Singular Spectrum Analysis

Jorge Ramos Adrián Rodríguez
Centro de Investigación en Matemáticas. Unidad Monterrey.

***Resumen*—El Análisis Espectral Singular (SSA) es una técnica de análisis y predicción de series de tiempo que combina elementos de análisis de series de tiempo clásicas, estadística multivariada, sistemas dinámicos y procesamiento de señales. El objetivo de SSA es descomponer la serie original en la suma de un pequeño numero de componentes interpretables tales como tendencia de variación lenta, componentes oscilatorios y ruido sin estructura. En este trabajo se muestra el uso básico de SSA univariado para realizar predicción de series univariadas así como unas breves ideas del uso del método para otras aplicaciones que permiten obtener mayor información de la serie.**

I. Introducción

Las series de tiempo siguen siendo hasta el dia de hoy un tema de investigación muy activo. Debido a esta constante investigación se han desarrollado diversos metodos para analizar dichas estructuras y que han tenido aplicaciones en una o mas areas. Muchos de estos metodos hacen uso de supuestos paramétricos, por lo que se requieren cumplir con ciertos requerimientos como linealidad o estacionariedad.

Algunos enfoques alternativos usan técnicas no paramétricas que son neutrales con respecto a areas problematicas de especificación como linealidad, estacionariedad y normalidad. Como resultado, dichas técnicas pueden proveer algunas formas de analizar series de tiempo de manera confiable y algunas veces mas atinados que las versiónes parámetricas. En este sentido, el Análisis Espectral Singular (SSA) es un método no paramétrico relativamente nuevo que ha mostrado tener una buena capacidad para trabajar con series de tiempo.

La idea básica de SSA es contruir una matriz a partir de la series de tiempo de manera que dicha matriz tenga ciertas propiedades, en el caso mas simple se genera una matriz de tipo Hankel, después a esta matriz se le hace una descomposición en valores singulares y se eligen los componentes sin ruido y después se procede a reconstruir la serie sin la presencia de dicho ruido.

Dado su capacidad de separar el ruido de la serie, es posible usar el algoritmo de SSA con procesos de información de alta frecuencia que generalmente tienen mucho ruido. La capacidad de SSA de descomponer la serie en componentes: estacional, tendencia y ciclicidad permite obtener información muy rica de la serie. También se ha probado que tiene una capacidad para manejar series de tiempo cortas que es algo que los métodos tradicionales no logran hacerlo debido a la falta de observaciones. En [ZMN11] se muestra una comparación de SSA con varios modelos ARIMA y se muestra que en general se obtienen mejores resultados estimando con SSA.

II. Singular Spectral Analysis

El algoritmo de SSA se establece en dos fases de dos pasos cada una. Primero supongamos que tenemos una serie de tiempo de longitud N

$$\mathbb{X}_N = (x_1, ..., x_N)$$

entonces tomamos un valor L, con $1 < L < N$ al que llamaremos la longitúd de ventana y definamos $K = N - L + 1$. La primera fase se conoce como de descomposición.

II-A. Fase de descomposición

En esta fase vamos a descomponer la serie en componentes con el fin de separar el ruido de la trayectoria que realmente sigue la serie de tiempo. Para ello requerimos de dos pasos: Inmersión (Embedding) y descomposición en valores singulares.

II-A1. ***Inmersión (Embedding):*** El proceso de inmersión en este método se logra mapeando la serie de tiempo original en una sucesión de vectores rezagados de tamaño L, con lo que tendremos un conjunto de K vectores rezagados

$$X_i = (x_i, ..., x_{i+L-1}), \quad (1 < i < K)$$

de tamaño L.

