

### 3. Singular Spectrum Analysis (SSA)

*Singular Spectrum Analysis (SSA)* é uma poderosa técnica que vem sendo aplicada em diversas áreas da Matemática, Física, Economia, Matemática Financeira, Meteorologia, Oceanografia, Ciências Sociais e em recentes Análises de Séries Temporais (HASSANI, 2007). A abordagem incorpora elementos de análise clássica de séries temporais, estatística multivariada, geometria multivariada, sistemas dinâmicos e processamentos de sinais (GOLYANDINA et al., 2001).

A motivação para o desenvolvimento de SSA encontra-se na decomposição clássica de Karhunen-Loève (LOÈVE, 1978), e outros resultados sobre a representação ortogonal de processos estocásticos. As primeiras referências da abordagem SSA são geralmente atribuídas aos trabalhos de BASILEVSKY & HUM (1979), BROOMHEAD & KING (1986a, 1986b) e BROOMHEAD et al. (1987). Entretanto, somente após 1989 as principais referências com aplicações foram publicadas por outros autores da Rússia, dos Estados Unidos e do Reino Unido que também desenvolviam estudos em SSA paralelamente e de forma independente, como os trabalhos de VAUTARD & GHIL (1989); GHIL & VAUTARD (1991) E VAUTARD et al. (1992).

O objetivo da SSA é fazer uma decomposição da série original em uma soma de componentes independentes e interpretáveis como componentes de tendência, componentes oscilatórias e componentes com estrutura de ruído. SSA é uma ferramenta que pode ser usada para resolver os seguintes problemas: 1) encontrar tendências de diferentes resoluções; 2) suavizar séries temporais; 3) extrair componentes sazonais; 4) extrair ciclos com pequenos e grandes períodos; 5) extrair periodicidades com amplitudes variáveis; 6) extrair tendências complexas e periodicidades; 7) encontrar estrutura em séries temporais curtas 8) detectar pontos de mudança na série e 9) fazer previsão de séries temporais.

No caso particular desta tese, a suavização de séries temporais será tratada através da remoção de ruídos. A princípio não é atribuído nenhum significado estocástico ao

termo ruído, e o conceito de um ruído estacionário determinístico é mais conveniente para SSA uma vez que ele é tratado com uma simples trajetória de uma série temporal melhor do que com uma amostra de tais trajetórias (GOLYANDINA et al., 2001).

Uma das vantagens do método proposto em SSA é sua abordagem não paramétrica. Apesar de alguns conceitos probabilísticos e estatísticos serem empregados nos métodos baseados em SSA, não é necessário conhecer o modelo paramétrico para a série temporal e não são requisitadas quaisquer suposições estatísticas tais como estacionariedade da série ou normalidade dos resíduos.

Outra característica importante do SSA é que ela pode ser usada para análise de séries relativamente curtas. Tem sido observado em muitas aplicações práticas que as características assintóticas são apresentadas para valores relativamente pequenos do tamanho da série  $T$ . Não é incomum aplicações de SSA com êxito a séries com  $T$  compreendido no intervalo de 20 a 30. Uma introdução elementar ao SSA pode ser encontrada em GOLYANDINA et al. (2001).

O método básico SSA consiste em quatro fases: 1) Decomposição da Série em uma Matriz Hankel, 2) Decomposição da Matriz Hankel via *Singular Value Decomposition* (SVD), 3) Reconstrução das subséries via Média Diagonal 4) Agrupamento das séries e identificação das componentes de tendência, componentes harmônicas e ruído, com vistas a retirada do ruído. As fases 1 e 2 constituem o estágio de Decomposição da série Temporal e as fases 3 e 4 garantem a Reconstrução da Série temporal. Esta série reconstruída pode ser modelada para previsão pelos métodos tradicionais ou pelo método de previsão SSA.

Na literatura outros termos podem ser encontrados para designar a metodologia SSA: *Caterpillar*, *Singular System Analysis*, *Empirical Orthogonal Function (EOF)*; *Singular Spectrum Approach* ou ainda, Expansão de Karhunen-Loeve. Nesta tese será usada somente a terminologia usual, *Singular Spectrum Analysis*, ou simplesmente SSA. ELSNER & TSONIS (1996) discutem sobre a etimologia do nome da técnica SSA. Nesta ênfase, o termo “*Spectrum*” tem sua origem na decomposição espectral de uma matriz  $X$  através do seu conjunto de autovalores  $\lambda$  denominado “*spectrum set*”. O termo singular remete a matriz singular  $(X - \lambda I)$ . Entretanto o termo *Singular Spectrum*

*Analysis* não é literalmente apropriado, pois não abrange a ideia das múltiplas técnicas incorporadas a abordagem SSA.

### 3.1 Decomposição SSA

O estágio de decomposição pode ser subdividido em duas etapas: incorporação e decomposição em valores singulares (SVD - *Singular Value Decomposition*).

#### 3.1.1 Incorporação

Seja  $Y_T = [y_1, \dots, y_T] \in \mathbb{R}^T$  uma série temporal unidimensional com cardinalidade igual a  $T$ . Por incorporação, entende-se como sendo um procedimento no qual, dado um valor  $L$ , conhecido como *window length*, uma série temporal  $Y_T \in \mathbb{R}^T$  é transformada, através de um mapa  $F: \mathbb{R}^T \rightarrow \mathbb{R}^{L \times K}$ , em uma matriz  $X = [X_1, \dots, X_K]_{L \times K} \in \mathbb{R}^{L \times K}$ , onde  $X_j = [y_j, \dots, y_{j+L-1}]^T \in \mathbb{R}^L, j = 1, \dots, K$ . Isto é,  $Y_T \in \mathbb{R}^T \xrightarrow{F} X \in \mathbb{R}^{L \times K}$ , onde  $K = T - L + 1$  e

$$X = F(Y_T) = [X_1, \dots, X_K]_{L \times K} = (x_{ij})_{i,j=1}^{L,K} = \begin{bmatrix} y_1 & y_2 & y_3 & \cdots & y_K \\ y_2 & y_3 & y_4 & \cdots & y_{K+1} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ y_L & y_{L+1} & y_{L+2} & \cdots & y_T \end{bmatrix}. \quad (20)$$

Por definição  $F$  é um mapa invertível e a matriz  $X$  em (20) é conhecida como **matriz trajetória**. O parâmetro  $L$ , que assume algum valor inteiro no intervalo  $2 \leq L \leq T$ , é o tamanho da janela da matriz trajetória e é o único parâmetro nesta fase. Os vetores  $X_j$ , com  $j = 1, \dots, K$ , colunas de  $X$  em (20), são chamados vetores  $L$ -defasados (ou simplesmente vetores defasados). A matriz  $X$  é uma matriz *Hankel*, ou seja, todos os elementos da diagonal onde  $i + j = \text{constante}$  são iguais.

Em resumo, na fase de Incorporação do método SSA a série  $Y_T$  é transformada por  $F$  em uma matriz  $X$  *Hankel* denominada **matriz trajetória**.

A título de ilustração, considere o exemplo citado em MAHMOUDVAND et al. (2013). Seja a série  $Y_6 = [1, 2, 3, 4, 5, 6]$ ,  $T = 6$ . Suponha  $L = 3$ , daí,  $K =$  número de colunas da matriz  $X = 6 - 3 + 1 = 4$ , e a matriz trajetória é dada por:

$$\mathbf{X} = \begin{bmatrix} 1 & 2 & 3 & 4 \\ 2 & 3 & 4 & 5 \\ 3 & 4 & 5 & 6 \end{bmatrix}. \quad (21)$$

Observa-se que primeira coluna  $[1, 2, 3]$  e a última linha  $[3 \ 4 \ 5 \ 6]$  formam a série original  $Y_6$ .

