II Congresso Anual da Sociedade Portuguesa de Estatística
Luso, 20 a 23 de Junho de 1994

Um teste de adequação dum modelo paramétrico de regressão sob condições de dependência

Carlos TENREIRO†

Resumo: Propomos neste artigo, um teste de adequação dum modelo paramétrico de regressão baseado no desvio quadrático integrado entre dois estimadores do núcleo: um paramétrico e outro não paramétrico. Os resultados que apresentamos são obtidos utilizando técnicas de U-estatísticas degeneradas, geradas por um processo estocástico fortemente estacionário e geometricamente absolutamente regular.

Palavras chave : estimador do núcleo, desvio quadrático integrado, teste de adequação, U-estatística.
Classificação AMS 1991 : 62G07, 62G10.

1 Introdução

Seja $((X_i, Y_i), i \in \mathbb{Z})$ um processo estocástico fortemente estacionário, onde Y_i é de dimensão 1 e X_i é de dimensão d e admite uma lei de probabilidade absolutamente contínua relativamente à medida de Lebesgue. Denotamos por f a sua densidade.

A adequação dum modelo paramétrico de regressão,

$$Y_i = \mu(X_i;\theta) + u_i, \text{ com } E[u_i|X_i] = 0,$$

onde o parâmetro θ pertence a um subconjunto Θ de $\mathbb{R}^{\mathrm{K}}$, para $\mathrm{K} \geq 1$, será neste trabalho julgada pela comparação global, de dois estimadores da função $E[Y_0|X_0 = \cdot]f(\cdot)$: um paramétrico e outro não paramétrico. Tais estimadores são, para $x \in \mathbb{R}^d$, definidos respectivamente por

$$m_n(x;\widehat{\theta}_n) = \frac{1}{nh_n^d}\sum_{i=1}^{n}\mu(X_i;\widehat{\theta}_n)K\left(\frac{x-X_i}{h_n}\right) \text{ e } m_n(x) = \frac{1}{nh_n^d}\sum_{i=1}^{n}Y_iK\left(\frac{x-X_i}{h_n}\right),$$

† Departamento de Matemática, Universidade de Coimbra.
O presente trabalho foi parcialmente subsidiado pelo acordo JNICT-Embaixada de França.

onde $\widehat{\theta}_n$ é na hipótese de existir $\theta \in \Theta$ tal que $E[Y_0|X_0 = \cdot] = \mu(\cdot;\theta)$, P_{X_0} quase certamente (q.c.), uma sucessão de estimadores de θ, (h_n) é uma sucessão de números reais estritamente positivos que converge para zero quando n tende para infinito, e K é um núcleo em $\mathbb{R}^d$, isto é, uma aplicação mensurável de $\mathbb{R}^d$ em $\mathbb{R}$ cujo integral sobre $\mathbb{R}^d$ é igual a 1. O desvio quadrático integrado é usado como medida da aproximação entre estes dois estimadores:

$$\int_{\mathbb{R}^d} \left(m_n(x) - m_n(x;\widehat{\theta}_n)\right)^2 dx$$
$$= \frac{1}{n^2 h_n^d} \sum_{i,j=1}^{n} (Y_i - \mu(X_i;\widehat{\theta}_n))(Y_j - \mu(X_j;\widehat{\theta}_n))\, K * \bar{K}\left(\frac{X_i - X_j}{h_n}\right),$$

onde designamos por $K * \bar{K}$ o produto de convolução de K por $\bar{K}$, com $\bar{K}$ definido, para $u \in \mathbb{R}^d$, por $\bar{K}(u) = K(-u)$.

Esta mesma ideia é desenvolvida num artigo recente de Härdle e Mammen (1993), do qual tomámos conhecimento numa fase adiantada do presente trabalho, no caso de independência de observações e onde a estatística de teste proposta por estes autores é baseada no desvio quadrático integrado entre $m_n(\cdot)/f_n(\cdot)$ e $m_n(\cdot;\widehat{\theta}_n)/f_n(\cdot)$, sendo $f_n(\cdot)$ o estimador do núcleo da densidade comum dos X_i. Apesar dos esperados pontos comuns, o estudo que agora apresentamos, segue uma linha orientadora diferente. Assim, o estudo do teste proposto será efectuado com base na teoria das U-estatísticas degeneradas de segunda ordem geradas por um processo estocástico fortemente estacionário e geometricamente absolutamente regular, e em particular num teorema central limite que apresentámos em Tenreiro (1994).

Procedimentos de teste baseados na comparação de estimadores paramétricos e não paramétricos são utilizados em Azzalini, Bowman e Härdle (1989), Eubank e Spiegelman (1990) e em Härdle e Marron (1990). A utilização do desvio quadrático integrado na construção de testes da hipótese simples $E[Y_0|X_0 = \cdot] = \mu_0(\cdot)$ contra a alternativa $E[Y_0|X_0 = \cdot] \neq \mu_0(\cdot)$, onde μ_0 é fixa à partida, é considerada nos trabalhos de Konakov (1977) e Nadaraya (1983).

O teste de adequação que estudamos neste artigo, é como referimos baseado na comparação a zero duma estatística cuja forma geral é dada por

$$J_n(\widehat{\theta}_n) = \frac{1}{n^2 h_n^d} \sum_{i,j=1}^{n} (Y_i - \mu(X_i;\widehat{\theta}_n))(Y_j - \mu(X_j;\widehat{\theta}_n))\, \omega\left(\frac{X_i - X_j}{h_n}\right),$$

onde ω é um núcleo simétrico em $\mathbb{R}^d$. Para $\theta \in \Theta$, o estudo assintótico de $J_n(\theta)$ que efectuamos na Secção 2, conduz-nos sob a hipótese $E[Y_0|X_0 = \cdot] = \mu(\cdot;\theta)$ a um desenvolvimento do tipo $EJ_n(\theta)+(nh_n^{d/2})^{-1}Z_n$, onde a variável aleatória Z_n é assintoticamente normal e centrada. Tal facto leva-nos a corrigir a estatística natural $J_n(\theta)$ deste efeito de viés assintótico. Somos conduzidos assim a uma estatística da forma

