

Uma implementação computacional do método de séries temporais funcionais

Vitória Martini Wendt (BIC-UFRGS), Eduardo Oliveira Horta (Orientador-UFRGS)
vitoriawendt@gmail.com, eduardo.horta@ufrgs.br

Introdução

O método de séries temporais funcionais tem recebido grande atenção nos últimos anos devido, entre outros fatores, ao aumento da capacidade computacional intrínseca a diversas áreas da ciência. Apesar de sua importância, o método só foi implementado de forma abrangente na plataforma R recentemente através do pacote *ftsa*, ainda sem contemplar uma importante contribuição feita por Bathia et al (2010).

Objetivos

Desenvolver e otimizar rotinas referentes a:

1. Estimção, predição e seleção de modelos: implementação da metodologia desenvolvida por Bathia et al (2010).
2. Performance: tempo de execução e armazenagem.
3. Visualização de dados funcionais.

Metodologia

Um conjunto de observações de dados funcionais $x_1, x_2, x_3, \dots, x_n$ é tipicamente observado com a presença de ruído ϵ_t . Deste modo, os dados aos quais teremos acesso são da forma:

$$y_t(u) = x_t(u) + \epsilon_t(u).$$

Em muitos casos, ϵ_t será diferente de zero e suas realizações serão dependentes em t . Assim, Bathia et al (2010) propôs uma representação alternativa para x_t :

$$x_t = \bar{y} + \sum_{j=1}^d W_{t,j} \psi_j(u),$$

em que d pode ser estimado através de modelos dinâmicos e $\psi_j(u)$ são auto-funções dos dados.

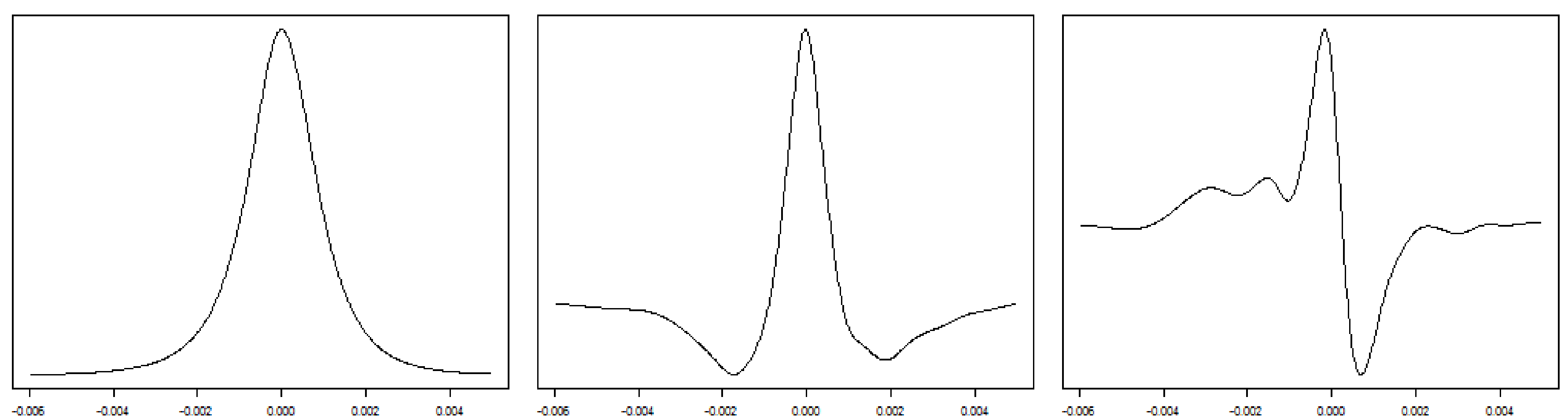


Imagem 1: Gráficos de \bar{y} (esquerda) e das auto-funções ψ_1 (centro) e ψ_2 (direita) retornados pela função *Ftsa2* na plataforma RStudio considerando dados funcionais de fechamento da Ibovespa em 2009;

Implementação computacional em ambiente R

Para unificar o uso computacional da metodologia desenvolvida por Bathia et al (2010), foi escolhida a linguagem de programação estatística R. Deste modo, desenvolveu-se um novo pacote, chamado *Ftsa2*, para a análise de séries temporais funcionais na plataforma *RStudio*.

