

# Análise de Séries Temporais

Fernando Lucambio

Departamento de Estatística  
Universidade Federal do Paraná

Julho, 2023

Nos últimos anos, uma técnica poderosa e computacionalmente intensiva conhecida como Análise de Espectro Singular (SSA) ou Análise da Séries Temporais Estruturais foi desenvolvida no campo da análise de séries temporais. SSA é uma técnica nova e poderosa aplicável a muitos problemas práticos, como o estudo de séries temporais clássicas, estatística multivariada, sistemas dinâmicos e processamento de sinais.

Possíveis áreas de aplicação do SSA são diversas: da matemática e física à economia e matemática financeira. Outras áreas podem incluir meteorologia e oceanologia às ciências sociais, pesquisa de mercado e medicina. Qualquer série temporal aparentemente complexa pode fornecer outro exemplo de aplicação bem-sucedida do SSA.

O SSA básico consiste em duas etapas complementares:

- ▶ decomposição e
- ▶ reconstrução;

ambos os estágios incluem duas etapas separadas.

No primeiro estágio decompomos a série e no segundo estágio reconstruímos a série original e usamos a série reconstruída para prever novos pontos de dados.

O conceito principal no estudo das propriedades do SSA é a separabilidade, que caracteriza o quão bem diferentes componentes podem ser separadas umas das outras.

A ausência de separabilidade aproximada é frequentemente observada em séries com estrutura complexa. Para essas séries e séries com estrutura especial, existem diferentes maneiras de modificar o SSA levando a diferentes versões, como SSA com centralização simples e dupla, Toeplitz SSA e SSA sequencial.

Notemos que, embora alguns conceitos probabilísticos e estatísticos sejam empregados nos métodos baseados em SSA, não precisamos fazer nenhuma suposição estatística, como estacionariedade das séries ou normalidade dos resíduos.

SSA é uma ferramenta muito útil que pode ser usada para resolver diferentes problemas como: encontrar tendências; alisamento; extração de componentes de sazonalidade; extração simultânea de ciclos com períodos pequenos e grandes, etc.

A resolução de todos estes problemas corresponde às chamadas capacidades básicas do SSA. Além disso, o método tem várias extensões essenciais.

Primeiro, a versão multivariada do método permite a expansão simultânea de várias séries temporais; ver, por exemplo

Danilov, D. and Zhigljavsky, A. (1997).

Em segundo lugar, as ideias do SSA levam a vários procedimentos de previsão para séries temporais; ver

Golyandina, N., Nekrutkin, V. and Zhigljavsky, A. (2001)

Danilov, D. and Zhigljavsky, A. (1997).

Além disso, as mesmas idéias são usadas em Golyandina, N., Nekrutkin, V., and Zhigljavsky, A. (2001) e Moskvina, V. G., and Zhigljavsky, A. (2003) para detecção de ponto de mudança em séries temporais.

Na área de análise de séries temporais não lineares, o SSA foi considerado como uma técnica que poderia competir com métodos clássicos (ARIMA, Holt-Winters, etc.).

Há uma série de pesquisas que consideraram o SSA como um método de filtragem ver, por exemplo, Hassani, H., Dionisio, A., and Ghodsi, M. (2010) e suas referências. Em outra pesquisa, a informação de ruído extraída pela técnica SSA, tem sido utilizada como teste de diagnóstico biomédico Ghodsi, M., Hassani, H., Sanei, S., and Hick, Y. (2009).

A técnica SSA também é utilizada como método de filtragem para medições longitudinais. Foi demonstrado que a redução de ruído é importante para o ajuste de curvas em modelos de curva de crescimento e que o SSA pode ser empregado como uma ferramenta poderosa para redução de ruído para medições longitudinais (Hassani, H. et al, 2009).

Estamos motivados a usar o SSA porque é uma técnica não paramétrica que trabalha com processos estatísticos arbitrários, sejam lineares ou não lineares, estacionários ou não estacionários, gaussianos ou não gaussianos.

