

Suavização por Meio de Ondasletas

O princípio do encolhimento por meio de ondasletas (*wavelet shrinkage*) tem por objetivo a **redução (e mesmo remoção) do ruído** presente num sinal, diminuindo (ou zerando) a magnitude dos coeficientes de ondasletas.

Neste capítulo introduzimos a técnica de encolhimento chamada **limiarização (*thresholding*)** por meio de ondasletas.

Os trabalhos principais nesta área são os artigos de Donoho e Johnstone (1994, 1995, 1998), Donoho *et al.* (1995, 1996a, 1996b), Johnstone e Silverman (1997), Vidakovic (1998, 1999) e Antoniadis *et al.* (2002). Os principais trabalhos concentram-se na área de regressão não paramétrica.

Filtro L suaviza uma série de entrada, enquanto H produz os detalhes.

Quando os coeficientes que representam detalhes são pequenos em grandeza, eles podem ser omitidos, sem afetar substancialmente a reconstrução da entrada. Assim, a ideia de limiarizar os coeficientes de ondasletas corresponde a eliminar (ou encolher) detalhes que correspondem a ruído.

O procedimento de limiarização corresponde a uma operação de suavização. No caso da análise de Fourier, limiarizar coeficientes de Fourier afeta o resultado globalmente, dado o caráter não local de senos e cossenos.

Uma transformada de ondasleta dos dados **preserva a “energia”**: a soma dos quadrados dos dados é igual à soma dos quadrados dos coeficientes de ondasletas, mas essa energia ficará concentrada em poucos coeficientes de ondasletas. Isso significa que uma função de interesse será descrita por um número pequeno de coeficientes de ondasletas. Por outro lado, ruído i.i.d. gaussiano é invariante por uma transformada ortogonal, e passa para o domínio de ondasletas não afetado.

Uma outra propriedade da transformada de ondaletas (como a de Fourier) é que os dados transformados tornam-se aproximadamente não correlacionados.

Escolha do Limiar

dois problemas:

- a escolha do esquema (ou política) de limiares;
- a escolha dos parâmetros que governam esse esquema.

5.2.1 Escolha do Esquema

Dois esquemas tradicionalmente utilizados são os seguintes:

(i) Limiar Duro (*Hard Threshold*), definido por

$$\delta_{\lambda}^H(x) = \begin{cases} 0, & \text{se } |x| \leq \lambda \\ x, & \text{se } |x| > \lambda. \end{cases} \quad (5.1)$$

(ii) Limiar Suave (*Soft Threshold*), definido por

$$\delta_{\lambda}^S(x) = \begin{cases} 0, & \text{se } |x| \leq \lambda \\ \text{sign}(x)(|x| - \lambda), & \text{se } |x| > \lambda. \end{cases} \quad (5.2)$$

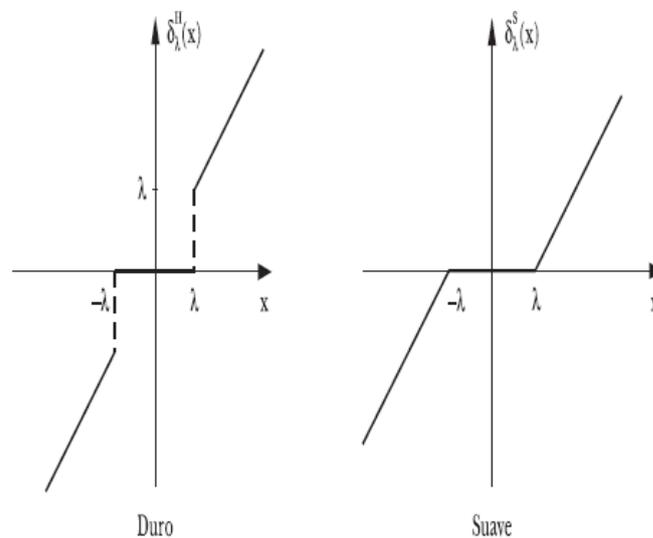


Figura 5.1: Limiar duro e suave.

A Figura 5.1 mostra como esses dois esquemas agem. Note que (i) é do tipo “mata” ou “preserva”, ao passo que (ii) é do tipo “mata” ou “encolhe” (reduz o tamanho da quantidade λ).

5.2.2 Escolha dos Parâmetros

Para a escolha do parâmetro λ , que aparece em (5.1) e (5.2), podemos usar um *limiar global*, aplicável a todos os níveis, ou então limiares que dependam do nível j , ou seja, para cada escala teremos um limiar λ_j .

