

Fecha de recepción: 28 de abril de 2017
Fecha de aceptación y versión final: 29 de julio de 2017

RInCE

Revista de Investigaciones del Departamento de Ciencias Económicas de La Universidad Nacional de la Matanza

Artículo de investigación:

El enfoque de espacio de estado en el análisis de las series de tiempo financieras

Autora: María de las Mercedes Abril¹

Resumen: El objetivo de este trabajo es examinar los métodos para tratar una gran variedad de datos con irregularidades que suceden en series de tiempo. Los modelos autorregresivos integrados de promedios móviles (ARIMA) son frecuentemente considerados como los que proveen la base principal para el modelado de series de tiempo. Ahora bien, dada la tecnología actual, puede haber alternativas más atractivas. Una nueva y poderosa solución fue ideada por Kalman (1960) y Kalman y Bucy (1961), usando la llamada representación de espacio de estado de una serie de tiempo. Esto provee una descripción muy compacta del modelo y está basado en el resultado conocido que dice que cualquier ecuación en diferencias (o diferencial) lineal de orden finito puede ser escrita como una ecuación vectorial en diferencias (o diferencial) lineal de primer orden. La ventaja de esta última representación es que involucra solamente dependencia de un paso, lo cual conduce a un algoritmo simple para calcular las predicciones de valores futuros de la serie conocido como el algoritmo del filtro y suavizador de Kalman. Presentaremos las ideas básicas de modelado estructural de series de tiempo y haremos notar la relación con los modelos autorregresivos integrados de promedios móviles.

Palabras claves: Espacio de estado, Filtro de Kalman, Suavizador.

¹ María de las Mercedes Abril. Facultad de Ciencias Económicas. Universidad Nacional de Tucumán. Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET). Correo de contacto: mabrilblanco@hotmail.com

Title: The state space approach for the analysis of financial time series

Abstract: The aim of this paper is to examine the methods for treating a variety of data with irregularities occurring in time series. The autoregressive integrated moving average models (or ARIMA models) are often considered to be the ones that provide the main basis for modeling any time series. However, given current technology, there may be more attractive alternatives. A powerful new solution was devised by Kalman (1960) and Kalman and Bucy (1961), using the so-called state space representation of a time series. This provides a very compact description of the model and is based on the known result that says any difference equation (or differential) that is linear and has a finite order can be written as a linear vector (or differential) equation in differences of first order. The advantage of the latter representation is that it involves only one step dependence, which leads to a simple algorithm to calculate predictions of future values of the series, known as the algorithm of the Kalman filter and smoother. We will introduce the basic ideas of structural modeling of time series and we will present the relationship with autoregressive integrated moving averages or ARIMA models.

Key words: State space, Kalman Filter, Smoother.

Título: A abordagem de espaço de estado na análise de séries temporais financeiras

Resumo: O objetivo deste trabalho é examinar os métodos para o tratamento de uma variedade de irregularidades de dados que ocorrem em séries temporais. O método ARIMA são muitas vezes considerados como fornecendo a base principal para a série temporal de modelagem. No entanto, tendo em conta a tecnologia actual, podem ser alternativas mais atractivas. Uma solução nova e poderosa foi concebido por Kalman (1960) e Kalman e Bucy (1961), usando o chamado espaço de estado de uma série temporal. Isso fornece uma descrição muito compacta do modelo com base no resultado conhecido que diz que qualquer equação de diferenças (ou diferencial) ordem finita linear pode ser escrita como uma equação de diferença de vectores (ou diferencial) de primeira ordem linear. A vantagem desta última representação é que ele envolve apenas uma dependência passo, o que leva a um algoritmo simples para calcular as previsões de valores futuros da série conhecida como o algoritmo do filtro

de Kalman e amaciador. Nós introduzimos as idéias básicas de modelagem estrutural de séries temporais e notamos o relacionamento com modelos de médias móveis integrados auto-regressivos.

Palavras chave: Espaço de estado, Filtro de Kalman, Suavizador

Cuerpo del artículo

1. Introducción

1.1 Definición del problema estudiado

En este trabajo cubriremos un amplio rango de situaciones en el análisis de los datos y mostraremos una gran variedad de problemas estadísticos que pueden ser tratados mediante el uso del enfoque de espacio de estado. Tal cual lo mencionamos con anterioridad, utilizaremos la llamada representación de espacio de estado de una serie de tiempo. Esta provee una descripción muy compacta del modelo y está basada en el resultado conocido que dice que cualquier ecuación en diferencias (o diferencial) lineal de orden finito puede ser escrita como una ecuación vectorial en diferencias (o diferencial) lineal de primer orden.

Sin embargo, debemos advertir antes de comenzar con nuestra discusión que resulta extremadamente difícil presentar una descripción breve del campo de las series de tiempo. La dificultad se basa en el hecho de que la materia es por sí misma muy compleja, siendo una rama de la estadística, pero con su metodología y su propio vocabulario peculiar. De cualquier manera, una vez que las ideas fundamentales se han entendido, la materia alcanza una elegante unidad y ofrece infinitas posibilidades para el desarrollo tanto de aspectos teórico como prácticos.

1.2 Justificación del estudio del problema

La ventaja clave del enfoque de espacio, que trataremos con detalle en el desarrollo de nuestra exposición es que está basado en el análisis estructural del problema. Los diferentes componentes que hacen a la serie, tales como tendencia, estacionalidad, ciclo y variaciones calendarias, junto con los efectos de variables explicativas e intervenciones, son modelados de forma separada antes de ponerlos conjuntamente en el modelo. En contraste, el enfoque clásico establecido por Box y Jenkins actúa como una "caja negra" en la cual el modelo adoptado depende

puramente de los datos sin ningún análisis previo de la estructura del sistema que generó los mismos.

1.3 Objetivos del trabajo

Como expertos en la materia y usuarios de las técnicas estadísticas podemos decir que existen dos aspectos a tener en cuenta en el estudio de las series de tiempo: el análisis y el modelado.

El objetivo del presente análisis será el de resumir las propiedades de una serie y remarcar sus características salientes, así como algunos enfoques alternativos para el análisis de las mismas.

Nuestra atención se centrará en el modelado de un conjunto de información, teniendo en cuenta que la principal razón para realizarlo es permitir la predicción de sus valores futuros.

1.4 Antecedentes

Modelado de series de tiempo. Conceptos generales

Antes de comenzar con nuestra investigación, resulta necesario definir una serie de conceptos que nos permitirán tener una idea de la forma en la que fue evolucionando esta materia.

El primero de ellos está dado por la definición del concepto de serie de tiempo. Este término, fundamental en esta investigación, puede ser visto como un conjunto de observaciones y_1, y_2, \dots, y_p ordenadas en el tiempo. El modelo básico para representar una serie de tiempo es el modelo aditivo, dado por

$$y_t = \mu_t + \gamma_t + \varepsilon_t, \quad t = 1, \dots, n \quad (1)$$

donde μ_t es un componente que cambia suavemente en el tiempo llamado tendencia, γ_t es un componente con período fijo llamado estacionalidad y ε_t es un componente irregular conocido con el nombre de error. En muchas aplicaciones, particularmente en economía, los componentes se combinan de modo multiplicativo, dando lugar a

$$y_t = \mu_t \gamma_t \varepsilon_t \quad (2)$$

No obstante, tomando logaritmos y trabajando con el logaritmo de los valores, el modelo (2) se puede reducir a la forma del modelo (1).

Para desarrollar un modelo adecuado para μ_t y γ_t necesitamos el concepto de camino aleatorio (o random walk en inglés). Este concepto consiste en establecer una serie z_t determinada por la relación $z_t = z_{t-1} + \eta_t$ donde η_t son variables aleatorias e idénticamente distribuidas con medias iguales a 0 y varianzas iguales a σ_η^2 .

