Departamento de Finanzas
Tecnológico de Monterrey, Campus Estado de México

# Documento de Trabajo
# (Working Paper)

TECNOLÓGICO
DE MONTERREY

WP-FZ-ITESM-CEM-No.01-2005

Serie en Econometría Financiera

## Modelos de Series de Tiempo: Una Introducción a la Técnica ARARMA-ARAR y ARIMA

Pablo López Sarabia
e-mail: plopezs@itesm.mx

TECNOLÓGICO DE MONTERREY

La serie **Documentos de Trabajo (Working Paper)** es una publicación del Departamento de Finanzas y la División de Negocios del Tecnológico de Monterrey, Campus Estado de México. Las opiniones y puntos de vista en cada una de las contribuciones son responsabilidad exclusiva de los autores. El fotocopiado del artículo y/o párrafos es permitido para fines educativos sin fines de lucro, siempre y cuando se haga mención de la fuente y el autor.

Toda la correspondencia relacionada con este documento se recibirá en:

Tecnológico de Monterrey, Campus Estado de México
Departamento de Finanzas
Aulas 5, 3er. Piso
Carretera Lago de Guadalupe Km. 3.5
Col. Margarita Maza de Juárez,
Atizapán de Zaragoza, 52926 Estado de México

Teléfono y Fax: 58-64-55-55 Ext. 3161

AUTHOR

[Autor:] Pablo López Sarabia, Profesor de Econometría del Instituto Tecnológico y de Estudios Superiores de Monterrey, Campus Estado de México

ABSTRACT

[Resumen:] El artículo describe la importancia del estudio de las series de tiempo en la economía y las finanzas; así como una introducción a la técnica de estimación ARARMA-ARAR que es una metodología alternativa a los modelos tradicionales de series de tiempo desarrollados por Box and Jenkins. La capacidad predictiva de los modelos ARARMAR-ARAR estan buena como las de los ARIMA, con la gran ventaja de que ya existe un software que permite desarrollar el algoritmo de manera sencilla y rápida a diferencia de los modelos ARIMA que siguen teniendo una parte de prueba y error a la hora de determinar el orden del modelo a estimar.

# 1. Pronóstico de Series de Tiempo: Su Necesidad e Importancia en Economía y Finanzas

## *1.1. Introducción.*

Conocer el futuro ha sido una constante desde el surgimiento del hombre, ya que este deseo forma parte de la curiosidad que todo ser humano tiene sobre su entorno presente, pero también de la interrogante permanente de si los hechos que hoy observa y vive se mantendrán de manera indefinida a lo largo del tiempo. Ironicamente, la busqueda por entender y explicar los eventos cotidianos logro un avance inimaginable para entender el futuro, ya que a través de la observación y experiencia pasada, el hombre encontro que ciertos fenomenos y hechos se repetian en determinados periodos de tiempo, pero otros se aparecian de manera intermitente, es decir, no había ninguna regularidad ni patrón especifico. Lo anterior permitio introducir la idea de incertidumbre o aleatoriedad sobre algunos fenomenos y hechos relevantes para la sobrevivencia y bienestar de los hombres y la sociedad en su conjunto; por ejemplo conocer el comportamiento del clima tomo un caracter fundamental para planificar los periodos de siembra en la agricultura y el almacenamiento de agua, así como para poder resguardarse

AUTHOR

[Autor:] Pablo López Sarabia, Profesor de Econometría del Instituto Tecnológico y de Estudios Superiores de Monterrey, Campus Estado de México

---

ABSTRACT

[Resumen:] El artículo describe la importancia del estudio de las series de tiempo en la economía y las finanzas; así como una introducción a la técnica de estimación ARARMA-ARAR que es una metodología alternativa a los modelos tradicionales de series de tiempo desarrollados por Box and Jenkins. La capacidad predictiva de los modelos ARARMAR-ARAR estan buena como las de los ARIMA, con la gran ventaja de que ya existe un software que permite desarrollar el algoritmo de manera sencilla y rápida a diferencia de los modelos ARIMA que siguen teniendo una parte de prueba y error a la hora de determinar el orden del modelo a estimar.

---

## 1. Pronóstico de Series de Tiempo: Su Necesidad e Importancia en Economía y Finanzas

### *1.1. Introducción.*

Conocer el futuro ha sido una constante desde el surgimiento del hombre, ya que este deseo forma parte de la curiosidad que todo ser humano tiene sobre su entorno presente, pero también de la interrogante permanente de si los hechos que hoy observa y vive se mantendrán de manera indefinida a lo largo del tiempo. Ironicamente, la busqueda por entender y explicar los eventos cotidianos logro un avance inimaginable para entender el futuro, ya que a través de la observación y experiencia pasada, el hombre encontro que ciertos fenomenos y hechos se repetian en determinados periodos de tiempo, pero otros se aparecian de manera intermitente, es decir, no había ninguna regularidad ni patrón especifico. Lo anterior permitio introducir la idea de incertidumbre o aleatoriedad sobre algunos fenomenos y hechos relevantes para la sobrevivencia y bienestar de los hombres y la sociedad en su conjunto; por ejemplo conocer el comportamiento del clima tomo un caracter fundamental para planificar los periodos de siembra en la agricultura y el almacenamiento de agua, así como para poder resguardarse del crudo invierno. Otro ejemplo que nos muestra lo relevante de conocer el futuro se ve claramente en las familias que desean pedir un crédito o invertir en una cuenta a largo plazo, para estas famiilias es relevante tener claro sus ingresos futuros y la tendencia general de la economía y sus variables claves para hacer una decisión eficiente.

Los ejemplos descritos en el parráfo anterior nos muestran la importancia que tiene para las familias, empresas y gobierno predecir el futuro de fenomenos y hechos que son considerados como aleatorios, y así poder realizar una buena planificación y control de dichos eventos que lleven a una toma de decisiones óptima. Al parecer hay un concenso generalizado sobre la importancia que tiene predecir el futuro, pero en cuanto a la manera de obtener dichos pronósticos se continua discutiendo la pertiencia y eficacia de las técnicas estadisticas y matematicas exisentes hoy día para tal fin. En áreas como la Economía y Finanzas la discusión se ha polarizado enormemente, ya que los pronosticos que se realizan presentan diferencias considerables con la realidad y en muhos casos estan fuera de toda lógica, orillando a planteamientos en los que se señala la imposibilidad de pronosticar el futuro y con ello estar condenados a una especie de ceguera que nunca podremos curar. Encontraposición encontramos posturas influenciadas considerablemente por un positivismo absoluto, en el cual se cree que los procedimientos matematicos permiten predecir ciertos fenomenos aleatorios con una gran precisión. Sin embargo, estas dos posturas extremas tienen claras deficiencias y pueden llevar a tomar decisiones incorrectas que generen daños innumerables y en muchos casos irreversibles.

En el caso de las personas e investigadores impregnados por un positivismo critico y con una nueva filosofía en la que las matemáticas juegan un papel importante, pero no determinante para la toma de decisiones, ya que se asumen y aceptan las limitantes de los modelos matematicos ante fenomenos aleatorios. Es dentro de esta lógica en la que surge una posición más ecleptica en las técnicas de pronóstico utilizadas en Economía y Finanzas, ya que al trabajar con fenomenos que presentan variabilidad en el tiempo, se reconoce que las predicciones tendrán cierta probabilidad de ocurrir con cierto nivel de confianza y con base en diferentes escenarios o bajo el supuesto de *ceteris paribus* (todo permanece constante) que los economistas utilizan regularmente con fines de aislar el fenomeno de análisis, pero poco realista en la práctica. Al mismo tiempo se tiene que tener en mente la posibilidad de estar ante muestras atipica y valores extremos generados por choques estocasticos que van influir nuestras predicciones, todo esto nos lleva a que cualquier técnica estadística y matematica conocida hasta el dia de hoy tiene implicita y explicitamente un nivel de error que debe ser tomado en cuenta al generar predicciones. Así, la mejor técnica de pronóstico será aquella que se acerque mejor a la realidad de manera eficiente y optima, pero también con el menor costo posible y maximizando el uso de la información disponible. Asumir esta postura permite dar un paso a delante y ser una luz en nuestra ceguera sobre el futuro, el cual no podemos predecir con seguridad, pero si con mayor certidumbre y con ello aspirar en un corto plazo a encontrar técnicas y teorías que nos permitan avisorar mejor nuestro futuro y el de la sociedad en su conjunto.

En particular el pronóstico de variables Económicas y Financieras se vuelve relevante para la toma de decisiones de los diferentes agentes económicos como las familias, empresas y el gobierno, debido a que hay una interdependencia entre ellos y su accionar, así la predicción de los niveles de desempleo, inflación, producción, ingreso, ventas, flujo de caja, tasa de interés, indice bursátil, precios de los citricos, costos de producción, número de productos defectuosos, etc.; se basan fundamentalmente en la información recopilada a través de registros sistemáticos efectuados en intervalos de tiempo fijos o permanentes que se conocen con el nombre de series de tiempo.

