

# Uma breve introdução à Análise de Espectro Singular

Andre Almeida de Mello

Dezembro de 2023

## 1 Introdução

O objetivo deste trabalho é apresentar, em um nível introdutório, a SSA (*Singular Spectrum Analysis*, ou Análise de Espectro Singular) como ferramenta de análise de séries temporais. A exposição que se segue baseia-se em artigo de Golyandina e Korobeynikov (2013), assim como em Cassiano (2014). Além dos apontamentos teóricos, ao final do relatório apresentamos um exemplo simples de aplicação da SSA no *software* R.

## 2 Preliminares

### 2.1 Decomposição Espectral

Antes de tratarmos da decomposição em valores singulares, técnica que embasa a SSA, trataremos da amplamente conhecida decomposição espectral, visto que a decomposição em valores singulares nada mais é do que a generalização desta técnica para matrizes não quadradas.

A decomposição espectral é um conceito importante em álgebra linear, onde uma matriz quadrada é decomposta com base em seus autovetores e autovalores associados. Dada uma matriz quadrada  $A$ , a decomposição espectral é expressa como:

$$A = P\Lambda P^{-1}$$

onde  $P$  é uma matriz cujas colunas são autovetores de  $A$  e  $\Lambda$  é a matriz diagonal cujos elementos não nulos são os autovalores associados às colunas de  $P$  (respeitando a ordem).

A matriz  $P$  é invertível, e  $P^{-1}$  é a inversa de  $P$ . Note que se  $P$  for uma matriz ortogonal, então a decomposição espectral de  $A$  pode ser escrita como:

$$A = P\Lambda P^T$$

A expansão baseada na decomposição espectral de uma matriz  $A$  envolve a expressão da matriz original como uma soma ponderada dos seus autovetores e autovalores associados. Essa expansão é expressa como:

$$A = \lambda_1 \mathbf{v}_1 \mathbf{v}_1^T + \lambda_2 \mathbf{v}_2 \mathbf{v}_2^T + \dots + \lambda_n \mathbf{v}_n \mathbf{v}_n^T$$

onde  $\lambda_i$  são os autovalores de  $A$  e  $\mathbf{v}_i$  são os autovetores correspondentes.

### 2.2 Decomposição em Valores Singulares

Conforme já indicado, a decomposição em valores singulares (SVD, do inglês *Singular Value Decomposition*), é uma generalização da decomposição espectral para matrizes não quadradas.

A SVD descreve uma matriz como o produto de três outras matrizes específicas, fornecendo uma representação útil e compacta da estrutura da matriz original. Matematicamente, para uma matriz  $A$  de dimensões  $m \times n$ , a decomposição singular pode ser expressa como:

$$A = U\Sigma V^T$$

onde  $U$  é uma matriz ortogonal  $m \times m$ ,  $\Sigma$  é uma matriz  $n \times n$  diagonal com os chamados “valores singulares” nas diagonais e  $V$  uma matriz ortogonal  $n \times n$ .

Os valores singulares na matriz  $\Sigma$  são números não negativos e representam a "força" ou a importância de cada componente na decomposição. Eles são frequentemente utilizados para redução de dimensionalidade e análise de componentes principais. A matriz  $U$  contém os autovetores associados à matriz  $AA^T$  (produto da matriz original com sua transposta), enquanto  $V$  contém os autovetores associados à matriz  $A^T A$  (produto da transposta da matriz original com a matriz original).

A expansão com base em valores singulares de  $A$  pode ser expressa como uma soma de produtos externos dos vetores das colunas de  $U$  e  $V$ , multiplicados pelos valores singulares correspondentes em  $\Sigma$ :

$$A = \sigma_1 \mathbf{u}_1 \mathbf{v}_1^T + \sigma_2 \mathbf{u}_2 \mathbf{v}_2^T + \dots + \sigma_r \mathbf{u}_r \mathbf{v}_r^T$$

onde  $\sigma_i$  são os valores singulares,  $\mathbf{u}_i$  são as colunas de  $U$ ,  $\mathbf{v}_i$  são as colunas de  $V$  e  $r$  é o posto de  $A$ .

Esta representação mostra como a matriz original pode ser aproximada pela soma ponderada de componentes, onde cada componente é um produto externo de um autovetor de  $U$ , um valor singular e um autovetor de  $V$ .