Consideremos la forma simple del modelo (1) en el cual μ_t es un camino aleatorio, no hay estacionalidad y todas las variables aleatorias se distribuyen normalmente. Esto nos da el modelo

$$\begin{aligned} y_t &= \mu_t + \varepsilon_t & \varepsilon_t &\sim N(0, \sigma_\varepsilon^2) \\ \mu_t &= \mu_{t-1} + \eta_t & \mu_t &\sim N(0, \sigma_\eta^2) \end{aligned} \quad (3)$$

para $t = 1, \dots, n$ donde ε_t y η_t son mutuamente independientes y $N(a, \sigma^2)$. Suponemos inicialmente que $\mu_0 \sim N(m_0, p_0)$ donde ambos parámetros son conocidos y donde σ_ε^2 y σ_η^2 también lo son. Este modelo se denomina modelo de nivel local. Dado que los caminos aleatorios no son estacionarios, éste también será un modelo no estacionario. Los valores de y_t son observados mientras que los valores de μ_t no pueden ser observados.

El modelo formulado en (3) es un ejemplo simple del modelo lineal gaussiano de espacio de estado (o MBEE de acuerdo a sus siglas) cuya forma general es

$$\begin{aligned} y_t &= \mathbf{Z}_t \alpha_t + \varepsilon_t, & \varepsilon_t &\sim N(0, \mathbf{H}_t), \\ \alpha_t &= \mathbf{T}_t \alpha_{t-1} + \mathbf{R}_t \eta_t, & \eta_t &\sim N(0, \mathbf{Q}_t), \end{aligned} \quad (4)$$

donde y_t es un vector de orden $p \times 1$ de observaciones α_t es un vector de orden $m \times 1$ inobservable, denominado vector de estado y $\mathbf{Z}_t, \mathbf{T}_t$ y \mathbf{R}_t son matrices conocidas. A la primera ecuación de (4) se la suele denominar ecuación de medida y a la segunda, ecuación de transición. Si uno tratara de desarrollar la teoría de los modelos de espacio de estado a partir del modelo general (4) encontraría que las ideas subyacentes no tienden a ser claras debido a arreglos de productos de matrices que son muy difíciles de comprender. Por lo tanto, en lugar de ello se prefiere trabajar por medio de diferentes partes de la teoría, como filtrado, iniciación, construcción de la verosimilitud, suavizado y estimación de parámetros en detalle para el modelo simple

expresado en (3). La ecuación (4) puede ser vista como un modelo de regresión lineal, en la cual Z_t es la matriz de regresores y α_t es el vector de coeficientes.

Modelos estructurales de series de tiempo

Una evolución dentro del análisis de series de tiempo que sentó las bases de la tarea que desarrollaremos está dado por los modelos estructurales de series de tiempo.

La idea básica de los modelos estructurales de series de tiempo (o STM de acuerdo a sus siglas en inglés) es que ellos pueden ser puestos como modelos de regresión en donde las variables explicativas son funciones del tiempo, con coeficientes que también cambian a través del tiempo. Entonces, dentro de un marco de regresión, una tendencia simple sería modelada en términos de una constante y el tiempo con un disturbio aleatorio aditivo, esto es

$$y_t = \alpha + \beta_t + \varepsilon_t, \quad t=1, \dots, n. \quad (5)$$

Este modelo es fácil de estimar usando mínimos cuadrados simples, pero sufre de la desventaja de que la tendencia es determinística. En general, esto es muy restrictivo. En efecto, en economía, por ejemplo, si una variable es considerada que tiene tendencia determinística como en (5) significaría que cualquier impulso económico de cualquier intensidad no tendrá efectos en el largo plazo, ya que todo retornará a su dada tendencia. Ahora bien, la flexibilidad es introducida permitiendo a los coeficientes α y β que evolucionen a través del tiempo como procesos estocásticos. De esta forma la tendencia se puede adaptar a los cambios subyacentes. La estimación actual, o filtrada, de la tendencia se la logra poniendo al modelo en su forma de espacio de estado y aplicándole luego el llamado filtro de Kalman. Esto último significa computar el mejor estimador en todos los puntos de la muestra usando al conjunto total de observaciones. La magnitud por la cual los parámetros pueden variar está gobernada por hiperparámetros. Ellos pueden ser estimados por máxima verosimilitud, pero, nuevamente el mejor enfoque está dado por la forma de espacio de estado y el filtro de Kalman.

La metodología de selección para los modelos estructurales es de alguna manera diferente a lo que usualmente se conoce, en el sentido que se pone menos

énfasis en la observación del correlograma de diversas transformaciones de la serie con el objeto de obtener una especificación inicial. Con esto no queremos decir que no se debe observar el correlograma, pero nuestra experiencia es que él puede ser difícil de interpretar sin un conocimiento previo de la naturaleza de la serie, y en muestras pequeñas y/o con datos desordenados puede conducir a conclusiones erróneas.

En lugar de ello, el énfasis está en la formulación del modelo en términos de componentes cuya presencia estaría sugerida por el conocimiento del fenómeno bajo estudio, de sus aplicaciones o por una inspección del gráfico de la serie. Por ejemplo, con observaciones mensuales, uno desearía incorporar desde un principio una parte estacional dentro del modelo, y solamente la sacará si luego prueba que no es significativa. Una vez que el modelo ha sido estimado, el mismo tipo de tests de diagnóstico que los usados para los modelos ARIMA pueden ser realizados con los residuos.

En particular el estadístico de Box-Ljung puede ser computado, siendo sus grados de libertad igual al número de autocorrelaciones residuales menos el número de hiperparámetros relativos. Los tests estándares de falta de normalidad y heteroscedasticidad pueden ser aplicados, como así también tests de la calidad predictiva en períodos posteriores a la muestra. Los gráficos de los residuos deben ser examinados, de la misma manera que lo apuntado por Box y Jenkins para la construcción de modelos ARIMA. En un modelo estructural de series de tiempo, estos gráficos pueden ser complementados con gráficos de los componentes suavizados. Estos, frecuentemente, suelen ser muy informativo ya que permiten al investigador constatar si los movimientos en los componentes corresponden a lo que podría esperarse sobre la base del conocimiento previo.

2. Metodología

Modelo básico de espacio de estado

El modelo lineal gaussiano de espacio de estado, centro de nuestra investigación o MBEE según sus siglas, tiene la forma ya establecida por la ecuación (4). La idea subyacente en el modelo es que el desarrollo del sistema en el tiempo está determinado por α_t , de acuerdo con la segunda ecuación de (4); pero debido a que α_t no puede ser observado directamente, debemos basar nuestro análisis en las

observaciones, es decir, y_t . Las matrices $\mathbf{Z}_t, \mathbf{T}_t, \mathbf{R}_t, \mathbf{H}_t$ y \mathbf{Q}_t se suponen inicialmente conocidas y los términos de error ε_t y η_t se suponen que son serialmente independientes e independientes entre sí en todo momento de tiempo. Las matrices \mathbf{Z}_t y \mathbf{T}_t pueden depender de y_1, \dots, y_{t-1} . El estado inicial α_0 se supone que es $N(\alpha_0, \mathbf{P}_0)$ e independiente de $\varepsilon_1, \dots, \varepsilon_n$ y de η_1, \dots, η_n , donde en principio se supone que α_0 y \mathbf{P}_0 son conocidos.

Denotemos como Y_{t-1} al conjunto y_1, \dots, y_{t-1} junto con toda la información anterior al tiempo $t=1$. Comenzando en $t=1$ y construyendo la distribución de α_t e y_t recursivamente, se puede mostrar que $p(y_t | \alpha_1, \dots, \alpha_t, Y_{t-1}) = p(y_t | \alpha_t)$ y $p(\alpha_t | \alpha_1, \dots, \alpha_{t-1}, Y_{t-1}) = p(\alpha_t | \alpha_{t-1})$.