A partir de esto generamos una matriz de trayectorias haciendo uso de estos vectores rezagados

$$\mathbf{X} = [X_1|X_2|...|X_K] = \begin{pmatrix} x_1 & x_2 & \dots & x_K \\ x_2 & x_3 & \dots & x_{K+1} \\ \vdots & \vdots & \ddots & \vdots \\ x_L & x_{L+1} & \dots & x_N \end{pmatrix}$$

Cabe notar que el (i, j)-esimo elemento de la matriz cumple que

$$x_{ij} = x_{i+j-1}$$

lo que implica que los elementos de las antidiagonales son iguales, por lo que la matriz de trayectorias contruida de esta manera es una matriz de Hankel. Algunas propiedades interesantes de este tipo de matrices y una explicación más tecnica de SSA puede encontrarse en [HMZG12].

II-A2. ***Descomposición en Valores Singulares (SVD):*** En este segundo paso se calcula la descomposición SVD de la matriz $\mathbf{S} = \mathbf{X}\mathbf{X}^t$ y denotamos por $\lambda_1, \lambda_2, ..., \lambda_L$ a los valores propios obtenidos de dicha matriz con

$$\lambda_1 \geq \lambda_2 \geq ... \geq \lambda_L \geq 0$$

y $U_1, U_2, ..., U_L$ sus respectivos vectores propios normalizados. Con esto definimos d =max{i, tal que $\lambda_i > 0$} y $V_i = \mathbf{X}^t U_i / \sqrt{\lambda_i}$ con lo que podemos obtener la descomposición espectral

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

donde $\mathbf{X}_i = \sqrt{\lambda_i} U_i V_i^t$.

II-B. Fase de reconstrucción

En esta fase se usa la descomposición espectral calculada anteriormente y mediante algunas manipulaciones se reconstruye la serie eliminando el ruido que puede alterar su comportamiento real.

II-B1. ***Agrupamiento de eigentriples:*** Una vez que se obtuvo la descomposición espectral se procede a particionar el conjunto de indices {1,...,d} en m subconjuntos disjuntos $I_1, ..., I_m$. Sea $I = \{i_1, ..., i_p\}$. Entonces la matriz resultante $\mathbf{X}_I$ correspondiente al grupo I esta definida como $\mathbf{X}_I = \mathbf{X}_{i_1} + \mathbf{X}_{i_2} + ... + \mathbf{X}_{i_p}$. Las matrices resultantes son computadas para los grupos $i = I_1, ..., I_m$ y usando la descomposición espectral generamos la descomposición

$$\mathbf{X} = \mathbf{X}_{I_1} + \mathbf{X}_{I_2} + ... + \mathbf{X}_{I_m}$$

La elección de la partición de los indices se puede hacer a través de varios criterios: inspeccionando graficos uno dimensionales de los vectores propios buscando componentes de variación lenta, inspeccionar graficos en dos dimensiones de los vectores propios para encontrar polígonos con p-vertices regulares en forma de espiral y revisar la w-correlación. Una explicación más detallada de estos criterios puede encontrarse en [GZ13], pero en esencia se trata de encontrar el numero de componentes adecuado que permita separar la información de la serie en distintos componentes.

II-B2. ***Promediado diagonal:*** Con la descomposición anterior procedemos a transformar cada matriz $\mathbf{X}_{I_j}$ en una nueva serie de longitúd N. Para esto, sea $\mathbf{Y}$ una matriz L$\times$ K con elementos y_{ij}, $1 \leq i \leq L$, $1 \leq j \leq K$. Definimos $L^* = min(L, K)$, $K^* = max(L, K)$, N=L+K-1, $y^*_{ij} = y_{ij}$ si $L < K$ y $y^*_{ij} = y_{ji}$ de otra manera. Ahora usando

$$y_k = \begin{cases} \frac{1}{k} \sum_{m=1}^{k} y^*_{m,k-m+1} & 1 \leq k < L^* \\ \frac{1}{L^*} \sum_{m=1}^{L^*} y^*_{m,k-m+1} & L^* \leq k \leq K^* \\ \frac{1}{N-k+1} \sum_{m=k-K^*+1}^{N-K^*+1} y^*_{m,k-m+1} & K^* < k \leq N \end{cases}$$

retransformamos la matriz $\mathbf{Y}$ en las N series $y_1, y_2, ..., y_N$. Esto corresponde a promediar la matriz de elementos sobre las antidiagonales $i + j = k + 1$, por ejemplo para $k = 2$ tenemos $y_2 = (y_{1,2} + y_{2,1})/2$.