### 3.1.2 Singular Value Decomposition (SVD)

Nesta etapa, a matriz  $X$  é decomposta em uma soma de matrizes elementares bi-ortogonais de posto 1. Seja  $S$  o produto entre a matriz trajetória e sua transposta,  $S = XX^T$ . De acordo com KUBRUSLY (2001), esta é uma matriz semi-definida positiva, de modo que  $\lambda_l \geq 0$ , para todo  $l$ .

Seja  $V_l := X^T U_l / \sqrt{\lambda_l}$  e considere a ordenação parcial:  $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_L \geq 0$  dos autovalores de  $S$ . Seja  $U_1, U_2, \dots, U_L$  o sistema ortonormal dos autovetores de  $S$  correspondentes àqueles autovalores.  $U_1, \dots, U_L$  formam uma base ortonormal do espaço coluna  $\mathcal{L}_d$  de  $X$  que é chamado **espaço trajetória** (GOLYANDINA, 2010) e a matriz trajetória  $X$  pode ser expandida através da decomposição em valores singulares (SVD), por :

$$X = \sum_{\lambda_l \in \sigma(S)} (\lambda_l)^{\frac{1}{2}} U_l V_l^T = E_1 + E_2 + \dots + E_L, \quad (22)$$

onde:

- $\{E_l\}_{l=1}^L = \lambda_l^{\frac{1}{2}} U_l V_l^T$  são chamadas de matrizes elementares,
- $\lambda_l^{\frac{1}{2}}$ ,  $i = 1, \dots, L$  são os valores singulares e o conjunto  $\{\lambda_l^{\frac{1}{2}}\}_{l=1}^L$  é denominado **espectro singular** de  $X$ .
- $\{U_l\}_{l=1}^L$  são os **vetores singulares à esquerda** da matriz trajetória  $X$ , ou ainda, factor empírico ortogonal function (*EOF*).
- $\{V_l\}_{l=1}^L$  são os **vetores singulares à direita** da matriz trajetória  $X$ , *factor vetor*, ou componente principal.

A coleção  $(\sqrt{\lambda_l}, U_l, V_l)$  é conhecida como  $l$ -ésima autotripla na SVD da matriz trajetória  $X$ . A contribuição de cada componente em (22) pode ser mensurada pela razão de valores singulares, dada por  $(\lambda_l)^{1/2} / \sum_{l=1}^L (\lambda_l)^{1/2}$ .

Seja  $d$  o posto, isto é, o número de autovalores não nulos da matriz trajetória  $X$ .

$$d = \max\{l, \text{tal que } \lambda_l > 0\} = \text{posto}(X) \quad (23)$$

Desta forma, a identidade descrita em (21) pode ser reescrita tal como:

$$X = E_1 + E_2 + \dots + E_d, \quad (24)$$

onde  $d \leq L$ .

Em resumo, a fase de *Singular Value Decomposition* do método SSA consiste em escrever a matriz de trajetória  $X$  como uma soma de  $d$  matrizes elementares ( $d \leq L$ ) obtidas por SVD de  $X$ .

Outra característica da SVD está relacionada com as propriedades das direções determinadas pelo autovetores. Especificamente, o primeiro autovetor determina a direção de tal forma que a variação das projeções dos vetores defasados nessa direção é máxima. Cada autovetor subsequente determina a direção que é ortogonal a todas as direções anteriores e a variação da projeção dos vetores defasados para esta direção também é máxima. Portanto, é natural chamar a direção do  $l$ -ésimo autovetor  $U_l$  a  $l$ -ésima direção principal. Note-se que as matrizes elementares  $E_l$  são construídas a partir das projeções dos vetores defasados nas  $l$ -ésimas direções particulares. Este ponto de vista sobre a SVD da matriz trajetória composto por vetores  $L$  - defasados e um apelo à associação com a análise de componentes principais leva a seguinte terminologia: podemos chamar o vetor  $U_l$  de o autovetor  $l$ , o vetor  $V_l$  será chamado o  $l$ -ésimo vetor fator e o vetor  $Z_l = \sqrt{\lambda_l} V_l$  o  $l$ -ésimo componente principal.

O comportamento dos valores singulares pode ser observado através de um gráfico similar ao gráfico de dispersão em que os seus  $L$  valores são apresentados de forma decrescente de magnitude, exibindo o espectro de  $X$ . O primeiro valor singular, aquele com maior valor absoluto sempre está associado ao componente de tendência. Assim como a tendência de uma série temporal pode ser descrita através de uma função

com variação lenta e estável no decorrer de um período longo de tempo, o comportamento esperado do primeiro fator vetor ou autovetor componente principal  $V_1$  deve apresentar também uma variação lenta e sem oscilações.

A presença de autovalores com valores próximos o suficiente para serem identificados como um par pode ser associado a um componente periódico de sinal. Sendo assim, pode-se identificar pares de valores singulares quase iguais no gráfico de espectro de  $X$ , retratando os componentes periódicos da série temporal. E teoricamente, uma série puramente residual produz uma sequência lentamente decrescente de valores singulares. E se um ruído é adicionado a um sinal, composto por autotriplas com valores singulares altos, então se pode observar uma quebra no espectro de  $X$ . Desta forma podem-se distinguir as autotriplas relacionadas ao sinal daquelas associadas ao ruído.

Para aplicar a *Decomposição em Valores Singulares* na matriz de trajetória  $X$  do exemplo citado, o primeiro passo é o cálculo de  $S$ , e dos autovalores e autovetores de  $S$ . Dada  $X$  obtida em (21),

$$S = \mathbf{X}\mathbf{X}^T = \begin{bmatrix} 1 & 2 & 3 & 4 \\ 2 & 3 & 4 & 5 \\ 3 & 4 & 5 & 6 \end{bmatrix} \begin{bmatrix} 1 & 2 & 3 \\ 2 & 3 & 4 \\ 3 & 4 & 5 \\ 4 & 5 & 6 \end{bmatrix} = \begin{bmatrix} 30 & 40 & 50 \\ 40 & 54 & 68 \\ 50 & 68 & 86 \end{bmatrix}$$

Os Autovalores de  $S$  são dados por  $\lambda_1 \cong 169,29$ ;  $\lambda_2 \cong 0,71$  e  $\lambda_3 \cong 0$ . E os autovetores de  $S$  associados a estes autovalores são:

$$U_1 = \begin{bmatrix} -0,418 \\ -0,565 \\ -0,712 \end{bmatrix}, \quad U_2 = \begin{bmatrix} 0,812 \\ 0,120 \\ -0,572 \end{bmatrix}, \quad \text{e} \quad U_3 = \begin{bmatrix} 0,408 \\ -0,816 \\ 0,408 \end{bmatrix}$$

Considerando que  $\lambda_1 \geq \lambda_2 \geq \lambda_3 \geq 0$  e os resultados dos autovalores, o posto da matriz de trajetória  $\mathbf{X}$  é  $d = 2$ , dado que  $\lambda_3 \cong 0$ . Portanto, a decomposição em valores singulares de  $\mathbf{X}$  pode ser escrita como:

$$\mathbf{X} = \begin{bmatrix} 1 & 2 & 3 & 4 \\ 2 & 3 & 4 & 5 \\ 3 & 4 & 5 & 6 \end{bmatrix} = \mathbf{X}_1 + \mathbf{X}_2$$

