Una aplicación de las series de tiempo visto desde los espacios de Hilbert.

Erika Valeria Rivera Jimenez
Director: Luis Fernando Villarraga Poveda

Universidad Distrital Francisco José de Caldas
Bogotá D.C.
2017

Dedicado a mis padres, hermano y novio.

Agradecimientos

Agradezco a mis padres y hermano por su apoyo y aliento en cada tropiezo o alegría vivida durante este proceso, a los profesores del proyecto curricular de matemáticas por los consejos y conocimientos transmitidos y a Pipesaubrio por su compresión, fortaleza, amor y paciencia.

Índice general

Introducción

Las series de tiempo surgen el siglo pasado gracias a los trabajos pioneros del matemático A.N. Kolgomorov en 1931 (Anales matemáticos de la academia de ciencias alemana) y del estadístico británico George Yule (1871-1951) y su obra (On the time-correlation Problem, (1921)); donde promueven la investigación y las primeras aplicaciones basadas en modelos autorregresivos de segundo orden, en base a los trabajos de Yule, el matemático y economista ruso Eugen Slutsky (1880-1948), integra una metodología, desde un punto de vista estocástico, fundamentando las series de tiempo y el análisis económico de las mismas, desde unas bases estadísticas.

En el presente trabajo se elegirá un modelo para estimar parametros y utilizar el modelo para mejorar y comprender el funcionamiento que genera la serie, visto desde los espacios de Hilbert $L^2(\Omega, \mathscr{F}, P)$. Inicialmente se mostrara que el espacio de todas las variables aleatorias X definidas en Ω son un espacio vectorial y con un producto interno definido será un espacio de Hilbert, soportado desde la teoría de la medida y el análisis funcional, se introducirá la relación entre espacios estocásticos y las series de tiempo encontrando que el modelo matemático para una serie de tiempo es el concepto de proceso estocástico, se enunciarán y demostrarán las propiedades del modelo autorregresivo. Finalmente, se aplicara la teoría ya mencionada a un ejemplo particular del modelo autorregresivo.

CAPÍTULO 1

Preliminares

1.1. Motivación

Las series de tiempo han tenido un rol muy importante en el análisis y predicción de eventos, son una secuencia de datos, valores u observaciones, que se miden en diferentes momentos cronológicamente ordenados donde los datos pueden estar en intervalos de tiempo iguales o desiguales.

Existen modelos matemáticos de las series de tiempo que permiten la predicción, comportamiento y comprensión de ciertos eventos que pueden utilizarse de diversas formas dependiendo del campo de aplicación particular como la economía, el marketing, la demografía y el medio ambiente, entre otros.
Las aplicaciones incluyen la separación (filtración) del ruido de las señales, la predicción de valores futuros de una serie y el control de valores futuros permitiendo un claro entendimiento de los resultados debido a su fácil modelización.

1.2. Espacio de Hilbert $L^2(\Omega, \mathscr{F}, P)$

El análisis de las series de tiempo es posible sin el estudio previo de los espacios de Hilbert. Pero, dada la relación que existe entre la geometría euclidiana y en particular con los conceptos de ortogonalidad y proyección ortogonal en espacios de dos y tres dimensiones. Daremos a conocer con el presenté trabajo el papel tan importante que tiene los espacios de Hilbert de dimensión infinita en el estudio de variables aleatorias con segundo momento finito y en especial en la teoría de predicción de procesos estacionarios.

1.2.1. Espacios de probabilidad

La teoría de la probabilidad es la parte de las matemáticas que se encarga del estudio de los fenómenos o experimentos aleatorios, se entiende por experimento aleatorio aquel que cuando se repite bajo las mismas condiciones iniciales, el resultado que se obtiene no siempre es el mismo,el espacio a trabajar consiste de una terna ordenada, denotada usualmente por $(\Omega, \mathscr{F}, P)$, en donde Ω es un espacio muestral, $\mathscr{F}$ es una σ-álgebra de subconjuntos de Ω, y P es una medida de probabilidad definida sobre $\mathscr{F}$ [Brezis, 2010, pág 89].A la pareja $(\Omega, \mathscr{F})$ se le llama espacio medible y a los elementos de $\mathscr{F}$ se les llama eventos o conjuntos medibles.
Como caso especial de σ-álgebra se tiene la σ-álgebra de Borel que esta generada por la colección de conjuntos abiertos de un espacio topológico. Los miembros de esta σ-álgebra son llamados conjuntos de Borel, para los espacios medibles $(\Omega, \mathscr{F})$ tendremos que si se le asocia una medida de probabilidad P, una función h de Ω sobre Y con Ω un espacio medible y Y un espacio topológico será llamada una *variable aleatoria* sí $h^{-1}(B)$ es un conjunto medible en Ω para cada conjunto abierto B en Y. La función h se dice medible de Ω a Y y se considera primeramente variables aleatorias correspondientes a $(Y, \mathscr{G}) = (\mathbb{R}^d, \mathscr{B}^d)$ y denotada con letras mayúsculas, por ejemplo $X : \Omega \to \mathbb{R}^d$ para $d > 1$ y $X : \Omega \to \mathbb{R}$ para $d = 1$. Para variables aleatorias con valores $d > 1$ se les llamará vectores aleatorios.[Grigoriu, 2002, pág 8].