A expansão com base em valores singulares é frequentemente usada em redução de dimensionalidade, compressão de dados e aproximação de matrizes, permitindo uma representação mais compacta dos dados mantendo as informações mais importantes.

## 3 O algoritmo SSA

### 3.1 Visão Geral

A SSA é uma poderosa técnica utilizada na análise de séries temporais. Desenvolvida inicialmente na década de 1980, a SSA tem se mostrado eficaz em decompor uma série temporal complexa em componentes mais simples, facilitando a identificação de padrões, tendências e ciclos.

A SSA é baseada na decomposição de uma série temporal em componentes chamados "componentes singulares". Cada componente singular representa uma parte específica da variabilidade da série temporal, como tendências, componentes periódicos ou ruído. A técnica é não paramétrica, o que significa que não faz suposições sobre a forma funcional dos padrões temporais.

O processo se dá em dois estágios, decomposição e agrupamento, que por sua vez podem ser subdivididos em quatro passos. No estágio da decomposição, é feita a chamada "incorporação" (*embedding*) da série temporal em uma matriz retangular denominada "matriz trajetória" (primeiro passo) e a decomposição em valores singulares desta mesma matriz trajetória (segundo passo). Conforme a expansão baseada nos valores singulares acima definida, resultará deste processo uma sequência de matrizes denominadas matrizes elementares. Caso a série temporal satisfaça a condição de separabilidade que definiremos adiante, cada matriz elementar representará um componente de tendência, periódico ou de ruído. No estágio do agrupamento, primeiramente as matrizes elementares são agrupadas conforme o componente da série temporal que representam (terceiro passo) e, caso seja de interesse, a partir das matrizes resultantes do agrupamento, uma série com a mesma dimensão da original, porém menos ruidosa, é reconstruída e possivelmente modelada (quarto passo).

Como já se pode entrever da breve apresentação dos passos do algoritmo, algumas das principais aplicações da SSA estão relacionadas à decomposição de séries temporais, como filtragem de ruídos, identificação de tendências e sazonalidade de diversos períodos. Além disso, como veremos, a SSA é útil também para a previsão de valores futuros.

Antes de discutir os parâmetros da SSA e a já mencionada questão da separabilidade, com a finalidade de facilitar o entendimento, voltaremos aos passos do algoritmo para explicá-los com maior detalhe.

### 3.2 Primeiro estágio: decomposição

#### 3.2.1 Primeiro passo: incorporação

Para realizar a incorporação (*embedding*), mapeamos a série temporal original em uma sequência de vetores defasados de tamanho  $L$  formando  $K = N - L + 1$  vetores defasados, onde  $X_i = (x_i, \dots, x_{i+L-1})^T$ , para  $i = 1, \dots, K$ .

$$\mathbf{X} = [X_1 : \dots : X_K] = (x_{ij})_{L,K}^{i,j=1}$$

$$\mathbf{X} = \begin{bmatrix} x_1 & x_2 & x_3 & \dots & x_K \\ x_2 & x_3 & x_4 & \dots & x_{K+1} \\ x_3 & x_4 & x_5 & \dots & x_{K+2} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ x_L & x_{L+1} & x_{L+2} & \dots & x_N \end{bmatrix}$$

Cabe mencionar duas importantes propriedades da matriz trajetória  $\mathbf{X}$ : a) tanto linhas quanto colunas são subséries da série original; e b) as antidiagonais de  $\mathbf{X}$  são constantes, o que equivale dizer que  $\mathbf{X}$  é uma matriz de Hankel.

### 3.2.2 Segundo passo: decomposição em valores singulares

Dada a base ortonormal  $\{P_i\}_{i=1}^L$ , e escolhendo  $\{P_i\}_{i=1}^L$  como sendo os autovetores de  $\mathbf{X}\mathbf{X}^T$ , é feita a decomposição em valores singulares da matriz de trajetória, isto é:

$$\mathbf{X} = \sum_{i=1}^L P_i Q_i^T = \sum_{i=1}^L \sqrt{\lambda_i} U_i V_i^T = \sum_i^L X_i$$

Chamamos de  $P_i = U_i$  os vetores singulares à esquerda de  $\mathbf{X}$  e de  $V_i$  os vetores singulares à direita, sendo  $Q_i = \sqrt{\lambda_i} V_i$ . Além disso,  $\lambda_i$  são os autovalores da matriz  $\mathbf{X}\mathbf{X}^T$ . Sendo esta positiva semidefinida, temos que  $\lambda_i \geq 0$ . Já a trinca  $(\sqrt{\lambda_i}, P_i, Q_i)$  é chamada de autotripla.