Lembrando que cada matriz  $\mathbf{X}_i$  é dada por  $\mathbf{X}_i = \sqrt{\lambda_i}U_iV_i^T$  e  $V_i = \mathbf{X}^T U_i / \sqrt{\lambda_i}$ ,  $\mathbf{X}_1 = \sqrt{\lambda_1}U_1V_1^T$  e  $V_1 = \mathbf{X}^T U_1 / \sqrt{\lambda_1}$ ;  $\mathbf{X}_2 = \sqrt{\lambda_2}U_2V_2^T$  e  $V_2 = \mathbf{X}^T U_2 / \sqrt{\lambda_2}$ . Para o exemplo em questão, fazendo o Cálculo de  $V_i$  obtém-se como resultados:

$$V_1 = \frac{\begin{bmatrix} 1 & 2 & 3 & 4 \\ 2 & 3 & 4 & 5 \\ 3 & 4 & 5 & 6 \end{bmatrix} \begin{bmatrix} -0,418 \\ -0,565 \\ -0,712 \end{bmatrix}}{\sqrt{169,29}} = \begin{bmatrix} -0,283 \\ -0,413 \\ -0,543 \\ -0,674 \end{bmatrix}$$

$$V_2 = \frac{\begin{bmatrix} 1 & 2 & 3 & 4 \\ 2 & 3 & 4 & 5 \\ 3 & 4 & 5 & 6 \end{bmatrix} \begin{bmatrix} 0,812 \\ 0,120 \\ -0,572 \end{bmatrix}}{\sqrt{0,71}} = \begin{bmatrix} -0,788 \\ -0,361 \\ 0,066 \\ 0,494 \end{bmatrix}$$

$$\begin{aligned} \mathbf{X}_1 &= \sqrt{169,29} \begin{bmatrix} -0,418 \\ -0,565 \\ -0,712 \end{bmatrix} \begin{bmatrix} -0,283 & -0,413 & -0,543 & -0,674 \end{bmatrix} \\ &= \begin{bmatrix} 1,539 & 2,247 & 2,953 & 3,661 \\ 2,079 & 3,036 & 3,993 & 4,950 \\ 2,621 & 3,827 & 5,033 & 6,239 \end{bmatrix} \end{aligned}$$

$$\begin{aligned} \mathbf{X}_2 &= \sqrt{0,71} \begin{bmatrix} 0,812 \\ 0,120 \\ -0,572 \end{bmatrix} \begin{bmatrix} -0,788 & -0,361 & 0,066 & 0,494 \end{bmatrix} \\ &= \begin{bmatrix} -0,539 & -0,247 & 0,047 & 0,339 \\ -0,079 & -0,036 & 0,007 & 0,050 \\ 0,379 & 0,173 & -0,033 & -0,239 \end{bmatrix} \end{aligned}$$

E pode-se verificar que

$$\mathbf{X} = \mathbf{X}_1 + \mathbf{X}_2$$

$$\begin{aligned} \mathbf{X} &= \begin{bmatrix} 1,539 & 2,247 & 2,953 & 3,661 \\ 2,079 & 3,036 & 3,993 & 4,950 \\ 2,621 & 3,827 & 5,033 & 6,239 \end{bmatrix} + \begin{bmatrix} -0,539 & -0,247 & 0,047 & 0,339 \\ -0,079 & -0,036 & 0,007 & 0,050 \\ 0,379 & 0,173 & -0,033 & -0,239 \end{bmatrix} \\ &= \begin{bmatrix} 1 & 2 & 3 & 4 \\ 2 & 3 & 4 & 5 \\ 3 & 4 & 5 & 6 \end{bmatrix} \end{aligned} \quad (25)$$

## 3.2 Reconstrução SSA

O estágio de reconstrução tem por objetivo reconstruir uma série temporal menos ruidosa a partir da decomposição da matriz  $X$  feita na segunda fase, SVD. A reconstrução é subdividida em duas etapas: média diagonal e agrupamento. Na etapa de média diagonal cada matriz da decomposição de  $X$  é transformada em uma série temporal por média diagonal, e por agrupamento, as componentes da série são separadas através de algum processo identificador destas componentes e a série aproximada pode ser reconstruída retirando as componentes de ruído identificadas.

### 3.2.1 Agrupamento

A etapa de agrupamento corresponde a dividir as matrizes elementares em vários grupos e somar as matrizes dentro de cada grupo. Este processo consiste em, dada a decomposição em (24), considerar a sequência  $\{E_l\}_{l=1}^d$  de matrizes elementares nesta decomposição. Agrupe-as em  $m \leq d$  grupos *disjuntos* utilizando algum método de agrupamento (por exemplo, com o auxílio da análise de componentes principais, análise de agrupamentos por clusterização hierárquica integrada com análise de componentes principais ou análise gráfica de vetores singulares) Assumindo que, após o agrupamento, o conjunto de índices gerado é dado por  $\{I_1, \dots, I_m\}$ , onde, para todo  $i$ ,  $I_i = \{I_{i1}, \dots, I_{ip_i}\}$  e  $p_i$  é a cardinalidade do grupo  $I_i$ . Note que  $\{E_l\}_{l=1}^d = \cup_{i=1}^m \{E_{I_{ij}}\}_{j=1}^{p_i}$ , onde  $m \leq d$ . A matriz elementar  $E_{I_i}$  gerada a partir do grupo  $\{E_{I_{ij}}\}_{j=1}^{p_i}$  é dada por  $E_{I_i} = \sum_{j=1}^{p_i} E_{I_{ij}}$ , de modo que a identidade em (23) pode ser reescrita como

$$X = \sum_{i=1}^m E_{I_i}. \quad (26)$$

O procedimento de agrupamento pode ser realizado também sob a sequência  $\cup_{i=1}^m \{E_{I_{ij}}\}_{j=1}^{p_i}$  de igual forma ao realizado na sequência  $\{E_l\}_{l=1}^d$  utilizando, inclusive, um método ou critério de agrupamento diferente do aplicado sob a sequência  $\{E_l\}_{l=1}^d$ . O objetivo do agrupamento é diminuir o número de componentes (ou matrizes

elementares) na SVD da matriz trajetória  $X$ . A contribuição da componente  $E_{I_i}$  pode ser mensurada pela razão de valores singulares dada por  $\sum_{j=1}^{p_i} (\lambda_{I_{ij}})^{1/2} / \sum_{l=1}^d (\lambda_l)^{1/2}$ .

Ao aplicar a fase do agrupamento das autotriplas para o exemplo da série temporal  $Y_6 = [1, 2, 3, 4, 5, 6]$ , note-se que na fase de SVD, obtiveram-se apenas dois componentes diferentes de zero e; portanto, nesta fase podem-se definir no máximo dois grupos:  $I_1 = \{1\}$ ,  $I_2 = \{2\}$ .

### 3.2.2 Média Diagonal

O objetivo da média diagonal é transformar uma matriz na forma de uma matriz Hankel em uma série temporal que é um componente aditivo da série original. Se  $z_{ij}$  representa um elemento de uma matriz  $Z$ , então, o  $k$ -ésimo termo da série resultante é obtido por uma média de  $z_{ij}$  sobre todos  $i, j$  tais que  $i + j = k + 1$ . Este procedimento é chamado média diagonal, ou Hankelização da matriz  $Z$ . O resultado da Hankelização de uma matriz  $Z$  é a matriz Hankel  $\mathcal{H}Z$ .

