SERIE RESEARCH MEMORANDA

NONPARAMETRIC TIME SERIES REGRESSION

Herman J. Bierens

Research Memorandum 1988-59 Dec. '88

VRIJE UNIVERSITEIT
FACULTEIT DER ECONOMISCHE WETENSCHAPPEN
EN ECONOMETRIE
AMSTERDAM

NONPARAMETRIC TIME SERIES REGRESSION *

Herman J.Bierens
Department of Econometrics
Free University, Amsterdam

December 1988

---

* Chapter 12 of:
PRINCIPLES OF NONLINEAR AND NONPARAMETRIC REGRESSION ANALYSIS
To be published by Marcel Dekker Inc., New York.

PRINCIPLES OF NONLINEAR AND NONPARAMETRIC REGRESSION ANALYSIS

TABLE OF CONTENTS:

PREFACE

This book deals with statistical inference of nonlinear regression models from two opposite points of view, namely the case where the functional form of the model is completely specified as a known function of regressors and unknown parameters, and the opposite case where the functional form of the model is completely unknown. First it is assumed that the response function of the regression model under review belongs to a certain well-specified parametric family of functional forms, by which estimation of the model merely amounts to estimation of the unknown parameters. For this class of models we review the asymptotic properties of the nonlinear least squares estimator for independent data as well as for time series.

In practice assumptions on the functional form are often made on the basis of computational convenience rather than on the basis of precise a priori knowledge of the empirical phenomenon under review. Therefore the linear regression model is still the most popular model specification in applied research. However, even if the specification of the functional form is based on sound theoretical considerations there is quite often a large range of functional forms that are theoretically admissible, so that there is no guarantee that the actually chosen functional form is true. Functional specification of a parametric nonlinear regression model should therefore always be verified by conducting model misspecification tests. Various model misspecification tests will therefore be discussed, in particular consistent tests which have asymptotic power 1 against all deviations from the null hypothesis that the model is correct.

The opposite case of parametric regression is nonparametric regression. Nonparametric regression analysis is concerned with estimation of a regression model without specifying in advance its functional form. Thus the only source of information about the functional form of the model is the data set itself. In this book we shall review various nonparametric regression approaches, with special emphasis on the kernel method, under various distributional assumptions.

This book is divided into three parts. In the first part we review the elements of abstract probability theory we need in part 2. Part 2 is devoted to the asymptotic theory of parametric and nonparametric regression analysis in the case of independent data generating processes. In part 3 we extend the analysis involved to time series.

The selection of the topics mainly reflexes my own interest in the subject. Instead of providing an encyclopedic survey of the literature, I have chosen for a setup which aims to fill the gap between intermediate statistics (including linear time series analysis) and the level necessary to get access to the recent literature on nonlinear and nonparametric regression analysis, with emphasis on my own contributions. The ultimate goal is to provide the student with the tools for his own independent research in this area, by showing what tools I and others have used and what they have been used for. Thus, this book may be viewed as an account of my own struggle with the material involved. I think this book is particularly suitable for self-tuition (at least it aims to be), and may prove useful in a graduate course in mathematical statistics and advanced econometrics.

Acknowledgements:

The first five chapters of this book have been disseminated in draft form as working papers. I am grateful to Anil Bera, Alexander Georgiev and Jan Magnus for suggesting additional references, and in particular to Lourens Broersma, Johan Smits and Ton Steerneman who suggested various improvements.

A large body of the material in chapter 6 has been published earlier in Truman F.Bewley (ed.), *Advances in Econometrics, Fifth World Congress*, Cambridge University Press. I am indebted to Cambridge University Press for granting permission to reprint it.

12. NONPARAMETRIC TIME SERIES REGRESSION

Recently the kernel regression approach has been extended to time series. Robinson (1983) shows strong consistency and asymptotic normality, using the α-mixing concept. In Bierens (1983) we proved uniform consistency under ν-stability in L^2 with respect to a φ-mixing base, and in Bierens (1987) we generalized this result to pointwise consistency and asymptotic normality. Collomb (1985) proves uniform strong consistency under the φ-mixing condition and Georgiev (1984) proves consistency in the case of a Markov data generating process. In this chapter we shall extend the results in chapter 6 to time series, employing the ν-stablity and φ-mixing concepts, on the basis of the results in Bierens (1987)*.

12.1 *Assumptions and preliminary lemmas*

In this section we state the assumptions we need to prove pointwise and uniform consistency and asymptotic normality of kernel time series regression estimators, and we prove three basic lemmas.

Assumption 12.1.1. The data generating process $\{(Y_t, X_t)\}$ is a strictly stationary ν-stable process in L^2 with respect to a strictly stationary φ-mixing base (W_t), where

$$\nu(m) = O(\exp(-c\cdot m)) \text{ for some } c > 0; \tag{12.1.1}$$

$$\Sigma_{m=0}^{\infty}\varphi(m)^{1/2} < \infty. \tag{12.1.2}$$

Also, we assume that $g(X_t)$ represents the conditional expectation of Y_t given the entire past of the data generating process:

Assumption 12.1.2. Let

$$g(X_t) = E(Y_t|X_t, X_{t-1}, X_{t-2}, \dots, Y_{t-1}, Y_{t-2}, \dots) \text{ a.s.} \tag{12.1.3}$$

The vector X_t may contain lagged Y_t's. Thus $g(X_t)$ is in fact a

---

*) Reprinted with permission of Cambridge University Press.