$$T_n(\widehat{\theta}_n) = \frac{2}{nh_n^{d/2}} \sum_{1\leq j<i\leq n} (Y_i - \mu(X_i;\widehat{\theta}_n))(Y_j - \mu(X_j;\widehat{\theta}_n))\,\omega\left(\frac{X_i - X_j}{h_n}\right),$$

sobre a qual baseamos um teste, de nível assintótico α e assintoticamente convergente, de adequação do modelo de regressão proposto, que apresentamos na Secção 3. Notemos que $T_n(\widehat{\theta}_n)$ é simples de calcular não necessitando de integrações numéricas, o que já não acontece se utilizarmos a medida global de comparação entre os estimadores paramétrico $m_n(\cdot;\widehat{\theta}_n)$ e não paramétrico $m_n(\cdot)$, dada por $\int_{\mathbb{R}^d} \left(m_n(x) - m_n(x;\widehat{\theta}_n)\right)^2 \pi(x)dx$, onde π é uma função de peso. No entanto, os resultados que apresentamos são facilmente obtidos neste caso.

Na Secção 4, por forma a avaliar a potência assintótica do teste proposto, estudamos o comportamento assintótico da estatística $T_n(\theta_0)$ para $\theta_0 \in \Theta$ fixo, sob uma sucessão de processos estocásticos $\left((X_{in}, Y_{in}), i \in \mathbb{Z}\right)_{n\in\mathbb{N}}$ cuja correspondente sucessão de funções de regressão μ_n para $n \in \mathbb{N}$, é da forma

$$\mu_n(x) = \mu(x;\theta_0) + a_n\nu_n(x),\ x\in\mathbb{R}^d,$$

onde $(a_n)_{n\in\mathbb{N}}$ é uma sucessão de números reais positivos que tende para zero quando n tende para infinito, e a sucessão de funções $(\nu_n)_{n\in\mathbb{N}}$, converge num sentido a precisar, para uma função limite não identicamente nula. Para esta sucessão de hipóteses alternativas locais $H_{1n} : E[Y_0|X_0 = \cdot] = \mu_n(\cdot)$, encontramos para o teste proposto uma velocidade de separação maximal entre H_{1n} e $H_1 : E[Y_0|X_0 = \cdot] = \mu(\cdot;\theta_0)$, igual a $\frac{1}{\sqrt{n}\, h_n^{d/4}}$.

Terminamos esta introdução notando que os resultados que obtemos podem em particular ser utilizados para testar a adequação dum modelo de regressão do tipo

$$g(Z_{i+1},\ldots,Z_{i+s}) = \mu(Z_i;\theta) + u_i,\ \text{com } E[u_i|Z_i] = 0 \text{ e } \theta\in\Theta,$$

onde $(Z_i, i \in \mathbb{Z})$ é um processo estocástico fortemente estacionário com valores em $\mathbb{R}^d$ e g uma aplicação mensurável de $(\mathbb{R}^d)^s$ em $\mathbb{R}$, fixa à partida.

As demonstrações dos resultados expostos são apresentadas na Secção 5.

2 Lei assintótica de $J_n(\theta)$

As hipóteses que introduzimos de seguida são designadas por (P), (N) e (S), consoante dizem respeito ao processo observado, ao núcleo ou à sucessão (h_n), respectivamente. Estas hipóteses são análogas às consideradas em Tenreiro (1994).

Hipóteses sobre o processo (P)

Seja $((X_i, Y_i), i \in \mathbb{Z})$ um processo estocástico de dimensão $d+1$ fortemente estacionário e F_n^m, $n \leq m$, $n, m \in \mathbb{Z} \cup \{-\infty, +\infty\}$, a tribo gerada pelas variáveis aleatórias $(X_n, Y_n), \ldots, (X_m, Y_m)$. Suponhamos que o processo $(X_i, Y_i), i \in \mathbb{Z}$, é geometricamente absolutamente regular, isto é, admitamos que existem $C > 0$ e $\rho \in]0,1[$ tais que,

$$\beta(i) = E\Big[\sup_{A \in F_i^{+\infty}} |P(A|F_{-\infty}^0) - P(A)|\Big] \leq C\rho^i, \text{ para } i \in \mathbb{N}.$$

A introdução desta medida de dependência entre o passado $F_{-\infty}^0$ e o futuro $F_i^{+\infty}$ é atribuída a A.N. Kolmogorov por Volkonskiĭ e Rozanov (1959).

Supomos ainda que as leis dos vectores (X_i, X_0), $i \geq 1$, admitem versões das densidades $f_{(X_i,X_0)}$ que satisfazem as condições,

$$\sup_{x \in \mathbb{R}^d} f(x) < \infty \text{ e}$$

$$\sup_{i \in \mathbb{N}} \sup_{\substack{x \in \{u \in \mathbb{R}^d | f(u) > 0\} \\ y \in \mathbb{R}^d}} f_i(y|x) < \infty,$$

onde $f_i(y|x) = \frac{f_{(X_i,X_0)}(y,x)}{f(x)}$.

Hipóteses sobre o núcleo (N)

ω é um núcleo simétrico em $\mathbb{R}^d$, isto é, $\omega(-u) = \omega(u)$, para $u \in \mathbb{R}^d$, e $\int_{\mathbb{R}^d} \omega(u)du = 1$, que supomos ainda limitado de suporte limitado.

Hipóteses sobre a sucessão (h_n) (S)

Suponhamos que

$$nh_n^d \to +\infty, \text{ quando } n \to +\infty,$$

e que existe $\gamma \in]0,1[$ tal que

$$\limsup_n n^\gamma h_n^d < \infty.$$

Seja $\mu(\cdot;\theta), \theta \in \Theta$ uma família paramétrica de aplicações mensuráveis de $\mathbb{R}^d$ em $\mathbb{R}$, onde $\Theta \subset \mathbb{R}^K$, para $K \geq 1$. Especificamos no Teorema seguinte a lei assintótica da variável aleatória

$$J_n(\theta) = \frac{1}{n^2 h_n^d} \sum_{i,j=1}^n (Y_i - \mu(X_i;\theta))(Y_j - \mu(X_j;\theta))\, \omega\left(\frac{X_i - X_j}{h_n}\right),$$

para $\theta \in \Theta$, na hipótese de se ter $E[Y_0|X_0 = \cdot] = \mu(\cdot;\theta)$, P_{X_0} $q.c.$.

