



Universidad Autónoma de Querétaro
Facultad de Ingeniería
Licenciatura en Matemáticas Aplicadas

MODELAJE ESTADÍSTICO PARA CASOS DE TIFOIDEA EN
MÉXICO
Y SU RELACIÓN CON LA TEMPERATURA MÁXIMA

Tesis

Que como parte de los requisitos para obtener el grado de
Licenciado en Matemáticas Aplicadas

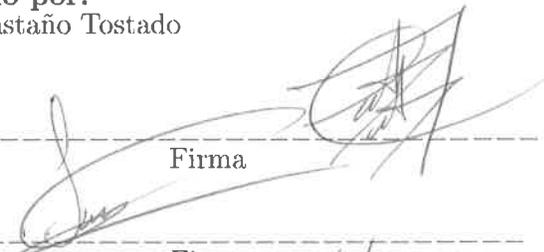
Presenta:

José Dolores Escolano Infante

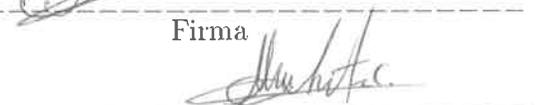
Dirigido por:

Dr. Eduardo Castaño Tostado

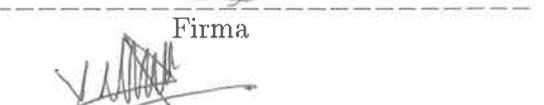
Dr. Eduardo Castaño Tostado
Presidente


Firma

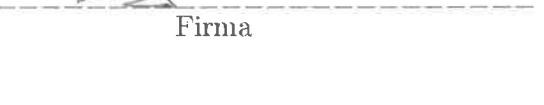
M. en C. Sara Silva Hernández
Secretario


Firma

Dr. Mario Santana Cibrian
Vocal


Firma

M. en C. Juan Antonio Villeda Reséndiz
Sinodal


Firma

Dr. Manuel Toledano Ayala
Director de la Facultad de Ingeniería

Centro Universitario
Querétaro, Qro.
Octubre 2018



Universidad Autónoma de Querétaro
Facultad de Ingeniería
Licenciatura en Matemáticas Aplicadas

MODELAJE ESTADÍSTICO PARA CASOS DE TIFOIDEA EN
MÉXICO
Y SU RELACIÓN CON LA TEMPERATURA MÁXIMA
Tesis

Que como parte de los requisitos para obtener el grado de
Licenciado en Matemáticas Aplicadas

Presenta:

José Dolores Escolano Infante

Dirigido por:

Dr. Eduardo Castaño Tostado

Dr. Eduardo Castaño Tostado
Presidente

Firma

M. en C. Sara Silva Hernández
Secretario

Firma

Dr. Mario Santana Cibrian
Vocal

Firma

L.M.A. Juan Antonio Villeda Reséndiz
Sinodal

Firma

Dr. Manuel Toledano Ayala
Director de la Facultad de Ingeniería

Centro Universitario
Querétaro, Qro.
Octubre 2018

Índice

1. Introducción	3
1.1. Antecedentes.	3
1.2. Justificación.	4
1.3. Descripción del problema.	4
1.4. Planteamiento teórico.	4
1.5. Objetivos.	4
1.6. Resultados esperados, posibles aplicaciones y uso del proyecto.	4
2. Regresión clásica en series de tiempo y análisis de datos exploratorios	5
2.1. Conceptos preliminares.	5
2.2. Antecedentes.	5
3. Modelos autorregresivos y de promedios móviles	12
3.1. Modelos autorregresivos	12
3.2. Modelos de promedios móviles	14
3.3. Modelos autorregresivos y de promedios móviles.	15
4. Análisis espectral y filtración	16
4.1. Periodicidad y comportamiento ciclado de una serie temporal	16
4.2. Densidad espectral	18
4.3. Periodograma y la discretización de la transformación de Fourier	20
4.4. Autocorrelación y autocorrelación parcial	23
4.5. Función de Autocorrelación parcial (PACF)	24
4.6. Predicción	25
4.7. Estimación	26
4.7.1. Máxima verosimilitud	27
4.8. Modelos integrados para datos no estacionarios	29
4.9. Identificación de modelos ARIMA	29
4.10. Regresión con errores autocorrelacionados	30
4.11. Modelos estacionales ARIMA	30
5. Regresión rezagada: Modelando la función de transferencia	32
6. Ejemplo del Southern Oscillation Index y el Reclutamiento	34
7. Modelaje de la relación del número de Casos de Tifoidea en México con la Temperatura Máxima	38
8. Conclusiones	43
9. Bibliografía	44
10. Programa	46

1. Introducción

1.1. Antecedentes.

Desde la antigüedad las enfermedades han ido cambiando teniendo efectos colaterales de mayor índole o en el mejor de los casos erradicándose. Pero uno de los casos que han resurgido en México es la fiebre tifoidea y se cree que uno de los factores determinantes para su posible propagación es debido al constante cambio climático que presenta todo el planeta. Aunado a ello sabemos que la bacteria que causa la fiebre tifoidea, *Salmonella typhi*, se propaga a través de alimentos, agua o bebidas contaminadas. Representa la quinta enfermedad infecciosa en viajeros por falta de higiene adecuada, se estima que se presentan de 1 a 10 casos nuevos por cada 100,000 habitantes a nivel mundial. La fiebre tifoidea es común en países en desarrollo, pero menos de 400 casos se notifican en los Estados Unidos cada año Who (2008). La mayoría de los casos en los Estados Unidos proviene de otros países en donde la fiebre tifoidea es común.

Checkley et al. (2003) reportaron que las temperaturas más altas aumentan las diarreas bacterianas y parasitarias al extender la supervivencia de las bacterias que causan enterogastritis en alimentos contaminados, así mismo Checkley et al. (2000) observaron que los ingresos hospitalarios diarios para la diarrea exhibían un incremento de dos por ciento por cada 5 ° C de aumento en la temperatura media.

Las comunidades costeras son más vulnerables al aumento de las inundaciones que pueden traer agua contaminada (procedente de fuentes tales como plantas de tratamiento de aguas residuales municipales y fábricas de alimento de animales que está en estrecha proximidad con las personas que viven en las comunidades costeras afectadas Semenza et al. (2012).

Los investigadores anteriores han utilizado análisis de series de tiempo para examinar la correlación entre las epidemias de diarrea y los factores climáticos Pascual et al (2000). Rodeo et al. (2002). Un modelo de regresión de series de tiempo se ha aplicado para evaluar el impacto del cambio climático a largo plazo, especialmente para las epidemias de diarrea extremas Kale et al. (2004), Hashizume et al (2007), De Magny et al (2008).

Los cambios climáticos y el aumento de los fenómenos meteorológicos extremos han sido fuertemente ligados a la incidencia de enfermedades gastrointestinales.

Existen muchos artículos basados en el estudio de diversas enfermedades gastrointestinales, tales como “Modeling the impact of climate variability on diarrhea-associated diseases in Taiwan (1996–2007)” elaborado por Wei-Chun Chou , Jiunn-Lin Wu, Yu-Chun Wang, Hsin Huang, Fung-Chang Sung y Chun-Yu Chuang en el 2010, el cual investiga y cuantifica la relación entre las variaciones del clima y la morbilidad asociada a la diarrea en Taiwán subtropical, aplicando un modelo de regresión de Poisson con factores climáticos (humedad relativa, temperatura máxima y la cantidad de precipitaciones extremas), autoregresiones, rezagos a largo plazo, estacionalidad y un efecto-tiempo de retraso. Los resultados indicaron que la temperatura máxima y los días de precipitaciones extremas estuvieron fuertemente relacionados con la morbilidad asociada a diarrea.

Otro artículo relacionado es “Effect of non-stationary climate on infectious gastroenteritis transmission in Japan” realizado por Daisuke Onozuka (2014), donde se analizaron los datos mensuales para los casos de gastroenteritis infecciosa en Fukuoka, Japón entre el 2000 y el 2012 usando el análisis de coherencia cruzado wavelet para evaluar el patrón de las asociaciones entre los índices para el Dipolo del Océano Índico (IOD) y El Niño Oscilación del Sur (ENOS). Donde la asociación era no estacionaria y parecía tener una influencia importante en la sincronía de la transmisión de la gastroenteritis infecciosa.

Así mismo, en el 2015, Daisuke Onozuka en colaboración con Akihito Hagihara en su artículo “Nationwide variation in the effects of temperature on infectious gastroenteritis incidence in Japan” investigaron los efectos de la temperatura sobre la incidencia de enfermedades infecciosas gastrointestinales utilizando datos a nivel nacional. Se empleó un análisis en dos etapas para evaluar la relación no lineal y el rezago entre la temperatura y la morbilidad, donde la primera etapa utilizó una regresión de Poisson permitiendo sobredispersión en un modelo no lineal de retraso para estimar los efectos específicos de la temperatura sobre la morbilidad. En la segunda etapa, se aplicó un meta-análisis multivariado para estimaciones de depósitos a nivel nacional. Así, sus resultados sugieren que las estrategias de salud pública dirigidas al control de la gastroenteritis infecciosa relacionada con la temperatura pueden ser más eficaces cuando se adaptan de acuerdo a las condiciones climáticas específicas de la región.

Por último, en el artículo “Climate change, extreme events and increased risk of salmonellosis in Maryland, USA: Evidence for coastal vulnerability” (2015) los autores Chengsheng Jiang, Kristi S. Shaw, Crystal R. Upperman, David Blythe, Clifford Mitchell, Raghu Murtugudde, Amy R. Sapkota y Amir Sapkota obtuvieron datos de casos de *Salmonella* que fue transmitida por los alimentos entre los años 2002-2012 y los datos de tiempo en el Centro Nacional de Datos Climáticos (1960-2012). Con éstos desarrollaron parámetros de exposición relacionados con eventos extremos de precipitación y temperatura utilizando una línea base de 30 años (1960-1989) vinculados con los datos de la salmonelosis a nivel de condado. Analizaron los datos mediante la estimación de ecuaciones binomiales

negativas generalizadas, donde observaron un aumento de 4,1 % en el riesgo de salmonelosis asociada con un aumento de una unidad en eventos térmicos extremos, y del cual, dicho aumento en el riesgo fue más pronunciado en las zonas costeras en comparación a las zonas no costeras. Del mismo modo, se observó un incremento del 5,6 % en el riesgo de salmonelosis asociada con un aumento de una unidad en eventos extremos de precipitación.

1.2. Justificación.

Muchos estudios se han realizado alrededor de este tema en diferentes partes del mundo, sin embargo no existen trabajos relacionados en países con economía emergente, como es el caso de México, es por esto la importancia del estudio de esta tesis, el cual es un estudio de los casos de salmonelosis en México.

Por otro lado, tener un modelo que pueda predecir el número de casos de salmonela ayudará al sistema de salud de México a estar preparado para poder atender a todo aquel que lo necesite, con ello se busca dar un sentido cuantitativo acerca de las incidencias que puedan haber de esta enfermedad y así darle herramientas a la parte química y médica para contrarrestar los estragos de esta enfermedad.

1.3. Descripción del problema.

Este estudio cuenta con ciertos factores que harán un poco complejo el desarrollo de la investigación, siendo algunas de ellas la calidad de los datos mismos que cuentan con un problema en su comportamiento, la falta de conocimiento acerca de cuál será el mejor modelaje de los datos para sacar un rendimiento óptimo de la información poseída y así mismo dar un buen rango de certeza, por mencionar algunas de las problemáticas.

1.4. Planteamiento teórico.

Con base en estudios previos se cree que los casos presentados de fiebre tifoidea se pueden explicar mediante diferentes variables, es por esto que se busca encontrar un modelo que sea adecuado para explicar esta enfermedad el cual se base directamente en los datos disponibles.

1.5. Objetivos.

Comparar diferentes modelos en los cuales se planea predecir el número de casos de salmonela a partir de diversas variables como la precipitación, la temperatura máxima, mínima y medía para así poder encontrar el mejor resultado posible.

Obtener la experiencia y los conocimientos suficientes para en un futuro poder modelar distintas series de tiempo y así lograr predecir distintas enfermedades con variables que no necesariamente estén relacionadas con el clima.

1.6. Resultados esperados, posibles aplicaciones y uso del proyecto.

Proporcionar el mejor modelo, i.e. el modelo con el menor AIC (Akaike information criterion), el cual logre predecir el número de casos de fiebre tifoidea (salmonelosis) en México con un menor margen de error.

Dicho modelo podría ser utilizado para prevenir a los sistemas de salud en México de futuros incrementos en dicha bacteria y así tener medicamentos suficientes para cubrir la futura demanda.

2. Regresión clásica en series de tiempo y análisis de datos exploratorios

2.1. Conceptos preliminares.

Durante este estudio se utilizarán las siguientes definiciones básicas de series de tiempo, por lo que no se profundizará en ellas.

Definición 2.1.1 La **función media** es definida como

$$\mu_{x_t} = E(x_t) = \int_{-\infty}^{\infty} x f_t(x) dx, \quad (1)$$

siempre y cuando exista.

Definición 2.1.2 El **ruido blanco** es una señal aleatoria que se caracteriza por el hecho de que sus valores de señal en dos tiempos diferentes no guardan correlación estadística

Definición 2.1.3 La **función de autocovarianza** es definida como el producto de segundo momento

$$\gamma_x(s, t) = \text{cov}(x_s, x_t) = E[(x_s - \mu_s)(x_t - \mu_t)], \quad (2)$$

para toda s y t .

Definición 2.1.4 La **función de covarianza cruzada** entre dos series, x_t y y_t , es

$$\gamma_{xy}(s, t) = \text{cov}(x_s, y_t) = E[(x_s - \mu_{x_s})(y_t - \mu_{y_t})]. \quad (3)$$

Definición 2.1.5 La **función de correlación cruzada (CCF)** es dada por

$$\rho_{xy}(s, t) = \frac{\gamma_{xy}(s, t)}{\sqrt{\gamma_x(s, s) \gamma_y(t, t)}}. \quad (4)$$

Definición 2.1.6 Una serie es **estrictamente estacionaria** cuando el comportamiento de la probabilidad para cualquier colección de valores

$$\{x_{t_1}, x_{t_2}, \dots, x_{t_k}\}$$

es idéntica a un conjunto desplazado

$$\{x_{t_1+h}, x_{t_2+h}, \dots, x_{t_k+h}\}.$$

Esto es,

$$P\{x_{t_1} \leq c_1, \dots, x_{t_k} \leq c_k\} = P\{x_{t_1+h} \leq c_1, \dots, x_{t_k+h} \leq c_k\} \quad (5)$$

para toda $k = 1, 2, \dots$, todo tiempo t_1, t_2, \dots, t_k , todo número c_1, \dots, c_k y todo desplazamiento $h = 0, \pm 1, \pm 2, \dots$.

Definición 2.1.7 Una serie de tiempo **débilmente estacionaria**, x_t , es un proceso tal que

(i) La función media, μ_t es constante y no depende de tiempo, y

(ii) la función de autocovarianza, $\gamma(s, t)$ depende únicamente de s y t sólo a través de la diferencia $|s - t|$.

Se utilizará el término **estacionario** para referirse a débilmente estacionario.

2.2. Antecedentes.

El contexto de la regresión lineal en series de tiempo supone que la salida o la serie de tiempo dependiente, x_t para $t = 1, \dots, n$ está relacionada con una colección de posibles entradas o series independientes, z_{t1}, \dots, z_{tq-1} considerando tales entradas fijas y conocidas. Dicho supuesto será modificado más adelante, ya que de momento es necesaria para la regresión lineal convencional. Expresando la relación anterior de la siguiente manera.

$$x_t = \beta_0 + \beta_1 z_{t1} + \beta_2 z_{t2} + \dots + \beta_{q-1} z_{tq-1} + w_t, \quad (6)$$

donde los coeficientes β son coeficientes desconocidos de la regresión fijos, y w_t es ruido blanco o un error aleatorio el cual es una variable normal con media cero y varianza σ_w^2 , independientes e idénticamente distribuidas (IID). Tal supuesto también será modificado posteriormente.

Para estimar β_0 y β_1 se utiliza el criterio de mínimos cuadrados ordinarios (MCO), minimizando la suma de los errores cuadrados, es decir,

$$Q = \sum_{t=1}^n w_t^2 = \sum_{t=1}^n (x_t - [\beta_0 + \beta_1 z_t])^2,$$

con respecto a β_0 y β_1 . Se evalúa $\frac{\partial Q}{\partial \beta_i} = 0$ para obtener las 2 ecuaciones

$$\hat{\beta}_1 = \frac{\sum_{t=1}^n (x_t - \bar{x})(z_t - \bar{z})}{\sum_{t=1}^n (z_t - \bar{z})^2},$$

y

$$\hat{\beta}_0 = \bar{x} - \hat{\beta}_1 \bar{z},$$

donde \bar{x} y \bar{z} son los respectivos promedios.

Ejemplo 2.2.1 Estimación de una tendencia lineal.

Sea x_t los datos de la temperatura global, sea como la Figura 1, con una tendencia a la alza. Mediante regresión lineal simple se ajustó el modelo de la forma (6) con $q = 1$ presentado a continuación.

$$x_t = \beta_0 + \beta_1 z_{t1} + w_t,$$

donde $z_t = 1880, \dots, 2009$ son los años y bajo el supuesto de que los errores w_t , son secuencias normales IID.

E:/tesis/Modelos de Salmonela/imagenes/Global temperatura modelado lienal.png

Figura 1: Temperatura global con ajuste lineal simple.