(non)linear ARX model with unknown functional form $g(.)$. We note that Robinson (1983) only assumes $E(Y_t|X_t) = g(X_t)$, which is weaker than (12.1.3) However, as has been argued in chapter 10, proper time series models should satisfy condition (12.1.3). The errors U_t are then martingale differences, so that the martingale difference central limit theorem 9.1.7 is applicable.

Next, we assume:

Assumption 12.1.3.
(I) If assumption 6.2.1 holds then in addition:
  (a) $\sigma_u^4(x)h(x)$ is uniformly bounded;
  (b) $g(x)^2h(x)$ has countinuous and bounded second derivatives.
(II) If assumption 6.4.2 holds then for every fixed $x^{(2)} \in \Delta_2$,
  (a) $\sigma_u^4(x^{(1)},x^{(2)})h(x^{(1)}|x^{(2)})$ is uniformly bounded on Δ_1;
  (b) $g(x^{(1)},x^{(2)})^2h(x^{(1)}|x^{(2)})$ has continuous and bounded second derivatives with respect to the components of $x^{(1)}$.

Finally, in order to prove theorem 6.5.1 for time series we need the following addition to assumption 12.1.1:

Assumption 12.1.4. The process $\{(Y_j,X_j)\}$ is ν-stable in L^4 with respect to the φ-mixing base considered in assumption 12.1.1.

The following lemma will be used to prove pointwise consistency and asymptotic normality of the kernel regression estimators considered in sections 6.1 through 6.4.

Lemma 12.1.1. Let $\{(Z_j,X_j)\}$ be a strictly stationary stochastic process in $R\times R^k$, with $E\,Z_j^4 < \infty$ and $E|X_j|^2 < \infty$. Let this process be ν-stable in L^2 with respect to a strictly stationary φ-mixing base, where ν satisfies condition (12.1.1) and φ satisfies condition (12.1.2). Let K be a Borel measurable real function on R^k such that

$$\int|K(x)|dx < \infty;\ \int|t\psi(t)|dt < \infty,$$

where

$$\psi(t) = \int \exp(i \cdot t'x) K(x) dx.$$

Denote for $x \in R^k$,

$$d_n(x) = \operatorname{var}\{(1/n)\Sigma_{j=1}^{n} Z_j K((x-X_j)/\gamma_n)\}$$

$$- (1/n^2)\Sigma_{j=1}^{n} \operatorname{var}\{Z_j K((x-X_j)/\gamma_n)\}, \qquad (12.1.4)$$

where $\gamma_n > 0$, $\gamma_n \downarrow 0$ as $n \to \infty$. Then

$$d_n(x) = O[\{\ln(n/\gamma_n)+\ln(1/(E\ Z_0^2 K((x-X_0)/\gamma_n)^2))\}$$

$$\times E\ Z_0^2 K((x-X_0)/\gamma_n)^2/n]$$

Proof: We can write

$$d_n(x) = 2(1/n^2)\Sigma_{\ell=1}^{n-1}\Sigma_{j=1}^{n-\ell} \operatorname{cov}\{Z_0 K((x-X_0)/\gamma_n), Z_j K((x-X_j)/\gamma_n)\}$$

$$= 2(1/n^2)\Sigma_{\ell=1}^{n-1}\Sigma_{j=1}^{n-\ell} E\{Z_0 Z_j K((x-X_0)/\gamma_n) K((x-X_j)/\gamma_n)\}$$

$$- 2(1/n^2)\Sigma_{\ell=1}^{n-1}\Sigma_{j=1}^{n-\ell} E\{Z_0 K((x-X_0)/\gamma_n)\}^2. \qquad (12.1.5)$$

Similarly, let

$d_n^{(m)}(x)$

$$= 2(1/n^2)\Sigma_{\ell=1}^{n-1}\Sigma_{j=1}^{n-\ell} \operatorname{cov}\{Z_0^{(m)} K((x-X_0^{(m)})/\gamma_n), Z_j^{(m)} K((x-X_j^{(m)})/\gamma_n)\}$$

$$= 2(1/n^2)\Sigma_{\ell=1}^{n-1}\Sigma_{j=1}^{n-\ell} E\{Z_0^{(m)} Z_j^{(m)} K((x-X_0^{(m)})/\gamma_n) K((x-X_j^{(m)})/\gamma_n)\}$$

$$- 2(1/n^2)\Sigma_{\ell=1}^{n-1}\Sigma_{j=1}^{n-\ell} E\{Z_0^{(m)} K((x-X_0^{(m)})/\gamma_n)\}^2, \qquad (12.1.6)$$

where

$$Z_j^{(m)} = E(Z_j | W_j, W_{j-1}, W_{j-2}, \ldots, W_{j-m});$$

$$X_j^{(m)} = E(X_j | W_j, W_{j-1}, W_{j-2}, \ldots, W_{j-m}).$$

We shall prove the lemma in three steps:

