만 이 방법을 쓰면 f_X 항이 나타나지 않기 때문이다. (2.7)식의 bias의 추정량으로는

$$\widehat{\text{bias}} = (X^T W(h) X)^{-1} X^T W(h) \tau,$$

이때

$$\tau = \begin{pmatrix} \beta_2(X_1 - x_0)^2 + & \cdots & +\beta_{1+\alpha}(X_1 - x_0)^{1+\alpha} \\ \vdots & \cdots & \vdots \\ \beta_2(X_n - x_0)^2 + & \cdots & +\beta_{1+\alpha}(X_n - x_0)^{1+\alpha} \end{pmatrix}$$

을 생각할 수 있겠다. 여기에서 $a=4$이면 최적 수렴율을 가진 추정량을 찾을 수 있지만 계산상의 편의와 속도를 고려한다면 $a=2$도 충분하다(Fan 과 Gijbels(1995)).

본 연구에서 제안하는 평활량을 선택하는 절차는 다음과 같다.

1) 교차타당성(cross-validation) 함수

$$CV(g)=\sum_{i=1}^{n}\{y_i-\hat{m}_{-i}(x_i)\}^2$$

를 최소화하는 예비평활량(pilot bandwidth) $\hat{g}$를 구한다. 여기에서 $\hat{m}_{-i}$는 i번째 표본 (X_i,Y_i)를 제외한 $p=3$인 국소 다항회귀 추정량이다. 그러니까 이 추정량에 사용되는 계획행렬과 가중치 행렬은

$$X^*=\begin{pmatrix}1 & (X_1-x_i) & \cdots & (X_1-x_i)^3\\ 1 & (X_2-x_i) & \cdots & (X_2-x_i)^3\\ \vdots & \vdots & \vdots & \vdots\\ 1 & (X_n-x_i) & \cdots & (X_n-x_i)^3\end{pmatrix},$$

$$W(g)=\mathrm{diag}\{K(\frac{X_j-x_i}{g})\}_{j\neq i,(n-1)\times(n-1)}$$

이다.

2) 1)에서 구하여진 $\hat{g}$를 이용하여

$$\begin{aligned}\hat{\beta}^* &= (\hat{\beta}_0^*,\hat{\beta}_1^*,\hat{\beta}_2^*,\hat{\beta}_3^*)^T\\ &= (X^*TW^*(\hat{g})X^*)^{-1}X^*TW^*(\hat{g})X^*\end{aligned}$$

를 구한다. 그리고 이를 이용하여 (2.7)식의 bias의 추정량

$$\widehat{\mathrm{bias}}=(X^TW(h)X)^{-1}X^TW(h)\hat{\tau}$$

을 구한다. 여기에서 X 와 $W(h)$ 는 (2.5)식에서 $p=1$일 때의 계획행렬과 가중치 행렬이고

$$\hat{\tau}=\begin{pmatrix}\hat{\beta}_2(X_1-x_0)^2+\hat{\beta}_3(X_1-x_0)^3\\ \vdots\\ \hat{\beta}_2(X_n-x_0)^2+\hat{\beta}_3(X_n-x_0)^3\end{pmatrix}$$

이다.

3) $\hat{g}$를 이용하여 $\sigma^2(x_0)$의 추정량인 표준화 잔차 가중제곱합(normalized weighted residual sum of squares),

$$\hat{\sigma}^2(x_0,\hat{g})=\frac{(Y-X^*\hat{\beta}^*)^TW^*(\hat{g})(Y-X^*\hat{\beta}^*)}{\mathrm{tr}\{W^*(\hat{g})-(X^{*2}W^*(\hat{g})X^*)^{-1}X^*W^{*2}(\hat{g})X^*\}}$$

을 구하여 (2.7)식의 var의 추정량

$$\widehat{\text{var}}(\hat{\beta}) = (X^T W(h) X)^{-1} X^T \hat{\Sigma} X (X^T W(h) X)^{-1}$$

을 구한다. 여기에서

$$\hat{\Sigma} = \text{diag}\{K^2(\frac{X_i - x_0}{h})\hat{\sigma}^2(X_i)\}_{n \times n}$$

이다.

