

## 8.5 Estimadores via Ondaletas

O objetivo aqui é estimar o espectro (3.14), utilizando técnicas de estimação não lineares, aplicando-se limiares aos coeficientes de ondaletas empíricos.

Se tivermos um processo estacionário discreto  $\{X_t, t \in \mathbb{Z}\}$ , o argumento acima pode ser repetido, com somas substituindo integrais. Se

$$\sum_{k=-\infty}^{\infty} |\gamma_k| < \infty \quad (3.13)$$

estiver satisfeita, obtemos o espectro de  $X_t$  como

$$f(\lambda) = \frac{1}{2\pi} \sum_{k=-\infty}^{\infty} \gamma_k e^{-i\lambda k}, \quad -\pi < \lambda < \pi. \quad (3.14)$$

Na estimação do espectro usando métodos de Fourier, procuramos por um compromisso entre resolução (viés) e estabilidade (variância). Os estimadores suavizados

utilizados, baseados em janelas espectrais, são apropriados para estimar funções com um grau homogêneo de regularidade. A presença de picos ou transitoriedades ocasiona os problemas conhecidos na análise espectral clássica.

Do ponto de vista teórico, bases de senos/cossenos são ótimas para espaços  $L^2$ -Sobolev, enquanto que bases de ondaletas, também ótimas para esses espaços, continuam ótimas para um conjunto amplo de espaços de Besov, que contêm funções que não são regulares de modo homogêneo, ou seja, podem ser regulares numa parte do domínio, mas menos regulares em outras partes.

Dois requisitos principais são necessários para estimadores baseados em expansões ortogonais, para que tenham uma adaptabilidade espacial alta (Neumann, 1996):

- (a) a base tem que ser suficientemente localizada;
- (b) o estimador tem que discriminar o ruído da variância.

Suponha as observações  $X_0, X_1, \dots, X_{2^J-1}$ , com  $T = 2^J$ ,  $J > 0$  inteiro. Wahba propôs o seguinte modelo:

$$\log I(\lambda_j) = \log f(\lambda_j) + \gamma + \epsilon_j, \quad j = 1, \dots, T-1, \quad (8.60)$$

com  $\epsilon_j = \log(\eta_j/2) - E\{\log(\eta_j/2)\}$ ,  $\eta_j \sim \chi_2^2$ . Prova-se que  $Var(\epsilon_j) = \pi^2/6$ .

O procedimento não linear proposto para estimar  $g(\lambda)$  é o seguinte:

[1 ] Calcule o log-periodograma,

$$g_\ell = \log I(\lambda_\ell), \quad \ell = 0, 1, \dots, T-1,$$

para frequências  $\lambda_\ell = \frac{2\pi\ell}{2^j}$ .

[2 ] Considere a transformada de ondaletas de  $g_\ell$ , obtendo-se os coeficientes empíricos de ondaletas  $\{y_{j,k}\}, j = 0, 1, \dots, J-1, k = 0, 1, \dots, 2^j-1$ .

[3 ] Aplique-se o limiar suave

$$\delta_\lambda^S(x) = \text{sign}(x)(|x| - \lambda)_+, \quad (8.61)$$

aos coeficientes  $\{y_{j,k}\}$ , com limiares dependentes do nível  $j$ ,  $\lambda = \lambda_{j,T}$ . Estes são calculados por

$$\lambda_{j,T} = \alpha_j \log T, \quad (8.62)$$

para níveis de resolução finos ( $j = J-1, J-2, \dots$ ), onde  $\alpha_j$  é uma constante, e

$$\lambda_{j,T} = \pi \sqrt{\frac{\log T}{3}}, \quad (8.63)$$

para níveis de resolução grosseiros ( $j \ll J-1$ ). Valores de  $\alpha_j$  para bases comumente usadas (bases de Daubechies, por exemplo), e alguns valores de  $J$  são dados na Tabela 8.1.

[4 ] Inverta a transformada de ondaletas, produzindo-se um estimador  $\hat{g}_\ell^*$  do log-espectro na frequência  $\omega_\ell$ . Se necessário, aplique exponenciação para obter um estimador de  $f(\lambda)$ .

