Introdução à Econometria Semiparamétrica

Aula 1 - Estimação Não Paramétrica Clássica

Luis A. F. Alvarez

8 de outubro de 2024

Estimando uma densidade

- Suponha que você, pesquisador, possui uma amostra aleatória $\{X_i\}_{i=1}^n$ de uma distribuição contínua F, cuja densidade denotamos por f.
    - Por exemplo, você possui dados da riqueza não financeira de uma amostra de indivíduos.
- Você está interessado em estimar a densidade f.
- Como você pode fazer isso?

Estimação paramétrica

- Um caminho consiste em realizar uma hipótese de forma funcional acerca de f.
- Por exemplo, se nossa população de interesse está restrita à cauda superior da distribuição de riqueza, a literatura sugere que deve valer a "lei de Pareto" (Benhabib e Bisin, 2018), segundo a qual:

$$f(x) = g(x|x_0, \alpha) = \alpha \frac{x_0^\alpha}{x^{\alpha+1}} \mathbf{1}\{x \geq x_0\}, \quad x_0, \alpha > 0$$

  - Distribuição compatível com modelos de agentes heterogêneos com mercados incompletos e ativos arriscados (Achdou et al., 2022).
- Sob a hipótese de forma funcional, a estimação de f se resume à estimação de dois parâmetros, $(x_0, \alpha) \in \mathbb{R}^2$.
- Por exemplo, podemos estimá-los através de máxima verossimilhança:

$$(\hat{x}_0, \hat{\alpha}) \in \operatorname{argmax}_{(z,a)\in\mathbb{R}^2} \frac{1}{n} \sum_{i=1}^{n} \log g(X_i|z, a)$$

Forma funcional e estimação não paramétrica

- E se a lei de Pareto não for válida?
  - Nesse caso, a densidade estimada não recuperará a distribuição correta, mesmo em amostras grandes.
- Como podemos estimar a distribuição correta sem requerer a hipótese de forma funcional?
- Para isso, necessitamos de métodos não paramétricos, que dispensem da hipótese de forma funcional.

Histograma

- O estimador não paramétrico mais simples é o histograma.
- Em um histograma, o suporte da distribuição é dividio em J células (*bins*) de comprimento $2h$, e a densidade para os pontos na j-ésima célula, $[a_j, b_j)$, é estimada como constante igual a:

$$\hat{f}_h(x_j) = \frac{1}{n2h} \sum_{i=1}^{n} \mathbf{1}\left\{|X_i - x_j| \leq h\right\} \ ,$$

onde o $x_j = \frac{a_j + b_j}{2}$ é o ponto intermediário da célula j.

Estimador de *kernel* para densidade

- O estimador $\hat{f}_h(x_j)$ usado na construção do histograma é um caso particular de um estimador de *kernel* para uma densidade em um ponto x, que é dado por:

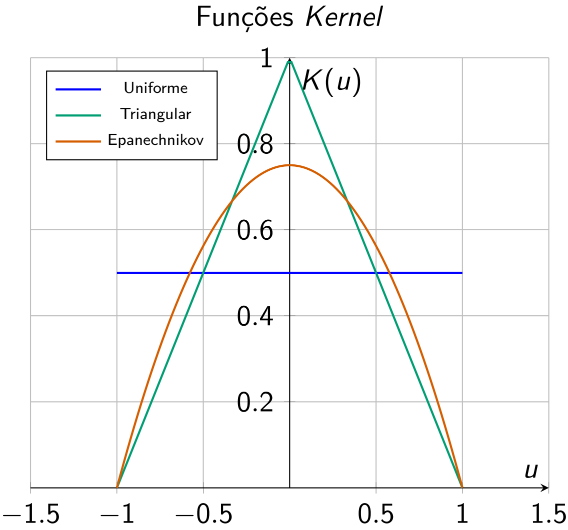
$$\hat{f}_h(x) = \frac{1}{nh} \sum_{j=1}^{n} K\left(\frac{X_j - x}{h}\right),$$

  onde $K : \mathbb{R} \mapsto \mathbb{R}$ é um *kernel*, i.e. uma função satisfazendo $\int_{-\infty}^{\infty} K(u)du = 1$ e simétrica em torno do zero, $K(u) = K(-u)$.
- No que segue, vamos supor que K é uma função não negativa com suporte no intervalo $[-1, 1]$, embora essas hipóteses não sejam necessárias para a maioria dos resultados.
- Sob a hipótese anterior, estimador de *kernel* propõe estimar $f(x)$ "contando" os pontos da amostra com distância a menos de h de x. K nos dá então o peso relativo de cada ponto dentro de essa região.

