UNIVERSIDADE FEDERAL DO RIO GRANDE DO NORTE
CENTRO DE CIÊNCIAS EXATAS E DA TERRA
PROGRAMA DE PÓS-GRADUAÇÃO EM MATEMÁTICA
APLICADA E ESTATÍSTICA

Estimador do Tipo Núcleo para Densidades Limites de Cadeias de Markov com Espaço de Estados Geral

Maria Aparecida da Silva Soares
Orientadora: Profa. Dra. Viviane Simioli Medeiros Campos

Natal, fevereiro de 2010

Livros Grátis

http://www.livrosgratis.com.br

Milhares de livros grátis para download.

UNIVERSIDADE FEDERAL DO RIO GRANDE DO NORTE

CENTRO DE CIÊNCIAS EXATAS E DA TERRA

PROGRAMA DE PÓS-GRADUAÇÃO EM MATEMÁTICA APLICADA E ESTATÍSTICA

Estimador do Tipo Núcleo para Densidades Limites de Cadeias de Markov com Espaço de Estados Geral

Maria Aparecida da Silva Soares

Dissertação de Mestrado apresentada ao Programa de Pós-Graduação em Matemática Aplicada e Estatística da Universidade Federal do Rio Grande do Norte (PPGMAE-UFRN) como parte dos requisitos necessários para obtenção do título de Mestre em Matemática Aplicada e Estatística.

Natal, fevereiro 2010

Dedicatória

"Este trabalho é dedicado a melhor mãe do mundo, a minha mãe Dona Maria, cujas mãos nunca souberam escrever mas sempre trabalharam para que hoje eu pudesse realizar este sonho. Obrigado mãe por tudo!!!"

Agradecimentos

A Deus acima de tudo por ter me ajudado a realizar este sonho, por ser o meu refúgio e minha fortaleza diante dos momentos de dificuldade.

A professora Viviane pela confiança em mim depositada mesmo sem me conhecer. Não foi apenas uma orientadora, foi uma excelente professora, uma amiga e uma mãe para mim, sempre me ajudou e aconselhou. Obrigado pela preocupação, carinho e respeito que sempre teve por mim. Professora Viviane obrigada por me ajudar a realizar esse sonho.

Ao professor André Gustavo pela orientação no primeiro ano de mestrado e por sempre ter ajudado e acreditado em mim.

A toda a minha família, em especial as minhas irmãs Lúcia e Ivanilda, ao meu pai seu José e a minha mãe dona Maria não só pelo apoio financeiro mas por terem me ensinado coisas que foram de fundamental importância para a realização deste sonho, me ensinaram a ter responsabilidade, ter dedicação e a correr atrás dos meus sonhos sem ter medo dos obstáculos.

A minha querida amiga Camila por sempre enxugar as minhas lágrimas, mesmo achando um mico chorar. Por comprar as minhas brigas e por ter me ajudado em uma fase ruim do mestrado. Por nunca ter me abandonado mesmo estando longe. Obrigado pela amizade!!!

A minha grande amiga Simone Pereira, que esteve ao meu lado nos momentos mais difíceis, por seus conselhos, por me indicar o melhor caminho a seguir e pela amizade a mim confiada todo esse tempo.

As minhas amigas, Elaine, Liliana, Simone Santos e Úrsula, por sempre torcerem por mim mesmo estando fora da realidade do mestrado.

A Patrícia pelas conversas na hora do café, por me escutar e por sempre ter a palavra certa pra me dizer, a que era inspirada na palavra de Deus.

A todos os meus colegas de pós graduação, em especial:

As minhas colegas de turma Enai e Tatiana Queiroz, pelas conversas e pelo companheirismo nos estudos.

A Renilma, que eu já conhecia, que se revelou uma grande amiga e companheira durante o mestrado. Espero ter conquistado sua amizade para sempre.

A Kelly pelo seu sorriso, sua alegria de viver e por ter torcido por mim sem ao menos me conhecer. Alguém que valeu a pena conhecer em 2010.

A Manassés e Renata Mendonça pela companhia nos estudos da disciplina de inferência, pelo respeito e carinho com que sempre me trataram.

Aos meus companheiros de estudo na minha segunda fase de probabilidade, Alysson, Kaline Juliana, Francinário, Claúdia, Tatiana Farache, Felipe e Jullyanna. A todos vocês, obrigado por estarem comigo neste momento.

Ao Hermes pela ajuda em métodos estatísticos.

Aos demais colegas Cecílio, Lenilson, Daniel, Marcônio, Kaline Andreza e aos que concluíram seus trabalhos, que fazem ou fizeram parte dessa família chamada PPGMAE.

A todos os professores da graduação e da pós graduação, em especial:

Ao professor Claudio Carlos Dias do DMAT-UFRN por ter sido meu mestre durante a graduação, pelas orientações e por toda a confiança que ele sempre teve em mim.

A professora Dione, pois além de coordenadora do programa e professora, tem um carinho e um cuidado especial por todos os alunos do PPGMAE.

Ao professor Damião pelo incentivo nos momentos de dificuldade e por ter me ajudado a gostar de estatística.

Ao professor Jaques, pelos conselhos e por ter me ensinado probabilidade.

Ao professor Pledson pelas orientações e incentivos durante a disciplina de seminários.

A professora Daniele por ter aceitado o convite para participar da banca.

Aos professores, Juan e Paulo, pelas correções e contribuições no Exame de Qualificação.

A Rafael, o secretário nota 10!!!

A Liandra pela amizade e pelo carinho.

Ao comitê gestor da bolsa reuni, em especial a professora Márcia Goretti.

A um alguém muito especial que fez parte de toda essa minha caminhada, que esta comigo desde a época da graduação, nos bons e maus momentos, e que foi muito importante para a realização deste sonho, pois ele participou de cada etapa, sua ajuda esta presente em cada detalhe desta dissertação e do que a antecedeu. Alguem que vai ficar para sempre em meu coração e que todas as paginas desse trabalho seriam poucas para expressar todos os agradecimento que ele merece. Ao meu Grande Amigo e companheiro de todas as horas, Moisés Medeiros. Obrigado pela sua amizade e por você existir em minha vida.

Resumo

Neste trabalho vamos estudamos a consistência para uma classe de estimadores núcleo de $f(\cdot)$ em cadeias de Markov com espaço de estados geral $E \subset \mathbb{R}^d$. Este estudo é dividido em duas partes: Na primeira $f(\cdot)$ é uma densidade estacionária de uma cadeia, e no segundo $f(x)\,\nu(dx)$ é a distribuição limite de uma cadeia geometricamente ergódica.

Palavras chaves: cadeias de Markov, estimadores núcleo, ergodicidade geométrica.

Abstract

In this work we studied the consistency for a class of kernel estimates of $f(\cdot)$ in the Markov chains with general state space $E \subset \mathbb{R}^d$ case. This study is divided into two parts: In the first one $f(\cdot)$ is a stationary density of the chain, and in the second one $f(x)\,\nu(dx)$ is the limit distribution of a geometrically ergodic chain.

Key words: Markov chains, kernel estimates, geometrically ergodic.

Sumário

Introdução

Neste trabalho iremos estudar uma classe de estimadores para uma função densidade de probabilidade desconhecida de uma população a partir de uma amostra dessa população. Para estimar a função densidade de uma cadeia de Markov com espaço de estados geral iremos fazer uso dos estimadores do tipo núcleo, pois trata-se de estimadores com boas propriedades estatísticas. Observe que, neste caso, estamos trabalhando com estimação não-paramétrica, pois estamos estimando funções e não parâmetros.

Inicialmente uma classe geral de estimadores do tipo núcleo definidos por

$$f_n(x) = \frac{1}{nh}\sum_{k=1}^{n} K\left(\frac{x - X_k}{h}\right),$$

com $h = h_n \downarrow 0$ quando $n \to \infty$, para estimar a função de densidade de probabilidade desconhecida, foi estudada por Parzen (1962) para uma sequência de variáveis aleatórias independentes e identicamente distribuídas (i.i.d.). Parzen (1962) mostrou as boas propriedades estatísticas para esses estimadores, ou seja, não vício assintótico, consistência e normalidade assintótica.

Em Prakasa Rao (1983) foi feito um grande estudo sobre os estimadores do tipo núcleo tanto no caso univariado como no caso multivariado.

Estimadores do tipo núcleo associados a cadeias de Markov estritamente estacionárias foram considerados por Roussas (1969, 1991), Rosenblatt (1970) e por Athreya e Atuncar (1998) com o objetivo de estimarem densidades em $\mathbb{R}^d$. Eles extenderam ao caso Markoviano os resultados obtidos por Parzen (1962). Roussas (1969) considerou cadeias satisfazendo a condição D_0 de Doob (1953) com um único conjunto ergódico o qual não possuia subconjuntos ciclicamente móveis. Rosenblatt (1970) substituiu essas hipóteses pela condição φ-mixing. Athreya e Atuncar (1998) substituiu a condição D_0 por uma mais fraca, a condição de *Harris* recorrência.

Nosso trabalho é baseado em Campos e Dorea (2005) onde foi considerado uma cadeia de Markov $\{X_n\}_{n\geq 0}$ com espaço de estados geral $E \subset \mathbb{R}^d$ e uma densidade $f(\cdot)$ com respeito a medida σ-finita ν de E. Para tanto usaremos os estimadores do tipo peso,

$$f_n(x) = \frac{1}{n}\sum_{k=1}^{n} W(h, x, X_k), h = h_n \downarrow 0,$$

onde $W(h, x, \cdot)$ é uma função peso apropriada também chamada de função núcleo, obtidos a partir dos estimadores do tipo núcleo definidos por Parzen (1962). Uma motivação para o uso desses estimadores é feita em Campos e Dorea (2001).

Verificamos que, com condições apropriadas para a sequência h e a função núcleo W, os estimadores $f_n(x)$ são assintoticamente não viciado, consistente em média quadrática, fracamente consistente e fortemente consistente.

Este trabalho é dividido em três partes. Na primeira parte, no Capítulo 1, damos alguns resultados e conceitos preliminares que serão de grande importância nos capítulos seguintes. Inicialmente, na seção 1.1, fazemos um breve estudo sobre cadeias de Markov com espaço de estados geral, por ser esse um conceito dificilmente encontrado em livros que tratam do assunto. Em seguida na seção 1.2 damos uma motivação para o uso dos estimadores do tipo núcleo em alguns casos a saber, ou seja, no caso em que temos uma sequência de variáveis aleatórias (i.i.d.) e no caso em que temos um processo de Markov.

O Capítulo 2, para estimar a densidade estacionária de uma cadeia de Markov, sob a condição da densidade inicial ser a estacionária e a cadeia de Markov satisfazendo a condição φ-mixing, ou seja,

$$|P(A\cap B) - P(A)P(B)| \leq \varphi(n)P(A) \;\; e \; \varphi(n) \downarrow 0 \; quando \; n \to \infty.$$

Vamos apresentar aqui alguns resultados importantes de Campos e Dorea (2005) e obteve-se a consistência desses estimadores de densidade.

Na seção 2.2 supondo apenas que a cadeia é estritamente estacionária verificamos que o estimador aqui estudado é assintoticamente não viciado, ou seja,

$$\lim_{n\to\infty} E(f_n(x)) = f(x), \; \forall x \in C_\nu(f).$$

Na seção 2.3 estudamos os resultados de consistência dos estimadores. Primeiramente verificamos a consistência em média quadrática, ou seja,

$$\lim_{n\to\infty} E\left\{[f_n(x) - f(x)]^2\right\} = 0,\ \forall x \in C_\nu(f),$$

consequentemente a consistência fraca

$$\lim_{n\to\infty} P(|f_n(x) - f(x)| > \epsilon) = 0,\ \forall x \in C_\nu(f),$$

e para finalizar verificamos a consistência forte, ou seja,

$$\lim_{n\to\infty} f_n(x) = f(x)\ \ \nu - q.c.\ (quase\ certamente),\ \forall x \in C_\nu(f).$$

E no Capítulo 3, estimamos a densidade limite, portanto a única estacionária de uma cadeia de Markov. Substituímos as hipóteses de cadeia de Markov estritamente estacionária e satisfazendo a condição φ-mixing pela suposição de que a cadeia é geometricamente ergódica, ou seja,

$$|P^n(x, A) - \pi(A)| \leq \alpha\rho^n,\ \forall\, x \in E,\ \forall\, A \in \mathcal{E},$$

onde α e ρ são constantes com $\alpha > 0$ e $0 < \rho < 1$. Consideramos a densidade inicial da cadeia como sendo a densidade limite e verificamos os mesmos resultados de consistência do Capítulo 2.

Na seção 3.2 verificamos que os estimadores definidos em Campos e Dorea (2001) são assintoticamente não viciado.

Na seção 3.3 usamos o não vício para verificar a consistência em média quadrática e consequentemente a consistência fraca, de maneira idêntica ao Capítulo 2.

E finalmente na seção 3.4 verificamos que uma cadeia de Markov geometricamente ergódica, sem a hipótese de estacionariedade estrita, satisfaz a condição φ-mixing com coeficiente mixing geométrico e desta forma verificamos a consistência forte como foi verificada no Capítulo 2, pois para a consistência forte utilizamos apenas a condição φ-mixing.

Capítulo 1

Preliminares

Neste capítulo iremos estudar alguns conceitos, definições e resultados preliminares que serão utilizados neste trabalho.

Inicialmente, na seção 1.1, vamos estudar cadeias de Markov com espaço de estados geral baseado em Athreya e Lahiri (2006). Começamos com as definições de cadeias de Markov, núcleo de transição, cadeias estritamente estacionária e densidades estacionárias. Ainda nesta seção apresentamos a Equação de Chapman-Kolmogorov.

Na seção 1.2 apresentamos um breve estudo dos estimadores do tipo núcleo com o objetivo de entendermos a motivação de Campos e Dorea (2005) ao trabalhar com uma classe mais geral desses estimadores, os estimadores do tipo peso.

1.1 Cadeias de Markov com espaço de estados geral

Seja $\{X_n\}_{n\geq 0}$ uma sequência de variáveis aleatórias com valores em um espaço E que é não necessariamente finito ou enumerável. A propriedade de Markov diz que condicionada a $X_0, X_1, ..., X_{n-1}, X_n$, a distribuição de X_{n+1}, só depende de X_n e não do passado, isto é, não depende de $\{X_j, j \leq n-1\}$. Quando E é não enumerável, definimos a propriedade de Markov, consideramos $(E, \mathcal{E}, P)$ um espaço de probabilidade onde $\mathcal{E}$ é a σ-álgebra de subconjuntos de E e escrevemos:

Definição 1.1 *A sequência $\{X_n\}_{n\geq 0}$ com valores em E é dita uma cadeia de Markov se para todo $A \in \mathcal{E}$,*

$$P\left(X_{n+1} \in A \,|\, \mathcal{F}_0^n\right) = P\left(X_{n+1} \in A \,|\, \sigma\left(X_n\right)\right), \tag{1.1}$$

com probabilidade 1, para todo $n \geq 0$, para alguma distribuição inicial de X_0, onde $\sigma\left(X_n\right)$ é a σ-algebra gerada por X_n.

Uma importante ferramenta para o estudo de cadeias de Markov é a função de transição de probabilidade (ou núcleo de transição de probabilidade).

Definição 1.2 *A função $P : E \times \mathcal{E} \longmapsto [0,1]$ é dita uma função de transição de probabilidade se,*
(i) para todo $x \in E$, $P\left(x, \cdot\right)$ é uma medida de probabilidade definida em $\mathcal{E}$,
(ii) para todo $A \in \mathcal{E}$, $P\left(\cdot, A\right)$ é uma função mensurável de $E \longmapsto [0,1]$.

Em outras palavras essa definição nos diz que fixado $x \in E$ temos que $P\left(x, \cdot\right)$ é uma medida de probabilidade em $\mathcal{E}$ e se variarmos x e fixarmos $A \in \mathcal{E}$ temos que $P\left(\cdot, A\right)$ é uma função mensurável em E.

Podemos afirmar que uma sequência $\{X_n\}_{n\geq 0}$ satisfaz (1.1) se para todo $n \in \mathbb{N}$ e
$A_0, A_1, ..., A_n \in \mathcal{E}$,

$$\begin{aligned} & P\left(X_j \in A_j,\, j = 0, 1, 2, ..., n\right) \\ = & \int_{A_0} \int_{A_{n-2}} \int_{A_{n-1}} P_{n-1}\left(x_{n-1}, A_n\right) P_{n-2}\left(x_{n-2,dx_{n-1}}\right) ... P_1\left(x_0, dx_1\right) \mu_0\left(dx_0\right), \end{aligned}$$

onde $\mu_0\left(A\right) = P\left(X_0 \in A\right)$, $A \in \mathcal{E}$.