Con los datos de la Figura 1, al sustituir los valores se obtiene que $\hat{\beta}_0 = -11,2$ y $\hat{\beta}_1 = 0,006$ con un error estándar de 0.0003. Lo cual indica un aumento de 0.51 a 0.63 grados centígrados por cada 100 años. Esto se muestra en la Figura 1, con el modelo $\hat{x}_t = -11,2 + 0,006z_t$ sobrepuesto con una línea roja. El modelo estimado puede evaluarse en R con los siguientes comandos:

```
summary(fit <- lm(gtemp ~ time(gtemp)))
plot(gtemp, type="o", ylab="Global Temperature Deviation")
abline(fit)
```

Por otro lado la forma multilineal (6) puede escribirse vectorialmente de la siguiente manera

$$x_t = \beta z_t + w_t, \quad (7)$$

donde $z_t = (1, z_{t1}, \dots, z_{tq})^T$, $\beta = (\beta_0, \beta_1, \dots, \beta_q)^T$ y $w_t \sim N(0, \sigma_w^2)$ IID. Y de la misma forma cuando $q = 1$ la estimación por MCO encuentra los coeficientes del vector β que minimiza la suma de los errores cuadrados

$$Q = \sum_{t=1}^n w_t^2 = \sum_{t=1}^n (x_t - \beta^T z_t)^2, \quad (8)$$

con respecto a $\beta_0, \beta_1, \dots, \beta_q$. Se minimiza Q diferenciando (8) con respecto a β , con lo que el mínimo, $\hat{\beta}$, debe cumplir $\sum_{t=1}^n (x_t - \hat{\beta}^T z_t) z_t = 0$. Lo cual lleva a las llamadas ecuaciones normales:

$$\left(\sum_{t=1}^n z_t z_t^T \right) \hat{\beta} = \sum_{t=1}^n z_t x_t. \quad (9)$$

Y si la matriz $\sum_{t=1}^n z_t z_t^T$ es no singular, la estimación por mínimos cuadrados de β será

$$\hat{\beta} = \left(\sum_{t=1}^n z_t z_t^T \right)^{-1} \sum_{t=1}^n z_t x_t.$$

La suma de los errores mínimos cuadrados (SSE por sus siglas en inglés), es obtenida por

$$SSE = \sum_{t=1}^n (x_t - \hat{\beta}^T z_t)^2. \quad (10)$$

Este estimador de mínimos cuadrados es insesgado, i.e $E(\hat{\beta}) = \beta$, y tiene una matriz de varianza y covarianza

$$\text{cov}(\hat{\beta}) = \sigma_w^2 C, \quad (11)$$

donde

$$C = \left(\sum_{t=1}^n z_t z_t^T \right)^{-1}. \quad (12)$$

Si los errores w_t son normales, $\hat{\beta}$ se vuelve el estimador máximo verosímil y tendrá una distribución normal. Un estimador insesgado de la varianza σ_w^2 es

$$s_w^2 = MSE = \frac{SSE}{n - (q + 1)}, \quad (13)$$

donde MSE (por sus siglas en inglés) es el error cuadrático medio.

Bajo la hipótesis de normalidad

$$t = \frac{(\hat{\beta}_i - \beta_i)}{s_w \sqrt{C_{ii}}}, \quad (14)$$

sigue una distribución t-Student con $n - k$ grados de libertad, donde c_{ii} denota el i -ésimo elemento de la diagonal de C (12). Dicho resultado es usado para probar la hipótesis nula $H_0 : \beta_i = 0$ para $i = 1, \dots, q$. En ocasiones es de interés seleccionar un subconjunto de las q variables independientes. Sea $r < q$ y $z_{t1:r} = \{z_{t1}, \dots, z_{tr}\}$ el subconjunto del cual se cree que depende la variable x_t . Obteniendo el modelo

$$x_t = \beta_0 + \beta_1 z_{t1} + \dots + \beta_r z_{tr} + w_t, \quad (15)$$

donde β_i con $i = 1, \dots, r$ es un subconjunto de las β 's originales.

En este caso la hipótesis nula será $H_0 : \beta_i = 0$ para $i = r + 1, \dots, q$. Se puede comparar el modelo completo (7) contra el modelo reducido (15) mediante el estadístico F obtenido como

$$F = \frac{(SSE_r - SSE) / (k - 1 - r)}{SSE / (n - k - 2)} = \frac{MSR}{MSE}, \quad (16)$$

donde SSE_r es la suma de los errores cuadrados bajo el modelo (15). Es claro que $SSE_r \geq SSE$ ya que el modelo completo tiene más parámetros, y si la hipótesis nula del modelo reducido es cierta, entonces $SSE_r \approx SSE$ pues los valores de β_i para $i = r + 1, \dots, q$ serán cercanos a 0. Por lo tanto, se rechaza H_0 si $SSE_r - SSE$ es grande. Bajo la hipótesis nula, (16) tiene una distribución central F con $q - r$ y $n - q - 1$ grados de libertad cuando (15) es el modelo correcto.

Dicho resultado suele resumirse en el análisis de varianza (ANOVA) que se muestra en la Tabla 1. Para este caso en particular, la hipótesis nula será rechazada a un nivel α si $F > F_{n-q-1}^{q-r}(\alpha)$.

Fuente de valoración	gl	Suma de Cuadrados	Cuadrados Medios	F
$a_{t,r+1:q}$	$q - r$	$SSR = SSE_r - SSE$	$MSR = SSR / (q - r)$	$F = \frac{MSR}{MSE}$
Error	$n - (q + 1)$	SSE	$MSE = SSE / (n - q - 1)$	

Cuadro 1: Análisis de varianza para regresión.

En el caso cuando $r = 0$, el modelo (15) se convierte en

$$x_t = \beta_0 + w_t,$$

donde la porción de variación de x_t explicada por todas las variables es calculada mediante

$$R^2 = \frac{SSE_0 - SSE}{SSE_0}, \quad (17)$$

y la suma de cuadrados residuales

$$SSE_0 = \sum (x_t - \bar{x})^2, \quad (18)$$

es la suma de residuales del modelo más básico, es decir, $x_t = \beta_0 + w_t$.

En este caso SSE_0 es la suma cuadrada de desviaciones del promedio \bar{x} y es conocida como la suma total de cuadrados ajustados. La medida R^2 es llamada coeficiente de determinación entre x_t y las variables z_{t1}, \dots, z_{tq} .

Si se considera el modelo de regresión normal con k coeficientes, se denota al estimador máximo verosímil de la varianza como

$$\hat{\sigma}_k^2 = \frac{SSE(q+1)}{n}, \quad (19)$$

donde $SSE(q+1)$ es la suma de cuadrados residuales bajo el modelo de regresión con $q+1$ coeficientes. Akaike (1969,1973,1974) sugirió una medida de bondad de ajuste que balanceara el error del ajuste contra el número de parámetros en el modelo.

Definición 2.2.1 *Criterio de información de Akaike (AIC)*

$$AIC = \log \hat{\sigma}_{q+1}^2 + \frac{n + 2(q+1)}{n}, \quad (20)$$

donde $\hat{\sigma}_{q+1}^2$ es dada en (19). El valor de $q+1$ que lleve al mínimo AIC se considera como el mejor modelo.

La idea de minimizar $\hat{\sigma}_{q+1}^2$ parecería un objetivo razonable, pero $\hat{\sigma}_{q+1}^2$ disminuirá monótonamente al ir incrementando $q+1$. Por ello, se dispone de otra forma de penalizar, sugerida por Sugiura (1978) y expandida por Hurvich y Tsai (1989), basada en los resultados distributivos con muestra pequeñas para los modelos de regresión lineal.

Definición 2.2.2 *AIC, sesgo corregido (AICc)*

$$AICc = \log \hat{\sigma}_k^2 + \frac{n+k}{n-k-2}, \quad (21)$$

donde $\hat{\sigma}_k^2$ es dada en 19, k es el número de parámetros en el modelo y n es el tamaño de la muestra.

También se derivó otra corrección basada en argumentos Bayesianos.

Definición 2.2.3 *Criterio de información bayesiana (BIC)*.

$$BIC = \log \hat{\sigma}_k^2 + \frac{k \log n}{n}, \quad (22)$$

con la misma notación que en la definición 2.2.2.

El BIC es conocido también como el criterio de información de Schwarz (SIC). Simulaciones diversas verifican que el BIC hace un mejor trabajo al obtener el orden correcto en muestras grandes, mientras que el AICc tiende a ser superior en muestras pequeñas donde el número de parámetros es relativamente grande Mcquarrie y Tsai (1998).

Ejemplo 2.2.2 Contaminación, temperatura y mortalidad.

Los datos mostrados en la Figura 2 son las series de un estudio realizado por Shumway Et Al. (1988) sobre los posibles efectos de la temperatura T_t y la contaminación P_t en el número de muertes por semana en Los Angeles M_t . Es posible notar componentes estacionales en todas las series, correspondientes a las variaciones invierno-verano y la tendencia a la baja en la mortalidad cardiovascular sobre un periodo de 10 años.



Figura 2: (Arriba) Número de muertes cardiovasculares. (Al centro) Temperatura en °F y (Abajo) Cantidad de partículas de contaminación.

La gráfica de dispersión mostrada en la Figura 3 indica las posibles relaciones lineales entre la mortalidad y las partículas de contaminación y la posible relación con la temperatura. Se puede apreciar una forma curvilínea entre la temperatura y la mortalidad, lo cual indica que tanto las bajas como las altas temperaturas están asociadas con el incremento de las muertes cardiovasculares.

imagenes/transfer tifo max/ejemplo1.1.2-2.png

Figura 3: Gráfica de dispersión entre la mortalidad, las partículas de contaminación y la temperatura.

Con base en las gráficas de dispersión, se proponen cuatro modelos de la forma tentativa.

$$M_t = \beta_1 + \beta_2 t + w_t, \quad (23)$$

$$M_t = \beta_1 + \beta_2 t + \beta_3 (T_t - \tilde{T}) + w_t, \quad (24)$$

$$M_t = \beta_1 + \beta_2 t + \beta_3 (T_t - \tilde{T}) + \beta_4 (T_t - \tilde{T})^2 + w_t, \quad (25)$$

$$M_t = \beta_1 + \beta_2 t + \beta_3 (T_t - \tilde{T}) + \beta_4 (T_t - \tilde{T})^2 + \beta_5 P_t + w_t, \quad (26)$$

donde M_t , denota las muertes cardiovasculares, T_t la temperatura, P_t el nivel de partículas y $\tilde{T} = 74,6$ es la temperatura promedio, misma que se ajustó para evitar problemas de escala. El modelo (23) es sólo la tendencia lineal, (24) es el modelo lineal en la temperatura, (25) es el modelo cuadrático en la temperatura y (26) es el modelo cuadrático en la temperatura y el número de partículas de contaminación. En la Tabla 2 se muestran los estadísticos principales de cada modelo. Los valores de la R^2 fueron computados usando $SSE_1 = 50,687$ y (18).

Es notorio que el mejor modelo es el que incluye a la temperatura, la temperatura cuadrada y el número de partículas, ya que recupera el 60% de la varianza consiguiendo el menor AIC y BIC. De lo que se obtiene el mejor modelo de predicción para la mortalidad

$$M_t = 81,59 - 0,027t - 0,473 (T_t - 74,6) + 0,023 (T_t - 74,6)^2 + 0,255P_t,$$

como era esperado, tiene una tendencia negativa en el tiempo, así como un coeficiente negativo para la temperatura ajustada. El efecto cuadrático de la temperatura se ve claramente en el gráfico de dispersión de la Figura 3. Se interpreta de la siguiente manera: Al aumentar una unidad de contaminación el número de muertes diarias aumenta en 0.255. Por otro lado es esencial revisar los residuales por la posible autocorrelación, la cual se discutirá más adelante.

Modelo	k	SSE	gl	MSE	R^2	AIC	BIC
(23)	2	40,020	506	79.0	0.21	5.38	5.40
(24)	3	31,413	505	62.2	0.38	5.14	5.17
(25)	4	27,985	504	55.5	0.45	5.03	5.07
(26)	5	20,508	503	40.8	0.60	4.27	4.77

Cuadro 2: Resumen de los modelos de la Mortalidad Cardiovascular

Dirección General de Bibliotecas de la UAQ

3. Modelos autorregresivos y de promedios móviles

Como se vio en el capítulo anterior, en la regresión clásica se busca ver cómo una variable dependiente es influenciada por los valores actuales contemporáneos de variables independientes. En el caso de las series de tiempo se permitirá que variables dependientes también sean influenciadas por valores pasados de variables tanto independientes como dependientes.

3.1. Modelos autorregresivos

La idea básica es explicar el valor actual de la serie x_t como una función de sus p valores pasados, $x_{t-1}, x_{t-2}, \dots, x_{t-p}$, con p determinando el número de pasos hacia atrás necesarios para predecir el valor actual.

Definición 3.1.1 Un modelo autorregresivo de orden p , denotado por $AR(p)$ es de la forma

$$x_t = \phi_1 x_{t-1} + \phi_2 x_{t-2} + \dots + \phi_p x_{t-p} + w_t, \quad (27)$$

donde x_t es estacionario, $w_t \sim N(0, \sigma_w^2)$, y $\phi_1, \phi_2, \dots, \phi_p$ son constantes con $\phi_p \neq 0$. La media de x_t en (27) es cero. En el caso que la media μ de x_t no sea cero se puede reemplazar x_t por $x_t - \mu$ en (27), obteniendo

$$x_t - \mu = \phi_1 (x_{t-1} - \mu) + \phi_2 (x_{t-2} - \mu) + \dots + \phi_p (x_{t-p} - \mu) + w_t,$$

o escribir

$$x_t = \alpha \phi_1 x_{t-1} + \phi_2 x_{t-2} + \dots + \phi_p x_{t-p} + w_t, \quad (28)$$

con $\alpha = \mu(1 - \phi_1 - \dots - \phi_p)$.

Para ver esto de una forma más simple se utilizará el operador rezago definido de la siguiente manera.

Definición 3.1.2 Se denota el operador rezago como

$$Bx_t = x_{t-1}$$

y extendido a sus potencias $B^2 x_t = B(Bx_t) = Bx_{t-1} = x_{t-2}$ y así sucesivamente. De modo que,

$$B^k x_t = x_{t-k}. \quad (29)$$

Definición 3.1.3 El operador inverso puede ser dado de tal forma que $B^{-1}B = 1$, así

$$x_t = B^{-1}Bx_t = B^{-1}x_{t-1},$$

el operador B^{-1} es llamado el operador prospectivo.

El modelo $AR(p)$ (27) puede reescribirse usando tales operadores como

$$(1 - \phi_1 B - \phi_2 B^2 - \dots - \phi_p B^p) x_t = w_t, \quad (30)$$

o de manera más concisa

$$\phi(B) x_t = w_t, \quad (31)$$

donde $\phi(B)$ es

$$\phi(B) = 1 - \phi_1 B - \phi_2 B^2 - \dots - \phi_p B^p. \quad (32)$$

Las propiedades de $\phi(B)$ son importantes para reexpresar (31).

Primero al considerar el modelo de primer orden $AR(1)$, dado por $x_t = \phi x_{t-1} + w_t$, iterando hacia atrás k veces se obtiene

$$\begin{aligned} x_t &= \phi x_{t-1} + w_t = \phi(\phi x_{t-2} + w_{t-1}) + w_t \\ &= \phi^2 x_{t-2} + \phi w_{t-1} + w_t \\ &\vdots \\ &= \phi^k x_{t-k} + \sum_{j=0}^{k-1} \phi^j w_{t-j}. \end{aligned}$$

Este método sugiere que, al continuar iterando hacia atrás y bajo los supuestos de que $|\phi| < 1$ y $\sup_t \text{var}(x_t) < \infty$, se puede representar un modelo AR (1) como un proceso lineal dado por

$$x_t = \sum_{j=0}^{\infty} \phi^j w_{t-j}. \quad (33)$$

Así, el proceso AR (1) definido anteriormente es estacionario con media

$$E(x_t) = \sum_{j=0}^{\infty} \phi^j E(w_{t-j}) = 0,$$

y función de autocovarianza

$$\begin{aligned} \gamma(h) &= \text{cov}(x_{t+h}, x_t) = E \left[\left(\sum_{j=0}^{\infty} \phi^j w_{t+h-j} \right) \left(\sum_{k=0}^{\infty} \phi^k w_{t-k} \right) \right] \\ &= E \left[(w_{t+h} + \dots + \phi^h w_t + \phi^{h+1} w_{t-1} + \dots) (w_t + \phi w_{t-1} + \dots) \right] \\ &= \sigma_w^2 \sum_{j=0}^{\infty} \phi^{h+j} \phi^j = \sigma_w^2 \phi^h \sum_{j=0}^{\infty} \phi^{2j} = \frac{\sigma_w^2 \phi^h}{1 - \phi^2}, \text{ con } h \geq 0. \end{aligned} \quad (34)$$

Además, ya que $\gamma(h) = \gamma(-h)$, sólo se presentará la autocovarianza para las $h \geq 0$.

Definiendo la función de autocorrelación del mismo modo que Shumway (2015):

Definición 3.1.4 La función de autocorrelación (ACF) es definida como

$$\rho(s, t) = \frac{\gamma(s, t)}{\sqrt{\gamma(s, s)\gamma(t, t)}}$$

Así, la ACF de un modelo AR (1) de (34), será

$$\rho(h) = \frac{\gamma(h)}{\gamma(0)} = \phi^h, \text{ con } h \geq 0, \quad (35)$$

y $\rho(h)$ satisface la recursión

$$\rho(h) = \phi \rho(h-1), \text{ para } h = 1, 2, \dots \quad (36)$$

La técnica de iteración hacia atrás para obtener la representación estacionaria de los modelos AR funciona bien cuando $p = 1$, pero para órdenes grandes no es así.

Sea el modelo AR (1) con el operador de la forma

$$\phi(B) x_t = w_t, \quad (37)$$

donde $\phi(B) = 1 - \phi B$, y $|\phi| < 1$. Así, se puede escribir el modelo de la forma (33) usando el operador de la forma

$$x_t = \sum_{j=0}^{\infty} \psi_j w_{t-j} = \psi(B) w_t, \quad (38)$$

donde $\psi(B) = \sum_{j=0}^{\infty} \psi_j B^j$ y $\psi_j = \phi^j$, y el supuesto de que se desconoce que $\psi_j = \phi^j$. Es posible sustituir $\psi(B) w_t$ de (38) en x_t de (37) obteniendo

$$\phi(B) \psi(B) w_t = w_t. \quad (39)$$

Lo cual llevaría a

$$(1 - \phi B) (1 + \psi B + \phi B^2 + \dots + \psi_j B^j + \dots) = 1. \quad (40)$$

Reorganizando

$$1 + (\psi_1 - \phi) B + (\psi_2 - \psi_1 \phi) B^2 + \dots + (\psi_j - \psi_{j-1} \phi) B^j + \dots = 1$$

por lo que para cada $j = 1, 2, \dots$, el coeficiente de B^j debe ser cero. De esta forma, el coeficiente para B es $(\psi_1 - \phi) = 0$, lo que implica que $\psi_1 = \phi$. Y así, el coeficiente para B^2 es $(\psi_2 - \psi_1 \phi) = 0$, luego $\psi_2 = \phi^2$. Y en general,

$$\psi_j = \psi_{j-1} \phi,$$

con $\psi_0 = 1$, obteniendo $\psi_j = \phi^j$.