## *1.2. Elementos Básicos de Series de Tiempo.*

Una serie de tiempo es un conjunto de observaciones $x_t$,cada una registrada en un periodo de tiempo especifico $t$. Se considerará que una serie de tiempo es discreta si el conjunto $T_0$ de veces en que la variable fue observada en el tiempo es finito o infinito numerable. Mientras que si el conjunto $T_0$ se encuentra en un intervalo cerrado o abierto, la serie de tiempo es continua. Es importante señalar que en algunos casos podemos encontrar series de variables de tipo cualitativo que nos generarian series de tiempo binarias al ser codificados sus valores o en otros casos con caracteristicas muy particulares, un ejemplo podría encontrarse en el hecho de llevar el registro de si un equipo particular en la temporada respectiva logro pasar a la postemporada o juegos de playoff, aquí sólo es posible encontrar dos posibles observaciones: paso a playoffs y no paso a playoffs, al codificar estas respuestas nos generaría una serie binaria, pero ahora imaginemos que registramos el resultado de cada partido de la temporada, es decir el resultado el partido en cada se mana de juego, aquí tendríamos tres opciones: gano, perdio o empato, generandose así series de tiempo de tipo cualitativo o categoricas y la cual requiere algunas técnicas de análisis diferentes a las empleadas con series de tiempo cuantitativas que se desarrollan en este trabajo.

Analizar y modelar una serie de tiempo como un proceso estocástico a permitido la aparición y unificación de ciertas técnicas estadísticas y matemáticas que han probado ser muy buenas para la construcción de modelos de series de tiempo y por ende mejorar la precisión en su pronóstico. Un proceso estocástico es una familia de variables aleatorias asociadas a un conjunto índice de números reales, de forma tal que a cada elemento del conjunto le corresponde una y sólo una variable aleatoria, es decir, $\{X(t); t \in T\}$ en donde $T$ es el conjunto índice y $X(t)$ es la variable aleatoria correspondiente al elemento $t$ de $T$. Con base en lo anterior, una serie de tiempo es una sucesión de observaciones generadas por un proceso estocástico cuyo conjunto índice se toma en relación al tiempo.

El análisis de series de tiempo se centra en el hecho de utilizar la información del pasado para construir un modelo estadístico que explique los eventos anteriores y pronóstique ese fenomeno en el futuro. Los modelos más conocidos en la literatura de series de tiempo son los modelos clásicos de tipo multiplicativo y aditivo; estos modelos establecen que cualquier valor observado $(x_t)$ en una serie de tiempo es producto de la influencia de los siguientes componentes: Tendencia, Ciclo, Estacionalidad e Irregularidad o Aleatoriedad.

Modelo Multiplicativo $x_t = m_t.s_t.Y_t$

Modelo Aditivo $x_t = m_t + s_t + Y_t$

La Tendencia es un componente sistemático y expresa el movimiento progresivo hacia arriba o hacia abajo en un periodo prolongado de muchos años; este puede deberse a cambios en la tecnología, población, costo y valor. Las oscilaciones repetidas hacia arriba y hacia abajo que muestran las siguientes cuatro fases: un pico (prosperidad), una contracción (recesión), un valle (depresión) y finalmente una expansión (recuperación-crecimiento) se le denomina Ciclo y también es una componente sistemática y puede deberse a numerosas combinaciones de factores que influyen en la economía como sería el caso de las elecciones federales. Es muy común que estas dos componentes se analicen de manera conjunta como Tendencia-Ciclo $(m_t)$. La componente de Estacionalidad $(s_t)$ es considerada como sistemática y se define como las fluctuaciones periódicas regulares que se presentan dentro de cada periodo de 12 meses, año tras año; esto se debe fundamentalmente a las condiciones climatologicas, costumbres sociales y religiosas, por ejemplo en invierno siempre se espera que la demanda de combustible sea mayor a la de otras estaciones del año, así mismo el periodo de la Navidad y día de las madres aumentan las ventas considerablemente respecto a otros días o meses. Finalmente, el componente de Irregularidad o Aleatoriedad $(Y_t)$ es no sistemático y se considera como las fluctuaciones erráticas o residuales que existen después de tomar en cuenta los efectos sistemáticos, y se deben en gran medida a sucesos no previstos como huelgas, huracanes, temblores, inundaciones, asesinatos políticos etc.; que son de corta duración y no repetitivos.

Es importante notar que una serie de tiempo observada no es más que una realización de un proceso estocástico, es decir, la observación $x_t$ es generada por cierta variable aleatoria $X_t$; el elemento probabilistico de la serie de tiempo radica en este hecho, ya que otra observación podría haber sido generada por el mismo proceso estocástico diferente de la que se observó en la realidad. La distinción entre un proceso estocástico y su realización no es trivial, ya que las realizaciones se utilizan para inferir el proceso estocástico subyacente a la serie de tiempo que se este analizando; así que si deseamos describir una serie de tiempo a través de un modelo matemático, será necesario conocer la función de densidad conjunta de todas las variables aleatorias que constituyen el proceso estocástico y al menos el primer y segundo momento potencial.

En el análisis de series de tiempo es común suponer que las observaciones no provienen de variables aleatorias independientes, ya que existe una estructura de correlación entre estas que dificulta encontrar la función de densidad conjunta. Además se busca que las series de tiempo sean estacionarias, es decir, que el valor de su media y varianza no varían sistemáticamente con el tiempo y el valor de la covarianza entre dos periodos depende solamente de la distancia o rezago entre dichos periodos y no del tiempo en el cual se ha calculado la covarianza. En sentido estricto la definición anterior se refiere a la estacionariedad en sentido débil, la cual es más que suficiene para los fines de nuestro análisis y la cual se puede expresar de manera formal como:

Sea $\{X_t\}$ un proceso estocástico que describe una serie de tiempo con $EX_t^2 < \infty$ y función de medias de $\{X_t\}$ igual a $\mu_x(t) = E(X_t)$ y función de covarianza o autocovarianza (ya que se esta utilizando la misma variable en rezagos diferentes) de $\{X_t\}$ dada por $\gamma_x(t+h,t) = Cov(X_{t+h}, X_t) = E\left[(X_{t+h} - \mu_x(t+h))(X_t - \mu_x(t)\right]$, se dice que es un proceso estocástico estacionario débil, si:

(i) $\mu_x(t)$ es independiente de $t$.

(ii) $\gamma_x(t+h,t)$ es independiente de $t$ para cada $h$. Cuando $h = 0$ se tiene la varianza de $Var(X_t) = \sigma^2$ donde $t$ y $h$ son números enteros.

La estacionariedad estricta de una serie de tiempo es definida por la condición que $(X_1, ..., X_n)$ y $(X_{1+h,...,}X_{n+h})$ tienen la misma función de densidad conjunta para todo entero $h$ y $n > 0$.

Los modelos más simples que podemos encontar al trabajar con series de tiempo son los modelos de media cero, dentro de los cuales está el ruido IID en el cual no hay tendencia-ciclo ni componente estacional y las observaciones son independientes e identicamente distribuidas (IID) de variables aleatoria con media cero. Este supuesto permite obtener de manera sencilla la función de densidad conjunta, ya que al ser IID sólo se necesita conocer la función de densidad de una de las variables y multiplicar está tanta veces como variables aleatorias se consideren. Además , en este modelo no hay correlación entre las variables, ya que al ser independientes la covarianza o en este caso la autocovarianza es cero y por tanto la corelación también es nulas entre los diferentes rezagos. Es decir, si $\{X_t\}$ es un ruido IID y $EX_t^2 < \infty$ , se cumple la definición de estacionalidad débil, ya que la media $EX_t = 0$ para todo $t$. Debido al supuesto de independencia el ruido IID no depende de $t$ y al existir el segundo momento finito se puede concluir que es estacionaria $\{X_t\} \sim IID(0, \sigma^2)$.

Un concepto fundamental en el análisis de series de tiempo es el de ruido blanco ($WN$ de las siglas en inglés) que se define de la siguiente manera: si $\{X_t\}$ es una secuencia de variables aleatorias no correlacionadas, cada una con medio cero y varianza $\sigma^2$, entonces $\{X_t\}$ es estacionaria con la misma función de covarianza o autocovarianza como el ruido IID. A esta secuencia se le denomina ruido blanco con media cero y varianza $\sigma^2$,es decir, $\{X_t\} \sim WN(0, \sigma^2)$. Es claro que cada secuencia $IID(0, \sigma^2)$ es siempre un $WN(0, \sigma^2)$,pero no de manera reciproca.

Por último, es necesario introducir la idea de una caminata aleatoria o random walk $(S_t; t = 0, 1, 2, ...)$ que se obtiene mediante la suma de variables aleatorias IID. Así, una caminata aleatoria con media cero se obtiene definiendo $S_0 = 0$ y $S_t = X_1 + X_2 +, ..., +X_t$ para $t = 1, 2, ...$ donde $\{X_t\}$ es ruido IID, $IID(0, \sigma^2)$. Si $\{S_t\}$ es una caminata aleatoria con las caracteristicas señaladas, entonces $ES_t = 0$, $ES_t^2 = t\sigma^2 < \infty$ para toda $t$ y para $h \geq 0$. Debido a que el segundo momento depende de $t$ la función de covarianza o autocovarianza también dependerá de $t$, violando la definición de estacionalidad débil, por lo tanto $\{S_t\}$ será no estacionaria.

