

XVIII Jornadas de Classificação e Análise de Dados
Universidade de Trás-os-Montes e Alto Douro, 7 a 9 de Abril de 2011

Vítimas mortais em acidentes de viação em Portugal Continental: uma abordagem através da Análise Espectral Singular

Fernando Sebastião¹, Irene Oliveira²

¹*Escola Superior de Tecnologia e Gestão, Instituto Politécnico de Leiria, CM-UTAD, fsebast@ipleiria.pt;*

²*Universidade de Trás-os-Montes e Alto Douro, CM-UTAD, ioliveir@utad.pt*

Sumário

Uma das preocupações da política nacional, nas últimas décadas, tem sido a segurança rodoviária, e em particular a intenção de reduzir os níveis de sinistralidade. Um dos agentes envolventes é a mortalidade resultante dos acidentes de viação.

Este trabalho tem como objectivo explorar os dados mensais do número de vítimas mortais em acidentes de viação ocorridos em Portugal Continental, durante os últimos doze anos, através da Análise Espectral Singular.

Palavras-chave Acidentes de viação, análise espectral singular, componentes principais, mortalidade rodoviária, séries temporais

1. Introdução

A Análise Espectral Singular (SSA), vista como uma extensão da Análise em Componentes Principais, geralmente exhibe resultados muito satisfatórios na análise de uma série temporal, nomeadamente quando os dados apresentam uma estrutura mais complexa, como é o caso de séries com várias componentes sazonais. O seu principal objectivo é decompor uma série de dados original num pequeno número de componentes para que se possam interpretar a tendência e as componentes oscilatórias, desprezando assim a estrutura de ruído.

As políticas de segurança rodoviária visam sobretudo contribuir para que a sinistralidade rodoviária em Portugal tenha cada vez menos impacto na sociedade. Analisar o comportamento do número de vítimas mortais que resultam dos acidentes de viação permite uma melhor percepção da influência das políticas adoptadas nos últimos anos, e pode contribuir para um planeamento de novas políticas de sensibilização, de prevenção e de aplicação de legislação.

Neste trabalho procedemos a uma breve descrição da técnica da SSA e apresentamos o estudo da série mensal do número de vítimas mortais em acidentes de viação que ocorreram em Portugal Continental entre os anos de 1998 e 2009. Em particular, são analisadas as componentes essenciais para a reconstrução dos dados originais e para aplicação posterior na previsão da mortalidade rodoviária mensal.

2. Análise espectral singular

Segundo Elsner e Tsonis (1996) e Golyandina *et al.* (2001), o método básico da SSA é constituído por duas etapas complementares: a decomposição e a reconstrução. A primeira etapa contempla dois passos fundamentais: a obtenção da matriz dos desfasamentos ao longo do tempo (ou matriz de trajectória) para um determinado tamanho do vector desfasado (comprimento da janela), e a decomposição em valores singulares que consiste em subdividir a série temporal na soma de pequenas componentes. A segunda etapa, também constituída por dois passos, tem como finalidade agrupar as componentes da decomposição em vários grupos e reconstruir a série original através da soma das componentes retidas como principais, eliminando aquelas que são essencialmente constituídas por ruído, de modo a que essa nova série reconstruída possa ser usada na previsão de novos valores. Um dos métodos de previsão, aplicado em muitos dos casos, é o algoritmo de previsão recorrente descrito detalhadamente em Golyandina *et al.* (2001).

3. Aplicação prática

Neste estudo é considerada a série mensal do número de vítimas mortais em acidentes de viação em Portugal Continental entre 1998 e 2009.

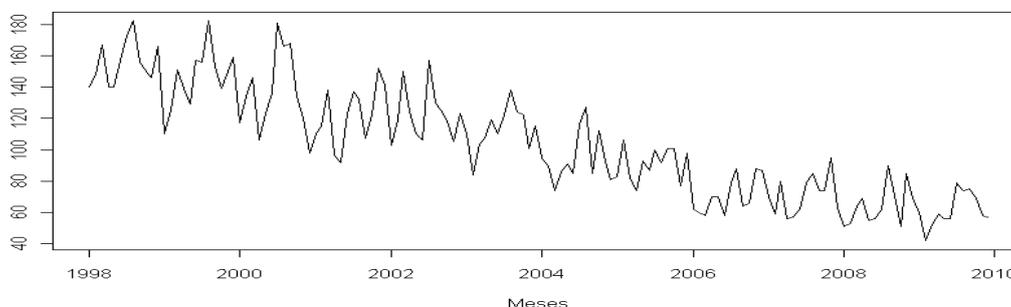


Gráfico 1: número de vítimas mortais em acidentes de viação em Portugal Continental.

Os dados do Gráfico 1, recolhidos ao longo destes 12 anos ($n = 144$ observações mensais), foram provenientes de entidades nacionais portuguesas detentoras das principais responsabilidades de trânsito e segurança rodoviária, nomeadamente a extinta Direcção Geral de Viação (DGV) que exerceu funções até 2006, dando origem à actual Autoridade Nacional de Segurança Rodoviária (ANSR), criada em 2007. No entanto, a entidade responsável pelo tratamento e ampla divulgação dos dados, a qual serviu de fonte de extracção dos mesmos, é o Instituto Nacional de Estatística (INE). Segundo o INE, a definição de vítima de acidente,

utilizada neste período em estudo, é aquela cujo óbito ocorra no local do evento ou no seu percurso até à unidade de saúde.