4) 2)와 3)의 절차에서 구해진 $\widehat{\text{bias}}$와 $\widehat{\text{var}}$을 이용하여 MSE의 추정량

$$\widehat{MSE}(x_0) = \{(1,0)\widehat{\text{bias}}\}^2 + (1,0)\widehat{\text{var}}(1,0)^T$$

를 최소화하는 $\hat{h}(x_0)$를 구한다.

다음절에서는 이러한 방법으로 구해진 평활량의 수행능력을 평가하기 위한 소표본 모의실험을 시행하였다.

4. 모의실험

본 논문에서 제안한 방법의 우수성을 입증하기 위하여 소표본 모의실험을 시행하였다. 등분산성 $\sigma^2(x_i) = \sigma^2$을 가정하고 다음의 4개의 전형적인 시험함수 $m(x)$와 σ에서 $MATLAB$을 이용하여 모의실험을 하였다.

$$\begin{aligned}
1) m(x) &= x + 2\exp(16x^2),\ \sigma = 0.4 \\
2) m(x) &= \sin(2x) + 2\exp(-16x^2),\ \sigma = 0.3 \\
3) m(x) &= 0.3\exp(-4(x+1)^2) + 0.7\exp(-16(x-1)^2),\ \sigma = 0.1 \\
4) m(x) &= 0.4x + 1,\ \sigma = 0.15
\end{aligned}$$

그리고 이 실험에서 사용된 계획점의 분포는 $X \sim U(-2,2)$로 하였다. Ruppert 등(1995)의 방법과의 비교를 위하여 크기 (n)가 50, 100 인 표본을 200번 반복$(nrep)$ 추출하여 $MISE$의 추정치

$$\widehat{MISE} = \frac{1}{nrep}\sum^{nrep}\frac{1}{n}\sum_{i=1}^{n} D_i^2$$

를 계산하였고, 이때 $D_i = (\hat{m}(x_i) - m(x_i))$, 또 추정치의 표본변동을 비교하기 위하여

$$\widehat{std} = \frac{1}{nrep}\text{std}(D_i^2)$$

를 계산하였다. $(\hat{MSE}(x_1), \cdots, \hat{MSE}(x_n))$ 를 최소화하는 $(\hat{h(x_1)}, \cdots, \hat{h(x_n)})$ 는 변동이 너무 심해서 $(\hat{m(x_1)}, \cdots, \hat{m(x_n)})$ 를 추정하는 평활량으로 바로 사용하기에는 무리가 있었다.

그래서 본 연구에서는 평활량의 값을 $3 \times 10/3$ 으로 하여 $(x_1, h(\hat{x}_1)), \cdots, (x_n, h(\hat{x}_n)))$ 을 평활하여 얻은 $(h(\hat{x}_1)^*, \cdots, h(\hat{x}_n)^*)$ 를 사용하여 $\hat{m}(x_1), \cdots, \hat{m}(x_n)$ 을 추정하였다.