Algunas propiedades de las variables aleatorias que son de importancia se mencionarán a continuación, una de ellas afirma que si, h es medible y P es una medida de probabilidad sobre Ω entonces una función $\mathscr{Q}$ de $\mathscr{G}$(σ-álgebra asignada al espacio topológico Y) a $[0,1]$ sera llamada la probabilidad inducida por h o la **distribución de** h sí $\mathscr{Q}$ definido como $P(h^{-1}(B))$ es una medida de probabilidad sobre Y para un conjunto abierto en $\mathscr{G}$. Probar que efectivamente $\mathscr{Q}$ es una medida, es inmediato ya que:

- $\mathscr{Q}(Y) = 1$

$$\begin{aligned}\mathscr{Q}(Y) =& P(h^{-1}(Y))\\ =& P(\Omega)\\ =& 1.\end{aligned}$$

- Aditividad contable.
  Para cualquier conjunto contable I y disjunto de eventos $B_i \in \mathscr{G}$, se tiene,

$$\begin{aligned}\mathscr{Q}\left(\bigcup_{i\in I} B_i\right) =& P\left(h^{-1}\left(\bigcup_{i\in I} B_i\right)\right)\\ =& \sum_{i=1}^{\infty} \mathscr{Q}(B_i).\end{aligned}$$

  Dado que $h^{-1}(B_i)$ son eventos disjuntos en $\mathscr{F}$.

Otra propiedad, dice que la σ-álgebra generada por una variable aleatoria X es,

$$\sigma(X) = X^{-1}(\mathscr{B}^d) = \left\{X^{-1}(B) \in \mathscr{F} : \forall B \in \mathscr{B}^d\right\}.$$

Y representa la más pequeña σ-álgebra con respecto a la cual X es medible [Grigoriu, 2002, pág 6]. La colección de subconjuntos $\sigma(X)$ es una σ-álgebra ya que:

- $X^{-1}(\mathbb{R}^d) = \Omega$ y dado que $\Omega \in \mathscr{F}$ entonces $X^{-1}(\mathbb{R}^d) \in \mathscr{F}$, obteniendo asi, $X^{-1}(\mathbb{R}^d) \in \sigma(X)$.

- Si $B \in \mathscr{B}^d$,

$$X^{-1}(\mathbb{R}^d - B) = X^{-1}(\mathbb{R}^d) - X^{-1}(B) = \Omega - X^{-1}(B).$$

Se tiene que $\Omega \in \mathscr{F}$ y $X^{-1}(B) \in \mathscr{F}$, por tanto, $X^{-1}(\mathbb{R}^d - B) \in \mathscr{F}$, obteniendo asi, $X^{-1}(\mathbb{R}^d - B) \in \sigma(X)$.

- Sea $B_i \in \mathscr{B}^d$ con $i = 1, 2, ...$, entonces,

$$X^{-1}(B_1 \cup B_2 \cup ...) = X^{-1}(B_1) \cup X^{-1}(B_2) \cup ...$$

Se tiene que $X^{-1}(B_i) \in \mathscr{F}$, por tanto, $\bigcup_{i=1}^{\infty} X^{-1}(B_i) \in \sigma(X)$.

Además para estudiar el espacio L^2 es necesario hacer el estudio de las *variables aleatorias simples* que se definen como una función de valores reales definida en un espacio de medida cuyo rango es finito, estas funciones tienen un papel muy importante en la teoría de la medida y sobre todo en los espacios de probabilidad y en el estudio general de los espacios L^p [Rudin, 1981, pág 61], un ejemplo de variable aleatoria simple es la función característica, definida como

$$\chi_{A_i} = \begin{cases} 1 & si \quad w \in A \\ \\ 0 & si \quad w \notin A \end{cases}$$

ya que si $\chi_{A_i}^{-1}(B) = \{w : \chi_{A_i}(w) \in B\}$ entonces

- Si $0, 1 \notin B$ entonces $\chi_{A_i}^{-1}(B) = \emptyset$.
- Si $1 \in B$ y $0 \notin B$ entonces $\chi_{A_i}^{-1}(B) = A$.
- Si $1 \notin B$ y $0 \in B$ entonces $\chi_{A_i}^{-1}(B) = A^c$.
- Si $0, 1 \in B$ entonces $\chi_{A_i}^{-1}(B) = \Omega$.