A dinâmica das séries temporais geralmente passa por mudanças estruturais durante o período considerado e é preciso ter certeza de que o método de previsão não é sensível às variações dinâmicas.

Além disso, ao contrário dos métodos tradicionais de previsão de séries temporais, tanto modelos autorregressivos quanto estruturais que assumem normalidade e estacionaridade da série, o método SSA é não paramétrico e não faz suposições prévias sobre os dados.

A série em tempo real geralmente possui uma estrutura complexa desse tipo; como consequência, podemos encontrar superioridade da SSA sobre as técnicas clássicas. Além disso, o método SSA decompõe uma série em suas partes componentes e a reconstrói, deixando o componente aleatório para trás.

Outro aspecto importante do SSA, que pode ser muito útil em economia, é que, ao contrário de muitos outros métodos, ele funciona bem mesmo para amostras pequenas ver, por exemplo, Hassani, H. and Zhigljavsky, A. (2009) e Hassani, H. (2007).

Deve-se ressaltar que, embora conceitos probabilísticos e estatísticos sejam empregados nos métodos baseados no SSA, não são necessárias suposições estatísticas como estacionariedade série ou normalidade dos resíduos e o SSA usa o bootstrapping para obter os intervalos de confiança para as previsões.

Consideramos uma série temporal  $Y_T = (y_1, \dots, y_T)$  unidimensional. Seja  $L$  um número inteiro fixo tal que  $L \leq T/2$  representando o comprimento da janela e também seja  $K = T - L + 1$ .

- **Passo 1. Computando a matriz de trajetória:** Isso transfere a série temporal  $Y_T$  na série multidimensional  $X_1, \dots, X_K$  com vetores  $X_i = (y_i, \dots, y_{i+L-1}) \in \mathbb{R}^L$ . O parâmetro único da incorporação é o comprimento da janela  $L$ . O resultado desta etapa é a matriz de trajetória  $\mathbf{X} = [X_1, \dots, X_K]$ :

$$\mathbf{X} = (x_{ij})_{i,j=1}^{L,K} = \begin{pmatrix} y_1 & y_2 & y_3 & \cdots & y_K \\ y_2 & y_3 & y_4 & \cdots & y_{K+1} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ y_L & y_{L+1} & y_{L+2} & \cdots & y_T \end{pmatrix}.$$

Observe que a matriz de trajetória  $\mathbf{X}$  é uma matriz de Hankel, o que significa que todos os elementos ao longo da diagonal  $i + j = \text{constante}$  são iguais.

- ▶ **Passo 2. Construindo uma matriz para aplicar SVD:** Calcular a matriz  $\mathbf{XX}^T$ .
- ▶ **Passo 3. Decomposição em valores singulares (SVD) da matrix  $\mathbf{XX}^T$ :** Calcular os autovalores e autovetores da matriz  $\mathbf{XX}^T$  e representá-la como  $\mathbf{XX}^T = P\Lambda P^T$ .

Aqui  $\Lambda = \text{diag}(\lambda_1, \dots, \lambda_L)$  é a matriz diagonal de autovalores de  $\mathbf{XX}^T$  de modo que

$$\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_L \geq 0$$

e

$$P = (P_1, P_2, \dots, P_L)$$

a matriz ortogonal correspondente de autovetores de  $\mathbf{XX}^T$ .

- ▶ **Passo 4. Seleção de autovetores:** Selecione um grupo de  $l$ ,  $1 \leq l \leq L$  autovetores  $P_{i_1}, P_{i_2}, \dots, P_{i_l}$ .