Definição 1.3 *Uma sequência de variáveis aleatórias $\{X_n\}_{n\geq 0}$ com valores em E é dita uma cadeia de Markov com função de transição $P\left(\cdot, \cdot\right)$ se (1.1) é válida e o lado direito da equação pode ser expresso como $P\left(X_n, A\right)$, ou seja,*

$$P\left(X_{n+1} \in A \,|\, \mathcal{F}_0^n\right) = P\left(X_{n+1} \in A \,|\, \sigma\left(X_n\right)\right) = P\left(X_n, A\right), \tag{1.2}$$

para alguma distribuição inicial da cadeia, isto é, distribuição de X_0, para todo $n \in \mathbb{N}$.

Desse modo temos que o processo $\{X_n\}_{n\geq 0}$ fica completamente determinado pelo núcleo P e por uma distribuição inicial de X_0 dada por $\mu_0(A) = P(X_0 \in A)$. Podemos assim, calcular todas as probabilidades de interesse.

Note que a Definição 1.3 simplifica a notação da propriedade de Markov e esta será usada de agora em diante.

Os exemplos a seguir ilustram de maneira bem simples, ou seja, para o caso i.i.d. e para o caso discreto as definições dadas anteriormente.

Exemplo 1.1 *Seja $\{X_n\}_{n\geq 0}$ uma sequência de variáveis aleatórias i.i.d. com valores em E e distribuição inicial μ. Então $\{X_n\}_{n\geq 0}$ é uma cadeia de Markov com função de transição $P(x, A) = \mu(A)$ e distribuição inicial μ.*

Exemplo 1.2 *Considere a cadeia de Markov com núcleo de transição dado por*

$$P = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix},$$

e distribuição inicial $f_4 = \left(\frac{1}{4}, \frac{1}{2}, \frac{1}{4}\right)$.

No Capítulo 2 faremos algumas observações a partir desse exemplo.

Seja $P(\cdot,\cdot)$ uma função de transição de probabilidade sobre $(E, \mathcal{E})$. Para cada $n \geq 0$ defina uma sequência de funções $\{P^n(\cdot,\cdot)\}_{n\geq 0}$ pelo esquema de iteração,

$$P^{n+1}(x, A) = \int_E P^n(y, A)\, P(x, dy), \tag{1.3}$$

onde $P^{(0)}(x, A) = I_A(x)$.

Pode-se verificar por indução que para cada n, $P^{(n)}(\cdot,\cdot)$ é uma função de transição de probabilidade, ou seja, fixado x, $P^{(n)}(x,\cdot)$ é uma medida de probabilidade e fixado A, $P^{(n)}(\cdot, A)$ é uma função mensurável.

Definição 1.4 *$P^{(n)}(\cdot,\cdot)$ definida por (1.3) é uma função de transição de n-passos gerada por $P(\cdot,\cdot)$.*

Verifica-se por indução que se $X_0 = x$ com probabilidade 1 então,

$$P\left(X_n \in A\right) = P^n\left(x, A\right), \ \forall n \geq 0. \tag{1.4}$$

Isso nos leva a Equação de Chapman-Kolmogorov, que é uma ferramenta muito útil no estudo de cadeias de Markov e é uma consequência da propriedade de Markov dada na Definição 1.1.

Proposição 1.1 (Equação de Chapman-Kolmogorov) *Sejam $P\left(\cdot,\cdot\right)$ uma função de transição de probabilidade sobre $\left(E,\mathcal{E}\right)$ e $P^n\left(\cdot,\cdot\right)$ definida em (1.3). Então para $n, m \geq 0$,*

$$P^{n+m}\left(x, A\right) = \int_E P^n\left(y, A\right) P^m\left(x, dy\right). \tag{1.5}$$

Demonstração: Seja E_x, P_x a esperança e a distribuiçao de probabilidade de $\left\{X_n\right\}_{n\geq 0}$ quando $X_0 = x$ com probabilidade 1.

De (1.4), pela definição de esperança condicional e pela propriedade de Markov temos,

$$\begin{aligned}
P^{n+m}\left(x, A\right) &= P_x\left(X_{n+m} \in A\right) = \int_E I_{\left(X_{n+m}\in A\right)} dP \\
&= \int_E E_x\left(I_{\left(X_{n+m}\in A\right)} | \mathcal{F}_0^m\right) dP = E_x\left(P\left(X_{n+m} \in A | \mathcal{F}_0^m\right)\right) \\
&= E_x\left(P\left(X_{n+m} \in A | \sigma\left(X_m\right)\right)\right) = E_x\left(P^n\left(X_m, A\right)\right) \\
&= \int_E P^n\left(y, A\right) P^m\left(x, dy\right).
\end{aligned}$$

■

Um dos objetivos deste trabalho, Capítulo 2, é estudar a estimação da densidade estacionária de uma cadeia de Markov com espaço de estados geral $E \subset \mathbb{R}^d$, considerando a densidade inicial como sendo a estacionária e obtendo assim uma cadeia estritamente estacionária como será visto na seção 2.1. Para isto vamos definir o que seja uma densidade estacionária e uma cadeia estritamente estacionária.

Definição 1.5 *Uma densidade $f\left(\cdot\right)$ é chamada estacionária para a cadeia de Markov com respeito a medida σ-finita ν de E se*

$$\int_A f\left(x\right)\nu\left(dx\right) = \int_E P^n\left(x, A\right) f\left(x\right)\nu\left(dx\right), \ \forall A \in \mathcal{E}, \ \forall n. \tag{1.6}$$

onde para $n \geq 2$

$$P^n(x, A) = \int_E P^k(y, A) P^{n-k}(x, dy), \; 1 \leq k \leq n-1. \tag{1.7}$$

Definição 1.6 *Dizemos que a cadeia de Markov é estritamente estacionária se*

$$P(X_n \in A) = \int_A f(x) \nu(dx), \; \forall A \in \mathcal{E}, \; n = 0, 1, 2, \ldots. \tag{1.8}$$

Em suma, uma cadeia de Markov $\{X_n\}_{n \geq 0}$ é estritamente estacionária quando as variáveis aleatórias têm a mesma distribuição.

Vamos ainda, no Capítulo 3, estudar a estimação da densidade limite, a única estacionária de uma cadeia de Markov com espaço de estados geral $E \subset \mathbb{R}^d$. Para tanto necessitamos de uma condição mais forte para a cadeia, pois não consideramos o processo como sendo estritamente estacionário, vamos supor que a cadeia é geometricamente ergódica.

Definição 1.7 *Dizemos que a cadeia é geometricamente ergódica se existe uma probabilidade* π *em* A *e constantes* $\alpha > 0$ *e* $0 < \rho < 1$ *tal que*

$$|P^n(x, A) - \pi(A)| \leq \alpha \rho^n, \; \forall x \in E, \; \forall A \in \mathcal{E}, \; \forall n. \tag{1.9}$$

Note que, necessariamente temos

$$\lim_{n \to \infty} P^n(x, A) = \pi(A), \tag{1.10}$$

e

$$\pi(A) = \int_A f(x) \nu(dx), \; \forall A \in \mathcal{E}. \tag{1.11}$$

Esta definição nos diz que além da cadeia convergir, ela converge com velocidade geométrica.

1.2 Estimadores do tipo núcleo para densidades

A necessidade de utilizar estimadores da função de densidade f aparece frequentemente em situações onde se descrevem aspectos como o comportamento nas caudas e a simetria, onde a amostra pode ser obtida experimentalmente ou gerada através de simulação, pois o gráfico resume convenientemente a informação relativa a forma da distribuição da amostra, dentre outros.

O histograma é o estimador mais simples de uma função de densidade porém, não possibilita na construção de uma função contínua. Esse fato incentivou diversos estudos na procura de estimadores contínuos com boas propriedades, como por exemplo, não vício assintótico e consistência.

Sejam $X_1, X_2, ..., X_n$ variáveis aleatórias independentes e identicamente distribuídas. Desejamos estimar $f(x)$ onde $x \in \mathbb{R}$ é um ponto de continuidade.

Quando se pretende estimar a função de distribuição $F(x)$ em um determinado ponto x, é natural tomarmos a função de distribuição da amostra, ou seja,

$$F_n(x) = \frac{1}{n} \# \{X_i \leq x, i = 1, 2, ..., n\},$$

onde # denota a cardinalidade. Essa distribuição é chamada de empírica.

Um estimador natural para a função densidade poderia ser dado pela derivada,

$$f_n(x) = \frac{F_n(x+h) - F_n(x-h)}{2h},$$

onde $h = h_n \downarrow 0$ quando $n \to \infty$ apropriadamente escolhida. Esse estimador da densidade é natural por ser a derivada da função de distribuição a menos de um conjunto de medida nula.

Assim, um estimador natural de $f(x)$ é dado por

$$f_n(x) = \frac{1}{2nh} \# \{X_i : X_i \in (x-h, x+h]\}$$

onde $h = h_n \downarrow 0$ quando $n \to \infty$. Esse estimador é chamado de estimador ingênuo e podemos reescrevê-lo como

$$f_n(x) = \frac{1}{2nh} \sum_{k=1}^{n} I_{[-1,1)} \left(\frac{x - X_k}{h} \right).$$

Parzen (1962) motivado por essas idéias definiu uma nova classe de estimadores

$$f_n(x) = \frac{1}{nh} \sum_{k=1}^{n} K \left(\frac{x - X_k}{h} \right), \tag{1.12}$$

onde $h = h_n \downarrow 0$ quando $n \to \infty$, onde K é denominada função do tipo núcleo que satisfaz algumas propriedades de regularidade que são essenciais para que $f_n(x)$ seja um estimador assintoticamente não viciado, consistente em média quadrática, fortemente consistente e assintoticamente normal distribuído. Esses estimadores são chamados de estimadores do tipo núcleo.

Campos e Dorea (2001) definiram uma nova classe de estimadores do tipo peso, que foram redefinidos a partir dos estimadores do tipo núcleo, dados por

$$f_n(x) = \frac{1}{n}\sum_{k=1}^{n} W(h, x, X_k), \tag{1.13}$$

onde $h = h_n \downarrow 0$ quando $n \to \infty$ é uma sequência apropriadamente escolhida, $W(h, x, \cdot)$ que é chamada de função peso ou função núcleo, satisfazendo algumas condições de regularidade e x o ponto em que se deseja estimar a densidade. Neste estudo iremos considerar essa classe de estimadores.

Capítulo 2

Estimação da Densidade Estacionária sob a Condição φ-mixing

2.1 Introdução

Neste capítulo, utilizando os estimadores do tipo peso definidos em (1.13), vamos estudar o problema de se estimar as densidades estacionárias de uma cadeia de Markov com espaço de estados geral.

Considere $\{X_n\}_{n\geq 0}$ uma cadeia de Markov com núcleo de transição $\{P(x,A) : x \in E, A \in \mathcal{E}\}$ onde $E \subset \mathbb{R}^d$ e $\mathcal{E}$ é uma σ-algebra de subconjuntos de E, assuma que a cadeia possui uma densidade estacionária $f(\cdot)$ com respeito a medida σ-finita ν em $\mathcal{E}$, ou seja,

$$\int_A f(x)\,\nu(dx) = \int_E P^n(x,A)\,f(x)\,\nu(dx)\,,\ \forall A \in \mathcal{E},\ \forall n. \tag{2.1}$$

Com o objetivo de estimar $f(\cdot)$, Campos e Dorea (2005) admitiram ser $f(\cdot)$ a densidade inicial da cadeia, obtendo assim uma cadeia de Markov estritamente estacionária como definido em (1.8). De fato, por (2.1) temos

$$\begin{aligned} P(X_0 \in A) &= \int_A f(x)\,\nu(dx) = \int_E P(x,A)\,f(x)\,\nu(dx) \\ &= \int_E P(X_1 \in A|X_0 = x)\,f(x)\,\nu(dx) = P(X_1 \in A)\,. \end{aligned}$$

E em geral temos

$$P(X_n \in A) = \int_A f(x)\,\nu(dx)\,,\ n = 0,1,...\ . \tag{2.2}$$

E admitiram ainda que a cadeia satisfaz a condição φ-mixing, ou seja, para todo $A \in \mathcal{F}_0^k$, para todo $B \in \mathcal{F}_{k+n}^{\infty}$, $k \geq 0$ e $n \geq 1$,

$$|P(A \cap B) - P(A)\,P(B)| \leq \varphi(n)\,P(A)\ \ e\ \varphi(n) \downarrow 0\ quando\ n \to \infty, \tag{2.3}$$

onde $\mathcal{F}_l^{l+m} = \sigma(X_l, X_{l+1}, ..., X_{l+m})$.

Foram obtidos o não vício assintótico e os resultados de consistência em pontos de ν-continuidade da função, ou seja, em pontos que satisfazem a seguinte definição:

Definição 2.1 *Para uma função g definida em E dizemos que x é um ponto de ν-continuidade de g, ou $x \in C_\nu(g)$, se dado $\epsilon > 0$ existe $\delta > 0$ tal que*

$$\nu\{y : |y - x| \leq \delta, |g(x) - g(y)| > \epsilon\} = 0,$$

onde $|\cdot|$ denota a norma euclidiana e a norma associada ao espaço E.

Na seção 2.2, usando apenas a estacionariedade estrita do processo, iremos verificar que o estimador é assintoticamente não viciado, ou seja,

$$\lim_{n\to\infty} E(f_n(x)) = f(x)\,,\ \forall x \in C_\nu(f)\,.$$

Já na seção 2.3 iremos verificar os resultados de consistência utilizando a condição do tipo mixing. Vamos verificar que o estimador é consistente em média quadrática, mais precisamente,

$$\lim_{n\to\infty} E\left\{[f_n(x) - f(x)]^2\right\} = 0,\ \forall x \in C_\nu(f)\,,$$

e consequentemente fracamente consistente, isto é,

$$\lim_{n\to\infty} P(|f_n(x) - f(x)| > \epsilon) = 0,\ \forall x \in C_\nu(f)\,.$$

E para finalizar iremos verificar ainda a consistência forte, ou seja,

$$\lim_{n\to\infty} f_n(x) = f(x) \;\; \nu - q.c.\ (quase\ certamente),\ \forall x \in C_\nu(f).$$

Nesse caso será usado apenas que a cadeia satisfaz a condição do tipo mixing, não será usado que a cadeia é estritamente estacionária.

2.2 Não vício assintótico

O objetivo desta seção é verificar que os estimadores dados por

$$f_n(x) = \frac{1}{n}\sum_{k=1}^{n} W(h, x, X_k),\ h = h_n \downarrow 0,$$

com a função peso $W(h, x, \cdot)$ e a sequência h satisfazendo as Condições 2.1 e 2.2 dadas mais adiante, são assintoticamente não viciado (ou assintoticamente não tendencioso).

E o que seria o não vício de um estimador?

De maneira natural defini-se o estimador $f_n(x)$ como *não viciado* se $E(f_n(x)) = f(x)$, para todo $x \in \mathbb{R}^d$. Em outras palavras, um estimador é não viciado se o seu valor esperado coincide com a função de interesse.

No entanto, essa definição não tem utilidade na estimação de funções de densidade pois não existem estimadores não viciado, existem sim estimadores assintoticamente não viciado para funções de densidade contínua, ver Rosenblatt (1956). Dizemos que $f_n(\cdot)$ é assintoticamente não viciado se $\lim_{n\to\infty} E(f_n(x)) = f(x)$, $\forall x \in \mathbb{R}^d$.

Antes de verificar o não vício será necessário que W e h satisfaçam a seguinte condição:

Condição 2.1 *Para $x \in C_\nu(f)$ e $h = h_n \downarrow 0$ assuma que $W(h, x, \cdot)$ são funções de densidade com respeito a ν. Assim*

$$\int_E |W(h, x, y)|\, \nu(dy) \leq K_0(x) < \infty. \tag{2.4}$$

Além disso, dado $\delta > 0$

$$|W_\delta(h, x, y)| \leq K_\delta(x) < \infty \tag{2.5}$$

e

$$\lim_{h\to 0} W_\delta(h,x,y) = 0, \tag{2.6}$$

onde

$$W_\delta(h,x,y) = W(h,x,y)\, I_{\{z:|z-x|>\delta\}}(y). \tag{2.7}$$

Para a demonstração do não vício vamos considerar a cadeia estritamente estacionária como em (2.2) e o seguinte lema:

Lema 2.1 (Campos e Dorea (2001)) *Seja $(E,\mathcal{E},\lambda)$ um espaço mensurável σ-finito. Para $x \in E$ fixado e $h > 0$ assuma que $V(h,x,\cdot)$ são funções a valores reais satisfazendo (2.4), (2.5) e (2.6). Então se ψ é uma função integrável e $x \in C_\lambda(\psi)$ temos*

$$\lim_{h\to 0}\left|\int_E V(h,x,y)\,\psi(y)\,\lambda(dy) - \psi(x)\int_E V(h,x,y)\,\lambda(dy)\right| = 0.$$

Observe que o não vício assintótico, mostrado a seguir, é uma aplicação do Lema 2.1.