Exemplos de *kernel*

Kernel	**Fórmula**
Uniforme	$K(u) = \frac{1}{2} \cdot \mathbf{1}(\lvert u \rvert \leq 1)$
Triangular	$K(u) = (1 - \lvert u \rvert) \cdot \mathbf{1}(\lvert u \rvert \leq 1)$
Epanechnikov	$K(u) = \frac{3}{4}(1 - u^2) \cdot \mathbf{1}(\lvert u \rvert \leq 1)$

Exemplos de *kernel*



Calculando o viés do estimador de *kernel*

- Vamos analisar as propriedades estatísticas do estimador de *kernel*.
- Considere um ponto x do suporte da distribuição tal que $f(\cdot)$ seja duas vezes continuamente diferenciável numa vizinhança de x. Neste caso, temos que, à medida que $h \to 0$:

$$\mathbb{E}[\hat{f}_h(x)] = \frac{1}{nh}\sum_{i=1}^{n}\mathbb{E}\left[K\left(\frac{X_i - x}{h}\right)\right] = \frac{1}{h}\mathbb{E}\left[K\left(\frac{X_1 - x}{h}\right)\right] =$$

$$\frac{1}{h}\int_{x-h}^{x+h} K\left(\frac{s-x}{h}\right) f(s)ds \overset{u=\frac{s-x}{h}}{=} \int_{-1}^{1} K(u)f(x+hu)du =$$

$$f(x)\underbrace{\int_{-1}^{-1} K(u)du}_{=1} + f'(x)h\underbrace{\int_{-1}^{-1} uK(u)du}_{=0} + \int_{-1}^{1}(uh)^2\frac{f''(\tilde{x}_{hu})}{2}K(u)du =$$

$$f(x) + \frac{h^2}{2}f''(x)\int_{-1}^{1} u^2K(u)du + o(h^2)$$

onde, da segunda para a terceira linha, aplicamos o teorema do valor médio de segunda ordem, e a última passagem usa continuidade (uniforme) da segunda derivada numa vizinhança de x.

Calculando a variância do estimador de *kernel*

- Do mesmo modo:

$$\mathbb{V}[\hat{f}_h(x)] = \frac{1}{nh^2}\mathbb{V}\left[K\left(\frac{X_1 - X}{h}\right)\right]$$

- Mas:

$$\mathbb{E}\left[K\left(\frac{X_1 - X}{h}\right)^2\right] = h\int_{-1}^{1} f(x)K^2(u)du + O(h^3)$$

e

$$\mathbb{E}\left[K\left(\frac{X_1 - X}{h}\right)\right]^2 = h^2 f(x)^2 + O(h^3)$$

- De modo que:

$$\mathbb{V}[\hat{f}_h(x)] = \frac{1}{nh}\int_{-1}^{1} f(x)K^2(u)du + \frac{1}{n}f(x)^2 + \underbrace{O\left(\frac{h}{n}\right)}_{=o\left(\frac{1}{nh}\right)}$$

Erro quadrático médio do estimador de *kernel*

- Combinando os pontos anteriores, temos que:

$$\text{EQM}[\hat{f}_h(x)] = \mathbb{E}[(\hat{f}_h(x) - f(x))^2] =$$

$$h^4 f''(x)^2 \left(\int_{-1}^{1} u^2 K(u) du\right)^2 + \frac{1}{nh}\int_{-1}^{1} f(x) K^2(u) du + \frac{1}{n} f(x)^2 + o\left(\frac{1}{nh} \vee h^4\right)$$

- Portanto, se $n \to \infty$ com $h \to 0$ e $nh \to \infty$.

$$\text{EQM}[\hat{f}_h(x)] \to 0\,,$$

  e o estimador de kernel é *consistente*.