Una prueba sencilla de estacionariedad está basada en la denominada función de autocorrelación (ACF) que se define de la siguiente manera: Sea $\{X_t\}$ una serie de tiempo estacionaria, la ACF de $\{X_t\}$ es $\rho_x(h) \equiv \frac{\gamma_x(h)}{\gamma_x(0)} = Cor(X_{t+h}, X_t)$ donde $\rho_x(h)$ indica la ACF del rezago $h$. Es claro que cuando $h = 0$, $\rho_x(0) = 1$. Puesto que la covarianza y la varianza están en las mismas unidades, $\rho_x(h)$ es un número sin unidad de medida, o puro, el cual se encuentra entre -1 y 1, al igual que cualquier coeficiente de correlación. Si se gráfica $\rho_x(h)$ frente a $h$ se obtiene el correlograma poblacional. Al tener en la práctica sólo las realizaciones de un proceso estocástico $x_1, x_{2,...,} x_n$ que conforman una serie de tiempo con media $\overline{x} = \frac{1}{n}\sum_{t=1}^{n} x_t$ y función de autocorrelación muestral $\widehat{\gamma}_x(h) = \sum_{t=1}^{n-|h|}(x_{t+|h|} - \overline{x})(x_t - \overline{x})$, $-n < h < n$. Entonces, la función de autocorrelación muestral esta dada por $\widehat{\rho_x}(h) \equiv \frac{\widehat{\gamma_x}(h)}{\widehat{\gamma}_x(0)}$ , $-n < h < n$.

Si se piensa que los datos de nuestra serie de tiempo provienen de un proceso estocástico estacionario $\{X_t\}$ puramente aleatorio, su autocorrelación en cualquier rezago $h > 0$ es cero. Al tener las realizaciones de la serie de tiempo podremos calcular la ACF muestral que es una buena estimación de la ACF poblacional $\{X_t\}$. Dicha estimación de la ACF puede sugerir cual de los muchos modelos estacionarios de series de tiempo posibles, es el candidato mas viable para representar la dependencia en los datos. Lo anterior, se verifica de manera visual al grafica la ACF muestral que se encuentra acotada entre -1 y 1, si los valores de la ACF muestral para rezagos $h > 0$ están muy cercanos a cero, es decir, $\widehat{\rho_x}(h) \cong 0$ y estos caen dentro del intervalo de confianza que a continuación se definirán, la serie de tiempo puede considerarse estacionaria y puramente aleatoria. Bartlett (1946) demostró que si una serie de tiempo es puramente aleatoria puede ser expresada como ruido $IID(0, \sigma^2)$ o ruido blanco $WN(0, \sigma^2)$ y que los coeficientes de autocorrelación muestral $\widehat{\rho_x}(h), h > 0$ están distribuidos en forma aproximadamente normal $N(0, \frac{1}{n})$ con media cero y varianza $\frac{1}{n}$ para un tamaño de muestra $n$ grande. Por lo tanto, si se quiere evaluar la significancia estadística de cada $\widehat{\rho_x}(h)$,a un nivel de confianza del 95%, se debe probar la hipótesis nula $H_0 : \rho_x(h) = 0$ y alternativa $H_1 : \rho_x(h) \neq 0$; bajo la regla de decisión de que $\rho_x(h) \in \left[-1.96(\frac{1}{\sqrt{n}}), 1.96(\frac{1}{\sqrt{n}})\right]$, donde 1.96 es el cuantil 0.975 de una distribución normal estándar y $(\frac{1}{\sqrt{n}})$ es la desviación estándar de $S.E.(\widehat{\rho_x}(h))$. Así que si un $\widehat{\rho_x}(h)$ particular está dentro del intervalo de confianza construido, no se rechaza la hipótesis nula; mientras que si sucede lo contrario se rechaza la hipotesis nula.

En el caso de que los valores $\widehat{\rho_x}(h)$ de la ACF tiendan muy rápidamente a cero, pero no todos se encuentren dentro del intervalo de confianza al 95% definido líneas arriba (no más del 5% de los valores de $\widehat{\rho_x}(h)$ deben salirse del intervalo) la serie de tiempo puede ser considerada como estacionaria, pero no puramente aleatoria y por tanto se deberá buscar un modelo diferente a simplemente considerar un proceso estocástico de ruido $IID(0, \sigma^2)$ o ruido blanco $WN(0, \sigma^2)$. Por todo lo comentado, la ACF se convierte en un elemento fundamental para encontrar y determinar nuestro modelo estadístico que representará y explicará la serie de tiempo de interés.

Si se desea probar la hipotesis conjunta de que todos los coeficientes de autocorrelaión $\rho_x(h) = 0$.se puede utilizar la estadística $Q$ desarollada por Box y Pierce y que se define como: $Q = n\sum_{h=1}^{m} \widehat{\rho_x}^2(h)$ , donde $n$ es el tamaño de la muestra y $m$ es la longitud del rezago. La estadística $Q$ está repartida aproximadamente ($n$ grande) como la distribución ji-cuadrada ($\chi^2_{m,1-\alpha}$) con $m$ grados de libertad y un nivel de confianza de $1-\alpha$, la regla de decisión consiste en comparar la $Q$ calculada contra el valor crítico de $Q_{m,1-\alpha}$ ; si $Q > Q_{m,1-\alpha}$ entonces se rechaza la hipotesis nula de que todos los $\rho_x(h)$ son cero, ya que alguno de ellos debe ser diferente de cero. Una variante de la estadística $Q$ de Box y Pierce es la estadística Ljung-Box (LB) desarrollada en 1978 se define como: $LB = n(n+2)\sum_{h=1}^{m}\left(\frac{\widehat{\rho_x}^2(h)}{n-h}\right) \sim \chi^2_{m,1-\alpha}$. Aunque la estadística $Q$ y $LB$ siguen una distribución $\chi^2_{m,1-\alpha}$ con $n$ grande, se ha encontrado que la estadística $LB$ posee propiedades que la hacen operar de manera muy eficiente ante muestras pequeñas, por lo que se considera que ésta es estadísticamente más potente que el estadístico $Q$.

A lo largo de este trabajo se utilizan varios operadores con el fin de facilitar la notación y el manejo de algunos polinomios. El primero de estos operadores es el de *retraso o backward (B),* el cual se define mediane la relación $BX_t = X_{t-1}$ para toda $t$. Este operador puede aplicarse sucesivamente sobre una variable, si el operador es utilizado $k$ veces llegaremos a la siguiente expresión general $B^kX_t = X_{t-k}$ para $k = 0, 1, 2, ...$ y toda $t$ . Es importante tener en cuenta que cada veces que aplicamos el operador de retraso perdemos una observación y cuando lo aplicamos $k$ veces se pierde el mismo número de observaciones, reduciendo nuestra serie de $N$ observaciones a $N-k$.

Otro operador de uso frecuente y que esta relacionado con el operador de retraso $B$ es el *operador diferencia* $(\bigtriangledown)$. Este operador expresa la siguiente relación $\bigtriangledown X_t = X_t - X_{t-1}$ para toda $t$. La relación que liga a los operadores de retraso y diferencia es $\bigtriangledown = 1 - B$ , es decir, $\bigtriangledown X_t = (1-B)X_t$ y de manera equivalente al operador de retraso, el operador diferencia también puede ser aplicado sucesivamente obteniendo la siguiente expresión para $\bigtriangledown^k$: $\bigtriangledown^k X_t = \sum_{h=0}^{k} \frac{k!}{h!(k-h)!}(-1)^h X_{t-h}$ para toda $k = 0, 1, 2, ...$ y toda $t$. Esta expresió puede obtenerse a través del Teorema del Binomio, ya que $\bigtriangledown^k$es un binomio elevado a la $k$-ésima potencia, es decir, $\bigtriangledown^k X_t = (1-B)^k X_t$.

## *1.3. Enfoques para Modelar Series de Tiempo.*

Existe un enfoque general para modelar y análizar series de tiempo y que se pueden resumir en los siguientes pasos:

1. Gráficar la serie y examinar de manera visual si hay una tendencia-ciclo, un componente estacional, algunos picos y cambios significativos en el comportamiento de la serie y detectar valores distantes (outlying-outliers).Además, la gráfica debe sugerir la posibilidad de representar los datos de la serie de tiempo como una realización del proceso estocástico $X_t = m_t + s_t + Y_t$, donde $m_t$ es la componente de tendencia-ciclo, $s_t$ es una función conocida y que representa el componente estacional, y $Y_t$ es una componente de ruido aleatorio que es estacionaria en sentido débil.

2. Remover la tendencia y el componente estacional para conseguir que los residuales sean estacionarios. Para lograrlo en muchas ocasiones es necesario transformar los datos de manera previa. Por ejemplo, si la magnitud de las fluctuaciones parecen crecer de manera lineal en forma considerable a lo largo el tiempo, se recomienda transformar la serie aplicandole un logaritmo natural a cada valor de la serie y con ello estabilizar las fluctuaciones. La manera de eliminar o aislar la tendencia-ciclo y la estacionalidad se basa generalmente en diferenciar los datos, pero esto puede variar según el enfoque particular que se tome y los cuales se analizará más adelante.