Ejemplo 3.1.2 La trayectoria de una muestra para un proceso AR (1)

La Figura 4 muestra dos procesos AR (1), uno usando a $\phi = 0,9$ y otro con $\phi = -0,9$ ambos con $\sigma_w^2 = 1$. Para el primero, $\rho(h) = (0,9)^h$, con $h \geq 0$ y así, las observaciones cercanas en tiempo están correlacionadas entre ellas de manera positiva. Esto significa que las observaciones en puntos de tiempos contiguos tenderán a valores cercanos entre sí. Este hecho se muestra en parte superior de la Figura 4 como una trayectoria de x_t , la cual se muestra muy suave. En contraste en el caso con $\phi = -0,9$, con una autocorrelación $\rho(h) = (-0,9)^h$, con $h \geq 0$. El resultado significa que las observaciones en puntos del tiempo contiguo están correlacionadas de forma negativa, pero las observaciones separadas en dos puntos del tiempo están correlacionadas de manera positiva. Lo anterior es mostrado en la parte inferior de la Figura 4, donde si la observación x_t es positiva, la siguiente observación x_{t+1} será negativa, y la siguiente observación x_{t+2} será positiva. Como resultado, en este caso, el camino de la muestra contiene numerosos picos contrastantes.

imagenes/transfer tifo max/ejemplo2.1.1.png

Figura 4: Procesos AR (1).

3.2. Modelos de promedios móviles

Definición 3.2.1 El modelo de promedios móviles de orden q denotado como $MA(q)$, está definido como

$$x_t = w_t + \theta_1 w_{t-1} + \theta_2 w_{t-2} + \dots + \theta_q w_{t-q}, \quad (41)$$

donde $w_t \sim N(0, \sigma_w^2)$ y $\theta_1, \theta_2, \dots, \theta_q$ son los parámetros con $\theta_q \neq 0$.

Escribiendo los modelos con promedios móviles de la misma manera que en el proceso lineal (38), tomando a $\psi_0 = 1$, $\psi_j = \theta_j$ para $j = 1, \dots, q$ y $\psi_j = 0$ para los otros valores. Así se podría escribir el proceso $MA(q)$ en la equivalente forma

$$x_t = \theta(B) w_t, \quad (42)$$

donde

$$\theta(B) = 1 + \theta_1 B + \theta_2 B^2 + \dots + \theta_q B^q. \quad (43)$$

A diferencia del proceso autorregresivo, los promedios móviles son un proceso estacionario para cualesquiera valores de los parámetros $\theta_1, \dots, \theta_q$.

Como en el caso del modelo AR, el polinomio $\theta(z)$, correspondiente al operador de promedios móviles $\theta(B)$, será útil para explorar las propiedades generales de los procesos MA. Por ejemplo, siguiendo los pasos de las ecuaciones (37)-(40) se puede escribir el modelo $MA(1)$ como $x_t = \theta(B) w_t$, donde $\theta(B) = 1 + \theta B$. Y si $|\theta| < 1$,

se tendrá el modelo como $\pi(B)x_t = w_t$, donde $\pi(B) = \theta^{-1}(B)$. Sea $\theta(z) = 1 + \theta z$, para $|z| \leq 1$, entonces $\pi(z) = \theta^{-1}(z) = \frac{1}{1+\theta z} = \sum_{j=0}^{\infty} (-\theta)^j z^j$, con lo que $\pi(B) = \sum_{j=0}^{\infty} (-\theta)^j B^j$.

3.3. Modelos autorregresivos y de promedios móviles.

Ahora se verá la mezcla de modelos AR y MA, los modelos autorregresivos con promedios móviles, para series de tiempo estacionarias.

Definición 3.3.1 Una serie de tiempo $\{x_t; t = 0, \pm 1, \pm 2, \dots\}$ es **ARMA** (p, q) si es estacionaria y

$$x_t = \phi_1 x_{t-1} + \dots + \phi_p x_{t-p} + w_t + \theta_1 w_{t-1} + \dots + \theta_q w_{t-q}, \quad (44)$$

con $\phi_p \neq 0 \neq \theta_q$, $w_t \sim N(0, \sigma_w^2)$ y $\sigma_w^2 > 0$. Los parámetros p y q son llamados órdenes autorregresivos y de promedios móviles respectivamente. Si x_t tiene media μ , se define $\alpha = \mu(1 - \phi_1 - \dots - \phi_p)$ y se escribe el modelo como

$$x_t = \alpha \phi_1 x_{t-1} + \dots + \phi_p x_{t-p} + w_t + \theta_1 w_{t-1} + \dots + \theta_q w_{t-q}, \quad (45)$$

donde $w_t \sim N(0, \sigma_w^2)$.

Es claro que, cuando $q = 0$ se obtiene un modelo autorregresivo de orden p , y cuando $p = 0$ es un modelo con promedios móviles de orden q . De la misma forma, al utilizar los operadores AR en (32) y MA en (43) el modelo ARMA (p, q) en (45) puede ser visto como

$$\phi(B)x_t = \theta(B)w_t. \quad (46)$$

Los modelos ARMA (p, q) en (44) o su equivalente en (46), pueden tener inconvenientes:

- Parámetros redundantes en el modelo.
- Modelos estacionarios AR que dependan del futuro.
- Modelos con MA que no sean únicos.

Para sobrellevar estos problemas, se añade una restricción adicional al modelo.

Definición 3.3.2 Los polinomios AR y MA están definidos como

$$\phi(z) = 1 - \phi_1 z - \dots - \phi_p z^p, \quad \phi_p \neq 0, \quad (47)$$

y

$$\theta(z) = 1 + \theta_1 z + \dots + \theta_q z^q, \quad \theta_q \neq 0, \quad (48)$$

respectivamente, donde z es un número complejo.

Para solucionar el primer problema, de ahora en adelante se hará referencia a un modelo ARMA (p, q) de la forma simple. Esto en adición a la definición original dada en (44), requiere que los polinomios $\phi(z)$ y $\theta(z)$ no tengan factores en común. De este modo el proceso $x_t = 0,5x_{t-1} - 0,5w_{t-1} + w_t$ no es un proceso ARMA $(1, 1)$ ya que en su forma reducida x_t es ruido blanco.

Para evitar el problema de que el modelo dependa del futuro, se introduce el concepto de causalidad.

Definición 3.3.3. Un modelo ARMA (p, q) se dice que es causal si la serie de tiempo $\{x_t : t = 0, \pm 1, \pm 2, \dots\}$ puede ser escrita como un proceso lineal de un solo lado como

$$x_t = \sum_{j=0}^{\infty} \psi_j w_{t-j} = \psi(B)w_t, \quad (49)$$

donde $\psi(B) = \sum_{j=0}^{\infty} \psi_j B^j$, $\sum_{j=0}^{\infty} |\psi_j| < \infty$, y $\psi_0 = 1$.

El proceso AR (1) , $x_t = \phi x_{t-1} + w_t$ es causal cuando $|\phi| < 1$. Equivalentemente, el proceso es causal, si y sólo si, la raíz de $\phi(z) = 1 - \phi z$ es más grande que 1 en valor absoluto. Esto es, la raíz, z_0 de $\phi(z)$ es $z_0 = \frac{1}{\phi}$, ya que $\phi(z_0) = 0$ y $|z_0| > 1$ ya que $|\phi| < 1$.

El estudio de estos proceso es importante ya que para crear un modelo que trate de predecir una variable con otra se utiliza la función de transferencia, la cual como parte del proceso utiliza el ajustar modelos ARMA.

4. Análisis espectral y filtración

En este capítulo se introducirán conceptos como el de periodicidad y el periodograma, el cual nos ayudará a descomponer una serie en subseries las cuales contribuyen de distinta manera para formar la serie original. Así mismo se verán las funciones de autocorrelación y autocorrelación parcial, las cuales nos ayudan a determinar el grado de los modelos ARMA. Y para finalizar el capítulo se verán los modelos ARIMA y SARIMA para el estudio de series de tiempo no estacionarias y estacionales.

4.1. Periodicidad y comportamiento ciclado de una serie temporal

Se define la frecuencia ω de una serie temporal como el número de ciclos por unidad de tiempo. Su periodo como la cantidad de observaciones en un ciclo, eso es, $\frac{1}{\omega}$.

Para determinar la tasa para la cual la serie oscila, se considerará un ciclo como un periodo completo de una función seno o coseno, definido sobre una unidad en un intervalo de tiempo.

Al tomar un proceso periódico

$$x_t = A \cos(2\pi\omega t + \phi) \quad (50)$$

para $t = 0, \pm 1, \pm 2, \dots$, donde ω es el índice de frecuencia, definido en número de ciclos por unidad de tiempo, A determina la amplitud de la función, y ϕ , llamada fase, determina el punto de inicio de la función coseno. Para facilitar su estudio se usará una identidad trigonométrica, con la que (50) es equivalente a

$$x_t = U_1 \cos(2\pi\omega t) + U_2 \sin(2\pi\omega t), \quad (51)$$

donde $U_1 = A \cos \phi$ y $U_2 = -A \sin \phi$ son comúnmente variables aleatorias normales. En este caso la amplitud es $A = \sqrt{U_1^2 + U_2^2}$ y la fase $\phi = \tan^{-1}\left(-\frac{U_2}{U_1}\right)$.

Se necesitan al menos 2 puntos para determinar un ciclo, así que la mayor frecuencia de interés será de $\frac{1}{2}$ ciclos por lapso. Esta llamada frecuencia plegable, define la mayor frecuencia que puede ser vista en una muestra discreta.

La generalización (51) permite la mezcla de periodos con múltiples frecuencias y amplitudes,

$$x_t = \sum_{k=1}^q [U_{k1} \cos(2\pi\omega_k t) + U_{k2} \sin(2\pi\omega_k t)] \quad (52)$$

donde U_{k1}, U_{k2} para $k = 1, 2, \dots, q$ son variables aleatorias independientes con media cero y varianza σ_k^2 y los ω_k son frecuencias distintas. Dicho proceso tiene como función de autocovarianza

$$\gamma(h) = \sum_{k=1}^q \sigma_k^2 \cos(2\pi\omega_k h) \quad (53)$$

la cual es una suma de componentes periódicos con pesos proporcionales a la varianza σ_k^2 . Por lo que x_t es un proceso estacionario con media cero y varianza

$$\gamma(0) = E(x_t^2) = \sum_{k=1}^q \sigma_k^2, \quad (54)$$

Ejemplo 4.1.1 Una serie periódica

Sea $t = 1, \dots, 100$, y las siguientes tres series

$$\begin{aligned} x_{t1} &= 2 \cos\left(2\pi t \frac{6}{100}\right) + 3 \sin\left(2\pi t \frac{6}{100}\right) \\ x_{t2} &= 4 \cos\left(2\pi t \frac{10}{100}\right) + 5 \sin\left(2\pi t \frac{10}{100}\right) \\ x_{t3} &= 6 \cos\left(2\pi t \frac{40}{100}\right) + 7 \sin\left(2\pi t \frac{40}{100}\right) \end{aligned}$$

Mismas que se grafican en la Figura 5 con sus correspondientes frecuencias y amplitudes cuadradas. Después se construye

$$x_t = x_{t1} + x_{t2} + x_{t3}$$

la cual también es graficada. Es preciso notar que la clasificación sistemática de los componentes de frecuencia es esencial en una serie cronológica, incluyendo sus contribuciones relativas.

imagenes/fig4_1.png

Figura 5: Componentes periódicos y su suma

El llamado periodograma permite descubrir componentes periódicos de una serie de tiempo. Aquel está dado por

$$P\left(\frac{j}{n}\right) = \left(\frac{2}{n} \sum_{t=1}^n x_t \cos\left(\frac{2\pi t j}{n}\right)\right)^2 + \left(\frac{2}{n} \sum_{t=1}^n x_t \sin\left(\frac{2\pi t j}{n}\right)\right)^2 \quad (55)$$

puede ser considerado una medida de la correlación cuadrada de los datos con oscilaciones senoidales a la frecuencia de $\omega_j = \frac{j}{n}$ o j ciclos en n tiempos. El periodograma de los datos simulados previamente y mostrado en la Figura (6), identifica los tres componentes x_{t1} , x_{t2} y x_{t3} de x_t .

E:/tesis/Modelos de Salmonela/imagenes/periodograma fig4_2.png

Figura 6: Periodograma de los datos generados en el ejemplo anterior

Nótese que

$$P\left(\frac{j}{n}\right) = P\left(1 - \frac{j}{n}\right), \quad j = 0, 1, \dots, n-1,$$

así que es un efecto de espejo, por lo que, comúnmente no se grafican las frecuencias mayores a la frecuencia del doblez.

También nótese que los pesos que se muestran en el periodograma escalado son

$$P\left(\frac{6}{100}\right) = 13, \quad P\left(\frac{10}{100}\right) = 41, \quad P\left(\frac{40}{100}\right) = 85,$$

con $P\left(\frac{j}{n}\right) = P\left(1 - \frac{j}{n}\right)$ y $P\left(\frac{j}{n}\right) = 0$ en otro caso. Estas son los valores exactos de las amplitudes cuadradas de los componentes generados en el ejemplo anterior. Lo que sugiere que el periodograma puede proveer una descomposición de los componentes de las varianzas de los datos reales.

Si se consideran los datos x_t como un color, hecho de una mezcla de colores primarios x_{t1}, x_{t2}, x_{t3} a diferentes intensidades, entonces se podría pensar al periodograma como un prisma que descompone el color x_t en sus colores primarios, de esta idea se deriva el término “análisis espectral”.

4.2. Densidad espectral

Propiedad 4.2.1 Representación espectral de un proceso estacionario

De forma no técnica, cualquier serie de tiempo estacionaria puede ser vista, aproximadamente, como la superposición aleatoria de senos y cosenos que oscilan a varias frecuencias.

Ejemplo 4.2.1 Un proceso periódico estacionario

Al tomar el proceso estacionario como en (51) con una frecuencia fija ω_0 ,

$$x_t = U_1 \cos(2\pi\omega_0 t) + U_2 \sin(2\pi\omega_0 t),$$

donde U_1 y U_2 son variables aleatorias independientes con media cero y varianza σ^2 . El número de periodos de tiempo que se necesitan para completar un ciclo es $\frac{1}{\omega_0}$, y al proceso le toma exactamente ω_0 ciclos por tiempo para $t = 0, \pm 1, \pm 2 \dots$ pero

$$\begin{aligned}\gamma(h) &= \sigma^2 \cos(2\pi\omega_0 h) = \frac{\sigma^2}{2} \exp(-2\pi i\omega_0 h) + \frac{\sigma^2}{2} \exp(2\pi i\omega_0 h) \\ &= \int_{-\frac{1}{2}}^{\frac{1}{2}} \exp(2\pi i\omega h) dF(\omega)\end{aligned}$$

Expresando en términos de la integral de Riemann-Stieltjes, con $F(\omega)$ una función definida por

$$F(\omega) = \begin{cases} 0 & \omega < -\omega_0 \\ \frac{\sigma^2}{2} & -\omega_0 < \omega < \omega_0 \\ \sigma^2 & \omega > \omega_0 \end{cases}$$

$F(\omega)$ se comporta como una función de distribución acumulativa para una variable aleatoria discreta, a excepción de que $F(\infty) = \sigma^2 = \text{var}(x_t)$ en vez de 1. Por lo cual, se le asigna el término de distribución espectral.

En general, si x_t es estacionaria con autocovarianza $\gamma(h) = E[(x_{t+h} - \mu)(x_t - \mu)]$, entonces existe una única función $F(\omega)$ monótonamente creciente y acotada, llamada función espectral de distribución, con $F(-\infty) = F(-\frac{1}{2}) = 0$, y $F(\infty) = F(\frac{1}{2}) = \gamma(0)$ tal que

$$\gamma(h) = \int_{-\frac{1}{2}}^{\frac{1}{2}} \exp(2\pi i\omega h) dF(\omega). \quad (56)$$

En el caso cuando la función de autocovarianza es absolutamente sumable, entonces la función de distribución espectral es absolutamente continua con $dF(\omega) = f(\omega) d\omega$.

Propiedad 4.2.2 La densidad espectral

Si la función de autocovarianza, $\gamma(h)$, de un proceso estacionario satisface

$$\sum_{h=-\infty}^{\infty} |\gamma(h)| < \infty, \quad (57)$$

entonces tiene por representación

$$\gamma(h) = \int_{-\frac{1}{2}}^{\frac{1}{2}} \exp(2\pi i\omega h) f(\omega) d\omega, \quad h = 0, \pm 1, \pm 2, \dots \quad (58)$$

como la transformación inversa de la densidad espectral, la cual se expresa como

$$f(\omega) = \sum_{h=-\infty}^{\infty} \gamma(h) \exp(-2\pi i\omega h) - \frac{1}{2} \leq \omega \leq \frac{1}{2}. \quad (59)$$

Esta densidad espectral es análoga a la función de probabilidad, el hecho de que $\gamma(h)$ sea definida no negativa asegura que $f(\omega) \geq 0$ para todo ω . Se sigue de (59) que f es par, i.e. $f(\omega) = f(-\omega)$ y tomando $h = 0$ en (58) lleva a que

$$\gamma(0) = \text{var}(x_t) = \int_{-\frac{1}{2}}^{\frac{1}{2}} f(\omega) d\omega,$$

la que expresa la varianza total como la densidad espectral integrada sobre todas las frecuencias.

Si la Propiedad 4.2 es satisfecha, la función de autocovarianza $\gamma(h)$ y la densidad espectral $f(\omega)$ contienen la misma información, pero expresada de diferente manera. La función de autocovarianza expresará la información

en término de rezagos, mientras que la densidad espectral expresará la misma información en términos de ciclos. Algunos problemas son mas fáciles de manejar cuando se considera la información rezagada y se tiende a manipular dichos problemas en el dominio del tiempo. Sin embargo, otros problemas son más fáciles de trabajar cuando se considera la información periódica y se tienden a manejar en el dominio del espectro. Los filtros lineales usan un conjunto de coeficientes específicos, a_j , $j = 0, \pm 1, \pm 2, \dots$, para transformar una serie de entrada x_t , y producir una serie de salida y_t de la forma

$$y_t = \sum_{j=-\infty}^{\infty} a_j x_{t-j}, \quad \sum_{j=-\infty}^{\infty} |a_j| < \infty. \quad (60)$$

Los coeficientes son llamados función de respuesta al impulso, y la transformación de Fourier

$$A(\omega) = \sum_{j=-\infty}^{\infty} a_j \exp[-2\pi i \omega j], \quad (61)$$

es llamada la función de respuesta a la frecuencia. Bajo el supuesto de que y_t tiene una densidad espectral $f_y(\omega)$ y x_t tiene una densidad espectral $f_x(\omega)$, entonces se tendrá el siguiente resultado:

Propiedad 4.2.3 Espectro de salida de una serie estacionaria filtrada

El espectro de una salida filtrada y_t en (60) está relacionado al espectro de una entrada x_t mediante

$$f_y(\omega) = |A(\omega)|^2 f_x(\omega), \quad (62)$$

donde la función de respuesta a la frecuencia $A(\omega)$ está definida como en (61)

Si x_t es ARMA, su densidad espectral puede ser obtenida de manera explícita usando el hecho de que es un proceso lineal, es decir, $x_t = \sum_{j=0}^{\infty} \psi_j w_{t-j}$, donde $\sum_{j=0}^{\infty} |\psi_j| < \infty$, que la densidad espectral del ruido blanco es $f_w(\omega) = \sigma_w^2$ y $\psi(z) = \frac{\theta(z)}{\phi(z)}$.