Apresentamos, a seguir, as principais propostas, onde $d_{j,k}$ representa um coeficiente de ondaleta genérico.

(P1) *Universal*

Donoho e Johnstone (1994) propuseram um estimador minimax para f no modelo (5.11) abaixo, denominado RiskShrink, baseado em limiar duro ou suave e num parâmetro λ , obtido pela minimização de um limite superior teórico de risco assintótico. Contudo, sugeriram depois um método alternativo, baseado no seguinte resultado para v.a.'s normais.

$$\lambda_n = \lambda_{j,n} = \sigma\sqrt{2\log n}, \quad (5.4)$$

que *não depende da escala* e é chamado *universal*. O nível de ruído σ tem que ser estimado a partir dos dados. Veja abaixo e Donoho *et al.* (1995).

Esse limiar conduz a estimadores que subestimam f , pois a tendência é eliminar ou reduzir muitos coeficientes, especialmente em níveis de resolução altos.

(P2) Donoho *et al.*(1995) usam uma variante do limiar universal, a saber

$$\lambda_n = \sigma\sqrt{\frac{2\log(n)}{n}}. \quad (5.5)$$

Esse limiar tem propriedades similares ao universal.

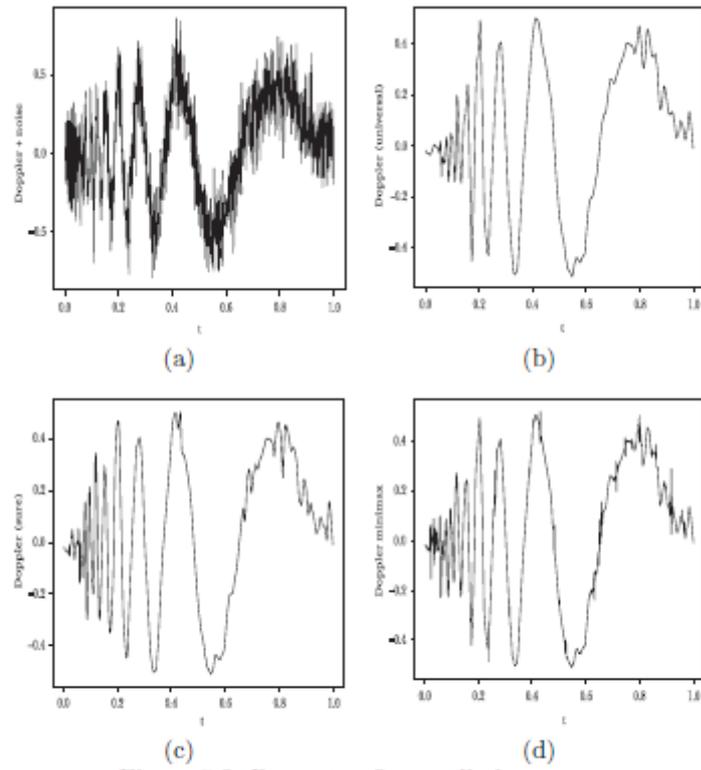
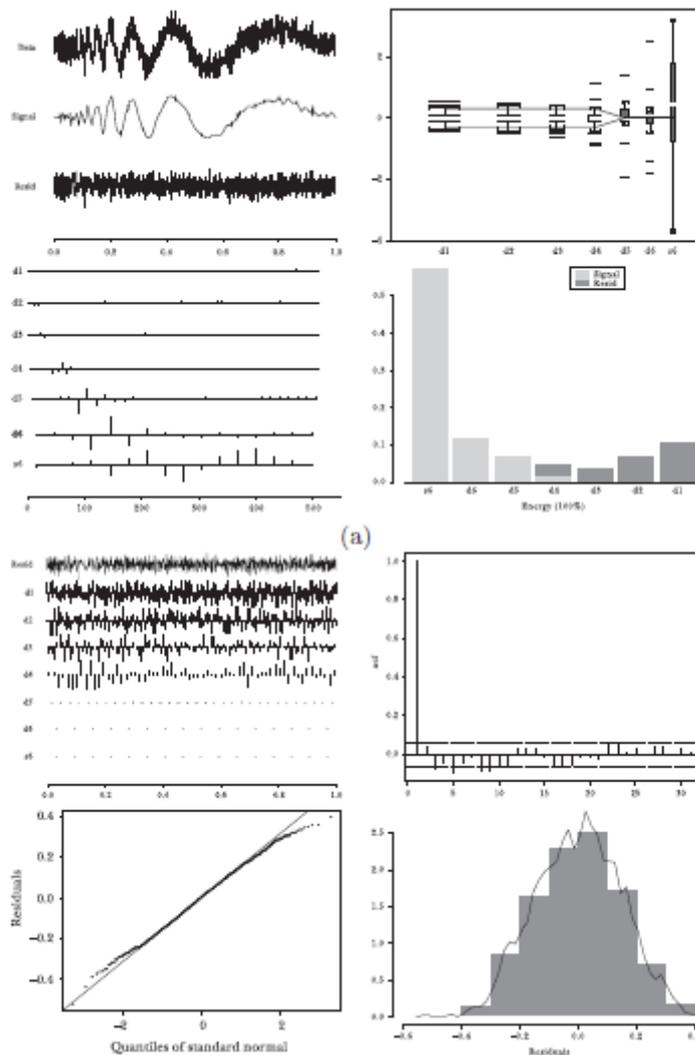