- Além disso, da expressão acima escolha h ótima, que balanceia *trade-off* viés-variância, é dada por:

$$h_n \propto n^{-1/5}$$

Estimação de densidades multidimensionais

- O estimador anteriormente introduzido é facilmente extensível para a estimação de uma densidade em $\mathbb{R}^d$. Seja $\boldsymbol{f}$ uma densidade em $\mathbb{R}^d$, e $\{\boldsymbol{X}_i\}_{i=1}^n$ uma amostra aleatória desta distribuição. Podemos estimar $\boldsymbol{f}(\boldsymbol{x})$ como:

$$\hat{\boldsymbol{f}}_{\boldsymbol{h}}(x) = \frac{1}{nh_1 \dots h_d} \sum_{i=1}^{n} \boldsymbol{K}_{\boldsymbol{h}}\left(\boldsymbol{X}_i - \boldsymbol{x}\right)$$

onde $\boldsymbol{K}_{\boldsymbol{h}}(\boldsymbol{u}) = \prod_{j=1}^{d} K(u_j/h_j)$, com K um *kernel* univariado.

- Nesse caso, viés é proporcional da ordem de $d \max_j h_j^2$ e variância a $\frac{1}{nh_1 \dots h_d}$. Portanto, se:

$$\max_j h_j \to 0, \quad nh_1 \dots h_d \to \infty\,,$$

estimador é consistente.

Estimador Nadaraya-Watson para a esperança condicional

- A lógica do estimador anterior pode ser estendida para, dada uma amostra aleatória $\{(Y_i, X_i)\}_{i=1}^n$ de (Y, X), estimar-se a função de esperança condicional $m(X) = \mathbb{E}[Y|X]$ num ponto x.
- Nesse caso, podemos tomar médias locais, dadas pelo estimador de Nadaraya-Watson:

$$\hat{m}(x) = \frac{\sum_{i=1}^n K\left((X_i - x)/h\right) Y_i}{\sum_{i=1}^n K\left((X_i - x)/h\right)}$$

Viés e variância do estimador de Nadaraya-Watson

- Para uma m duas vezes continuamente diferenciável numa vizinhança de x, com densidade f positiva e continuamente diferenciável numa vizinhança de x, temos, se $n \to \infty$ e $nh \to \infty$

$$\text{Viés}(\hat{m}(x)) = h^2 \left(\frac{1}{2} m''(x) + \frac{m'(x) f'(x)}{f(x)} \right) \int_{-\infty}^{\infty} u^2 K(u) du + o(h^2)$$

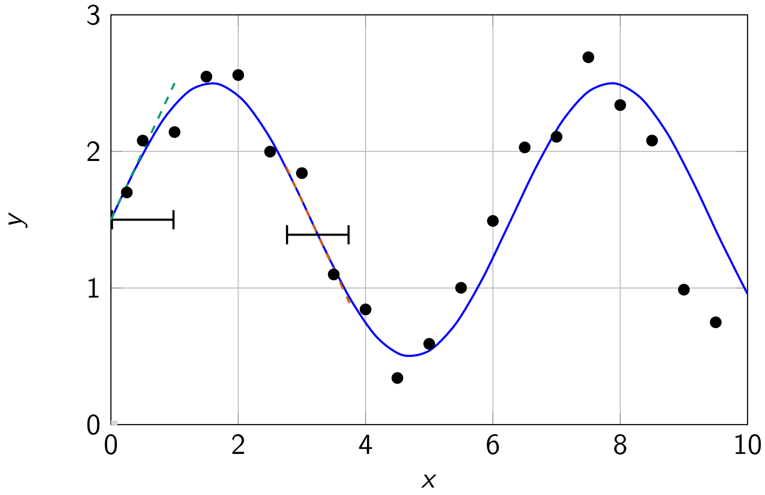
$$\mathbb{V}(\hat{m}(x)) = \frac{\sigma^2(x)}{f(x) nh} \int_{-\infty}^{\infty} K^2(u) du + o\left(\frac{1}{nh} \right)$$

  onde $\sigma^2(X) = \mathbb{V}[Y|X]$.
- Estimador é consistente para $m(x)$ se $n \to \infty$, $h \to 0$ e $nh \to \infty$.
- Banda ótima em termos de erro quadrático médio é dada por: $h_n \propto n^{-1/5}$

Viés de fronteira

- A derivação anterior pressupõe um ponto interior ao suporte de X,
- Para pontos na fronteira do suporte, a expansão de Taylor é unilateral, de modo que haverá um componente de viés adicional em $\hat{m}(x)$.
    - Esse componente adicional de viés tem taxa de convergência mais lenta, de h e não h^2, e será proporcional a $m'(x)$.