Propiedad 4.2.4 La densidad espectral de un ARMA. Si x_t es ARMA (p, q) , $\phi(B)x_t = \theta(B)w_t$, su densidad espectral está dada por

$$f_x(\omega) = \sigma_w^2 \frac{|\theta[\exp(-2\pi i \omega)]|^2}{|\phi[\exp(-2\pi i \omega)]|^2}, \quad (63)$$

donde $\phi(z) = 1 - \sum_{k=1}^p \phi_k z^k$ y $\theta(z) = 1 + \sum_{k=1}^q \theta_k z^k$.

4.3. Periodograma y la discretización de la transformación de Fourier

Definición 4.3.1 Dado un conjunto de datos x_1, \dots, x_n , se define la discretización de la transformada de Fourier, ((DFT) por sus siglas en inglés), como

$$d(\omega_j) = n^{-\frac{1}{2}} \sum_{t=1}^n x_t \exp(-2\pi i \omega_j t) \quad (64)$$

para $j = 0, 1, \dots, n-1$, donde las frecuencias $\omega_j = \frac{j}{n}$ son llamadas frecuencias fundamentales.

En ocasiones es de utilidad explorar la inversa del DFT, dada por

$$x_t = n^{-\frac{1}{2}} \sum_{j=0}^{n-1} d(\omega_j) \exp(2\pi i \omega_j t) \quad (65)$$

para $t = 1, \dots, n$.

Ahora se definirá el periodograma en términos de los módulos cuadrados de la DFT, recordando que si $z = a + ib$, entonces su conjugado es $\bar{z} = a - ib$, y su módulo $|z|^2 = z\bar{z} = a^2 + b^2$.

Definición 4.3.2 Dada una base de datos x_1, \dots, x_n , se define el periodograma como

$$I(\omega_j) = |d(\omega_j)|^2 \quad (66)$$

para $j = 0, 1, 2, \dots, n-1$.

Notando que $I(0) = n\bar{x}^2$, donde \bar{x} es el promedio de la muestra. Y que $\sum_{t=1}^n \exp(-2\pi i t \frac{j}{n}) = 0$ para $j \neq 0$, el DFT se puede ver como

$$d(\omega_j) = n^{-\frac{1}{2}} \sum_{t=1}^n (x_t - \bar{x}) \exp(-2\pi i \omega_j t) \quad (67)$$

para $j \neq 0$. Así, para $j \neq 0$,

$$\begin{aligned} I(\omega_j) &= |d(\omega_j)|^2 \\ &= n^{-1} \sum_{t=1}^n \sum_{s=1}^n (x_t - \bar{x})(x_s - \bar{x}) \exp(-2\pi i \omega_j (t-s)) \end{aligned} \quad (68)$$

$$\begin{aligned} &= n^{-1} \sum_{h=-(n-1)}^{n-1} \sum_{t=1}^{n-|h|} (x_{t+|h|} - \bar{x})(x_t - \bar{x}) \exp(-2\pi i \omega_j h) \\ &= \sum_{h=-(n-1)}^{n-1} \hat{\gamma}(h) \exp(-2\pi i \omega_j h) \end{aligned} \quad (69)$$

donde $h = t - s$, con $\hat{\gamma}(h)$ como la función de autocovarianza.

Se puede pensar en el periodograma, $I(\omega_j)$, como una muestra de la densidad espectral de x_t .

En ocasiones es útil trabajar con la parte real e imaginaria del DFT por partes separadas. Para esto, se definen las siguientes transformaciones.

Definición 4.3.3 Dada una base de datos x_1, \dots, x_n , definimos la transformada coseno como

$$d_c(\omega_j) = n^{-\frac{1}{2}} \sum_{t=1}^n x_t \cos(2\pi \omega_j t) \quad (70)$$

y la transformada seno como

$$d_s(\omega_j) = n^{-\frac{1}{2}} \sum_{t=1}^n x_t \sin(2\pi \omega_j t) \quad (71)$$

donde $\omega_j = \frac{j}{n}$ para $j = 0, 1, \dots, n-1$.

Y como $d(\omega_j) = d_c(\omega_j) - i d_s(\omega_j)$, entonces

$$I(\omega_j) = d_c^2(\omega_j) + d_s^2(\omega_j). \quad (72)$$

Propiedad 4.3.1 Causalidad en un proceso ARMA (p, q) . Un modelo ARMA (p, q) es causal si y sólo si $\phi(z) \neq 0$ para $|z| \leq 1$. Los coeficientes de un proceso lineal dado en (49) pueden ser determinados resolviendo

$$\psi(z) = \sum_{j=0}^{\infty} \psi_j z^j = \frac{\theta(z)}{\phi(z)}, \quad |z| \leq 1.$$

Otra forma de ver la propiedad anterior, es que un proceso ARMA es causal sólo cuando las raíces $\phi(z)$ yacen fuera del círculo unitario.

Por último, para atacar el problema de unicidad, se escoge un modelo que permita una representación infinita autorregresiva.

Definición 4.3.4 Se dice que un modelo ARMA (p, q) es invertible, si la serie de tiempo $\{x_t : t = 0, \pm 1, \pm 2, \dots\}$ puede ser escrita como

$$\pi(B)x_t = \sum_{j=0}^{\infty} \pi_j x_{t-j} = w_t, \quad (73)$$

donde $\pi(B) = \sum_{j=0}^{\infty} \pi_j B^j$ y $\sum_{j=0}^{\infty} |\pi_j| < \infty$, fijando $\pi_0 = 1$.

De manera análoga a la propiedad 4.3 se tiene la siguiente propiedad

Propiedad 4.3.2 Invertibilidad de un proceso ARMA (p, q) .

Un modelo ARMA (p, q) es invertible, si y sólo si $\theta(z) \neq 0$ para $|z| \leq 1$. El coeficiente π_j de $\pi(B)$ dado en (73) puede ser determinado resolviendo

$$\pi(z) = \sum_{j=0}^{\infty} \pi_j z^j = \frac{\phi(z)}{\theta(z)}, \quad |z| \leq 1.$$

Otra forma de ver la propiedad 4.3 es que un proceso ARMA será invertible cuando las raíces de $\theta(z)$ se encuentran fuera del círculo unitario.

Ejemplo 4.3.1 Parámetros redundantes, causalidad e invertibilidad.

Considérese el proceso

$$x_t = 0,4x_{t-1} + 0,45x_{t-2} + w_t + w_{t-1} + 0,25w_{t-2},$$

o, en forma de operador,

$$(1 - 0,4B - 0,45B^2)x_t = (1 + B + 0,25B^2)w_t.$$

En primera instancia, x_t parece ser un proceso ARMA $(2, 2)$, pero se puede notar que

$$\phi(B) = 1 - 0,4B - 0,45B^2 = (1 + 0,5B)(1 - 0,9B)$$

y

$$\theta(B) = (1 + B + 0,25B^2) = (1 + 0,5B)^2$$

tienen un factor en común que puede ser cancelado; después de cancelarlo, los operadores son $\phi(B) = (1 - 0,9B)$ y $\theta(B) = (1 + 0,5B)$ y el modelo es un ARMA $(1, 1)$, $(1 - 0,9B)x_t = (1 + 0,5B)w_t$, o

$$x_t = 0,9x_{t-1} + 0,5w_{t-1} + w_t.$$

El modelo es causal ya que $\phi(z) = (1 - 0,9z) = 0$ cuando $z = \frac{10}{9}$, lo cual yace fuera del círculo unitario. Este modelo también es invertible, ya que la raíz de $\theta(z) = 1 + 0,5z$ es $z = -2$, el cual también yace fuera del círculo unitario.

Para escribir el modelo como un proceso lineal, se pueden obtener los pesos ψ usando la Propiedad 4.3.1, $\phi(z)\psi(z) = \theta(z)$, o

$$(1 - 0,9z)(1 + \psi_1 z + \psi_2 z^2 + \dots + \psi_j z^j + \dots) = 1 + 0,5z.$$

Reordenando, se tiene que

$$1 + (\psi_1 - 0,9)z + (\psi_2 - 0,9\psi_1)z^2 + \dots + (\psi_j - 0,9\psi_{j-1})z^j + \dots = 1 + 0,5z.$$

Igualando los coeficientes de z se obtiene que $\psi_1 - 0,9 = 0,5$ y $\psi_j - 0,9\psi_{j-1} = 0$ para $j > 1$. Así, $\psi_j = 1,4(0,9)^{j-1}$ para $j \geq 1$ y el modelo puede ser escrito como

$$x_t = w_t + 1,4 \sum_{j=0}^{\infty} 0,9^{j-1} w_{t-j}.$$

La representación invertible usando la Propiedad 4.3.1 es obtenida mediante la igualación de los coeficientes $\theta(z)\pi(z) = \phi(z)$

$$(1 + 0,5z)(1 + \pi_1 z + \pi_2 z^2 + \pi_3 z^3 + \dots) = 1 - 0,9z.$$

En este caso los pesos π están dados por $\pi_j = (-1)^j 1,4(0,5)^{j-1}$, para $j \geq 1$, así se puede escribir como

$$x_t = -1,4 \sum_{j=1}^{\infty} (-0,5)^{j-1} x_{t-j} + w_t.$$

4.4. Autocorrelación y autocorrelación parcial

Sea la ACF de un proceso MA (q), $x_t = \theta(B) w_t$, así x_t es una combinación lineal de ruido blanco; dicho proceso es estacionario con media

$$E(x_t) = \sum_{j=0}^q \theta_j E(w_{t-j}) = 0$$

y autocovarianza

$$\begin{aligned} \gamma(h) &= \text{cov}(x_{t+h}, x_t) \\ &= \text{cov}\left(\sum_{j=0}^q \theta_j w_{t+h-j}, \sum_{k=0}^q \theta_k w_{t-k}\right) \\ &= \begin{cases} \sigma_w^2 \sum_{j=0}^{q-h} \theta_j \theta_{j+h}, & 0 \leq h \leq q \\ 0 & h > q \end{cases} \end{aligned} \quad (74)$$

como $\gamma(h) = \gamma(-h)$, sólo se mostrarán los valores positivos. Note que $\gamma(q)$ no puede ser cero debido a que $\theta_q \neq 0$. Y el corte de $\gamma(h)$ después del rezago q es la forma en la que el modelo MA (q) se comporta. Por otro lado, dividiendo (74) por $\gamma(0)$ lleva al ACF para un MA (q):

$$\rho(h) = \begin{cases} \frac{\sum_{j=0}^{q-h} \theta_j \theta_{j+h}}{1 + \theta_1^2 + \dots + \theta_q^2} & 1 \leq h \leq q, \\ 0 & h > q. \end{cases} \quad (75)$$

Y para un modelo ARMA (p, q) de la forma $\phi(B) x_t = \theta(B) w_t$,

$$x_t = \sum_{j=0}^{\infty} \psi_j w_{t-j}. \quad (76)$$

Se sigue inmediatamente que $E(x_t) = 0$. También, la función de autocovarianza de x_t puede ser escrita como

$$\gamma(h) = \text{cov}(x_{t+h}, x_t) = \sigma_w^2 \sum_{j=0}^{\infty} \psi_j \psi_{j+h}, \quad h \geq 0, \quad (77)$$

se puede resolver para $\gamma(h)$ y la ACF es $\rho(h) = \gamma(h) / \gamma(0)$. Así

$$\gamma(h) = \sum_{j=1}^p \phi_j \gamma(h-j) + \sigma_w^2 \sum_{j=h}^q \theta_j \psi_{j-h}, \quad h \geq 0, \quad (78)$$

donde se usa el hecho de que para $h \geq 0$

$$\text{cov}(w_{t+h}, w_t) = \psi_{j-h} \sigma_w^2.$$

De (78), se puede escribir una ecuación general homogénea para la ACF de un proceso causal ARMA:

$$\gamma(h) - \phi_1 \gamma(h-1) - \dots - \phi_p \gamma(h-p) = 0, \quad h \geq \max(p, q+1) \quad (79)$$

con condiciones iniciales

$$\gamma(h) - \sum_{j=1}^p \phi_j \gamma(h-j) = \sigma_w^2 \sum_{j=h}^q \theta_j \psi_{j-h}, \quad 0 \leq h \leq \max(p, q+1) \quad (80)$$

Dividiendo (79) y (80) por $\gamma(0)$ permite resolver para la ACF, $\rho(h) = \gamma(h) / \gamma(0)$.

4.5. Función de Autocorrelación parcial (PACF)

Como se vio en (75), para modelos con promedios móviles de orden q , la ACF tiene que ser cero para rezagos mayores a q . Aún más, debido a que $\theta_q \neq 0$, el ACF no será cero en el rezago q . Así el ACF provee una cantidad considerable de información sobre el orden de dependencia cuando es un proceso de promedios móviles. Sin embargo, si es un proceso ARMA o AR la ACF indica muy poco sobre el orden de dependencia. Considerando esto, se obtiene una función para estos modelos, que actúe de la misma manera que el ACF para los modelos MA. Así, para los modelos AR tenemos la función de autocorrelación parcial (PACF).

Es pertinente recordar que si X , Y y Z son variables aleatorias, entonces la correlación parcial entre X y Y dado Z es obtenida mediante la regresión de X sobre Z para obtener \hat{X} , la regresión de Y sobre Z para obtener \hat{Y} , y así calculando

$$\rho_{XY|Z} = \text{corr}(X - \hat{X}, Y - \hat{Y}).$$

La idea es que $\rho_{XY|Z}$ mida la correlación entre X y Y con el efecto lineal de Z removido (o parcialmente removido). Ahora considérese el modelo AR (1), $x_t = \phi x_{t-1} + w_t$. Entonces

$$\begin{aligned} \gamma_x(2) &= \text{cov}(x_t, x_{t-2}) \\ &= \text{cov}(\phi^2 x_{t-2} + \phi w_{t-1} + w_t, x_{t-2}) \\ &= \phi^2 \gamma_x(0) \end{aligned}$$

Este resultado se sigue de la causalidad, ya que x_{t-2} involucra $\{w_{t-2}, w_{t-3}, \dots\}$, los cuales no están correlacionados con w_t y w_{t-1} . La correlación entre x_t y x_{t-2} no es cero, debido a que x_t es dependiente de x_{t-2} a través de x_{t-1} . Al suponer que se rompe la cadena de dependencia removiendo (o quitando parcialmente) el efecto de x_{t-1} . Esto es considerar la correlación entre $x_t - \phi x_{t-1}$ y $x_{t-2} - \phi x_{t-1}$, lo cual sería la correlación entre x_t y x_{t-2} con la dependencia lineal de x_{t-1} removida en ambos. Considerando esto se tendría que

$$\text{cov}(x_t - \phi x_{t-1}, x_{t-2} - \phi x_{t-1}) = \text{cov}(w_t, x_{t-2} - \phi x_{t-1}) = 0.$$

Ya con la PACF se puede ver la correlación entre x_s y x_t con los efectos lineales entre ellos removidos.

Definiendo formalmente el PACF para series de tiempo estacionarias. Sea \hat{x}_{t+h} para $h \geq 2$, se denota la regresión de x_{t+h} sobre $\{x_{t+h-1}, x_{t+h-2}, \dots, x_{t+1}\}$ el cual se escribe como

$$\hat{x}_{t+h} = \beta_1 x_{t+h-1} + \beta_2 x_{t+h-2} + \dots + \beta_{h-1} x_{t+1}. \quad (81)$$

En lo cual no se necesita intercepto alguno debido a que x_t tiene media cero; de no ser el caso, se puede reemplazar x_t por $x_t - \mu$. Además, \hat{x}_t denota la regresión de x_t sobre $\{x_{t+1}, x_{t+2}, \dots, x_{t+h-1}\}$. Entonces

$$\hat{x}_t = \beta_1 x_{t+1} + \beta_2 x_{t+2} + \dots + \beta_{h-1} x_{t+h-1}. \quad (82)$$

Definición 4.5.1 La función de autocorrelación parcial (PACF) de un proceso estacionario x_t denotado por ϕ_{hh} para $h = 1, 2, \dots$ es

$$\phi_{11} = \text{corr}(x_{t+1}, x_t) = \rho(1) \quad (83)$$

y

$$\phi_{hh} = \text{corr}(x_{t+h} - \hat{x}_{t+h}, x_t - \hat{x}_t), \quad h \geq 2. \quad (84)$$

Ambos $(x_{t+h} - \hat{x}_{t+h})$ y $(x_t - \hat{x}_t)$ están correlacionados con $\{x_{t+1}, \dots, x_{t+h-1}\}$. La PACF, ϕ_{hh} , es la correlación entre x_{t+h} y x_t con la dependencia lineal de $\{x_{t+1}, \dots, x_{t+h-1}\}$ removida.

La PACF para modelos MA se comporta como el ACF para los modelos AR. También, la PACF para los AR tienen un comportamiento similar al ACF para los modelos MA. Ya que un modelo invertible ARMA tiene una representación infinita de AR, el PACF no se cortará. El comportamiento de la ACF y PACF se resumen en la tabla 3

	AR(p)	MA(q)	ARMA(p, q)
ACF	disminuyen lentamente	se corta después del retraso q	disminuyen lentamente
PACF	se corta después del retraso p	disminuyen lentamente	disminuyen lentamente

Cuadro 3: Comportamiento del ACF y PACF para modelos ARMA

4.6. Predicción

La meta es predecir los valores futuros de una serie de tiempo x_{n+m} , para $m = 1, 2, \dots$ basados en los datos recolectados hasta el presente $\mathbf{x} = \{x_n, x_{n-1}, \dots, x_1\}$. Bajo el supuesto de que el modelo de x_t es conocido y que ya la serie ya ha sido estacionarizada.