Resultados

A implementação do modelo apresentado anteriormente em linguagem R atentou para duas principais questões na análise de dados funcionais:

- Performance: A questão de performance é um dos principais fatores que impedem a disseminação de metodologias voltadas a dados funcionais. Constatou-se que o desempenho computacional do pacote *Ftsa* era fraco em aspectos de tempo de execução de suas funções.

Tabela 1: Comparação de desempenho entre os pacotes *Ftsa* e *Ftsa2* utilizando bancos de dados de diferentes tamanhos.

n	Ftsa		Ftsa2	
	Tempo (seg.)	Memória (kbs)	Tempo (seg.)	Memória (kbs)
100	0.26	-	0.00	-
1000	2.16	-	0.39	-
10000	19.54	75.89	4.55	74.98

- Visualização: outra questão importante que ainda não é tratada de forma padronizada é a visualização de dados funcionais.

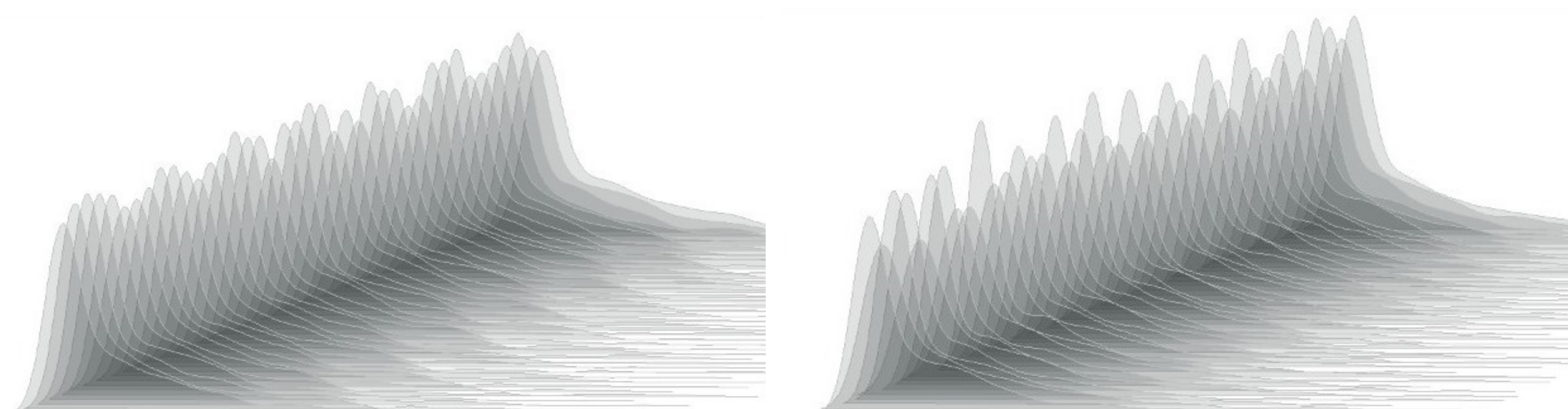


Imagem 2: Gráfico Waterfall para a visualização do $\log(\hat{y})[1 : m,]$ (esquerda) e $\log(\hat{y})[m : 1,]$ (direita).

Conclusão

O modelo desenvolvido por Bathia et al (2010) trouxe para o campo estatístico de análise de séries temporais funcionais uma importante contribuição, sendo o único modelo capaz de tratar séries temporais funcionais com a presença de ruído diferente de zero.

Deste modo, o pacote *Ftsa2* é proposto para a padronização do uso computacional da metodologia, atentando para a resolução de problemas de implementação comuns em métodos anteriores de análise de dados funcionais.

A partir dos resultados expostos ao lado, é possível dimensionar a diferença de desempenho computacional entre o pacote desenvolvido neste trabalho (*Ftsa2*) e o principal pacote de análise de dados funcionais presente na plataforma RStudio (*Ftsa*).