Por tanto $\chi_{A_i}^{-1}(B) \in \mathscr{F}$, lo que implica que la función característica es una variable aleatoria simple.
Antes de entrar en detalle con el espacio $L^2(\Omega, \mathscr{F}, P)$ se deben tener en cuenta algunas propiedades de las variables aleatorias simples que se utilizaran más adelante, tales

como: sí X es una variable aleatoria simple tal que la norma de x_i es finita con $i \in I$, su esperanza o valor esperado se define como

$$E[X] = \sum_{i \in I} x_i P(A_i).$$

y si la esperanza de una variable aleatoria X existe, esta es también denotada por

$$E[X] = \int_{\Omega} X \ dP.$$

1.2.2. Completitud del espacio $L^2(\Omega, \mathscr{F}, P)$

Sea $(\Omega, \mathscr{F}, P)$ un espacio de probabilidad y denotamos por $L^q(\Omega, \mathscr{F}, P)$ la colección de variables aleatorias X a valor real definidas sobre $(\Omega, \mathscr{F}, P)$ tal que $E[|X|^q] < \infty$ para $q \geq 1$. Si X está en $L^q(\Omega, \mathscr{F}, P)$ se escribíra $X \sim L^q(\Omega, \mathscr{F}, P)$. Con $q = 2$ es el caso relevante para las aplicaciones dado que está relaciona el calculo del segundo momento y la teoría de la estimación.
Consideramos el espacio $L^2(\Omega, \mathscr{F}, P)$ donde dos variables aleatorias X y Y son estrictamente iguales si para cada ω se cumple $X(\omega) = Y(\omega)$. Existen, sin embargo, otras formas más debíles de igualdad donde cambiaremos las variables aleatorias individuales $X \in L^2$ a clases de equivalencia de variables aleatorias que coinciden en casi toda parte definidas como,

$$[X] = \left\{ Y \in L^2 : P(Y = X) = 1 \right\}.$$

De tal forma, que se pueden cumplir satisfactoriamente las propiedades para que L^2 sea un espacio de Hilbert principalmente, se debe cumplir que la relación $\sim$ definida por $X \sim Y \iff X \overset{c.s}{=} Y$ es una relación de equivalencia y esto se tiene facilmente ya que, si se cumplen las relaciones

- Reflexiva.
  $X \sim X \iff P(\{\omega : X(\omega) \neq X(\omega)\}) = P(\oslash) = 0.$
- Simetrica.
  $X \sim Y \iff P(\{\omega : X(\omega) = Y(\omega)\}) = 1 = P(\{\omega : Y(\omega) = X(\omega)\})$.
  Por tanto $Y \sim X$.

- Transitiva.
  Si $X \sim Y$ y $Y \sim Z$ entonces $X \sim Z$.
  Sea, $P(\underbrace{\{\omega : X(\omega) = Y(\omega)\}}_{A}) = 1$ Y $P(\underbrace{\{\omega : Y(\omega) = Z(\omega)\}}_{B}) = 1$, es decír, $P(A) = 1$ y $P(B) = 1$.
  Luego,
  $$1 = P(A \cap B) = P(\{\omega : X(\omega) = Z(\omega)\}) = P(X = Z).$$
  Puesto que la intersección de dos conjuntos de probabilidad uno tiene probabilidad uno, luego $X \sim Z$.

Dicho lo anterior se probara que efectivamente sí $(\Omega, \mathscr{F}, P)$ es un espacio de probabilidad y se denotá por $L^2(\Omega, \mathscr{F}, P)$ el espacio de las clases de todas las variables aleatorias X definidas en Ω y que satisfacen la condición

$$E(X^2) = \int_{\Omega} X^2(\omega) \ dP < \infty.$$

Entonces con la notación usual de multiplicación por un escalar $\alpha \in \mathbb{R}$ y la suma de variables aleatorias, L^2 es un espacio vectorial. Además L^2 es un espacio de Hilbert con el producto interno $\langle X, Y \rangle = E[X, Y]$ para $X, Y \in L^2$ y la norma $\|X\|_{L_2} = \left(E(X^2)\right)^{1/2}$.
L^2 sera un espacio vectorial, sí probamos, inicialmente que $\alpha X + \beta Y$ con $X, Y \in L^2$ y $\alpha, \beta \in \mathbb{R}$ pertenecen a L^2 basta ver que,

$$\begin{aligned} \|(X+Y)\|^2 + \|(X-Y)\|^2 &= E((X+Y)^2) + E((X-Y)^2) \\ &= E(X^2) + 2E(XY) + E(Y^2) + E(X^2) - 2E(XY) + E(Y^2) \\ &= 2(E(X^2) + E(Y^2)). \end{aligned}$$

Luego
$$E((X+Y)^2) \leq 2(E(X^2) + E(Y^2)) < \infty$$
Y por tanto $X + Y \in L^2$