La predicción de mínimos cuadrados de x_{n+m} es

$$x_{n+m}^n = E(x_{n+m} | \mathbf{x}) \quad (85)$$

debido a que la esperanza condicional minimiza el error cuadrático medio

$$E[x_{n+m} - g(\mathbf{x})]^2, \quad (86)$$

donde $g(\mathbf{x})$ es una función de las observaciones \mathbf{x} .

Tomando los predictores lineales de los datos de la forma

$$x_{n+m}^n = \alpha_0 + \sum_{k=1}^n \alpha_k x_k \quad (87)$$

donde α_i son números reales, se ve que las α 's dependen de n y m . Por ejemplo si $m = n = 1$, entonces x_2^1 es un paso adelante de la predicción lineal de x_2 dado x_1 . En términos de (87) $x_2^1 = \alpha_0 + \alpha_1 x_1$. Pero si $n = 2$, x_3^2 es un paso adelante de la predicción lineal de x_3 dado x_1 y x_2 . Nuevamente en términos de (87) $x_3^2 = \alpha_0 + \alpha_1 x_1 + \alpha_2 x_2$, y en general, las α 's en x_2^1 y x_3^2 serán diferentes.

El predictor lineal de la forma (87) que minimice el error cuadrático medio (86) es llamado el mejor predictor lineal.

Propiedad 4.6.1 *El mejor predictor lineal para procesos estacionarios (BLP)*

Dado los datos x_1, \dots, x_n , el mejor predictor lineal, $x_{n+m}^n = \alpha_0 + \sum_{k=1}^n \alpha_k x_k$, de x_{n+m} para $m \geq 1$, se encuentra resolviendo

$$E[(x_{n+m} - x_{n+m}^n) x_k] = 0, \quad k = 0, 1, \dots, n, \quad (88)$$

donde $x_0 = 1$, para $\alpha_0, \alpha_1, \dots, \alpha_n$.

La ecuación especificada anteriormente es llamada la ecuación de predicción, y es usada para encontrar los coeficientes $\{\alpha_0, \alpha_1, \dots, \alpha_n\}$. Si $E(x_t) = \mu$, la primera ecuación ($k = 0$) de (88) implica

$$E(x_{n+m}^n) = E(x_{n+m}) = \mu.$$

Así, calculando la esperanza en (86), se tiene

$$\mu = \alpha_0 + \sum_{k=1}^n \alpha_k \mu \quad \text{ó} \quad \alpha_0 = \mu \left(1 - \sum_{k=1}^n \alpha_k \right).$$

Por lo tanto, la forma del BLP es

$$x_{n+m}^n = \mu + \sum_{k=1}^n \alpha_k (x_k - \mu).$$

Primero se considerará la predicción un paso adelante. Esto es, dado los datos $\{x_1, \dots, x_n\}$, se desea predecir el valor de la serie de tiempo en el siguiente punto, x_{n+1} . El BLP de x_{n+1} es de la forma

$$x_{n+1}^n = \phi_{n1} x_n + \phi_{n2} x_{n-1} + \dots + \phi_{nn} x_1, \quad (89)$$

considerando el caso general donde $\mu = 0$ y $\alpha_0 = 0$, y se muestra la dependencia de los coeficientes en n ; en este caso, α_k en (87) es $\phi_{n, n+1-k}$ en (89), para $k = 1, \dots, n$. Usando la Predicción, los coeficientes $\{\phi_{n1}, \phi_{n2}, \dots, \phi_{nn}\}$ satisfacen

$$E \left[\left(x_{n-1} - \sum_{j=1}^n \phi_{nj} x_{n+1-j} \right) x_{n+1-k} \right] = 0, \quad k = 1, \dots, n,$$

ó

$$\sum_{j=1}^n \phi_{nj} \gamma(k-j) = \gamma(k), \quad k = 1, \dots, n. \quad (90)$$

La ecuación de predicción anterior se puede escribir de manera matricial como

$$\Gamma_n \phi_n = \gamma_n, \quad (91)$$

donde $\Gamma_n = (\gamma(k-j))_{j,k=1}^n$ es una matriz de $n \times n$, $\phi_n = (\phi_{n1}, \dots, \phi_{nn})'$ es un vector de $n \times 1$, y $\gamma_n = (\gamma(1), \dots, \gamma(n))'$ es un vector de $n \times 1$.

La matriz Γ_n es no negativa definida. Si Γ_n es singular, existen muchas soluciones para (91), pero, por el teorema de proyección, x_{n+1}^n es única. Y si Γ_n es no singular, los elementos de ϕ_n son únicos, y están dados por

$$\phi_n = \Gamma_n^{-1} \gamma_n. \quad (92)$$

Para modelos ARMA, el hecho de que $\sigma_w^2 > 0$ y $\gamma(h) \rightarrow 0$ cuando $h \rightarrow \infty$ es suficiente para asegurar que Γ_n es definida positiva. La predicción a un paso adelante, escrita en notación vectorial, será

$$x_{n+1}^n = \phi_n^T \mathbf{x} \quad (93)$$

y donde $\mathbf{x} = (x_n, x_{n-1}, \dots, x_1)^T$.

El error cuadrático medio de la predicción a un paso es

$$P_{n+1}^n = E(x_{n+1} - x_{n+1}^n)^2 = \gamma(0) - \gamma_n^T \Gamma_n^{-1} \gamma_n. \quad (94)$$

Para verificar (94) usando (93) y (92),

$$\begin{aligned} E(x_{n+1} - x_{n+1}^n)^2 &= E(x_{n+1} - \phi_n^T \mathbf{x})^2 = E(x_{n+1} - \gamma_n^T \Gamma_n^{-1} \gamma_n)^2 \\ &= E(x_{n+1}^2 - 2\gamma_n^T \Gamma_n^{-1} \mathbf{x} x_{n+1} + \gamma_n^T \Gamma_n^{-1} \mathbf{x} \mathbf{x}^T \Gamma_n^{-1} \gamma_n) \\ &= \gamma(0) - 2\gamma_n^T \Gamma_n^{-1} \gamma_n + \gamma_n^T \Gamma_n^{-1} \Gamma_n \Gamma_n^{-1} \gamma_n \\ &= \gamma(0) - \gamma_n^T \Gamma_n^{-1} \gamma_n. \end{aligned}$$

En general para modelos ARMA las ecuaciones de predicción no serán tan simples como un caso AR. Además, para n grandes, el uso de (92) se vuelve complicado, debido a que se requiere invertir una matriz muy grande. Para esto se tienen otras alternativas que no requieren invertir matrices.

Proposición 4.6.2 El algoritmo de Durbin-Levinson

Las ecuaciones (92) y (94) pueden ser resueltas de manera iterativa de la siguiente manera:

$$\phi_{00} = 0, P_1^0 = \gamma(0) \quad (95)$$

para $n \geq 1$,

$$\phi_{nn} = \frac{\rho(n) - \sum_{k=1}^{n-1} \phi_{n-1,k} \rho(n-k)}{1 - \sum_{k=1}^{n-1} \phi_{n-1,k} \rho(k)}, P_{n+1}^n = P_n^{n-1} (1 - \phi_{nn}^2), \quad (96)$$

donde, para $n \geq 2$,

$$\phi_{nk} = \phi_{n-1,k} - \phi_{nn} \phi_{n-1,n-k}, k = 1, 2, \dots, n-1. \quad (97)$$

4.7. Estimación

Ahora se supondrá que se tienen n observaciones x_1, \dots, x_n de un proceso invertible ARMA(p, q) en el cual, inicialmente se supondrá que p y q son conocidos, y la meta es estimar los parámetros $\phi_1, \dots, \phi_p, \theta_1, \dots, \theta_q$ y σ_w^2 .

Una manera de hacerlo es mediante el método de momentos, pero a pesar de que dicho método es bueno, puede dar estimaciones no óptimas. Para evitar esto se usará un método que lleve al estimador óptimo

Cuando el proceso es AR(p), las primeras $p+1$ ecuaciones de (79) y (80) llevan a lo siguiente.

Definición 4.7.1 La ecuación Yule-Walker está dada por

$$\gamma(h) = \phi_1 \gamma(h-1) + \dots + \phi_p \gamma(h-p), h = 1, 2, \dots, p \quad (98)$$

$$\sigma_w^2 = \gamma(0) - \phi_1 \gamma(1) - \dots - \phi_p \gamma(p). \quad (99)$$

En notación matricial

$$\Gamma_P \phi = \gamma_p, \sigma_w^2 = \gamma(0) - \phi' \gamma_p, \quad (100)$$

donde $\Gamma_p = \{\gamma(k-j)\}_{j,k=1}^p$ una matriz de $p \times p$, $\phi = (\phi_1, \dots, \phi_p)'$, un vector de $p \times 1$, y $\gamma_p = (\gamma(1), \dots, \gamma(p))'$ un vector de $p \times 1$. Usando el método de momentos, se reemplaza $\gamma(h)$ por $\hat{\gamma}(h)$ y se resuelve

$$\hat{\phi} = \hat{\Gamma}_p^{-1} \hat{\gamma}_p, \hat{\sigma}_w^2 = \hat{\gamma}(0) - \hat{\gamma}'_p \hat{\Gamma}_p^{-1} \hat{\gamma}_p. \quad (101)$$

Obteniendo los estimadores comúnmente llamados estimadores Yule-Walker.

Para modelos AR(p), si la muestra es grande, los estimadores Yule-Walker siguen una distribución normal, y $\hat{\sigma}_w^2$ es cercana al valor de σ_w^2 .

En el caso de los modelos AR(p), los estimadores Yule-Walker dan una estimación óptima, pero si se usa el método de momentos para modelos MA o ARMA, no se obtendrán estimadores óptimos, ya que tales procesos son no lineales.

4.7.1. Máxima verosimilitud

Se partirá de un modelo AR(1)

$$x_t = \mu + \phi(x_{t-1} - \mu) + w_t \quad (102)$$

donde $|\phi| < 1$ y $w_t \sim N(0, \sigma_w^2)$ IID. Dado los datos x_1, x_2, \dots, x_n , la verosimilitud resulta en

$$\begin{aligned} L(\mu, \phi, \sigma_w^2) &= f(x_1, x_2, \dots, x_n | \mu, \phi, \sigma_w^2) \\ &= f(x_1) f(x_2 | x_1) \cdots f(x_n | x_{n-1}), \end{aligned}$$

y como $x_t | x_{t-1} \sim N(\mu + \phi(x_{t-1} - \mu), \sigma_w^2)$, se tiene que

$$f(x_t | x_{t-1}) = f_w((x_t - \mu) - \phi(x_{t-1} - \mu)),$$

donde $f_w(\cdot)$ es la densidad de w_t ; es normal con media cero y varianza σ_w^2 . Así,

$$L(\mu, \phi, \sigma_w^2) = f(x_1) \prod_{t=2}^n f_w[(x_t - \mu) - \phi(x_{t-1} - \mu)].$$

Y para encontrar $f(x_1)$ se puede usar la representación

$$x_1 = \mu + \sum_{j=0}^{\infty} \phi^j w_{1-j},$$

con lo que x_1 es normal con media μ y varianza $\sigma_w^2 / (1 - \phi^2)$. Finalmente, para un AR(1) la verosimilitud es

$$L(\mu, \phi, \sigma_w^2) = (2\pi\sigma_w^2)^{-\frac{n}{2}} (1 - \phi^2)^{\frac{1}{2}} \exp\left[-\frac{S(\mu, \phi)}{2\sigma_w^2}\right], \quad (103)$$

donde

$$S(\mu, \phi) = (1 - \phi^2)(x_1 - \mu)^2 + \sum_{t=2}^n [(x_t - \mu) - \phi(x_{t-1} - \mu)]^2, \quad (104)$$

llamada la suma incondicional de cuadrados.

Tomando la derivada parcial del log de (103) respecto de σ_w^2 e igualando a cero se obtiene que para cualesquiera parámetros μ y ϕ , $\sigma_w^2 = n^{-1} S(\mu, \phi)$ maximiza la verosimilitud. Así el estimador máximo verosímil de σ_w^2 es

$$\hat{\sigma}_w^2 = n^{-1} S(\hat{\mu}, \hat{\phi}) \quad (105)$$

donde $\hat{\mu}$ y $\hat{\phi}$ son los estimadores MV de μ y ϕ respectivamente. Y, si se sustituye n en la ecuación anterior por $n - 2$, se obtiene un estimador incondicional de mínimos cuadrados para σ_w^2 .

Tomando los logaritmos en (103) reemplazando σ_w^2 por $\hat{\sigma}_w^2$ e ignorando las constantes, $\hat{\mu}$ y $\hat{\phi}$ son los valores que minimizan los criterios de la función

$$l(\mu, \phi) = \log [n^{-1} S(\mu, \phi)] - n^{-1} \log (1 - \phi^2); \quad (106)$$

esto es, $l(\mu, \phi) \propto -2 \log L(\mu, \phi, \hat{\sigma}_w^2)$. La minimización de $l(\mu, \phi)$ se logra mediante una aproximación numérica. Para los modelos AR se tiene la ventaja de que se puede condicionar a valores iniciales. Esto es, que podemos hacer a un lado el término que causa no linealidad de la verosimilitud, condicionado a x_1 ; así la verosimilitud resulta en

$$\begin{aligned} L(\mu, \phi, \sigma_w^2 | x_1) &= \prod_{t=2}^n f_w [(x_t - \mu) - \phi(x_{t-1} - \mu)] \\ &= (2\pi\sigma_w^2)^{-(n-1)/2} \exp \left[-\frac{S_c(\mu, \phi)}{2\sigma_w^2} \right], \end{aligned} \quad (107)$$

donde la suma de cuadrados condicional es

$$S_c(\mu, \phi) = \sum_{t=2}^n [(x_t - \mu) - \phi(x_{t-1} - \mu)]^2. \quad (108)$$

Y la condicional de máxima verosimilitud de σ_w^2 es

$$\hat{\sigma}_w^2 = S_c(\hat{\mu}, \hat{\phi}) / (n - 1) \quad (109)$$

donde $\hat{\mu}$ y $\hat{\phi}$ son los valores que minimizan la condicional de la suma de cuadrados $S_c(\mu, \phi)$. Finalmente, tomando a $\alpha = \mu(1 - \phi)$, la suma de cuadrados condicional puede ser escrita como

$$S_c(\mu, \phi) = \sum_{t=2}^n [x_t - (\alpha + \phi x_{t-1})]^2. \quad (110)$$

Para modelos generales AR (p), la estimación MV se sigue de manera análoga como para el ejemplo AR (1).

Para modelos ARMA es un tanto más complicado escribir la verosimilitud como una función explícita de los parámetros. En lugar de eso, se usarán los errores de predicción de un paso adelante, $x_t - x_t^{t-1}$.

Para modelos ARMA (p, q), sea $\beta = (\mu, \phi_1, \dots, \phi_p, \theta_1, \dots, \theta_q)'$ un vector $(p + q + 1)$ dimensional de los parámetros del modelo. La verosimilitud puede ser escrita como

$$L(\beta, \sigma_w^2) = \prod_{t=1}^n f(x_t | x_{t-1}, \dots, x_1).$$

La distribución condicional de x_t dada por x_{t-1}, \dots, x_1 es Normal con media x_t^{t-1} y varianza $P_t^{t-1} = \gamma(0) \prod_{j=1}^{t-1} (1 - \phi_{jj}^2)$.

Para modelos ARMA, $\gamma(0) = \sigma_w^2 \sum_{j=0}^{\infty} \psi_j^2$, que también se puede escribir como

$$P_t^{t-1} = \sigma_w^2 \left\{ \left[\sum_{j=0}^{\infty} \psi_j^2 \right] \left[\prod_{j=1}^{t-1} (1 - \phi_{jj}^2) \right] \right\} = \sigma_w^2 r_t.$$

Nótese que el término r_t es una función sólo de los parámetros de regresión, mismos que pueden ser computados recursivamente mediante $r_{t+1} = (1 - \phi_{tt}^2) r_t$ con la condición inicial $r_1 = \sum_{j=0}^{\infty} \psi_j^2$. Así, la verosimilitud de los datos puede ser escrita como

$$L(\beta, \sigma_w^2) = (2\pi\sigma_w^2)^{-n/2} [r_1(\beta) r_2(\beta) \cdots r_n(\beta)]^{-1/2} \exp \left[-\frac{S(\beta)}{2\sigma_w^2} \right], \quad (111)$$

donde

$$S(\beta) = \sum_{t=1}^n \left[\frac{(x_t - x_t^{t-1}(\beta))^2}{r_t(\beta)} \right]. \quad (112)$$

Como x_t^{t-1} y r_t son funciones de β solamente, dados los valores de β y σ_w^2 la verosimilitud puede ser evaluada.

Ahora el estimador máximo verosímil se calculará maximizando (111) con respecto de β y σ_w^2 . Al igual que en el ejemplo con AR (1) se tiene que

$$\sigma_w^2 = n^{-1} S(\hat{\beta}), \quad (113)$$

donde $\hat{\beta}$ es el valor de β que minimiza la verosimilitud

$$l(\beta) = \log [n^{-1} S(\beta)] + n^{-1} \sum_{t=1}^n \log r_t(\beta). \quad (114)$$

Como $x_1^0 = \mu$ y $x_t^{t-1} = \mu + \phi(x_{t-1} - \mu)$, para $t = 2, \dots, n$, al considerar el hecho que $\phi_{11} = \phi$ y $\phi_{hh} = 0$ para $h > 1$, tenemos que $r_1 = \sum_{j=0}^{\infty} \phi^{2j} = (1 - \phi^2)^{-1}$, $r_2 = (1 - \phi^2)^{-1} (1 - \phi^2) = 1$, y en general, $r_t = 1$ para $t = 2, \dots, n$.

4.8. Modelos integrados para datos no estacionarios

EL ARMA integrado, o ARIMA, es una ampliación del modelo de la clase ARMA que incluye la diferencia.

Definición 4.8.1 *Un proceso x_t se dice ser ARIMA (p, d, q) si*

$$\nabla^d x_t = (1 - B)^d x_t$$

es ARMA (p, q) . En general, se escribe el modelo como

$$\phi(B) (1 - B)^d x_t = \theta(B) w_t. \quad (115)$$

Si $E(\nabla^d x_t) = \mu$ se puede escribir el modelo como

$$\phi(B) (1 - B)^d x_t = \delta + \theta(B) w_t,$$

donde $\delta = \mu(1 - \phi_1 - \dots - \phi_p)$.