Visualização



Regressão linear local

- O estimador de Nadaraya-Watson resolve:

$$\hat{m}(x) = \text{argmin}_{a \in \mathbb{R}} \sum_{i=1}^{n} K\left(\frac{X_i - x}{h}\right)(Y_i - a)^2$$

- Uma expansão de Taylor $m(X_i) \approx m(x) + m'(x)(X_i - x)$, e a expressão acima sugerem a estimação de $m(x)$ como o intercepto da regressão linear local:

$$\min_{a,b} \sum_{i=1}^{n} K\left(\frac{X_i - x}{h}\right)(Y_i - a - b(X_i - x))^2$$

- Coeficiente associado a $(X_i - x)$, $\hat{b}$, estima $m'(x)$, embora esse estimador não seja “bom” (regressão quadrática local é preferível).

Viés e variância da regressão linear local e ajuste automático à fronteira

- Sob as mesmas condições do estimador de Nadaraya-Watson, o estimador $\tilde{m}(x)$ satisfaz, para os pontos interiores do suporte de X.

$$\text{Viés}(\tilde{m}(x)) = \frac{h^2}{2} m''(x) \int_{-\infty}^{\infty} u^2 K(u) du + o(h^2)$$

$$\mathbb{V}(\tilde{m}(x)) = \frac{\sigma^2(x)}{f(x)nh} \int_{-\infty}^{\infty} K^2(u) du + o\left(\frac{1}{nh}\right)$$

- Ademais, uma fórmula análoga (com as mesmas taxas), vale para os pontos de fronteira do suporte (Fan e Gijbels, 1996, Teorema 3.3).
- Banda ótima em termos de erro quadrático médio é dada por $h_n \propto n^{-1/5}$

Inferência

- Sob condições de regularidade, temos que:

$$\sqrt{nh}\left(\frac{\tilde{m}(x) - m(x) - \text{Viés}(\tilde{m}(x))}{\sqrt{\mathbb{V}[x]}}\right) \xrightarrow{d} \mathcal{N}(0,1)\,.$$

- Estimador é assintoticamente normal, embora não necessariamente seja centrado no parâmetro de interesse.
- Se $\sqrt{nh}\text{Viés}(\tilde{m}(x)) \to 0$, mau centramento desaparece assintoticamente, e intervalo de confiança

$$[\tilde{m}(x) + z_{\alpha/2}\sqrt{\widehat{\mathbb{V}[\tilde{m}(x)}]}, \tilde{m}(x) + z_{1-\alpha/2}\sqrt{\widehat{\mathbb{V}[\tilde{m}(x)}]}]\,,$$

  tem cobertura assintótica de $1-\alpha$.
  - $\sqrt{nh}\text{Viés}(\tilde{m}(x)) \to 0 \iff nh^5 \to 0$ (condição de *undersmoothing*).
  - Banda ótima para estimação pontual não satisfaz essa condição.
  - Condição também não é "ótima" sob diferentes critérios inferenciais (Armstrong e Kolesár, 2020; Calonico, Cattaneo e Farrell, 2022)

Alternativas inferenciais

- Calonico, Cattaneo e Titiunik (2014), Calonico, Cattaneo e Farrell (2018) e Calonico, Cattaneo e Farrell (2022): estime Viés($\tilde{m}(x)$) e ajuste erro padrão pela incerteza desta etapa, considerando o intervalo de confiança:

$$\tilde{m}(x) - \widehat{\text{Viés}(\tilde{m}(x))} \pm z_{1-\alpha/2}\sqrt{\mathbb{V}[\tilde{m}(x) \widehat{+ \text{Viés}(\tilde{m}(x))}]},\ ,$$

- Armstrong e Kolesár (2020): considere o intervalo de confiança

$$\tilde{m}(x) \pm c_{1-\alpha/2}(M)\sqrt{\mathbb{V}[\widehat{\tilde{m}(x)}]},\ ,$$

  onde o valor crítico ajustado $c_q(M)$ incorpora o efeito do pior viés possível, numa classe de funções "suaves", com parâmetro de suavidade M, sobre o valor crítico.