## 3.3 Reconstrução

### 3.3.1 Terceiro passo: agrupamento das autotriplas

Seja  $d = \max\{j : \lambda_j \neq 0\}$ . Uma vez obtida a expansão acima descrita, o procedimento de agrupamento consiste em particionar o conjunto de índices  $\{1, \dots, d\}$  em  $m$  subconjuntos disjuntos  $I_1, \dots, I_m$ . Defina  $\mathbf{X}_I = \sum_{i \in I} X_i$ . Temos agora que

$$\mathbf{X} = \mathbf{X}_{I_1} + \dots + \mathbf{X}_{I_m},$$

O procedimento de escolher os conjuntos  $I_1, \dots, I_m$  é chamado de agrupamento de autotriplas. Se  $m = d$  e  $I_j = \{j\}$ ,  $j = 1, \dots, d$ , então o agrupamento correspondente é chamado de elementar. Para a SSA básica, a escolha de várias autotriplas principais corresponde à aproximação da série temporal em vista da conhecida propriedade de otimalidade da SVD.

### 3.3.2 Quarto passo: média diagonal

Neste passo, transformamos cada matriz  $X_{I_j}$  da decomposição agrupada em uma nova série de comprimento  $N$ . Seja  $Y$  uma matriz  $L \times K$  com elementos  $y_{ij}$ ,  $1 \leq i \leq L$ ,  $1 \leq j \leq K$ , e suponha, para simplificar, que  $L \leq K$ . Ao realizar a média diagonal, transferimos a matriz  $Y$  para a série  $(y_1, \dots, y_N)$  usando a fórmula

$$\tilde{y}_s = \sum_{(l,k) \in A_s} \frac{y_{lk}}{|A_s|}$$

onde  $A_s = \{(l, k) : l+k = s+1, 1 \leq l \leq L, 1 \leq k \leq K\}$  e  $|A_s|$  denota o número de elementos no conjunto  $A_s$ . Isso equivale a calcular a média aritmética dos elementos da matriz ao longo das antidiagonais. A média diagonal aplicada a uma das matrizes  $X_{I_k}$  produz uma série reconstruída  $X_e(k) = (x_{e(k)1}, \dots, x_{e(k)N})$ . Portanto, a série inicial  $(x_1, \dots, x_N)$  é decomposta em uma soma de  $m$  séries reconstruídas:

$$x_n = \sum_{k=1}^m \tilde{x}_n^{(k)}, n = 1, \dots, N$$

### 3.4 Separabilidade e escolha de parâmetros

Duas questões importantes têm relação com a possibilidade de uma decomposição apropriada da série temporal observada e os parâmetros envolvidos.

A noção de separabilidade responde a essa pergunta. A separabilidade de duas séries temporais  $X_N^{(1)}$  e  $X_N^{(2)}$  significa a possibilidade de extrair  $X_N^{(1)}$  a partir da soma observada  $X_N^{(1)} + X_N^{(2)}$ . A SSA pode separar aproximadamente, por exemplo, sinal e ruído, ondas senoidais com diferentes frequências, tendência e sazonalidade, entre outros.

Se duas séries temporais forem aproximadamente separáveis, surge o problema de identificação dos termos correspondentes, por exemplo, a  $X_N^{(1)}$ . Os componentes das séries temporais podem ser identificados com base no seguinte princípio: a forma de um autovetor replica a forma do componente da série temporal que produz esse autovetor. Assim, os gráficos dos autovetores podem auxiliar no processo de identificação. Além disso, uma senoide gera, exata ou aproximadamente, dois componentes de ondas senoidais com a mesma frequência e defasagem de  $\pi/2$ . Portanto, o gráfico de dispersão de um par de autovetores que produz um polígono de vértice T mais ou menos regular, pode ajudar a identificar uma senoide de período T. Para os problemas de extração de sinal, suavização e redução de ruído, são escolhidos várias autotriplas principais.