Como $y_t = \nabla^d x_t$ es ARMA, se usan los métodos de predicción para obtener y_t , lo cual a su vez lleva a predecir a x_t . Por ejemplo, si $d = 1$, dadas las predicciones de y_{n+m}^n para $m = 1, 2, \dots$ se tiene $y_{n+m}^n = x_{n+m}^n - x_{n+m-1}^n$, así

$$x_{n+m}^n = y_{n+m}^n + x_{n+m-1}^n$$

con la condición inicial de $x_{n+1}^n = y_{n+1}^n + x_n$.

4.9. Identificación de modelos ARIMA

Para la construcción de los modelos ARIMA se siguen varios pasos básicos. Graficar los datos, pensar en una posible transformación de los mismos, identificar los órdenes de dependencia del modelo, estimar los parámetros, diagnosticar y elegir el modelo.

Primero, como en cualquier análisis de datos, se debe construir la gráfica de los datos en el tiempo, e inspeccionar en busca de anomalías. Luego de observarse un fenómeno, como que la varianza de los datos va aumentando con respecto del tiempo, indicaría la necesidad de transformar los datos para estabilizar la varianza. El siguiente paso es identificar los valores preliminares para el orden autorregresivo, p , el orden diferencial, d , y el orden de promedio móvil, q . La gráfica de los datos en el tiempo comúnmente sugiere si alguna diferencia es necesaria. Ahora, de suceder, la primera diferencia de los datos, $d = 1$, se deberá inspeccionar y graficar con respecto del tiempo. Ya con el valor preliminar de d , lo siguiente es mirar la ACF y PACF de $\nabla^d x_t$ para escoger los valores de p y q .

El siguiente paso es diagnosticar el ajuste del modelo. Esto incluye el análisis de residuales así como la compresión del modelo. Nuevamente, el primer paso es graficar los residuales con respecto del tiempo, $x_t - \hat{x}_t^{t-1}$, o de los residuales estandarizados

$$e_t = \frac{x_t \hat{x}_t^{t-1}}{\sqrt{\hat{P}_t^{t-1}}} \quad (116)$$

donde \hat{x}_t^{t-1} es la predicción de un paso adelante de x_t basado en el ajuste del modelo y \hat{P}_t^{t-1} es el estimador paso adelante del error de la varianza. Si el modelo ajusta bien, los residuales estandarizados deberían ser una secuencia IID con media cero y varianza uno.

El paso final es escoger el modelo. Esto es, qué modelo se usará para la predicción. Las técnicas más populares son el AIC, AICc, y el BIC, en el contexto de regresión.

4.10. Regresión con errores autocorrelacionados

Considerese el modelo de regresión

$$y_t = \sum_{j=1}^r \beta_j z_{tj} + x_t \quad (117)$$

donde x_t es un proceso con alguna función de covarianza $\gamma_x(s, t)$.

Escribiendo en notación vectorial. Sea $\mathbf{y} = Z\boldsymbol{\beta} + \mathbf{x}$ y $\Gamma = \{\gamma_x(s, t)\}$. donde $\Gamma^{-1/2}\mathbf{y} = \Gamma^{-1/2}Z\boldsymbol{\beta} + \Gamma^{-1/2}\mathbf{x}$. Entonces, se puede escribir el modelo como $\mathbf{y}^* = Z^*\boldsymbol{\beta} + \boldsymbol{\delta}$, con una matriz identidad de covarianza $\boldsymbol{\delta}$. Se sigue que la ponderación estimada de $\boldsymbol{\beta}$ es $\hat{\boldsymbol{\beta}}_w = (Z'\Gamma^{-1}Z)^{-1}Z'\Gamma^{-1}\mathbf{y}$, y la matriz de varianza y covarianza del estimador es $\text{var}(\hat{\boldsymbol{\beta}}_w) = (Z'\Gamma^{-1}Z)^{-1}$.

Es común suponer estacionaridad en la estructura de la covarianza para el proceso del error x_t .

Si el proceso del error es ARMA, entonces se usará la transformación $\pi(B)x_t = w_t$, donde $\pi(B) = \theta(B)^{-1}\phi(B)$:

$$S(\phi, \theta, \beta) = \sum_{t=1}^n w_t^2 = \sum_{t=1}^n \left[\pi(B)y_t - \sum_{j=1}^r B_j \pi(B)z_{tj} \right]^2$$

Ahora, el principal problema es que no se conoce cómo es el comportamiento del ruido x_t a priori. Una forma sencilla de resolver esto es la siguiente:

- Primero, correr una regresión ordinaria de y_t sobre z_{t1}, \dots, z_{tr} , como si los errores no estuviesen correlacionados. Conservando los residuales, $\hat{x}_t = y_t - \sum_{j=1}^r \hat{\beta}_j z_{tj}$.
- Identificar el modelo o los modelos ARMA para los residuales \hat{x}_t .
- Correr las ponderaciones de mínimos cuadrados o MV para los modelos de regresión con errores autocorrelacionados usando el modelo especificado en el paso anterior.
- Inspeccionar los residuales \hat{w}_t para asegurar que sean ruido blanco y ajusta el modelo de ser necesario.

4.11. Modelos estacionales ARIMA

Se van a modificar los modelos ARIMA para poder considerar la estacionalidad y el comportamiento no estacionario. En ocasiones, la dependencia del pasado tiende a ocurrir fuertemente en múltiplos de algún rezago estacional s , ya sea debido a una variabilidad natural, física, biológica o un proceso económico, tiende a tener fluctuaciones estacionales. Debido a esto se introduce un polinomio autorregresivo y con promedios móviles que pueda identificar los rezagos estacionales. El resultado puramente estacional autorregresivo con promedios móviles, ARMA $(P, Q)_s$, toma la forma

$$\Phi_P(B^s)x_t = \Theta_Q(B^s)w_t, \quad (118)$$

donde los operadores

$$\Phi_P(B^s) = 1 - \Phi_1 B^s - \Phi_2 B^{2s} - \dots - \Phi_P B^{Ps} \quad (119)$$

y

$$\Theta_Q(B^s) = 1 + \Theta_1 B^s + \Theta_2 B^{2s} + \dots + \Theta_Q B^{Qs} \quad (120)$$

son el operador estacional autorregresivo y el operador estacional de promedios móviles, de orden P y Q respectivamente, con periodo estacional s .

Ejemplo 4.11.1 Serie estacional AR

Se toma una serie estacional autorregresiva de primer orden con un período mensual escrita como

$$(1 - \Phi B^{12}) x_t = w_t$$

ó

$$x_t = \Phi x_{t-12} + w_t$$

Este modelo presenta una serie x_t en términos de un rezago pasado múltiplo de un periodo estacional anual, $s = 12$ meses. Claramente la estimación y la predicción para tal proceso involucra solamente las modificaciones directas en la unidad de rezago ya tratada. En particular, la condición causal requiere que $|\Phi| < 1$.

Definición 3.11.1 *El modelo multiplicativo estacional autorregresivo integrado y con promedios móviles, SARIMA está dado por*

$$\Phi_P(B^s) \phi(B) \nabla_s^D \nabla^d x_t = \delta + \Theta_Q(B^s) \theta(B) w_t, \quad (121)$$

donde w_t es ruido blanco. Se suele denotar como $ARIMA(p, d, q) \times (P, D, Q)_s$. Los componentes autorregresivos y de promedios móviles ordinarios son representados por los polinomios $\phi(B)$ y $\theta(B)$ de órdenes p y q respectivamente, y los componentes estacionales autorregresivos y de promedios móviles como $\Phi_P(B^s)$ y $\Theta_Q(B^s)$ de órdenes P y Q ; y los componentes diferenciales ordinarios y estacionales por $\nabla^d = (1 - B)^d$ y $\nabla_s^D = (1 - B^s)^D$.

5. Regresión rezagada: Modelando la función de transferencia

Para el interés de predecir la serie de salida, y_t , a partir de la serie de entrada, x_t , se podría considerar el modelo de regresión

$$y_t = \sum_{j=0}^{\infty} \alpha_j x_{t-j} + \eta_t = \alpha(B) x_t + \eta_t, \quad (122)$$

donde $\sum_j |\alpha_j| < \infty$. Se supone que el proceso de entrada, x_t y el de ruido, η_t , en (122) son ambos estacionarios e independientes pero no necesariamente. Los coeficientes $\alpha_0, \alpha_1, \dots$ describen los pesos asignados a los valores pasados de x_t usados para predecir y_t , por lo cual se usará la siguiente notación

$$\alpha(B) = \sum_{j=0}^{\infty} \alpha_j B^j, \quad (123)$$

donde B^j es un operador de rezago, el cual cumple que $B^j x_t = x_{t-j}$.

Se deben identificar modelos ARIMA, digase ARIMA (p, d, q) y ARIMA (p_η, d_η, q_η) , a las series x_t y η_t respectivamente. Para el caso de modelar un ARMA (p, q) para la entrada x_t y para el ruido η_t , se tendrá la siguiente representación

$$\phi(B) x_t = \theta(B) w_t \quad (124)$$

y

$$\phi_\eta(B) \eta_t = \theta_\eta(B) z_t, \quad (125)$$

donde w_t y z_t son ruidos blancos independientes con varianzas σ_w^2 y σ_z^2 , respectivamente.

Box y Jenkins (1970) propone que en algunos casos $\alpha(B)$ podría ser expresado como un cociente de polinomios involucrando un número pequeño de coeficientes, a su vez con un rezago específico, d , como

$$\alpha(B) = \frac{\delta(B) B^d}{\omega(B)}, \quad (126)$$

donde

$$\omega(B) = 1 - \omega_1 B - \omega_2 B^2 - \dots - \omega_r B^r \quad (127)$$

y

$$\delta(B) = \delta_0 + \delta_1 B + \dots + \delta_s B^s \quad (128)$$

con d indicando el rezago en el cual la serie x_t no es influyente sobre y_t hasta el periodo $t + d$, o

$$y_t = 0x_t + 0x_{t-1} + 0x_{t-2} + \dots + \omega_0 x_{t-d}.$$

El valor s indica en qué medida la serie de salida y_t continua siendo influenciada por la serie de entrada x_t , i.e. y_t es influenciada por $(x_{t-d}, x_{t-d-1}, \dots, x_{t-d-s})$. Y finalmente, el valor r indica en qué medida y_t es influenciada por sus propios valores pasados, i.e. y_t es influenciada por $(y_{t-1}, y_{t-2}, \dots, y_{t-r})$. Por cuestiones de comodidad, se representará el operador inverso de un polinomio, $[\omega(B)]^{-1}$ como $\frac{1}{\omega(B)}$.

Especificar un modelo parsimonioso que involucre una forma simple para $\alpha(B)$ y estimar los parámetros en el modelo anterior son la tarea principal de la metodología en la función de transferencia.

Suponiendo que encontramos un modelo ARIMA para la serie de entrada x_t y se aplicará este operador a ambos lados de (122), se obtendría el siguiente modelo

$$\tilde{y}_t = \frac{\phi(B)}{\theta(B)} y_t = \alpha(B) w_t + \frac{\phi(B)}{\theta(B)} \eta_t = \alpha(B) w_t + \tilde{\eta}_t, \quad (129)$$

donde w_t y el ruido transformado $\tilde{\eta}_t$ son independientes.

La serie w_t es una versión preblanqueada de la serie de entrada, y su correlación cruzada con la serie transformada \tilde{y}_t será

$$\gamma_{\tilde{y}w}(h) = E(\tilde{y}_{t+h} w_t) = E \left[\sum_{j=0}^{\infty} \alpha_j w_{t+h-j} w_t \right] = \sigma_w^2 \alpha_h, \quad (130)$$

debido a que la función de autocovarianza de un ruido blanco será cero excepto cuando $j = h$ en (130).

En algunos casos se escribirán los componentes de $\delta(B)$ y $\omega(B)$ de manera separada, así, se podrá encontrar la ecuación

$$y_t = \frac{\delta(B) B^d}{\omega(B)} x_t + \eta_t$$

como

$$\omega(B) y_t = \delta(B) B^d x_t + \omega(B) \eta_t,$$

o en su forma de regresión

$$y_t = \sum_{k=1}^r \omega_k y_{t-k} + \sum_{k=0}^s \delta_k x_{t-d-k} + u_t, \quad (131)$$

donde

$$u_t = \omega(B) \eta_t. \quad (132)$$

La expresión (131) sugiere realizar una regresión en versiones rezagadas tanto de la entrada x_t y la salida y_t , para obtener $\hat{\beta}$, el vector de $(r + s + 1) \times 1$ de los estimados de la regresión

$$\beta = (\omega_1, \dots, \omega_r, \delta_0, \delta_1, \dots, \delta_s)'$$

Los residuales de la regresión anterior,

$$\hat{u}_t = y_t - \hat{\beta}' z_t,$$

donde

$$z_t = (y_{t-1}, \dots, y_{t-r}, x_{t-d}, \dots, x_{t-d-s})',$$

podrían ser usados para especificar un ARMA para el proceso de ruido η_t , ya que se puede estimar ese proceso de (132), usando \hat{u}_t y $\hat{\omega}(B)$ y aplicando el operador de promedios móviles para obtener $\hat{\eta}_t$. Así, ajustar un ARMA (p_η, q_η) para estimar el proceso de ruido completaría la especificación.

En el capítulo siguiente se reproducirá un ejemplo del libro de Shumway y Stoffer (2015), con la serie ambiental Southern Oscillation Index (SOI) y su Reclutamiento (Recruitment).

6. Ejemplo del Southern Oscillation Index y el Reclutamiento

En el libro de Shumway (2015) se presenta el ejemplo del Índice de Oscilación del Sur y el reclutamiento de nuevos peces donde se consideran los datos durante un periodo de 453 meses entre los años 1950 a 1987. El estudio de dicho ejemplo es relevante para la futura comprensión de un caso más complejo, por ello se reproduce a continuación.

Se tienen los datos del Índice de Oscilación del Sur, SOI, por sus siglas en inglés, el cual mide los cambios en la presión del aire relacionada a las temperaturas superficiales del Océano Pacífico. Y el reclutamiento, que es el valor estimado del número de nuevos peces, Reclutamiento, los cuales se muestran en la Figura 7

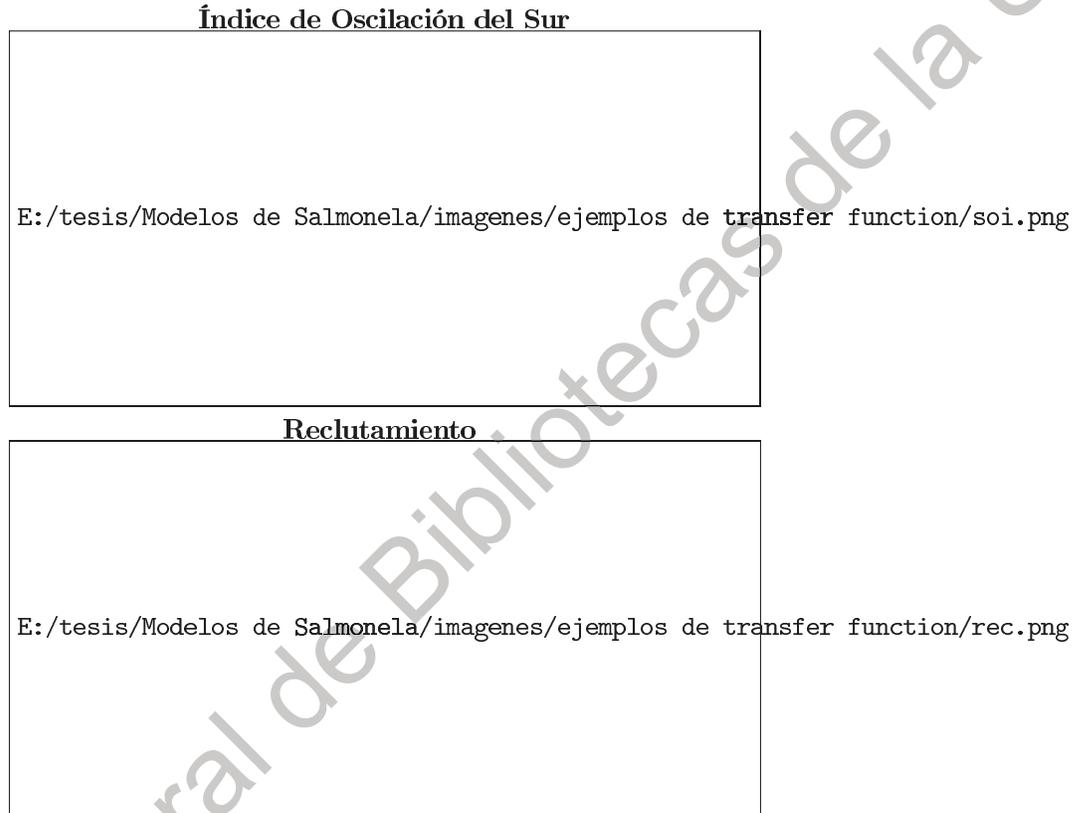


Figura 7: SOI y Reclutamiento mensuales (1950-1987)

Lo primero que se hace es tratar de ajustar un modelo ARMA, como en la ecuación (124), a la serie de entrada, SOI. Para lograr esto, se fijaron en la ACF y la PACF del SOI, los cuales se muestran en la Figura 8 y son calculados con la función de R

```
acf2(soi)
```

E:/tesis/Modelos de Salmonela/imagenes/ejemplos de transfer function/ACF y PACF SOI.png

Figura 8: ACF y PACF del SOI

En la Figura 8 se logra apreciar que el modelo AR(1) haría un buen trabajo para SOI ya que en la PACF después del primer rezago los valores se cortan, mientras que en el ACF los valores van cayendo lentamente.

Se continúa ajustando el ARMA(1,0) al SOI, obteniendo como valor autorregresivo $\hat{\phi} = 0,5875$ con una $\hat{\sigma}_w^2 = 0,0379$ y un p -valor menor a 10^{-16} , al igual que un intercepto de 13.7507 con un p -valor de 0.024 y una tendencia en el tiempo de -0.0069. Dichos valores fueron obtenidos con la función

```
(fit = arima(soi, xreg=time(soi), order=c(1, 0, 0)))
```

i.e.

$$\begin{aligned}\phi(B) \nabla x_t &= \theta(B) w_t \\ (1 - 0,588B)(SOI_t - SOI_{t-1}) &= w_t\end{aligned}$$

Donde la función arima devuelve un modelo ARIMA, recibiendo a “SOI” como una serie de tiempo univariada, «order» será un vector con 3 componentes enteros (p, d, q) , donde p es el orden del polinomio autorregresivo, d es el grado de diferenciación y q es el orden del polinomio con respecto a los promedios móviles. A su vez se agregó como entrada “xreg”, la cual calcula la primera diferencia de la serie antes de ajustarla, lo que es equivalente a $\nabla SOI = SOI_t - SOI_{t-1}$.