Teorema 1 *Suponhamos satisfeitas as hipóteses* (P), (N), *e* (S), *e admitamos que existe* $\delta > 0$ *tal que* $E|Y_0|^{8+\delta} < \infty$. *Se para algum* $\theta \in \Theta$, $E[Y_0|X_0 = \cdot] = \mu(\cdot;\theta)$, P_{X_0} *q.c., então*

$$nh_n^{d/2}\{J_n(\theta) - EJ_n(\theta)\},$$

é assintoticamente normal de média zero e de variância $2\tau^2$, *onde*

$$\tau^2 = \int_{\mathbb{R}^d} \sigma^4(x) f^2(x)dx \int_{\mathbb{R}^d} \omega^2(z)dz,$$

com $\sigma^2(\cdot) = Var[Y_0|X_0 = \cdot] \neq 0$. *Além disso,*

$$EJ_n(\theta) = \frac{\omega(0)}{nh_n^d}\frac{1}{n}\sum_{i=1}^{n}(Y_i - \mu(X_i;\theta))^2 + o_p\Big(\frac{1}{nh_n^{d/2}}\Big),$$

e o parâmetro τ^2 *pode ser estimado em probabilidade por*

$$\tau_n^2(\theta) = \frac{2}{n^2h_n^d}\sum_{1\leq j<i\leq n}(Y_i - \mu(X_i;\theta))^2(Y_j - \mu(X_j;\theta))^2\omega^2\Big(\frac{X_i - X_j}{h_n}\Big).$$

3 O teste da hipótese H_0 de adequação dum modelo paramétrico de regressão

Seja $\{\mu(\cdot;\theta) : \theta \in \Theta\}$ a família paramétrica de curvas considerada na secção anterior, onde Θ é um aberto convexo de $\mathbb{R}^K$. Propomos nesta secção um teste de adequação do modelo de regressão

$$Y_i = \mu(X_i;\theta) + u_i, \text{ com } E[u_i|X_i] = 0 \text{ e } \theta \in \Theta, \tag{3.1}$$

ou seja, um teste da hipótese

$$H_0 : \exists\theta_0 \in \Theta, P_{X_0}(\{x : E[Y_0|X_0 = x] = \mu(x;\theta_0)\}) = 1,$$

contra

$$H_0^c : \forall\theta \in \Theta, P_{X_0}(\{x : E[Y_0|X_0 = x] \neq \mu(x;\theta_0)\}) > 0,$$

baseado na estatística $T_n(\widehat{\theta}_n)$, onde para $\theta \in \Theta$, a variável aleatória

$$T_n(\theta) = \frac{2}{nh_n^{d/2}}\sum_{1\leq j<i\leq n}(Y_i - \mu(X_i;\theta))(Y_j - \mu(X_j;\theta))\,\omega\Big(\frac{X_i - X_j}{h_n}\Big),$$

é obtida por correcção da estatística natural $J_n(\theta)$ do efeito de viés assintótico dado no Teorema 1, e $\widehat{\theta}_n \in \Theta$ é uma sucessão de estimadores do parâmetro desconhecido θ_0.

É no entanto necessário introduzir algumas hipóteses sobre o modelo de regressão considerado bem como sobre a sucessão de variáveis aleatórias $\widehat{\theta}_n$.

Hipóteses sobre o modelo paramétrico de regressão (M)

Para todo o $x \in \mathbb{R}^d$, supomos que a aplicação $\theta \longrightarrow \mu(x;\theta)$ admite derivadas parciais de segunda ordem contínuas em Θ, com

$$E\Big[\sup_{\theta\in\Theta}|\mu(X_0;\theta)|\Big]^{8+\delta} < \infty,$$

$$E\Big[\sup_{\theta\in\Theta}\Big|\frac{\partial\mu}{\partial\theta_k}(X_0;\theta)\Big|\Big]^{8+\delta} < \infty, \text{ e } E\Big[\sup_{\theta\in\Theta}\Big|\frac{\partial^2\mu}{\partial\theta_k\partial\theta_{k'}}(X_0;\theta)\Big|\Big]^{4+\delta} < \infty,$$

para algum $\delta > 0$ e $k, k' = 1, \ldots, \mathrm{K}$.

Hipóteses sobre a sucessão de estimadores $\widehat{\theta}_n$ (E)

Admitimos que $\widehat{\theta}_n$ é uma sucessão de variáveis aleatórias que converge em probabilidade para um elemento de Θ, e que sob a hipótese H_0 satisfaz

$$\widehat{\theta}_n - \theta_0 = o_p\Big(\frac{1}{\sqrt{n}\, h_n^{d/4}}\Big).$$

Supondo que o parâmetro θ do modelo de regressão (3.1) é identificável, isto é, $\mu(\cdot;\theta_1) = \mu(\cdot;\theta_2)$, $P_{X_0}\, q.c. \Rightarrow \theta_1 = \theta_2$, podemos em particular e sob certas condições de regularidade, considerar $\widehat{\theta}_n$ o estimador dos mínimos quadrados não lineares ponderados de θ, definido como uma solução do problema de optimização

$$min_\theta \sum_{i=1}^{n}\big(Y_i - \mu(X_i;\theta)\big)^2 a(X_i),$$

onde a é uma dada função estritamente positiva [cf. Gouriéroux e Monfort (1989), vol. I, pg. 236-251]. Tal não é no entanto admitido.

Designemos por ϕ a função de repartição da lei normal centrada e reduzida sobre $\mathbb{R}$.

Teorema 2 *Admitamos satisfeitas as hipóteses* (P), (N), (S), (M) *e* (E), *e suponhamos que* $E|Y_0|^{8+\delta} < \infty$, *para algum* $\delta > 0$. *Então,*

$$\left\{\frac{T_n(\widehat{\theta}_n)}{\sqrt{2\tau_n^2(\widehat{\theta}_n)}} \geq \phi^{-1}(1-\alpha)\right\},$$

é a região critica dum teste de nível assintótico α *e assintoticamente convergente para testar* H_0 *contra* H_0^c.