Proposição 2.1 (Não vício assintótico) *Se a Condição 2.1 é satisfeita então*

$$\lim_{n\to\infty} E(f_n(x)) = \lim_{h\to 0} EW(h,x,X_1) = f(x),\ \forall x \in C_\nu(f). \tag{2.8}$$

Demonstração: Pela estacionariedade estrita do processo, sabendo que quando $n \to \infty$ $h \to 0$ e pelo Lema 2.1 temos

$$\begin{aligned}
\lim_{n\to\infty} E(f_n(x)) &= \lim_{n\to\infty} E\left(\frac{1}{n}\sum_{k=1}^{n} W(h,x,X_k)\right) = \lim_{n\to\infty}\frac{1}{n}E\left(\sum_{k=1}^{n} W(h,x,X_k)\right)\\
&= \lim_{n\to\infty}\frac{1}{n}\sum_{k=1}^{n} E(W(h,x,X_k)) = \lim_{n\to\infty}\frac{1}{n}nE(W(h,x,X_1))\\
&= \lim_{n\to\infty} EW(h,x,X_1) = \lim_{h\to 0} EW(h,x,X_1)\\
&= \lim_{h\to 0}\int_E W(h,x,y)\,f(y)\,\nu(dy) = \lim_{h\to 0} f(x)\int_E W(h,x,y)\,\nu(dy) = f(x),
\end{aligned}$$

pois, pela Condição 2.1, $W(h,x,\cdot)$ é uma função densidade.

Logo,

$$\lim_{n\to\infty} E(f_n(x)) = \lim_{h\to 0} EW(h,x,X_1) = f(x),\ \forall\, x \in C_\nu(f).$$

Portanto, $f_n(x)$ é um estimador assintoticamente não viciado. ■

Agora vamos dar um exemplo que mostra o não vício assintótico dessa classe de estimadores sendo f a densidade de uma variável aleatória com distribuição Cauchy padrão.

Exemplo 2.1 *Sejam $X_1, X_2, ..., X_n$ variáveis aleatórias i.i.d. . Sejam $f(x) = \frac{1}{\pi(1+x^2)}$ com $x \in \mathbb{R}$ a densidade de uma variável aleatória com distribuição Cauchy e $W(h,x,y) = \frac{1}{h\sqrt{2\pi}} e^{\frac{-(x-y)^2}{2h^2}}$. Vamos mostrar que,*

$$\lim_{n\to\infty} E(f_n(x)) = f(x).$$

Inicialmente vamos verificar que W satisfaz a Condição 2.1.

Com efeito, sabemos que por definição W é uma densidade, pois está definida como sendo a densidade de uma normal com média x e variância h^2. Por outro lado temos que dado $\delta > 0$ para $|x-y| > \delta$ tem-se que $\frac{|x-y|}{h} > \frac{\delta}{h} \overset{h\to 0}{\to} \infty$ e assim

$$\begin{aligned} W_\delta(h,x,y) &= \frac{1}{h\sqrt{2\pi}} e^{\frac{-(x-y)^2}{2h^2}} I_{\{y:|x-y|>\delta\}}(y) \\ &\leq \frac{1}{h\sqrt{2\pi}} e^{\frac{-\delta^2}{2h^2}} I_{\{y:|x-y|>\delta\}}(y) \overset{h\to 0}{\to} 0, \end{aligned}$$

logo W satifaz (2.6).

E derivando $W_\delta(h,x,y)$ com relação a h podemos observar que a função é crescente em h para $h < \delta$ e tomando $h_0 \leq \delta$ temos para $0 < h \leq h_0$ que

$$\begin{aligned} W_\delta(h,x,y) &= \frac{1}{h\sqrt{2\pi}} e^{\frac{-(x-y)^2}{2h^2}} I_{\{y:|x-y|>\delta\}}(y) \\ &\leq \frac{1}{h_0\sqrt{2\pi}} e^{\frac{-(x-y)^2}{2h_0^2}} I_{\{y:|x-y|>\delta\}}(y) \\ &\leq \frac{1}{h_0\sqrt{2\pi}} e^{\frac{-\delta^2}{2h_0^2}} I_{\{y:|x-y|>\delta\}}(y) = K_\delta(x), \end{aligned}$$

então W satisfaz (2.5).

Agora vamos mostrar que

$$f_n(x) = \frac{1}{n}\sum_{k=1}^{n} \frac{1}{h\sqrt{2\pi}} e^{\frac{-(x-X_k)^2}{2h^2}},$$

é assintoticamente não viciado.

De fato,

$$\begin{aligned}
|E(f_n(x)) - f(x)| &= \left| E\left(\frac{1}{n}\sum_{k=1}^{n}\frac{1}{h\sqrt{2\pi}}e^{\frac{-(x-X_k)^2}{2h^2}}\right) - \frac{1}{\pi(1+x^2)}\right| \\
&= \left| \frac{1}{n}\sum_{k=1}^{n} E\left(\frac{1}{h\sqrt{2\pi}}e^{\frac{-(x-X_k)^2}{2h^2}}\right) - \frac{1}{\pi(1+x^2)}\right| \\
&= \left| E\left(\frac{1}{h\sqrt{2\pi}}e^{\frac{-(x-X_1)^2}{2h^2}}\right) - \frac{1}{\pi(1+x^2)}\right| \\
&= \left| \int_{-\infty}^{\infty}\frac{1}{h\sqrt{2\pi}\pi(1+y^2)}e^{\frac{-(x-y)^2}{2h^2}}dy - \frac{1}{\pi(1+x^2)}\right| \\
&= \left| \frac{1}{\pi}\left[\int_{-\infty}^{\infty}\frac{1}{h\sqrt{2\pi}(1+y^2)}e^{\frac{-(x-y)^2}{2h^2}}dy - \frac{1}{(1+x^2)}\right]\right|,
\end{aligned}$$

fazendo uma mudança de variável $t = \frac{x-y}{\sqrt{2}h}$ temos que $\frac{dt}{dy} = \frac{-1}{\sqrt{2}h}$ assim tem-se que $dt = \frac{-dy}{\sqrt{2}h}$, então

$$\begin{aligned}
&\left| \frac{1}{\pi}\left[\int_{-\infty}^{\infty}\frac{1}{h\sqrt{2\pi}(1+y^2)}e^{\frac{-(x-y)^2}{2h^2}}dy - \frac{1}{(1+x^2)}\right]\right| \\
= &\left| \frac{1}{\pi}\left[\int_{\infty}^{-\infty}\frac{-e^{-t^2}}{\sqrt{\pi}\left(1+\left(x-\sqrt{2}ht\right)^2\right)}dt - \frac{1}{(1+x^2)}\right]\right| \\
= &\left| \frac{1}{\pi}\left[\int_{-\infty}^{\infty}\frac{e^{-t^2}}{\sqrt{\pi}\left(1+\left(x-\sqrt{2}ht\right)^2\right)}dt - \frac{1}{(1+x^2)}\int_{-\infty}^{\infty}\frac{e^{-t^2}}{\sqrt{\pi}}dt\right]\right|,
\end{aligned}$$

passando o limite quando $n \to \infty$ temos que $h \to 0$, então segue-se que

$$\begin{aligned}
\lim_{n\to\infty}|E(f_n(x)) - f(x)| &= \lim_{h\to 0}\left| \frac{1}{\pi}\left[\int_{-\infty}^{\infty}\frac{e^{-t^2}}{\sqrt{\pi}\left(1+\left(x-\sqrt{2}ht\right)^2\right)}dt - \int_{-\infty}^{\infty}\frac{e^{-t^2}}{\sqrt{\pi}(1+x^2)}dt\right]\right| \\
&= \frac{1}{\pi}\left| \lim_{h\to 0}\left[\int_{-\infty}^{\infty}\frac{e^{-t^2}}{\sqrt{\pi}\left(1+\left(x-\sqrt{2}ht\right)^2\right)}dt - \int_{-\infty}^{\infty}\frac{e^{-t^2}}{\sqrt{\pi}(1+x^2)}dt\right]\right| \\
&= \frac{1}{\pi}\left| \lim_{h\to 0}\int_{-\infty}^{\infty}\frac{e^{-t^2}}{\sqrt{\pi}\left(1+\left(x-\sqrt{2}ht\right)^2\right)}dt - \lim_{h\to 0}\int_{-\infty}^{\infty}\frac{e^{-t^2}}{\sqrt{\pi}(1+x^2)}dt\right|,
\end{aligned}$$

e pelo Teorema da Convergência Dominada tem-se

$$\begin{aligned}
&\frac{1}{\pi}\left| \lim_{h\to 0}\int_{-\infty}^{\infty}\frac{e^{-t^2}}{\sqrt{\pi}\left(1+\left(x-\sqrt{2}ht\right)^2\right)}dt - \lim_{h\to 0}\int_{-\infty}^{\infty}\frac{e^{-t^2}}{\sqrt{\pi}(1+x^2)}dt\right| \\
= &\frac{1}{\pi}\left| \int_{-\infty}^{\infty}\lim_{h\to 0}\frac{e^{-t^2}}{\sqrt{\pi}\left(1+\left(x-\sqrt{2}ht\right)^2\right)}dt - \lim_{h\to 0}\int_{-\infty}^{\infty}\frac{e^{-t^2}}{\sqrt{\pi}(1+x^2)}dt\right| \\
= &\frac{1}{\pi}\left| \int_{-\infty}^{\infty}\frac{e^{-t^2}}{\sqrt{\pi}(1+x^2)}dt - \int_{-\infty}^{\infty}\frac{e^{-t^2}}{\sqrt{\pi}(1+x^2)}dt\right| = 0.
\end{aligned}$$

Portanto $E(f_n(x)) \to f(x)$, ou seja, $f_n(x)$ é assintoticamente não viciado.

2.3 Consistência

O objetivo desta seção é verificar que os estimadores dados por

$$f_n(x) = \frac{1}{n}\sum_{k=1}^{n} W(h, x, X_k), \;\; h = h_n \downarrow 0$$

e a função peso $W(h, x, \cdot)$, são consistentes.

Em geral, juntamente com a propriedade de não vício do estimador $f_n(x)$, estuda-se a sua consistência. Dependendo da forma de convergência, observa-se diferentes formas de consistência.

Estimadores consistentes são bastante estudados por que a medida que o tamanho da amostra aumenta os estimadores ficam tão próximos quanto se deseja.

Se $f_n(x) \to f(x)$ em probabilidade para todo $x \in \mathbb{R}^d$, afirma-se que $f_n(x)$ é fracamente consistente. Se a convergência acontece quase certamente, $f_n(x)$ é fortemente consistente. Outros tipos de convergência dependem do que se entenda por critério de erro.

Se $E\left\{[f_n(x) - f(x)]^2\right\} \to 0$ para todo $x \in \mathbb{R}^d$, quando $n \to \infty$, dizemos que a sequência $f_n(x)$ é consistente em erro quadrático médio.

Além de considerarmos que a cadeia é estritamente estacionária, vamos assumir ainda que satisfaz a condição φ-mixing dada em (2.3).

Nos exemplos a seguir iremos discutir a importância dessas hipóteses. Nestes exemplos consideraremos m a medida de Lebesgue.

Exemplo 2.2 *Considere $E = [0, 2]$ e o núcleo de transição dado por*

$$P(x, dy) = \left[I_{[0,1]^2}(x, y) + I_{(1,2]^2}(x, y)\right] dy.$$

Observe que este exemplo também poderia ser enunciado da seguinte forma:

Sejam $U_1, U_2, U_3, \ldots$ variáveis aleatórias com distribuição $U[0, 1]$ e considere a cadeia $X_0, X_1, X_2, \ldots$ com $X_0 \in [0, 2]$,

$$X_i = U_i \text{ com probabilidade } 1 \text{ se } X_{i-1} \in [0,1]$$

$$X_i = 1 + U_i \text{ com probabilidade } 1 \text{ se } X_{i-1} \notin [0,1].$$

Iremos considerar os três casos a seguir:

a) Considere $f_1(x) = I_{[0,1]}(x)$.

Assim temos que

$$f_1(A) = \int_A f_1(x)\,dx = \int_A I_{[0,1]}(x)\,dx = \int_{A\cap[0,1]} dx = m(A\cap[0,1]),$$

e

$$\begin{aligned} f_1P(A) &= \int_E P(x,A) f_1(x)\,dx = \int_E P(x,A) I_{[0,1]}(x)\,dx = \int_{[0,1]} P(x,A)\,dx \\ &= \int_{[0,1]}\int_A P(x,dy)\,dx = \int_{[0,1]}\int_{[0,1]\cap A} dydx = m(A\cap[0,1]) = f_1(A), \end{aligned}$$

ou seja, $f_1(x)$ é uma densidade estacionária e portanto

$$\begin{aligned} P_{f_1}(X_0\in A) = P_{f_1}(X_n\in A) &= \int_A f_1(x)\,dx = \int_A I_{[0,1]}(x)\,dx \\ &= \int_{A\cap[0,1]} dx = m(A\cap[0,1]), \forall n\geq 1. \end{aligned} \quad (2.9)$$

Por outro lado temos

$$\begin{aligned} P_{f_1}(X_0\in A,\, X_1\in B) &= \int_B\int_A P(x,dy) f_1(x)\,dx = \int_B\int_A P(x,dy) I_{[0,1]}(x)\,dx \\ &= \int_B\int_{A\cap[0,1]} P(x,dy)\,dx = \int_{B\cap[0,1]}\int_{A\cap[0,1]} dydx \\ &= m(A\cap[0,1])\, m(B\cap[0,1]), \end{aligned}$$

e

$$\begin{aligned} P_{f_1}(X_0\in A,\, X_2\in B) &= \int_B\int_E\int_A P(y,dz) P(x,dy) f_1(x)\,dx \\ &= \int_B\int_E\int_A P(y,dz) P(x,dy) I_{[0,1]}(x)\,dx \\ &= \int_B\int_E\int_{A\cap[0,1]} P(y,dz) P(x,dy)\,dx = \int_B\int_{[0,1]}\int_{A\cap[0,1]} P(y,dz)\,dydx \\ &= \int_{B\cap[0,1]}\int_{[0,1]}\int_{A\cap[0,1]} dzdydx = m(A\cap[0,1])\, m(B\cap[0,1]), \end{aligned}$$

prosseguindo dessa maneira obtemos

$$P_{f_1}(X_0 \in A,\, X_n \in B) = m(A \cap [0,1])\, m(B \cap [0,1]),\ \forall n > 2. \tag{2.10}$$

Assim pela estacionariedade estrita do processo, pela propriedade de Markov, por (2.9) e (2.10) temos

$$\begin{aligned}
P_{f_1}(X_i \in A,\, X_{i+k} \in B) &= P_{f_1}(X_{i+k} \in B |\, X_i \in A)\, P_{f_1}(X_i \in A) \\
&= P_{f_1}(X_k \in B |\, X_0 \in A)\, P_{f_1}(X_i \in A) \\
&= P_{f_1}(X_k \in B |\, X_0 \in A)\, P_{f_1}(X_0 \in A) \\
&= P_{f_1}(X_k \in B,\, X_0 \in A) = m(A \cap [0,1])\, m(B \cap [0,1]),\ \forall i \geq 0 \ e\ \forall k \geq 1,
\end{aligned}$$

que juntamente com (2.9) nos garante que a cadeia é φ_{f_1}-mixing com $\varphi_{f_1}(n) = 0$.

b)Considere $f_2(x) = I_{(1,2]}(x)$.

