비모수 회귀함수 추정에서의 점근적 L_1거리의 최소화

金 鐵 洙*

Asymptotic Minimization of L_1 Distance in Nonparametric Regression Function Estimation

*Kim, Chul-soo**

Summary

The method of bandwidth Choice minimizing the normalized integrated absolute error is considered for the fixed design regression model with higher order kernel function. The results are derived using a modification of the result of Myoungshic Jhun (1988).

1. Introduction

Nonparametric regression and curve fitting method for the estimation of regression functions are useful when the dynamics underlying a measured time course are of interested. There are several curve-fitting methods available for the nonparametric estimation of a regression function. Kernel estimators in nonparametric regression for the fixed disign case were introduced by Priestly and Chao (1972). We consider a modified version proposed by Gasser and Müller (1979).

Myoungshic Jhun (1988) proposed a L_1 badwidth selection method in kernel regression for the stocastic design case. In this article, we consider the fixed design regression and apply the results of Myoungshic Jhun (1988) to the fixed design case with high order kernel.

2. Asymptotic properties of regression function estimate

Consider the fixed design regression model

(1) $Y_i = m(x_i) + \epsilon_i \quad i = 1, 2, \cdots, n,$

where the known x_i for $0 \le x_1 \le x_2 \le \cdots \le x_n = 1$ and the unknown regression function $m \in C[0, 1]$ is to be estimated. The errors (ϵ_i) are assumed to be independently and identically distributed with $E(\epsilon_1) = 0$, $E(\epsilon_1^2) = \sigma^2 < \infty$.

Given $x \in (0, 1)$, a kernel estimate of $m(x)$ by Gasser and Müller (1979) is defined by

(2) $$m_n(x) = \frac{1}{h} \sum_{i=1}^{n} \int_{s_{i-1}}^{s_i} K\left(\frac{x-u}{h}\right) du \, Y_i$$

* 自然科學大學 專任講師

where h is the bandwidth now depending on n, K is a kernel function satisfying some regularity conditions and $s_0 = s_n = 1$, $s_i = \frac{x_i + x_{i+1}}{2}$, $i = 1, \cdots, n-1$

Basic requriements for kernel estimate (2) to be consistent are

$$h \to 0, nh \to \infty$$

as $n \to \infty$ and

$$\int K(x)\,dx = 1.$$

We say that a kernel K is of order p if

$$\int K(x)^j dx = 0 \text{ for all } 1 < j < p$$

but

$$\int K(x) x^p dx \neq 0.$$

For further study, we need the following assumptions :

A1. K is of order p

A2. K has compact support $[-1, 1]$

A3. m is p times continuously differentiable.

For fixed p, under the assumptions A1-A3, we obtain for bias and variance of the estimate (1), assuming $h \to 0$ and $nh \to \infty$:

$$\text{(3)} \quad Em_n(x) - m(x) = \frac{(-1)^p}{p!} h^p (m^{(p)} B(K) + 0(1)) + O\left(\frac{1}{n}\right)$$

$$\text{(4)} \quad Var(m_n(x)) = \frac{\sigma^2}{nh} (v(K) + 0(1))$$

where $B(K) = \int x^p K(x)\,dx$ and $v(K) = \int K^2(x)\,dx$. By (3) and (4), the MSE optimal bandwidth sequence is seen to be $h \sim n^{-\frac{1}{2p+1}}$, and this yields the rate of convergence $MSE \sim n^{-\frac{2p}{2p+1}}$

LEMMA 1. For fixed h and x, under the assumptions A1-A3,

$$m_n(x) - m(x) = \frac{(-1)^p}{p!} h^p m^{(p)}(x) B(K) + W_n + 0(h^p)$$

where W_n is asymptotically normal with mean zero and variance

$\frac{\sigma^2}{nh} v(K)$ as $n \to \infty$ provided $nh \to \infty$ as $n \to \infty$.

PROOF : Note that

$$m_n(x) - m(x) = Em_n(x) - m(x) + m_n(x) - Em_x(x)$$
$$= Bias + m_n(x) - Em_n(x)$$

Using (4) and central limit theorem, $m_n(x) - Em_n(x)$ is asymptotically normal with mean zero and variance $\frac{\sigma^2}{nh} v(K)$.

Thus we otain the result.

LEMMA 2. For fixed $c > 0$ and $x \in R$, let

$$Z_n(x, c) = n^{\frac{p}{2p+1}} [m_n(x, cn^{-\frac{1}{2p+1}}) - m(x)].$$

Then, under the assumptions A1-A3, $Z_n(x, c)$ weakly converges to a normal random variable with mean

$$\text{(5)} \quad \mu(K, x, c) = \frac{(-1)^p}{p!} c^p m^{(p)}(x) B(K)$$

and variance

$$\text{(6)} \quad s^2(K, x, c) = \frac{\sigma^2}{c} v(K)$$

PROOF : From Lemma 1, the asymptotic normality of $Z_n(x, c)$ follows.

3. Minimization of IAE

Consider integrated absolute error loss

(7) $$IAE(h)=\int|m_n(x,h)-m(x)|dx$$

as a criterion for the selection of the bandwidth h.

LEMMA 3. Let Z be a standard normal random variable and y be a real number.

Define $\psi(y)=E|Z-y|$.
Then $\psi(y)=\psi(-y)$.