Además,
$$E((\alpha X)^2) = \alpha^2 E(X^2) < \infty \quad \alpha \in \mathbb{R}, X \in L^2,$$

obteniendo así que L^2 es un espacio vectorial.
Luego para que L^2 sea un espacio de Hilbert, se probara que $\langle X, Y\rangle = E[X, Y]$ cumple las propiedades de producto interno.
Por tanto,

1. $\langle X, Y\rangle = \overline{\langle Y, X\rangle}$.
$$\langle X, Y\rangle = E(XY) = E(YX) = \langle Y, X\rangle$$

2. $\langle X+Y, Z\rangle = \langle X, Z\rangle + \langle Y, Z\rangle$.
$$\langle X+Y, Z\rangle = E((X+Y)Z) = E(XZ+YZ) = E(XZ) + E(YZ) = \langle X, Z\rangle + \langle Y, Z\rangle$$

3. $\langle \alpha X, Y\rangle = \alpha \langle X, Y\rangle$.
$$\langle \alpha X, Y\rangle = E(\alpha XY) = \alpha E(XY) = \alpha \langle X, Y\rangle$$

4. $\langle X, X\rangle \geq 0$.
$$\langle X, X\rangle = E(XX) = E(X^2) \geq 0$$

5. $\langle X, X\rangle = 0$ si y solo si $X \overset{c.s}{=} 0$.
Esta condición no se satisface en un sentido estricto. La condición se mantiene si no se distingue entre variables aleatorias que difieren sobre un conjunto de probabilidad cero. Teniendo en cuenta la relación de equivalencia definida anteriormente se tendrá,
$\Leftarrow$) Si $X \overset{c.s}{=} 0$ y $N = \{\omega \in \Omega : X(\omega) = 0\}$ tal que $N^c = \{\omega \in \Omega : X(\omega) > 0\}$ con $P(N^c) = 0$.Por tanto,
$$\langle X, X\rangle = \int_\Omega X^2 dP = \int_N X^2 dP + \int_{N^c} X^2 dP = 0.$$
$\Rightarrow$) Si $\langle X, X\rangle = \int_\Omega X^2 dP = 0$ y $A_n = \{\omega \in \Omega : X^2(\omega) > \frac{1}{n}\}$. Se tendra que,
$$X^2(\omega) \geq \frac{1}{n}\chi_{A_n}(\omega)$$

$$\int_{\Omega} X^2 dP \geq \frac{1}{n}\int_{\Omega} \chi_{A_n}(\omega) dP$$

$$\int_{\Omega} X^2 dP \geq \frac{1}{n} P(A_n)$$

$$0 \geq \frac{1}{n} P(A_n)$$

Por tanto, $P(A_n) = 0$. Ahora sí, $X = \bigcup_{n\in\mathbb{N}} A_n$, obtendremos,

$$P(\{\omega \in \Omega : X(\omega) \neq 0\}) = P(\bigcup_{n\in\mathbb{N}} A_n) = \sum_{n\in\mathbb{N}} P(A_n) = 0.$$

Cumpliendo, la definición de dos variables aleatorias son iguales casi seguramente es decir, $X \stackrel{c.s}{=} Y$, si y solo si $P(\{\omega : X(\omega) \neq Y(\omega)\}) = 0$ concluyendo, $X \stackrel{c.s}{=} 0$.

Por último se necesita probar que si $X_n \in L^2$ con $n = 1, 2, ...,$ y $\|X_n - X_m\| \to 0$ cuando $m, n \to \infty$, entonces existe $X \in L^2$ tal que X_n converge casi seguramente a X, es decir, L^2 es completo. Para esto se necesita la proposición que afirma que si $X_n \in L^2$ y $\|X_{n+1} - X_n\| \leq 2^{-n}, n = 1, 2, ...,$ entonces existe una variable aleatoria X en $(\Omega, \mathscr{F}, P)$ tal que $X_n \to X$ con probabilidad uno [Brockwell, 2006, pág 62].
Para verificar la proposición, sea $X_0 = 0$ entonces $X_n = \sum_{j=1}^{n}(X_j - X_{j-1})$ ahora $\sum_{j=1}^{\infty} |X_j - X_{j-1}|$ es finito con probabilidad uno, ya que, por el teorema de la convergencia monótona [Rudin, 1981, pág 16],se tiene que $\lim_{n\to\infty} \sum_{j=1}^{n} |X_j - X_{j-1}|$ existe y es finito con probabilidad uno.