4 Potência local do teste

Designando por θ_0 o verdadeiro valor do parâmetro θ sob a hipótese nula, a potência assintótica do teste proposto é obtida a partir do comportamento assintótico da estatística $\dfrac{T_n(\theta_0)}{\sqrt{2\tau_n^2(\theta_0)}}$ sob uma sucessão de processos $\left((X_{in}, Y_{in}), i \in \mathbb{Z}\right)_{n\in\mathbb{N}}$, cuja correspondente sucessão de funções de regressão $\mu_n(\cdot), n \in \mathbb{N}$, converge para $\mu(\cdot;\theta_0) \equiv \mu_0(\cdot)$, quando n tende para infinito. Consideremos as seguintes

Hipóteses sobre a sucessão de alternativas locais (AL)

Seja $\left((X_{in}, Y_{in}), i \in \mathbb{Z}\right)_{n\in\mathbb{N}}$ uma sucessão de processos fortemente estacionários de dimensão $d+1$, cujos coeficientes de mistura de tipo β, convergem para zero de forma exponencial, uniformemente em relação a n. Para $n\in\mathbb{N}$, seja f_n a densidade marginal do processo $X_{in}, i \in \mathbb{Z}$ e $f_{in}(\cdot|x)$ a densidade da lei condicional de $X_{in}|X_{0n} = x$, que supomos satisfazerem as condições

$$\sup_{n\in\mathbb{N}} \sup_{x\in\mathbb{R}^d} f_n(x) < \infty, \text{ e}$$

$$\sup_{n,i\in\mathbb{N}} \sup_{\substack{x\in\{u\in\mathbb{R}^d | f_n(u)>0\} \\ y\in\mathbb{R}^d}} f_{in}(y|x) < \infty.$$

Suponhamos então que a sucessão de alternativas locais $\mu_n(\cdot), n \in \mathbb{N}$, é da forma

$$\mu_n(x) = \mu_0(x) + a_n\nu_n(x), \text{ para } x\in\mathbb{R}^d, \tag{4.1}$$

onde (a_n) é uma sucessão de números reais que tende para zero quando n tende para infinito. Admitamos que existe uma função γ de quadrado integrável e não identicamente nula tal que

$$\int_{\mathbb{R}^d} |\nu_n f_n(x) - \gamma(x)|^2 dx \to 0, \text{ quando } n\to+\infty, \tag{4.2}$$

onde

$$\sup_{n\in\mathbb{N}} \sup_{x\in\mathbb{R}^d} |\nu_n f_n(x)| < \infty. \tag{4.3}$$

Supomos ainda que existe uma função η de quadrado integrável tal que

$$\int_{\mathbb{R}^d} |\sigma_n^2 f_n(x) - \eta(x)|^2 dx \to 0, \text{ quando } n\to+\infty, \tag{4.4}$$

onde $\sigma_n^2(\cdot) = Var[Y_{0n}|X_{0n} = \cdot]$.

A potência assintótica do teste proposto é descrita no teorema seguinte. Sejam para $n \in \mathbb{N}$,

$$T_{nn} = \frac{2}{nh_n^{d/2}} \sum_{1\le j<i\le n} (Y_{in} - \mu_0(X_{in}))(Y_{jn} - \mu_0(X_{jn}))\,\omega\left(\frac{X_{in} - X_{jn}}{h_n}\right), \text{ e}$$

$$\tau_{nn}^2 = \frac{2}{n^2 h_n^d} \sum_{1 \leq j < i \leq n} (Y_{in} - \mu_0(X_{in}))^2 (Y_{jn} - \mu_0(X_{jn}))^2 \, \omega^2\left(\frac{X_{in} - X_{jn}}{h_n}\right).$$

Teorema 3 *Suponhamos satisfeitas as hipóteses* (AL), (N) *e* (S) *e que para algum* $\delta > 0$, $\sup_{n \in \mathbb{N}} E[|Y_{0n}| + |\mu_0(X_{0n})|]^{8+\delta} < \infty$. *Se*

$$a_n = \frac{1}{\sqrt{n} h_n^{d/4}},$$

então $T_{nn}/\sqrt{2\tau_{nn}^2}$ *é assintoticamente normal de média*

$$\left\{\int_{\mathbb{R}^d} \gamma^2(x)dx\right\}\left\{2\int_{\mathbb{R}^d} \eta^2(x)dx \int_{\mathbb{R}^d} \omega^2(z)dz\right\}^{-\frac{1}{2}},$$

e de variância 1. Se $a_n = o\Big(\frac{1}{\sqrt{n} h_n^{d/4}}\Big)$, *então* $T_{nn}/\sqrt{2\tau_{nn}^2}$ *é assintoticamente normal de média zero e variância 1. Se* $\frac{1}{\sqrt{n} h_n^{d/4}} = o(a_n)$, *então* $T_{nn}/\sqrt{2\tau_{nn}^2} \to +\infty$, *em probabilidade, quando* $n \to +\infty$.

Nas condições do teorema anterior a potência assintótica do teste é, para $a_n = \frac{1}{\sqrt{n} h_n^{d/4}}$, igual a

$$\begin{aligned} &\lim_{n \to +\infty} P\Big(\frac{T_{nn}}{\sqrt{2\tau_{nn}^2}} \geq \phi^{-1}(1-\alpha)\Big) \\ &= 1 - \phi\Big(\phi^{-1}(1-\alpha) - \left\{\int_{\mathbb{R}^d} \gamma^2(x)dx\right\}\left\{2\int_{\mathbb{R}^d} \eta^2(x)dx \int_{\mathbb{R}^d} \omega^2(z)dz\right\}^{-\frac{1}{2}}\Big). \end{aligned}$$

Tendo em conta a expressão da potência assintótica, observemos que:

(i) A potência assintótica do teste proposto é uma função crescente de $\int_{\mathbb{R}^d} \gamma^2(x)dx$. A positividade estrita desta quantidade para toda a função γ não identicamente nula, permite concluir que o teste é uniformemente estritamente cêntrico [cf. Bickel e Rosenblatt (1973), pg. 1082].