Step 1:

$$|d_n^{(m)}(x)| \le 4n^{-1}(m+\Sigma_{\ell=0}^{\infty}\varphi(\ell)^{½})E\{Z_0^{(m)}K((x-X_0^{(m)})/\gamma_n)\}^2;$$

Step 2: For sufficiently large n and some constant $c_1 > 0$, independent of j and m, we have

$$E\ \sup_x|\{Z_0K((x-X_0)/\gamma_n)\}\{Z_jK((x-X_j)/\gamma_n)\}$$

$$-\ \{Z_0^{(m)}K((x-X_0^{(m)})/\gamma_n)\}\{Z_j^{(m)}K((x-X_j^{(m)})/\gamma_n)\}|$$

$$\le c_1\gamma_n^{-1}\nu(m);$$

Step 3: For sufficiently large n and some constant $c_2 > 0$, $c_3 > 0$, independent of x, j and m, we have

$$|E\{Z_0K((x-X_0)/\gamma_n)\}E\{Z_jK((x-X_j)/\gamma_n)\}$$

$$-\ E\{Z_0^{(m)}K((x-X_0^{(m)})/\gamma_n)\}E\{Z_j^{(m)}K((x-X_j^{(m)})/\gamma_n)\}|$$

$$\le c_2\{\gamma_n^{-1}\nu(m)\}^2 + c_3\gamma_n^{-1}\nu(m).$$

We first show that the results of these three steps imply the lemma. Let (m_n) be a sequence of positive integers such that $m_n \to \infty$ and $m_n/n \to 0$, and let n be so large that

$$m_n \ge \Sigma_{\ell=0}^{\infty}\varphi(\ell)^{½}.$$

It follows from step 1 and step 2 with j=0 that for sufficiently large n,

$$|d_n^{(m)}(x)| \le 8(m_n/n)E\{Z_0K((x-X_0)/\gamma_n)\}^2 + 8(m_n/n)c_1\gamma_n^{-1}\nu(m_n).$$

Moreover, it follows from steps 2 and 3 that

$$|d_n(x)-d_n^{(m)}(x)| \le 2(c_1+c_3)\gamma_n^{-1}\nu(m_n) + 2c_2\{\gamma_n^{-1}\nu(m_n)\}^2.$$

Without loss of generality we may assume that for sufficiently large n,

$$m_n/n \le (c_1+c_2+c_3)/4 \text{ and } \gamma_n^{-1}\nu(m_n) \le 1 \qquad (12.1.7)$$

(as will appear). Then

$$|d_n(x)| \le 8(m_n/n)E\{Z_0K((x-X_0)/\gamma_n)\}^2 + 4(c_1+c_2+c_3)\gamma_n^{-1}\nu(m_n)$$

$$= O(\rho_n(x)),$$

where

$$\rho_n(x) = (m_n/n)E\{Z_0K((x-X_0)/\gamma_n)\}^2+\gamma_n^{-1}\exp(-c\cdot m_n). \quad (12.1.8)$$

Minimizing the right hand side of (12.1.8) to m_n yields

$$\nu(m_n) = \exp(-c\cdot m_n) = (1/c)(\gamma_n/n)\{E\ Z_0K((x-X_0)/\gamma_n)\}^{-2},$$

with

$$m_n = (1/c)[\log(c) + \log(n/\gamma_n)-2\cdot\log(E\ Z_0K((x-X_0)/\gamma_n)),$$

(thus observe that indeed (12.1.7) holds). Hence

$$\rho_n(x) \le (1/c)\big|1 + \log(c) + \ln(n/\gamma_n)$$

$$+ \log(\{E\ Z_0^2K((x-X_0)/\gamma_n)^2\}^{-1})\big|\{E\ Z_0^2K((x-X_0)/\gamma_n)^2/n\}$$

$$= O[\{\ln(n/\gamma_n)+\ln(E\ Z_0^2K((x-X_0)/\gamma_n)^2)^{-1}\}E\ Z_0^2K((x-X_0)/\gamma_n)^2/n],$$

as was to be shown.

Proof of step 1:
Since $\{Z_j^{(m)}K((x-X_j^{(m)})/\gamma_n)\}$ is a φ^*-mixing sequence, where

$$\varphi^*(\ell) = 1 \text{ if } \ell < m,\ \varphi^*(\ell) = \varphi(\ell-m) \text{ if } \ell \ge m,$$

it follows from (9.2.3) and theorem 9.2.1 that

$$\text{cov}\{Z_0^{(m)}K((x-X_0^{(m)})/\gamma_n),Z_j^{(m)}K((x-X_j^{(m)})/\gamma_n)\}$$

$$\le 2\varphi^*(j)^{\frac{1}{2}}E\{Z_0^{(m)}K((x-X_0^{(m)})/\gamma_n)\}^2.$$

Step 1 follows now from the fact that

$$\Sigma_{j=0}^{\infty}\varphi^*(j)^{\frac{1}{2}} \le m + \Sigma_{j=0}^{\infty}\varphi(j)^{\frac{1}{2}}.$$

Proof of step 2:

From the inversion formula for Fourier transforms we have

$$K(x) = (1/(2\pi))^k \int \exp(-i\cdot t'x)\psi(t)dt \qquad (12.1.9)$$

From (12.1.9) it follows that for all $x \in R^k$

$$|Z_0 Z_j K((x-X_0)/\gamma_n)K((x-X_j)/\gamma_n)$$

$$- Z_0^{(m)} Z_j^{(m)} K((x-X_0^{(m)})/\gamma_n)K((x-X_j^{(m)})/\gamma_n)|$$

$$\leq (1/(2\pi))^{2k}\gamma_n^{2k}\int\int|Z_0 Z_j \exp(i\cdot t_1{}'X_0+i\cdot t_2{}'X_j)$$

$$- Z_0^{(m)} Z_j^{(m)} \exp(i\cdot t_1{}'X_0^{(m)}+i\cdot t_2{}'X_j^{(m)})||\psi(\gamma_n t_1)\psi(\gamma_n t_2)|dt_1 dt_2 .$$

$$\leq (1/(2\pi))^{2k}\gamma_n^{2k}\int\int|Z_0 Z_j \exp(i\cdot t_1{}'X_0+i\cdot t_2{}'X_j)$$

$$- Z_0 Z_j \exp(i\cdot t_1{}'X_0^{(m)}+i\cdot t_2{}'X_j^{(m)})||\psi(\gamma_n t_1)\psi(\gamma_n t_2)|dt_1 dt_2$$

$$+ (1/(2\pi))^{2k}\gamma_n^{2k}\int\int|Z_0 Z_j \exp(i\cdot t_1{}'X_0^{(m)}+i\cdot t_2{}'X_j^{(m)})$$

$$- Z_0^{(m)} Z_j^{(m)} \exp(i\cdot t_1{}'X_0^{(m)}+i\cdot t_2{}'X_j^{(m)})||\psi(\gamma_n t_1)\psi(\gamma_n t_2)|dt_1 dt_2 .$$

$$\leq (1/(2\pi))^{2k}\gamma_n^{2k}|Z_0 Z_j||(X_0,X_j)-(X_0^{(m)},X_j^{(m)})|$$

$$\times \int\int|(t_1,t_2)||\psi(\gamma_n t_1)\psi(\gamma_n t_2)|dt_1 dt_2$$

$$+ (1/(2\pi))^{2k}\gamma_n^{2k}|Z_0 Z_j - Z_0^{(m)} Z_j^{(m)}|\int\int|\psi(\gamma_n t_1)\psi(\gamma_n t_2)|dt_1 dt_2 .$$

$$\leq (1/(2\pi))^{2k}\gamma_n^{-1}|Z_0 Z_j||(X_0,X_j)-(X_0^{(m)},X_j^{(m)})|$$

$$\times \int\int|(t_1,t_2)||\psi(t_1)\psi(t_2)|dt_1 dt_2$$

$$+ (1/(2\pi))^{2k}|Z_0 Z_j - Z_0^{(m)} Z_j^{(m)}|\int\int|\psi(t_1)\psi(t_2)|dt_1 dt_2 .$$

$$\leq (1/(2\pi))^{2k}\gamma_n^{-1}|Z_0 Z_j|(|X_0 - X_0^{(m)}|^2+|X_j - X_j^{(m)}|^2)^{1/2}$$

$$\times \int|t||\psi(t)|dt\int|\psi(t)|dt$$

$$+ (1/(2\pi))^{2k}(|Z_0||Z_j - Z_j^{(m)}| + |Z_0 - Z_0^{(m)}||Z_j^{(m)}|)(\int|\psi(t)|dt)^2. \tag{12.1.10}$$

Applying Cauchy-Schwarz' and Liapounov's inequalities and using the inequality $E(Z_j^{(m)})^2 \le E\, Z_j^2$ [cf. exercise 1] it follows from (12.1.10) that there exist constants $c_*^{(1)} > 0$, $c_*^{(2)} > 0$, independent of j and m, such that

$$\begin{aligned}
&E\, \sup_x |Z_0 Z_j K((x-X_0)/\gamma_n)K((x-X_j)/\gamma_n) \\
&\quad - Z_0^{(m)} Z_j^{(m)} K((x-X_0^{(m)})/\gamma_n)K((x-X_j^{(m)})/\gamma_n)| \\
&\le (1/(2\pi))^{2k}\gamma_n^{-1}(E[Z_0 Z_j]^2)^{\frac{1}{2}}|(E|X_0 - X_0^{(m)}|^2 + E|X_j - X_j^{(m)}|^2)^{\frac{1}{2}} \\
&\qquad\qquad \times \int|t||\psi(t)|dt\int|\psi(t)|dt \\
&\quad + (1/(2\pi))^{2k}(E\, Z_0^2)^{\frac{1}{2}}(E[Z_j - Z_j^{(m)}]^2)^{\frac{1}{2}}(\int|\psi(t)|dt)^2 \\
&\quad + (1/(2\pi))^{2k}(E[Z_0 - Z_0^{(m)}]^2)^{\frac{1}{2}}(E\, Z_j^2)^{\frac{1}{2}}(\int|\psi(t)|dt)^2 \\
&\le c_*^{(1)}\gamma_n^{-1}\nu(m) + c_*^{(2)}\nu(m) \quad (12.1.11)
\end{aligned}$$

For n so large that $c_*^{(1)}\gamma_n^{-1} \ge c_*^{(2)}$ we may take c_1 in step 2 equal to $2\cdot c_*^{(1)}$.