Aplicando este procedimiento a alguna de las matrices resultantes X_{I_k} produce una serie reconstruida $\tilde{X}^{(k)} = (\tilde{x}_1^{(k)}), ..., \tilde{x}_N^{(k)}$. Por lo que, la serie inicial $x_1, ..., x_N$ la terminamos descomponiendo en la suma de m series reconstruidas

$$x_n = \sum_{k=1}^{m} \tilde{x}_n^{(k)} \quad (n = 1, 2, ..., N)$$

II-B3. ***w-Correlación:*** Como se menciono anteriormente una forma de establecer el agrupamiento es haciendo uso de la w-correlación por lo que se vuelve importante establecer su definición. Sea L^* y K^* como los definimos anteriormente y

$$w_i = \begin{cases} i & 0 \leq i < L^* \\ L^* & L^* \leq i \leq K^* \\ N - i + 1 & K^* < i \leq N \end{cases}$$

El peso w_i es igual al numero de veces que el elemento x_i aparece en la matriz de trayectoria $\mathbf{X}$. Definimos el producto interno de dos series $X^{(1)}$ y $X^{(2)}$ de longitud N como

$$(X^{(1)}, X^{(2)})_w \coloneqq \sum_{i=1}^{N} w_i x_i^{(1)} x_i^{(2)}$$

y diremos que las series son w-ortogonales si $(X^{(1)}, X^{(2)})_w = 0$. Entonces para medir el grado de separabilidad aproximada entre dos series calculamos su w-correlación

$$\rho^{(w)}(X^{(1)}, X^{(2)}) \coloneqq \frac{(X^{(1)}, X^{(2)})_w}{\|X^{(1)}\|_w \|X^{(2)}\|_w}$$

por lo que dos series seran aproximadamente separables si la w-correlación es cercana a cero. Cabe notar por lo establecido anteriormente que la longitúd de la ventana L es parte de la definición de la w-correlación, por lo que los resultados de la w-correlación dependen de elegir una L razonable.

III. PREDICIENDO CON SSA

Para realizar la predicción usando SSA requerimos que la series satisfaga una fórmula lineal recurrente de tamaño L-1 de la forma

$$y_t = a_1 y_{t-1} + a_2 y_{t-2} + ... + a_{L-1} y_{t-L+1}, \quad t = L + 1, ..., N$$

donde los pesos de la relación recurrente son obtenidos a partir de los vectores propios. Existen dos manera de realizar la predicción usando SSA: Método de predicción recurrente y el método de predicción por vector.

Definimos $U_j^{\triangledown}$ como el vector de las primeras L-1 componentes del vector propio U_j y por π_j al último componente de U_j. Y definimos el vector de coeficientes $A \equiv (a_1, a_2, ..., a_{L-1})$ de la siguiente manera

$$A \equiv \frac{1}{1 - v^2} \sum_{j=1}^{r} \pi_j U_j^{\triangledown}$$

con $v^2 = \sum_{j=1}^{r} \pi_j^2$.

III-A. *Método de predicción recurrente (RSSA)*

Las predicciones $(\hat{y}_{N+1}, ..., \hat{y}_{N+h})$ usando este método se resuelven usando la siguiente

$$\hat{y}_i = \begin{cases} \tilde{y}_i & i = 1, ..., N \\ A^t(\hat{y}_{i-L+1}, ..., \hat{y}_{i-1}) & i = N+1, ..., N+h \end{cases}$$

donde $\tilde{y}_1, ..., \tilde{y}_N$ son los valores de las series reconstruidas y h es el horizonte de predicción.