3. Utilizar algunas técnicas para aislar y estimar la componente de tendencia-ciclo, estacionalidad y de ruido aleatorio que se espera que la componente aleatoria sea estacionaria.Escoger un modelo para ajustar los residuales haciendo uso de la muestra o realizaciones y de la función de autocorrelación muestral.

4. El pronóstico de algunos valores se logrará con base en la predicción de los residuales y entonces invertir la transformación en caso de haberla hecho y así obtener el pronóstico de la serie original $\{X_t\}$.

5. Un enfoque alternativo que ha tomado mucha fuerza en el campo de la Ingeniería es expresar la serie de tiempo en términos de sus componentes de Fourier, es decir, en movimientos sinoidales de diferente frecuencia que pueden ser expresados a través de la función seno y coseno.

6. Otro enfoque desarollado de manera éxitosa por Box y Jenkins (1976), se basa en aplicar el operador diferencia repetidamente a la serie $\{X_t\}$ hasta que las observaciones diferenciadas se parezcan a una realización de alguna serie de tiempo estacionaria $\{W_t\}$. En el momento que se cuente con una serie de tiempo estacionaria se puede aplicar la teoria de procesos estocásticos estacionarios para modelar, analizar y pronósticar $\{W_t\}$ y por lo tanto, obtener el proceso original y analizar la y modelar diferencias

# 2. Pronóstico de Series de Tiempo: El Enfoque de los Modelos ARARMA-ARAR.

## *2.1. Introducción.*

El algortimo que se utiliza en esta sección y el cual es conocido como ARARMA o ARAR, y fue elaborado en primera instancia por Parzen en 1982 y perfeccionado por Newton y el mismo Parzen en 1984. Este metódo de modelación y pronóstico es aplicado automaticamente, pero esto no significa que haya formulas rutinarias, ya que el modelo se basa en una filosofía flexible que ofrece varios modelos y diagnosticos cualitativos y cuantitativos que pueden ser considerados para verificar el ajuste del modelo. El enfoque de modelos ARARMA o ARAR es aplicable en todos los campos de la ciencia donde se utilicen series de tiempo de tipo cuantitativo.

Una discusión relevante y aún no conclusa al pronosticar series de tiempo, se centra en el hecho de si las predicciones de corto y largo alcance, requieren de metodos diferentes para obtener un pronostico satisfactorio. La técnica de los modelos ARARMA o ARAR permite hacer los dos tipos de pronosticos, ya que utiliza un modelo autorregresivo parsimonioso no estacionario cuyos residuales $\widetilde{Y}_t$ son ajustados a través de una autorregresión estacionaria

## *2.2. Enfoque de Modelos Iterativos en el Análisis de Series de Tiempo (Newton y Parzen 1984)*

El problema de pronosticar los valores de una serie de tiempo a través de sus valores pasados tiene una amplia literatura, que propone diferentes enfoques. Dentro de estos destacan los denominados modelos iterativos que permiten hacer un ajuste automatico de una serie muestral, entre diversos modelos, y los cuales han tenido un gran desarrollo en los últimos años.

Una serie de memoria larga requiere un modelo no estacionario con periodo, ciclo y componente tendencial . Una serie de memoria corta requiere un modelo estacionario, el cual es un filtro lineal que relaciona la serie de tiempo con sus choques aleatorios o innovations.

El filtro lineal es un AR, MA o ARMA. El modelo que nosotros ajustamos a una serie de tiempo $Y_t$ es un modelo iterativo de la forma:

$$Y_t \rightarrow \widetilde{Y}_t \rightarrow \varepsilon_t$$

Si es necesario transformar una serie de memoria larga $Y_t$ a una serie de memoria corta $\left\{\widetilde{Y}_t\right\}$ , La $\widetilde{Y}_t$ seleccionada debe satisfacer una de las siguientes formas:

1. $\widetilde{Y}_t = Y_t - \widehat{\phi}(\widehat{\tau})Y_{t-\tau}$
2. $\widetilde{Y}_t = Y_t - \widehat{\phi}_1 Y_{t-1} - \widehat{\phi}_2 Y_{t-2}$
3. $\widetilde{Y}_t = Y_t - \widehat{\phi}_1 Y_{t-\tau-1} - \widehat{\phi}_2 Y_{t-\tau-2}$

Usualmente, $\widetilde{Y}_t$ es de memoria corta y puede ser transformada a un ruido blanco o no memory generandose una nueva serie de tiempo $\{\varepsilon_t\}$ que se obtuvo aproximando un esquema autorregresivo AR($m$) cuyo orden $m$ es determinado por un criterio de orden. En este artículo se asume que $\widetilde{Y}_t$ siempre es de memoria corta. Por lo que se puede modelar mediante un proceso autorregresivo estacionario.

Determinar el mejor rezago o retraso de $\widehat{\tau}$, se uso una autorregresión no estacionaria, por lo que ninguno fija un máximo de rezagos $m$ y escoge $\widehat{\tau}$ como el rezago que minimiza sobre todos los $\tau$. $\sum_{t=m+1}^{T}\left[Y_t - \phi_\tau Y_{t-\tau}\right]^2$ escoger una $\widehat{\tau}$ como el rezago que minimiza sobre todos los $\widehat{\tau}$ : $\frac{\sum_{t=\tau+1}^{T}[Y_t-\phi_\tau Y_{t-\tau}]^2}{\sum_{t=\tau+1}^{T} Y_t^2}$ para cada $\tau$ , uno determina $\phi_\tau$ , y entonces uno determina $\widehat{\tau}$ (el valor optimo de $\tau$) como el valor que minimiza : $ERR(\phi,\tau) = \frac{\sum_{t=\tau+1}^{T}[Y_t-\phi Y_{t-\tau}]^2}{\sum_{t=\tau+1}^{T} Y_t^2}$ ó $ERR(\phi,\tau) = \frac{\sum_{t=m+1}^{T}[Y_t-\phi Y_{t-\tau}]^2}{\sum_{t=m+1}^{T} Y_t^2}$

La decisión de si la serie de tiempo es de memoria larga o no, se basa en el valor del $ERR(\phi,\tau)$ . Una regla que puede ser usada es ver si $ERR(\phi,\tau) < \frac{8}{T}$ , la serie de tiempo es considerada como de memoria larga.

Para el proposito de pronosticar es suficiente adoptar para $\widetilde{Y}_t$ como un modelo autorregresivo estacionario conveniente de orden $m$ cuyos coeficientes $\alpha_1, ..., \alpha_m$ son estimados por medio de las ecuaciones Yule-Walker en la función de correlación $\widehat{\rho}(v)$ de $\widetilde{Y}_t$. En este trabajo el modelo adoptado para todas las series de tiempo fue de la forma: $\widetilde{Y}_t = Y_t - \phi(\tau)Y_{t-\tau}$ y $\widetilde{Y}_t + \alpha_1\widetilde{Y}_{t-1} + ... + \alpha_m\widetilde{Y}_{t-m} = \varepsilon_t$ , la varianza de los residuales es denotada por $RVY = \frac{\sum_{t=m+1}^{T}\widetilde{Y}_t^2}{\sum_{t=m+1}^{T} Y_t^2}$ y $RVYT = \frac{\sum_{t=\tau+1}^{T}\varepsilon_t^2}{\sum_{t=\tau+1}^{T}\widetilde{Y}_t^2}$, los últimos 18 puntos de la gráfica de $Y$ y $\widetilde{Y}$ no representa valores observados de estas series pero los valores pronosticados de horizonte $h = 1$ a $18$ .

## 2.3. *Modelación de una Serie de Tiempo Univariada y Pronóstico, Enfoque Automatico Usando Modelos ARARMA-ARAR.*

El modelo que proponemos se ajusta en general a una serie de tiempo $Y_t$ a través de un modelo iterativo (cuyas funciones de transformación se simbolizan con G y g$_\infty$ $Y_t \rightarrow G \rightarrow \widetilde{Y}_t \rightarrow g_\infty \rightarrow \varepsilon_t$ ruido blanco (WN) donde $\widetilde{Y}_t$ es el resultado de una transformación de memoria corta seleccionada para transformar una serie de memoria larga en una serie de memoria corta, y $g_\infty$ es un whitening filter , el cual es de alguna manera una aproximación de un filtro AR o un filtro ARMA. Parzen (1982) introdujo la terminologia ARARMA para las series de tiempo iterativas modeladas con G que se determina mediante una estimación de autorregresión no estacionaria. Un esquema ARIMA introducido por Box y Jnekins corresponden a un puro operador de diferencias para G. Un análisis autorregresivo por medio de las ecuaciones de Yule-Walker producen un esquema autorregresivo estacionario; y un esquema autorregresivo no estacionario es uno en el cual es ajustado a través de estimar sus coeficientes por minimos cuadrados ordinarios. Identificar el modelo final o el whitening filter general de una serie de tiempo, uno debería determinar el tipo de memoria de su modelo e identificar un modelo iterativo para una serie de tiempo.