En muchas aplicaciones \mathbf{R}_t es la matriz identidad. En otras, uno puede definir $\eta_t^* = \mathbf{R}_t \eta_t$ y $\mathbf{Q}_t^* = \mathbf{R}_t \mathbf{Q}_t \mathbf{R}_t'$ y proceder sin la inclusión explícita de \mathbf{R}_t , con lo cual el modelo es mucho más simple.

Filtrado de los modelos de espacio de estado

El objetivo del filtro de Kalman es actualizar nuestro conocimiento del sistema cada vez que una nueva observación es obtenida. Una vez que el modelo ha sido puesto en la forma de espacio de estado, el camino está abierto para la aplicación de un número importante de algoritmos. En el centro de ellos está el filtro de Kalman. Este filtro es un procedimiento recursivo para computar el estimador óptimo del vector de estado en el momento t , basado en la información disponible hasta ese tiempo t .

En esta parte de nuestro trabajo derivaremos el filtro de Kalman para el modelo (4) para el caso donde el estado inicial α_0 es $N(a_0, P_0)$ con a_0 y P_0 conocidos. Supóngase que queremos obtener la distribución a posteriori de α_{t+1} dado Y_t . Como todas las distribuciones son normales, las distribuciones condicionales son también normales. Supongamos que α_t dado Y_{t-1} es $N(a_t, \mathbf{P}_t)$ y la de α_t dado Y_t es $N(a_{t|t}, \mathbf{P}_{t|t})$. Nuestro objetivo es calcular recursivamente $a_{t|t}$, $\mathbf{P}_{t|t}$, a_{t+1} y \mathbf{P}_{t+1} dado a_t y \mathbf{P}_t .

Sea $v_t = y_t - \mathbf{Z}_t' \alpha_t$. Entonces v_t es el error de predicción un paso adelante $y_t - E(y_t | Y_{t-1})$. Denotemos su matriz de varianzas como \mathbf{F}_t . Entonces,

$$\mathbf{F}_t = \mathbf{Z}_t \mathbf{P}_t \mathbf{Z}_t' + \mathbf{H}_t \quad t=1, \dots, n. \quad (8)$$

Mediante la teoría elemental de regresión se puede demostrar que la regresión de α_t en Y_t es igual a la regresión en Y_{t-1} más la regresión en $y_t - E(y_t | Y_{t-1})$. Consecuentemente,

$$a_{t|t} = a_t + \mathbf{B}_t' v_t, \quad (9)$$

donde

$$\begin{aligned} \mathbf{B}_t &= E(\alpha_t v_t') [\text{Var}(v_t)]^{-1} \\ &= E[\alpha_t (\alpha_t - a_t)' \mathbf{Z}_t' + \alpha_t \varepsilon_t'] \mathbf{F}_t^{-1} \\ &= \mathbf{P}_t \mathbf{Z}_t' \mathbf{F}_t^{-1} \quad t=1, \dots, n. \end{aligned} \quad (10)$$

Nuevamente, mediante la teoría elemental de la regresión

$$\text{Var}(a_{t|t}) = \text{Var}(a_t) - \mathbf{B}_t' \text{Var}(v_t) \mathbf{B}_t,$$

así

$$\mathbf{P}_{t|t} = \mathbf{P}_t - \mathbf{P}_t \mathbf{Z}_t' \mathbf{F}_t^{-1} \mathbf{Z}_t \mathbf{P}_t, \quad t=1, \dots, n. \quad (11)$$

Estos resultados también pueden obtenerse mediante la técnica de completar el cuadrado usada para el modelo de nivel local, y basándose en la teoría de la distribución multivariante y sus distribuciones condicionales resultantes. La manipulación de matrices es mucho más intrincada; para ello se recomienda ver las aplicaciones presentes en la sección 3.2.2 del libro de Harvey (1989). Ellos también surgen de la teoría Bayesiana estándar. Para obtener a_{t+1} y \mathbf{P}_{t+1} simplemente es necesario notar que,

$$\begin{aligned} a_{t+1} &= E(\alpha_{t+1} | Y_t) = E(\mathbf{T}_{t+1} \alpha_t + \mathbf{R}_{t+1} \eta_{t+1} | Y_t) \\ &= \mathbf{T}_{t+1} a_{t|t} \end{aligned} \quad (12)$$

y

$$\begin{aligned}
 \mathbf{P}_{t+1} &= \text{Var}(\alpha_{t+1} | Y_t) \\
 &= \text{Var}(\mathbf{T}_{t+1} \alpha_t + \mathbf{R}_{t+1} \eta_{t+1} | Y_{t+1}) \\
 &= \mathbf{T}_{t+1} \mathbf{P}_{t|t} \mathbf{T}_{t+1}' + \mathbf{R}_{t+1} \mathbf{Q}_{t+1} \mathbf{R}_{t+1}', \quad t = 0, \dots, n-1.
 \end{aligned} \tag{13}$$

con $a_{0|0} = a_0$ y $\mathbf{P}_{0|0} = \mathbf{P}_0$.

Podemos combinar las fórmulas (9) a la (13) para obtener un solo par de ecuaciones de actualización.

$$a_{t+1} = \mathbf{T}_{t+1} a_t + \mathbf{K}_t v_t, \tag{14}$$

donde

$$\begin{aligned}
 \mathbf{K}_t &= \mathbf{T}_{t+1} \mathbf{B}_t \\
 &= \mathbf{T}_{t+1} \mathbf{P}_t \mathbf{Z}_t' \mathbf{F}_t^{-1},
 \end{aligned} \tag{15}$$

y

$$\mathbf{P}_{t+1} = \mathbf{T}_{t+1} \mathbf{P}_t (\mathbf{T}_{t+1}' - \mathbf{Z}_t' \mathbf{K}_t) + \mathbf{R}_{t+1} \mathbf{Q}_{t+1} \mathbf{R}_{t+1}', \quad t = 0, \dots, n-1. \tag{16}$$

El conjunto de fórmulas (9) a la (13) o alternativamente de la fórmula (14) a la (16), constituyen el filtro de Kalman para el modelo (4).

Se puede mostrar que cuando las observaciones no están normalmente distribuidas, restringiendo a estimaciones que son lineales en las y_t y bajo supuestos apropiados, los valores de a_{t+1} dados por el filtro minimizan el error cuadrático medio de estimación de cada componente de α_{t+1} .

Iniciación en los modelos de espacio de estado

Ahora consideraremos como iniciar el filtrado cuando a_0 y \mathbf{P}_0 son desconocidos y α_0 es tomada con distribución difusa. Una aproximación adecuada puede, con frecuencia, alcanzarse numéricamente tomando $a_0 = 0$ y $\mathbf{P}_0 = K \mathbf{I}_m$ donde K es un número finito pero grande.

Una buena alternativa que opera en casos suficientemente simples se explica a continuación. Esto necesita la solución de los primeros n pares de ecuaciones en (4) para α_m en términos de $Y_m, \varepsilon_1, \dots, \varepsilon_m$ y η_1, \dots, η_m . Expresando la solución de la siguiente manera, $\alpha_m = AY_m^* + u$ donde Y_m^* es Y_m escrito como un vector columna y u depende linealmente de los ε_t 's y los η_t 's se tratará a Y_m como fijo y a u como si tuviera su distribución incondicional, y se inicia el filtro tomando, $a_{m|m} = AY_m$ y $\mathbf{P}_{m|m} = \text{Var}(u)$.