III-B. *Método de predicción por vector (VSSA)*

El método de predicción por vector es otro enfoque utilizado para realizar predicción usando SSA. Aunque en general RSSA arroja buenos resultados de predicción, VSSA es mas robusto en la predicción que RSSA cuando se presentan outliers o grandes choques en las series [HMOS14]. El problema cuando encontramos un punto de cambio estructural es que se rompe en cierta forma el supuesto de que la serie es representada por una formula lineal recurrente, ya que dicho punto de cambio estructural hace que dicho punto de cambio altera la relación y se debe generar una nueva formula lineal recurrente para representar la parte de la serie que se encuentra después de dicho punto.

Definimos la siguiente matriz

$$\Pi = U^{\triangledown} U^{\triangledown t} + (1 - v^2) A A^t$$

Ahora definimos el operador lineal

$$\mathfrak{p}^{(v)} : gen\{U_1, ..., U_r\} \mapsto \mathbb{R}^L$$

$$\mathfrak{p}^{(v)} Y = \begin{pmatrix} \Pi Y_{\triangle} \\ A^t Y_{\triangle} \end{pmatrix}$$

donde $Y_{\triangle}$ es un vector que contiene los ultimos L-1 elementos de Y_N. Sea

$$Z_j = \begin{cases} \tilde{X}_j & j = 1, ..., K \\ \mathfrak{p}^{(v)} Z_{j-1} & j = K+1, ..., K+h+L-1 \end{cases}$$

donde las $\tilde{X}_j$'s son las columnas reconstruidas de la matriz de trayectoria de la i-esima series después de agrupar y eliminando los componentes de ruido. Contruimos la matriz $Z = [Z_1|Z_2|...|Z_{K+h+L-1}]$ y aplicando promediado diagonal, obtenemos una nueva serie $\hat{y}_1, ..., \hat{y}_{N+h+L-1}$, donde $\hat{y}_{N+1}, ..., \hat{N+h}$ forman los h términos del horizonte de predicción que buscamos.

IV. SSA AUTOMATIZADO

De la forma en que se ha planteado el esquema de construcción del método de SSA se puede observar que requiere dos paramétros que son los que van a determinar si la predicción tiene sentido o no. A continuación se plantea un esquema para automatizar el proceso de elección de dichos parámetros y calculo de la predicción a un horizonte h. Aunque debe recordarse que el horizonte debe ser corto 1 o máximo 2 pues como cualquier método de estimación tiende a generar predicciones no confiables con un mayor horizonte.

El pseudo algoritmo para automatizar el proceso de predicción usando SSA es el siguiente

1. Dada nuestra serie de tiempo y_N, verificamos su longitud (N) y definimos el horizonte a utilizar para la prueba.
2. Dividimos la serie en 2 partes: conjunto de entrenamiento y conjunto de validación.
3. Usando el conjunto de entrenamiento generamos la matriz de trayetoria **X** eligiendo K=N-L+1. Inicialmente se utiliza L=2 pero se va a variar hasta L=N/2 verificando su desempeño.
4. Obtenemos la descomposición SVD de **X**.
5. Evaluamos el desempeño para todas las combinaciones de valores propios r $(1 \leq r \leq L-1)$, para cada L fija generamos las matrices elementales y agrupamos.
6. Realizamos el promedio diagonal para cada agrupamiento de matrices obtenido para transformarlas en matrices de Hankel, y se reconstruyen las series.
7. Seleccionamos el enfoque a usar: RSSA o VSSA.
8. Definimos una función de pérdida para que funcione como nuestro criterio de selección
9. Usamos la función de costos aplicada a la diferencia entre el valor real de la serie y la reconstruida.
10. Se comparan todos los valores obtenidos con la función de costos y elegimos los valores ed L y r que minimizaron la función.
11. Por último usando la L y r óptimas para con toda la serie se aplica RSSA o VSSA a la serie.
12. Calculamos las bandas de confianza al nivel α usando bootstrap.