Una teoría ratificada de inferencia estadística es disponible sólo para series de memoria corta, la cual es ergodica. La modelación de una serie de memoria corta a través de un filtro blanqueador puede ser vista como una ciencia, y puede hacerse de manera semiautomatica. Dada una muestra de una serie estacionaria de memoria corta $\widetilde{Y}_t$, nuestro procedimiento para modelar en el domino de tiempo es calcular a través de aproximar un esquema autorregresivo.

1. La forma de la función de autocorrelación muestral es de la forma: $\widehat{\rho}(v) = \frac{\sum_{t=1}^{T-v} \widetilde{Y}_t \widetilde{Y}_{t+v}}{\sum_{t=1}^{T} \widetilde{Y}_t^2}$, pero no basada alguna decisión sobre este, o sobre las correlaciones parciales. Mejor dicho , calcular un esquema que se acerque un autorregresivo.
2. Resolver de manera sucesiva el orden $m = 1, 2, ...$ de las ecuaciones de Yule-Walker para los coeficientes autorregresivos $\alpha_{1m}, ...\alpha_{mm}$ y la varianza residual $\sigma_m^2$ .
3. Usanso un autorregresivo criterio de determinación de orden.que puede ser el AIC o CAT, para determinar $\widehat{m}(1)$, y $\widehat{m}(2)$, el mejor y segundo mejor orden de la aproximación de un esquema autorregresivo.
4. Calcular PVH(h), la función horizonte de la varianza pronosticada para esa idea y suministrar sobre el tipo de memoria y el tipo ARMA

de la serie de tiempo. Calcular los horizontes HOR1, HOR2 usando una aproximación a un esquema AR de orden $\widehat{m}(1)$, y $\widehat{m}(2)$.

5. Calcular un subconjunto de modelos AR.

6. Calcular un subconjunto de modelos ARMA.

Uno también puede calcular varias funciones de densidad espectral y funciones de distribución espectral, uno necesita una idea adicional de domino espectral.

El diagnostico de que exista una serie de tiempo de memoria larga puede ser hecha de manera semiautomatica, Muchos criterios pueden ser posibles para diagnosticar el tipo de memoria de una serie usando a)correlaciones, b)densidad espectral, c)varianzas pronosticadas de un proceso autorregresivo , d) función horizonte de varianzas pronosticadas , e)función de densidad espectral y f) S-PLAY diagnostico. Las definiciones de estas son dadas a bajo en terminos de los parametros poblacionales, asumiendo una serie de tiempo estacionaria. En la práctica, el diagnostico está basado sobre una muestra de manera analoga de estos parametros.

El horizonte de la varianza pronosticada PVH(h), h=1,2,..., es definido en términos del error cuadratico medio normalizado de la predicción de memoria infinita de h pasos adelante:

$\sigma^2_{h,\infty} = E\left\{|Y^v(t+h/t)|^2\right\}$, $Y^v(t+h/t) = Y_t - Y^v(t+h/t)$, $Y^\mu(t+h/t) = E\left\{Y(t+h)/Y(t), Y(t-1), ...\right\}$

Una formula para $\sigma^2_{h,\infty}$ es obtenido a través de introducir una representación MA($\infty$) de :

$Y_t = \varepsilon_t + \beta_1\varepsilon_{t-1}+, ...$ . Entonces $\sigma^2_{h,\infty} = \sigma^2_\infty\left\{1+\beta_1^2+...+\beta_{h-1}^2\right\}$ . La gráfica de $\sigma^2_{h,\infty}$ aumenta monotonamente desde $\sigma^2_\infty$ en $h=1$ a 1 cuando $h$ tiende a $\infty$. Nosotros definimos PVH(h)=$1-\sigma^2_{h,\infty}$ , $h=1,2,...$ y definimos el horizonte HOR como el valor más pequeño de $h$ para el cual PVH(h)$\leq$ 0.05, por lo cual $\sigma^2_{h,\infty} \geq 0.95$ .

Los coeficientes del MA infinito $\beta_k$ son estimadas a través de invertir la función de transferencia $g_m(z)$ de una aproximación autorregresiva esquema obtenido para $k=1,2,...$

$\alpha_0\beta_k + \alpha_1\beta_{k-1} + ... + \alpha_k\beta_0 = 0$

Por que HOR=$\infty$ , nuestra media HOR es comparativamente grande: experimentos conducen a nosotros a concluir que uno debería comparar HOR con el orden ORD del esquema que aproximación autorregresivo. Sea HOR/ORD que denota la razón de HOR a ORD; identificar la serie de tiempo como sigue: si HOR/ORD$\leq$ 1, entonces MA(q), con q$\leq$HOR$-1$, Si

HOR/ORD$\geq 4$ y PVH decae lentamente entonces memoria larga, Si PVH declina suavemente y exponencialmente, entonces un AR(p) es indicado. Si PVH tiene inclinación o dobla BENDS, entonces un ARMA. Si PVH tiene muchos niveles extendidos con periodo $\tau$, entonces un modelo ARMA es indicado de la forma:
$Y_t = \frac{I+B_1L+B_2L^2+...+B_qL^q}{I-\alpha_\tau L_\tau}\varepsilon_t$.

La identificación final del orden p y q debe ser mediane la estimación de los parametros o mediante el uso de S-arreglos. La determinación de la más apropiada y bondadosa transformación de $Y$ a $\widetilde{Y}$, donde $Y$ es una serie de memoria larga y $\widetilde{Y}$ es una de memoria corta debe ser inevitablemente involve la naturaleza isica de las observaciones de la serie de tiempo. Un enfoque semiautomatico puede ser desarrollado a través de considerar los siguientes ejemplos de series de memoria larga.

Una serie de tiempo $Y_t, t = 0, 1, ...,$ es llamada periodica con periodo $\tau$, Si $Y_{t+\tau} - Y_t = 0$, para toda $t$. Si sigue una tendencia lineal $Y_t = a + bt$ , si para toda $t$. $Y_{t+1} - Y_t = b$ una constante. Es una harmonica pura de periodo $\tau$ si para toda $t$ . $Y_t - \phi Y_{t-1} + Y_{t-2} = 0$ donde $\phi = 2\cos\frac{2\pi}{\tau}$, entonces: $Y_t = A\cos\frac{2\pi}{\tau}t + B\text{sen}\frac{2\pi}{\tau}t$.

Unas transformaciones bondadosa de memoria corta , es natural considerar:

1. $\widetilde{Y}_t = Y_t - \widehat{\phi}(\widehat{\tau})Y_{t-\widehat{\tau}}$
2. $\widetilde{Y}_t = Y_t - \widehat{\phi}_1 Y_{t-1} - \widehat{\phi}_2 Y_{t-2}$
3. $\widetilde{Y}_t = Y_t - \widehat{\phi}_1 Y_{t-(m-1)} - \widehat{\phi}_2 Y_{t-m}$

Cuyos coeficientes $\tau, \widehat{\phi}(\tau), \widehat{\phi}_1, \widehat{\phi}_2$ son determinados adaptively de los datos. Nuestra primera elección es (1); el rezago $\widehat{\tau}$ es seleccionado para minimizar sobre $\tau$. $ERR(\phi, \tau) = \frac{\sum_{t=\tau+1}^{T}\left[Y_t - \widehat{\phi}(\tau)Y_{t-\tau}\right]^2}{\sum_{t=\tau+1}^{T} Y_t^2}$ y $\widehat{\phi}(\tau)$ es seleccionado para minimizar sobre $\phi(\tau)$. $\sum_{t=\tau+1}^{T}\left[Y_t - \phi(\tau)Y_{t-\tau}\right]^2$ . La función de correlación estacionaria $\rho(\tau)$ de $\{Y_t, t = 1, 2, ..., T\}$ es definido por medio de : $\widehat{\rho}(\tau) = \frac{\sum_{t=1}^{T-\tau} Y_t Y_{t+\tau}}{\sum_{t=1}^{T} Y_t^2}$. Definimos SSQ($v$)= $\sum_{t=1}^{v} Y_t^2$. Uno puede mostrar que: $\widehat{\phi}(\tau) = \widehat{\rho}(\tau)\frac{SSQ(T)}{SSQ(T-\tau)}, ERR(\tau) = 1 - \left|\widehat{\phi}(\tau)\right|^2 \frac{SSQ(T-\tau)}{SSQ(T)-SSQ(\tau)}$, el más significativo rezago $\widehat{\tau}$ es definido como el valor que minimiza el $ERR(\tau)$. :

*2.4. Filosofía de los Modelos ARARMA-ARAR.*

A lo largo del artículo me he referido al concepto de pronóstico, el cual debe entenderse como la proyección del pasado (observaciones hechas en $t_0, t_{0-1,...}$) hacia el futuro ( a través de calcular $t_0 + 1, t_0 + 2, ...$), a través de un modelo que se basa y ajusta en los datos observados en el pasado. Una serie de tiempo en la que las obervaciones pasadas no dan información acerca del futuro se denomina ruido blanco ($WN$), el cual ya ha sido definido previamente de manera formal.