(ii) Sendo a potência assintótica uma função decrescente de $I(\omega) = \int_{\mathbb{R}^d} \omega^2(z)dz$, o núcleo ω deverá ser escolhido por forma a minimizar a funcional I. Designando por $\Omega(\sigma)$, com $\sigma > 0$, a classe dos núcleos em $\mathbb{R}^d$ da forma $\omega(x) = \prod_{i=1}^d \omega_0(x_i)$, para $x = (x_1, \ldots, x_d) \in \mathbb{R}^d$, onde ω_0 é um núcleo em $\mathbb{R}$, positivo, simétrico, limitado de suporte limitado, e de variância σ^2, então [cf. Epanechnikov (1969) e Bosq e Lecoutre (1987)]

$$\min_{\omega \in \Omega(\sigma)} I(\omega) = I(\omega^*) = \Big(\frac{3}{5\sigma\sqrt{5}}\Big)^d,$$

onde

$$\omega^*(x_1, \ldots, x_d) = \Big(\frac{3}{4\sigma\sqrt{5}}\Big)^d \prod_{i=1}^d \Big(1 - \frac{x_i^2}{5\sigma^2}\Big) \mathbb{I}_{\{|x_i| \leq \sigma\sqrt{5}\}}(x_i).$$

(iii) O termo $2\int_{\mathbb{R}^d}\eta^2(x)dx$ está relacionado com a variância assintótica da estatística de teste que estimámos por $2\tau_n^2(\widehat{\theta}_n)$, onde $\widehat{\theta}_n$ satisfaz as hipóteses descritas em (E). Tal facto leva-nos a propôr o estudo, que julgamos não realizado, das propriedades assintóticas do estimador

$$\widehat{\theta}_n = Arg\, min_\theta\, \tau_n^2(\theta).$$

5 Demonstrações

5.1 Demonstração do Teorema 1

i) Normalidade assintótica de $nh_n^{d/2}\{J_n(\theta)-EJ_n(\theta)\}$

Para $\theta\in\Theta$ nas condições do enunciado, temos

$$nh_n^{d/2}\{J_n(\theta)-EJ_n(\theta)\} = \frac{2}{n}\sum_{1\le j<i\le n}\{H_n(Z_i,Z_j)-EH_n(Z_i,Z_j)\}+o_p(1),$$

onde $Z_i=(X_i,Y_i), i\in\mathbb{Z}$, e para $u=(u_1,u_2)$, $v=(v_1,v_2)\in\mathbb{R}^{d+1}$,

$$H_n(u,v)=\frac{1}{h_n^{d/2}}(u_2-\mu(u_1;\theta))(v_2-\mu(v_1;\theta))\,\omega\left(\frac{u_1-v_1}{h_n}\right).$$

Como $E[Y_0|X_0=\cdot]=\mu(\cdot;\theta), P_{X_0}, q.c.$, a variável aleatória em estudo é assintoticamente equivalente a uma U-estatística degenerada de segunda ordem, cuja normalidade assintótica obtemos como aplicação do Teorema 2.1 de Tenreiro (1994). Sejam para $r>0$ e $n\in\mathbb{N}$, $u_n(r)$, $v_n(r)$, $w_n(r)$ e $z_n(r)$ as sucessões aí definidas relativamente ao processo $Z_i, i\in\mathbb{Z}$. Tendo em conta que $E|Y_0|^{8+\delta}<\infty$, e pelas hipóteses (P), (N) e (S), concluímos que para $r\ge 1-\frac{2}{10+\delta}$ existe $C>0$ tal que, para $n\in\mathbb{N}$ se tem: $u_n(r)\le C(h_n^d)^{\frac{1}{r}-\frac{2}{8+\delta}-\frac{1}{2}}$, se $r<4+\frac{\delta}{2}$, e $v_n(r)\le C(h_n^d)^{\frac{1}{r}-\frac{2}{8+\delta}}$, $w_n(r)\le C$ e $z_n(r)\le C(h_n^d)^{\frac{4+\delta}{8+\delta}}$, se $r<2+\frac{\delta}{4}$. Da hipótese (S), as quatro primeiras condições do Teorema 2.1 de Tenreiro (1994) são então satisfeitas. Precisemos agora a forma da variância assintótica. Temos

$$E[H_n(Z_0,\overline{Z}_0)]^2=\frac{1}{h_n^d}\int_{\mathbb{R}^d}(\sigma^2 f*\overline{\sigma^2 f})(v)\,\omega^2\left(\frac{v}{h_n}\right)dv,$$

onde $\overline{Z}_0$ e Z_0 são variáveis aleatórias independentes e identicamente distribuídas, e para $F:\mathbb{R}^d\to\mathbb{R}$, $\overline{F}(u)=F(-u)$, com $u\in\mathbb{R}^d$. O lema de Bochner [cf. Bosq e Lecoutre (1987), pg. 61] permite agora concluir, tendo em conta a continuidade do produto de convolução de funções de quadrado integrável, que $nh_n^{d/2}\{J_n(\theta)-EJ_n(\theta)\}$ converge, quando n tende para infinito, para uma lei normal de média zero e variância $2\tau^2 = 2\int_{\mathbb{R}^d}\sigma^4(x)f^2(x)dx\int_{\mathbb{R}^d}\omega^2(z)dz$.

ii) Desenvolvimento assintótico de $EJ_n(\theta)$

Temos

$$\begin{aligned} EJ_n(\theta) &= \frac{\omega(0)}{nh_n^d} E\big(Y_0 - \mu(X_0;\theta)\big)^2 + \frac{1}{nh_n^{d/2}} \frac{2}{n} \sum_{1\le j<i\le n} EH_n(Z_i, Z_j) \\ &= \frac{\omega(0)}{nh_n^d} E\big(Y_0 - \mu(X_0;\theta)\big)^2 + o\Big(\frac{1}{nh_n^{d/2}}\Big), \end{aligned}$$

pois pelo Lema 1 de Yoshihara (1976), $\left|\frac{2}{n}\sum_{1\le j<i\le n} EH_n(Z_i,Z_j)\right| = o(1)$ [ver demonstração do Teorema 4.1 de Tenreiro (1994)]. O desenvolvimento pretendido obtem-se agora por aplicação do teorema central limite.