Ao aplicar o processo de Hankelização a cada elemento  $E_{I_i}$ ,  $i = 1, \dots, m$ , em (24), obtém-se uma nova expansão:

$$X \cong \tilde{E}_{I_1} + \tilde{E}_{I_2} + \dots + \tilde{E}_{I_m}. \quad (27)$$

onde  $\tilde{E}_{I_i} = \mathcal{H}E_{I_i}$ . A cada matriz desta expansão o processo de Hankelização é aplicado através da média diagonal, apresentado a seguir, e que equivale à aplicação bijetora inversa  $F^{-1}$  do mapa invertível  $F$  sobre a matriz trajetória  $X$ .

Considere a matriz trajetória  $X$  e assuma que  $L^* = \min(L, K)$  e que  $K^* = \max(L, K)$ . Considere também que  $x_{l,k}^{(i)}$  seja um elemento na linha  $l$  e coluna  $k$  na matriz  $E_{I_i}$ . O elemento  $y_t^{(i)}$  da componente SSA  $[y_t^{(i)}]_{1 \times T}$  é calculado por meio da média diagonal, definida em (28), a partir da matriz elementar  $E_{I_i}$ .

$$\tilde{y}_t^{(i)} = \begin{cases} \frac{\sum_{l=1}^t x_{l,t-l+1}^{(i)}}{t}, & \text{se } 1 \leq t < L^* \\ \frac{\sum_{l=1}^{L^*} x_{l,t-l+1}^{(i)}}{L^*}, & \text{se } L^* \leq t < K^* \\ \frac{\sum_{l=t-K^*+1}^{T-K^*+1} x_{l,t-l+1}^{(i)}}{T-K^*+1}, & \text{se } K^* \leq t \leq T \end{cases} \quad (28)$$

Este processo equivale à decomposição da série original  $Y_T = [y_1, \dots, y_T]$  em uma soma de  $m$  séries.

$$y_t = \sum_{i=1}^m \tilde{y}_t^{(i)}, \quad (29)$$

onde  $\tilde{E}_{I_i} = (\tilde{y}_1^{(i)}, \tilde{y}_2^{(i)}, \dots, \tilde{y}_T^{(i)})$ ,  $i = 1, \dots, m$ . A série aproximada será  $\tilde{Y}_T = (\tilde{y}_1, \tilde{y}_2, \dots, \tilde{y}_T)$ , onde  $y_t = \sum_{l=1}^r \tilde{y}_t^{(l)}$ ,  $t = 1, \dots, T$  e  $r < m$ .

Observa-se que em termos de  $F$ , o processo de média diagonal é feito através da sequência:

$$\begin{aligned} F^{-1}(X) &= F^{-1}\left(\sum_{i=1}^m E_{I_i}\right) = F^{-1}\left(\sum_{i=1}^m \left(\sum_{j=1}^{p_i} E_{I_{ij}}\right)\right) = \sum_{i=1}^m F^{-1}\left(\sum_{j=1}^{p_i} E_{I_{ij}}\right) \Rightarrow \\ &\Rightarrow \sum_{i=1}^m \left(\sum_{j=1}^{p_i} F^{-1}(E_{I_{ij}})\right) = \sum_{i=1}^m \left(\sum_{j=1}^{p_i} [y_t^{(ij)}]_{1 \times T}\right) = \sum_{i=1}^m [y_t^{(i)}]_{1 \times T} \quad (30) \\ &= \sum_{i=1}^m [y_t^{(i)}]_{1 \times T} = [y_t]_{1 \times T}. \end{aligned}$$

Cada componente  $[y_t^{(i)}]_{1 \times T}$  concentra parte da energia da série temporal original  $[y_t]_{1 \times T}$  que pode ser mensurada pela razão de valores singulares  $\sum_{j=1}^{p_i} (\lambda_{I_{ij}})^{1/2} / \sum_{l=1}^d (\lambda_l)^{1/2}$ .

De acordo com GOLYANDINA et al. (2001), as componentes SSA  $[y_t^{(i)}]_{1 \times T}$  podem ser classificadas em três categorias: *tendência*, *harmônica* e *ruído*.

Com o intuito de exemplificar esta última fase, foram calculadas as médias dos elementos diagonais da seguinte forma:

$$\mathbf{X}_1 = \begin{bmatrix} 1,539 & 2,247 & 2,953 & 3,661 \\ 2,079 & 3,036 & 3,993 & 4,950 \\ 2,621 & 3,827 & 5,033 & 6,239 \end{bmatrix}$$

Como  $L = 3$ ,  $K = 4$  e adicionalmente,  $L^* = \min(L, K) = \min(3, 4) = 3$  e  $K^* = \max(L, K) = \max(3, 4) = 4$ .

- Para  $k = 1$  escolhe-se a primeira expressão da Equação (28), já que está cumprida a restrição  $1 \leq k < L^*$ .

$$y_1 = \frac{x_{1,1}}{1} = x_{1,1} = 1,538$$

- Para  $k = 2$  escolhe-se a primeira expressão da Equação (28), já que está cumprida a restrição  $1 \leq k < L^*$ .

$$y_2 = \frac{(x_{1,2} + x_{2,1})}{2} = 2,162$$

- Para  $k = 3$  escolhe-se a segunda expressão da Equação (28), já que está cumprida a restrição  $L^* \leq k \leq K^*$ .

$$y_3 = \frac{(x_{1,3} + x_{2,2} + x_{3,1})}{3} = 2,870$$

- Para  $k = 4$  escolhe-se a segunda expressão da Equação (28), já que está cumprida a restrição  $L^* \leq k \leq K^*$ .

$$y_4 = \frac{(x_{1,4} + x_{2,3} + x_{3,2} + x_{4,1})}{3} = 3,826$$

- Para  $k = 5$  escolhe-se a terceira expressão da Equação (28), já que está cumprida a restrição  $K^* < k \leq T$ .

$$y_5 = \frac{(x_{2,4} + x_{3,2})}{2} = 4,991$$

- Para  $k = 6$  escolhe-se a terceira expressão da Equação (28), já que está cumprida a restrição  $K^* < k \leq T$ .

$$y_6 = \frac{x_{3,4}}{1} = 6,239$$

Logo

$$\tilde{\mathbf{X}}_1 = \begin{bmatrix} 1,538 \\ 2,162 \\ 2,870 \\ 3,827 \\ 4,991 \\ 6,239 \end{bmatrix} \quad (31)$$

Similarmente é feito o cálculo da média diagonal para  $\mathbf{X}_2$ :

$$\mathbf{X}_2 = \begin{bmatrix} -0,538 & -0,246 & 0,047 & 0,339 \\ -0,080 & -0,036 & 0,007 & 0,050 \\ 0,379 & 0,173 & -0,033 & -0,239 \end{bmatrix}$$

Obtendo como resultado:

$$\tilde{\mathbf{X}}_2 = \begin{bmatrix} -0,538 \\ -0,162 \\ 0,130 \\ 0,173 \\ 0,009 \\ -0,239 \end{bmatrix} \quad (32)$$

Desta forma, a série reconstruída é a seguinte:

$$\mathbf{Y} = \tilde{\mathbf{X}}_1 + \tilde{\mathbf{X}}_2 = \begin{bmatrix} 1,538 \\ 2,162 \\ 2,870 \\ 3,827 \\ 4,991 \\ 6,239 \end{bmatrix} + \begin{bmatrix} -0,538 \\ -0,162 \\ 0,130 \\ 0,173 \\ 0,009 \\ -0,239 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \\ 3 \\ 4 \\ 5 \\ 6 \end{bmatrix}$$

### 3.3 Métodos de Agrupamentos

Um elemento muito importante no estágio de reconstrução SSA é o método utilizado para se fazer o agrupamento. Resultados distintos podem acontecer dependendo do método usado. Dentre os diversos métodos de agrupamento, destacam-se alguns que vem sendo utilizados em trabalhos científicos tratando este assunto.