**Exemplo 8.6.** (Gao, 1993). A Figura 8.8 mostra o resultado de aplicar os limiares (8.62) e (8.63), para o caso de um processo AR(24). Na Figura (a) temos o verdadeiro log-espectro do processo, em (b) o log-periodograma, em (c) o estimador usando o método proposto e, em (d), o estimador obtido usando o limiar baseado na suposição de normalidade. Note os picos espúrios no último estimador.

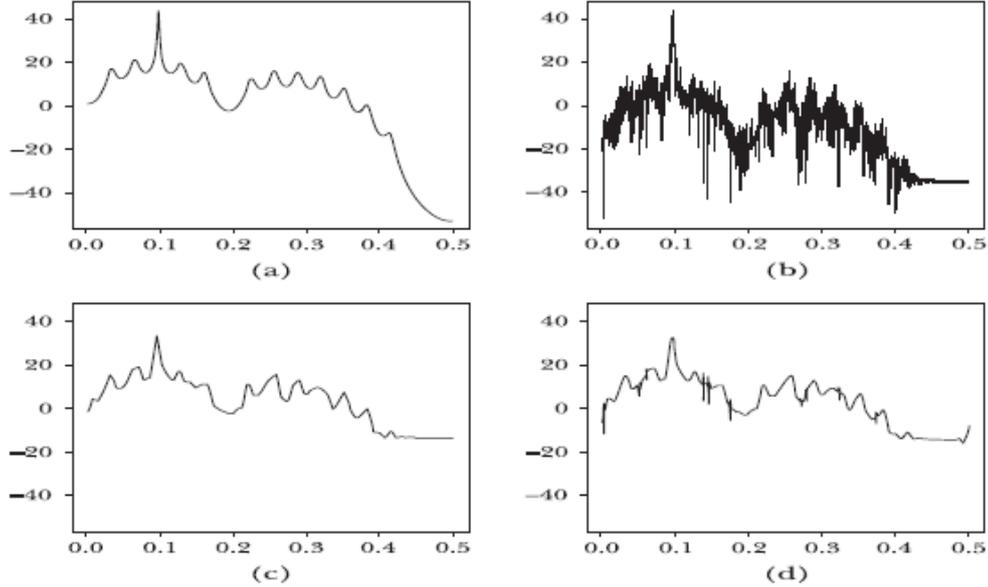


Figure 8.8: (a) Log-espectro de um processo AR(24); (b) Log-periodograma; (c) Estimador baseado na proposta de Gao; (d) Estimador baseado na normalidade.

Vejamos, agora, o caso de um processo estacionário não gaussiano, considerado por Neumann (1996). Considere, inicialmente, o periodograma modificado (*tapered*)

$$I_T(\lambda) = \frac{1}{2\pi H_2^{(T)}} \left| \sum_{t=0}^{T-1} h_t X_t e^{-it\lambda} \right|^2, \quad (8.65)$$

onde  $h_t$  é uma janela de dados (*taper*), com  $h_t = h(t/T)$  e  $H_k^{(T)} = \sum_{t=0}^{T-1} h_t^k$ .

Como o periodograma habitual ( $h_t \equiv 1$ ),  $I_T(\lambda)$  é assintoticamente não viesado, não consistente, tem uma distribuição assintótica qui-quadrado e, para  $\lambda_1 \neq \lambda_2$ ,  $I_T(\lambda_1)$  e  $I_T(\lambda_2)$  são assintoticamente não correlacionadas.

A partir da ondaleta pai,  $\phi$ , e ondaleta mãe,  $\psi$ , consideramos a base ortonormal em  $L^2(\mathbb{R})$ ,

$$\{\tilde{\phi}_{\ell,k}\}_{k \in \mathbb{Z}} \cup \{\tilde{\psi}_{j,k}\}_{j \geq \ell, k \in \mathbb{Z}}, \quad (8.66)$$

geradas da maneira usual. Como queremos uma base ortonormal em  $L^2(\Pi)$ ,  $\Pi = [-\pi, \pi]$ , a partir de (8.66) geramos

$$\{\phi_{\ell,k}\}_{k \in I_\ell} \cup \{\psi_{j,k}\}_{j \geq \ell, k \in I_j}, \quad (8.67)$$

com  $I_j = \{1, 2, \dots, 2^j\}$ ,

$$\phi_{\ell,k}(t) = \sum_{n \in \mathbb{Z}} (2\pi)^{-1/2} \tilde{\phi}_{\ell,k}((2\pi)^{-1}t + n)$$

e

$$\psi_{j,k}(t) = \sum_{n \in \mathbb{Z}} (2\pi)^{-1/2} \tilde{\psi}_{j,k}((2\pi)^{-1}t + n).$$

Aqui,  $L^2(\Pi)$  é o espaço das funções de período  $2\pi$  sobre  $\Pi$ , de quadrado integrável. Ou seja, a partir de ondaletas sobre a reta, obtemos ondaletas “periodizadas” sobre  $[-\pi, \pi]$ .