iii) Convergência em probabilidade de $\tau_n^2(\theta)$

Temos

$$\tau_n^2(\theta) = \frac{1}{nh_n^{d/2}} \frac{2}{n} \sum_{1\le j<i\le n} U_n(Z_i, Z_j),$$

onde para $u = (u_1, u_2)$, $v = (v_1, v_2) \in \mathbb{R}^{d+1}$,

$$U_n(u,v) = \frac{1}{h_n^{d/2}} (u_2 - \mu(u_1;\theta))^2 (v_2 - \mu(v_1;\theta))^2 \, \omega^2\left(\frac{u_1 - v_1}{h_n}\right).$$

Consideremos agora a decomposição seguinte [cf. Yoshihara (1989), pg. 111], cujo primeiro termo é uma U-estatística degenerada de segunda ordem:

$$\begin{aligned} nh_n^{d/2}\tau_n^2(\theta) &= \frac{2}{n} \sum_{1\le j<i\le n} \{U_n^*(Z_i,Z_j) - EU_n^*(Z_i,Z_j)\} \\ &\quad + (1 - \frac{1}{n}) \sum_{i=1}^{n} \{V_n(Z_i) - EV_n(Z_i)\} + \frac{2}{n} \sum_{i=1}^{n-1} (n-i) EU_n(Z_i, Z_0), \end{aligned}$$

onde $V_n(u) = EU_n(u, Z_0)$ e $U_n^*(u,v) = U_n(u,v) - V_n(u) - V_n(v) - EV_n(Z_0)$. Sendo o primeiro termo da decomposição anterior um $O_p(1)$, e o segundo um $O_p(\sqrt{n}\, h_n^{d/2})$, concluímos pelo Lema 1 de Yoshihara (1976) que,

$$\tau_n^2(\theta) = \frac{n-1}{n} \frac{1}{h_n^{d/2}} EU_n(Z_0, \overline{Z}_0) + o_p(1).$$

A convergência desejada obtem-se agora do ponto i), pois

$$\lim_{n\to+\infty} \frac{1}{h_n^{d/2}} EU_n(Z_0, \overline{Z}_0) = \lim_{n\to+\infty} E[H_n(Z_0, \overline{Z}_0)]^2 = \tau^2.$$

5.2 Demonstração do Teorema 2

i) Normalidade assintótica da estatística de teste sob a hipótese nula

Suponhamos que existe $\theta_0 \in \Theta$, tal que $E[Y_0|X_0 = \cdot] = \mu(\cdot;\theta_0), P_{X_0}$ *q.c.*. Pelo Teorema 1, basta mostrar que $T_n(\widehat{\theta}_n)$ e $T_n(\theta_0)$, bem como $\tau_n^2(\widehat{\theta}_n)$ e $\tau_n^2(\theta_0)$, são assintoticamente equivalentes, o que faremos nos pontos seguintes.

i$_1$) $T_n(\widehat{\theta}_n) - T_n(\theta_0) = o_p(1)$

Um desenvolvimento de Taylor permite concluir que para $i = 1, \ldots, n$, existe $\alpha \in]0,1[$ tal que,

$$\begin{aligned} Y_i - \mu(X_i;\widehat{\theta}_n) &= Y_i - \mu(X_i;\widehat{\theta}_n) - \sum_{k=1}^{K}(\widehat{\theta}_n - \theta_0)_k \frac{\partial \mu}{\partial \theta_k}(X_i;\theta_0) \\ &\quad -\frac{1}{2}\sum_{k,k'=1}^{K}(\widehat{\theta}_n - \theta_0)_k(\widehat{\theta}_n - \theta_0)_{k'}\frac{\partial^2 \mu}{\partial \theta_k \partial \theta_{k'}}(X_i;\theta_0 + \alpha(\widehat{\theta}_n - \theta_0)), \end{aligned}$$

ou ainda, pela hipótese (M),

$$T_n(\widehat{\theta}_n) - T_n(\theta_0) = -\sum_{k=1}^{K}(\widehat{\theta}_n - \theta_0)_k \frac{2}{n}\sum_{1 \le j < i \le n} U_n^k(Z_i, Z_j) + O_p\left(\|\widehat{\theta}_n - \theta_0\|^2 n h_n^{d/2}\right),$$

onde para $k = 1, \ldots, K$,

$$U_n^k(u,v) = \frac{1}{h_n^{d/2}}\left\{(u_2 - \mu(u_1;\theta_0))\frac{\partial \mu}{\partial \theta_k}(v_1;\theta_0) + (v_2 - \mu(v_1;\theta_0))\frac{\partial \mu}{\partial \theta_k}(u_1;\theta_0)\right\}\omega\left(\frac{u_1 - v_1}{h_n}\right),$$

para $u = (u_1, u_2)$ e $v = (v_1, v_2) \in \mathbb{R}^{d+1}$. Para $k = 1, \ldots, K$, $\frac{2}{n}\sum_{1 \le j < i \le n} U_n^k(Z_i, Z_j)$ é uma U-estatística de segunda ordem não degenerada que, procedendo como no ponto iii) da demonstração do Teorema 1, pode ser decomposta na soma de três termos, sendo o mais significativo da ordem de $\sqrt{n}\, h_n^{d/4}$, atendendo uma vez mais à hipótese (M) e ao facto de $EU_n^k(Z_0, \overline{Z}_0) = 0$.