Figura 5.2: Reconstrução com limiar suave.
 (a) *Doppler* mais ruído gaussiano, $\sigma = 0,12$; (b) *Universal*;
 (c) *Sureshrink*; (d) *Minimax*.



Na Figura 5.3(a), temos os dados (*Data*), a reconstrução (*Signal*) e os resíduos (*Resid*) no canto superior esquerdo. No canto superior direito temos um *box-plot* dos coeficientes da transformada de ondaletas dos dados originais, com os limiares minimax superpostos como retas horizontais. Supõe-se que qualquer coeficiente entre as retas seja igual a zero. No canto inferior esquerdo temos os coeficientes da transformada discreta, e no canto inferior direito um gráfico em barras, decompondo, para cada cristal, a energia devida ao sinal e ao ruído.

Na Figura 5.3(b). No canto superior esquerdo temos os coeficientes de ondaletas dos resíduos e no canto superior direito um gráfico da função de autocorrelação dos resíduos, mostrando que algumas correlações são não nulas. No canto inferior esquerdo temos um gráfico quantil-quantil, com os quantis da normal padrão no eixo horizontal. Finalmente, no canto inferior direito temos um histograma e um estimador da densidade dos resíduos. As duas últimas figuras mostram a assimetria dos resíduos.

Regressão Não Paramétrica

5.3.1 Modelo com delineamento regular fixo, erros i.i.d. normais

Considere o modelo

$$y_i = f(t_i) + \varepsilon_i, \quad i = 1, 2, \dots, n, \quad (5.11)$$

onde $\varepsilon_i \sim i.i.d. N(0, \sigma^2)$, $\sigma^2 < \infty$ e $t_i = i/n$.

No caso usual de regressão paramétrica, supõe-se que f obedeça a um modelo particular, por exemplo, uma função linear de p parâmetros $\theta_1, \dots, \theta_p$. A estimação dos parâmetros é feita usando-se mínimos quadrados ordinários. Na situação do modelo (5.11), iremos supor que f pertença a uma classe de funções satisfazendo certas condições de suavidade.

O enfoque tradicional usa estimadores baseados em **núcleos (*kernels*)**, **splines** e **séries de Fourier**. Uma revisão excelente sobre métodos não paramétricos em séries temporais é dada por Härdle *et al.* (1997). Veja também Ogden (1997).

Contudo, esses métodos **não são adaptativos**, no sentido que os estimadores resultantes podem alcançar taxas de convergência **mais baixas** se a regularidade (suavidade) da função f for mal especificada. Para obter estimadores que **são adaptativos** pode-se usar, por exemplo, o método **kernel com largura de faixa (*bandwidth*) variável**, ou então, usar **ondaletas**, que conduzem a procedimentos que são adaptativos e quase ótimos sobre classes de funções com regularidades não homogêneas e isso é conseguido por meio do procedimento de limiarização.

Se $f_i = f(t_i)$, o objetivo é estimar $\mathbf{f} = (f_1, \dots, f_n)'$, com o menor erro quadrático médio, ou seja, queremos minimizar $R_2(\hat{f}, f) = E\|\hat{f} - f\|_2^2$ para alguma classe de funções dada, \mathcal{F} , com $\|f\|_2^2 = \sum_{i=1}^n f_i^2$.