Un modelo para una serie de tiempo $\{Y_t\}$ se puede encontrar aplicando un filtro que transforme la serie $\{Y_t\}$ en una serie de ruidos blancos $\{\varepsilon_t\}$, este filtro se conoce como "whitening filter" (puede traducirse como filtro blanqueador). Simbolicamente se puede describir este proceso de la siguiente manera:

$\{Y_t\} \longrightarrow \boxed{Whitening\ filter} \longrightarrow \{\varepsilon_t\}$

Para estimar el whitening filter se utiliza frecuentemente una formula parametrica, en el caso particular de un esquema ARMA(p,q), tenemos lo siguiente: $\sum_{j=0}^{p} a_j Y_{t-j} = \mu + \sum_{k=0}^{q} b_k \varepsilon_{t-k}$ donde $a_0 = b_0 = 1$. Algunos estadísticos e investigadores buscan y escogen el filtro que tenga el menor número de parametros posible, este hecho se conoce como el principio de parsimonia de un modelo. La importancia de que el filtro utilizado cumpla con la parsimonia se basa en el hecho de evitar un sobre ajuste, ya que la experiencia indica que un ajuste demasiado bueno es una señal o presagio de que los criterios para medir la precisión de los pronósticos fallarán.

En la teoría y en la práctica un esquema AR y ARMA son modelos apropiados solamente para series de tiempo que se denominan como "short memory" (memoria corta). Una serie de tiempo de memoria corta es aquella para la cual sólo es posible obtener pronósticos de corto alcance; matemáticamente es una serie de tiempo que puede ser modelada como una serie estacionaria no deterministica con media cero. Si una serie de tiempo presenta una componente de tenencia-ciclo, entonces es posible hacer pronósticos de largo o mediano alcance. Estas series de tiempo se llaman "long memory" (memoria larga).

El enfoque utilizado en esta sección para modelar y prónosticar una serie de tiempo se basa en el concepto de "memory" (memoria). Parzen (1982) describe tres tipos de memoria: a) *no memory o white noise*, b) *short memory* y c) *long memory*. Así, el primer paso en la construcción de un modelo es clasificar a nuestra serie en uno de los tres tipos de memoria propuestos.

Para determinar la memoria de una serie de tiempo en la práctica se considera una transformación lineal de la serie de la forma $\widetilde{Y}_t = Y_t + a_1 Y_{t-1} + \ldots + a_m Y_{t-m}$ , y se calcula $\widetilde{\sigma}^2 = \frac{\sum_{t=1+m}^{T} |\widetilde{Y}_t|^2}{\sum_{t=1+m}^{T} |Y_t|^2}$ . Entonces tendremos una serie de *no memory* si todos los $\widetilde{Y}_t$ satisfacen la siguiente regla que podemos considerar como adecuada: $\log\widetilde{\sigma}^2 + \frac{2m}{T} > \frac{-1}{T}$; mientras que tendremos un proceso *long memory* si algunos $\widetilde{Y}_t$ satisfacen la siguiente regla: $\widetilde{\sigma}^2 < \frac{8}{T}$. Finalmente, una serie de tiempo se considera *short memory*, si no cumple ninguna de las reglas señaladas, es decir, sucede de otra manera.

El valor minímo de $\widetilde{\sigma}^2$(sobre todos los coeficientes $a_{1,\ldots,}a_m$) se escribe como $\widehat{\sigma}^2_m$, el subíndice $m$ describe el número de variables predictoras $Y_{t-1},\ldots, Y_{t-m}$. $\widehat{\sigma}^2_m$ puede interpretarse como un estimador del minímo error cuadrático medio de memoria $m$ , lo que permite dar un paso adelante en la predicción del error. La secuencia $\widehat{\sigma}^2_m$ tiene valor 1 en $m = 0$ y decrece monotonamente (como $m$ tiende a $\infty$) a un límite definido por: $\widehat{\sigma}^2_\infty = lim_{m\to\infty}\widetilde{\sigma}^2_m$ , el cual estima el minímo error cuadratico medio de memoria infinita. El valor de $\widehat{\sigma}^2_\infty$, la curva de $\widehat{\sigma}^2_m$ como una función de $m$ y la manera en la cual $\widehat{\sigma}^2_m$ converge a $\widehat{\sigma}^2_\infty$ proporciona los elementos necesarios para identificar posibles modelos para nuestra serie de tiempo con parametros finitos que se acercan aproximadamente a los verdaderos modelos de parametros infinitos.

Para una serie de tiempo de memoria corta, un modelo de whitening filter es simplemente aplicar un esquema AR(p) o en su caso un ARMA (p,q). La literatura de series de tiempo parece respaldar la idea de que una serie de memoria corta, puede ser modelada muy bien a través de un AR($\widehat{m}$), es decir, un esquema autorregresivo finito que se aproxima en gran medida a un proceso autorregresivo infinito AR($\infty$). El orden $\widehat{m}$ es determinado a través de diversos criterios como el AIC, AICC o CAT.

El proceso de ajuste que nos aproxime a un AR($m$) se apoya en las siguientes pruebas:

1. Prueba de *no memory* o ruido blanco, se basa en el hecho de que $\widehat{m} = 0$
2. Prueba de *long memory* o memoria larga, se basa en los siguientes criterios: que $\widehat{\sigma}^2_m$ debe estar muy cerca de cero o el horizonte HOR debe ser grande o la forma del CAT debe indicar un *long memory*.
3. Prueba para determinar el orden de $m$, se basa en determinar los valores de $\widehat{m}$ y $\widehat{m}(2)$ que nos aproximan a un esquema AR cuyos residuales deben verificarse para garantizar que son ruido blanco.

Para una serie de tiempo de memoria larga, un modelo de whitening filter se basa en aplicar dos filtros de manera sucesiva, es decir, un proceso de memoria larga se convierte primero en uno de memoria corta y a ésta serie transformada se le aplica otro filtro que permita obtener ruido blanco. Dicho proceso se puede describir en el siguiente esquema:

$\boxed{\text{Long memory o memoria larga } Y_t} \rightarrow \boxed{\text{Short memory o memoria corta } \widetilde{Y}_t} \rightarrow$

$\boxed{\text{White noise o ruido blanco } \varepsilon_t}$

La definición $\widetilde{Y}_t$ no es única, en el sentido de que el filtro definido transforma $Y_t$ a $\widetilde{Y}_t$ a través de escoger el mejor filtro bajo el principio de parsimonia, haciendo que $\widetilde{Y}_t$ sea una serie de memoria corta. Es importante señalar que no existe una prueba matemática rigurosa para garantizar la existencia de una serie de memoria corta. Sin embargo, un whitening filter para una serie de memoria larga es unica, esto se interpreta como si una serie de filtros sucesivos quitarán la tendencia-ciclo y la estacionalidad generando ruido blanco. De esta manera se tiene una técnica automatica para modelar series de tiempo de memoria corta, donde la arbitrariedad aplicada en la transformación de $\widetilde{Y}_t$ será compensada por la transformación que se hace a $\varepsilon_t$ .

Una manera recomendable para encontrar una transformación de $Y_t$ en $\widetilde{Y}_t$ es ajustar a uno o dos rezagos un proceso autorregresivo no estacionario. El modelo para $Y_t$ consiste de un AR no estacionario seguido de un ARMA estacionario. Por lo anterior, esté enfoque para modelar una serie de tiempo se denomina ARARMA, aunque usualmente el ARMA será parte de una aproximación de un proceso AR, por lo que también estos modelos son conocidos como ARAR.

Los modelos ARIMA basados en la metodología Box-Jenkins, son casos especiales de un modelo ARARMA o ARAR, en el cual la transformación de $Y_t$ a $\widetilde{Y}_t$ están restringidas únicamente al uso de operadores de diferencias. En general debe tenerse en cuenta que al trabajar con series de tiempo, uno podría considerar otras formas de transformar $Y_t$ en $\widetilde{Y}_t$ aunque estás no esten contempladas en los modelos ARARMA. Uno de estos procedimiento consiste en construir un estimador de $\widehat{\mu}_t$ a través de la función de valores de la media $\mu_t = E(Y_t)$, asumiendo que es una función periodica de cierto periodo (en el caso de datos mensuales, periodo 12). Lo que permitirá obtener $\widetilde{Y}_t$ mediante el siguiente calculó dependiendo de si el modelo es aditivo o multiplicativo:

$\widetilde{Y}_t = Y_t - \widehat{\mu}_t$ si hay estacionalidad aditiva.

$\widetilde{Y}_t = \frac{Y_t}{\widehat{\mu}_t}$ si hay estacionalidad multiplicativa.

Así la nueva serie $\widetilde{Y}_t$ obtenida con base en las expresiones anteriores puede ser considerada como una serie de tiempo ajustada estacionalmente.