Después se aplicó el operador estimado AR(1) i.e. $(1 - 0,588B)$, como en la ecuación (129) a la serie Reclutamiento para luego calcular la correlación cruzada.

Para obtener los datos del operador aplicados a la serie de entrada SOI, simplemente se toman los residuales del ajuste con la siguiente función

```
soi.pw = resid(fit)
```

i.e. $\frac{\phi(B)}{\theta(B)} x_t = w_t$

Y para aplicarlo a la serie de salida se usa el siguiente código

```
ar1 = as.numeric(fit$coef[1])
rec.d = resid(lm(rec~time(rec), na.action=NULL))
rec.fil = filter(rec.d, filter=c(1, -ar1), method="conv", sides=1)
```

donde primero guardan el coeficiente autorregresivo, ϕ_1 (línea 1), luego remueven la tendencia de la serie de salida Recruitment, ∇y_t (línea 2), y por último aplican el operador a la serie de salida, $\frac{\phi(B)}{\theta(B)} y_t$ (línea 3).

Una vez hecho esto, se calcula la correlación cruzada entre ambas con la siguiente función

```
ccf(soi.pw, rec.fil, main="", ylab="CCF", na.action=na.omit)
```

Misma que calcula y grafica la correlación cruzada de dos series de tiempo, ubicadas como dos primeras entradas de la función. La gráfica se muestra en la Figura 9.

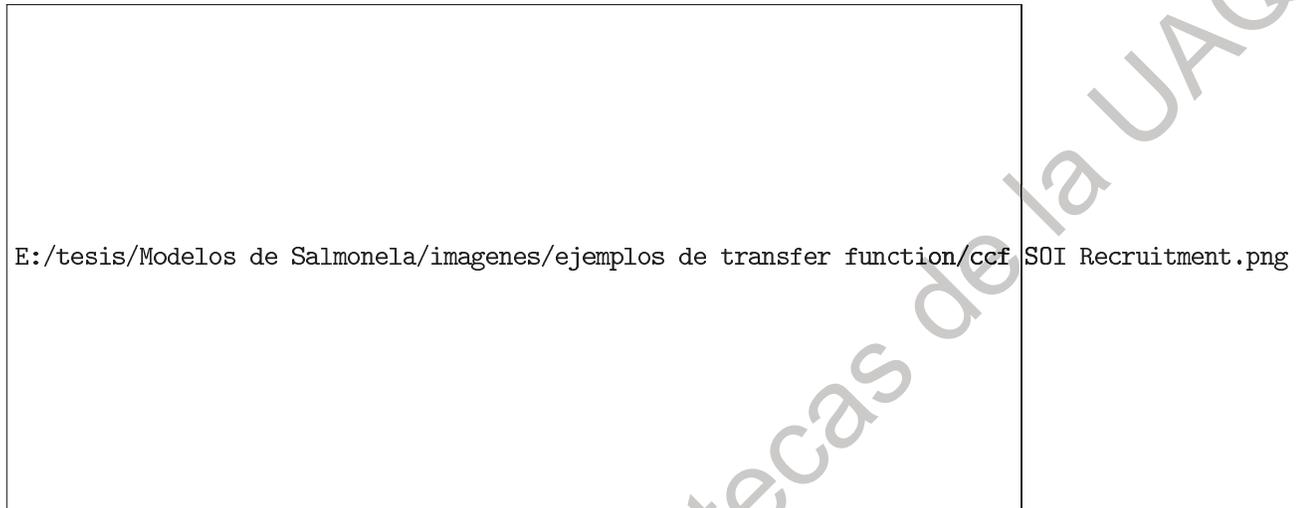


Figura 9: CCF entre el SOI preblanqueado y el Recruitment transformado

Viendo la CCF se aprecia que el 5° rezago es el más dominante, así $d = 5$, y que los rezagos siguientes a éste decaen, por lo que $s = 0$ y entonces $\delta(B) = \delta_0$. Se sugiere que $\omega(B) = 1 - \omega_1 B$ por la autocorrelación del Reclutamiento. Con esto se tiene el modelo para la ecuación (126) como

$$\begin{aligned}\alpha(B) &= \frac{\delta(B) B^d}{\omega(B)} \\ &= \frac{\delta_0 B^5}{1 - \omega_1 B}\end{aligned}$$

el cual se sustituye en el modelo (122)

$$\begin{aligned}y_t &= \alpha(B) x_t + \eta_t, \\ y_t &= \frac{\delta_0 B^5}{1 - \omega_1 B} x_t + \eta_t, \\ (1 - \omega_1 B) y_t &= \delta_0 B^5 x_t + u_t, \\ y_t &= \omega_1 B y_t + \delta_0 B^5 x_t + u_t, \\ y_t &= \omega_1 y_{t-1} + \delta_0 x_{t-5} + u_t\end{aligned}\tag{133}$$

donde

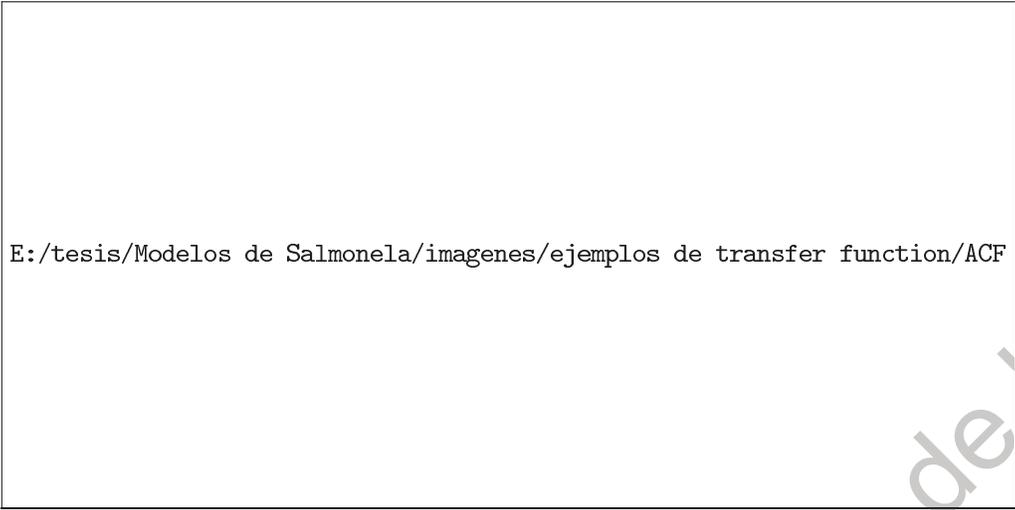
$$u_t = (1 - \omega_1 B) \eta_t$$

con (133) se realiza la regresión obteniendo los siguiente valores $\hat{\omega}_1 = 0,853$ y $\hat{\delta}_0 = -18,86$ con

$$\hat{u}_t = (1 - 0,853B) \eta_t$$

Para calcular dichos coeficientes se ejecutaron las siguientes funciones en R

```
rec.d = resid(lm(rec~time(rec), na.action=NULL))
soi.d = resid(lm(soi~time(soi), na.action=NULL))
fish = ts.intersect(rec.d, rec.d1=lag(rec.d, -1), soi.d5=lag(soi, -5))
summary(fish.fit <- lm(rec.d~0+rec.d1+soi.d5, data=fish))
```



E:/tesis/Modelos de Salmonela/imagenes/ejemplos de transfer function/ACF Y PACF eta.png

Figura 10: ACF y PACF para η_t

donde primero se remueve la tendencia de ambas series, entrada y salida. Después se crea una base con los datos rezagados para ser usados en la regresión, y por último calculan los coeficientes de la regresión ω_1 y δ_0 .

Por último, se aplica el operador de promedios móviles a \hat{u}_t para estimar el ruido original η_t , al cual se le ajusta un modelo AR(2), basados en el ACF y en el PACF que se muestra en la Figura 10.

De dicho ajuste se obtiene que

$$(1 + 0,427B - 0,328B^2) \eta_t = z_t.$$

Dichos datos se calculan con el siguiente código

```
om1 = as.numeric(fish.fit$coef[1])
eta.hat = filter(resid(fish.fit), filter=om1, method="recur", sides=1)
acf2(eta.hat)
(eta.fit <- arima(eta.hat, order=c(2,0,0)))
```

donde primeramente se guarda el valor del coeficiente $\hat{\omega}_1$ para después aplicarlo en los residuales de la regresión $u_t = (1 - \hat{\omega}_1 B) \eta_t$ y obtener el ruido original $\hat{\eta}_t$. Así se calculó el ACF y PACF y se ajustó finalmente el modelo para el ruido $\phi_w(B) \eta_t = \theta_w(B) z_t$ con $\phi_w(B) = \phi_{w1}B + \phi_{w2}B^2$ y a $\theta_w(B) = 1$.

La función “filter” aplica un filtrado línea a una serie de tiempo univariada, la cual es la primera entrada de la función; la segunda entrada, “filter” es un vector de coeficientes para filtrar la serie. La entrada “method” puede ser “convolution” para ser aplicados con promedios móviles, o “recursive” donde una autorregresión será aplicada.

Ahora, se verá una aplicación de un caso donde el modelo de función de transferencia no es conocido.

7. Modelaje de la relación del número de Casos de Tifoidea en México con la Temperatura Máxima

Se tiene la serie de salida (Y_t) como el número de casos de tifoidea por mes, y la serie de entrada (X_t) como las temperaturas máximas por mes, estas a nivel nacional en México desde el 2004 hasta el 2015, mismos que se muestran en la Figura 11.

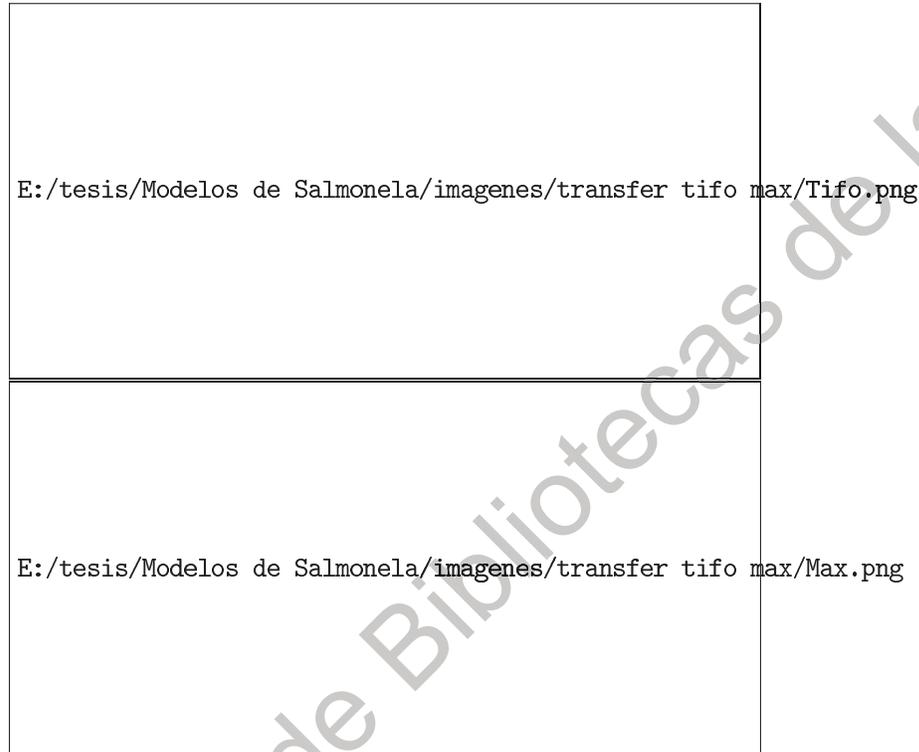


Figura 11: En la parte superior se muestra el número de casos de tifoidea y en la parte inferior se muestran las temperaturas máximas en México en el periodo 2004 - 2015

Para aplicar la metodología de la función de transferencia lo primero que se debe hacer es volver ambas series, de entrada y salida, en estacionarias, es decir, eliminar el comportamiento estacional anual. Lo cual se logra calculando la diferencia de primer orden con un rezago de 12 meses dado el comportamiento anual observado. Los datos, ya con la tendencia a 12 meses removida, i.e. $y_t - y_{t-12} = (1 - B^{12}) y_t$ y $x_t - x_{t-12} = (1 - B^{12}) x_t$, se muestran en la Figura 12.

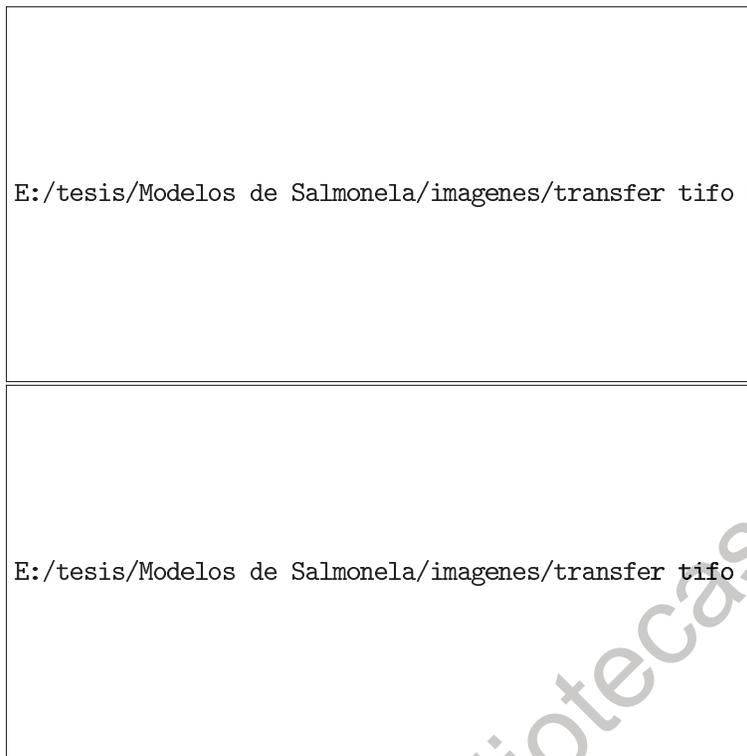


Figura 12: (Superior) Número de casos de tifoidea (y_t); y (inferior) temperatura máxima (x_t), ambas diferenciadas a 12 meses

Una vez que las series ya son estacionarias, se prosigue a ajustar un modelo SARIMA a la serie de entrada. Las temperaturas máximas, como en la ecuación (124). Para ello se observa el ACF y PACF de la serie, los cuales se muestran en la Figura 13. Se aprecia que el modelo SARIMA $(1, 0, 0) \times (1, 1, 0)_{12}$ haría un estupendo trabajo, ya que en los valores de autocorrelación parcial (PACF), en los rezagos 1 y 12 los valores se cortan, y en dichos rezagos las autocorrelaciones (ACF) van cayendo lentamente.

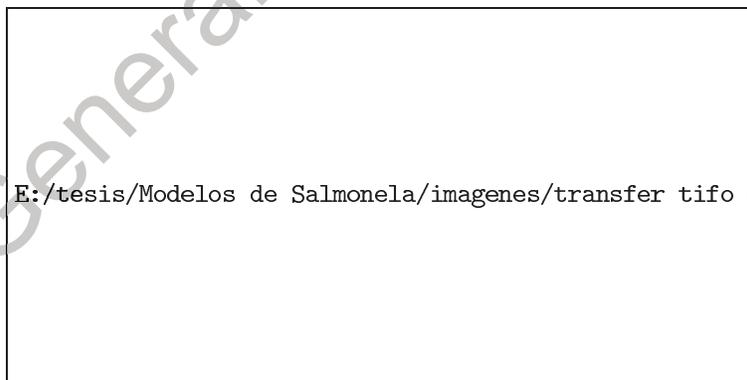


Figura 13: ACF y PACF de las temperaturas máximas

Después se ajustó el modelo SARIMA $(1, 0, 0) \times (1, 1, 0)_{12}$ a los datos de la temperatura máxima, obteniendo como valor autorregresivo $\phi_1 = 0,5064$ y como valor autorregresivo estacional $\Phi_1 = -0,5130$ con una $\hat{\sigma}_w^2 = 0,7466$. Luego se aplica el operador SARMA $(1, 0, 0) \times (1, 1, 0)_{12}$, i.e. $(1 - 0,5064B) (1 + 0,513B^{12}) (1 - B^{12})$, a ambas series como en la ecuación (129), y se calcula la correlación cruzada entre ellas la cual se muestra en la Figura 14.

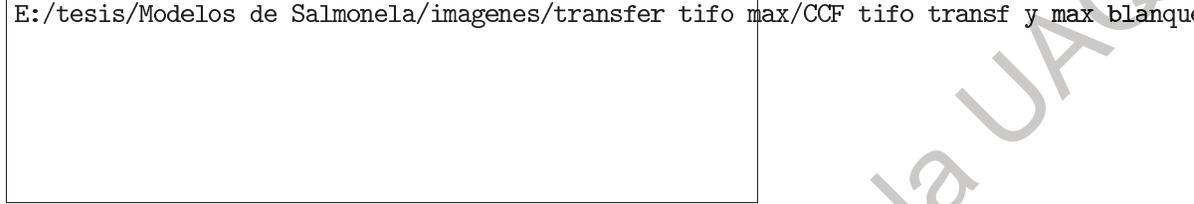


Figura 14: CCF entre la temperatura Máxima preblanqueada y los casos de tifoidea transformados

En dicha Figura 14 se observa que el 10° rezago es el más dominante, por lo que $d = 10$, que los siguientes rezagos a este, no son significativos, así $s = 0$ y $\delta(B) = \delta_0$.

Observando los ACF y PACF de la serie de Tifoidea transformada, es decir, desestacionalizada, mostrada en la Figura 15, se aprecia que el rezago 1 es más grande, con lo que se toma a $r = 1$ y $\omega(B) = 1 - \omega_1 B$.