Assim, como pela hipótese (E), $\sqrt{n}\, h_n^{d/4}(\widehat{\theta}_n - \theta_0) = o_p(1)$, obtemos

$$T_n(\widehat{\theta}_n) - T_n(\theta_0) = O_p\left(\|\widehat{\theta}_n - \theta_0\|\sqrt{n}\, h_n^{d/4}\right) + O_p\left(\|\widehat{\theta}_n - \theta_0\|^2 n h_n^{d/2}\right) = o_p(1).$$

i$_2$) $\tau_n^2(\widehat{\theta}_n) - \tau_n^2(\theta_0) = o_p(1)$

Um novo desenvolvimento de Taylor permite escrever, para $i = 1, \ldots, n$,

$$Y_i - \mu(X_i;\widehat{\theta}_n) = Y_i - \mu_0(X_i;\theta_0) - \sum_{k=1}^{K}(\widehat{\theta}_n - \theta_0)_k \frac{\partial \mu}{\partial \theta_k}(X_i;\theta_0 + \alpha(\widehat{\theta}_n - \theta_0)),$$

para algum $\alpha\in]0,1[$. De forma análoga ao ponto anterior obtemos,

$$\tau_n^2(\widehat{\theta}_n)-\tau_n^2(\theta_0)=O_p\big(\|\widehat{\theta}_n-\theta_0\|\big)=o_p(1).$$

ii) Convergência do teste

Suponhamos que para todo o $\theta\in\Theta$, $P_{X_0}\big(\{x: E[Y_0|X_0=x]\neq\mu(x;\theta)\}\big)>0$. Pela hipótese (E), seja $\widetilde{\theta}\in\Theta$ tal que $\lim_n\widehat{\theta}_n=\widetilde{\theta}$, em probabilidade. Para $\theta_0=\widetilde{\theta}$, o desenvolvimento obtido em i_2) permite escrever $T_n(\widehat{\theta}_n)/nh_n^{d/2}$ sob a forma

$$\frac{1}{n^2h_n^d}\sum_{1\le j<i\le n}(Y_i-\mu(X_i;\widetilde{\theta}))(Y_j-\mu(X_j;\widetilde{\theta}))\,\omega\Big(\frac{X_i-X_j}{h_n}\Big)+O_p\big(\|\widehat{\theta}_n-\widetilde{\theta}\|\big).$$

Procedendo como no ponto iii) da demonstração do Teorema 1, concluímos que

$$\frac{1}{nh_n^{d/2}}T_n(\widehat{\theta}_n)=\int_{\mathbb{R}^d}\big(E[Y_0|X_0=x]-\mu(x;\widetilde{\theta})\big)^2f^2(x)dx+o_p(1),$$

e assim, $T_n(\widehat{\theta}_n)\longrightarrow+\infty$, em probabilidade, quando n tende para infinito.

De forma análoga concluiriamos que $\tau_n^2(\widehat{\theta}_n)$ converge em probabilidade para uma constante estritamente positiva, quando n tende para infinito, e portanto, $\frac{T_n(\widehat{\theta}_n)}{\sqrt{2\tau_n^2(\widehat{\theta}_n)}}$, converge em probabilidade para $+\infty$, o que implica a convergência assintótica do teste.

5.3 Demonstração do Teorema 3

Consideremos a decomposição (ver ponto iii) da demonstração do Teorema 1),

$$\begin{aligned}T_{nn}&=\frac{2}{n}\sum_{1\le j<i\le n}\{H_n^*(Z_{in},Z_{jn})-EH_n^*(Z_{in},Z_{jn})\}\\&\quad+(1-\frac{1}{n})\sum_{i=1}^{n}\{G_n(Z_{in})-EG_n(Z_{in})\}+\frac{2}{n}\sum_{i=1}^{n-1}(n-i)EH_n(Z_{in},Z_{0n}),\end{aligned}$$

onde $Z_{in}=(X_{in},Y_{in})$, $i\in\mathbb{Z}$, e para $u=(u_1,u_2)$, $v=(v_1,v_2)\in\mathbb{R}^{d+1}$,

$$H_n(u,v)=\frac{1}{h_n^{d/2}}(u_2-\mu_0(u_1))(v_2-\mu_0(v_1))\,\omega\Big(\frac{u_1-v_1}{h_n}\Big),$$

$G_n(u)=EH_n(u,Z_0)$ e $H_n^*(u,v)=H_n(u,v)-G_n(u)-G_n(v)-EG_n(Z_0)$.

Cada um dos termos da decomposição anterior é estudado de seguida.

i) Estudo do primeiro termo

Temos

$$E[H_n^*(Z_{0n},\overline{Z}_{0n})]^2=\frac{1}{h_n^d}\int_{\mathbb{R}^d}(\sigma'^2_nf_n*\overline{\sigma'^2_nf_n})(z)\,\omega^2\Big(\frac{z}{h_n}\Big)dz,$$

onde ${\sigma'}_n^2(\cdot) = E[(Y_{0n} - \mu_0(X_{0n}))^2 | X_{0n} = \cdot]$. Tendo em conta que a convergência L_2 de $\sigma_n^2 f_n$ para η (hipótese (4.4)), implica a de ${\sigma'}_n^2 f_n$ para η, obtemos como aplicação da proposição seguinte que,

$$\lim_{n\to+\infty} E[H_n^*(Z_{0n}, \overline{Z}_{0n})]^2 = \int_{\mathbb{R}^d} \eta^2(x)dx \int_{\mathbb{R}^d} \omega^2(z)dz.$$

Proposição 5.1 *Sejam $(\alpha_n)_{n\in\mathbb{N}}$ uma sucessão de aplicações mensuráveis de $\mathbb{R}^d$ em $\mathbb{R}$, e x um ponto de $\mathbb{R}^d$. Suponhamos que existe $\alpha \in \mathbb{R}$, tal que,*

$$\forall \epsilon > 0\ \exists \delta > 0\ \exists n_0 \in \mathbb{N}^* : n \geq n_0, \sup_{\|u-x\|<\delta} |\alpha_n(u) - \alpha| < \epsilon.$$

Seja ainda φ uma aplicação integrável de $\mathbb{R}^d$ em $\mathbb{R}$, de suporte limitado. Se $h_n \to 0, n \to +\infty$, então

$$\lim_{n\to+\infty} \frac{1}{h_n^d} \int_{\mathbb{R}^d} \alpha_n(u) \varphi\left(\frac{x-u}{h_n}\right) du = \alpha \int_{\mathbb{R}^d} \varphi(u)du.$$