Proof of Step 3:
By stationarity and the trivial inequality

$$|a^2 - b^2| \le (a-b)^2 + |a||a-b|$$

we have

$$\begin{aligned}
&|E(Z_0 K((x-X_0)/\gamma_n))E(Z_j K((x-X_j)/\gamma_n)) \\
&\qquad - E(Z_0^{(m)} K((x-X_0^{(m)})/\gamma_n))E(Z_j^{(m)} K((x-X_j^{(m)})/\gamma_n))| \\
&= |(E\, Z_0 K((x-X_0)/\gamma_n))^2 - (E\, Z_0^{(m)} K((x-X_0^{(m)})/\gamma_n))^2|
\end{aligned}$$

$$\leq \{E|Z_0K((x-X_0)/\gamma_n) - Z_0^{(m)}K((x-X_0^{(m)})/\gamma_n)|\}^2$$

$$+ |E\ Z_0K((x-X_0)/\gamma_n)|E|Z_0K((x-X_0)/\gamma_n)-Z_0^{(m)}K((x-X_0^{(m)})/\gamma_n)| \tag{12.1.12}$$

From (12.1.9) it follows, similarly to (12.1.11), that there exist constants $c_*^{(1)} > 0$, $c_*^{(2)} > 0$, independent of x and m, such that

$$E|Z_0K((x-X_0)/\gamma_n) - Z_0^{(m)}K((x-X_0^{(m)})/\gamma_n)|$$

$$\leq (1/(2\pi))^k\gamma_n^k\int E|Z_0\exp(i\cdot t'X_0) - Z_0^{(m)}\exp(i\cdot t'X_0^{(m)})||\psi(\gamma_n t)|dt.$$

$$\leq (1/(2\pi))^k\gamma_n^k E|Z_0 - Z_0^{(m)}|\int|\psi(\gamma_n t)|dt.$$

$$+ (1/(2\pi))^k\gamma_n^k E|X_0 - X_0^{(m)}||Z_0^{(m)}|\int|t\psi(\gamma_n t)|dt.$$

$$\leq (1/(2\pi))^k\{E|Z_0 - Z_0^{(m)}|\}^{\frac{1}{2}}\int|\psi(t)|dt.$$

$$+ (1/(2\pi))^k\gamma_n^{-}\{E|X_0 - X_0^{(m)}|^2\}^{\frac{1}{2}}\{E[Z_0^{(m)}]^2\}^{\frac{1}{2}}\int|t\psi(t)|dt.$$

$$\leq c_*^{(1)}\gamma_n^{-1}\nu(m) + c_*^{(2)}\nu(m). \tag{12.1.13}$$

Realizing that by (12.1.9)

$$|E\ Z_0K((x-X_0)/\gamma_n)| \leq (1/(2\pi))^k E|Z_0|\int|\psi(t)|dt < \infty,$$

step 3 now easily follows from (12.1.12) and (12.1.13). This completes the proof. Q.E.D.

The following lemma enables us prove uniform consistency

Lemma 12.1.2. Let the conditions of lemma 12.3.1 hold, except that now $E\ Z_j^2 < \infty$ suffices, and let

$$a_n(x) = (1/n)\Sigma_{j=1}^n Z_jK((x-X_j)/\gamma_n)$$

Then

$$E\ \sup_x|a_n(x) - E\ a_n(x)| = O[(\log(n/\gamma_n^2))^{\frac{1}{2}}/\sqrt{n}]$$

Proof: Let

$$a_n^{(m)}(x) = (1/n)\Sigma_{j=1}^n Z_j^{(m)} K((x-X_j^{(m)})/\gamma_n)$$

It follows from (12.1.9), similarly to (12.1.18), that there exist constants $c_*^{(1)} > 0$, $c_*^{(2)} > 0$, independent of x and m, such that

$$E\ \sup_x |a_n(x) - a_n^{(m)}(x)| \le c_*^{(1)}\gamma_n^{-1}\nu(m) + c_*^{(2)}\nu(m). \quad (12.1.14)$$

Moreover, it follows from (12.1.9), the well-known formula

$$\exp(i\cdot u) = \cos(u) + i\cdot\sin(u),$$

Liapounov's inequality and inequality (1.4.4) that

$$E\ \sup_x |a_n^{(m)}(x) - E\ a_n^{(m)}(x)|$$

$$\le (1/(2\pi))^k \gamma_n^k \int E|(/n)\Sigma_{j=1}^n (Z_j^{(m)}\exp(i\cdot t'X_j^{(m)})$$

$$- E\ Z_j^{(m)}\exp(i\cdot t'X_j^{(m)})||\psi(\gamma_n t)|dt.$$

$$\le (\surd 2)(1/(2\pi))^k \gamma_n^k \int \{E[(/n)\Sigma_{j=1}^n (Z_j^{(m)}\cos(t'X_j^{(m)})$$

$$- E\ Z_j^{(m)}\cos(t'X_j^{(m)}))]^2\}^{\frac{1}{2}}|\psi(\gamma_n t)|dt.$$

$$+ (\surd 2)(1/(2\pi))^k \gamma_n^k \int \{E[(/n)\Sigma_{j=1}^n (Z_j^{(m)}\cos(t'X_j^{(m)})$$

$$- E\ Z_j^{(m)}\cos(t'X_j^{(m)}))]^2\}^{\frac{1}{2}}|\psi(\gamma_n t)|dt. \quad (12.1.15)$$