Assim temos que

$$f_2(A) = \int_A f_2(x)\, dx = \int_A I_{(1,2]}(x)\, dx = \int_{A \cap (1,2]} dx = m(A \cap (1,2])$$

e

$$\begin{aligned}
f_2 P(A) &= \int_E P(x, A)\, f_2(x)\, dx = \int_E P(x, A)\, I_{(1,2]}(x)\, dx = \int_{(1,2]} P(x, A)\, dx \\
&= \int_{(1,2]} \int_A P(x, dy)\, dx = \int_{(1,2]} \int_{(1,2] \cap A} dy dx = m(A \cap (1,2]) = f_2(A),
\end{aligned}$$

ou seja, $f_2(x)$ é uma densidade estacionária e portanto

$$\begin{aligned}
P_{f_2}(X_0 \in A) = P_{f_2}(X_n \in A) &= \int_A I_{(1,2]}(x)\, dx = \int_{A \cap (1,2]} dx \\
&= \int_A f_2(x)\, dx = m(A \cap (1,2]),\forall n \geq 1. \qquad (2.11)
\end{aligned}$$

Por outro lado temos

$$\begin{aligned}
P_{f_2}(X_0 \in A,\, X_1 \in B) &= \int_B \int_A P(x, dy)\, f_2(x)\, dx = \int_B \int_A P(x, dy)\, I_{(1,2]}(x)\, dx \\
&= \int_B \int_{A \cap (1,2]} P(x, dy)\, dx = \int_{B \cap (1,2]} \int_{A \cap (1,2]} dy dx \\
&= m(A \cap (1,2])\, m(B \cap (1,2]),
\end{aligned}$$

e

$$\begin{aligned}
P_{f_2}(X_0 \in A,\, X_2 \in B) &= \int_B \int_E \int_A P(y,dz)\, P(x,dy)\, f_2(x)\, dx \\
&= \int_B \int_E \int_A P(y,dz)\, P(x,dy)\, I_{(1,2]}(x)\, dx \\
&= \int_B \int_E \int_{A\cap(1,2]} P(y,dz)\, P(x,dy)\, dx = \int_B \int_{(1,2]} \int_{A\cap(1,2]} P(y,dz)\, dydx \\
&= \int_{B\cap(1,2]} \int_{(1,2]} \int_{A\cap(1,2]} dzdydx = m(A\cap(1,2])\, m(B\cap(1,2])\,,
\end{aligned}$$

prosseguindo dessa maneira obtemos

$$P_{f_2}(X_0 \in A,\, X_n \in B) = m(A\cap(1,2])\, m(B\cap(1,2])\,,\ \forall n > 2. \tag{2.12}$$

Assim pela estacionariedade estrita do processo, pela propriedade de Markov, por (2.11) e (2.12) temos

$$\begin{aligned}
P_{f_2}(X_i \in A,\, X_{i+k} \in B) &= P_{f_2}(X_{i+k} \in B|\, X_i \in A)\, P_{f_2}(X_i \in A) \\
&= P_{f_2}(X_k \in B|\, X_0 \in A)\, P_{f_2}(X_i \in A) \\
&= P_{f_2}(X_k \in B|\, X_0 \in A)\, P_{f_2}(X_0 \in A) \\
&= P_{f_2}(X_k \in B,\, X_0 \in A) = m(A\cap(1,2])\, m(B\cap(1,2])\,,\ \forall i \geq 0 \ e\ \forall k \geq 1,
\end{aligned}$$

que juntamente com (2.11) nos garante que a cadeia é φ_{f_2}-mixing com $\varphi_{f_2}(n) = 0$.

Exemplo 2.3 *Sob as mesmas condições do Exemplo 2.2, consideremos* $f_3(x) = 2I_{[0,\frac{1}{2}]}(x)$.

Vamos considerar $f_3(x)$ a densidade inicial da cadeia, sendo assim,

$$P_{f_3}(X_0 \in A) = \int_A f_3(x)\, dx = \int_A 2I_{[0,\frac{1}{2}]}(x)\, dx = 2\int_{A\cap[0,\frac{1}{2}]} dx = 2m\left(A \cap \left[0, \frac{1}{2}\right]\right). \tag{2.13}$$

Verificamos que $f_3(x)$ não é uma densidade estacionária pois,

$$f_3(A) = \int_A f_3(x)\, dx = \int_A 2I_{[0,\frac{1}{2}]}(x)\, dx = 2\int_{A\cap[0,\frac{1}{2}]} dx = 2m\left(A \cap \left[0, \frac{1}{2}\right]\right)$$

e no entanto

$$\begin{aligned} f_3P(A) &= \int_E P(x,A)\, f_3(x)\, dx = \int_E P(x,A)\, 2I_{\left[0,\frac{1}{2}\right]}(x)\, dx \\ &= 2\int_{\left[0,\frac{1}{2}\right]} P(x,A)\, dx = 2\int_{\left[0,\frac{1}{2}\right]}\int_A P(x,dy)\, dx \\ &= 2\int_{\left[0,\frac{1}{2}\right]}\int_{A\cap[0,1]} dydx = m(A\cap[0,1]). \end{aligned}$$

Logo $f_3(x)$ não é uma densidade estacionária.

Mas por outro lado temos que

$$\begin{aligned} P_{f_3}(X_1\in A) &= \int_E P(x,A)\, f_3(x)\, dx = \int_E 2I_{\left[0,\frac{1}{2}\right]}(x)\, P(x,A)\, dx \\ &= 2\int_{\left[0,\frac{1}{2}\right]} P(x,A)\, dx = 2\int_{\left[0,\frac{1}{2}\right]}\int_A P(x,dy)\, dx \\ &= 2\int_{\left[0,\frac{1}{2}\right]}\int_{A\cap[0,1]} dydx = m(A\cap[0,1]). \end{aligned}$$

Para $x\in[0,1]$

$$\begin{aligned} P^2(x,A) &= \int_E P(x,dy)\, P(y,A) = \int_{[0,1]} P(y,A)\, dy \\ &= \int_{[0,1]}\int_A P(y,dz)\, dy = m(A\cap[0,1]) \end{aligned}$$

o que implica em

$$\begin{aligned} P_{f_3}(X_2\in A) &= \int_E P^2(x,A)\, f_3(x)\, dx = \int_E P^2(x,A)\, 2I_{\left[0,\frac{1}{2}\right]}(x)\, dx \\ &= 2\int_{\left[0,\frac{1}{2}\right]} P^2(x,A)\, dx = m(A\cap[0,1]) \end{aligned}$$

e de maneira análoga,

$$P_{f_3}(X_n\in A) = m(A\cap[0,1]),\ \forall n>2, \tag{2.14}$$

$\forall n\geq 1$, ou seja, as variáveis aletórias $X_1, X_2, ...$ têm distribuição $U[0,1]$ (Uniforme no intervalo $[0,1]$).

Para a distribuição conjunta de X_0 e X_2 temos

$$\begin{aligned}
P_{f_3}(X_0 \in A, X_2 \in B) &= \int_B \int_E \int_A P(y,dz)\, P(x,dy)\, f_3(x)\, dx \\
&= \int_B \int_E \int_A P(y,dz)\, P(x,dy)\, 2I_{\left[0,\frac{1}{2}\right]}(x)\, dx \\
&= 2\int_B \int_E \int_{A\cap\left[0,\frac{1}{2}\right]} P(y,dz)\, P(x,dy)\, dx = 2\int_B \int_{[0,1]} \int_{A\cap\left[0,\frac{1}{2}\right]} P(y,dz)\, dydx \\
&= 2\int_{B\cap[0,1]} \int_{[0,1]} \int_{A\cap\left[0,\frac{1}{2}\right]} dzdydx = 2m\left(A\cap\left[0,\frac{1}{2}\right]\right) m\left(B\cap[0,1]\right).
\end{aligned}$$

Da mesma forma podemos mostrar que

$$P_{f_3}(X_0 \in A, X_n \in B) = 2m\left(A\cap\left[0,\frac{1}{2}\right]\right) m\left(B\cap[0,1]\right),\ \forall n > 2,$$

$\forall n \geq 1$ e assim por (2.13) e (2.14) temos

$$P_{f_3}(X_0 \in A)\, P_{f_3}(X_n \in B) = P_{f_3}(X_0 \in A, X_n \in B),\ \forall n \geq 1. \tag{2.15}$$

Para a distribuição conjunta de X_1 e X_3 temos

$$\begin{aligned}
P_{f_3}(X_1 \in A, X_3 \in B) &= \int_B \int_E \int_{A\cap[0,1]} P(y,dz)\, P(x,dy)\, dx = \int_B \int_{[0,1]} \int_{A\cap[0,1]} P(y,dz)\, dydx \\
&= \int_{B\cap[0,1]} \int_{[0,1]} \int_{A\cap[0,1]} dzdydx = m\left(A\cap[0,1]\right) m\left(B\cap[0,1]\right).
\end{aligned}$$

Da mesma forma podemos mostrar que

$$P_{f_3}(X_1 \in A, X_n \in B) = m\left(A\cap[0,1]\right) m\left(B\cap[0,1]\right),\ \forall n > 3$$

$\forall n \geq 2$.

Sendo assim, usando (2.14) e a propruedade de Markov

$$\begin{aligned}
P_{f_3}(X_i \in A, X_{i+k} \in B) &= P_{f_3}(X_{i+k} \in B|\, X_i \in A)\, P_{f_3}(X_i \in A) \\
&= P_{f_3}(X_{k+1} \in B|\, X_1 \in A)\, P_{f_3}(X_i \in A) \\
&= P_{f_3}(X_{k+1} \in B|\, X_1 \in A)\, P_{f_3}(X_1 \in A) \\
&= P_{f_3}(X_{k+1} \in B,\ X_1 \in A) = m\left(A\cap[0,1]\right) m\left(B\cap[0,1]\right),\ \forall i \geq 1\ e\ \forall k \geq 1.
\end{aligned}$$

Novamente por (2.14)

$$P_{f_3}(X_i \in A)\, P_{f_3}(X_{i+k} \in B) = P_{f_3}(X_i \in A, X_{i+k} \in B),\ \forall i \geq 1\ e\ \forall k \geq 1, \tag{2.16}$$

que juntamente com (2.15) nos garante que a cadeia é φ_{f_3}-mixing com $\varphi_{f_3}(n) = 0$.

Considere o estimador definido por

$$f_n(x) = \frac{1}{n}\sum_{k=1}^{n} W(h, x, X_k)$$

com

$$W(h, x, y) = \frac{1}{h}K\left(\frac{x-y}{h}\right)$$

e

$$K(z) = \begin{cases} \frac{1}{2} & \text{se} \quad z \in [-1, 1) \\ 0 & \text{se} \quad z \notin [-1, 1) \end{cases},$$

ou seja, o estimador ingenuo.

Por (2.16) temos que as variáveis $X_1, X_2, ..., X_n$ são i.i.d. com distribuição $U[0, 1]$ e sabemos, pela Lei Forte dos Grandes Números (Kintchine), que o estimador ingenuo converge em probabildade para a densidade comum das variáveis, ou seja,

$$f_n(x) = \frac{1}{n}\sum_{k=1}^{n} W(h, x, X_k) \xrightarrow{P} I_{[0,1]}(x).$$

Mas por outro lado, se as condições do exemplo fossem suficientes para provarmos a consistência fraca, teriamos que ter

$$\frac{1}{n}\sum_{k=1}^{n} W(h, x, X_k) \xrightarrow{P} f_3(x) = 2I_{[0,\frac{1}{2}]}(x).$$

c)Considere a cadeia de Markov com núcleo de transição dado por

$$P = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix},$$

e distribuição inicial $f_4 = \left(\frac{1}{4}, \frac{1}{2}, \frac{1}{4}\right)$.

Podemos verificar que f_4 é uma distribuição estacionária para a cadeia, isto é, $f_4 P = f_4$ e portanto

$$P_{f_4}(X_1 = 1) = P_{f_4}(X_2 = 1) = ... = P_{f_4}(X_n = 1) = P_{f_4}(X_0 = 1) = \frac{1}{4}.$$

Mas a cadeia não é φ_{f_4}-mixing, pois $P_{f_4}(X_1 = 1, X_2 = 1) = 0$ e assim

$$|P_{f_4}(X_1 = 1, X_2 = 1) - P_{f_4}(X_1 = 1) P_{f_4}(X_2 = 1)| = \frac{1}{8}.$$

Considere o estimador

$$f_n(i) = \frac{1}{n} \sum_{k=1}^{n} I_{(X_k = i)}.$$

Observe que $f_n(1) = \frac{1}{n} \sum_{k=1}^{n} I_{(X_k=1)}$ dependendo do ponto em que a cadeia começar, convergira para valores diferentes de $f_4(1) = \frac{1}{4}$. De fato, se a cadeia começar no estado 1 ou 3 teremos

$$f_n(x) = \frac{1}{n} \sum_{k=1}^{n} I_{(X_k=1)} \to \frac{1}{2}$$

e se começar no estado 2 teremos

$$f_n(x) = \frac{1}{n} \sum_{k=1}^{n} I_{(X_k=1)} \to 0.$$

Ou seja, as condições do exemplo não são suficientes para garantir a consistência fraca.

Observação: Como pode ser observado no Exemplo 2.2 (a) a cadeia é φ_{f_1}-mixing e se considerarmos $f_1(x)$ como sendo a densidade inicial da cadeia estaremos, através dos resultados de Campos e Dorea (2005), estimando assim $f_1(x)$. De maneira análoga, no Exemplo 2.2 (b) a cadeia é φ_{f_2}-mixing e considerando $f_2(x)$ como sendo a densidade inicial, estimamos $f_2(x)$. Já no exemplo 2.2 (c) verificamos que a condição φ-mixing não é suficiente para verificar os resultados de consistência se não supormos a densidade inicial como sendo a estacionária. E no exemplo 2.4 verificamos que a hipótese de distribuição inicial como sendo a estacionária não garante os resultados de consistência, será necessário que a cadeia satisfaça a condição φ-mixing.

Observe que a condição φ-mixing é colocada nas estruturas de dependência da cadeia, mais precisamente, esta definição nos permite usar uma desigualdade de momentos que será muito importante para a verificação do próximo resultado, pois não foi assumido independência para as variáveis.

Lema 2.2 (Roussas e Ioannides (1987)) *Seja $\{X_n\}_{n\geq 0}$ φ-mixing e assuma que ξ e η são variáveis aleatórias $\mathcal{F}_0^k$-mensurável e $\mathcal{F}_{k+n}^{\infty}$-mensurável respectivamente tais que*

$$E\left|\xi\right|^p < \infty,\ E\left|\eta\right|^q < \infty,$$

para $p > 1$, $q > 1$ com $\frac{1}{p} + \frac{1}{q} = 1$. Então a covariância satisfaz

$$\left|cov\left(\xi, \eta\right)\right| \leq 2\varphi^{1/p}\left(n\right)\left(E\left|\xi\right|^p\right)^{1/p}\left(E\left|\eta\right|^q\right)^{1/q}. \tag{2.17}$$

E vamos considerar ainda que $W\left(h, x, \cdot\right)$ satisfaz também a seguinte condição.

Condição 2.2 *Assuma que a Condição 2.1 é satisfeita e que para $\gamma_n\left(x\right) = \nu\left\{y : \left|y - x\right| \leq h\right\}$ temos*

$$\lim_{n\to\infty} \gamma_n\left(x\right) = \gamma\left(x\right) < \infty,\ \lim_{n\to\infty} n\gamma_n\left(x\right) = \infty \tag{2.18}$$

e

$$\gamma_n\left(x\right) W\left(h, x, y\right) \leq K_1\left(x\right) < \infty \tag{2.19}$$

para h pequeno.

Agora temos ferramentas suficientes para verificar a consistência em média quadrática e a consistência fraca dos estimadores definidos em (1.13), que é verificada a partir da consistência em média quadrática usando a desigualdade de Tchebyschev.

Teorema 2.1 (Consistência em média quadrática e consistência fraca) *Se a Condição 2.2 é satisfeita e $\sum_{n=1}^{\infty} \varphi^{1/2}\left(n\right) < \infty$, temos a consistência em média quadrática*

$$\lim_{n\to\infty} E\left\{\left[f_n\left(x\right) - f\left(x\right)\right]^2\right\} = 0,\ \forall\ x \in C_\nu\left(f\right) \tag{2.20}$$

e a consistência fraca, dado $\epsilon > 0$

$$\lim_{n\to\infty} P\left(\left|f_n\left(x\right) - f\left(x\right)\right| > \epsilon\right) = 0,\ \forall\ x \in C_\nu\left(f\right). \tag{2.21}$$

Demonstração: Observe que

$$
\begin{aligned}
E\left\{\left[f_n(x)-f(x)\right]^2\right\} &= E\left[f_n^2(x)-2f_n(x)f(x)+f^2(x)\right]\\
&= E\left[f_n^2(x)\right]-E\left[2f_n(x)f(x)\right]+E\left[f^2(x)\right]\\
&= E\left[f_n^2(x)\right]-2f(x)E\left[f_n(x)\right]+f^2(x)\\
&= E\left[f_n^2(x)\right]-2f(x)E\left[f_n(x)\right]+f^2(x)+\left\{E\left[f_n(x)\right]\right\}^2-\left\{E\left[f_n(x)\right]\right\}^2\\
&= \left\{E\left[f_n^2(x)\right]-\left\{E\left[f_n(x)\right]\right\}^2\right\}\\
&+ \left\{\left\{E\left[f_n(x)\right]\right\}^2-2f(x)E\left[f_n(x)\right]+f^2(x)\right\}\\
&= var\left(f_n(x)\right)+\left[E\left(f_n(x)\right)-f(x)\right]^2,
\end{aligned}
$$

mas pela Proposição 2.1, $E\left(f_n(x)\right)\to f(x)$, então basta mostrar que $var\left(f_n(x)\right)\to 0$.