Presentaremos una versión simplificada del tratamiento de un método general desarrollado por de Jong (1991) y que figura en la obra de Harvey y Shepard (1993). El mismo está basado en la idea de encontrar la densidad condicional de α_0 dadas unas pocas observaciones iniciales. Sean a_t^0 y v_t^0 los valores de a_t y v_t calculados por el filtro de Kalman, definido en (9) y (11) para un valor de α_0 , suponiendo que $\mathbf{P}_0 = 0$, y sean a_t^* , v_t^* valores similares considerando que $\alpha_0 = 0, \mathbf{P}_0 = 0$. Entonces

$$\begin{aligned} a_{t+1}^0 &= \mathbf{T}_{t+1} a_t^0 + \mathbf{K}_t v_t^0 \\ &= (\mathbf{T}_{t+1} - \mathbf{K}_t \mathbf{Z}_t) a_t^0 + \mathbf{K}_t y_t. \end{aligned}$$

De forma similar

$$a_{t+1}^* = (\mathbf{T}_{t+1} - \mathbf{K}_t \mathbf{Z}_t) a_t^* + \mathbf{K}_t y_t.$$

En consecuencia

$$a_{t+1}^0 - a_{t+1}^* = (\mathbf{T}_{t+1} - \mathbf{K}_t \mathbf{Z}_t) (a_t^0 - a_t^*)$$

Como $a_1^0 = \mathbf{T}_1 \alpha_0$ y $a_1^* = 0$, tenemos que,

$$a_{t+1}^0 = \mathbf{A}_t \alpha_0 + a_{t+1}^*, \quad t = 0, \dots, n-1, \quad (17)$$

con $\mathbf{A}_t = (\mathbf{T}_{t+1} - \mathbf{K}_t \mathbf{Z}_t) \mathbf{A}_{t-1}$ donde $\mathbf{A}_0 = \mathbf{T}_1$. Se desprende entonces que,

$$v_t^* = v_t^0 + \mathbf{Z}_t \mathbf{A}_{t-1} \alpha_0.$$

Sean $\mathbf{V}_t^0 = (\mathbf{v}_1^{0'}, \dots, \mathbf{v}_t^{0'})'$, $\mathbf{V}_t^* = (\mathbf{v}_1^{*'}, \dots, \mathbf{v}_t^{*'})'$, sea \mathbf{C}_t una matriz en bloque con bloque fila igual a $\mathbf{Z}_s \mathbf{A}_{s-1}$ y sea $\tilde{\mathbf{F}}_t$ una matriz diagonal en bloque con bloques \mathbf{F}_s . Como $E(v_t^0 | \alpha_0) = 0$, entonces, \mathbf{V}_t^* dado α_0 es $N(\mathbf{C}_t \alpha_0, \tilde{\mathbf{F}}_t)$

Ahora tomemos a α_0 como un vector aleatorio con densidad a priori $N(a_0, \mathbf{P}_0)$. La densidad a posteriori de α_0 dado \mathbf{Y}_t es la misma que la densidad a posteriori dado \mathbf{V}_t^* , por el teorema de Bayes para densidades gaussianas

$$p(\alpha_0 | \mathbf{Y}_t) \sim N\left[\left(\mathbf{P}_0^{-1} + \mathbf{S}_t\right)^{-1} \left(\mathbf{P}_0^{-1} a_0 + s_t\right), \left(\mathbf{P}_0^{-1} + \mathbf{S}_t\right)^{-1}\right]$$

donde $\mathbf{S}_t = \mathbf{C}'_t \tilde{\mathbf{F}}_t^{-1} \mathbf{C}_t$ y $s_t = \mathbf{C}'_t \tilde{\mathbf{F}}_t^{-1} \mathbf{V}_t^*$. Esas cantidades pueden calcularse de forma recursiva, utilizando,

$$\mathbf{S}_t = \mathbf{S}_{t-1} + \mathbf{A}'_{t-1} \mathbf{Z}'_t \mathbf{F}_t^{-1} \mathbf{Z}_t \mathbf{A}_{t-1}, \text{ y} \tag{18}$$

$$s_t = s_{t-1} + \mathbf{A}'_{t-1} \mathbf{Z}'_t \mathbf{F}_t^{-1} v_t^*,$$

para $t=1, \dots, n$ con $\mathbf{S}_0 = 0$ y $s_0 = 0$. En el caso difuso $\mathbf{P}_0^{-1} = 0$, por lo tanto tenemos,

$$p(\alpha_0 | \mathbf{Y}_t) \sim N(\mathbf{S}_t^{-1} s_t, \mathbf{S}_t^{-1})$$

Sea τ el primer valor de t para el cual existe \mathbf{S}_t^{-1} . Como $E(\alpha_0 | \mathbf{Y}_\tau) = \mathbf{S}_\tau^{-1} s_\tau$, de (16) se puede ver que $a_{\tau+1} = E(a_{\tau+1}^0 | \mathbf{Y}_\tau)$ está dado por

$$a_{\tau+1} = \mathbf{A}_\tau \mathbf{S}_\tau^{-1} s_\tau + a_{\tau+1}^*, \tag{19}$$

y que $\mathbf{P}_{\tau+1} = \text{Var}(\alpha_{\tau+1} - a_{\tau+1} | \mathbf{Y}_\tau) = E[\text{Var}(\mathbf{A}_\tau \alpha_0 + a_{\tau+1}^* | \mathbf{Y}_\tau)]$ está dado por,

$$\mathbf{P}_{\tau+1} = \mathbf{A}_\tau \mathbf{S}_\tau^{-1} \mathbf{A}'_\tau + \mathbf{P}_{\tau+1}^*, \tag{20}$$

donde $\mathbf{P}_{\tau+1}^*$ es el valor dado por el filtro de Kalman suponiendo $a_0 = 0$ y $\mathbf{P}_0 = \mathbf{0}$. Como $a_{\tau+1}$ y $\mathbf{P}_{\tau+1}$ están adecuadamente definidos, podemos continuar con el filtro de Kalman, que fue presentado en el apartado anterior, en forma normal para $t = (\tau+1), \dots, (n-1)$. Cuando α_0 tiene una distribución a priori difusa, este método brinda los valores de $a_t = E(\alpha_t | \mathbf{Y}_{t-1})$ para $t = (\tau+1), \dots, (\tau+2)$; sin embargo, los valores para $t \leq m$ no existen. Para modelos adecuadamente comportados $\tau = m$, pero esto no siempre necesita ser así.

Los cálculos pueden ser implementados sistemáticamente aumentando el vector de estado de la siguiente forma. Sea $\mathbf{Y}_t^+ = [y_t \quad \mathbf{0}]$, donde $\mathbf{0}$ es una matriz de orden $p \times m$ cuyos elementos son todos iguales a cero, y sea \mathbf{A}_t^+ una matriz de orden $m \times (m+1)$ obtenida a partir de la siguiente recursión,

$$\mathbf{A}_{t+1}^+ = \mathbf{T}_{t+1} \mathbf{A}_t^+ + \mathbf{K}_t \mathbf{N}_t^+, \quad t = 1, \dots, n-1. \quad (21)$$

la cual es análoga a la recursión del filtro de Kalman obtenida en el apartado anterior, donde

$$\mathbf{N}_t^+ = \mathbf{Y}_t^+ - \mathbf{Z}_t \mathbf{A}_t^+, \quad (22)$$

y donde el valor inicial es tomado como $\mathbf{A}_1^+ = \mathbf{0}$. La primera columna de \mathbf{A}_t^+ es a_t^* y las restantes m son los vectores obtenidos al aplicar el mismo filtro de Kalman que se utilizó para calcular a_t^* a los últimos m vectores nulos de \mathbf{Y}_t^+ . También definimos la matriz S_t^+ de orden $(m+1) \times (m+1)$ mediante la recursión,

$$S_t^+ = S_{t-1}^+ + N_t^{+'} F_t^{-1} N_t^+, \quad t = 1, \dots, n, \quad (23)$$

con $S_0^+ = 0$.