Sem perda de generalidades, para alguns metodos de agrupamentos é mais conveniente primeiramente fazer a média diagonal das  $d$  matrizes resultantes da SVD, e obter  $d$  series componentes da série temporal, e depois fazer agrupamento destas  $d$

séries temporais. Ou seja, pode-se, apropriadamente, inverter a ordem dos procedimentos na fase de reconstrução SSA.

### 3.3.1 Agrupamento por Análise de Componentes Principais

Considere cada valor singular  $\lambda_l$  na SVD, em ordem decrescente de magnitude, que é expressa pela contribuição na SVD através da razão dos valores singulares  $\sum_{j=1}^{p_i} (\lambda_{l_{ij}})^{1/2} / \sum_{l=1}^d (\lambda_l)^{1/2}$ . De acordo VASCONCELLOS (2011), através da análise de componentes principais (ACP), pode-se escolher o valor ótimo  $R$  no conjunto de índices  $\{1, \dots, m\}$  de modo que apenas as  $R$  primeiras autotriplas são consideradas na expansão (24). Assim, a série temporal aproximada filtrada SSA através de ACP é dada por:

$$[\tilde{y}_t]_{1 \times T} = \sum_{i=1}^R [y_t^{(i)}]_{1 \times T} \quad (33)$$

Neste caso, a série aproximada gerada é menos ruidosa e SSA não trabalha com remoção de componentes com tendência ou sazonalidades.

### 3.3.2 Agrupamento por Análise Gráfica dos Vetores Singulares

A análise das coordenadas da série temporal na base definida pelos vetores singulares resultantes da SVD permite identificar as componentes de tendência e da sazonalidade da série. O problema geral aqui consiste em identificar e separar as componentes oscilatórias das componentes que fazem parte da tendência. De acordo com GOLYANDINA et al. (2001) a análise gráfica de tais coordenadas aos pares permite identificar por meio visual as componentes harmônicas da série.

Considere um harmônico puro com frequência igual a  $\omega$ , fase igual a  $\delta$ , amplitude igual a  $\xi$  e período  $\rho = \frac{1}{\omega}$  definido como um divisor do tamanho da janela  $L$  e  $K$ . Se o

parâmetro  $\rho$  assume um valor inteiro, então  $\rho$  é classificado como *período do harmônico* (MORETTIN & TOLOI, 2006).

As funções seno e o cosseno com frequências, amplitudes e fases iguais resultam em um diagrama de dispersão que exhibe um padrão circular (MORETTIN & TOLOI, 2006). Por sua vez, se  $\rho = \frac{1}{\omega}$  é um inteiro, então o diagrama de dispersão exhibe um polígono regular com  $\rho$  vértices. Para uma frequência  $\omega = m/n < 0.5$  com  $m$  e  $n$  inteiros e primos, os pontos são vértices de um polígono regular de  $n$  vértices (GOLYANDINA et al., 2001). As coordenadas da série temporal em duas componentes ortogonais podem ser dispostas em um diagrama de dispersão. Dessa forma, a identificação dos componentes que são gerados por um harmônico é reduzida à análise pictórica do padrão determinado nos diferentes pares de componentes.

Em CASSIANO et al. (2013) há uma aplicação da filtragem SSA aplicada à série de vazão que compara os métodos de agrupamento citados e os resultados mostram que o poder preditivo melhora quando se utiliza a análise gráfica dos vetores singulares.

### 3.3.2.1 Análise de Periodograma

A análise de periodograma também auxilia na identificação de uma componente harmônica geral. De acordo com MIRANIAN et al. (2013), o periodograma do vetor singular de cada autotripla fornece informações sobre o comportamento periódico do componente e frequência (período) das oscilações. Portanto, o agrupamento adequado pode ser feito com a ajuda de análise de periodograma.

Para a série  $Y_T = (y_1, \dots, y_T)$ , o periodograma  $\Pi_{Y_T}^T(\omega)$  é definido por:

$$\Pi_{Y_T}^T(\omega) = \frac{1}{T} \left| \sum_{t=0}^{T-1} e^{-2\pi i \omega t} y_{t+1} \right|, \quad \omega = (-0.5, 0.5] \quad (34)$$

onde  $i = \sqrt{-1}$ .

Para as componentes periódicas, o periodograma tem picos em torno de frequência do componente (período). Por isso, a identificação visual é simples.

### 3.3.3 Agrupamento por Clusterização

A Clusterização visa agrupar um conjunto com  $N$  objetos em  $K$  clusters mutuamente excludentes, de tal forma que os objetos em um mesmo cluster apresentem similaridades entre si e dissimilaridades em relação aos objetos pertencentes aos outros clusters. Sendo assim pode-se tratar os vetores singulares resultantes na SVD ou as séries resultantes da fase média Diagonal, como um conjunto a ser clusterizado em grupos disjuntos de características específicas. Logo, métodos de análise de clusters podem ser utilizados na classificação dos vetores singulares na SVD.

### 3.4 Separabilidade

As etapas de SVD e Agrupamento apoiam-se na propriedade denominada separabilidade, que caracteriza o quão bem os diferentes componentes podem ser separados uns dos outros em termos de dependência linear. A decomposição e reconstrução da série original terá êxito se os seus componentes aditivos forem separáveis um do outro. A ausência de separabilidade aproximada é frequentemente observada em série com estrutura complexa. Para estas séries e séries com estrutura especial, há diferentes maneiras de modificar SSA levando a diferentes versões, como SSA com centragem individuais e duplos, Toeplitz SSA e SSA sequencial (ver ELSNER & TSONIS, 1996, secção 1.7).

Conforme declararam GOLYANDINA et al. (2001), na prática não ocorre separabilidade exata, fala-se somente em separabilidade aproximada. Comumente, fala-se somente de separabilidade através de uma medida natural de dependência entre as subséries. Essa medida é conhecida como correlação ponderada ou  $w$  – correlação, que é uma função que quantifica a dependência linear entre duas componentes SSA  $Y_t^{(i)}$  e  $Y_t^{(j)}$ , definido por

$$\rho_{ij}^{(w)} = \frac{(Y_t^{(i)}, Y_t^{(j)})_w}{\|Y_t^{(i)}\|_w \|Y_t^{(j)}\|_w}, \quad (35)$$

onde  $\|\cdot\|_w$  é a norma euclidiana,  $(\cdot)_w$  é o produto interno tal que:

$$\|Y_t^{(i)}\|_w = \sqrt{(Y_t^{(i)}, Y_t^{(i)})_w} \text{ é a norma da } i\text{-ésima subsérie;}$$

$$(Y_t^{(i)}, Y_t^{(j)})_w = \sum_{k=1}^T w_k y_k^{(i)} y_k^{(j)} \text{ é o produto interno entre um par de subséries;}$$

e  $w_k = \min\{k, L, T - k\}$  são os ponderadores.

Se o valor absoluto da correlação ponderada é pequeno então as componentes SSA correspondentes são classificadas como  $w$ -ortogonais (ou quase  $w$ -ortogonais) e tem separabilidade aproximada. Caso contrário, são ditas mal separadas. Em outras palavras, os componentes altamente correlacionados se tornam pertencentes a um mesmo grupo de autotriplas. Por isso, comumente a correlação ponderada é usada também na fase de agrupamento para geração de novas matrizes elementares na SVD (GOLYANDINA et al., 2001).