Tal como no ponto i) da demonstração do Teorema 1, tendo em conta as hipóteses (AL), (N) e (S), e que $\sup_{n\in\mathbb{N}} E[|Y_{0n}| + |\mu_0(X_{0n})|]^{8+\delta} < \infty$, concluímos que o termo em estudo é assintoticamente normal de média zero e de variância $2\int_{\mathbb{R}^d} \eta^2(x)dx \int_{\mathbb{R}^d} \omega^2(z)dz$.

ii) <u>Estudo do segundo termo</u>

Como para $u = (u_1, u_2) \in \mathbb{R}^{d+1}$,

$$G_n(u) = a_n h_n^{d/2}(u_2 - \mu_0(u_1)) \frac{1}{h_n^d} \int_{\mathbb{R}^d} \nu_n f_n(x)\, \omega\left(\frac{u_1 - x}{h_n}\right) dx,$$

concluímos pela hipótese (4.3) que, $E|G_n(Z_{0n})|^2 = O(a_n^2 h_n^d)$, e então o termo em estudo é um $O_p(\sqrt{n}\, a_n h_n^{d/2})$.

iii) <u>Estudo do terceiro termo</u>

Pelo Lema 1 de Yoshihara (1976), concluímos que o terceiro termo é assintoticamente equivalente a $nE[H_n(Z_{0n}, \overline{Z}_{0n})]$, onde pela hipótese (4.2) e pela Proposição 5.1,

$$\begin{aligned} E[H_n(Z_{0n}, \overline{Z}_{0n})] &= h_n^{d/2} a_n^2 \frac{1}{h_n^d} \int_{\mathbb{R}^d} (\nu_n f_n * \overline{\nu_n f_n})(z)\, \omega\left(\frac{z}{h_n}\right) dz \\ &= h_n^{d/2} a_n^2 \int_{\mathbb{R}^d} \gamma^2(x)dx + o(h_n^{d/2} a_n^2). \end{aligned}$$

Dos pontos anteriores, e tendo em conta que de forma análoga à do ponto iii) da demonstração do Teorema 1, τ_{nn}^2 converge em probabilidade para $\int_{\mathbb{R}^d} \eta^2(x)dx \int_{\mathbb{R}^d} \omega^2(z)dz$, quando n tende para infinito, concluímos que

$$\begin{aligned}\frac{T_{nn}}{\sqrt{2\tau_{nn}^2}} &= Z_n + nh_n^{d/2}a_n^2\Big\{\int_{\mathbb{R}^d}\gamma^2(x)dx\Big\}\Big\{2\int_{\mathbb{R}^d}\eta^2(x)dx\int_{\mathbb{R}^d}\omega^2(z)dz\Big\}^{-1/2}\\ &\quad +o\big(nh_n^{d/2}a_n^2\big) + O_p\big(\sqrt{n}\,a_nh_n^{d/2}\big),\end{aligned}$$

onde Z_n é assintoticamente normal centrada e reduzida. O desenvolvimento anterior permite obter o resultado pretendido.

Agradecimentos: Queremos deixar expresso o nosso agradecimento sincero ao Professor Christian Gouriéroux, pelas sugestões e comentários feitos sobre uma versão preliminar do presente trabalho, aquando da nossa estada no Centre de Recherche en Economie et Statistique (Paris) em Janeiro de 1994.

Bibliografia

Azzalini, A., Bowman, A. e Härdle, W. (1989). On the use of nonparametric regression for model checking. *Biometrika*, **76**, 1-12.

Bosq, D. e Lecoutre, J-P. (1987). *Théorie de l'Estimation Fonctionnelle*. Economica, Paris.

Bradley, R.C. (1986). Basic properties of strong mixing conditions. *In: Dependence in Probability and Statistics, a survey of recent results*, ed.: Eberlein, E. e Taqqu, M., Birkhäuser, Boston, 165-192.

Epanechnikov, V. (1969). Nonparametric estimates of a multivariate probability density. *Theory Probab. Appl.*, **14**, 153-158.

Eubank, R.L. e Spiegelman, C.H. (1990). Testing the goodness of fit of a linear model via nonparametric regression techniques. *J. Amer. Statist. Assoc.*, **85**, 387-392.

Gouriéroux, C. e Monfort, A. (1989). *Statistique et Modèles Econométriques*. Economica, Paris.

Hall, P. (1984). Central limit theorem for integrated square error properties of multivariate nonparametric density estimators. *J. Multivariate Anal.*, **14**, 1-16.

Härdle, W. e Mammen, E. (1993). Comparing nonparametric versus parametric regression fits. *Ann. Statist.*, **21**, 1926-1947.

Härdle, W. e Marron, J.S. (1990). Semiparametric comparison of regression curves. *Ann. Statist.*, **18**, 63-89.

Konakov, V.D. (1977). On a global measure of deviation for an estimate of the regression line. *Theory Probab. Appl.*, 22, 858-868.

Nadaraya, E.A. (1983). A limit distribution of the square error deviation of nonparametric estimators of the regression function. *Z. Wahrsch. Verw. Gebiete*, 64, 37-48.

Tenreiro, C. (1994). O teste de ajustamento de Bickel-Rosenblatt generalizado a processos misturadores. *In: A Estatística e o Futuro, e o Futuro da Estatística (Actas do I Congresso Anual da SPE, 1994, Vimeiro)*, ed. Pestana, D. e al., Edições Salamandra, Lisboa, 89-98.

Tenreiro, C. (1995). Loi asymptotique des erreurs quadratiques intégrées des estimateurs à noyau de la densité et de la régression, sous des conditions de dépendance. Pré-publicação 95-04, Departamento de Matemática, Universidade de Coimbra.

Volkonskiĭ, V.A. e Rozanov, Yu.A. (1959). Some limit theorems for random functions. I. *Theory Probab. Appl.*, **4**, 178-197.

Yoshihara, K. (1976). Limiting behavior of U-statistics for stationary, absolutely regular processes. *Z. Wahrsch. Verw. Gebiete*, **35**, 237-252.

Yoshihara, K. (1989). Limiting behavior of generalized quadratic forms generated by absolutely regular sequences II. *Yokohama Math. J.*, **37**, 109-123.