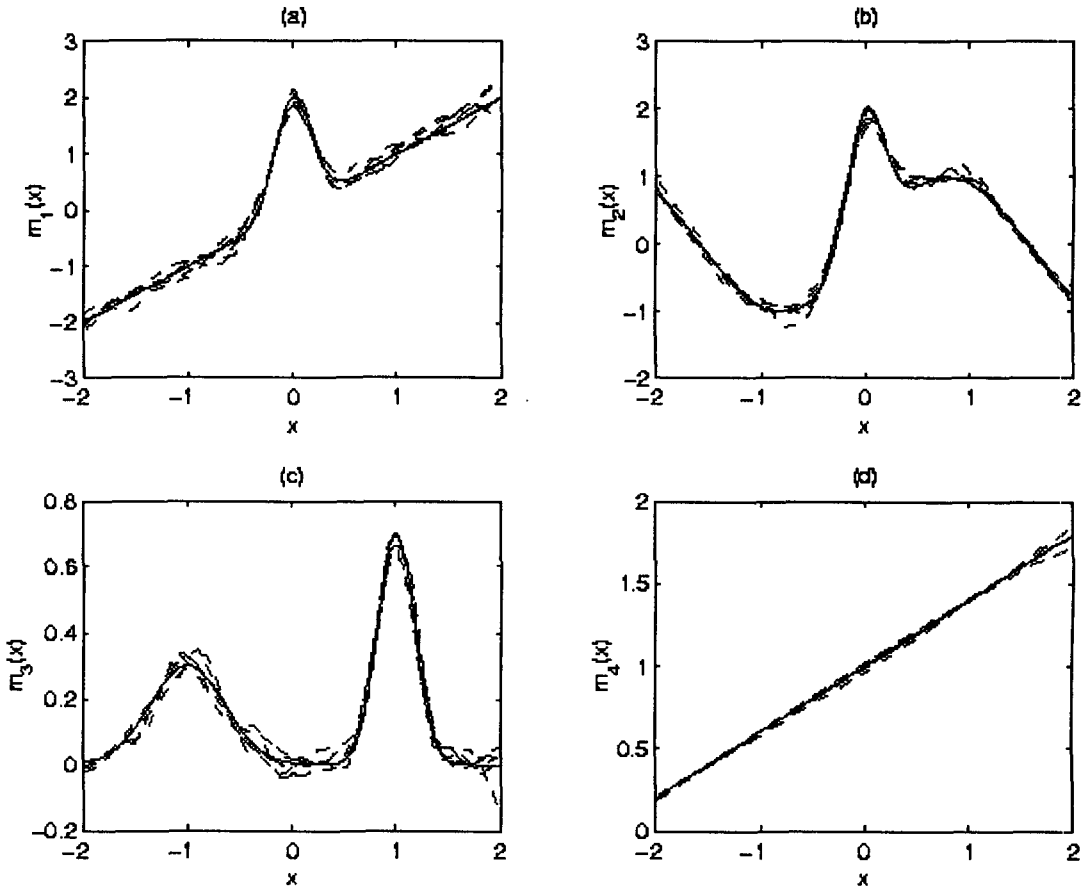

그림 1. 시험 회귀함수(실선)와 변수평활량 국소 선형회귀함수 추정치(점선)

표 1에는 모의실험의 결과가 나타나 있다. 이 표에 의하면 본 연구에서 제안된 방법이 Ruppert 등(1995)의 삽입방법보다 전반적으로 더 우수함을 알 수가 있다. 그리고 표본의 크기가 50일 때 보다 100일 때 제안된 방법이 더 우수함을 볼 수가 있다. 이 사실로 미루어 제안된 방법의 수렴속도도 더 빠를 것으로 추측된다. 이 표로부터 국소 회귀함수 추정법은 선형회귀함수모형을 추정하는데 있어서 아주 우수하다는 것을 알 수가 있다. 우리가 모의실험을 위하여 사용한 모형은 추정하기 어려운 정도를 표시하는 noise-to-signal의 비, $\sigma^2/[\text{var}\{m(X)\} + \sigma^2]$가 1/3 정도로 모두 비슷하기 때문이다. 제안된 방법은 선형함수 모형, 4)번 모형에서 더 우수하고, 함수 모형 2)번에서도 더 우수함을 보였다.

얼마만큼 실제의 함수에 가깝게 추정 할 수 있는지를 알아보기 위하여 $n = 200$인 표본

을 5번 반복 추출하여 추정된 함수를 겹쳐서 그려보았다. 그림 1이 이러한 실험결과를 보여주고 있다. 이 그림으로부터 제안된 방법이 실제의 함수를 잘 추정하는 것을 볼 수가 있다. 특히 2)번과 4)번 함수를 추정한 결과가 아주 만족스러운데 이는 표 1에서 나타난 결과와 일치한다. 이상의 모의실험 결과를 종합하면 제안된 평활량 추정법은 여러 가지 함수를 추정하는데 있어서 좋은 선택법으로 사료된다. 이외에도 회귀함수를 추정하는 과정에서 생길 수 있는 $X'X$의 비정칙성 문제는 평활량 선택과정에서 해결하기 때문에 회귀함수를 추정하는 과정에서는 이러한 문제점이 전혀 발생하지 않는다는 큰 장점을 가지고 있다. 그러나 제안된 방법이 교차타당성 방법을 이용하기 때문에 계산하는데 많은 시간이 필요하다는 단점을 가지고 있다.

