



Séries Temporais Funcionais: uma Aplicação à Estatística Criminal

Autor: Guilherme Rodrigues Boff (BIC – UFRGS)
Orientador: Prof. Dr. Eduardo de Oliveira Horta (UFRGS)

Introdução

A Estatística Criminal tem grande importância no contexto de auxiliar o trabalho de inteligência das forças policiais. Uma particular contribuição está na construção de *heatmaps* de ocorrências de diversos tipos de crimes, a fim de aumentar a capacidade de organização das instituições de segurança pública com uma distribuição mais eficiente de efetivos em locais com maior concentração de crimes. Nesse sentido, a teoria de dados funcionais tem ampla utilidade na estimação, modelagem e previsão desses mapas de intensidade.

Dados Funcionais

Dados funcionais são dados em que as unidades observacionais podem ser curvas, superfícies, etc. Uma característica prevalente nesse tipo de dado é a presença de erros observacionais. Assim, no contexto de séries temporais funcionais, sendo (λ_t) a série de interesse (um *heatmap*), só se tem acesso a ela acompanhada de um ruído ε_t , conforme a equação abaixo:

$$\hat{\lambda}_t(\mathbf{u}) = \lambda_t(\mathbf{u}) + \varepsilon_t(\mathbf{u}), \mathbf{u} \in S$$

onde S é o quadrado unitário $[0,1] \times [0,1]$.

Metodologia

Baseada na metodologia desenvolvida por Bathia *et al.* (2010), e nas contribuições apresentadas em Horta & Ziegelmann (2016), a proposta desta pesquisa consiste em usar a representação espectral

$$\lambda_t(\mathbf{u}) = \mu(\mathbf{u}) + \sum_{j=1}^{\infty} \eta_{tj} \psi_j(\mathbf{u}), \mathbf{u} \in S^1$$

em que μ é a média da série λ_t , onde o objetivo é estimar as variáveis aleatórias latentes η_{tj} , obtendo, então, uma redução de dimensão. Em seguida, aplica-se uma modelagem de séries temporais multivariadas com base em η_{tj} para, enfim, fazer previsões para a série temporal funcional λ_t .

¹ Ver Horta & Ziegelmann (2016).

Implementação Parcial

Na área da Estatística Criminal, λ_t pode ser interpretada como o mapa de intensidade de ocorrências de certo tipo de crime no mês t na região S . Através da interface *RStudio*, pode-se construir esses mapas. Para tanto, é necessário converter o endereço de cada ocorrência em coordenadas geográficas. A função *geocode* do pacote *ggmap* faz os *requests* diretamente à plataforma *Maps* do Google, uma vez que se tenha uma chave de acesso (*API*) registrada na plataforma. A Figura 1 abaixo mostra um exemplo de *heatmap*:

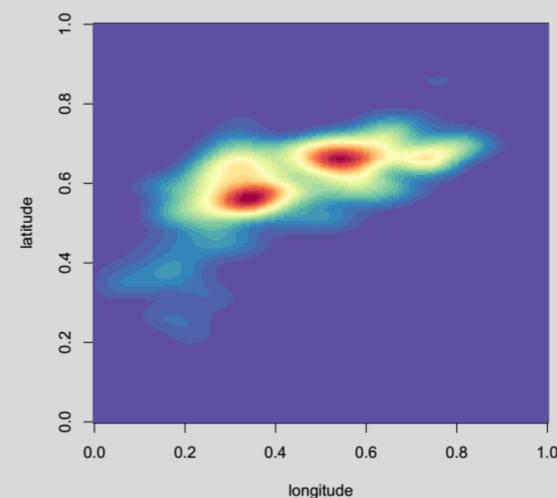


Figura 1: *heatmap* de ocorrências de roubo de veículos em um mês em uma cidade brasileira.

Há uma limitação na quantidade de *requests* que cada usuário pode realizar por dia. Assim sendo, esse processo ainda está em andamento para as observações do banco de dados que está sendo trabalhado. Desse modo, futuramente serão obtidas as previsões dos *heatmaps* com base nesses dados.

Referências

BATHIA, Neil; YAO, Qiwei; ZIEGELMANN, Flávio A. Identifying the Finite Dimensionality of Curve Time Series. *Annals of Statistics*. v. 38, p. 3352-3386, 2010.

HORTA, Eduardo O.; ZIEGELMANN, Flávio A. Identifying the spectral representation of Hilbertian time series. *Statistics & Probability Letters*. v. 118, p. 45-49, 2016.