Para isto vamos mostrar que $n\gamma_n(x)\,var\left(f_n(x)\right)$ é limitado e como $n\gamma_n(x)\overset{n\to\infty}{\to}\infty$ pela Condição 2.2 teremos que $var\left(f_n(x)\right)\to 0$.

Com efeito, pela estacionariedade estrita do processo temos

$$
\begin{aligned}
n\gamma_n(x)\,var\left(f_n(x)\right) &= n\gamma_n(x)\,var\left(\frac{1}{n}\sum_{k=1}^{n}W\left(h,x,X_k\right)\right)=n\gamma_n(x)\frac{1}{n^2}var\left(\sum_{k=1}^{n}W\left(h,x,X_k\right)\right)\\
&= \gamma_n(x)\frac{1}{n}var\left(\sum_{k=1}^{n}W\left(h,x,X_k\right)\right)=\gamma_n(x)\frac{1}{n}\sum_{k=1}^{n}var\left[W\left(h,x,X_k\right)\right]\\
&+ \gamma_n(x)\frac{2}{n}\sum_{k=1}^{n-1}\sum_{j=k+1}^{n}cov\left[W\left(h,x,X_k\right),W\left(h,x,X_j\right)\right]\\
&= \gamma_n(x)\frac{1}{n}nvar\left[W\left(h,x,X_1\right)\right]\\
&+ \gamma_n(x)\frac{2}{n}\sum_{k=1}^{n-1}\sum_{j=k+1}^{n}cov\left[W\left(h,x,X_k\right),W\left(h,x,X_j\right)\right]\\
&= \gamma_n(x)\,var\left(W\left(h,x,X_1\right)\right)+\gamma_n(x)\frac{2}{n}\sum_{j=1+1}^{n}cov\left[W\left(h,x,X_1\right),W\left(h,x,X_j\right)\right]\\
&+ \ldots+\gamma_n(x)\frac{2}{n}\sum_{j=(n-1)+1}^{n}cov\left[W\left(h,x,X_{n-1}\right),W\left(h,x,X_j\right)\right]\\
&= \gamma_n(x)\,var\left(W\left(h,x,X_1\right)\right)\\
&+ \gamma_n(x)\frac{2}{n}(n-1)\,cov\left[W\left(h,x,X_1\right),W\left(h,x,X_2\right)\right]\\
&+ \ldots+\gamma_n(x)\frac{2}{n}(n-(n-1))\,cov\left[W\left(h,x,X_{n-1}\right),W\left(h,x,X_n\right)\right]
\end{aligned}
$$

$$
\begin{aligned}
&= \gamma_n(x)\, var\left(W(h,x,X_1)\right) \\
&+ \gamma_n(x)\,\frac{2}{n}\sum_{k=1}^{n-1}(n-k)\,cov\left[W(h,x,X_1), W(h,x,X_{1+k})\right]. \qquad (2.22)
\end{aligned}
$$

Vamos mostrar que (2.22) é limitada. Na primeira parcela de (2.22) temos

$$
\begin{aligned}
\gamma_n(x)\, var\left(W(h,x,X_1)\right) &= \gamma_n(x)\left\{E\left(W^2(h,x,X_1)\right) - \left[E\left(W(h,x,X_1)\right)\right]^2\right\} \\
&= \gamma_n(x)\, E\left(W^2(h,x,X_1)\right) - \gamma_n(x)\left[E\left(W(h,x,X_1)\right)\right]^2
\end{aligned}
$$

mas, por (2.18) e pelo Lema (2.1) temos

$$
\begin{aligned}
\lim_{n\to\infty}\gamma_n(x)\left[E\left(W(h,x,X_1)\right)\right]^2 &= \lim_{n\to\infty}\left\{\gamma_n(x)\left[\int_E W(h,x,y)\,f(y)\,\nu(dy)\right]^2\right\} \\
&= \gamma(x)\lim_{h\to 0}\left[f(x)\int_E W(h,x,y)\,\nu(dy)\right]^2 \\
&= \gamma(x)\, f^2(x) < \infty. \qquad (2.23)
\end{aligned}
$$

Agora defina,

$$
V(h,x,y) = \gamma_n(x)\, W^2(h,x,X_1)
$$

então por (2.7), (2.19) e (2.5) temos,

$$
\begin{aligned}
V_\delta(h,x,y) &= \gamma_n(x)\, W_\delta^2(h,x,X_1) = \gamma_n(x)\, W_\delta(h,x,X_1)\, W_\delta(h,x,X_1) \\
&\leq \gamma_n(x)\, W(h,x,X_1)\, W_\delta(h,x,X_1) \leq K_1(x)\, W_\delta(h,x,X_1) \leq K_1(x)\, K_\delta(x) < \infty.
\end{aligned}
$$

Além disso, usando (2.6) e (2.18) temos

$$
\begin{aligned}
\lim_{h\to 0} V_\delta(h,x,y) &= \lim_{h\to 0}\left\{\gamma_n(x)\, W_\delta^2(h,x,X_1)\right\} \\
&= \lim_{n\to\infty}\gamma_n(x)\lim_{h\to 0} W_\delta^2(h,x,X_1) = 0.
\end{aligned}
$$

E ainda,

$$
\begin{aligned}
\int_E V(h,x,y)\,\nu(dy) &= \int_E \gamma_n(x)\, W^2(h,x,y)\,\nu(dy) \\
&\leq \int_E K_1(x)\, W(h,x,y)\,\nu(dy) = K_1(x)\int_E W(h,x,y)\,\nu(dy) \\
&= K_1(x) < \infty. \qquad (2.24)
\end{aligned}
$$

Assim como $V(h,x,y)$ satisfaz as hipóteses do Lema 2.1 temos,

$$\begin{aligned}
\lim_{n\to\infty}\left\{\gamma_n(x)\,E\left[W^2(h,x,X_1)\right]\right\} &= \lim_{n\to\infty} E\left[\gamma_n(x)\,W^2(h,x,X_1)\right]\\
&= \lim_{n\to\infty}\int_E \gamma_n(x)\,W^2(h,x,y)\,f(y)\,\nu(dy)\\
&= \lim_{h\to 0}\int_E V(h,x,y)\,f(y)\,\nu(dy)\\
&= f(x)\lim_{h\to 0}\int_E V(h,x,y)\,\nu(dy)\\
&= f(x)\lim_{h\to 0}\int_E \gamma_n(x)\,W^2(h,x,y)\,\nu(dy)\\
&= f(x)\lim_{h\to 0}\int_E \gamma_n(x)\,W(h,x,y)\,W(h,x,y)\,\nu(dy)\\
&\le f(x)\,K_1(x)\lim_{h\to 0}\int_E W(h,x,y)\,\nu(dy)\\
&= f(x)\,K_1(x)<\infty. \qquad (2.25)
\end{aligned}$$

Segue que

$$\limsup_{n\to\infty}\gamma_n(x)\,E\left(W^2(h,x,X_1)\right)<\infty.$$

Isto junto com (2.23) assegura que

$$\limsup_{n\to\infty}\gamma_n(x)\,var\left(W(h,x,X_1)\right)<\infty,$$

mais precisamente

$$\limsup_{n\to\infty}\gamma_n(x)\,var\left(W(h,x,X_1)\right)<M_1<\infty. \qquad (2.26)$$

Agora vamos mostrar que a segunda parcela de (2.22) também é limitada. Tomando $\xi=\gamma_n^{1/2}(x)\,W(h,x,X_1)$ e $\eta=\gamma_n^{1/2}(x)\,W(h,x,X_{1+k})$ e por (2.24) podemos fazer uso do Lema 2.2 com $p=q=2$.

Assim, pela estacionariedade estrita do processo e por (2.25) temos

$$\begin{aligned}
cov(\xi,\eta) &= cov\left(\gamma_n^{1/2}(x)\,W(h,x,X_1),\gamma_n^{1/2}(x)\,W(h,x,X_{1+k})\right)\\
&\le 2\varphi^{1/2}(k)\left(E\left(\gamma_n^{1/2}(x)\,W(h,x,X_1)\right)^2\right)^{1/2}\left(E\left(\gamma_n^{1/2}(x)\,W(h,x,X_{1+k})\right)^2\right)^{1/2}\\
&= 2\varphi^{1/2}(k)\,\gamma_n^{1/2}(x)\left(E\left(W^2(h,x,X_1)\right)\right)^{1/2}\gamma_n^{1/2}(x)\left(E\left(W^2(h,x,X_{1+k})\right)\right)^{1/2}\\
&= 2\varphi^{1/2}(k)\,\gamma_n(x)\left(E\left(W^2(h,x,X_1)\right)\right)\le 2\varphi^{1/2}(k)\,f(x)\,K_1(x)\\
&= 2\varphi^{1/2}(k)\,M_2, \qquad (2.27)
\end{aligned}$$

onde $M_2 = f(x) K_1(x)$.

Logo por (2.26) e (2.27) temos,

$$\begin{aligned} n\gamma_n(x)\, var(f_n(x)) &\leq M_1 + \frac{2}{n}\sum_{k=1}^{n-1}(n-k)\, 2\varphi^{1/2}(k)\, M_2 \\ &= M_1 + \frac{4M_2}{n}\sum_{k=1}^{n-1}(n-k)\,\varphi^{1/2}(k) \\ &\leq M_1 + 4M_2\sum_{k=1}^{n-1}\varphi^{1/2}(k) < \infty, \end{aligned}$$

pois, por hipótese, $\sum_{k=1}^{n-1}\varphi^{1/2}(k) < \infty$.

Assim, como $n\gamma_n(x)\, var(f_n(x))$ é limitada, dividindo ambos os lados de (2.22) por $n\gamma_n(x)$ temos,

$$var(f_n(x)) = \frac{\gamma_n(x)\, var(W(h,x,X_1)) + \gamma_n(x)\frac{2}{n}\sum_{k=1}^{n-1}(n-k)\, cov[W(h,x,X_1), W(h,x,X_{1+k})]}{n\gamma_n(x)},$$

e passando o limite quando $n \to \infty$ fica

$$\lim_{n\to\infty} var(f_n(x)) = 0,$$

pois por (2.18) $\lim_{n\to\infty} n\gamma_n(x) = \infty$.

Portanto, $f_n(x)$ é consistente em média quadrática.

Já (2.21) segue de (2.20) pela desigualdade de Tchebyschev, ou seja, dado $\epsilon > 0$ temos

$$0 \leq P(|f_n(x) - f(x)| > \epsilon) \leq \frac{E\left\{[f_n(x) - f(x)]^2\right\}}{\epsilon^2}.$$

Assim,

$$\lim_{n\to\infty} P(|f_n(x) - f(x)| > \epsilon) \leq \lim_{n\to\infty}\frac{E\left\{[f_n(x) - f(x)]^2\right\}}{\epsilon^2} = 0.$$

Portanto, $f_n(x)$ é um estimador fracamente consistente. ■

Para verificarmos a consistência forte, além das condições dadas anteriormente iremos precisar dos seguintes resultados.

Lema 2.3 (Roussas e Ioannides (1987)) *Seja $\{X_n\}_{n\geq 0}$ φ-mixing e assuma que η é uma variável aleatória $\mathcal{F}_{k+n}^{\infty}$-mensurável tal que $|\eta| < M$ q.c. então*

$$\left|E\left(\eta|\mathcal{F}_0^k\right) - E\eta\right| \leq 2\varphi(n)\, M, \text{ q.c. }.$$

Lema 2.4 (Devroye (1991)) *Seja* $\mathcal{G}_0 = \{\emptyset, \Omega\} \subset \mathcal{G}_1 \subset \ldots \subset \mathcal{G}_l$ *uma sequência de* σ*-álgebras encaixadas. Seja* U *uma variável aleatória* $\mathcal{G}_l$*-mensurável e integrável. Defina uma Doob martingale* $U_l = E(U|\mathcal{G}_l)$*. Assuma que existe uma variável aleatória* $\mathcal{G}_{l-1}$*-mensurável* V_l *e constantes* a_l *tais que* $V_l \leq U_l \leq V_l + a_l$*. Então dado* $\epsilon > 0$

$$P(|U - EU| \geq \epsilon) \leq 4exp\left\{-\frac{2\epsilon^2}{\sum_{l=1}^{n} {a_l}^2}\right\}.$$

Outro resultado que iremos usar é a primeira parte da proposição a seguir que é uma consequência do Lema de Borel-Cantelli.

Proposição 2.2 *Seja* $\{X_n\}_{n\geq 1}$ *uma sequência de variáveis aleatórias sob algum espaço de probabilidade* $(\Omega, \mathcal{F}, P)$*.*

(i) Se $\sum_{n=1}^{\infty} P(|X_n| > \epsilon) < \infty$ *para cada* $\epsilon > 0$ *então*

$$P\left(\lim_{n\to\infty} X_n = 0\right) = 1.$$

(ii) Se $\{X_n\}_{n\geq 1}$ *são duas a duas independentes e* $P(\lim_{n\to\infty} X_n = 0) = 1$ *então*

$$\sum_{n=1}^{\infty} P(|X_n| > \epsilon) < \infty,$$

para cada $\epsilon > 0$*.*

Para a prova deste resultado consultar Durret (2005).

Teorema 2.2 (Consistência forte) *Se a Condição 2.2 é satisfeita,* $\sum_{n\geq 1} \varphi(n) < \infty$ *e se para todo* $\beta > 0$

$$\sum_{n\geq 1} exp\left\{-n\beta\gamma_n^2(x)\right\} < \infty. \tag{2.28}$$

Então

$$\lim_{n\to\infty} f_n(x) = f(x), \ \nu - q.c., \ \forall x \in C_\nu(f) \tag{2.29}$$

e

$$P\left(\lim_{n\to\infty} f_n(x) = f(x)\right) = 1.$$

Demonstração: A demonstração feita em Campos e Dorea (2005) é baseada nas idéias de Dorea e Zhao (2002).

Pela Proposição 2.1 temos

$$\lim_{n\to\infty} E\left(f_n\left(x\right)\right) = f\left(x\right),$$

assim é suficiente mostrar que

$$\lim_{n\to\infty} \left[f_n\left(x\right) - E\left(f_n\left(x\right)\right)\right] = 0, \ \nu - q.c. \ .$$

Com efeito, defina as seguintes funções auxiliares,

$$\psi_k\left(X_k\right) = \sum_{j\geq 0} \left[E\left(\gamma_n\left(x\right) W\left(h, x, X_{k+j}\right) | \mathcal{F}_1^k\right) - E\left(\gamma_n\left(x\right) W\left(h, x, X_{k+j}\right)\right)\right]$$

e

$$\psi_{k+1}\left(X_k\right) = \sum_{j\geq 0} \left[E\left(\gamma_n\left(x\right) W\left(h, x, X_{k+j+1}\right) | \mathcal{F}_1^k\right) - E\left(\gamma_n\left(x\right) W\left(h, x, X_{k+j+1}\right)\right)\right],$$

onde $\mathcal{F}_1^k = \sigma\left(X_1, X_2, ..., X_k\right)$.