### *2.5. El Algoritmo ARARMA-ARAR*

Dado un conjunto de datos $\{Y_t, t = 1, 2, ...n\}$ ,el primer paso es determinar si el proceso subyacente es de memoria larga y por tanto aplicar una transformación que los convierta en una serie de memoria corta, antes de intentar ajustar un modelo autorregresivo. Las transformaciones que permite el algoritmo ARAR son más generales que las obtenidas por el operador de diferencias, y se resumen en dos tipos:

$\widetilde{Y}_t = Y_t - \widehat{\phi}(\widehat{\tau})Y_{t-\widehat{\tau}}$ ...(1)

$\widetilde{Y}_t = Y_t - \widehat{\phi}_1 Y_{t-1} - \widehat{\phi}_2 Y_{t-2}$ ...(2)

Con la ayuda del siguiente algoritmo que se describe en cinco pasos, se debe clasificar $\{Y_t\}$ y tomar uno de los siguientes cursos de acción:

- **L**. Declarar $\{Y_t\}$ como una serie de memoria larga y transformarla en una serie $\left\{\widetilde{Y}_t\right\}$ usando la ecuación (1)
- **M**. Declarar $\{Y_t\}$ como una serie moderadamente de memoria larga y transformarla en una serie $\left\{\widetilde{Y}_t\right\}$ usando la ecuación (2)
- **S**. Declarar $\{Y_t\}$ como una serie de memoria corta.

El algoritmo para decidir si a $\{Y_t\}$ se le aplica **L**, **M** o **S**, es el siguiente:

1. Para cada $\tau = 1, 2, ..., 15$ nosotros encontramos el valor de $\widehat{\phi}(\tau)$ de $\phi$ que minimiza $ERR(\phi, \tau) = \frac{\sum_{t=\tau+1}^{n}[Y_t - \phi Y_{t-\tau}]^2}{\sum_{t=\tau+1}^{n} Y_t^2}$. Nosotros definimos $ERR(\tau) = ERR(\widehat{\phi}(\tau), \tau)$ y escogemos el rezago $\widehat{\tau}$ que es el valor de $\tau$ que minimiza los $ERR(\tau)$.
2. Si $ERR(\widehat{\tau}) \leq \frac{8}{n}$, entonces tomamos el curso de acción **L**.
3. Si $\widehat{\phi}(\widehat{\tau}) \geq 0.93$ y $\widehat{\tau} > 2$, entonces tomamos el curso de acción **L**.

4. Si $\widehat{\phi}(\widehat{\tau}) \geq 0.93$ y $\widehat{\tau} = 1$ o 2 ,determine los valores de $\widehat{\phi}_1$ y $\widehat{\phi}_2$ de $\phi_1$ y $\phi_2$ que minimice $\sum_{t=3}^{n} [Y_t - \phi_1 Y_{t-1} - \phi_2 Y_{t-2}]^2$ , entonces tomamos el curso de acción **M**.

5. Si $\widehat{\phi}(\widehat{\tau}) < 0.93$, entonces tomamos el curso de acción **S**.

Sea $\{S_t, t = 1, 2, ...T\}$ una serie de memoria corta derivada de $\{Y_t\}$ por medio de nuestro algoritmo y $\overline{S}$ la media de $S_1, S_2, ..., S_T$ . El primer paso en nuestro procedimiento de modelación es ajustar un proceso autorregresivo a la serie corregida por su media, $X_t = S_t - \overline{S}$, $t = 1, ..., T$ . El modelo ajustado tiene la siguiente forma: $X_t = \phi_1 X_{t-1} + \phi_{l_1} X_{t-l_1} + \phi_{l_2} X_{t-l_2} + \phi_{l_3} X_{t-l_3} + Z_t$, donde $\{Z_t\} \sim WN(0, \sigma^2)$ y para rezagos dados $l_1, l_2$ y $l_3$ ; los coeficientes $\phi_j$ y la varianza del ruido blanco $\sigma^2$ son encontrados de las ecuaciones de Yule-Walker:

$$\begin{bmatrix} 1 & \widehat{\rho}(l_1 - 1) & \widehat{\rho}(l_2 - 1) & \widehat{\rho}(l_3 - 1) \\ \widehat{\rho}(l_1 - 1) & 1 & \widehat{\rho}(l_2 - l_1) & \widehat{\rho}(l_3 - l_1) \\ \widehat{\rho}(l_2 - 1) & \widehat{\rho}(l_2 - l_1) & 1 & \widehat{\rho}(l_3 - l_2) \\ \widehat{\rho}(l_3 - 1) & \widehat{\rho}(l_3 - l_1) & \widehat{\rho}(l_3 - l_2) & 1 \end{bmatrix} \begin{bmatrix} \phi_1 \\ \phi_{l_1} \\ \phi_{l_2} \\ \phi_{l_3} \end{bmatrix} = \begin{bmatrix} \widehat{\rho}(1) \\ \widehat{\rho}(l_1) \\ \widehat{\rho}(l_2) \\ \widehat{\rho}(l_3) \end{bmatrix}$$

y

$\sigma^2 = \widehat{\gamma}(0) \left[1 - \phi_1 \widehat{\rho}(1) - \phi_{l_1} \widehat{\rho}(l_1) - \phi_{l_2} \widehat{\rho}(l_2) - \phi_{l_3} \widehat{\rho}(l_3)\right]$ , donde $\widehat{\gamma}(j)$ y $\widehat{\rho}(j)$ , $j = 0, 1, 2, ...$, son las autocovarianzas y autocorrelaciones muestrales de la serie $\{X_t\}$.

El algoritmo ARAR se puede programar en S-Plus o cualquier otro lenguaje o paquete que lo permita. Para la realización de este trabajo se eligió el software ITSM (Interactive Time Series Modelling, Brockwell y Davis (1994)), ya que tiene integrado un modulo exclusivo para construir modelos ARAR con base en el algoritmo aquí desarollado. Así, ITSM puede calcular los coeficientes $\phi_j$ para cada conjunto de rezagos tal que : $1 < l_1 < l_2 < l_3 \leq m$ , donde $m$ puede ser seleccionada de entre 13 o 26. Una vez que se ha escogido el valor de $m$, se elige el modelo para el cual las estimaciones de las ecuaciones de Yule-Walker hacen que $\sigma^2$ sea minima, y entonces registrar los rezagos $(l_j)$, coeficientes y varianza del ruido blanco para el modelo ajustado. Un procedimiento más lento es escoger los rezagos y coeficientes calculados de las ecuaciones Yule-Walker que maximiza la Verosimilitud Gaussiana de las observaciones. Para está opción el máximo de rezagos $m$ es 13, si se utiliza ITSM.

Si el filtro de memoria corta encontrado en el primer paso de nuestro algoritmo tiene coeficientes $\psi_0(= 1)$, $\psi_1, ..., \psi_k$ $(k \geq 0)$, entonces la serie de memoria corta puede ser expresada como: $S_t = \psi(B)Y_{t+k} = Y_{t+k} + \psi_1 Y_{t+k-1} + ... + \psi_k Y_t$, ...(3) donde $\psi(B)$ es el polinomio de retraso, $\psi(B) = 1 + \psi_1 B + ... + \psi_k B^k$. Similarmente, si los coeficientes de el subconjunto autorregresivo encontrado en el segundo paso son $\phi_1, \phi_{l_1}, \phi_{l_2}$, y $\phi_{l_3}$, entonces el subconjunto de modelos AR para la serie corregida por su media $\{X_t = S_t - \overline{S}\}$ es $\phi(B)X_t = Z_t$ ...(4), donde $\{Z_t\} \sim WN(0, \sigma^2)$ y $\phi(B) = 1 - \phi_1 B - \phi_{l_1} B^{l_1} - \phi_{l_2} B^{l_2} - \phi_{l_3} B^{l_3}$ . De las ecuaciones (3) y (4) obtenemos: $\xi(B)Y_t = \phi(1)\overline{S} + Z_t$ ...(5), donde $\xi(B) = \psi(B)\phi(B) = 1 + \xi_1 B + ... + \xi_{k+l_3} B^{k+l_3}$ .

Asumiendo ahora que en el modelo ajustado (5) el término de ruido blanco $Z_t$ es no correlacionado con $\{Y_j, j < t\}$ para cada $t$, podemos determinar el minímo error cuadrático medio de los predictores lineales $P_n Y_{n+h}$ de $Y_{n+h}$ en periodos de $\{1, Y_1, ..., Y_n\}$, para $n > k + l_3$, de las repeticiones $P_n Y_{n+h} = -\sum_{j=1}^{k+l_3} \xi_j P_n Y_{n+h-j} + \phi(1)\overline{S}, h \geq 1$ ...(6) con las condiciones iniciales $P_n Y_{n+h} = Y_{n+h}$, para $h \leq 0$ ...(7). El error cuadrático medio del predictor $P_n Y_{n+h}$ se encuentra a través de : $E\left[(Y_{n+h} - P_n Y_{n+h})^2\right] = \sum_{j=1}^{h-1} \tau_j^2 \sigma^2$ ...(8), donde $\sum_{j=0}^{\infty} \tau_j z^j$ es la expansión de Taylor de $\frac{1}{\xi(z)}$ en una vecindad de $z = 0$. Equivalentemente, la secuencia $\{\tau_j\}$ puede ser encontrada de la repetición $\tau_0 = 1, \sum_{j=0}^{n} \tau_j \xi_{n-j} = 0$, $n = 1, 2, ...$ ...(9).