Uma ferramenta complementar utilizada para o processo de agrupamento é a análise gráfica da “matriz”  $w$ -correlação. A matriz  $w$ -correlação indica as correlações entre as componentes da SVD através de uma escala de tons de cinza, variando do branco ao preto, correspondendo aos valores absolutos de 0 a 1. A forma da matriz correlação fornece um indicativo de como realizar um agrupamento adequado. Espera-se uma divisão clara dos componentes em duas partes. A primeira parte é constituída por componentes altamente correlacionados (tonalidade mais escura), o que caracteriza o grupo de autotriplas relacionado ao sinal de série. Por outro lado, a segunda parte retrata o ruído ao exibir muitas componentes com correlações baixas (tonalidades claras).

### 3.5 Escolha do parâmetro $L$ de defasagem

Desde o início dos estudos em SSA, a determinação de um valor para o parâmetro  $L$  de defasagem vem sendo alvo de discussão em diversos trabalhos. Em GOLYANDINA (2010) há uma longa discussão acerca do comprimento ideal da janela

de defasagem partindo do princípio que este valor pode ser fixado ou variável, dependendo de seu comportamento assintótico quando  $T \rightarrow \infty$ .

Valores entre 2 e  $T/2$  são suficientes para a realização da decomposição da série. Contudo, naturalmente a decomposição da série emporal torna-se mais detalhada à medida que o  $L$  cresce, pois  $L$  suficientemente grande faz com que cada vetor  $L$  – *defasado* incorpore essencialmente o comportamento da série.

Entretanto, nos caso onde  $T$  é muito grande, usar  $T/2$  pode demandar um custo operacional que pode comprometer a análise SSA, muito embora haja recursos computacionais avançados recentes como os mencionados em KOROBEYNIKOV (2010), que podem fazer cálculos muito rápidos. Observando o comportamento assintótico do comprimento da janela (*window length*) a considerar a matriz  $L \times K$ , com  $K \rightarrow \infty$ , a escolha de um número fixo de linhas  $L$  menor que  $T/2$  tem um custo operacional menor (GOLYANDINA, 2010).

GOLYANDINA (2010) sugere que ao trabalhar com séries periódicas, para alcançar separabilidade suficiente dos componentes, utilize-se um comprimento  $L$  proporcional ao período de sazonalidade dos dados. A autora também verificou que no processo SSA, no qual se pode separar bem o sinal  $S_T$  do ruído  $R_T$ , se o sinal for do tipo

$$S_T = a^T \cos\left(\frac{\pi T}{5}\right), \quad (36)$$

a escolha de um nível de ruído é tal que se tem alguma razão sinal-ruído (SNR – *signal-to-noise ratio*) dada por:  $SNR = \frac{\sum_{t=1}^T S_t^2}{\sum_{t=1}^T R_t^2}$  ou  $SNR = \frac{\sum_{t=1}^T S_t^2}{\sigma_{R_T}^2}$  se os resíduos são aleatórios.

SNR pode ser usado para comparar a qualidade do processamento de séries de igual tamanho. Entretanto, não pode-se afirmar que SSA separa sinal de ruído apenas se SNR for maior que um determinado valor, por exemplo, para qualquer SNR pequeno, um sinal de onda senoidal é assintoticamente separável de um ruído branco com  $T \rightarrow \infty$  e  $L \sim \beta T$  com  $0 < \beta < 1$ . Este mesmo artigo mostra, através de experimentos, que para séries aplicadas com sinal como em (32), a perturbação do sinal cresce quando  $L$  cresce e decresce quando  $K = T - L + 1$  cresce. Então os erros resultantes são causados por estas tendências contraditórias. Mostra também que o crescimento da razão dos erros é maior para  $L > N/2$ . No entanto, não é surpresa que os erros reconstruídos são grandes

para comprimento de janela  $L$  pequeno. Com isso, o artigo mantém a premissa que a questão sobre o comprimento da janela permanece em aberto.

Segundo GOLYANDINA & VLASSIEVA (2009), ao tentar resolver o problema de minimizar a raiz quadrada da soma dos quadrados dos erros (*RMSE – Root Mean Square Error*) do estimador de  $S_l$  em um ponto fixo  $l$ , o comprimento da janela varia entre  $T/3$  e  $T/2$ . Isto significa que sempre que se está com o caso com o sinal constante, o comprimento da janela, que minimiza os erros de reconstrução da série como um todo, depende do peso de cada ponto na série temporal. Em todo caso, a recomendação é a de se escolher o comprimento da janela um pouco menor que a metade do comprimento da série  $T$ . Assim, escolher o comprimento da janela usando  $T/3 \leq L \leq T/2$  faz sentido.

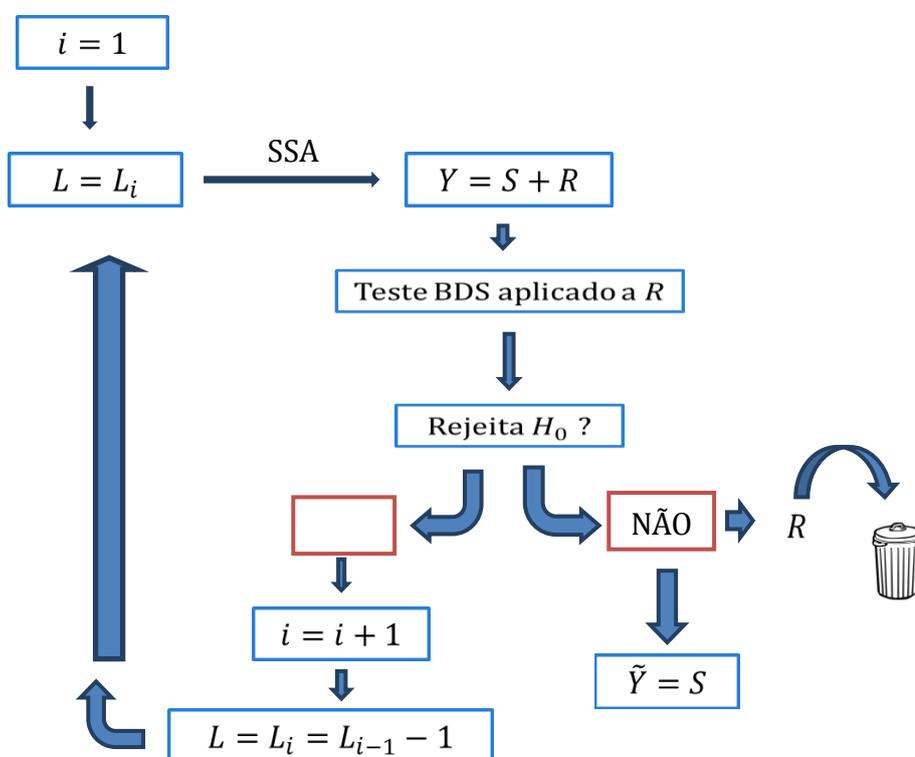
Mesmo com todas as especulações e programas que tentem simular um valor para o comprimento da janela, os autores indicam que a escolha deste parâmetro sempre vai depender da natureza dos dados. Portanto, sempre haverá um intervalo de confiança em que o valor de  $L$  deverá se adequar.

Considerando a necessidade de se obter este valor de  $L$  e sabendo destas premissas anteriores sobre o tema, neste trabalho é usada uma metodologia baseada em técnicas de reamostragem *Bootstrap* (EFRON & TIBSHIRANI, 1993), proposta em MENEZES (2014) para se obter uma estimativa de  $L$ .