  - Intervalos *undersmothened* ou corrigidos por viés podem ser vistos como escolhendo um M implicitamente e, nesse caso, construindo um intervalo de maior comprimento.´

Aproximação uniforme

- A discussão até aqui focou na qualidade da aproximação local em um único ponto.
- Entretanto, para o uso dos estimadores locais em análises posteriores, é interessante entender suas propriedades de aproximação, uniformemente no suporte.
- Einmahl e Mason (2005): para o estimador de Nadaraya-Watson com vetor aleatório $\mathbf{X}$ em $\mathbb{R}^d$, e para qualquer subconjunto compacto I em $\mathbb{R}^d$:

$$\sup_{\boldsymbol{x}\in I} |\hat{m}(\boldsymbol{x}) - m(\boldsymbol{x}) - \text{Viés}(\hat{m}(\boldsymbol{x}))| = O_{\mathbb{P}}\left(\sqrt{\frac{(-\sum_{j=1}^{d}\log(h_j)) \vee \log\log n}{nh_1\ldots h_d}}\right)$$

- Para as bandas ótimas, $h_j \propto n^{-1/(d+4)}$, taxa é de $\sqrt{\frac{\log(n)}{n^{\frac{4}{4+d}}}}$.
- Resultados similares para dados com dependência temporal (Alvarez e Pinto, 2023).

Aproximação "global"

- Recorde-se que, sob condições bastante gerais, o estimador de MQO de Y num vetor $\boldsymbol{X}$ é consistente para o minimizador de

$$\min_{b \in \mathbb{R}^d} \mathbb{E}[(Y - \boldsymbol{X}'b)^2],$$

e que esse mesmo minimizador também resolve

$$\min_{b \in \mathbb{R}^d} \mathbb{E}[(\mathbb{E}[Y|\boldsymbol{X}] - \boldsymbol{X}'b)^2],$$

de modo que o estimador de MQO recupera a melhor aproximação linear à função de esperança condicional $\mathbb{E}[Y|\boldsymbol{X}]$ na norma $L_2(\mathbb{P}_{\boldsymbol{X}})$.

- No entanto, essa aproximação pode ser bastante ruim quando o objetivo não é "resumir" $\mathbb{E}[Y|\mathbf{X}]$, e sim analisar seu comportamento "global".

Aproximando a esperança condicional

- Recorde-se que a função esperança condicional resolve:

$$\min_{h \in L_2(\mathbb{P}_{\boldsymbol{X}})} \mathbb{E}[(Y - h(\boldsymbol{X}))^2]$$

- Quando o suporte $\mathcal{X}$ de $\boldsymbol{X}$ é finito e pequeno relativamente a n, análogo amostral do problema acima é factível, visto que $L_2(\mathbb{P}_{\boldsymbol{X}}) = \{\sum_{s \in \mathcal{X}} a_s \mathbf{1}\{x = s\} : (a_s)_{s \in \mathcal{X}} \in \mathbb{R}^{\mathcal{X}}\}$.
  - Regressão "saturada", com indicadores de todos os pontos do suporte, é um estimador não paramétrico factível de $\mathbb{E}[Y|\boldsymbol{X}]$.
- No entanto, se $|\mathcal{X}|$ é grande em relação a n, análogo amostral torna-se impraticável.
  - Pouco ou nenhum (se $|\mathcal{X}| \geq n$) grau de liberdade para estimar a esperança condicional.
  - Em particular, se $\boldsymbol{X}$ contiver alguma variável aleatória contínua, existirão infinitas soluções que se adequarão aos dados perfeitamente, e poderão distar de $\mathbb{E}[Y|\boldsymbol{X}]$ tão mal quanto se queira.

Estimador de séries

- Dada a impraticabilidade de se estimar:

$$\min_{h \in L_2(\mathbb{P}_{\boldsymbol{X}})} \sum_{i=1}^{n} (Y_i - h(\boldsymbol{X}_i))^2 \,,$$

o estimador não paramétrico de séries propõe-se a estimar:

$$\min_{h \in \Theta_{J_n}} \sum_{i=1}^{n} (Y_i - h(\boldsymbol{X}_i))^2 \,,$$

onde $(\Theta_j)_{j \in \mathbb{N}}$ são uma sequência crescente de subespaços lineares de $L^2(\mathbb{P}_{\boldsymbol{X}})$ de dimensão finita, tais que a função de esperança condicional $m(\boldsymbol{X})$ satisfaz:

$$m \in \mathcal{A} \,,$$

onde $(\mathcal{A}, d)$ é um espaço métrico com $\cup_{i \in \mathbb{N}} \Theta_j \subseteq A \subseteq L^2(\mathbb{P}_{\boldsymbol{X}})$ e $A = \overline{\cup_{i \in \mathbb{N}} \Theta_j}$.
  - Em palavras, m é aproximada tão bem quanto se queira na métrica d, por um elemento de $\cup_{i \in \mathbb{N}} \Theta_j$.