O passo de agrupamento corresponde a dividir as matrizes elementares  $X_i$  em vários grupos e somando as matrizes em cada grupo. Seja  $I = \{i_1, \dots, i_l\}$  um grupo de índices. Em seguida, a matriz  $X_I$  correspondente ao grupo  $I$  é definida como

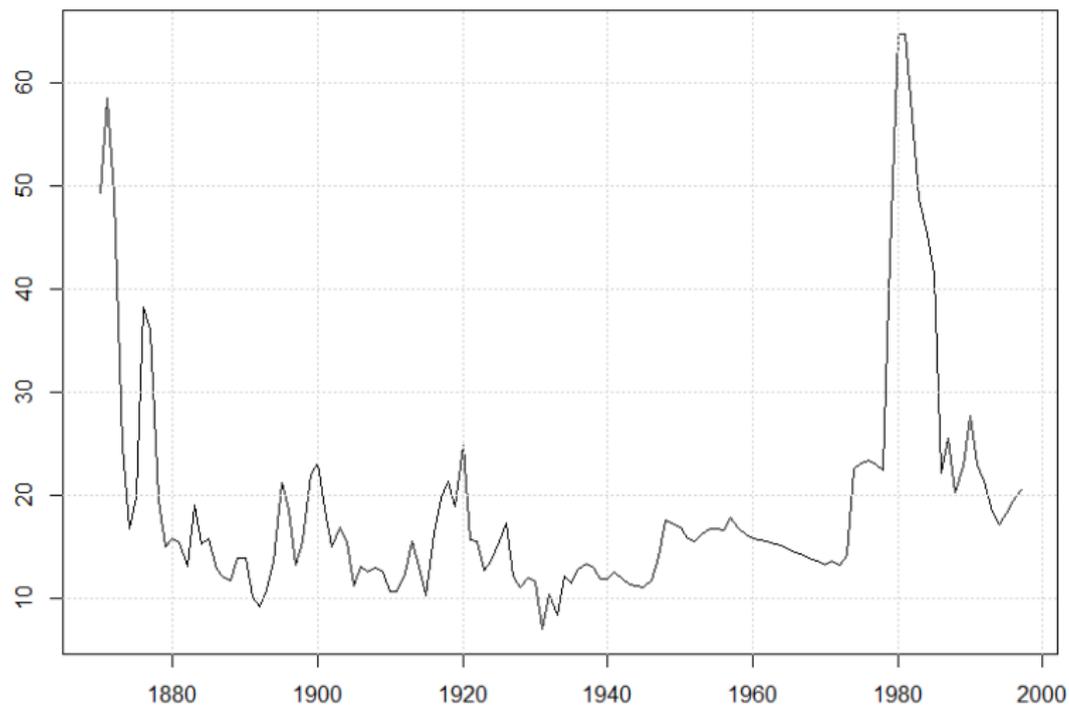
$$X_I = X_{i_1} + \dots + X_{i_l}.$$

- ▶ **Passo 5. Reconstrução da série unidimensional:** Calcule a matriz

$$\tilde{\mathbf{X}} = \|\tilde{X}_{ij}\| = \sum_{k=1}^l P_{i_k} P_{i_k}^T \mathbf{X},$$

como uma aproximação para  $\mathbf{X}$ . A transição para a série unidimensional pode agora ser alcançada pela média das diagonais da matriz  $\tilde{\mathbf{X}}$ .