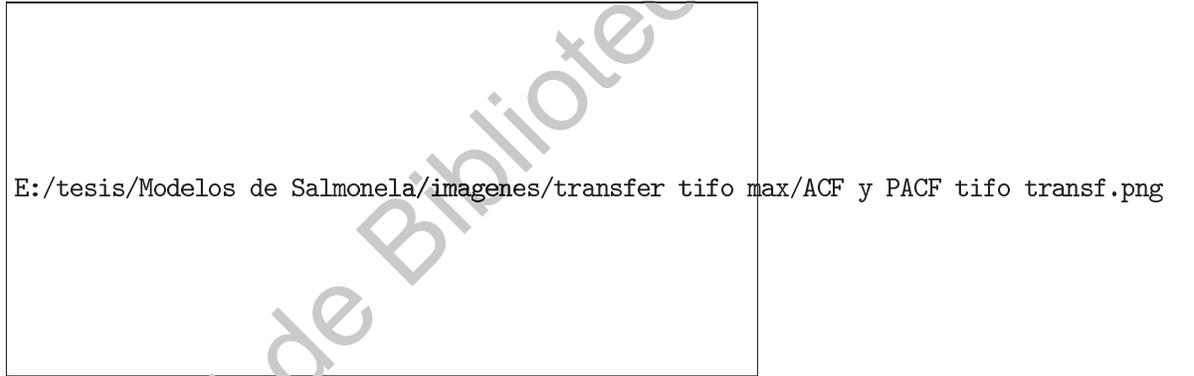


Figura 15: ACF y PACF de la serie de Tifoidea transformada

Con todo esto, se tiene el modelo para la ecuación (126) como

$$\alpha(B) = \frac{\delta(B) B^d}{\omega(B)} = \frac{\delta_0 B^{10}}{1 - \omega_1 B},$$

el cual se sustituye en el modelo (122)

$$\begin{aligned} y_t &= \alpha(B) x_t + \eta_t, \\ y_t &= \frac{\delta_0 B^{10}}{1 - \omega_1 B} x_t + \eta_t, \\ (1 - \omega_1 B) y_t &= \delta_0 B^{10} x_t + (1 - \omega_1 B) \eta_t, \\ y_t &= \omega_1 B y_t + \delta_0 B^{10} x_t + u_t, \\ y_t &= \omega_1 y_{t-1} + \delta_0 x_{t-10} + u_t, \end{aligned} \tag{134}$$

donde

$$u_t = (1 - \omega_1 B) \eta_t.$$

Con la ecuación (134) se corre la regresión obteniendo los siguientes valores $\hat{\omega}_1 = -0,15707$, con un p -valor de 0.1109, y $\hat{\delta}_0 = 124,67104$, con un p -valor de 0.0141.

Debido a que el parámetro ω_1 es no significativo, se tomará $\omega_1 = 0$, y se vuelve a realizar la siguiente regresión

$$y_t = \delta_0 x_{t-10} + \eta_t, \tag{135}$$

donde $\hat{\delta}_0 = 100,98$, con un p -valor de 0.0381, por lo que la temperatura máxima (x_t) sigue siendo significativa para el modelo.

Por último al ruido $\hat{\eta}_t$ se le intentará ajustar un modelo SARMA, y para ello se observan los ACF y PACF mostrados en la Figura 16

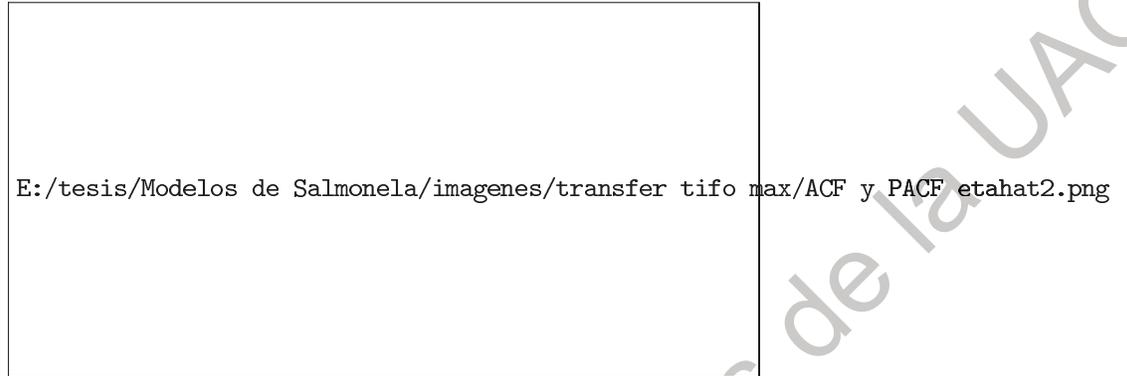


Figura 16: ACF y PACF de $\hat{\eta}_t$

Se ajustaron varios modelos para η_t , mismos que se muestran en la Tabla 4. Cada modelo podía ser considerado como bueno en términos del ACF de los residuales, los gráficos q-q normales y los p -valores Ljung-Box. Debido a ello, se escogió el modelo con un mejor AIC, AICc y BIC, que en este caso fue el SARMA $(3, 0, 4) \times (0, 0, 1)_{12}$.

Modelo SARMA	AIC	AICc	BIC	# parámetros	significativos
$(3, 0, 3) \times (1, 0, 1)_{12}$	13.3923	13.43075	12.6131	9	4
$(3, 0, 5) \times (1, 0, 1)_{12}$	13.3291	13.37653	12.59916	11	4
$(3, 0, 5) \times (0, 0, 1)_{12}$	13.2775	13.32022	12.52305	10	6
$(3, 0, 3) \times (0, 0, 1)_{12}$	13.4022	13.43678	12.59864	8	4
$(3, 0, 4) \times (0, 0, 1)_{12}$	13.2615	13.2999	12.48246	9	6

Cuadro 4: Modelos ajustados para $\hat{\eta}_t$, con los AIC, AICc, BIC, el número de parámetros que tiene y cuantos de ellos eran significativos.

Así

$$\begin{aligned}
 \phi_{\eta}(B)\eta_t &= \alpha + \Theta_{\eta}(B^{12})\theta_{\eta}(B)z_t, \\
 (1 - \phi_{\eta_1}B - \phi_{\eta_2}B^2 - \phi_{\eta_3}B^3)\eta_t &= \alpha + (1 + \Theta_{\eta_1}B^{12})(1 + \theta_{\eta_1}B + \theta_{\eta_2}B^2 + \theta_{\eta_3}B^3 + \theta_{\eta_4}B^4)z_t, \\
 \eta_t &= \phi_{\eta_1}\eta_{t-1} + \phi_{\eta_2}\eta_{t-2} + \phi_{\eta_3}\eta_{t-3} + \alpha \\
 &\quad + z_t + \theta_{\eta_1}z_{t-1} + \theta_{\eta_2}z_{t-2} + \theta_{\eta_3}z_{t-3} + \theta_{\eta_4}z_{t-4} \\
 &\quad + \Theta_{\eta_1}z_{t-12} + \Theta_{\eta_1}\theta_{\eta_1}z_{t-13} + \Theta_{\eta_1}\theta_{\eta_2}z_{t-14} + \Theta_{\eta_1}\theta_{\eta_3}z_{t-15} + \Theta_{\eta_1}\theta_{\eta_4}z_{t-16}.
 \end{aligned}$$

Ajustamos el modelo, donde la estimación de los parámetros se puede observar en la Tabla 5

Parámetro	Est	s.e.	p -valor
AR1 ϕ_{η_1}	-0.198	0.056	0.0004537
AR2 ϕ_{η_2}	0.053	0.063	0.4055368
AR3 ϕ_{η_3}	0.929	0.056	$< 2.2 \times 10^{-16}$
MA1 θ_{η_1}	-0.056	0.105	0.5937877
MA2 θ_{η_2}	0.017	0.079	0.8282098
MA3 θ_{η_3}	-0.731	0.075	$< 2.2 \times 10^{-16}$
MA4 θ_{η_4}	0.415	0.093	9.567×10^{-6}
SMA1 Θ_{η_1}	-0.805	0.127	2.435×10^{-10}
intercepto α	182.64	35.354	2.392×10^{-7}

Cuadro 5: Estimación de los parámetros del modelo SARMA $(3, 0, 4) \times (0, 0, 1)_{12}$ para $\hat{\eta}_t$, con sus errores estándar y sus p -valores

Se prosiguió a diagnosticar el mejor modelo ajustado, para lo que se usaron los residuales estandarizados, el factor de autocorrelación de los residuales, el gráfico normal Q-Q y los p-valores del estadístico Ljung-Box los cuales se muestran en la Figura 17



Figura 17: Diagnósticos del modelo SARMA(3, 0, 4) × (0, 0, 1)₁₂ para η_t

Observando los residuales se puede ver que se encuentran alrededor del cero. Los factores de autocorrelación de los residuales se encuentran dentro de la banda, por lo que se puede suponer que los residuales no están autocorrelacionados. El gráfico normal Q-Q nos indica que los residuales son normales. Y al ver los p-valores de Ljung-Box de nuestro ajuste, se puede apreciar que casi todos ellos son altos a un nivel de confianza de 0.95, por lo que es razonable afirmar que los residuales no están autocorrelacionados entre sí.

Así, el modelo completo será

$$\begin{aligned}
 y_t = & y_{t-12} + \delta_0 x_{t-10} - \delta_0 x_{t-22} + \phi_{\eta 1} \eta_{t-1} + \phi_{\eta 2} \eta_{t-2} + \phi_{\eta 3} \eta_{t-3} + \alpha \\
 & + z_t + \theta_{\eta 1} z_{t-1} + \theta_{\eta 2} z_{t-2} + \theta_{\eta 3} z_{t-3} + \theta_{\eta 4} z_{t-4} \\
 & + \Theta_{\eta 1} z_{t-12} + \Theta_{\eta 1} \theta_{\eta 1} z_{t-13} + \Theta_{\eta 1} \theta_{\eta 2} z_{t-14} + \Theta_{\eta 1} \theta_{\eta 3} z_{t-15} + \Theta_{\eta 1} \theta_{\eta 4} z_{t-16},
 \end{aligned} \tag{136}$$

donde y_t son los casos de tifoidea, y x_t son las temperaturas máximas.

8. Conclusiones

Se estacionarizó las series de salida (casos de tifoidea por mes) y entrada (temperaturas máximas por mes) para después ajustar el modelo SARIMA $(1, 0, 0) \times (1, 1, 0)_{12}$ a la serie de entrada. Dicho modelo se aplica a ambas series para después calcular la correlación y así calcular el orden de regresión d de (126), siendo en este caso 10, y los rezagos significativos siguientes.

Luego de observar los ACF y PACF de la serie de salida y determinar los rezagos significativos, se prosiguió a sustituir en el modelo (122), con el cual se ejecutó la regresión y se identificó que no todos los valores eran significativos, para lo que se estimaron nuevamente los valores de la regresión haciendo caso omiso a los factores no significativos. Por último se ajustó y diagnosticó el modelo SARMA $(3, 0, 4) \times (0, 0, 1)_{12}$ para el ruido $\hat{\eta}_t$.

Con el análisis anterior se muestra que y_t , el número de casos de tifoidea, depende de sí misma 12 meses atrás, así como de 100 veces la temperatura hace 10 meses, 100 veces la temperatura hace 22 meses atrás y una función compleja de ruido.

Sin embargo no se logró graficar la predicción con sus intervalos de confianza debido a la falta de conocimientos de algún paquete de programación en R, el cual pudiese con un ruido tan complejo.

Se podría continuar este estudio modelándolo con la temperatura mínima en vez de la máxima, la cual es más complicada de modelar con un SARIMA ya que las temperaturas mínimas tardaban más en bajar cuando estaban en sus puntos más altos, lo cual no se observa en las temperaturas máximas.

9. Bibliografía

- Akaike, H. (1969). Fitting autoregressive models for prediction. *Ann. Inst. Stat. Math.*, 21, 243-247.
- Akaike, H. (1973). Information theory and an extension of the maximum likelihood principle. In *2nd Int. Symp. Inform. Theory*, 267-281. B.N. Petrov and F. Csaki, eds. Budapest: Akademia Kiado.
- Akaike, H. (1974). A new look at statistical model identification. *IEEE Trans. Automat. Contr.*, AC-19, 716-723.
- Box, G.E.P. y G.M. Jenkins (1970). *Time Series Analysis, Forecasting, and Control*. Oakland, CA:Holden-Day.
- Checkley W., Epstein L.D., Gilman R.H., Cabrera L. y R.E. Black. (2003). Effects of acute diarrhea on linear growth in Peruvian children. *Am J Epidemiol*, 157(2), 166-75.
- Checkley W., Epstein L.D., Gilman R.H., Figueroa D., Cama R.I., Patz J.A. y R.E. Black (2000). Effects of EI Niño and ambient temperature on hospital admissions for diarrhoeal disease in Peruvian children. *Lancet*, 355, 422-50.
- Chengsheng J., Kristi S.S., Upperman C.R., Blythe D., Clifford M., Murtugudde R., Sapkota A.R. y A. Sapkota (2015) Climate change, extreme events and increased risk of salmonellosis in Maryland, USA: Evidence for coastal vulnerability. *Environment International*, 83, 58-62.
- de Magny G.C., Murtugudde R., Sapiano M.R.P., Nizam A., Brown C.W., Busalacchi A.J., Yunus M., Nair G.B., Gil A.I., Lanata C.F., Calkins J., Manna B., Rajendran K., Bhattacharya M.K., Huq A., Sack R.B. y R.R. Colwell (2008). Environmental signatures associated with cholera. *Proceedings of the National Academy of Sciences of the United States of America*, 105, 17676-17681.
- Hashizume M., Armstrong B., Hajat S., Wagatsuma Y., Faruque A.S., Hayashi T. y D.A. Sack (2007) Associated between climate variability and hospital visit for non-cholera diarrhoea in Bangladesh: effects and vulnerable groups. *Int J Epidemiol*, 36, 1030-7.
- Hurvich, C.M y C.-L. Tsai (1989). Regression and time series model selection in small samples. *Biometrika*, 76, 297-307.
- Kale P.L., Hinde J.P. y F.F. Nobre (2004). Modeling diarrhea disease in children less than 5 years old. *Ann Epidemiol*, 14, 371-7.
- McQuarrie, A.D.R. y C-L. Tsai (1998). *Regression and Time Series Model Selection*, Singapore:WorldScientific.
- Onozuka D. (2014). Effect of non-stationary climate on infectious gastroenteritis transmission in Japan. *Scientific Reports*, 4, 5157.
- Onozuka D. y A. Hagihara (2015). Nationwide variation in the effects of temperature on infectious gastroenteritis incidence in Japan. *Scientific Reports*, 5, 12932.
- Pascual M., Rodó X., Ellner S.P., Colwell R. y M.J. Bouma (2000). Cholera dynamics and EI Niñosouthern oscillation. *Science*, 289, 1766-9.
- Rodó X., Pascual M., Fuchs G. y A.S.G. Faruque (2002). ENSO and cholera: a nonstationary link related to climate change? *Proc Natl Acad Sci U S A*, 99, 129016.
- Semenza J.C., Herbst S., Rechenburg A., Suk J.E., Hoser C. y C. Schreiber (2012). Climate change impact assessment of food- and waterborne diseases. *Crit. Rev. Environ. Sci. Technol.* 42, 857-890.
- Shumway, R.H. (1988). *Applied Statistical Time Series Analysis*. Englewood Cliffs, NJ: Prentice-Hall.
- Shumway, R.H. y D.S. Stoffer (2015). *Time Series Analysis and Its Applications*. Blue Printing, 3rd edition.
- Sugiura, N. (1978). Further analysis of the data by Akaike's information criterion and the finite corrections, *Commun. Statist, A, Theory Methods*, 7, 13-26.
- Wei-Chun C., Jiunn-Lin W., Yu-Chun W., Hsin H. Fung-Chang S. y C. Chun-Yu. (2010). Modeling the impact of climate variability on diarrhea-associated diseases in Taiwan. *Science of The Total Environment*, 409(1), 43-51.

- Who. (2008). Water-related diseases. *Water sanitation and health (WSH)*, Vol 2008. World Health Organization.

Dirección General de Bibliotecas de la UAQ

10. Programa

```
library(astsa)
library(forecast)
library(lmtest)
Max #Xt
Max.d12 = diff(Max,12) #xt (Xt blanqueada)
acf2(Max.d12) #ACF y PACF de xt
(fit = arima(Max.d12, order=c(1, 0, 0),seasonal = list(order = c(1, 0, 0),
period = 12),include.mean = FALSE)) # ajuste sarima
coefstest(fit)#significancia de los estimadores
Max.d12.pw = resid(fit) #wt
Tifo #Yt
Tifo.d12 = diff(Tifo,12)#yt
recfit <- Arima(Tifo.d12, model=fit)
rec.fil <- recfit$residuals
#aplicar el SARMA a yt = ~yt(yt gorro)
ccf(Max.d12.pw[-(1:13)], rec.fil[-(1:13)], main="", ylab="CCF")
#CCF entre ~yt y wt para sacar los ordenes d y s
acf2(rec.fil[-(1:13)]) #ACF de ~yt para sacar el orden r
fish = ts.intersect(Tifo.d12, Tifo.d12_1=lag(Tifo.d12,-1), Max.d12_10=lag(Max.d12,-10))
#crea una base de datos del mismo tamaño
summary(fish.fit <- lm(Tifo.d12~0+ Tifo.d12_1+ Max.d12_10, data=fish))
#crea el ajuste de regresión para los coeficientes del polinomio
summary(fish.fit <- lm(Tifo.d12~0+ Max.d12_10, data=fish))
#nuevo ajuste si Tifo.d12_1
eta.hat =resid(fish.fit) #calculo de los etas
acf2(eta.hat) #ACF y PACF de los etas para luego modelar con un SARMA
#Ajuste y evaluación de varios modelos para las etas——
sarima(eta.hat,3,0,3,1,0,1,12)
(fit.eta1 = arima(eta.hat, order=c(3, 0, 3),include.mean = TRUE,seasonal =
list(order = c(1, 0, 1), period = 12)))
coefstest(fit.eta1)
sarima(eta.hat,3,0,5,1,0,1,12)
(fit.eta2 = arima(eta.hat, order=c(3, 0, 5),include.mean = TRUE,seasonal =
list(order = c(1, 0, 1), period = 12)))
coefstest(fit.eta2)
sarima(eta.hat,3,0,5,0,0,1,12)
(fit.eta3 = arima(eta.hat, order=c(3, 0, 5),include.mean = TRUE,seasonal =
list(order = c(0, 0, 1), period = 12)))
coefstest(fit.eta3)
sarima(eta.hat,3,0,3,0,0,1,12)
(fit.eta4 = arima(eta.hat, order=c(3, 0, 3),include.mean = TRUE,seasonal =
list(order = c(0, 0, 1), period = 12)))
coefstest(fit.eta4)
sarima(eta.hat,3,0,4,0,0,1,12)
(fit.eta3.5 = arima(eta.hat, order=c(3, 0, 4),include.mean = TRUE,seasonal =
list(order = c(0, 0, 1), period = 12)))
coefstest(fit.eta3.5)
#——
```