Ahora, particionando \mathbf{A}_t^+ y S_t^+ de modo que sean conformables como

$$\mathbf{A}_t^+ = \begin{pmatrix} a_t^* & \mathbf{A}_t \end{pmatrix} \quad \text{y} \quad S_t^+ = \begin{pmatrix} s_t^0 & s_t^{' } \\ s_t & S_t \end{pmatrix},$$

donde s_t^0 es un escalar, vemos que $\mathbf{A}_t, \mathbf{S}_t$ y s_t son las mismas dadas en (18). Por lo tanto, para realizar cálculos es necesario iniciar utilizando (21) hasta que S_t^+ se hace no singular en el tiempo $t = \tau$, luego se calcula $a_{\tau+1}$ y $\mathbf{P}_{\tau+1}$ utilizando (18) y (19) y a partir de allí se continúa utilizando el filtro normal. Por supuesto que no es necesario dejar de usar (21) a partir del momento $t = \tau$; el proceso podría continuarse calculando $a_{\tau+1}$ mediante una relación análoga a (19). Harvey (1989), en la sección 3.4.3 de su trabajo, hace referencia a que estos resultados fueron obtenidos de de Jong (1991), sin que existiera una prueba, pero incluye una extensión para el caso en el que algunos elementos de α_t son estacionarios. El tratamiento de de Jong (1991) es considerablemente más general.

Suavizado de las series

En este apartado presentaremos recursiones para calcular $\hat{\alpha}_t = E(\alpha_t | \mathbf{Y}_n)$.

Comenzaremos suponiendo que a_0 y \mathbf{P}_0 son conocidos.

Definamos,

$$\hat{\alpha}_t = a_t + \sum_{j=1}^n B_{ij} \mathbf{v}_j, \quad t = 1, \dots, n \quad (24)$$

donde B_{ij} es el coeficiente de la regresión de α_t en \mathbf{v}_t y está dado por

$$B_{ij} = E\left(\alpha_t \mathbf{v}_j'\right) \mathbf{F}_j^{-1}. \quad \text{Ahora bien,} \quad \mathbf{v}_j = \mathbf{Z}_j (\boldsymbol{\alpha}_j - \mathbf{a}_j) + \boldsymbol{\varepsilon}_j \quad \text{y, además,}$$

$\alpha_s - a_s = \mathbf{L}_{s-1} (\alpha_{s-1} - a_{s-1}) + \mathbf{R}_s \eta_s - \mathbf{K}_{s-1} \boldsymbol{\varepsilon}_{s-1}$, donde $\mathbf{L}_{s-1} = \mathbf{T}_s \mathbf{K}_{s-1} \mathbf{Z}_{s-1}$ para $s = j, j-1, \dots$. Como

$E\left[(\boldsymbol{\alpha}_t - \mathbf{a}_t)(\boldsymbol{\alpha}_t - \mathbf{a}_t)'\right] = \mathbf{P}_t$ tenemos, en consecuencia,

$$E\left(\alpha_t \mathbf{v}_j'\right) = \mathbf{P}_t \mathbf{L}_t' \mathbf{L}_{t+1}' \dots \mathbf{L}_{j-1}' \mathbf{Z}_j', \quad j = t+1, \dots, n.$$

Sustituyendo en (24) obtenemos la siguiente expresión

$$\mathbf{r}_{t-1} = \mathbf{Z}_t' \mathbf{F}_t^{-1} \mathbf{v}_t + \mathbf{L}_t' \mathbf{r}_t, \quad t = n, n-1, \dots, 1, \quad (25)$$

donde \mathbf{r}_t satisface la recursión hacia atrás

$$\mathbf{r}_{t-1} = \mathbf{Z}_t' \mathbf{F}_t^{-1} \mathbf{v}_t + \mathbf{L}_t' \mathbf{r}_t, \quad t = n, n-1, \dots, 1, \quad (26)$$

con $r_n = 0$. La varianza de r_{t-1} está dada por

$$N_{t-1} = \mathbf{Z}_t' \mathbf{F}_t^{-1} \mathbf{Z}_t + \mathbf{L}_t' N_t \mathbf{L}_t, \quad t = n, n-1, \dots, 1, \quad (27)$$

con $N_n = 0$. Los residuos suavizados tal como lo definen de Jong (1989) y de Jong y Penzer (1998), son estimados mediante la recursión,

$$u_t = \mathbf{F}_t^{-1} \mathbf{v}_t - \mathbf{K}_t' \mathbf{r}_t. \quad (28)$$

La varianza de u_t está dada por

$$\mathbf{M}_t = \mathbf{F}_t^{-1} + \mathbf{K}_t' N_t \mathbf{K}_t. \quad (29)$$

Las cantidades u_t y r_t juegan el rol de pivotes en la construcción del test de diagnóstico para observaciones atípicas y cambios estructurales.

El suavizador, definido en (26) y (27) fue introducido por de Jong (1989) en un esquema más general que el considerado aquí. En la literatura, el suavizador $E(\alpha_t | Y_n)$ para t variables con n fijo se llama suavizador de intervalo fijo. Si t es fijo y $n = (t+1), (t+2), \dots$ se llama suavizador de punto fijo, y si ambos varían, con $n-t$ fijo se llama suavizador de rezago fijo. En su trabajo, de Jong (1989), presenta algoritmos

análogos para los suavizadores de punto fijo y de rezago fijo. Un suavizador planteado con anterioridad a los que acabamos de describir, se denominará suavizador clásico; se encuentra en el trabajo de Harvey (1989) y es derivado a partir de la obra de Anderson y Moore (1979). Este último toma la forma,

$$\hat{\alpha}_t = T_1 a_1 + P_t^* (\hat{\alpha}_{t+1} - a_{t+1}), \quad t = (n-1), \dots, 1; \quad (30)$$

$$P_t^* = P_{t|t} T_{t+1}' P_t^{-1},$$

con $\hat{\alpha}_n = a_{n|n}$. Comparando esta expresión con (26) y (27) se observa que el suavizador clásico requiere la inversión de $n-1$ matrices P_t , las que posiblemente pueden ser grandes, mientras que el suavizador de de Jong no requiere de esas inversiones salvo la de F_t , la cual ya fue invertida durante el proceso de filtrado de Kalman. Esto puede ser una ventaja considerable cuando tratamos con modelos grandes.

Si las varianzas y covarianzas no son requeridas, una variante del suavizador de de Jong debida a Koopman (1993), es computacionalmente más rápida. Sea $\hat{\eta}_t = E(\eta_t | Y_n)$, $t=1, \dots, n$. Como $E(\eta_t v_j') = 0$ para $j < t$, tenemos,

$$\hat{\eta} = \sum_{j=t}^n E(\eta_t v_j') F_j^{-1} v_j. \quad (31)$$

Como antes, $v_j = Z_j(\alpha_j - a_j) + \varepsilon_j$ y $\alpha_s - a_s = L_{s-1}(\alpha_{s-1} - a_{s-1}) + R_s \eta_s - K_{s-1} \varepsilon_{s-1}$, así para $s > t$, $E[\eta_t (\alpha_s - a_s)'] = E[\eta_t (\alpha_{s-1} - a_{s-1})'] L_{s-1}'$ mientras que $E[\eta_t (\alpha_s - a_s)'] = Q_t R_t'$. De esta manera,

$$E(\eta_t v_j') = Q_t R_t' L_t' L_{t+1}' \dots L_{j-1}' Z_j', \quad j = t+1, \dots, n.$$

Sustituyendo en (31) obtenemos

$$\hat{\eta} = Q_t R_t' r_{t-1}, \quad (32)$$

donde r_t satisface la recursión presentada en (26). Los $\hat{\alpha}_t$ son entonces calculados mediante la recursión hacia delante

$$\hat{\alpha}_t = T_t \hat{\alpha}_{t-1} + R_t \hat{\eta}_t, \quad t = 1, \dots, n-1, \quad (33)$$

junto con $\hat{\alpha}_n = a_{n|n}$. El valor inicial es $\hat{\alpha}_0 = a_0 + \hat{\eta}_0$ con $\hat{\eta}_0 = P_0 r_0$.