$$\begin{aligned}
E\left(\sum_{j=1}^{\infty} |X_j - X_{j-1}|\right) &= \int_X \sum_{j=1}^{\infty} |X_j - X_{j-1}| dP \\
&= \int_X \lim_{n\to\infty} \sum_{j=1}^{n} |X_j - X_{j-1}| dP \\
&= \lim_{n\to\infty} \int_X \sum_{j=1}^{n} |X_j - X_{j-1}| dP
\end{aligned}$$

$$
\begin{aligned}
&= \lim_{n\to\infty} \sum_{j=1}^{n} \int_X |X_j - X_{j-1}| dP \\
&= \sum_{j=1}^{\infty} E|X_j - X_{j-1}||1| \\
&\leq \sum_{j=1}^{\infty} \|X_j - X_{j-1}\| \\
&= \|X_1\| + \sum_{m=1}^{\infty} \|X_{m+1} - X_m\| \\
&\leq \|x_1\| + \sum_{j=1}^{\infty} \frac{1}{2^j} \\
&< \infty
\end{aligned}
$$

Concluyendo, que si X_n es una sucesión de Cauchy en L^2 entonces se podran encontrar $n_1, n_2, ...,$ tal que $n_1 < n_2 < ...,$ y

$$
\|X_n - X_m\| \leq 2^{-k} \ \ para \ \ n, m > n_k
$$

Por la proposición anterior, existe una variable aleatoria X tal que $X_n \to X$ con probabilidad uno cuando $n \to \infty$.Ahora dado $\epsilon > 0$, para m y n suficientemente grandes, tendremos que $\|X_m - X_n\| < \epsilon$ luego por lema de Fatou [Rudin, 1981, pág 23], para valores grande de m.

$$
\begin{aligned}
E[(X_m - X)^2] &= E[\liminf_{n->\infty}(X_m - X_n)^2] \\
&\leq \liminf_{n\to\infty} E[(X_m - X_n)^2] \\
&= \liminf_{n\to\infty} \|X_m - X_n\|^2 \\
&< \epsilon^2.
\end{aligned}
$$

De modo que $\limsup_{m\to\infty} E[(X_m - X)^2] < \epsilon^2$, y por tanto, $X_m \xrightarrow{m.c} X$.

Así, L^2 es un espacio de Hilbert, completo con un producto interno definido.

1.2.3. Aproximación lineal en $L^2(\Omega, \mathscr{F}, P)$

Se comenzara dando un ejemplo del uso del teorema de la proyección particularmente en espacios de Hilbert. El resultado general es entonces establecido más adelante. Para esto suponga que X_1, X_2 y Y son variables aleatorias en $L^2(\Omega, \mathscr{F}, P)$ donde, solamente X_1 y X_2 son conocidas podríamos desear estimar el valor de Y utilizando la combinación lineal $\widehat{Y} = \alpha_1 X_1 + \alpha_2 X_2$ la cual minimiza el error cuadrático medio,

$$S = E|Y - \alpha_1 X_1 - \alpha_2 X_2|^2 = \|Y - \alpha_1 X_1 - \alpha_2 X_2\|^2.$$

el objetivo es encontrar un elemento $\widehat{Y}$ en el conjunto

$$\mathscr{M} = \left\{ X \in L^2(\Omega, \mathscr{F}, P) : X = a_1 X_1 + a_2 X_2 \ \text{ para algún } \ a_1, a_2 \in \mathbb{R} \right\}$$

cuya distancia cuadrada de Y, $\|Y - \widehat{Y}\|^2$ es tan pequeña como sea posible. Se esperaría que $\widehat{Y}$ tenga la propiedad que $Y - \widehat{Y}$ sea ortogonal a todos los elementos de $\mathscr{M}$, esto se podría aplicar en casos más generales, que sera establecido en el *teorema de la proyección*. Aplicando esto a la estimación de Y descrita más arriba, se obtiene

$$\langle Y - \alpha_1 X_1 - \alpha_2 X_2, X \rangle = 0 \ \text{ para todo } \ X \in \mathscr{M} \tag{1.1}$$

o equivalentemente, por la linealidad del producto interior,

$$\langle Y - \alpha_1 X_1 - \alpha_2 X_2, X_i \rangle = 0, \quad i = 1, 2.$$

Dada la definición del producto interior en L^2, lo anterior puede quedar expresado en términos de la esperanza como,

$$\alpha_1 E(X_1^2) + \alpha_2 E(X_2 X_1) = E(Y X_1),$$

$$\alpha_1 E(X_1 X_2) + \alpha_2 E(X_2^2) = E(Y X_2),$$

del cual α_1 y α_2 se pueden encontrar.
El teorema clave que soporta toda la teoría del siguiente capítulo llamado el *Teorema de la proyección* afirma, que sí $\mathscr{M}$ es un subespacio cerrado del espacio de Hilbert $\mathscr{H}$ y