V. SSA UNIVARIADO CON DATOS REALES

Para probar la efectividad del método SSA se realizo la predicción del valor del indice de apertura del SP500 para el lunes 19 de Noviembre de 2018 haciendo uso de RSSA y un enfoque que incorpora información de las series que componenen dicho indice haciendo uso de regresión PCA y un enfoque de matrices aleatorias para la elección de los componentes a usar y se obtuvieron los siguientes resultados

	19-Nov-2018	**Diferencia absoluta con el real**
Real	2730.74	-
PCA reg	2657.66	73.08
RSSA	2755.11	24.37

Cuadro I
COMPARACIÓN EN EL CALCULO DEL VALOR DEL INDICE DE APERTURA DEL SP500

Puede notarse que haciendo uso de RSSA obtuvimos una mejor predicción que con el enfoque usando información de otras series y usando regresión PCA. Por otro lado, también se realizo un contraste para medir la inflación interanual para octubre de 2019 haciendo uso de RSSA y un enfoque utilizando un modelo VAR cumpliendo todas las pruebas estaadísticas necesarias y haciendo uso de dos series de tiempo

correspondientes al M0 y a la tasa de desocupación, es decir igual que en el enfoque anterior incorporando información extra a la serie de inflación. Los resultados son los siguientes

	INPC interanual	Diferencia absoluta con el real
Real	4.9	-
VAR	5.037	0.137
RSSA	4.924	0.024

Cuadro II
COMPARACIÓN EN EL CALCULO DEL VALOR DEL INPC INTERANUAL PARA OCTUBRE DE 2018

También se utilizo una serie del PIB trimestral a precios de 2013, y se estimaron el último trimestre de 2017 y los primeros dos de 2018, los resultados se resumen en la siguiente tabla

PIB trimestral	**2017/04**	**2018/01**	**2018/02**
Predicción RSSA	18384466	18486997	18594765
Real	18291126	18469975	18441674
Diferencia absoluta	93340	17022	153091

Cuadro III
PREDICIENDO EL PIB TRIMESTRAL USANDO RSSA

Puede observarse que la diferencia absoluta es pequeña por lo que podemos indicar que se obtuvieron buenos resultados aunque cuando se tiene la presencia de cambio estructural en 1994 y 2008, aunque también el efecto de empacar en paquetes trimestrales el PIB ayuda a que el efecto del cambio estructural sea menor.

Por último se probó la eficiencia de los dos métodos de predicción RSSA y VSSA con el indice de la Bolsa Mexicana de Valores. Debido a que esta serie presenta un cambio estructural bien marcado en 2008 como se puede observar en la figura 1. Debido a este cambio estructural sperariamos que VSSA arroje mejores resultados debido a la supuesta robustez que maneja comparado con RSSA. Para hacer esta comparativa se establecio un horizonte de seis meses para la predicción, los resultados se muestran en la siguiente tabla

Real	46124.85	48358.16	44662.55
RSSA	50618.30	50768.03	50914.40
VSSA	**50231.49**	**50396.16**	**50556.94**
Real	47663.20	49698.01	49547.68
RSSA	51060.40	51209.62	51364.48
VSSA	**50713.82**	**50866.76**	**51015.75**

Cuadro IV
COMPARACIÓN DE LOS MÉTODOS DE PREDICCIÓN RSSA Y VSSA ANTE LA PRESENCIA DE UN CAMBIO ESTRUCTURAL

Como puede verse usando VSSA se obtienen mejores resultados que con RSSA cuando existe la presencia de cambio estructural en la serie.