Informações muito úteis para a separação estão contidas na chamada matriz de w-correlação. Esta é a matriz composta por correlações ponderadas entre os componentes das séries temporais reconstruídas. Os pesos refletem o número de entradas dos termos das séries temporais em sua matriz de trajetória. Componentes bem separados têm correlação pequena, enquanto componentes mal separados têm correlação grande. Portanto, ao observar a matriz de w-correlação, é possível encontrar grupos de séries reconstruídas elementares correlacionadas e usar essas informações para o agrupamento consequente. Uma das regras é não incluir em grupos diferentes os componentes correlacionados.

As condições de separabilidade (aproximada) fornecem recomendações para a escolha do comprimento da janela  $L$ : ela deve ser grande o suficiente ( $L \sim N/2$ ) e, se desejamos extrair um componente periódico com período conhecido, então os comprimentos de janela que são divisíveis pelo período proporcionam uma melhor separabilidade. A SSA com  $L$  pequeno realiza a suavização da série por meio de um filtro de ordem  $2L - 1$  (Golyandina e Zhigljavsky (2013)), se escolhermos algumas autotriplas principais. Geralmente, a escolha do comprimento da janela é importante, mas o resultado é estável em relação a pequenas alterações de  $L$ .

Se a série temporal possui uma estrutura complexa, é recomendado a chamada SSA Sequencial. A SSA Sequencial consiste em duas etapas, na primeira etapa a tendência é extraída com um pequeno comprimento de janela e depois os componentes periódicos são detectados e extraídos do resíduo com  $L \sim N/2$ .

Se utilizarmos a SSA como uma técnica exploratória e livre de modelo (*model free*), então a justificativa da decomposição não é formal. Ela é baseada na teoria de separabilidade e interpretabilidade dos resultados. O processamento em tempo real ou em lotes pela SSA é possível se a classe de séries for suficientemente especializada, permitindo-nos estabelecer a regra para escolher parâmetros adequados. Para realizar testes estatísticos, um modelo concreto deve ser especificado.

### 3.5 Previsões recursivas

Seja  $I$  o conjunto escolhido de autotriplas,  $P_i \in R^L$ ,  $i \in I$ , os autovetores correspondentes,  $\underline{P_i}$  as suas primeiras  $L - 1$  coordenadas,  $\pi_i$  seja a última coordenada de  $P_i$ ,  $\nu^2 = \sum_i \pi_i^2$ . Defina  $R = (a_{L-1}, \dots, a_1)^T$  como

$$R = \frac{1}{1 - \nu^2} \sum_{i \in I} \pi_i \underline{P_i}$$

O algoritmo para previsões recursivas pode ser formulado como se segue.

Primeiro, a série temporal  $\{y_{N+M} = (y_1, \dots, y_{N+M})$  é definida como:

$$y_i = \begin{cases} \tilde{x}_i, & \text{para } i = 1, \dots, N \\ \sum_{j=1}^{L-1} a_j y_{i-j}, & \text{para } i = N + 1, \dots, N + M \end{cases}$$

Então, os números  $y_{N+1}, \dots, y_{N+M}$  são os  $M$  termos da previsão. Em outras palavras, as previsões são obtidas pelo uso direto de uma relação de recorrência linear com coeficientes  $\{a_j, j = 1, \dots, L - 1\}$ .

### 3.6 Um exemplo simples de SSA

Para ilustrar o procedimento, apresentamos um exemplo simples de SSA sobre uma série temporal de pequena dimensão. Seja a série temporal  $\{x_t\} = \{1, 2, 3, 4, 5, 6\}$ ,  $t = 1, \dots, 6$ .

#### 3.6.1 Incorporação

Tomando  $L = 3$ , a matriz trajetória é dada por

$$\mathbf{X} = \begin{bmatrix} 1 & 2 & 3 & 4 \\ 2 & 3 & 4 & 5 \\ 3 & 4 & 5 & 6 \end{bmatrix}$$

#### 3.6.2 Decomposição

Vamos primeiro determinar  $\mathbf{X}\mathbf{X}^T$ , seus autovalores e autovetores.