Esta abordagem consiste essencialmente em explorar as potencialidades da técnica da SSA aplicada à série temporal em causa. Não se pretende efectuar outro tipo de análise, nomeadamente verificar se existe ou não a influência de variados factores no número de vítimas mortais em acidentes de viação em Portugal.

3.1. Decomposição, reconstrução e previsão

Uma questão inicial importante na aplicação da SSA é a decisão acerca do comprimento apropriado da janela, m , para obter a série temporal desfasada ao longo do tempo. Este deve ter em conta o problema em estudo e a informação preliminar acerca da série temporal. Nesta aplicação prática é expectável que a série temporal tenha pelo menos a componente sazonal anual, pelo que é aconselhável usar um comprimento da janela proporcional a esse período para obtermos uma melhor separação das componentes periódicas (Hassani, 2007). Alguns resultados sugerem que m deverá ser elevado mas não ultrapassar $n/2$.

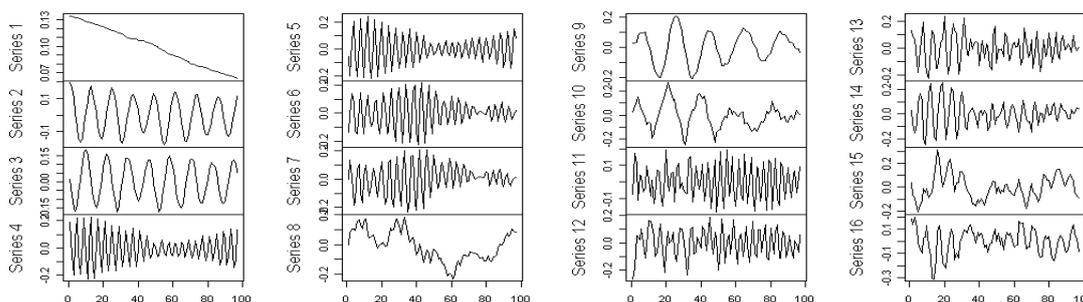


Gráfico 2: primeiras 16 componentes principais.

Por estas razões, utilizou-se $m = n/3 = 48$ meses. Tendo em conta a matriz de trajectória e a decomposição em valores singulares, seleccionaram-se as primeiras 16 componentes principais (CPs), ilustradas no Gráfico 2, de acordo com os valores próprios associados. A estas 16 CPs associaram-se sete pares de componentes oscilatórias (2-3, 4-5, 6-7, 9-10, 11-12, 13-14, 15-16), por ostentarem valores próprios similares, enquanto a CP1 corresponde à tendência decrescente dos dados e a CP8 exhibe uma mistura de tendência com componente oscilatória.

Foram analisados os espectros de cada uma das 16 CPs retidas, donde se concluiu que os períodos dos respectivos pares indicados são aproximadamente iguais a 12, 4, 4, 20, 4, 6 e 33 meses. Tendo em conta a natureza da série em estudo, é possível efectuar uma interpretação dos ciclos temporais mais relevantes no contexto da sinistralidade rodoviária. Para além da

compreensível sazonalidade anual, destaca-se o ciclo de 4 meses, uma vez que encontramos 3 principais picos de frequência (março a maio, julho/agosto e novembro/dezembro) em que o número de vítimas mortais é elevado. Estes meses coincidem com as férias escolares nas épocas festivas e com as férias da maioria dos portugueses, o que geralmente provoca uma maior circulação automóvel e como consequência um maior número de acidentes com vítimas mortais. Também as condições meteorológicas podem ser factores responsáveis pelo aumento do número de mortos, principalmente no pico do Inverno. Contudo há que ter em consideração que os dados em causa correspondem à totalidade dos mortos registados em todo o território nacional continental e não desagregados ao nível dos distritos, nos quais se verificam geralmente discrepâncias significativas entre as condições meteorológicas.

Um outro período a realçar na análise dos dados é o de 33 meses, que leva a pressupor uma relação com as alterações à legislação em vigor introduzidas na última década. Em 2003, para fazer face à elevada sinistralidade rodoviária, o Conselho Nacional de Segurança Rodoviária publicou o Plano Nacional de Prevenção Rodoviária; em 2005, introduziu-se a obrigatoriedade do uso do colete retrorreflector e o aumento das sanções pecuniárias; e em 2007 arrancou a campanha de sensibilização de segurança rodoviária “Vamos Travar Este Drama”.

Posteriormente reconstruímos a série original, sem ruído, à custa das 16 CPs identificadas anteriormente que permitem descrever o sinal através das componentes oscilatórias e da tendência, as quais foram utilizadas na aplicação do algoritmo de previsão recorrente com o intuito de prever novos valores mensais da série para o ano de 2010.

4. Conclusão

Este trabalho pretende ilustrar como a abordagem à técnica da SSA permite extrair as componentes oscilatórias e a tendência numa série temporal de dados relacionados com a sinistralidade rodoviária, área onde muitas das vezes apenas é feito um estudo exploratório ou gráfico. Os valores obtidos resultantes da previsão, mostram como a SSA desempenha também um papel fundamental neste campo.

Referências:

- ELSNER, J. B. & TSONIS, A. A. (1996) *Singular Spectrum Analysis. A New Tool in Time Series Analysis*, New York, Plenum Press.
- GOLYANDINA, N. E., NEKRUKTIN, V. V. & ZHIGLJAVSKY, A. A. (2001) *Analysis of Time Series Structure: SSA and Related Techniques*, Boca Raton, Chapman & Hall.
- HASSANI, H. (2007) Singular Spectrum Analysis: Methodology and Comparison. *Journal of Data Science*, 5, 239-257.