La ventaja de este suavizador sobre el de de Jong es que es significativamente menos costoso en términos computacionales y requiere menos espacio de almacenamiento. Los costos relativos de computación son examinados con todo detalle por Koopman en el desarrollo de su trabajo. Por otra parte, el suavizador de Koopman tiene la ventaja de que las matrices de varianzas y covarianzas de los errores de estimación no pueden calcularse a partir de él; así, si ellas son necesarias, se deberá usar el suavizador de de Jong o en su defecto, el clásico.

Estimación de los hiperparámetros por máxima verosimilitud

Suponiendo que a_0 y P_0 son conocidos, la densidad conjunta de y_1, \dots, y_n es

$$p(Y_n) = \prod_{t=1}^n p(y_t | Y_{t-1}),$$

donde $p(y_t | Y_t) = N(Z_t \alpha_t, F_t)$. Por lo tanto, tomando logaritmos obtenemos

$$\log L = -\frac{np}{2} \log(2\pi) - \frac{1}{2} \sum_{t=1}^n \log |F_t| - \frac{1}{2} \sum_{t=1}^n v_t' F_t^{-1} v_t, \quad (34)$$

donde $v_t = y_t - Z_t \alpha_t$ y F_t se encuentra expresada en (8). A la expresión (34) se la conoce como descomposición del error de predicción del logaritmo de la verosimilitud. Esta última fórmula debe maximizarse con respecto a los elementos del vector ψ de hiperparámetros desconocidos. Estableciendo $H_t = \sigma^2 H_t^*$, $Q_t = \sigma^2 Q_t^*$ encontramos que $F_t = \sigma^2 F_t^*$ donde F_t^* depende solamente de H_t^* y de Q_t^* junto con cualquier hiperparámetro desconocido que se encuentre en las restantes matrices del modelo. El parámetro de escala σ^2 puede concentrarse de forma tal que termine reduciendo la dimensión de la búsqueda numérica en una unidad.

En el caso difuso, cuando a_0 y P_0 son desconocidos, como se lo planteó en el apartado anterior, consideremos que τ es el menor valor de t para el cual $p(\alpha_t | Y_t)$ existe. Entonces tomamos la verosimilitud condicional con Y_τ dando lugar a la expresión

$$\log L = -\frac{(n-\tau)}{2} \log(2\pi) - \frac{1}{2} \sum_{t=\tau+1}^n \log |F_t| - \frac{1}{2} \sum_{t=\tau+1}^n v_t' F_t^{-1} v_t. \quad (35)$$

Las últimas dos expresiones, es decir, las fórmulas (34) y (35) pueden maximizarse numéricamente luego de haber sido concentradas.

Observaciones faltantes

Si todos los elementos de los vectores de observaciones y_{r+1}, \dots, y_u faltan, podemos utilizar el mismo algoritmo de filtrado y suavizado que se usaría si todas las observaciones estuvieran presentes pero tomando $v_t = 0, K_t = 0$ en todos los puntos en los cuales no hay observaciones. Los vectores de observaciones faltantes pueden estimarse por $Z_t \hat{\alpha}_t$, donde $\hat{\alpha}_t$ es el valor suavizado de α .

No obstante, supongamos que en el momento t faltan algunos pero no todos los componentes del vector de observaciones. Sea y_t^* el vector de valores realmente observados y sea y_t el vector completo de orden $p \times 1$ del cual algunos de sus elementos no son observados. Entonces $y_t^* = W_t y_t$ donde W_t es una matriz conocida cuyas filas son un subconjunto de las filas de I_p . Consecuentemente, en los puntos de tiempo donde no todos los elementos de y_t están disponibles, la primera ecuación del modelo básico de espacio de estado, la cual es

$$y_t = Z_t \alpha_t + \varepsilon_t$$

es reemplazada por

$$y_t^* = Z_t^* \alpha_t + \varepsilon_t^*, \quad \varepsilon_t^* \sim N(0, H_t^*), \quad (36)$$

donde $Z_t^* = W_t Z_t$, $\varepsilon_t^* = W_t \varepsilon_t$, y $H_t^* = W_t H_t W_t'$. El filtro y el suavizador de Kalman operarán exactamente como en el caso estándar, siempre que y_t, Z_t y H_t sean reemplazados por y_t^*, Z_t^* y H_t^* en los puntos relevantes. Por supuesto, la dimensión del vector de observaciones varía en el tiempo, pero ello no afecta la validez de las fórmulas. Los elementos faltantes pueden ser estimados por los elementos apropiados de $Z_t \hat{\alpha}_t$, donde $\hat{\alpha}_t$ es el valor suavizado. Es obvio que estos métodos no puede usarse cuando ocurren valores rezagados de y_t en el vector de estado, como, por ejemplo, en el caso de modelos ARMA. Algunos métodos para el tratamiento de

valores faltantes en ese tipo de modelos han sido presentados en las obras de Harvey y Pierse (1984), Kohn y Ansley (1986) y Gómez y Maravall (1993).

3. Resultados

Ilustremos este análisis con un breve ejemplo que se encuentra presente en detalle en otros trabajos de investigación. Para ello tomemos la siguiente situación. Consideraremos un brevísimo análisis de las cargas transportadas, en especial las despachadas a través del Ferrocarril Central Norte Argentino para la década de 1920, a partir de cifras obtenidas de las Memorias Ministerio de Obras Públicas de la Nación. Utilizaremos las técnicas que hemos desarrollado para estudiar cómo se organizaba este ferrocarril. Para mayores detalles sobre el desarrollo de este enfoque podemos dirigirnos a Abril y Blanco (2015).

Para finales del siglo XIX el Central Norte tenía 1.049 kilómetros de longitud y estaba organizado en dos secciones. Este fue el puntapié inicial del posterior desarrollo de la red del Estado que se llevó a cabo durante las primeras décadas del nuevo siglo. A partir de entonces, el Central Norte extendió sus rieles a nuevas regiones. Sin embargo, para 1914 la expansión de las líneas del Estado detuvo su ímpetu por falta del tan necesario financiamiento externo. Pero al iniciarse la década de 1920 se inició un nuevo período en la vida de estas líneas férreas. Al gobierno del presidente Irigoyen, le quedaba aún un tercio de su mandato. Entre las preocupaciones del momento, el profundo desequilibrio económico regional era una de ellas.

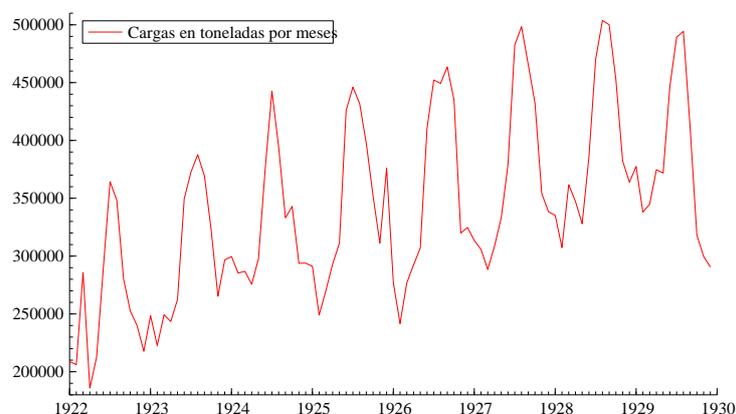
A fin de remediar estas cuestiones, el gobierno se propuso en materia ferroviaria llevar adelante dos proyectos importantes; la finalización de las obras interrumpidas por la guerra y las conexiones fronterizas en el norte. Algunas de las líneas que se construyeron con este objetivo en mente tardaron más tiempo del necesario en completarse y otras definitivamente se paralizaron. Pero a pesar de las dificultades de orden financiera para 1930, las líneas del Estado habían alcanzado 8.979 kilómetros de extensión. Más de dos mil kilómetros nuevos de vías férreas se terminaron sumaron a las ya existentes.