$$\mathbf{X}\mathbf{X}^T = \begin{bmatrix} 30 & 40 & 50 \\ 40 & 54 & 68 \\ 50 & 68 & 86 \end{bmatrix}$$

$$\lambda_1 \approx 169.29$$

$$\lambda_2 \approx 0.71$$

$$\lambda_3 \approx 0$$

$$U_1 = \begin{bmatrix} -0.418 \\ -0.565 \\ -0.712 \end{bmatrix}$$

$$U_2 = \begin{bmatrix} 0.812 \\ 0.12 \\ -0.572 \end{bmatrix}$$

$$U_3 = \begin{bmatrix} 0.408 \\ -0.816 \\ 0.408 \end{bmatrix}$$

Como  $\lambda_3 = 0$ , o posto de  $\mathbf{X}$  é  $d = 2$ . Logo, precisamos agora de  $V_1$  e  $V_2$ :

$$V_1 = \begin{bmatrix} -0.283 \\ -0.413 \\ -0.543 \\ -0.674 \end{bmatrix}$$

$$V_2 = \begin{bmatrix} -0.788 \\ -0.361 \\ -0.066 \\ -0.494 \end{bmatrix}$$

#### 3.6.3 Agrupamento

Podemos agora escrever  $\mathbf{X} = \mathbf{X}_1 + \mathbf{X}_2$ :

$$\mathbf{X}_1 = \sqrt{\lambda_1} U_1 V_1^T = \begin{bmatrix} 1.539 & 2.247 & 2.953 & 3.661 \\ 2.079 & 3.036 & 3.993 & 4.950 \\ 2.261 & 3.827 & 5.033 & 6.239 \end{bmatrix}$$

$$\mathbf{X}_2 = \sqrt{\lambda_2} U_2 V_2^T = \begin{bmatrix} -0.539 & -0.247 & 0.047 & 0.339 \\ -0.079 & -0.036 & 0.007 & 0.050 \\ 0.379 & 0.173 & -0.033 & -0.239 \end{bmatrix}$$

## 4 Exemplo de aplicação no software R

Nesta seção iremos apresentar brevemente um exemplo simples de aplicação da SSA em uma série temporal. Trata-se do conjunto de dados CO2, disponível no R 1). Foi utilizado nas análises o pacote **Rssa**.

Após a decomposição em valores singulares da matriz trajetória da série com  $L = 120$ , obtivemos os autovalores (Figura 2) e autovetores (Figura 3).

Uma inspeção visual indica que os componentes 1 e 4 compõem a tendência, enquanto os pares 2-3 e 5-6 são periódicos.

Isto se confirma com a inspeção da Figura 4.

Na etapa de reconstrução, agrupando os componentes 1-4, 2-3 e 5-6, obtivemos a série reconstruída conforme mostrada na Figura 6.

Por fim, a Figura mostra um exemplo de previsão 36 passos a futuro, com base no componente de tendência, obtido pelo método recursivo acima descrito.

Esta aplicação, ainda que simples, ilustra o potencial da SSA em decompor uma série temporal de forma precisa.

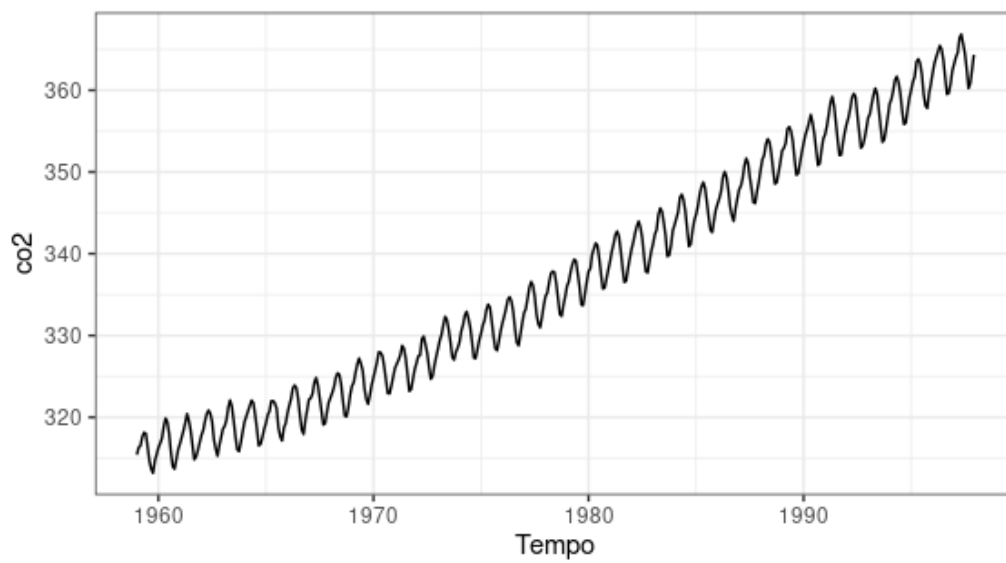


Figura 1: Série CO2.