PROOF: Note that $\psi(y)=\int_{-\infty}|z-y|\phi(y)dz$. Then $\psi(y)=[2\Phi(y)-1]y+2\phi(y)=[2(1-\Phi(-y))-1]y+2\phi(y)=[2\Phi(-y)-1](-y)+2\phi(-y)=E|Z+y|$ where Φ is the standard normal distribution function and ϕ standard normal density function.

With $h=cn^{-\frac{1}{2p+1}}$, where c>0, let

$$J_n(c)=n^{-\frac{2p}{2p+1}}\int_0^1|m_n(x,cn^{-\frac{1}{2p+1}})-m(x)|dx$$

and

$$T_n(c)=E(J_n(c))$$

LEMMA 4. Under the assumptions A1-A3,

$$\lim_{n\to\infty}T_n(c)=\int_0^1 s(x,c)\psi\left\{\frac{\mu(x,c)}{s(x,c)}\right\}dx.$$

PROOF : Let Z be a standard normal random variable. Then by Lemma 2, $Z_n(x,c)$ converges in distribution to a random variable

$$s(x,c)\left\{Z+\frac{\mu(x,c)}{s(x,c)}\right\}$$

and

$$\lim_{n\to\infty}E(Z_n(x,c))=s(x,c)\psi\left\{\frac{\mu(x,c)}{s(x,c)}\right\}$$

as $n\to\infty$ for a.e. Hence by bounded convergence theorem, we have the above result.

We will find the otimal c which minimizes limit of the normalized IAE

$$\lim_{n\to\infty}n^{2p+1}\int^1|m_n(x,cn^{-\frac{1}{2p+1}})-m(x)|dx.$$

THEOREM. Let $H(c)=\lim_{n-\infty}T_n(c)$. Then H(c) attains its minimum at a unique point c^* for the kernel of order p where p is even positive integer.

PROOF: Note that

$$H(c)=\int_0^1 s(x,c)\psi\left\{\frac{\mu(x,c)}{s(x,c)}\right\}dx$$

where μ and s^2 are given as in (5), (6).

Since $H(c)\to\infty$ as $c\to0$ and is continuous in c, it is sufficient to show that

$$\frac{d}{dc}H(c)=0 \text{ only once.}$$

Now,

$$\frac{d}{dc}H(c)=(-\frac{1}{2})C^{-\frac{3}{2}}\ \sigma v(K)^{\frac{1}{2}}\int_0^1\psi(\frac{\mu}{s})dx$$
$$+\int_0^1 s\psi'(\frac{\mu}{s})\ \frac{d}{dc}(\frac{\mu}{s})dx=C^{-\frac{3}{2}}\delta(c)$$

where $k_p=\frac{(-1)^p}{p!}$ and

$$\delta(c)=\frac{-1}{2}\sigma v(K)^{\frac{1}{2}}\int_0^1\psi(\frac{\mu}{s})dx+(p+\frac{1}{2})\kappa_p c^{p+\frac{1}{2}}$$
$$B(K)\int_0^1\psi'(\frac{\mu}{s})m^{(p)}(x)dx.$$

But,

$$\delta'(c)=p(p+\frac{1}{2})\kappa_p c^{p-\frac{1}{2}}B(K)^2\int_0^1\psi'(\frac{\mu}{s})m^{(p)}(x)$$

$$dx + \frac{(p+\frac{1}{2})^2 \kappa_p^2 c^{2p}}{\sigma v^{\frac{1}{2}}(K)} B(K)^2 \int_0^1 \psi''\left(\frac{\mu}{s}\right) m^{(p)}(x)\,dx$$

For even positive integer p, $\delta'(c)>0$, so $\delta(c)$ is an increasing function of c for each $x \in (0,1)$. Also $\delta(c)$ converges to $\delta(0)<0$ as $c \to 0$ and $\delta(c) \to \infty$ as $c \to \infty$.

Thus there exists exactly one value c such that $\delta(c)=0$.

Therfore there exists only one c^* such that $H(c)=0$

REFERENCES

Christine Jennen-Steinmeth and Theo Gasser. 1988. A unifying approach to nonparametric regression estimation, *JASA.* 83, 1084~1089.

Peter Hall and M.P. Wand. 1988. On the minimization of absolute distance in kernel density estimation, *St. and Prob. letters,* 6, 311~314.

Hans-Georg Müller. 1987, On the asymptotic MSE of L_1 kernel estimates, *J. Approxi. Th.* 51, 193~201.

Joan G. Staniswalis. 1989. Local bandwidth selection for kernel estimates, *JASA.* 84, 284~288.

Myoungshic Jhun. 1988, L^1 bandwidth selection in kernel regression function estimation, *J. KOR. STAT. Sco.* 17, 1~8.

M.B. Priestly and Chao. 1972. Nonparametric function fitting, *J. Roy. St. Soc.* B 34, 385~392.

Theo Gasser and hans-Georg Müller. 1979. Kernel estimation of regression functions in "Smoothing techniques for curve estimation", *Lec. Noet in Math.*, 757, 5~22.

Theo Gasser and Hans-Georg Müller. 1984. Estimating regression functions and their derivatives by the kernel method, *Scand. J. Stat.*, 11, 171~185.

국문요약

고계 kernel을 갖는 kernel 회귀함수 추정량에서의 L_1거리를 최소화하는 bandwidth를 선택하는 방법을 다뤘다. Kernel함수가 짝수위수를 갖는 경우 L_1거리를 최소화하는 해는 일의적임을 보였다.