El Central Norte Argentino tuvo un rol fundamental en la expansión económica de la región del noroeste de nuestro país. Su importancia fue determinante para

consolidar su integración con los puertos del litoral y de Buenos Aires. Este hecho se pone en evidencia cuando observamos de acuerdo a las cifras obtenidas de las fuentes consultadas que el porcentaje que representa la carga transportada sobre el ingreso total es de un 77% en promedio para la década en consideración.

Es necesario destacar que el principal transporte de esta línea, constituido por caña de azúcar y sus derivados, hacía que en la gran mayoría de las situaciones los máximos se presenten hacia mitad de año, es decir, por lo general, en los meses de julio y agosto, los meses que por lo general son de zafra. Esto nos da un indicio de la existencia de un componente estacional que se observa con suma claridad en el siguiente gráfico.

Gráfico 1: Total de carga transportada por mes (expresadas en toneladas) por el Ferrocarril Central Norte Argentino durante la década de 1920.

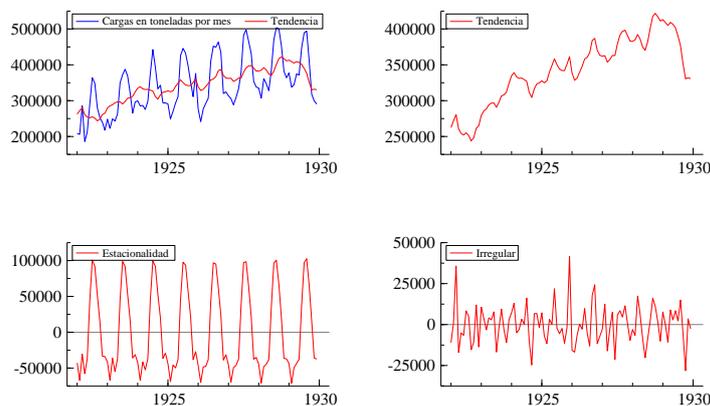


Fuente: Elaboración propia.

La idea básica del enfoque utilizado es que la información disponible puede ser analizada por medio de un modelo similar a una regresión lineal, en donde las variables explicativas son funciones del tiempo, pero con coeficientes que también cambian con el transcurso del tiempo.

Nos encontramos con dos ecuaciones, cuando tratamos con un modelo de este tipo; una correspondiente a aquellos elementos no observables pero estimables y otra que plantea una relación lineal entre esos elementos y el conjunto de información bajo estudio. A su vez, tenemos la gran ventaja de poder observar efectivamente cada componente que hace a la serie tal cual lo muestra el siguiente gráfico

Gráfico 2: Descomposición por medio de un modelo de espacio de estado de la carga total transportada por mes (expresadas en toneladas) por el Ferrocarril Central Norte Argentino durante la década de 1920.

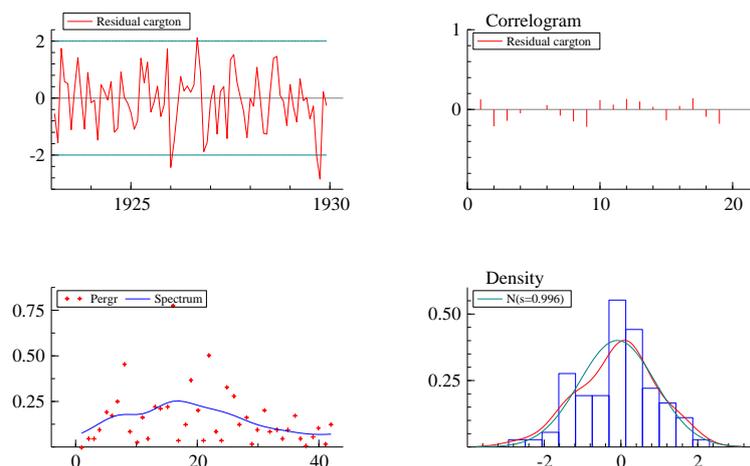


Fuente: Elaboración propia.

En este caso las cargas transportadas por mes se pueden descomponer en una serie de elementos como ser la tendencia, la estacionalidad y un componente irregular que no son observables a simple vista, pero que pueden estimarse y estudiarse de forma separada si el problema así lo demanda. La tendencia, es claramente ascendente, indicador en muchos casos hasta de posibles sobreproducciones de ciertos elementos transportados, aunque la misma experimenta un marcado descenso cerca del año 1929, fin del período que hemos estudiado. Por otro lado, este conjunto de datos presenta un claro componente estacional. Dicho componente se encuentra relacionado a la principal carga que transportaba esta línea y se encuentra atado a los correspondientes períodos de siembra y cosecha del mismo.

Un segundo análisis que se puede realizar es observar el comportamiento de los residuos que surgen de la aplicación de un modelo de este tipo al conjunto de datos que hemos elegido para estudiar. Esto nos permite observar si hemos dejado algo de lado en nuestra primera estimación de los componentes de ese conjunto de datos. Dicha situación se puede observar en el siguiente gráfico

Gráfico 3: Residuos de la aplicación de un modelo de series de tiempo a la carga total transportada por mes (expresadas en toneladas) por el Ferrocarril Central Norte Argentino durante la década de 1920.



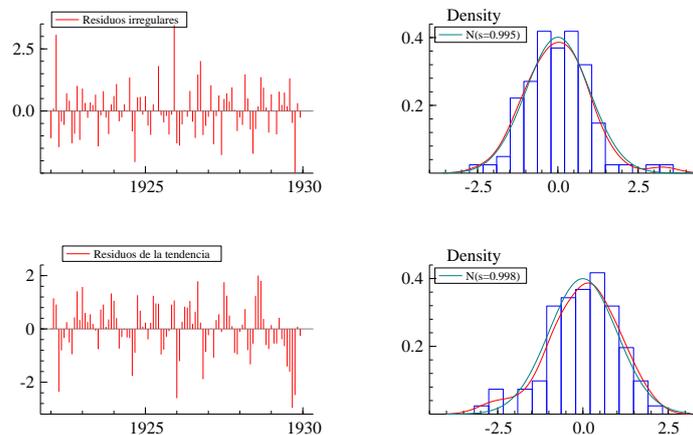
Fuente: Elaboración propia.

La mayoría de los residuos se encuentran dentro de la banda de confianza del 95% y se ajustan a una densidad del tipo normal. Además, si tomamos la tendencia y el componente irregular, los modelamos, y nos centramos en los residuos que surgen de este modelo, veremos que sucede una situación bastante similar, es decir, ambos residuos se ajustan a una distribución de tipo normal.

Todo esto tomado en conjunto nos permite inferir que, si tomamos un modelo de este tipo, en donde es posible separar los diversos componentes presentes en un conjunto de datos como éste y modelarlos de un modo análogo al de una regresión

lineal con coeficientes estocásticos, el mismo describirá adecuadamente sus principales características, como ser la existencia de una tendencia ascendente y la presencia de estacionalidad, relacionada al principal bien que transportaba el mismo. Esta situación se observa a continuación

Gráfico 4: Residuos del componente irregular y residuos de la tendencia que surgen de la aplicación de un modelo de series de tiempo a la carga total transportada por mes (expresadas en toneladas) por el Ferrocarril Central Norte Argentino durante la década de 1920.