$x \in \mathscr{H}$ [Bartle, 1995], entonces, hay un único elemento $\widehat{x} \in \mathscr{M}$ tal que

$$\|x - \widehat{x}\| = \inf_{y \in \mathscr{M}} \|x - y\|$$

y $\widehat{x} \in \mathscr{M}$ y $\|x - \widehat{x}\| = \inf_{y \in \mathscr{M}} \|x - y\|$ si y solo si $\widehat{x} \in \mathscr{M}$ y $(x - \widehat{x}) \in \mathscr{M}^{\perp}$.(El elemento $\widehat{x}$ es llamado la proyección de x sobre $\mathscr{M}$). La demostración del presente teorema se puede ver en [Kreyszig, 1989]. Dado un espacio de Hilbert $\mathscr{H}$, un subespacio cerrado $\mathscr{M}$, y un elemento $x \in \mathscr{M}$, por el teorema de la proyección se muestra que el elemento de $\mathscr{M}$ más cercano a x es el único elemento $\widehat{x} \in \mathscr{M}$ tal que

$$\langle x - \widehat{x}, y \rangle = 0 \quad \text{para todo} \quad y \in \mathscr{M} \tag{1.2}$$

Se interpretara $\widehat{x} = P_{\mathscr{M}} x$ como la mejor predicción de x en el subespacio $\mathscr{M}$.

El teorema de la proyección nos deja que $\widehat{x} = P_{\mathscr{M} x}$ está únicamente determinada por esta condición para cualquier espacio de Hilbert $\mathscr{H}$ y un subespacio cerrado $\mathscr{M}$.
Esto justifica en particular el uso de la ecuación 1.1, el teorema de la proyección juega un rol fundamental en todos los problemas que involucran aproximación o predicción de variables aleatorias con varianza finita.

CAPÍTULO 2

Modelos autorregresivos

El modelo matemático para una serie temporal es el concepto de proceso estocástico supondremos que el valor observado de la serie en el instante t es una extracción al azar de una variable aleatoria definida en dicho instante, en consecuencia, una serie de n datos será una muestra de un vector de n variables aleatorias ordenadas en el tiempo $(z_1, ..., z_t, ..., z_n)$.

2.1. Procesos estocásticos

Se denomina proceso estocástico al conjunto de variables $\{z_t\}$, $t = 1.....n$ y la serie observada se considera una realización o trayectoria del proceso [Peña Sanchez, 1995], más formalmente sea $(\Omega, \mathscr{F}, P)$ un espacio de probabilidad, $z = (z_t, t \in T)$ una familia de variables aleatorias definidas sobre $(\Omega, \mathscr{F}, P)$ y con valores en un espacio medible (E, B).
$\{z_t\}$ es llamado un proceso estocástico con espacio muestral $(\Omega, \mathscr{F}, P)$ espacio condicional (o espacio de datos) (E, B) y conjunto tiempo T [Bosq, 2000].

Se debe tener en cuenta que sí T es contable, entonces el proceso de tiempo es discreto mientras, que sí T es un intervalo en $\mathbb{R}$ ($T = [a, b], \mathbb{R}^+, \mathbb{R}$), $\{z_t\}$ es un proceso de tiempo continuo. Por ejemplo, se quiere saber la cantidad de autos rojos vendidos en una agencia de autos en los próximos 6 meses, teniendo en exhibición 4 colores negro, rojo, gris y azul.[Coleman, 1986],donde el espacio muestral sera $0, 1, 2, 3, 4$ que consideraremos una realización del proceso $\{z_t\}$, $t = 1, 2, .., 6$, siendo z_t la variable aleatoria número de autos rojos vendidos en los t meses.
La forma probabilística de un proceso estocástico se conocerá si tenemos la *distribución conjunta* de las n variables aleatorias $\{z_t\}$, determinar prácticamente esta distribución lleva consigo un buen número de realizaciones, para lograr simplificar esta determinación, se supone que la distribución conjunta es normal multivariante, ya que, entonces, quedará determinada por el vector de medias y la matriz de varianzas y covarianzas entre las variables [Peña Sanchez, 1995].
Para esto, una función del tiempo que proporciona las medias de las distribuciones marginales z_t para cada instante:

$$\mu_t = E[z_t]$$

se le llamará *función de medias*.
La *función de varianzas* del proceso proporciona las varianzas en cada instante temporal:

$$\sigma_t^2 = Var(z_t)$$

y se dira que el proceso es estable en la varianza si ésta es constante en el tiempo. La función que describe las covarianzas en dos instantes cualesquiera:

$$Cov(t, t + j) = Cov(z_t, z_{t+j}) = E[(z_t - \mu_t)(z_{t+j} - \mu_{t+j})]$$

se le llamará *función de autocovarianza*. A la estandarización de la función de covarianzas:

$$\varrho(t, t + j) = \frac{Cov(t, t + j)}{\sigma_t \sigma_{t+j}}$$

se le llamará *función de autocorrelación*.

Diremos que un proceso puede ser estable en la media y no en la varianza y al contrario,

además, la estructura de dependencia lineal entre las variables aleatorias del proceso se representa por las funciones de covarianza y correlación.
En general, la función de autocorrelacion y autocovarianzas dependen de dos parametros (t, j), siendo t el instante inicial y j el intervalo de observaciones.