Similarly to step 1 in the proof of lemma 12.1.1 we have, uniformly in t,

$$E[(/n)\Sigma_{j=1}^n (Z_j^{(m)}\cos(t'X_j^{(m)}) - E\ Z_j^{(m)}\cos(t'X_j^{(m)}))]^2$$

$$\le \{(1/n) + 4n^{-1}(m+\Sigma_{\ell=0}^{\infty}\varphi(\ell)^{\frac{1}{2}})\}E[Z_0^{(m)}]^2$$

$$\le \{(1/n) + 4n^{-1}(m+\Sigma_{\ell=0}^{\infty}\varphi(\ell)^{\frac{1}{2}})\}E[Z_0]^2$$

$$\le c_*(m/n), \quad (12.1.16)$$

say, for sufficiently large m, and similarly,

$$E[(/n)\Sigma_{j=1}^{n}(Z_j^{(m)}\sin(t'X_j^{(m)}) - E\, Z_j^{(m)}\sin(t'X_j^{(m)}))]^2$$

$$\leq c_*(m/n), \qquad (12.1.17)$$

Combining (12.1.15), (12.1.16) and (12.1.17) it follows

$$E\, \sup_x |a_n^{(m)}(x) - E\, a_n^{(m)}(x)|$$

$$\leq [(2\surd 2)(1/(2\pi))^k \int |\psi(t)|dt\surd c_*](m/n)^{1/2} = c_1(m/n)^{1/2} \qquad (12.1.18)$$

say. Combining (12.1.14) and (12.1.18) it follows that for sufficiently large n

$$E\, \sup_x |a_n(x) - E\, a_n(x)| \leq 2\, E\, \sup_x |a_n^{(m_n)}(x) - E\, a_n^{(m_n)}(x)|$$

$$+ E\, \sup_x |a_n^{(m_n)}(x) - E\, a_n^{(m_n)}(x)|$$

$$\leq 2c_*^{(1)}\gamma_n^{-1}\nu(m_n) + 2c_*^{(2)}\nu(m_n) + c_1(m_n/n)^{1/2}$$

$$= O[\{(m_n/n)+\gamma_n^{-2}\exp(-2c\cdot m_n)\}^{1/2}] \qquad (12.1.19)$$

say. Minimizing the right hand side of (12.1.19) to m_n, the lemma follows. Q.E.D.

Finally, the following lemma will enable us to extend the results in section 6.5 to time series.

Lemma 12.1.3. Let (Z_j) be a strictly stationary stochastic process in R satisfying $E\, Z_j^4$. Let (Z_j) be ν-stable in L^4 with respect to a strictly stationary φ-mixing base, where

$$\nu(m) = O(\exp(-c\cdot m)) \text{ for some } c > 0,\ \Sigma_{\ell=0}^{\infty}\varphi(\ell)^{1/2} < \infty.$$

Then for every $\varepsilon > 0$,

$$\text{plim}_{n\to\infty} n^{1/2-\varepsilon}(1/n)\Sigma_{j=1}^{n}(Z_j - E\, Z_j) = 0 \qquad (12.1.21)$$

and

$$\text{plim}_{n\to\infty} n^{1/2-\varepsilon}(1/n)\Sigma_{j=1}^{n}(Z_j{}^2 - E\ Z_j{}^2) = 0. \qquad (12.1.22)$$

Proof: Let (W_j) be the base and let

$$Z_j^{(m)} = E(Z_j \,|\, W_j, W_{j-1}, W_{j-2}, \ldots, W_{j-m}).$$

Denote similarly to (12.1.4), (12.1.5) and (12.1.6)

$$d_n = \text{var}\{(1/n)\Sigma_{j=1}^{n} Z_j{}^2\} - (1/n^2)\Sigma_{j=1}^{n}\text{var}(Z_j{}^2),$$

$$d_n^{(m)} = \text{var}\{(1/n)\Sigma_{j=1}^{n}(Z_j^{(m)})^2\} - (1/n^2)\Sigma_{j=1}^{n}\text{var}((Z_j^{(m)})^2).$$

Then it follows similarly to the proof of lemma 12.1.1 that for sufficiently large n

$$|d_n^{(m)}| \le 4((m + \Sigma_{\ell=0}^{\infty}\varphi(\ell)^{1/2})/n)E\ Z_j^4 \le c_1 m/n,$$

and

$$|d_n - d_n^{(m)}| \le c_2 \nu(m),$$

for some $c_1 > 0$, $c_2 > 0$. Thus

$$d_n = O((m_n/n) + \exp(-c\cdot m_n))$$

Now choose $m = n^{\varepsilon}$. Then

$$n^{1-2\varepsilon}\text{var}\{(1/n)\Sigma_{j=1}^{n} Z_j{}^2\}$$

$$\le n^{1-2\varepsilon}\{c_1 n^{\varepsilon-1} + c_2\exp(-(1/4)cn^{\varepsilon})\} + n^{-2\varepsilon}E\ Z_0{}^4 \to 0$$

as $n \to \infty$. This proves (12.1.22). The proof of (12.1.21) is nearly the same. Q.E.D.