Inicialmente vamos mostrar que essas funções estão bem definidas. Considerando $\eta = \gamma_n\left(x\right) W\left(h, x, X_{k+l+j}\right)$ no Lema (2.3), observamos que por (2.19) $|\eta| \leq K_1\left(x\right)$. Então pelo Lema 2.3 temos ,

$$\begin{aligned}
\sum_{j\geq 0} \left|E\left(\eta | \mathcal{F}_1^k\right) - E\eta\right| &= \sum_{j\geq 1} \left|E\left(\gamma_n\left(x\right) W\left(h, x, X_{k+l+j}\right) | \mathcal{F}_1^k\right) - E\left(\gamma_n\left(x\right) W\left(h, x, X_{k+l+j}\right)\right)\right| \\
&\leq \sum_{j\geq 0} 2\varphi\left(l+j\right) K_1\left(x\right) < \infty.
\end{aligned}$$

Assim considerando

$$L\left(x\right) = \sum_{j\geq 0} 2\varphi\left(j+l\right) K_1\left(x\right),$$

com $l = 0, 1$, temos

$$\psi_k\left(X_k\right) \leq L\left(x\right) \ \ e \ \psi_{k+1}\left(X_k\right) \leq L\left(x\right). \tag{2.30}$$

Assim,

$$\begin{aligned}
\psi_k(X_k) - \psi_{k+1}(X_k) &= \sum_{j\geq 0}\left[E\left(\gamma_n(x)W(h,x,X_{k+j})\,|\mathcal{F}_1^k\right) - E\left(\gamma_n(x)W(h,x,X_{k+j})\right)\right] \\
&- \sum_{j\geq 0}\left[E\left(\gamma_n(x)W(h,x,X_{k+j+1})\,|\mathcal{F}_1^k\right) - E\left(\gamma_n(x)W(h,x,X_{k+j+1})\right)\right] \\
&= E\left(\gamma_n(x)W(h,x,X_k)\,|\mathcal{F}_1^k\right) - E\left(\gamma_n(x)W(h,x,X_k)\right) \\
&+ \sum_{j\geq 1}\left[E\left(\gamma_n(x)W(h,x,X_{k+j})\,|\mathcal{F}_1^k\right) - E\left(\gamma_n(x)W(h,x,X_{k+j})\right)\right] \\
&- \sum_{j\geq 0}\left[E\left(\gamma_n(x)W(h,x,X_{k+j+1})\,|\mathcal{F}_1^k\right) - E\left(\gamma_n(x)W(h,x,X_{k+j+1})\right)\right] \\
&= E\left(\gamma_n(x)W(h,x,X_k)\,|\mathcal{F}_1^k\right) - E\left(\gamma_n(x)W(h,x,X_k)\right) \\
&= \gamma_n(x)\left[W(h,x,X_k) - E\left(W(h,x,X_k)\right)\right]
\end{aligned}$$

pois $W(h,x,X_k)$ é $\mathcal{F}_1^k$ mensurável, e

$$\begin{aligned}
n\gamma_n(x)\left[f_n(x) - E\left(f_n(x)\right)\right] &= n\gamma_n(x)\left[\frac{1}{n}\sum_{k=1}^{n}W(h,x,X_k) - E\left(\frac{1}{n}\sum_{k=1}^{n}W(h,x,X_k)\right)\right] \\
&= \gamma_n(x)\left[\sum_{k=1}^{n}W(h,x,X_k) - E\left(\sum_{k=1}^{n}W(h,x,X_k)\right)\right] \\
&= \sum_{k=1}^{n}\gamma_n(x)\left[W(h,x,X_k) - E\left(W(h,x,X_k)\right)\right] \\
&= \sum_{k=1}^{n}\left[\psi_k(X_k) - \psi_{k+1}(X_k)\right] \\
&= \psi_1(X_1) - \psi_2(X_1) + \psi_2(X_2) - \psi_3(X_2) + ... \\
&+ \psi_n(X_n) - \psi_{n+1}(X_n) \\
&= \psi_1(X_1) - \psi_{n+1}(X_n) + \sum_{k=2}^{n}\left[\psi_k(X_k) - \psi_k(X_{k-1})\right].
\end{aligned}$$

Além disso,

$$\begin{aligned}
&P\left(|f_n(x) - E\left(f_n(x)\right)| \geq \epsilon\right) = P\left(\left|\frac{n\gamma_n(x)}{n\gamma_n(x)}\left[f_n(x) - E\left(f_n(x)\right)\right]\right| \geq \epsilon\right) \\
= \; &P\left(|n\gamma_n(x)\left[f_n(x) - E\left(f_n(x)\right)\right]| \geq \epsilon n\gamma_n(x)\right) \\
= \; &P\left(\left|\psi_1(X_1) - \psi_{n+1}(X_n) + \sum_{k=2}^{n}\left[\psi_k(X_k) - \psi_k(X_{k-1})\right]\right| \geq n\gamma_n(x)\,\epsilon\right) \\
\leq \; &P\left(|\psi_1(X_1) - \psi_{n+1}(X_n)| \geq \frac{n\epsilon\gamma_n(x)}{2}\right) + P\left(\left|\sum_{k=2}^{n}\left[\psi_k(X_k) - \psi_k(X_{k-1})\right]\right| \geq \frac{n\epsilon\gamma_n(x)}{2}\right),
\end{aligned}$$

logo

$$\begin{aligned}
& P\left(\left|f_n(x)-E\left(f_n(x)\right)\right| \geq \epsilon\right) \\
\leq\ & P\left(\left|\psi_1\left(X_1\right)-\psi_{n+1}\left(X_n\right)\right| \geq \frac{n \epsilon \gamma_n(x)}{2}\right) \quad (2.31)\\
+\ & P\left(\left|\sum_{k=2}^n\left[\psi_k\left(X_k\right)-\psi_k\left(X_{k-1}\right)\right]\right| \geq \frac{n \epsilon \gamma_n(x)}{2}\right). \quad (2.32)
\end{aligned}$$

Desse modo, vamos mostrar que dado $\epsilon>0$, existe $\beta(\epsilon)>0$ tal que,

$$P\left(\left|f_n(x)-E\left(f_n(x)\right)\right| \geq \epsilon\right)=4 exp\left\{n \gamma_n^2(x) \beta(\epsilon)\right\},$$

então por (2.27) e pela Proposição 2.2 obtemos a convergência desejada.

De fato, como as funções ψ estão bem definidas segue de (2.30) que,

$$\begin{aligned}
\left|\psi_1\left(X_1\right)-\psi_{n+1}\left(X_n\right)\right| & =\left|\psi_1\left(X_1\right)+\left(-\psi_{n+1}\left(X_n\right)\right)\right| \leq\left|\psi_1\left(X_1\right)\right|+\left|\left(-\psi_{n+1}\left(X_n\right)\right)\right| \\
& =\left|\psi_1\left(X_1\right)\right|+\left|\psi_{n+1}\left(X_n\right)\right| \leq L(x)+L(x)=2 L(x).
\end{aligned}$$

Assim, para n grande de (2.31) temos,

$$P\left(\left|\psi_1\left(X_1\right)-\psi_{n+1}\left(X_n\right)\right| \geq \frac{\epsilon n \gamma_n(x)}{2}\right)=0,$$

pois, por (2.18) $n \gamma_n(x) \rightarrow \infty$ quando $n \rightarrow \infty$.

Para (2.32) vamos usar o Lema 2.4. Seja $\mathcal{G}_l=\mathcal{F}_1^l$, assim pelo Teorema da Convergência Monótona (ver Magalhães (2006)) e pelas propriedades de esperança condicional, para $k>l$ temos

$$\begin{aligned}
& E\left(\psi_k\left(X_k\right) \mid \mathcal{G}_l\right)=E\left(\sum_{j \geq 0}\left[E\left(\gamma_n(x) W\left(h, x, X_{k+j}\right) \mid \mathcal{G}_k\right)-E\left(\gamma_n(x) W\left(h, x, X_{k+j}\right)\right)\right] \mid \mathcal{G}_l\right) \\
= & \sum_{j \geq 0} E\left(\left[E\left(\gamma_n(x) W\left(h, x, X_{k+j}\right) \mid \mathcal{G}_k\right)-E\left(\gamma_n(x) W\left(h, x, X_{k+j}\right)\right)\right] \mid \mathcal{G}_l\right) \\
= & \sum_{j \geq 0}\left(E\left[E\left(\gamma_n(x) W\left(h, x, X_{k+j}\right) \mid \mathcal{G}_k\right) \mid \mathcal{G}_l\right]-E\left[E\left(\gamma_n(x) W\left(h, x, X_{k+j}\right)\right) \mid \mathcal{G}_l\right]\right) \\
= & \sum_{j \geq 0}\left[E\left(\gamma_n(x) W\left(h, x, X_{k+j}\right) \mid \mathcal{G}_l\right)-E\left(\gamma_n(x) W\left(h, x, X_{k+j}\right)\right)\right]=\psi_k\left(X_l\right).
\end{aligned}$$

Analogamente,

$$\begin{aligned}
& E\left(\psi_k\left(X_{k-1}\right)|\mathcal{G}_l\right) \\
= \; & E\left(\sum_{j\geq 0}\left[E\left(\gamma_n\left(x\right)W\left(h,x,X_{k-1+j}\right)|\mathcal{G}_{k-1}\right)-E\left(\gamma_n\left(x\right)W\left(h,x,X_{k-1+j}\right)\right)\right]|\mathcal{G}_l\right) \\
= \; & \sum_{j\geq 0} E\left(\left[E\left(\gamma_n\left(x\right)W\left(h,x,X_{k-1+j}\right)|\mathcal{G}_{k-1}\right)-E\left(\gamma_n\left(x\right)W\left(h,x,X_{k-1+j}\right)\right)\right]|\mathcal{G}_l\right) \\
= \; & \sum_{j\geq 0}\left(E\left[E\left(\gamma_n\left(x\right)W\left(h,x,X_{k-1+j}\right)|\mathcal{G}_{k-1}\right)|\mathcal{G}_l\right]-E\left[E\left(\gamma_n\left(x\right)W\left(h,x,X_{k-1+j}\right)\right)|\mathcal{G}_l\right]\right) \\
= \; & \sum_{j\geq 0}\left[E\left(\gamma_n\left(x\right)W\left(h,x,X_{k-1+j}\right)|\mathcal{G}_l\right)-E\left(\gamma_n\left(x\right)W\left(h,x,X_{k-1+j}\right)\right)\right]=\psi_k\left(X_l\right).
\end{aligned}$$

Agora defina $U_l = E\left(U|\mathcal{G}_l\right)$, então segue-se que

$$\begin{aligned}
U_l &= E\left(U|\mathcal{G}_l\right)=E\left(\sum_{k=2}^{n}\left[\psi_k\left(X_k\right)-\psi_k\left(X_{k-1}\right)\right]|\mathcal{G}_l\right) \\
&= \sum_{k=2}^{n}\left[E\left(\psi_k\left(X_k\right)|\mathcal{G}_l\right)-E\left(\psi_k\left(X_{k-1}\right)|\mathcal{G}_l\right)\right] \\
&= \sum_{k=2}^{l}\left[E\left(\psi_k\left(X_k\right)|\mathcal{G}_l\right)-E\left(\psi_k\left(X_{k-1}\right)|\mathcal{G}_l\right)\right]+\sum_{k=l+1}^{n}\left[E\left(\psi_k\left(X_k\right)|\mathcal{G}_l\right)-E\left(\psi_k\left(X_{k-1}\right)|\mathcal{G}_l\right)\right] \\
&= \sum_{k=2}^{l}\left[\psi_k\left(X_k\right)-\psi_k\left(X_{k-1}\right)\right]+\sum_{k=l+1}^{n}\left[\psi_k\left(X_l\right)-\psi_k\left(X_l\right)\right]=\sum_{k=2}^{l}\left[\psi_k\left(X_k\right)-\psi_k\left(X_{k-1}\right)\right]
\end{aligned}$$

e

$$\begin{aligned}
\left|U_l-U_{l-1}\right| &= \left|\sum_{k=2}^{l}\left[\psi_k\left(X_k\right)-\psi_k\left(X_{k-1}\right)\right]-\sum_{k=2}^{l-1}\left[\psi_k\left(X_k\right)-\psi_k\left(X_{k-1}\right)\right]\right| \\
&= \left|\psi_2\left(X_2\right)-\psi_2\left(X_1\right)+\psi_3\left(X_3\right)-\psi_3\left(X_2\right)+...+\psi_l\left(X_l\right)-\psi_l\left(X_{l-1}\right)\right. \\
&- \left.\psi_2\left(X_2\right)+\psi_2\left(X_1\right)-\psi_3\left(X_3\right)+\psi_3\left(X_2\right)+...-\psi_{l-1}\left(X_{l-1}\right)+\psi_{l-1}\left(X_{l-2}\right)\right| \\
&= \left|\psi_l\left(X_l\right)-\psi_l\left(X_{l-1}\right)\right|=\left|\psi_l\left(X_l\right)+\left(-\psi_l\left(X_{l-1}\right)\right)\right| \\
&\leq \left|\psi_l\left(X_l\right)\right|+\left|\left(-\psi_l\left(X_{l-1}\right)\right)\right|=\left|\psi_l\left(X_l\right)\right|+\left|\psi_l\left(X_{l-1}\right)\right|\leq L\left(x\right)+L\left(x\right)=2L\left(x\right).
\end{aligned}$$

Por outro lado, definindo $U=\sum_{k=2}^{n}\left[\psi_k\left(X_k\right)-\psi_k\left(X_{k-1}\right)\right]$ e usando novamente o Teorema da Convergência Monótona temos

$$
\begin{aligned}
EU &= E\left(\sum_{k=2}^{n}\left[\psi_k\left(X_k\right)-\psi_k\left(X_{k-1}\right)\right]\right)=\sum_{k=2}^{n}E\left[\psi_k\left(X_k\right)-\psi_k\left(X_{k-1}\right)\right]\\
&= \sum_{k=2}^{n}E\left[\psi_k\left(X_k\right)\right]-\sum_{k=2}^{n}E\left[\psi_k\left(X_{k-1}\right)\right]\\
&= \sum_{k=2}^{n}E\left(\sum_{j\geq 0}\left[E\left(\gamma_n\left(x\right)W\left(h,x,X_{k+j}\right)|\mathcal{F}_1^k\right)-E\left(\gamma_n\left(x\right)W\left(h,x,X_{k+j}\right)\right)\right]\right)\\
&- \sum_{k=2}^{n}E\left(\sum_{j\geq 0}\left[E\left(\gamma_n\left(x\right)W\left(h,x,X_{k+j+1}\right)|\mathcal{F}_1^k\right)-E\left(\gamma_n\left(x\right)W\left(h,x,X_{k+j+1}\right)\right)\right]\right)\\
&= \sum_{k=2}^{n}\left\{\sum_{j\geq 0}E\left[E\left(\gamma_n\left(x\right)W\left(h,x,X_{k+j}\right)|\mathcal{F}_1^k\right)-E\left(\gamma_n\left(x\right)W\left(h,x,X_{k+j}\right)\right)\right]\right\}\\
&- \sum_{k=2}^{n}\left\{\sum_{j\geq 0}E\left[E\left(\gamma_n\left(x\right)W\left(h,x,X_{k+j+1}\right)|\mathcal{F}_1^k\right)-E\left(\gamma_n\left(x\right)W\left(h,x,X_{k+j+1}\right)\right)\right]\right\}\\
&= \sum_{k=2}^{n}\left\{\sum_{j\geq 0}\left[E\left[E\left(\gamma_n\left(x\right)W\left(h,x,X_{k+j}\right)|\mathcal{F}_1^k\right)\right]-E\left[E\left(\gamma_n\left(x\right)W\left(h,x,X_{k+j}\right)\right)\right]\right]\right\}\\
&- \sum_{k=2}^{n}\left\{\sum_{j\geq 0}\left[E\left[E\left(\gamma_n\left(x\right)W\left(h,x,X_{k+j+1}\right)|\mathcal{F}_1^k\right)\right]-E\left[E\left(\gamma_n\left(x\right)W\left(h,x,X_{k+j+1}\right)\right)\right]\right]\right\}\\
&= \sum_{k=2}^{n}\left\{\sum_{j\geq 0}\left[E\left(\gamma_n\left(x\right)W\left(h,x,X_{k+j}\right)\right)-E\left(\gamma_n\left(x\right)W\left(h,x,X_{k+j}\right)\right)\right]\right\}\\
&- \sum_{k=2}^{n}\left\{\sum_{j\geq 0}\left[E\left(\gamma_n\left(x\right)W\left(h,x,X_{k+j+1}\right)\right)-E\left(\gamma_n\left(x\right)W\left(h,x,X_{k+j+1}\right)\right)\right]\right\}=0.
\end{aligned}
$$

Logo tomando $V_l = U_{l-1} - 2L\left(x\right)$ e $a_l = 4L\left(x\right)$ no Lema 2.4 temos,

$$
\begin{aligned}
P\left(\left|U-EU\right|\geq\epsilon\right) &= P\left(\left|\sum_{k=2}^{n}\left[\psi_k\left(X_k\right)-\psi_k\left(X_{k-1}\right)\right]\right|\geq\frac{n\gamma_n\left(x\right)\epsilon}{2}\right)\leq 4exp\left\{\frac{-2n^2\gamma_n\left(x\right)^2\epsilon^2}{4\sum_{k=1}^{n}16L\left(x\right)^2}\right\}\\
&= 4exp\left\{\frac{-2n^2\gamma_n\left(x\right)^2\epsilon^2}{4n16L\left(x\right)^2}\right\}=4exp\left\{\frac{-n\gamma_n\left(x\right)^2\epsilon^2}{32L\left(x\right)^2}\right\}=4exp\left\{-n\gamma_n\left(x\right)^2\beta\left(\epsilon\right)\right\},
\end{aligned}
$$

onde $\beta\left(\epsilon\right)=\frac{\epsilon^2}{32L^2(x)}$.