Neste método, valores decrescentes para  $L$  (a partir de  $(T + 1)/2$  se  $T$  for ímpar ou a partir de  $T/2$  se  $T$  for par) são testados até se encontrar um valor  $L$  tal que a série de resíduos obtida na fase de agrupamento SSA, quando submetida ao teste BDS, estabelecido em BROCK et al. (1987), não rejeite a hipótese de independência dos dados ao nível de 5% de significância.

Mais detalhadamente, o procedimento inicia-se com a escolha do parâmetro  $L$  inicial: se  $T$  é par, então  $L$  inicial é igual à  $T/2$  e se  $T$  é ímpar, o  $L$  inicial será  $(T + 1)/2$ . A partir desta definição, procede-se com a filtragem SSA com o objetivo de reduzir o ruído das séries. Do procedimento SSA, obtém-se para cada série envolvida, uma componente de sinal ( $S$ ) composta pelas componentes de tendência e harmônica e uma

componente de ruído ( $R$ ). Testa-se a componente  $R$  via BDS e, se o teste não rejeitar a hipótese nula de independência, então a componente  $R$  é classificada como ruído e é removida da série para que a mesma possa ser modelada. Caso o teste BDS rejeite a hipótese nula, então um novo valor de  $L$  é estimado subtraindo o valor anterior de uma unidade:  $L_i = L_{i-1} - 1$ , onde  $i = 2, \dots$  indica o número de tentativas de estimar este parâmetro. Este procedimento se repete tantas vezes quantas forem necessárias até que o teste BDS não rejeite a hipótese nula. A figura 3.1 ilustra o procedimento proposto. Contudo, tem-se verificado que o valor ótimo de  $L$  coincide com o seu tamanho máximo  $T/2$ .



**Figura 3.1.** Estimação de  $L$  para filtragem SSA da série.

### 3.6 Previsão SSA

O método de previsão SSA descrito a seguir é desenvolvido para gerar previsões de séries que satisfazem aproximadamente a fórmula recorrente linear (LRF – Linear Recurrente Formulae). Diz-se que essa série satisfaz uma LRF de ordem ou dimensão  $L - 1$  se existirem coeficientes  $a_1, \dots, a_{L-1}$ , tais que

$$y_{T-1} = \sum_{k=1}^{L-1} a_k y_{T-i-k}, \quad 0 \leq i \leq T - L. \quad (37)$$

Para verificar se uma série temporal satisfaz esta condição, os coeficientes  $a_1, \dots, a_{L-1}$  podem ser encontrados através dos autovetores obtidos a partir do passo SVD, estes e mais detalhes podem ser obtidos em GOLYANDINA et al. (2001). Esta não é uma condição restritiva, pois a classe de séries temporais que satisfazem alguma LFR é bastante ampla incluindo séries polinomiais, harmônicas e exponenciais.

Seja  $Y_t = (y_1, y_2, \dots, y_T)$  a série original e  $\tilde{Y}_t = (\tilde{y}_1, \tilde{y}_2, \dots, \tilde{y}_T)$  a série reconstituída por SSA. Para um autovetor  $U \in \mathbb{R}^L$  denota-se o vetor das primeiras  $L - 1$  componentes do vetor  $U$  como  $U^\nabla \in \mathbb{R}^{L-1}$ . Seja  $v^2 = \pi_1^2 + \pi_2^2 + \dots + \pi_r^2 < 1$  onde  $\pi_i$  é o último componente do autovetor  $U_i$  ( $i = 1, \dots, r$ ). Pode ser provado que o último componente  $y_L$  de algum vetor  $Y = (y_1, y_2, \dots, y_L)^T$  é uma combinação linear dos primeiros componentes  $(y_1, y_2, \dots, y_{L-1})$ , ou seja,  $y_L = a_1 y_1 + a_2 y_2 + \dots + a_{L-1} y_{L-1}$ , onde o vetor de coeficientes pode ser expressado de forma vetorial por  $A = (a_1, a_2, \dots, a_{L-1})^T = \frac{1}{1-v^2} \sum_{i=1}^r \pi_i U_i^\nabla$ . Seguindo MAHMOUDVAND et al. (2013), assume-se que o vetor  $U_j^\nabla$  denota as primeiras  $L - 1$  componentes do autovetor  $U_j$  e que  $\pi_j$  é a última componente de  $U_j$  ( $j = 1, \dots, r$ ). As previsões  $\hat{y}_{T+1}, \hat{y}_{T+2}, \dots, \hat{y}_{T+h}$  são obtidas pela fórmula recorrente

$$\hat{y}_i = \begin{cases} \tilde{y}_i, & \text{para } i = 1, \dots, T \\ \sum_{j=1}^{L-1} a_j \hat{y}_{i-j} & \text{para } i = T + 1, T + 2, \dots, T + h \end{cases} \quad (38)$$

Vale salientar que quanto maior o número de passos à frente ( $h$ ) as previsões se tornam mais dependentes das predições anteriores. Para que a predição considere pelo menos um valor aproximado  $\tilde{Y}_j$  da série original, recomenda-se utilizar um horizonte de no máximo  $M = L - 1$ .

Exemplificando este estágio, para a série  $Y_6$  considerada, tem-se que o vetor  $U_j^V$  contém os  $L - l$  elementos do vetor  $U_j$ , com  $j = 1, \dots, r$ . Considerando-se  $r = 1$  e  $M = 1$ , tem-se  $U_1^V = \begin{bmatrix} -0,418 \\ -0,565 \end{bmatrix}$ .

O valor de  $\pi_1$  corresponde ao último elemento do vetor  $U_j$  é  $\pi_1 = -0,712$ .

O valor de  $v^2$  é dado por  $v^2 = \sum_{j=1}^r \pi_j^2 = \pi_1^2 = (-0,712)^2 = 0,5069$

O vetor  $R$  é obtido por:

$$\begin{aligned} R &= \frac{1}{1 - v^2} \sum_{j=1}^r \pi_j U_j^V = \frac{1}{1 - 0,5069} (\pi_1 U_1^V) = 2,028 \left( -0,712 \begin{bmatrix} -0,418 \\ -0,565 \end{bmatrix} \right) \\ &= 2,028 \begin{bmatrix} 0,298 \\ 0,402 \end{bmatrix} = \begin{bmatrix} 0,604 \\ 0,815 \end{bmatrix} \end{aligned}$$

Para efetuar a previsão é utilizada Equação (38). Para obtermos a previsão  $\hat{y}_7$ ,

$$Q_7 = [\hat{y}_{7-3+1}, \dots, \hat{y}_{7-1}]^T = [\hat{y}_5, \hat{y}_6]^T = [4,991 \quad 6,239]^T$$

$\hat{y}_5$  e  $\hat{y}_6$  correspondem à posição 5 e 6 respectivamente da série  $\tilde{X}_1$ , exemplificada na parte final da fase 2 do estágio 2 de agrupamento. Desta forma  $R$  e  $Q_7$  produzem a primeira previsão um passo à frente para  $\hat{y}_7 = 0,815\hat{y}_6 + 0,604\hat{y}_5$ :

$$\hat{y}_7 = R^T Q_7 = [0,604 \quad 0,815] \begin{bmatrix} 4,991 \\ 6,239 \end{bmatrix} = 3,01 + 5,08 = 8,09 \approx 8$$

E assim por diante, os  $M$  valores podem ser previstos. Como neste exemplo  $M = 1$ , a previsão finaliza com o cálculo de  $\hat{y}_7$

O algoritmo recorrente de previsão SSA é encontrado mais detalhadamente em GOLYANDINA et al. (2001) e pode resultar em acuradas previsões se razoáveis análises forem feitas nas autotriplas geradas e a série for reconstruída por uma eficiente técnica de agrupamento que identifique bem as componentes de ruído. Este algoritmo está implementado no software Caterpillar.