Viés de aproximação e *trade-off* viés-variância

- As propriedades de aproximação do estimador de séries dependem crucialmente do viés de aproximação :

$$\rho_j = \inf_{s \in \Theta_j} d(m, s)\,.$$

- Essas quantidades são conhecidas teoricamente para diferentes classes de séries, com base em resultados de teoria da aproximação.
- A escolha de J_n pode ser feita contrastando uma estimativa de $\hat{\rho}_j$ com a variância estimada do estimador, de modo a equalizar o trade-off viés-variância ou garantir inferência correta (Belloni, Chernozhukov, Chetverikov e Kato, 2015; Chen, T. Christensen e Kankanala, 2024).

Polinômios

- Considere a situação em que o suporte de $\boldsymbol{X}$ é dado por um intervalo $[a, b]$ de $\mathbb{R}$.
- Nesse caso, podemos considerar a classe de polinômios:

$$\Theta_j = \left\{ \sum_{l=0}^{j} a_l x^l : (a_l)_{l=0}^{j} \in \mathbb{R}^{j+1} \right\},$$

- Pelo teorema de Stone-Weierstrass, sabemos que, se m é contínua em $[a, b]$, então m é tão bem aproximada quanto se queira por um polinômio, na norma uniforme.
  - Sob condições de diferenciabilidade e Hölder-continuidade nas derivadas de m, podemos encontrar uma função explícita para ρ_j como função de j e m.

Splines

- Os polinômios da seção anterior podem funcionar bastante mal, especialmente na fronteira do suporte (fenômeno de Runge).
  - Isso se deve à convexidade de x^p.
- Uma alternativa bastante comum consiste no uso de *splines*.
- Um *spline* cúbico com *nós* $a < t_1 < t_2 < \dots t_{j-4} < b$ é dado por:

$$\Theta_j = \text{span}\{1, x, x^2, x^3, \max(x - t_1, 0)^3, \dots, , \max(x - t_{j-4}, 0)^3\}$$

- Função $3 - 1 = 2$ vezes continuamente diferenciável em $[a, b]$.
- Ordem do polinômio é *fixa*, o que mudam é o número de nós (uma convenção é usar percentis igualmente espaçados do suporte).
- Melhores propriedades de aproximação para funções em espaços Hölder que polinômios (veja Belloni, Chernozhukov, Chetverikov e Kato, 2015).
- Um *B-Spline* é uma base do espaço gerado por um *spline* com propriedades computacionais desejáveis.

Tensor products

- Os resultados anteriores pressupunham que $\boldsymbol{X}$ era univariado.
- O caso multivariado pode ser lidado através de tensores. Construímos séries univariadas para cada uma das entradas de $\boldsymbol{X}$, e depois consideramos o espaço gerado pelos produtos das bases de cada um dos espaços.
- Por exemplo, se $\boldsymbol{X} = (X_1, X_2)$ é bidimensional, e consideramos polinômios de grau 2 para cada uma das dimensões, o espaço-produto é:

$$\text{span}\{1, x_1, x_2, x_1 x_2, x_1^2 x_2, x_2^2, x_1 x_2^2, x_1^2, x_2^2, x_1^2 x_2^2\}$$

- Resultado de aproximação uniforme para estimador de séries baseado em tensores de *splines* (Belloni, Chernozhukov, Chetverikov e Kato, 2015) e escolha "ótima" de J_n (velocidade de convergência mais rápida para o estimador):

$$\sup_{x \in \mathcal{X}} |\widehat{h}(x) - m(x)| = O_{\mathbb{P}}\left(\left(\frac{\log n}{n}\right)^{s/(2s+d)}\right),$$

onde s é o "grau de suavidade" de m, dado pelas restrições sobre sua derivada.

Resultados adicionais de estimação de séries

1. Outras bases (veja Chen, 2007; Belloni, Chernozhukov, Chetverikov e Kato, 2015; Cattaneo, Farrell e Feng, 2020):
   - Polinômios particionados.
   - Bases de Fourier.
   - Ondaletas.
2. Aproximações com dependência temporal (Chen e T. M. Christensen, 2015).
3. Outras funções objetivo.
   - Quantil (Belloni, Chernozhukov, Chetverikov e Fernández-Val, 2019).