Para poder estimar las características del proceso como medias, varianzas, entre otros a partir de su cambio en el tiempo es necesario suponer que las distribuciones de las variables en cada instante son estables a lo largo del tiempo. Obteniedo así el concepto de estacionaridad.

2.1.1. Procesos estacionarios

Se dice que un proceso estocástico (serie temporal) es *estacionario en sentido débil* si existen y son estables la media, la varianza y las covarianzas, es decir, si para todo t

1. $\mu_t = \mu = cte$
2. $\sigma_t^2 = \sigma^2 = cte$
3. $Cov(t, t+k) = Cov(t, t-k) = \gamma_k \qquad k = 0, \pm 1, \pm 2, ...$

Para un proceso estacionario la función de autocorrelación se calcula mediante;

$$\varrho_k = \gamma_k / \gamma_0$$

teniendo en cuenta que $\gamma_0 = \sigma^2$. Una propiedad importante de los procesos estacionarios es tener incrementos estacionarios; es decir, si z_t es estacionario, el proceso ω_t definido por:

$$\omega_t = z_t - z_{t-1}$$

es también estacionario.

- Media constante.

$$E[\omega_t] = E[z_t - z_{t-1}] = E[z_t] - E[z_{t-1}] = 0$$

- Varianza constante.
  Teniendo en cuenta lo demostrado en el item anterior.

$$\begin{aligned}
Var(\omega_t) &= E[(\omega_t - \mu)^2] \\
&= E[(z_t - z_{t-1} - \mu)(z_t - z_{t-1} - \mu)] \\
&= E[-2z_t\mu + 2z_{t-1}\mu - 2(z_t)(z_{t-1}) + \mu^2] + E[z_t^2] + E[z_{t-1}^2] \\
&= E[-2(z_t)(z_{t-1})] + E[z_t^2] + E[z_{t-1}^2] \\
&= -2Cov(z_t, z_{t-1}) + Var(z_t) + Var(z_{t-1}) \\
&= 2(\sigma^2 - \gamma_1).
\end{aligned}$$

  llamando γ a la covarianza entre observaciones contiguas.

- Función de autocovarianza.

$$Cov(\omega_t, \omega_{t+k}) = E[(z_t - z_{t-1})(z_{t+k} - z_{t+k-1})] = 2\gamma_k - \gamma_{k+1} - \gamma_{k-2}.$$

Un proceso estacionario muy importante es el ruido blanco, que se define como,

1. $E[z_t] = 0$
2. $Var(z_t) = \sigma^2$
3. $Cov(z_t, z_{t-k}) = 0, \quad k = \pm 1, \pm 2, ...$

Por ejemplo si tiramos una moneda en instantes $t = 1, 2, ...$ y definimos $z_t = -1$ si se obtiene cara y $z_t = +1$ si se obtiene cruz, se obtiene un proceso de ruido ya que la esperanza es cero, la varianza es constante, es decir, igual a uno y las covarianzas son cero [Peña Sanchez, 1995].
Los procesos estacionarios juegan un papel crucial en el análisis de series de tiempo además, muchas series de tiempo observadas no son estacionarias en apariencia. Con frecuencia, tales conjuntos de datos pueden ser transformados por ciertas técnicas, que razonablemente pueden modelarse como realizaciones de un proceso estacionario.
Método 1 realiza una estimación por mínimos cuadrados de m_t.
Método 2 es para un suavizado por medio de una media móvil.

Método 3 es una diferenciación para generar datos estacionarios.
La teoría de procesos estacionarios se utiliza para el análisis, ajuste y predicción de la serie resultante. En todo esto la función de autocovariancia es una herramienta primaria.

Existe un modelo clasico de descomposición de un proceso estacionario que permite la posibilidad de representar los datos como una realización del proceso escrita:

$$z_t = m_t + s_t + Y_t$$

donde m_t es una función de cambio lento conocida como *componente tendencia*, s_t es una función con el periodo conocido d conocido como *componente estacional* y Y_t es una *componente de ruido blanco* que se sabe que es estacionario por lo definido anteriormente.
El objetivo es estimar y extraer las componentes deterministicas m_t y s_t con la esperanza de que la componente residual o ruido blanco resultará ser un proceso aleatorio estacionario. Podemos utilizar la teoria de tales procesos para encontrar un modelo probabilistico satisfactorio para el proceso y analizar sus propiedades y utilizarlo en conjugación con m_t y s_t con fines de prediccion modelización y analisis de z_t.