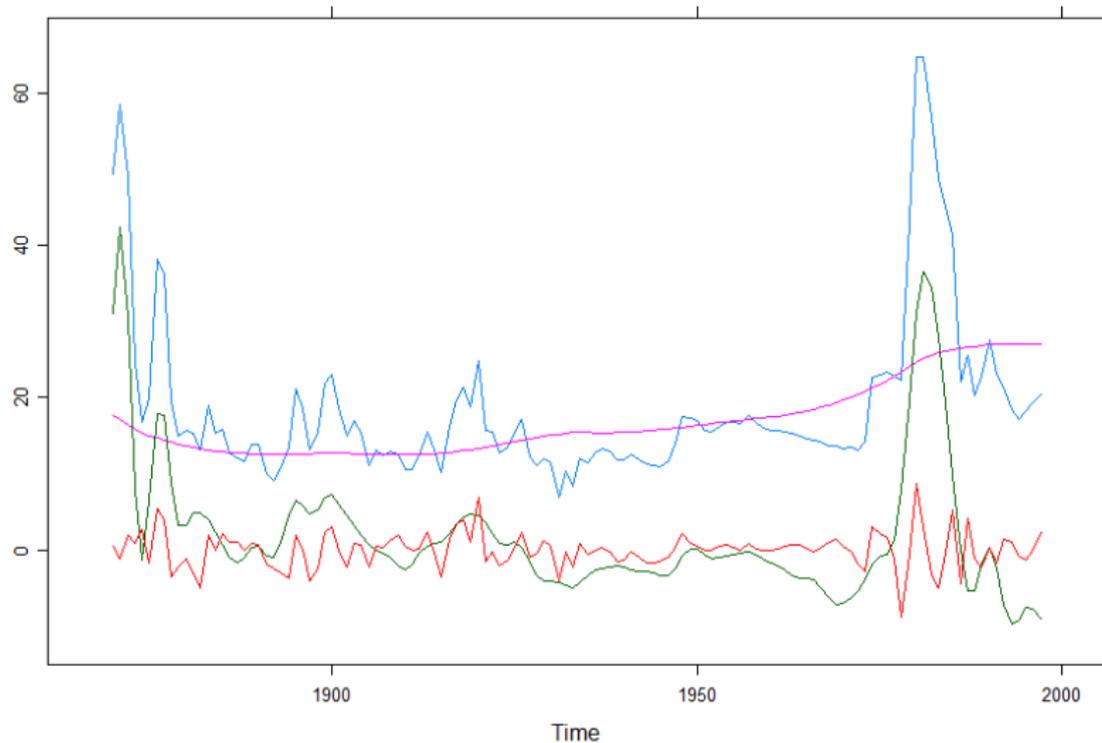
### Preços do petróleo em dólares constantes de 1997: 1870-1997



```
library(Rssa)
library(fma)
plot(oilprice, main="Preços do petróleo em dólares
constantes de 1997: 1870-1997", ylab = "", xlab = "")
grid() ajuste = ssa(oilprice)
ajuste1 = reconstruct(ajuste,
groups = list(Trend = 1, Seasonality = 2:11))
plot(ajuste1, add.residuals = TRUE, add.original = TRUE,
plot.method = "xyplot",superpose = TRUE,
auto.key = list(columns = 2))
grid()
```