Como,

$$\begin{aligned} P\left(\left|f_n(x)-E\left(f_n(x)\right)\right|\geq\epsilon\right) &\leq P\left(\left|\psi_1\left(X_1\right)-\psi_{n+1}\left(X_n\right)\right|\geq\frac{n\gamma_n(x)\,\epsilon}{2}\right)\\ &+ P\left(\left|\sum_{k=2}^{n}\left[\psi_k\left(X_k\right)-\psi_k\left(X_{k-1}\right)\right]\right|\geq\frac{n\gamma_n(x)\,\epsilon}{2}\right), \end{aligned}$$

e sendo

$$\lim_{n\to\infty}P\left(\left|\sum_{k=2}^{n}\left[\psi_k\left(X_k\right)-\psi_k\left(X_{k-1}\right)\right]\right|\geq\frac{n\gamma_n(x)\,\epsilon}{2}\right)\leq\lim_{n\to\infty}4exp\left\{-n\gamma_n^2(x)\,\beta(\epsilon)\right\}=0,$$

temos

$$\lim_{n\to\infty}P\left(\left|\sum_{k=2}^{n}\left[\psi_k\left(X_k\right)-\psi_k\left(X_{k-1}\right)\right]\right|\geq\frac{n\gamma_n(x)\,\epsilon}{2}\right)=0,$$

e já vimos que

$$\lim_{n\to\infty}P\left(\left|\psi_1\left(X_1\right)-\psi_{n+1}\left(X_n\right)\right|\geq\frac{n\gamma_n(x)\,\epsilon}{2}\right)=0,$$

então

$$\lim_{n\to\infty}P\left(\left|f_n(x)-E\left(f_n(x)\right)\right|\geq\epsilon\right)=0.$$

Logo, $f_n(x)\overset{\nu-q.c}{\to}f(x)$.

Portanto, $f_n(x)$ é um estimador fortemente consistente. ∎

Observação: Note que, nesta prova a suposição de que a cadeia é estritamente estacionária não foi usada, exceto no não vício assintótico (Proposição 2.1).

Capítulo 3

Estimação da Densidade Limite sob a Ergodicidade Geométrica

3.1 Introdução

Neste capítulo continuamos estudando o problema de estimar a densidade limite, a única estacionária, de uma cadeia de Markov com espaço de estados geral.

Seja $\{X_n\}_{n\geq 0}$ uma cadeia de Markov com valores em E, com núcleo de transição $\{P(x,A) : x \in E, A \in \mathcal{E}\}$ e seja $f(\cdot)$ a densidade limite da cadeia relativa a uma medida σ-finita ν em $\mathcal{E}$, ou seja, existe uma medida de probabilidade π tal que

$$\lim_{n\to\infty} P^n(x,A) = \pi(A), \ \forall x \in E, \ \forall A \in \mathcal{E},$$

onde

$$\pi(A) = \int_A f(y)\,\nu(dy), \ \forall A \in \mathcal{E}.$$

Campos e Dorea (2005) substituiu as suposições, de que a cadeia é φ-mixing e estritamente estacionária, consideradas no Capítulo 2, pela hipótese de ergodicidade geométrica, ou seja,

$$|P^n(x,A) - \pi(A)| \leq \alpha\rho^n, \ \forall\, x \in E, \ \forall\, A \in \mathcal{E}, \tag{3.1}$$

onde α e ρ são constantes com $\alpha > 0$ e $0 < \rho < 1$. E iremos verificar que $f_n(x)$ definido em (1.13) é um bom estimador para $f(\cdot)$.

Na seção 3.2 verificamos que os estimadores definidos em (1.13) são assintoticamente não viciados, ou seja,

$$\lim_{n\to\infty} E(f_n(x)) = f(x), \ \forall x \in C_\nu(f).$$

Na seção 3.3 verificamos que esses estimadores são consistentes em média quadrática, mais precisamente,

$$\lim_{n\to\infty} E\left\{[f_n(x) - f(x)]^2\right\} = 0, \ \forall x \in C_\nu(f),$$

e consequentemente fracamente consistentes, isto é,

$$\lim_{n\to\infty} P(|f_n(x) - f(x)| > \epsilon) = 0, \ \forall x \in C_\nu(f).$$

E na seção 3.4 verificamos a consistência forte, ou seja,

$$\lim_{n\to\infty} f_n(x) = f(x) \ \ \nu - q.c. \ (quase\ certamente), \ \forall x \in C_\nu(f).$$

Em Campos e Dorea (2005) a consistência forte foi obtida a partir da verificação que se a cadeia é geometricamente ergódica então ela satisfaz a condição φ-mixing com coeficiente mixing geométrico. Logo a consistência forte é verificada sob as mesmas condições do Capitulo 2, pois como podemos observar na demonstração do Teorema 2.2 não utilizamos a hipótese de estacionariedade estrita, utilizamos apenas a condição φ-mixing.

3.2 Não vício assintótico

Nosso objetivo nesta seção é verificar que o estimador

$$f_n(x) = \frac{1}{n}\sum_{k=1}^{n} W(h, x, X_k), \ h = h_n \downarrow 0$$

e função peso $W(h, x, \cdot)$ satisfazendo a condição (2.2), é assintoticamente não viciado supondo que a cadeia é geometricamente ergódica como em (3.1).

Seja $p_n(x)$ a densidade em relação a medida ν, da distribuição de X_n, isto é,

$$P_{p_0}(X_n \in A) = \int_A p_n(y)\,\nu(dy)\,,\ \forall\ A \in \mathcal{E}\ e\ \forall\, n \geq 0, \tag{3.2}$$

onde P_{p_0} indica que a densidade inicial é p_0. De forma similar vamos fazer uso da notação E_{p_0} para a esperança.

Para o não vicio assintótico vamos fazer uso da desigualdade a seguir,

$$\left|\int_A [p_n(y) - f(y)]\,\nu(dy)\right| \leq \alpha\rho^n,\ \forall\, A \in \mathcal{E}, \tag{3.3}$$

que é uma consequência da ergodicidade geométrica.

Com efeito, como a cadeia é geometricamente ergódica temos

$$\begin{aligned} \left|P_{p_0}(X_n \in A) - \pi(A)\right| &= \left|\int_E P^n(x,A)\,p_0(x)\,\nu(dx) - \int_E \pi(A)\,p_0(x)\,\nu(dx)\right| \\ &= \left|\int_E [P^n(x,A) - \pi(A)]\,p_0(x)\,\nu(dx)\right| \\ &\leq \int_E \left|P^n(x,A) - \pi(A)\right| p_0(x)\,\nu(dx) \leq \alpha\rho^n, \end{aligned}$$

mas por outro lado temos

$$\left|P_{p_0}(X_n \in A) - \pi(A)\right| = \left|\int_A [p_n(y) - f(y)]\,\nu(dy)\right|,$$

logo,

$$\left|\int_A [p_n(y) - f(y)]\,\nu(dy)\right| \leq \alpha\rho^n.$$

Proposição 3.1 (Não vício assintótico) *Se a Condição 2.2 é satisfeita então temos que $f_n(x)$ é assintoticamente não viciado, ou seja,*

$$\lim_{n\to\infty} E_{p_0}(f_n(x)) = f(x)\,,\ \forall\, x \in C_\nu(f)\,. \tag{3.4}$$

Demonstração: Seja $x \in C_\nu(f)$. Pelo Lema 2.1 dado $\epsilon > 0$ existe N tal que para todo $n \geq N$,

$$\left|\int_E W(h,x,y)\,f(y)\,\nu(dy) - f(x)\right| < \epsilon. \tag{3.5}$$

Mas por (2.19) sabemos que

$$\gamma_n(x)\,W(h,x,y) \leq K_1(x) < \infty,$$

logo tem-se que

$$W\left(h,x,y\right)\leq\frac{K_{1}\left(x\right)}{\gamma_{n}\left(x\right)}<\infty, se\ \gamma_{n}\left(x\right)>0. \tag{3.6}$$

Assim por (3.6) e (3.3) temos

$$\begin{aligned}
&\left|\int_{E}W\left(h,x,y\right)\left[p_{k}\left(y\right)-f\left(y\right)\right]\nu\left(dy\right)\right|\leq\left|\int_{E}\frac{K_{1}\left(x\right)}{\gamma_{n}\left(x\right)}\left[p_{k}\left(y\right)-f\left(y\right)\right]\nu\left(dy\right)\right|\\
= &\frac{K_{1}\left(x\right)}{\gamma_{n}\left(x\right)}\left|\int_{E}\left[p_{k}\left(y\right)-f\left(y\right)\right]\nu\left(dy\right)\right|\leq\frac{K_{1}\left(x\right)}{\gamma_{n}\left(x\right)}\alpha\rho^{k}.
\end{aligned}\tag{3.7}$$

Desse modo segue-se que

$$\begin{aligned}
&\left|E_{p_{0}}\left(f_{n}\left(x\right)\right)-f\left(x\right)\right|=\left|E\left(\frac{1}{n}\sum_{k=1}^{n}W\left(h,x,X_{k}\right)\right)-f\left(x\right)\right|\\
= &\left|\frac{1}{n}\sum_{k=1}^{n}E\left(W\left(h,x,X_{k}\right)\right)-f\left(x\right)\right|=\left|\frac{1}{n}\sum_{k=1}^{n}\int_{E}W\left(h,x,y\right)p_{k}\left(y\right)\nu\left(dy\right)-\frac{1}{n}\sum_{k=1}^{n}f\left(x\right)\right|\\
= &\left|\frac{1}{n}\sum_{k=1}^{n}\left[\int_{E}W\left(h,x,y\right)p_{k}\left(y\right)\nu\left(dy\right)-f\left(x\right)\right]\right|\leq\frac{1}{n}\sum_{k=1}^{n}\left|\int_{E}W\left(h,x,y\right)p_{k}\left(y\right)\nu\left(dy\right)-f\left(x\right)\right|\\
= &\frac{1}{n}\sum_{k=1}^{n}\left|\int_{E}W\left(h,x,y\right)p_{k}\left(y\right)\nu\left(dy\right)+\int_{E}W\left(h,x,y\right)f\left(y\right)\nu\left(dy\right)-\int_{E}W\left(h,x,y\right)f\left(y\right)\nu\left(dy\right)-f\left(x\right)\right|\\
= &\frac{1}{n}\sum_{k=1}^{n}\left|\int_{E}W\left(h,x,y\right)\left[p_{k}\left(y\right)-f\left(y\right)\right]\nu\left(dy\right)+\int_{E}W\left(h,x,y\right)f\left(y\right)\nu\left(dy\right)-f\left(x\right)\right|\\
\leq &\frac{1}{n}\sum_{k=1}^{n}\left\{\left|\int_{E}W\left(h,x,y\right)\left[p_{k}\left(y\right)-f\left(y\right)\right]\nu\left(dy\right)\right|+\left|\int_{E}W\left(h,x,y\right)f\left(y\right)\nu\left(dy\right)-f\left(x\right)\right|\right\}.
\end{aligned}$$

Assim, por (3.5) e (3.7) temos

$$\left|E_{p_{0}}\left(f_{n}\left(x\right)\right)-f\left(x\right)\right|<\frac{K_{1}\left(x\right)}{n\gamma_{n}\left(x\right)}\sum_{k=1}^{n}\alpha\rho^{k}+\epsilon,$$

sendo $\sum_{k=1}^{n}\alpha\rho^{k}<\infty$, pois $\alpha>0$ e $0<\rho<1$. Então quando $n\rightarrow\infty$, como ϵ é arbitrário e por (2.18) $n\gamma_{n}\left(x\right)\rightarrow\infty$, temos

$$E_{p_{0}}\left(f_{n}\left(x\right)\right)\rightarrow f\left(x\right).$$

Portanto $f_{n}\left(x\right)$ é um estimador assintoticamente não viciado. ■

3.3 Consistência em média quadrática e consistência fraca

Nesta seção vamos verificar a consistência em média quadrática dos estimadores aqui estudados e como consequência desse resultado a consistência fraca.

Vamos assumir, mesmo tendo por hipótese apenas a existência da densidade marginal das variáveis aleatórias X_1 e X_{k+1}, que o vetor aleatório (X_1, X_{1+k}) têm densidade conjunta denotada por $p_{1,1+k}$ com $k = 1, ..., n-1$. Pois se as densidades marginais existirem não implica que a densidade conjunta existe como pode ser visto no exemplo a seguir:

Exemplo 3.1 *Considere $E = [0,1] \times [0,1]$, $\mathcal{B}$ a σ-álgebra de E e seja $G = \{(x,y) \in E; x = y\}$, ou seja, a diagonal do quadrado unitário no plano. Para todo $B \in \mathcal{B}$ definimos*

$$P_{X,Y}(B) = P((X,Y) \in B) = \frac{\textit{comprimento de } (G \cap B)}{\sqrt{2}}.$$

Pode-se verificar que X e Y têm distribuição $U[0,1]$. No entanto a distribuição $P_{X,Y}$ não possui densidade. De fato, pela definição de distribuição conjunta temos que

$$P_{X,Y}(G) = 1.$$

Por outro lado, supondo que existe uma densidade conjunta $p(x,y)$, temos

$$P_{X,Y}(G) = \int\int_G p(x,y)\, dxdy = 0.$$

já que a medida de Lebesgue de de G é zero.

Teorema 3.1 (Consistência em média quadrática e consistência fraca) *Sob a Condição 2.2 temos a consistência em média quadrática*

$$\lim_{n\to\infty} E_{p_0}\left\{[f_n(x) - f(x)]^2\right\} = 0, \ \forall\, x \in C_\nu(f) \tag{3.8}$$

e a consistência fraca dado $\epsilon > 0$,

$$\lim_{n\to\infty} P_{p_0}(|f_n(x) - f(x)| > \epsilon) = 0, \ \forall\, x \in C_\nu(f). \tag{3.9}$$

Demonstração: Já vimos que

$$E_{p_0}\left[f_n(x) - f(x)\right]^2 = var_{p_0}\left(f_n(x)\right) + \left[E_{p_0}\left(f_n(x)\right) - f(x)\right]^2,$$

e que por (2.18) $n\gamma_n(x) \to \infty$, então basta mostrar que $n\gamma_n(x)\, var_{p_0}\left(f_n(x)\right)$ é limitado, pois como já vimos na Proposição 3.1 $E_{p_0}\left(f_n(x)\right) \to f(x)$.

Com efeito, temos que

$$\begin{aligned}
n\gamma_n(x)\, var_{p_0}\left(f_n(x)\right) &= n\gamma_n(x)\, var_{p_0}\left(\frac{1}{n}\sum_{k=1}^{n} W(h,x,X_k)\right) \\
&= \frac{\gamma_n(x)}{n} var_{p_0}\left(\sum_{k=1}^{n} W(h,x,X_k)\right) \\
&= \frac{\gamma_n(x)}{n}\sum_{k=1}^{n} var_{p_0}\left(W(h,x,X_k)\right) && (3.10) \\
&+ \frac{2\gamma_n(x)}{n}\sum_{1\le j\le k\le n} cov_{p_0}\left[W(h,x,X_j)\, W(h,x,X_k)\right]. && (3.11)
\end{aligned}$$

Para a limitação de (3.10) temos por (2.19) que

$$\begin{aligned}
\frac{\gamma_n(x)}{n}\sum_{k=1}^{n} var_{p_0}\left(W(h,x,X_k)\right) &= \frac{\gamma_n(x)}{n}\sum_{k=1}^{n}\left\{E_{p_0}\left[W^2(h,x,X_k)\right] - \left[E_{p_0}\left[W(h,x,X_k)\right]\right]^2\right\} \\
&\le \frac{\gamma_n(x)}{n}\sum_{k=1}^{n} E_{p_0}\left[W^2(h,x,X_k)\right] \\
&= \frac{\gamma_n(x)}{n}\sum_{k=1}^{n} E_{p_0}\left[W(h,x,X_k)\, W(h,x,X_k)\right] \\
&= \frac{1}{n}\sum_{k=1}^{n} E_{p_0}\left[\gamma_n(x)\, W(h,x,X_k)\, W(h,x,X_k)\right] \\
&\le \frac{1}{n}\sum_{k=1}^{n} E_{p_0}\left[K_1(x)\, W(h,x,X_k)\right] \\
&= \frac{K_1(x)}{n}\sum_{k=1}^{n} E_{p_0}\left[W(h,x,X_k)\right],
\end{aligned}$$

mas pela Proposição 3.1 $\frac{1}{n}\sum_{k=1}^{n} E_{p_0}\left[W(h,x,X_k)\right]$ é limitado. Logo $\frac{\gamma_n(x)}{n}\sum_{k=1}^{n} var_{p_0}\left(W(h,x,X_k)\right)$ é limitado.