12.2 *Consistency and asymptotic normality of time series kernel regression function estimators*

In this section we generalize the approach in chapter 6 to time series.

Theorem 12.2.1. Let

$$E\, U_j^8 < \infty \text{ and } E\, g(X_j)^4 < \infty \tag{12.2.1}$$

and let the kernel K be such that for $\ell = 1,2$,

$$\int |t\psi_\ell(t)|dt < \infty, \text{ where } \psi_\ell(t) = \int \exp(i\cdot t'x)K(x)^\ell dx.^{*)} \tag{12.2.2}$$

Moreover,

let $\int zz'K(z)^2 dz$ be finite in the continuous case, and let $\int z_1 z_1{}'K(z_1,0)^2 dz_1$ be finite in the mixed continuous-discrete case, respectively. *) (12.2.3)

With assumptions 12.1.1, 12.1.2 and 12.1.3 and the conditions (12.2.1), (12.2.2) and (12.2.3) the asymptotic normality results in sections 6.2 and and 6.4 go through.

Proof: We only prove the theorem for the continuous case, leaving the proof for the discrete and mixed continuous-discrete cases as an exercise. We now have to show that (6.2.2) and (6.2.4) go through and that

$$\mathrm{var}(\hat{h}(x)) \to 0 \tag{12.2.4}$$

in the time series case under review, as only in these steps the independence assumption has been employed.

Step 1: Proof of (6.2.2). Since now the $v_{n,j}(x)$ defined by (6.2.7) are martingale differences, it suffices to show

---

*) Note that the conditions (12.2.2) and (12.2.3) hold for kernels of the type (6.2.36).

$$\text{plim}_{n\to\infty}(1/n)\Sigma_{j=1}^{n}(v_{n,j}(x)^2 - E\, v_{n,j}(x)^2) = 0, \qquad (12.2.5)$$

as then (6.2.2) follows from theorem 9.1.7. Thus consider lemma 12.1.1 with $Z_j = U_j^2$ and $K(x)$ replaced by $K(x)^2$. Since

$$E\, U_j^4 K((x-X_j)/\gamma_n)^4 = \gamma_n^k \int \sigma_u^4(x-\gamma_n z)h(x-\gamma_n z)K(z)^4 dz = O(\gamma_n^k),$$

it follows from lemma 12.1.1 that

$$\text{var}((1/n)\Sigma_{j=1}^{n} U_j^2 K((x-X_j)/\gamma_n)^2/\gamma_n^k)$$

$$= (1/n^2)\Sigma_{j=1}^{n}\text{var}(U_j^2 K((x-X_j)/\gamma_n)^2/\gamma_n^k) + \gamma_n^{-2k} d_n(x)$$

$$= O(1/(n\gamma_n^k)) + O((1/(n\gamma_n^k))\ln(n/\gamma_n^{k+1})) \to 0 \qquad (12.2.6)$$

as $n \to \infty$, provided γ_n is proportional to $n^{-\tau}$ with $\tau < 1/k$. Therefore, (12.2.4) follows from (12.2.6) and Chebishev's inequality.

Step 2: Proof of (6.2.4). Let Z_j in lemma 12.1.1 be

$$Z_j = Y_j - U_j - g(x) = g(X_j) - g(x).$$

Then

$$\text{var}(\sqrt{(n\gamma_n^k)}(1/n)\Sigma_{j=1}^{n}(g(X_j)-g(x))K((x-X_j)/\gamma_n)/\gamma_n^k)$$

$$- (1/n^2)\Sigma_{j=1}^{n}\text{var}(\sqrt{(n\gamma_n^k)}(g(X_j)-g(x))K((x-X_j)/\gamma_n)/\gamma_n^k)$$

$$= (n/\gamma_n^k)d_n(x) = O(\gamma_n^2\ln(n/\gamma_n^{k+3})) + o(1),$$

where the last conclusion follows from the fact that by assumption 12.1.3 and Taylor's theorem

$$E(g(X_j)-g(x))^2K((x-X_j)/\gamma_n)^2$$

$$= \gamma_n^k\int(g(x-\gamma_n z)-g(x))^2 h(x-\gamma_n z)K(z)^2 dz$$

$$\approx \gamma_n^{k+2}\int\{z'(\partial/\partial x')(g(x)^2h(x))\}^2K(z)^2dz = O(\gamma_n^{k+2}).$$

Since $\gamma_n^2\ln(n/\gamma_n^{k+3}) \to 0$ as $n \to \infty$ if γ_n is proportional to $n^{-\tau}$ with $\tau > 0$, (6.2.4) follows.