## Reconstructed Series

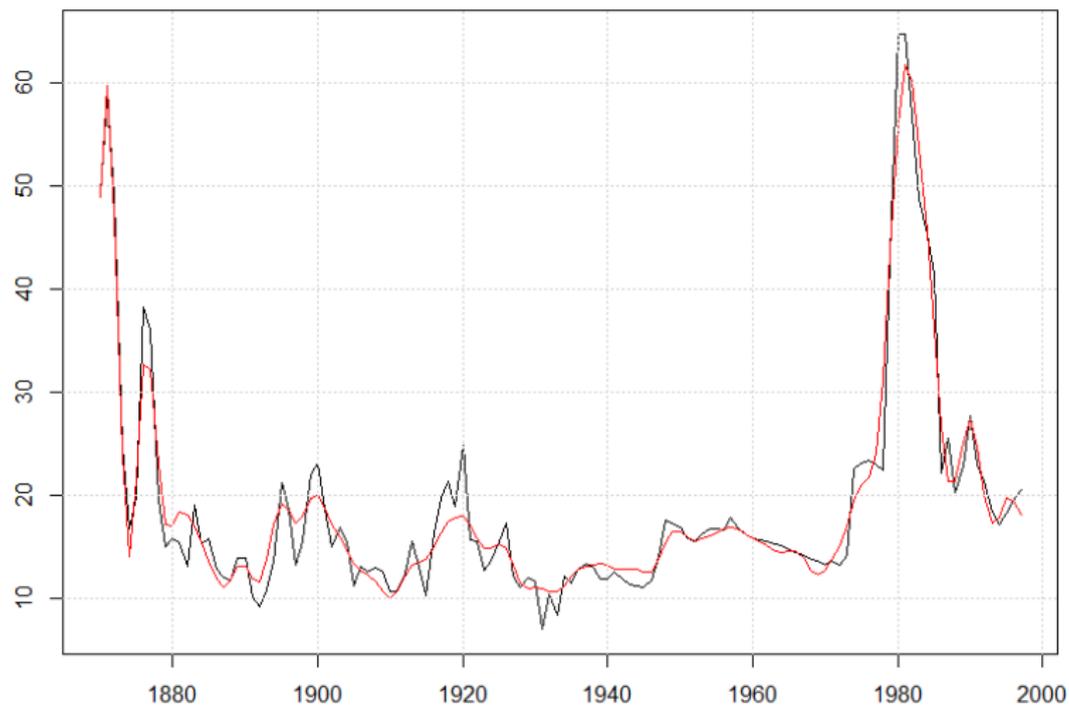
Original ——— Seasonality ———  
Trend ——— Residuals ———



```
plot(oilprice, main="Preços do petróleo em dólares  
constantes de 1997: 1870-1997", ylab = "", xlab = "")
```

```
lines(ajuste1$Trend+ajuste1$Seasonality, col = "red")  
grid()
```

### Preços do petróleo em dólares constantes de 1997: 1870-1997



- ▶ Golyandina, N., Nekrutkin, V., and Zhigljavsky, A. (2001). Analysis of Time Series Structure: SSA and related techniques, Chapman & Hall/CRC, New York - London.
- ▶ Broomhead, D. S., and King, G. (1986). Extracting qualitative dynamics from experimental data, Physica D, 20, pp.217-236.
- ▶ Danilov, D., and Zhigljavsky, A. (1997) (Eds.). Principal Components of Time Series: the 'Caterpillar' method, University of St. Petersburg, St. Petersburg. (In Russian.).
- ▶ Elsner, J. B., and Tsonis, A. A. (1996). Singular Spectrum Analysis, A New Tool in Time Series Analysis, Plenum Press, New York and London.

- ▶ Hassani, H., and Zhigljavsky, A. (2009). Singular Spectrum Analysis: Methodology and Application to Economics Data, Journal of System Science and Complexity, 22(3), pp. 372-394.
- ▶ Vautard, R., Yiou, P., and Ghil, M. (1992). Singular spectrum analysis: a toolkit for short, noisy and chaotic series, Physica D, 58, pp. 95-126.
- ▶ Moskvina, V. G., and Zhigljavsky, A. (2003). An algorithm based on singular spectrum analysis for change-point detection, Communication in Statistics - Simulation and Computation, 32(4), pp. 319-352.
- ▶ Hassani, H. (2007). Singular Spectrum Analysis: Methodology and Comparison, Journal of Data Science, 5(2), pp.239-257.

- ▶ Hassani, H., Heravi, S., and Zhigljavsky, A. (2009). Forecasting European Industrial Production with Singular Spectrum Analysis, *International Journal of Forecasting*, 25(1), pp. 103-118.
- ▶ Patterson, K; Hassani, H; Heravi, S; Zhigljavsky, A. (2010). Forecasting the Final Vintage of the Index of Industrial Production, *Journal of Applied Statistics*, Forthcoming.
- ▶ Alexandrov, Th., and Golyandina, N. (2004). The automatic extraction of time series trend and periodical components with the help of the Caterpillar-SSA approach. *Exponenta Pro* 3-4 (In Russian.), pp. 54-61.