Aqui, f é estimada por

$$\hat{f}(x) = \sum_{k=0}^{2^{j_0}-1} \hat{c}_{j_0,k} \phi_{j_0,k}(x) + \sum_{j=j_0}^J \sum_{k=0}^{2^j-1} \hat{d}_{j,k} \psi_{j,k}(x), \quad (5.12)$$

na qual $\hat{c}_{j_0,k}$ e $\hat{d}_{j,k}$ são os coeficientes de ondaleta empíricos. A seguir aplica-se um limiar (duro ou suave) aos coeficientes $\hat{d}_{j,k}$.

O procedimento do encolhimento consiste nos três estágios seguintes:

- (1) Tome a transformada de ondaletas discreta de y_1, \dots, y_n , obtendo-se os n coeficientes de ondaletas $d_{j,k}$, que são contaminados por ruído.
- (2) Use limiares (*thresholds*) para reduzir ou anular aqueles coeficientes abaixo de certo valor. Obtemos, nesse estágio, os coeficientes desprovidos de ruído.
- (3) Tome a transformada de ondaletas inversa dos coeficientes do estágio (2) para obter as estimativas \hat{f}_i .

A transformada (4.46) aplicada ao modelo (5.11) produz

$$\mathbf{W}\mathbf{y} = \mathbf{W}\mathbf{f} + \mathbf{W}\boldsymbol{\varepsilon} \quad (5.13)$$

e como \mathbf{W} é ortogonal, ela transforma ruído branco em ruído branco, isto é, se $w_{j,k}$ são os coeficientes de ondaletas de $f(t_i)$, podemos escrever

$$d_{j,k} = w_{j,k} + \sigma z_{j,k} \quad (5.14)$$

onde $z_{j,k} \sim i.i.d. N(0, 1)$, $j = 0, 1, \dots, J - 1$, $k = 0, 1, \dots, 2^j - 1$, $d_{-1,0} = c_{00}$.

Ou seja, (5.14) nos diz que os coeficientes de ondaletas de uma amostra com ruído podem ser escritos como os coeficientes de ondaletas sem ruído adicionados a ruído branco.

5.4.2 Estimação de Densidades

O interesse aqui é estimar uma função densidade de probabilidade $f(x)$ com base nas observações X_1, \dots, X_n , que constituem uma amostra de f . Para uma revisão dos métodos usuais, veja Silverman (1986).

Considere a expansão

$$f(x) = \sum_k \alpha_{\ell,k} \phi_{\ell,k}(x) + \sum_{j \geq \ell} \sum_k \beta_{j,k} \psi_{j,k}(x), \quad (5.53)$$

onde os coeficientes são estimados por

$$\hat{\alpha}_{\ell,k} = n^{-1} \sum_{i=1}^n \phi_{\ell,k}(X_i), \quad (5.54)$$

$$\hat{\beta}_{j,k} = n^{-1} \sum_{i=1}^n \psi_{j,k}(X_i). \quad (5.55)$$

Para estimar f , usamos limiares para obter

$$\tilde{\beta}_{j,k} = \delta^{(\cdot)}(\hat{\beta}_{j,k}, \lambda_j), \quad (5.56)$$

onde $\delta^{(\cdot)}$ representa uma política de limiar dura ou suave e λ_j um parâmetro que varia de nível para nível. Em geral, não aplicamos limiares aos coeficientes $\hat{\alpha}_{\ell,k}$.

Assim como em regressão não paramétrica, estamos interessados em medidas de erros globais do desempenho dos estimadores sobre classes de regularidade suficientemente amplas. O estimador de f , obtido substituindo-se os $\tilde{\beta}_{j,k}$ em (5.53), atinge um desempenho quase ótimo, em termos de taxas de convergência.

Donoho *et al.* (1996b) utilizam limiares dados por

$$\tilde{\beta}_{j,k} = \begin{cases} \hat{\beta}_{j,k}, & \text{se } |\hat{\beta}_{j,k}| > KC(j)\sqrt{n} \\ 0, & \text{caso contrário,} \end{cases} \quad (5.57)$$

com $C(j) = \sqrt{j}$ e K constante.

Delyon e Juditsky (1993) usam os limiares de (5.57) com $C(j) = \sqrt{j - \ell}$. Johnstone *et al.* (1992) usam $\lambda_j = A\sqrt{j}$, com A constante. Veja também Hall e Patil (1995).