Step 3: Proof of (12.2.4). Let

$$\sigma_n^2(x) = E\, K((x-X_0)/\gamma_n)^2\gamma_n^{-k}.$$

Observe that similarly to (6.1.17),

$$\sigma_n^2(x) \to h(x)\int K(z)^2 dz.$$

Now let $Z_j = 1$ in lemma 12.1.1 and let x be such that $h(x) > 0$. Then

$$\begin{aligned} \mathrm{var}(\hat{h}(x)) &\le \gamma_n^{-2k} d_n(x) + \sigma_n^2(x)/(n\gamma_n^k) \\ &= O\{[\log(n/\gamma_n)+\log(\gamma_n^{-k})+1]/(n\gamma_n^k)\} \\ &= O[(\log(n\gamma_n^{-k-1}))/(n\gamma_n^k)] \to 0 \end{aligned}$$

if γ_n is proportional to $n^{-\tau}$ with $\tau < 1/k$. Since the latter condition is satisfied throughout section 6.2, the proof of theorem 12.2.1 for the continuous case is completed. Q.E.D.

Theorem 12.2.2. Let assumption 12.1.1 hold. Then:
(a) the conclusion of theorem 6.3.1 becomes:

$$(n/\log(n))^{m/(2m+2k)} \sup_{x\in\{x\in R^k : h(x)\ge\delta\}} |\hat{g}(x) - g(x)|$$

is stochastically bounded, with corresponding optimal window width of the form

$$\gamma_n = c\cdot(\log(n) / n)^{1/(2m+2k)};$$

(b) the conclusion of theorem 6.4.1 regarding uniform consistency becomes:

$$(n/\log(n))^{m/(2m+2k_1)} \sup_{x\in\{x\in R^k : h(x)\ge\delta\}} |\hat{g}(x) - g(x)|$$

is stochastically bounded, with corresponding optimal window width

$$\gamma_n = c\cdot(\log(n) / n)^{1/(2m+2k_1)},$$

where $h(x)$ is defined by (6.4.13).

Proof: Again we confine our attention to the continuous case, i.e., part (a). The only places in section 6.3 where the independence assumption has been employed are (6.3.6) and subsequently (6.3.7) and (6.3.10). From lemma 12.1.2 with $Z_j = Y_j$ it follows that in the present case (6.3.7) becomes:

$$E\ \sup_x |\hat{g}(x)\hat{h}(x) - E\ \hat{g}(x)\hat{h}(x)|$$

$$= \gamma_n^{-k} E\ \sup_x |a_n(x) - E\ a_n(x)| = O[(\log(n/\gamma_n^2))^{\frac{1}{2}}/(\gamma_n^k\sqrt{n})] \tag{12.2.7}$$

Combining (12.2.7) with (6.3.9), there exist constants $c_1 > 0$, $c_2 > 0$ such that (6.3.10) becomes:

$$E\ \sup_x |\hat{g}(x)\hat{h}(x) - g(x)h(x)|$$

$$= O(c_1(\log(n/\gamma_n^2))^{\frac{1}{2}}/(\gamma_n^k\sqrt{n}) + c_2\gamma_n^m)$$

$$= O(c_1(\log(n))^{\frac{1}{2}}/(\gamma_n^k\sqrt{n}) + c_2\gamma_n^m) \tag{12.2.8}$$

provided

$$\log(1/\gamma_n^2)\ /\ \log(n) \to 0. \tag{12.2.9}$$

Minimizing (12.2.8) to γ_n yields an optimal window width of the form $\gamma_n = c\cdot(\log(n)\ /\ n)^{1/(2m+2k)}$ for which indeed (12.2.9) holds. With this window width, (12.2.8) becomes

$$E\ \sup_x |\hat{g}(x)\hat{h}(x) - g(x)h(x)| = O[(\log(n)\ /\ n)^{m/(2m+2k)}].$$

This proves part (a) of the theorem. Q.E.D.

Finally, we have

Theorem 12.2.3. Under the conditions of theorems 12.2.1 and 12.2.2 and the additional assumption 12.1.4 the conclusions of theorem 6.5.1 carry over.

Proof: It follows straightforwardly from lemma 12.1.3 that (6.5.22) goes through, which proves the theorem. Q.E.D.

Remark 1: Along the same lines as in this section we can generalize the SMINK estimation approach in chapter 7 to time series.

Remark 2: We note that the φ-mixing condition on the base can be relaxed to the weaker α-mixing condition, but at the expense of stronger conditions on the moments of Y_j and X_j. This follows from theorem 9.2.1. This extension is left as an exercise.

Exercises:

1. Prove $E(Z_j^{(m)})^2 \leq E\, Z_j^2$. (Cf. lemma 12.1.1)
2. Prove theorem 12.2.1 for the discrete and mixed continuous-discrete case.
3. Prove part (b) of theorem 12.2.2.

References

Bierens,H.J. (1983), "Uniform consistency of kernel estimators of a regression function under generalized conditions", *Journal of the American Statistical Association* 77, 699-707.

Bierens,H.J. (1987), "Kernel Estimators of Regression Functions", in: Truman F.Bewley (ed.), *Advances in Econometrics 1985*, New York: Cambridge University Press, 99-144

Collomb,G. (1985), "Non-parametric time series analysis and prediction: Uniform almost sure convergence of the window and K-NN autoregression estimates", *Statistics* 16, 297-307

Georgiev,A.A. (1984), "Nonparametric system identification by kernel methods", *IEEE Transact. Aut.Cont.* 29, 356-358.

Robinson,P.M. (1983), "Nonparametric estimators for time series", *Journal of Time Series Analysis* 4, 185-207.