VI. MUCHO MAS SSA

El análisis Espectral Singular tiene también una versión multivariada que en esencia funciona de la misma forma que el algoritmo básico de SSA para mas detalles puede revisarse [GZ13]. Una aplicación indirecta del algoritmo de

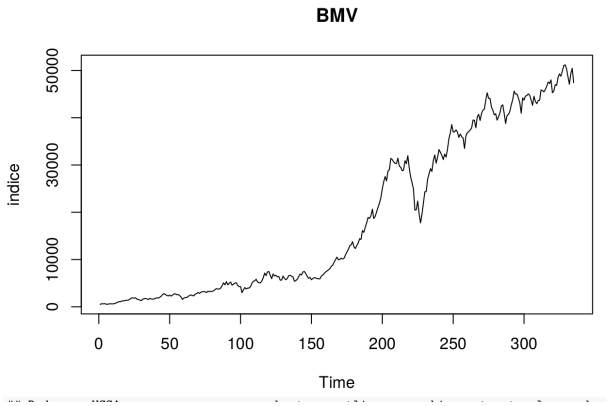

Figura 1. Serie del indice de la Bolsa Mexicana de Valores

SSA univariado es que pueden detectarse puntos de cambio estructural haciendo uso de un estadístico sobre las series reconstruidas pero tiene muchos parámetros por lo que no es necesariamente práctico su uso, para mas detalles se pueden consultar [GZ13] y [MZ03], un simple ejemplo de su funcionamiento se muestran en las figuras 2 y 3. También se han comenzado a establecer pruebas de causalidad en el mismo sentido de Granger bajo este enfoque [HZPS10] y tambien puede ser usado para rellenar datos faltantes en una serie [GZ13].

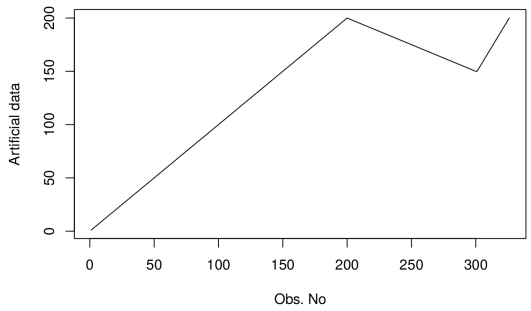

Figura 2. Generamos una serie artificial con dos puntos de cambio estructural.

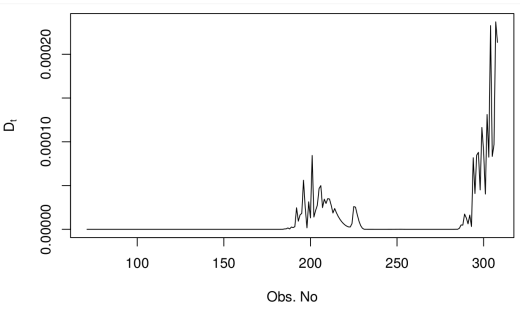

Figura 3. La serie al aplicarsele el método de detección de cambio estructural. Puede observarse el crecimiento del estadístico en los puntos de cambio estructural.

Referencias

[GZ13] Nina Golyandina and Anatoly Zhigljavsky. *Singular Spectrum Analysis for time series*. Springer Science & Business Media, 2013.

[HMOS14] Hossein Hassani, Rahim Mahmoudvand, Hardi Nabe Omer, and Emmanuel Sirimal Silva. A preliminary investigation into the effect of outlier (s) on singular spectrum analysis. *Fluctuation and Noise Letters*, 13(04):1450029, 2014.

[HMZG12] Hossein Hassani, Rahim Mahmoudvand, Mohammad Zokaei, and Mansoureh Ghodsi. On the separability between signal and noise in singular spectrum analysis. *Fluctuation and Noise Letters*, 11(02):1250014, 2012.

[HZPS10] Hossein Hassani, Anatoly Zhigljavsky, Kerry Patterson, and A Soofi. A comprehensive causality test based on the singular spectrum analysis. *Causality in Science*, pages 379–406, 2010.

[MZ03] Valentina Moskvina and Anatoly Zhigljavsky. An algorithm based on singular spectrum analysis for change-point detection. *Communications in Statistics-Simulation and Computation*, 32(2):319–352, 2003.

[ZMN11] Mohammad Zokaei, Rahim Mahmoudvand, and Nader Najari. Comparison of singular spectrum analysis and arima models. 2011.