Fuente: Elaboración propia.

4. Discusión

Con respecto a la línea férrea que hemos analizado podemos decir que luego de revisar las fuentes históricas pertinentes llegamos a la conclusión de que las líneas férreas estatales, como las construidas en nuestro territorio, fueron más el resultado de una demanda previa que el generador de una demanda nueva.

Desde el punto de vista histórico se observa que el Central Norte Argentino tuvo un rol fundamental en la expansión económica del noroeste de nuestro país. Su importancia fue determinante para consolidar su integración comercial con los puertos del litoral y de Buenos Aires. Este hecho se pone en evidencia cuando determinamos de acuerdo al análisis de las fuentes históricas que el porcentaje que representaba la

carga transportada sobre el ingreso total que era un promedio de un 77% para la década bajo estudio.

Por otro lado, otro hecho surge cuando analizamos las cifras referidas a esta línea. A partir de ellas corroboramos que la caña de azúcar era uno de los principales productos que transportaba y el volumen acarreado era de tal magnitud que el desempeño de este ferrocarril se encontraba atado a los ciclos agrícolas de la misma, es decir, la cantidad aumentaba en época de cosecha y disminuía en caso contrario tal como se lo puede ver a la hora de aplicar un modelo como el que hemos presentado en forma teórica. En este caso elementos tales como la tendencia, la estacionalidad y el ciclo, junto con los efectos de otras variables explicativas e intervenciones se pueden modelar de un modo muy simple tomarse por separado de ser necesario tal cual lo hemos visto a la hora de aplicar este enfoque.

Con respecto al enfoque que hemos utilizado en este caso, podemos decir que los modelos de espacio de estado tienen el gran atractivo de ser muy generales. Cubren una amplia gama de situaciones y posibilidades de análisis e incluyen a todos los modelos tradicionales.

Debido a la naturaleza Markoviana de los mismos, los cálculos necesarios para su implementación pueden ser puestos de una manera recursiva. Esto permite que grandes cantidades de información puedan ser efectivamente manejadas sin incrementos desproporcionados en la carga computacional, lo cual resulta sumamente útil a la hora de enfrentar un problema práctico como el que hemos presentado muy brevemente.

5. Conclusiones

Tal como lo hemos mencionado al comienzo en este trabajo hemos cubierto un amplio rango de situaciones en el análisis de los datos y hemos mostrado una gran variedad de problemas estadísticos que pueden ser tratados mediante el uso del enfoque de espacio de estado al presentar un ejemplo en el que se vio gráficamente el uso de este enfoque.

La ventaja clave del enfoque de espacio de estado es que está basado en el análisis estructural del problema. Los diferentes componentes que hacen a la serie, tales como tendencia, estacionalidad, ciclo y variaciones calendarias, junto con los

efectos de variables explicativas e intervenciones, son modelados de forma separada antes de ponerlos conjuntamente en el modelo de espacio de estado. Es de la responsabilidad del investigador identificar y modelar cualquier rasgo que requiera especial tratamiento en situaciones particulares. En contraste, el enfoque tradicional es una "caja negra" en la cual el modelo adoptado depende puramente de los datos sin ningún análisis previo de la estructura del sistema que generó los datos.

Una segunda ventaja de los modelos de espacio de estado es que son muy flexibles. Debido a su naturaleza recursiva y a las técnicas computacionales usadas para analizarlos, es inmediata la posibilidad de incorporar dentro de la estructura del sistema los cambios conocidos que suceden en el tiempo. Por el contrario, los modelos tradicionales son homogéneos a través del tiempo porque están basados en el supuesto que la serie luego de ciertas transformaciones (usualmente diferenciada) sea estacionaria.

Sin embargo, la principal desventaja que este enfoque posee respecto a otros utilizados con mayor asiduidad es la falta de información, conocimiento y de programas de computación relativos a estos modelos dentro de la comunidad estadística. Pero, a pesar de ello, podemos decir que los modelos de espacio de estado son muy generales y cubren una amplia gama de situaciones y enfoques, incluyendo a todos los modelos tradicionales. En el caso de tener observaciones multivariadas, las mismas pueden ser manejadas mediante una extensión directa de la teoría univariada, lo que no sucede con los modelos mayormente utilizados en situaciones prácticas.

Es muy fácil incorporar ciclos y observaciones perdidas a los modelos de espacio de estado; además las variables explicativas pueden ser agregadas a ellos sin ninguna dificultad. Más aún, los coeficientes de regresión asociados pueden variar estocásticamente en el tiempo si ello fuera necesario en las aplicaciones. Ajustes por días laborables y otras variaciones calendarías pueden fácilmente ser incorporadas lo cuales los hacen muy atractivos en ámbitos tales como el análisis de cuestiones económicas.

6. Referencias

ABRIL, MARÍA DE LAS MERCEDES (2014). *El Enfoque de Espacio de Estado de las Series de Tiempo para el Estudio de los Problemas de Volatilidad*. Tesis Doctoral en Estadística. Universidad Nacional de Tucumán. Argentina.

ABRIL, MARÍA DE LAS MERCEDES; BLANCO MARÍA BEATRIZ (2015). *Los caminos de hierro del Norte Argentino. El Ferrocarril Central Norte Argentino durante la década de 1920. Un análisis de series de tiempo para las cargas despachadas*. Revista RinCE, N°11 Vol. 6 (septiembre 2015). Departamento de Ciencias Económicas. UNLaM. San Justo. Bs. As. Argentina.

ANDERSON, B. D. O. Y MOORE, J. B. (1979). *Optimal Filtering*. Prentice- Hall: Englewood Cliffs, New Jersey.

BOX, G. E. P. Y JENKINS, G. M. (1976), *Time Series Analysis: Forecasting and Control* (Revised edition), Holden-Day Inc.: San Francisco.

DE JONG, P. (1991). The diffuse Kalman filter. *Ann. Statist.*, 19, 1073-83.

DE JONG, P. (1989). Smoothing and interpolation with the state-space model. *J. Amer. Statist. Assoc.*, 84, 1085-88.

DE JONG, P. Y PENZER, J. R. (1998). Diagnosing shocks in time series. *J. Amer. Statist. Ass.*, 93, 796-806.

GOMEZ, V. Y MARAVALL, A. (1993). Initializing the Kalman filter with incompletely specified initial conditions. En *Approximate Kalman Filtering (Series on Approximation and Decomposition)* (G. R. Chen, Ed.). World Scientific Publ. Co: London.

HARVEY, A. C. (1989). *Forecasting, Structural Time Series Models and the Kalman Filter*. Cambridge University Press: Cambridge.

HARVEY, A. C. Y PIERSE, R. G. (1984). Estimating missing observations in economic time series. *J. Amer. Statist. Ass.*, 79, 125-31.

HARVEY, A. C. Y SHEPHARD, N. (1993). Structural time series models. En *Handbook of Statistics*, Vol. 11: Econometrics (G. S. Maddala, C. R. Rao y H. D. Vinod, Eds.), 261-302.

KALMAN, R. E. (1960). A new approach to linear filtering and prediction problems. *Trans. ASME, J. Basic Eng.*, 82D, 35-45.

KALMAN, R. E. Y BUCY, R. S. (1961). New results in linear filtering and prediction problems. *Trans. ASME, J. Basic Eng.*, 83D, 95-108.

KOHN, R. Y ANSLEY, C. F. (1986). Estimation, prediction and interpolation for ARIMA models with missing data. *J. Amer. Statist. Ass.*, 81, 751-61.

KOOPMAN, S. J. (1993). Disturbance smoother for state space models. *Biometrika*, 80, 117-26.