Já para (3.11) temos que

$$
\begin{aligned}
& \frac{2\gamma_n(x)}{n} \sum_{1\leq j\leq k\leq n} cov_{p_0}\left[W(h,x,X_j), W(h,x,X_k)\right] \\
= \; & \frac{2\gamma_n(x)}{n} \sum_{1\leq j\leq k\leq n} \left\{E_{p_0}\left[W(h,x,X_j)W(h,x,X_k)\right] - E_{p_0}\left[W(h,x,X_j)\right]E_{p_0}\left[W(h,x,X_k)\right]\right\} \\
= \; & \frac{2\gamma_n(x)}{n} \sum_{1\leq j\leq k\leq n} \int_E \int_E W(h,x,u)W(h,x,v)\, p_{j,k}(u,v)\,\nu(du)\,\nu(dv) \\
- \; & \frac{2\gamma_n(x)}{n} \sum_{1\leq j\leq k\leq n} \int_E W(h,x,u)\, p_j(u)\,\nu(du) \int_E W(h,x,v)\, p_k(v)\,\nu(dv) \\
= \; & \frac{2\gamma_n(x)}{n} \sum_{1\leq j\leq k\leq n} \int_E \int_E W(h,x,u)W(h,x,v)\left[p_{j,k}(u,v) - p_j(u)\,p_k(v)\right]\nu(du)\,\nu(dv) \\
= \; & \frac{2\gamma_n(x)}{n} \sum_{1\leq j\leq k\leq n} \int_E W(h,x,u)\, p_j(u) \int_E p_j(y) \int_E W(h,x,v)\, P^{k-j}(u,v)\,\nu(dv)\,\nu(dy)\,\nu(du) \\
- \; & \frac{2\gamma_n(x)}{n} \sum_{1\leq j\leq k\leq n} \int_E W(h,x,u)\, p_j(u) \int_E p_j(y) \int_E W(h,x,v)\, P^{k-j}(y,v)\,\nu(dv)\,\nu(dy)\,\nu(du) \\
\leq \; & \frac{2\gamma_n(x)}{n} \sum_{1\leq j\leq k\leq n} \int_E W(h,x,u)\, p_j(u) \int_E p_j(y) \left|\int_E W(h,x,v)\left[P^{k-j}(u,v) - P^{k-j}(y,v)\right]\nu(dv)\right| \nu(dy)\,\nu(du)
\end{aligned}
$$

Por outro lado temos que (3.3) é equivalente a

$$
\left|\int_A \left[P^n(x,v) - f(v)\right]\nu(dv)\right| \leq \alpha\rho^n, \ \forall x \in E, \ \forall A \in \mathcal{E}. \tag{3.12}
$$

Logo por (2.19) e (3.12) temos que

$$
\begin{aligned}
& \left|\int_E W(h,x,v)\left[P^{k-j}(u,v) - P^{k-j}(y,v)\right]\nu(dv)\right| \\
= \; & \left|\int_E W(h,x,v)\left[P^{k-j}(u,v) - P^{k-j}(y,v)\right]\nu(dv) + \int_E W(h,x,v) f(v)\,\nu(dv) - \int_E W(h,x,v) f(v)\,\nu(dv)\right| \\
\leq \; & \left|\int_E W(h,x,v)\left[P^{k-j}(u,v) - f(v)\right]\nu(dv)\right| + \left|\int_E W(h,x,v)\left[P^{k-j}(y,v) - f(v)\right]\nu(dv)\right| \\
\leq \; & \left|\int_E \frac{K_1(x)}{\gamma_n(x)}\left[P^{k-j}(u,v) - f(v)\right]\nu(dv)\right| + \left|\int_E \frac{K_1(x)}{\gamma_n(x)}\left[P^{k-j}(y,v) - f(v)\right]\nu(dv)\right| \\
= \; & \frac{K_1(x)}{\gamma_n(x)}\left|\int_E \left[P^{k-j}(u,v) - f(v)\right]\nu(dv)\right| + \frac{K_1(x)}{\gamma_n(x)}\left|\int_E \left[P^{k-j}(y,v) - f(v)\right]\nu(dv)\right| \\
\leq \; & \frac{K_1(x)}{\gamma_n(x)}\alpha\rho^{k-j} + \frac{K_1(x)}{\gamma_n(x)}\alpha\rho^{k-j} = 2\frac{K_1(x)}{\gamma_n(x)}\alpha\rho^{k-j}
\end{aligned}
$$

Desse modo tem-se que

$$\begin{aligned}
&\frac{2\gamma_n(x)}{n}\sum_{1\leq j\leq k\leq n} cov_{p_0}\left[W(h,x,X_j),W(h,x,X_k)\right]\\
\leq\ &\frac{2\gamma_n(x)}{n}\sum_{1\leq j\leq k\leq n} 2\frac{K_1(x)}{\gamma_n(x)}\alpha\rho^{k-j}\int_E W(h,x,u)\,p_j(u)\,\nu(du)\\
=\ &\frac{4K_1(x)}{n}\sum_{1\leq j\leq k\leq n}\alpha\rho^{k-j}\int_E W(h,x,u)\,p_j(u)\,\nu(du)\\
=\ &\frac{4K_1(x)}{n}\sum_{j=1}^{n-1}\sum_{k=j+1}^{n}\alpha\rho^{k-j}E_{p_0}\left[W(h,x,X_j)\right]\\
\leq\ &4K_1(x)\,\alpha\frac{\rho}{1-\rho}\frac{1}{n}\sum_{j=1}^{n-1}E_{p_0}\left[W(h,x,X_j)\right]
\end{aligned}$$

que é limitada usando mais uma vez a Proposição 3.1.

Assim como $n\gamma_n(x)\,var_{p_0}(f_n(x))$ é limitado e, por (2.18), $n\gamma_n(x)\to\infty$ então fazendo como no Teorema 2.1 segue que

$$var_{p_0}(f_n(x))\to 0.$$

Portanto $f_n(x)$ é consistente em média quadrática.

Já a demonstração da consistência fraca é idêntica a demonstração feita no Capítulo 2, ou seja, é consequência imediata da consistência em média quadrática pela desigualdade de Tchebyschev. ∎

3.4 Consistência forte

Iniciamos esta seção verificando que uma cadeia de Markov geometricamente ergódica satisfaz a condição φ-mixing. Para Roussas e Ioannides (1987) provar que uma cadeia de Markov satisfaz a condição φ-mixing é equivalente a mostrar que

$$\left|P(\mathcal{B}\,|\,\mathcal{F}_0^n)-P(\mathcal{B})\right|\leq\varphi(n)\downarrow 0\ q.c.\ . \tag{3.13}$$

.

Desse modo, vamos verificar o seguinte resultado:

Lema 3.1 *Se $\{X_k\}_{k\geq 0}$ é um processo de Markov geometricamente ergódico como em (3.1) então ele é φ-mixing com coeficiente $\varphi(n) = \bar{\alpha}\rho^n$ onde $\bar{\alpha} = 2\alpha$ e ρ é o coeficiente ergódico.*

Demonstração: Sabemos pela propriedade de Markov, Definição 1.3, (3.1) e (3.3) que

$$\begin{aligned}
\left|P_{p_0}\left(X_{k+n} \in B \,|\, \mathcal{F}_0^k\right) - P_{p_0}\left(X_{k+n} \in B\right)\right| &= \left|P_{p_0}\left(X_{k+n} \in B \,|\, \sigma\left(X_k\right)\right) - P_{p_0}\left(X_{k+n} \in B\right)\right| \\
&= \left|P^n\left(X_k, B\right) - P_{p_0}\left(X_{k+n} \in B\right)\right| \\
&= \left|P^n\left(X_k, B\right) - P_{p_0}\left(X_{k+n} \in B\right) + \pi\left(B\right) - \pi\left(B\right)\right| \\
&= \left|\left[P^n\left(X_k, B\right) - \pi\left(B\right)\right] - \left[P_{p_0}\left(X_{k+n} \in B\right) - \pi\left(B\right)\right]\right| \\
&\leq \left|P^n\left(X_k, B\right) - \pi\left(B\right)\right| + \left|P_{p_0}\left(X_{k+n} \in B\right) - \pi\left(B\right)\right| \\
&\leq \alpha\rho^n + \alpha\rho^{k+n} \leq 2\alpha\rho^n = \bar{\alpha}\rho^n = \varphi\left(n\right),
\end{aligned}$$

que é exatamente (3.13). Logo a sequência é φ-mixing.

Portanto se a cadeia de Markov é geomericamente ergódica então satisfaz a condição φ-mixing com coeficiente mixing geométrico dado por $\varphi(n) = 2\alpha\rho^n$. ∎

A consistência forte é obtida sob as seguintes condições.

Teorema 3.2 (Consistência forte) *Seja $x \in C_\nu(f)$. Assuma que a Condição 2.2 é satisfeita. Se para todo $\beta > 0$*

$$\sum_{n\geq 1} exp\left\{-n\beta\gamma_n^2\left(x\right)\right\} < \infty,$$

então temos a consistência forte

$$\lim_{n\to\infty} f_n\left(x\right) = f\left(x\right),\ \nu - q.c.\ ,$$

e

$$P_{p_0}\left(\lim_{n\to\infty} f_n\left(x\right) = f\left(x\right)\right) = 1.$$

Demonstração: Pela Proposição 3.1 basta mostrar que

$$\lim_{n\to\infty}\left[f_n\left(x\right) - E_{p_0}\left(f_n\left(x\right)\right)\right] = 0,\ q.c., \tag{3.14}$$

mas a demonstração é idêntica à do Teorema 2.2, apesar de no Capítulo 2 trabalharmos com um processo estritamente estacionário e φ-mixing, a única hipótese utilizada foi a condição do tipo mixing.

Assim a demonstração de (3.14) segue do Lema (3.2). $\blacksquare$

Conclusão

Nosso trabalho foi baseado em Campos e Dorea (2005) e nosso objetivo aqui foi estimar as densidades associadas a uma cadeia de Markov com espaço de estados geral $E \subset \mathbb{R}^d$. Para estimar a densidade $f(\cdot)$ usamos os estimadores do tipo núcleo definidos por,

$$f_n(x) = \frac{1}{n}\sum_{k=1}^{n} W(h, x, X_k), \ h = h_n \downarrow 0,$$

com a função peso $W(h, x, \cdot)$ e a sequência h satisfazendo condições de regularidade.

Foi verificado que esses estimadores são assintoticamente não viciados, ou seja,

$$\lim_{n\to\infty} E(f_n(x)) = f(x), \ \forall x \in C_\nu(f),$$

consistentes em média quadrática, mais precisamente

$$\lim_{n\to\infty} E\left\{[f_n(x) - f(x)]^2\right\} = 0, \ \forall x \in C_\nu(f),$$

consequentemente fracamente consistente, isto é,

$$\lim_{n\to\infty} P(|f_n(x) - f(x)| > \epsilon) = 0, \ \forall x \in C_\nu(f).$$

e fortemente consistente, ou seja,

$$\lim_{n\to\infty} f_n(x) = f(x) \ \nu - q.c. \ (quase\ certamente), \ \forall x \in C_\nu(f).$$

Esses resutados foram verificados em dois casos: No primeiro caso, com o objetivo de estimar a densidade estacionária, foi considerada uma cadeia de Markov estritamente estacionária e satisfazendo a condição φ-mixing. E no segundo caso, com o objetivo de estimar a densidade limite, foi considerado que a cadeia era geometricamente ergódica.

Em Campos e Dorea (2005) foi mostrado ainda a normalidade assintótica desses estimadores, ou seja,

$$\sqrt{n\gamma_n(x)}\left[f_n(x) - E(f_n(x))\right] \xrightarrow{D} N\left(0, \sigma^2(x)\right).$$

Para mais detalhes ver Campos e Dorea (2005).

Desse modo concluimos que os estimadores do tipo núcleo possuem boas propriedades estatísticas, ou seja, são assintoticamente não viciados, consistentes e assintoticamente normal distribuídos.

Referências Bibliográficas

ATHREYA, K. B.; ATUNCAR, G. S. Kernel estimations for real-valued Markov chains. *Sankhya: The Indian Journal of Statistics*. vol. 60, p. 1-17, 1998.

ATHREYA, K. B.; LAHIRI, S. N. *Measure theory and probability theory*. 1. ed. Springer. 2006. 618 p.

CAMPOS, V. S. M. (2001). *Análise assintótica de estimadores tipo núcleo para densidades associadas a cadeias de Markov gerais.*

CAMPOS, V. S. M.; DOREA, C. C. Y. Kernel density estimation: the general case. *Statistics and Probability Letters*. vol. 55, p. 173-180, 2001.

CAMPOS, V. S. M.; DOREA, C. C. Y. Kernel estimation for stationary density of Markov chains with general state space. *Ann. Inst. Statist. Math.*. vol. 57, p. 443-453, 2005.

DEVROYE, Luc. Exponential inequalities in nonparametric estimation, *Nonparametric Functional Estimations and Related Topics*. (ed. G. G. Roussas). *Kluwer Academic Publishers*. 31-44. 1991.

DOOB, J. L. *Stochastic Processes*. John Wiley & Sons. New York. 1953.

DOREA, Chang C. Y.; ZHAO, L. C. Nonparametric density estimation in hidden Markov models. *Statistical Inference for Stochastic Processes*. vol. 5, p. 55-64, 2002.

DURRET, Rick *Probability: Theory and Examples*. 3.ed. Thomson. 2005.

MAGALHÃES, Marcos Nascimento *Probabilidade e Variáveis Aleatórias*. 2.ed. Edusp. São Paulo. 2006.

PARZEN, E. Estimation of a probability function and its mode. *Annals of Mathematical Statistics*. vol. 33, p. 1065-1076, 1962.

PRAKASA RAO, B.L.S. Nonparametric functional estimation. *Academic Press*. New York, 1983.

ROSENBLATT, M. Density estimates and Markov sequences. *Annals Math. Statistic*. vol. 27, p. 832-837, 1970.

ROUSSAS, G. G. Estimation of transition distribution function and its quantiles in Markov processes: strong consistency and Asymptotic Normality. *Nonparametric Functional Estimation and Related Topics*. (ed. G. G. Roussas). Kluwer Academic Publishers. Dordrecht. p. 443-462, 1991.

ROUSSAS, G. G.; IONNIDES, D. . Moment inequalities for mixing sequences of random variables. *Stochastic analysis and Applications*. vol. 5, p. 61-120, 1987.

ROUSSAS, G. G. Nonparametric estimation in Markov processes. *Annals of the Institute of Statistical Mathematics*. vol. 21, p. 73-87, 1969.

Livros Grátis

(http://www.livrosgratis.com.br)

Milhares de Livros para Download:

Baixar livros de Administração
Baixar livros de Agronomia
Baixar livros de Arquitetura
Baixar livros de Artes
Baixar livros de Astronomia
Baixar livros de Biologia Geral
Baixar livros de Ciência da Computação
Baixar livros de Ciência da Informação
Baixar livros de Ciência Política
Baixar livros de Ciências da Saúde
Baixar livros de Comunicação
Baixar livros do Conselho Nacional de Educação - CNE
Baixar livros de Defesa civil
Baixar livros de Direito
Baixar livros de Direitos humanos
Baixar livros de Economia
Baixar livros de Economia Doméstica
Baixar livros de Educação
Baixar livros de Educação - Trânsito
Baixar livros de Educação Física
Baixar livros de Engenharia Aeroespacial
Baixar livros de Farmácia
Baixar livros de Filosofia
Baixar livros de Física
Baixar livros de Geociências
Baixar livros de Geografia
Baixar livros de História
Baixar livros de Línguas

Baixar livros de Literatura
Baixar livros de Literatura de Cordel
Baixar livros de Literatura Infantil
Baixar livros de Matemática
Baixar livros de Medicina
Baixar livros de Medicina Veterinária
Baixar livros de Meio Ambiente
Baixar livros de Meteorologia
Baixar Monografias e TCC
Baixar livros Multidisciplinar
Baixar livros de Música
Baixar livros de Psicologia
Baixar livros de Química
Baixar livros de Saúde Coletiva
Baixar livros de Serviço Social
Baixar livros de Sociologia
Baixar livros de Teologia
Baixar livros de Trabalho
Baixar livros de Turismo