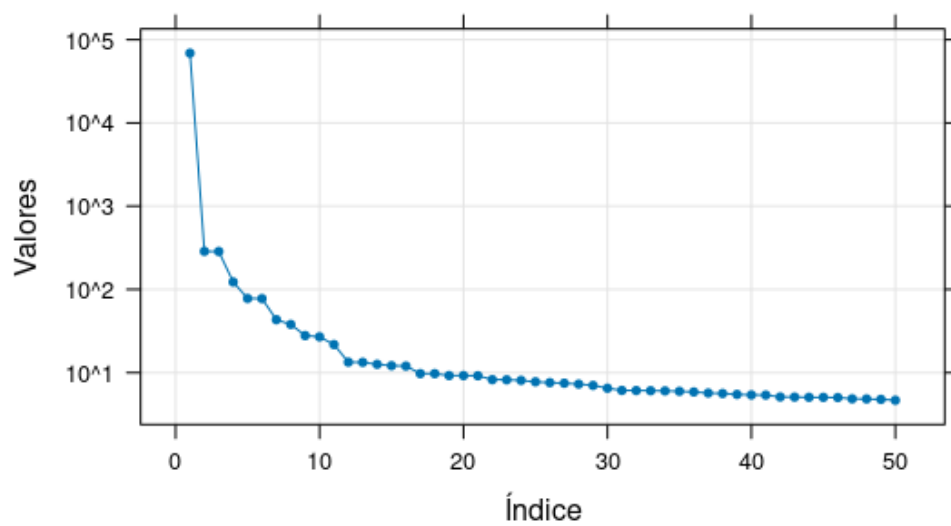


Figura 2: Autovalores da SVD da série CO2.

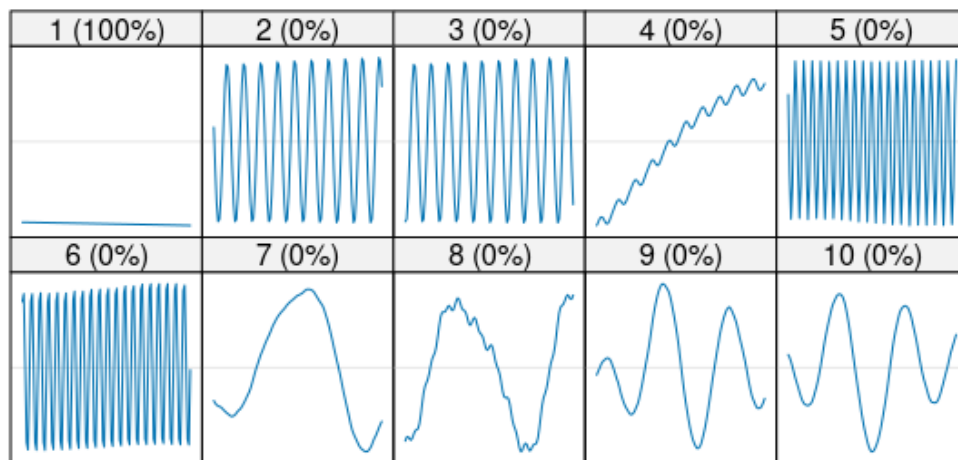


Figura 3: Autovetores da SVD da série CO2.