표 1. 제안된 방법과 삽입방법의 비교.

표본크기	$m(x)$모형	proposed ($\hat{std}$)	plug-in($\hat{std}$)
50	1)	0.0494 (0.0914)	0.0524 (0.0887)
	2)	0.0332 (0.0587)	0.0407 (0.0758)
	3)	0.0044 (0.0087)	0.0045 (0.0085)
	4)	0.0016 (0.0022)	0.0024 (0.0034)
100	1)	0.0233 (0.0377)	0.0305 (0.0496)
	2)	0.0158 (0.0267)	0.0214 (0.0369)
	3)	0.0019 (0.0033)	0.0022 (0.0035)
	4)	0.0008 (0.0011)	0.0012 (0.0017)

5. 결론 및 향후과제

제안된 추정량의 수렴율을 규명하여 기존의 방법과 비교를 해야겠다. 그리고 다른 변수 평활량 추정법과의 비교를 통해 제안된 방법의 우수성을 규명해야겠다. 또한 계산속도를 줄일 수 있는 algorithm의 개발을 계속해 나가야 하겠다.

참고문헌

1. Fan, J. and Gijbels, I. (1995). Data-driven bandwidth selection in local polynomial fitting: Variable bandwidth and spatial adaptation, *Journal of Royal Statistical Society* , Ser. B, Vol. 57, 371-394.

2. Fan, J. and Gijbels, I. (1996). *Local Polynominal Modelling and Its Applications.* Chapman and Hall.

3. Fan, J. and Hunag, L. S. (1998). Rates of convergence for the pre-asymptotic substitution bandwidth selector, unpublished manuscript.

4. Fan, J., Gijbels, I. , Hu, T. C. and Huang, L. S. (1996). A study of variable bandwidth selection for local polynomial regression, *Statistica Sinica*, 6, 113-127.

5. Fan, J., Hall, P., Martin, M. and Patil, P. (1996). On local smoothing of nonparametric curve estimators, *Journal of the American Statistical Association*, 91, 258-266.

6. Ruppert, D., Sheather, S. J. and Wand, M. P. (1995). An effective bandwidth selectors for local least square regression, *Journal of the American Statistical Association*, 90, 1257-1270.

7. Hart, J. D. and Yi, S. (1998). One-sided cross-validation,*Journal of the American Statistical Association*, 93, 620-631.

8. Simonoff, J. S. (1996). *Smoothing Methods in Statistics*, Springer.

On variable bandwidth Kernel Regression Estimation[5]

Kyungha Seog [6] · Sungsuk Chung [7] · Daehak Kim [8]

Abstract

Local polynomial regression estimation is the most popular one among kernel type regression estimator. In local polynomial regression function esimation bandwidth selection is crucial problem like the kernel estimation. When the regression curve has complicated structure variable bandwidth selection will be appropriate. In this paper, we propose a variable bandwidth selection method fully data driven. We will choose the bandwdith by selecting minimising estiamted MSE which is estimated by the pilot bandwidth study via croos-validation method. Monte carlo simulation was conducted in order to show the superiority of proposed bandwidth selection method.

[5]This research was supported by NON DIRECTED RESEARCH FUND, Korea Research Foundation, 1997

[6]Associate Professor, Department of applied statistics, Inje University, Kimhae, Kyungnam, 621-749, Korea

[7]Assistant Professor, Department of statistics, Chonbuk University, Chonju, Chunbuk, 560-756, Korea

[8]Associate Professor, Department of statistical information, Catholic University of Taegu-Hyosung, Kyungsan, Kyungbuk, 712-702, Korea