Consideremos a expansão, em ondaletas periódicas, do espectro  $f(\lambda)$ ,

$$f = \sum_{k \in I_\ell} \alpha_\ell \phi_{\ell,k} + \sum_{j \geq \ell} \sum_{k \in I_j} \alpha_{j,k} \psi_{j,k}, \quad (8.68)$$

sendo

$$\alpha_k = \int f(t) \phi_{\ell,k}(t) dt, \quad (8.69)$$

e

$$\alpha_{j,k} = \int f(t) \psi_{j,k}(t) dt \quad (8.70)$$

são os coeficientes de ondaletas.

Definamos o estimador de  $\alpha_{j,k}$

$$\tilde{\alpha}_{j,k} = \int \psi_{j,k}(\lambda) I_T(\lambda) d\lambda, \quad (8.71)$$

e uma definição análoga para  $\tilde{\alpha}_k$ .

O objetivo é considerar um estimador da função densidade espectral baseados em (8.68). Para um conjunto restrito de índices

$$\mathcal{I} = \{(j, k) : 2^j \leq CT^{1-\alpha}, k \in I_j\},$$

na qual  $C < \infty$  e  $0 < \alpha \leq 1/3$  são constantes fixas, a normalidade assintótica dos coeficientes  $\tilde{\alpha}_{j,k}$  pode ser provada. Seja  $\sigma_{j,k}^2$  a variância de  $\tilde{\alpha}_{j,k}$ .

Para estimar o espectro  $f(\lambda)$ , a idéia é usar (8.68) com algum tipo de limiar para os coeficientes de ondaletas. Podemos usar, por exemplo, os limiares duros e suaves

vistos anteriormente e escolher  $\lambda = \lambda_j(\sigma_{j,k}, \mathcal{F})$ , que depende da escala, do desvio padrão de  $\tilde{\alpha}_{j,k}$  e da classe  $\mathcal{F}$  de funções à qual  $f(\lambda)$  pertence.

Neumann (1996) propõe dois tipos de limiares que satisfazem certas condições desejáveis, a saber

$$\lambda_{j,k} = (\sigma_{j,k} \vee \sigma_T) \sqrt{2 \log(\#\mathcal{I})} \quad (8.72)$$

e

$$\lambda_{j,k} = (\sigma_{j,k} \vee \sigma_T) \sqrt{2 \log((\#\mathcal{I})/2^\ell)}, \quad (8.73)$$

sendo

$$\sigma_T = \max\left\{ \max_{(j,k) \in \mathcal{I}} \{\sigma_{j,k}\}, C_0 T^{-1/2} \right\},$$

para algum  $C_0 > 0$  fixado.

Se  $\mathcal{F}$  for um espaço com norma essencialmente equivalente a um espaço de Besov, então o estimador

$$\hat{f} = \sum_{k \in I_\ell} \tilde{\alpha}_{\ell,k} \phi_{\ell,k} + \sum_{(j,k) \in \mathcal{I}} \delta^{(\cdot)}(\tilde{\alpha}_{j,k}, \lambda_{j,k}) \psi_{j,k} \quad (8.74)$$

conduz à taxa de convergência adequada para o risco do estimador, especificamente

$$\sup_{f \in \mathcal{F}} \{E \|\hat{f} - f\|_{L^2(\Pi)}^2\} = O((\log T/T)^{2m/(2m+1)}). \quad (8.75)$$

Sabe-se que  $T^{-2m/(2m+1)}$  é a taxa de convergência ótima em classes com grau de regularidade  $m$ , se  $X_t$  for estacionário gaussiano. Logo, o estimador acima é aproximadamente minimax, a menos de um fator de ordem  $\log T$ . Veja Neumann (1996), para um estimador que atinge essa taxa ótima e sugestões para melhorar estes estimadores.