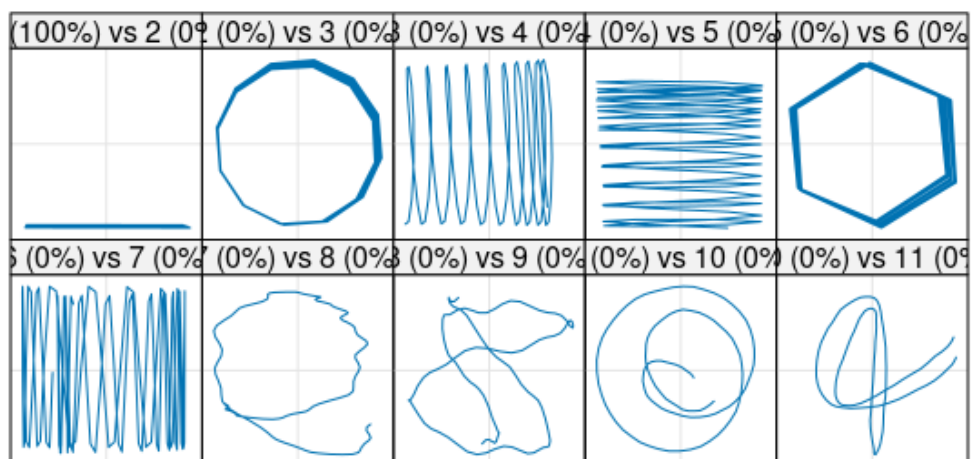


Figura 4: Pares de autovetores da SVD da série CO2.



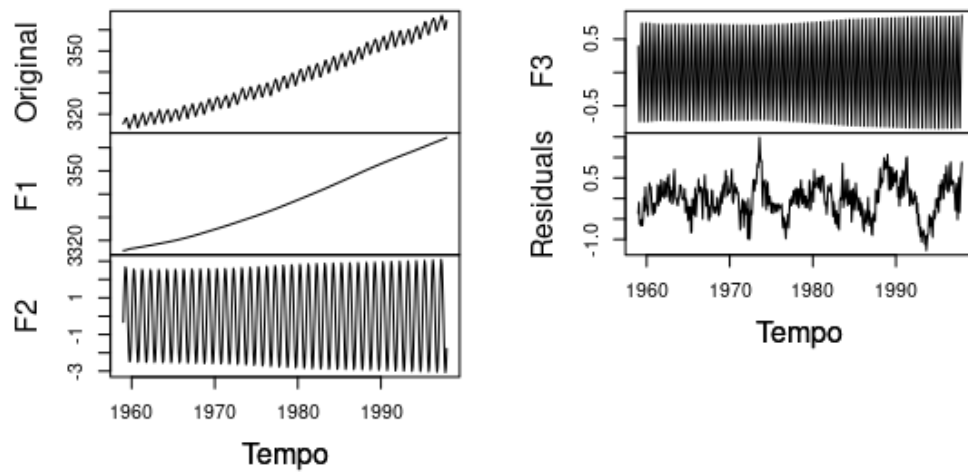


Figura 5: Série CO2 reconstruída.

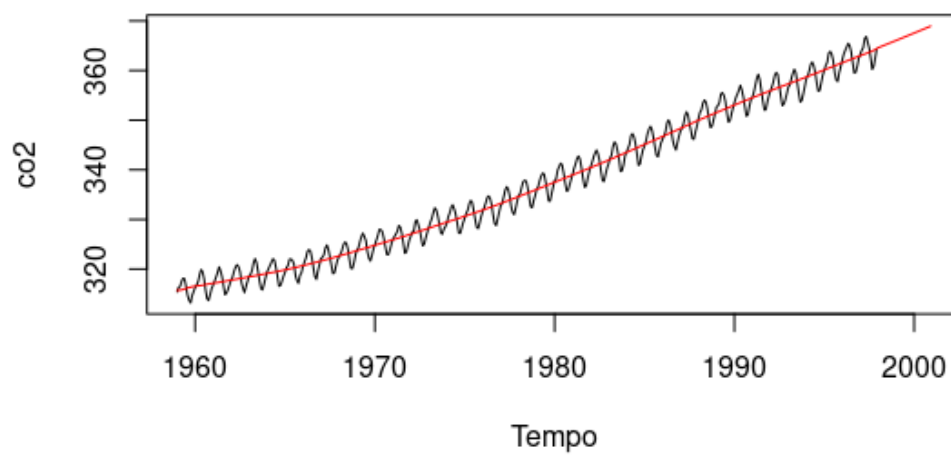


Figura 6: Previsões 36 passos a futuro.

## 5 Referências

1. CASSIANO, Keila Mara. Análise de Séries Temporais Usando Análise Espectral Singular (SSA) e Clusterização de Suas Componentes Baseada em Densidade. Tese (doutorado) – Pontifícia Universidade Católica do Rio de Janeiro, Departamento de Engenharia Elétrica, 2014.
2. GOLYANDINA, Nina & KOROBEYNIKOV, Anton. Basic Singular Spectrum Analysis and Forecasting with R. Preprint submitted to Computational Statistics & Data Analysis January 22, 2013.