Para avaliar a precisão e confiabilidade das previsões, pode-se usar uma adequada combinação das três abordagens: (a) construção de intervalos de confiança; (b) avaliação das previsões retrospectivas e (c) verificação da estabilidade das previsões. Para mais informações sobre a abordagem (a), nos referimos a GOLYANDINA et al. (2001). Para abordagem (c), PEPELYSHEV & ZHIGLJAVSKY (2010) formalizaram o problema de avaliar a estabilidade de longo horizonte de previsão de uma determinada família de técnicas de previsão e argumentou que SSA poderia ser uma família adequada de técnicas a serem consideradas em algumas aplicações.

Previsões retrospectivas, abordagem (b), são realizadas por truncar a série e os valores de previsão nos pontos removido temporariamente. Estas previsões podem então ser comparados com os valores observados da série temporal para fazer uma avaliação da qualidade das previsões. Seja  $e_{T,h}(x) = y_{T+h}(x) - \hat{y}_{T,h}(x)$  denotam o erro de previsão, onde  $\hat{y}_{T,h}(x)$  são as previsões de  $y_{T+h}(x)$  usando SSA ( $h = 1, \dots, M$ ). Então, define-se o Erro de Previsão Quadrado Integrado como uma medida de acurácia dada por:

$$ISFE_{T,h} = \sum_x e_{T,h}^2(x). \quad (39)$$

### 3.7 SSA na Análise de Séries Temporais

SSA é uma técnica relativamente recente. A partir da publicação de “*Singular Spectrum Analysis: A New Tool in Time Series Analysis*” (ELSNER & TSONIS, 1996) as aplicações nesta área começaram a se intensificar em grupos de pesquisa formados principalmente nos Estados Unidos, Reino Unido e Rússia.

Aplicações de SSA nas áreas de Meteorologia e Geofísica podem ser vistas em ALLEN & SMITH (1996); GHIL & TARICCO (1997), DANILOV & ZHIGLJAVSKY (1997); YIOU et al. (2000), GOLYANDINA et al. (2001) e HASSANI (2007). Na área Farmaceutica SSA foi usado por BAPTISTA (2008).

Vários estudos envolvendo SSA e MSSA (SSA para dados multivariados) foram feitos com aplicações a economia como em HASSANI (2009), HASSANI

(2010), HASSANI et al., (2013 a, b,c), CASSIANO et al., (2013), MENEZES et al. (2011), MENEZES et al. (2012), BENEKI et al. (2011) e BENEKI & LEON (2012).

BENEKI et al. (2011) aplicam SSA a uma série trimestral de PIB dos EUA ajustada sazonalmente. Os resultados indicam que a etapa de decomposição deixa uma quantidade significativa de informação dentro do componente de ruído. No entanto, é realizado SSA sobre o componente de ruído extraído, que é capaz de extrair a informação desejada. Isso também ajuda a ter um melhor processo de extração de sinal.

Em ZOKAEI & MAHMOUDVAND (2011) um novo critério para medir a separabilidade forte com base no coeficiente de variação dos valores singulares da matriz de trajetória no SSA é proposto e as propriedades desta medida são discutidas nos casos uni e multivariado. Também é considerada uma abordagem alternativa: SSA é usada como uma ferramenta de filtragem de dados perturbados, a fim de reduzir o efeito da existência de outliers e o desempenho do SSA é comparado com o desempenho utilizando o método de mínimos múltiplos quadrados – (*Least Squares Multiple*) - LSM em situações de presenças de valores extremos.

Os trabalhos mais recentes envolvendo SSA e MSSA estão baseados em aplicações usando técnicas já conhecidas. MAHMOUDVAND & ZOKAEI (2012) usam o modelo de Lee-Carter como um modelo de referência para avaliar desempenhos do SSA para previsões de mortalidade em alguns conjuntos de dados reais. Este trabalho é aprimorado em MAHMOUDVAND et al. (2013).

HADDAD et al., (2012) analisam séries temporais da média global do nível do mar para determinar a sua tendência e componentes sazonais no âmbito da técnica SSA. BENEKI & LEON (2012) mostram uma série de critérios espectrais para um bom ajuste sazonal. Os critérios referem-se ao espectro de coerência e de fase entre as séries não ajustadas e as séries sazonalmente ajustadas, bem como para o espectro e a coerência entre o ruído e a série ajustada e o espectro entre a tendência-ciclo e o ruído.

MOHAMMAD (2012) fornece um quadro geral para a aplicação de SSA para problemas de detecção de ponto em série em tempo de valores reais, compara SSA com várias outras técnicas e propõe um novo processo de aprendizado por robôs.

Ainda sobre SSA, em (HADDAD et al., 2013) há uma abordagem de extração de sinal automático, com base em SSA, utilizada para análise e ajuste sazonal da série do nível médio do mar Mediterrâneo a partir de conjuntos de dados de altimetria de Satélite, que se estendem desde 1993.

SSA filtra apenas uma série temporal, contudo muitas aplicações práticas envolvem a análise de múltiplas séries temporais. GOLYANDINA & STEPANOV (2005) comparam os métodos CSSA e MSSA, duas abordagens baseadas em análise espectral singular que permitem analisar um conjunto de séries temporais. No método CSSA, a SSA é aplicada a várias séries temporais separadamente e, portanto, neste método a estrutura de correlação entre as séries não é considerada. Por sua vez na abordagem MSSA a estrutura de correlações entre as séries é considerada. GOLYANDINA & USEVICH (2005) e ZHIGLJAVSKCKY et. al (2011) mostram que a MSSA é tão eficiente quanto a SSA.

As aplicações mais recentes em SSA/MSSA são análise e previsão de preço de comercialização de energia (MIRANIAN et al., 2013), economia e previsão de inflação (HASSANI et al., 2013a; HASSANI & MAHMOUVAND, 2013; HASSANI et al, 2013b; HASSANI et al, 2012; HASSANI et al., 2013c; MENEZES et al., 2012), Geologia (HADDAD, et al., 2012), previsão de vazão de afluentes (CASSIANO et al., 2013) e consumo de energia (JUNIOR et al., 2013). MENEZES (2014) apresentou uma proposta de modelagem combinada dos modelos autorregressivos periódicos de ordem  $p$  com SSA, numa abordagem univariada e também em versão multivariada (MSSA).

No Brasil, além das referências citadas neste trabalho dos autores CASSIANO, et al. (2013) ou MENEZES, M.L. et al. (2012) ou JÚNIOR et al. (2013), aplicações de SSA podem ser vistas nos trabalhos de MACIEL (2000), PEREIRA & MACIEL (2001); PEREIRA et al. (2002); PEREIRA et al. (2004); MACHADO & PEREIRA (2006); SAITO (2003); BAKITA (2005), MACHADO (2005) e em JESUS (2008), que traz aplicações de SSA a series de consumo de energia elétrica. A autora brasileira ESQUIVEL (2012) também usou SSA para previsão de séries temporais de vento e da série IBOVESPA e observou que o algoritmo de previsão SSA consegue representar melhor as variações nos dados como flutuações sazonais e picos e reproduziu resultados

tão bons ou superiores aos métodos clássicos de previsão (Holt Winters e Box & Jenkins).