Bibliografia I

Achdou, Yves et al. (2022). "Income and wealth distribution in macroeconomics: A continuous-time approach". Em: *The review of economic studies* 89.1, pp. 45–86.

Alvarez, Luis e Cristine Pinto (2023). *A maximal inequality for local empirical processes under weak dependence*. arXiv: 2307.01328 [econ.EM]. URL: https://arxiv.org/abs/2307.01328.

Armstrong, Timothy B e Michal Kolesár (2020). "Simple and honest confidence intervals in nonparametric regression". Em: *Quantitative Economics* 11.1, pp. 1–39.

Belloni, Alexandre, Victor Chernozhukov, Denis Chetverikov e Iván Fernández-Val (2019). "Conditional quantile processes based on series or many regressors". Em: *Journal of Econometrics* 213.1, pp. 4–29.

Bibliografia II

Belloni, Alexandre, Victor Chernozhukov, Denis Chetverikov e Kengo Kato (2015). "Some new asymptotic theory for least squares series: Pointwise and uniform results". Em: *Journal of Econometrics* 186.2, pp. 345–366.

Benhabib, Jess e Alberto Bisin (2018). "Skewed wealth distributions: Theory and empirics". Em: *Journal of Economic Literature* 56.4, pp. 1261–1291.

Calonico, Sebastian, Matias D Cattaneo e Max H Farrell (2018). "On the effect of bias estimation on coverage accuracy in nonparametric inference". Em: *Journal of the American Statistical Association* 113.522, pp. 767–779.

— (2022). "Coverage error optimal confidence intervals for local polynomial regression". Em: *Bernoulli* 28.4.

Calonico, Sebastian, Matias D Cattaneo e Rocio Titiunik (2014). "Robust nonparametric confidence intervals for regression-discontinuity designs". Em: *Econometrica* 82.6, pp. 2295–2326.

Bibliografia III

Cattaneo, Matias D, Max H Farrell e Yingjie Feng (2020). "Large sample properties of partitioning-based series estimators". Em: *The Annals of Statistics* 48.3, pp. 1718–1741.

Chen, Xiaohong (2007). "Chapter 76 Large Sample Sieve Estimation of Semi-Nonparametric Models". Em: ed. por James J. Heckman e Edward E. Leamer. Vol. 6. Handbook of Econometrics. Elsevier, pp. 5549–5632. DOI: https://doi.org/10.1016/S1573-4412(07)06076-X. URL: https://www.sciencedirect.com/science/article/pii/S157344120706076X.

Bibliografia IV

Chen, Xiaohong, Timothy Christensen e Sid Kankanala (mar. de 2024). "Adaptive Estimation and Uniform Confidence Bands for Nonparametric Structural Functions and Elasticities". Em: *The Review of Economic Studies*, rdae025. ISSN: 0034-6527. DOI: 10.1093/restud/rdae025. eprint: https://academic.oup.com/restud/advance-article-pdf/doi/10.1093/restud/rdae025/57021176/rdae025.pdf. URL: https://doi.org/10.1093/restud/rdae025.

Chen, Xiaohong e Timothy M. Christensen (2015). "Optimal uniform convergence rates and asymptotic normality for series estimators under weak dependence and weak conditions". Em: *Journal of Econometrics* 188.2. Heterogeneity in Panel Data and in Nonparametric Analysis in honor of Professor Cheng Hsiao, pp. 447–465. ISSN: 0304-4076. DOI: https://doi.org/10.1016/j.jeconom.2015.03.010. URL: https://www.sciencedirect.com/science/article/pii/S0304407615000792.

Bibliografia V

Einmahl, Uwe e David M. Mason (2005). "Uniform in bandwidth consistency of kernel-type function estimators". Em: *The Annals of Statistics* 33.3, pp. 1380–1403. DOI: 10.1214/009053605000000129. URL: https://doi.org/10.1214/009053605000000129.

Fan, Jianqing e Irene Gijbels (mar. de 1996). *Local polynomial modelling and its applications*. en. Chapman & Hall/CRC Monographs on Statistics and Applied Probability. Philadelphia, PA: Chapman & Hall/CRC.