## *2.6. Utilización del Programa ARAR del paquete ITSM.*

Para determinar un modelo ARAR para una serie de datos $\{Y_t\}$ con base el el paquete ITSM y pronosticar con este diferentes valores de la serie. Primero se debe cargar el conjunto de datos. Después aparecera en la pantalla un menu principal en el cual se puede calcular el polinomio de memoria corta, que permitirá encontrar el mejor filtro de memoria corta. Enseguida el programa proporcionará los coeficientes $1, \psi_1, ..., \psi_k$ del filtro seleccionado. La serie de memoria corta que va a generar el paquete es de la siguiente forma: $S_t = Y_t + \psi_1 Y_{t-1} + ... + \psi_k Y_{t-k}$ .Al oprimer la tecla enter, aparece un submenu que permite hacer un AR . La primera opción ajusta un modelo autorregresivo con 4 coeficientes diferentes de cero o significativos a una serie de media corregida del tipo $X_t = S_t - \overline{S}$ , escogiendo los rezagos y coeficientes que minimizan la varianza del ruido blanco generado por la ecuaciones de Yule-Walker.

$X_t = \phi_1 X_{t-1} + \phi_{l_1} X_{t-l_1} + \phi_{l_2} X_{t-l_2} + \phi_{l_3} X_{t-l_3}$ .Al oprimir la tecla F y enter, los rezagos optimosy los coeficientes del modelo correspondiente aparecen en pantalla. Los coeficientes $\xi_j$ de $B^j$ en el whitening filter general donde $B$ es el operador de retraso, $\xi(B) = (1 + \psi_1 B + ... + \psi_k B^k)(1 - \phi_1 B - \phi_{l_1} B^{l_1} - \phi_{l_2} B^{l_2} - \phi_{l_3} B^{l_3})$, también son calculados y mostrados en la pantalla. Al teclear enter otra vez, el programa te pregunta el número de valores que deseas pronosticar de $\{Y_t\}$, y entonces muestra una gráfica con los datos originales más los valores futuros que fueron solicitados.

$P_n Y_{n+h-j} + \phi(1)\overline{S}, h \geq 1$ predictores $+\phi(1)\overline{S}, h \geq 1Y + Z_t$...de media coregida ${}_tY_{t+k}Y_{t+k}$transformada contrado para transformar la serie de memoria larga en una de memoria corta, tiene los coeficientes utiliza para está opción un máximo de rezagos $m$ =ajustado ajustado estimadas $\widehat{\rho}(j)$ aunque ya existe un software que tiene integrado el algoritmo es decir, un

modelo autorregresivo.

Las transformaciones que permite el algoritmo ARAR son más generales que las obtenidas por el operador $\{S_t, t = 1, 2, ...T\}$ una serie de memoria

corta derivada de $\{Y_t\}$ por medio de nuestro algoritmo y $\overline{S}$ la media ipor tanto aplicar una transformación que los convierta en una serie de memoria corta, antes de intentar ajustar un modelo autorregresivo. Las transformaciones que permite el algoritmo ARAR son más generales que las obtenidas por el operador muestral

## 3. Conclusiones

A lo largo de artículo se describio la metodología de los modelos ARARMA-ARAR, concluyendo que los modelos ARIMA elaborados por Box y Jenkins son un caso particular de la familia general de modelos denominados ARARMA-ARAR. La ventaja de los modelos general es su buena capacidad predictiva, así como que existe un software que permite realizar las estimaciones respectivas. Los modelos ARARMA-ARAR utilizan un filtro que permite modelar las series como de memoria larga o corto.

## 4. Referencias Bibliograficas

Brockwell y Davis (1994), "Introduction to Time Series", Editorial Spring-Verlang.

aparecen en pantalla. Los coeficientes $\xi_j$ de $B^j$ en el whitening filter general donde $B$ es el operador de retraso, $\xi(B) = (1 + \psi_1 B + ... + \psi_k B^k)(1 - \phi_1 B - \phi_{l_1} B^{l_1} - \phi_{l_2} B^{l_2} - \phi_{l_3} B^{l_3})$, también son calculados y mostrados en la pantalla. Al teclear enter otra vez, el programa te pregunta el número de valores que deseas pronosticar de $\{Y_t\}$, y entonces muestra una gráfica con los datos originales más los valores futuros que fueron solicitados.

$P_n Y_{n+h-j} + \phi(1)\overline{S}, h \geq 1$ predictores $+\phi(1)\overline{S}, h \geq 1Y$ $+Z_t$...de media coregida ${}_tY_{t+k}Y_{t+k}$transformada contrado para transformar la serie de memoria larga en una de memoria corta, tiene los coeficientes utiliza para está opción un máximo de rezagos $m$ =ajustado ajustado estimadas $\widehat{\rho}(j)$ aunque ya existe un software que tiene integrado el algoritmo es decir, un

modelo autorregresivo.

Las transformaciones que permite el algoritmo ARAR son más generales que las obtenidas por el operador $\{S_t, t = 1, 2, ...T\}$ una serie de memoria

corta derivada de $\{Y_t\}$ por medio de nuestro algoritmo y $\overline{S}$ la media ipor tanto aplicar una transformación que los convierta en una serie de memoria corta, antes de intentar ajustar un modelo autorregresivo. Las transformaciones que permite el algoritmo ARAR son más generales que las obtenidas por el operador muestral

## 3. Conclusiones

A lo largo de artículo se describio la metodología de los modelos ARARMA-ARAR, concluyendo que los modelos ARIMA elaborados por Box y Jenkins son un caso particular de la familia general de modelos denominados ARARMA-ARAR. La ventaja de los modelos general es su buena capacidad predictiva, así como que existe un software que permite realizar las estimaciones respectivas. Los modelos ARARMA-ARAR utilizan un filtro que permite modelar las series como de memoria larga o corto.

## 4. Referencias Bibliograficas

Brockwell y Davis (1994), "Introduction to Time Series", Editorial Spring-Verlang.

Hamilton, James D. (1994), "Time Series Analysis", Princeton University Press.

Robert H. Shumway, Robert H. y Stoffer David S. (2000), "Time Series Analysis and Its Applications", Editorial Spring-Verlang.

Documento de Trabajo (Working Paper)

**WP-FZ-ITESM-CEM-No.01-2005**



Toda la correspondencia relacionada con la Serie de Documentos de Trabajo se recibirá en:

Tecnológico de Monterrey, Campus Estado de México
Departamento de Finanzas, Aulas 5, 3er. Piso
Carretera Lago de Guadalupe Km. 3.5
Col. Margarita Maza de Juárez, Atizapán de Zaragoza,
52926 Estado de México. Correo electrónico: plopezs@itesm.mx
Teléfono y Fax: 58-64-55-55 Ext. 3161

*Diseño y Redacción:* MF. Pablo López Sarabia

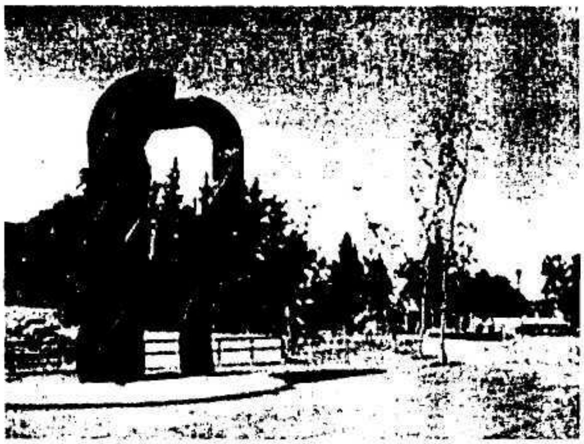

**Serie en Econometría Financiera**

# Modelos de Series de Tiempo: Una Introducción a la Técnica ARARMA-ARAR y ARIMA

**Pablo López Sarabia**
**e-mail: plopezs@itesm.mx**

Documento de Trabajo (Working Paper)

TECNOLÓGICO
DE MONTERREY



La serie **Documentos de Trabajo (Working Paper)** es una publicación del Departamento de Finanzas y la División de Negocios del Tecnológico de Monterrey, Campus Estado de México. Las opiniones y puntos de vista en cada una de las contribuciones son responsabilidad exclusiva de los autores. El fotocopiado del artículo y/o párrafos es permitido para fines educativos sin fines de lucro, siempre y cuando se haga mención de la fuente y el autor.

Toda la correspondencia relacionada con este documento se recibirá en:


Tecnológico de Monterrey, Campus Estado de México
Departamento de Finanzas, Aulas 5, 3er. Piso
Carretera Lago de Guadalupe Km. 3.5, Col. Margarita Maza de Juárez,
Atizapán de Zaragoza, 52926 Estado de México
Tel. y Fax: 58-64-55-55 Ext. 3161




Editor y creador de la serie Documentos de Trabajo en Econometría Financiera ***M.F. Pablo López Sarabia***