Se usará en el presenté trabajo el **método 3**, donde se intenta eliminar el término tendencia m_t mediante la diferenciación, definido el operador derivada con la notación ∇ tal que

$$\nabla z_t = z_t - z_{t-1} = (1 - B)z_t$$

donde B es el operador de desplazamiento hacia atras,

$$Bz_t = z_{t-1}.$$

y,

$$B^j(z_j) = z_{t-j} \quad \text{y} \quad \nabla^j(z_t) = \nabla(\nabla^{j-1}(z_t)), \quad j \geq 1 \text{ con } \nabla^0(z_t) = z_t.$$

Para j=2,

$$\begin{aligned}\nabla(\nabla z_t) &= \nabla((1-B)z_t)\\ &= (1-B)\nabla(z_t)\\ &= (1-B)(1-B)z_t\\ &= (1-2B+B^2)z_t\\ &= z_t - 2z_{t-1} + z_{t-2}\end{aligned}$$

Si el operador ∇ se aplica a una función de tendencia lineal $m_t = at + b$, se obtiene la función constante $\nabla m_t = a$.
De la misma manera cualquier polinomio con tendencia de grado k puede ser reducido a una constante por la aplicación del operador ∇^k.
Comenzando con el modelo $z_t = m_t + Y_t$ donde $m_t = \sum_{j=0}^{k} a_j t^j$ y Y_t estacionario con media cero, se obtiene que

$$\nabla^k z_t = k!a_k + \nabla^k Y_t$$

un proceso estacionario con media $k!a_k$.

La técnica de diferenciación que aplicamos anteriormente a los datos no estacionarios se pueden adaptar para hacer frente a la estacionalidad del periodo d introduciendo el operador de diferencia ∇d definido por

$$\nabla d \;\; z_t = z_t - z_{t-d} = (1-B^d)z_t.$$

Aplicando el operador ∇d al modelo

$$z_t = m_t + s_t + Y_t$$

donde $\{s_t\}$ tiene periodo d se obtiene

$$\nabla d \;\; z_t = m_t - m_{t-d} + Y_t - Y_{t-d}$$

la cual da una descomposición del diferencial $\nabla_d \;\; z_t$, en una componente tendencia $(m_t - m_{t-d})$ y un ruido blanco $(Y_t - Y_{t-d})$.

Para ver la utilidad del método de diferenciación, sea $\{Y_t\}$ un proceso estacionario con media cero y a, b constantes.

(a) Si $z_t = a + bt + s_t + Y_t$ donde s_t es una componente estacional con periodo 12, probaremos que

$$\nabla\nabla_{12}z_t = (1 - B)(1 - B^{12})z_t$$

es estacionario.

Dado que muchas series no son estacionarias, pero existen métodos que transforman las series en procesos estacionarios el objetivo será estimar y extraer las componentes m_t y s_t de tal forma que la componente residual o ruido llegará a ser un proceso aleatorio estacionario, por lo cual primero se aplica el operador diferencial para eliminar la componente de estacionalidad $\{s_t\}$ de la serie de tiempo.

$$\begin{aligned}
\nabla_{12}z_t &= z_t - z_{t-12} \\
&= (1 - B^{12})z_t \\
&= a + bt + s_t + Y_t - a - b(t - 12) - s_{t-12} - Y_{t-12} \\
&= Y_t + 12b - Y_{t-12}.
\end{aligned}$$

Ahora aplicando el diferencial para eliminar la componente de tendencia $a + bt$.

$$\begin{aligned}
\nabla\nabla_{12}z_t &= \nabla\left[(1 - B^{12})z_t\right] \\
&= (1 - B^{12})\nabla z_t \\
&= (1 - B^{12})(1 - B)z_t \\
&= (1 - B - B^{12} + B^{13})z_t \\
&= z_t - z_{t-1} - z_{t-12} + z_{t-13} \\
&= Y_t - Y_{t-12} + Y_{t-13}.
\end{aligned}$$

Obteniendo asi un residual o ruido que se sabe que es estacionario y centrado en cero.

(b) Si $z_t = (a + bt)s_t + Y_t$ donde s_t es de nuevo una componente estacional con periódo

12 probaremos que

$$\nabla_{12}^2 z_t = (1 - B^{12})(1 - B^{12}) z_t$$

es estacionaria.
Aplicando diferenciación para eliminar la componente de estacionaridad.

$$\begin{aligned}\nabla_{12} z_t = (1 - B^{12}) z_t &= z_t - z_{t-12} \\ &= (a - bt)s_t + Y_t - (a + b(t-12))s_{t-12} - Y_{t-12} \\ &= Y_t - (12b)s_t - Y_{t-12}.\end{aligned}$$

Por último, aplicando diferenciación para eliminar la componente de tendencia.

$$\begin{aligned}\nabla_{12}^2 z_t &= (1 - B^{12})(1 - B^{12}) z_t \\ &= (1 - B^{12} - B^{12} + B^{24}) z_t \\ &= z_t - 2z_{t-12} + z_{t-24} \\ &= Y_t + 2Y_{t-12} + Y_{t-24}.\end{aligned}$$

Se obtiene a Y_t que es un proceso estacionario con media cero.

2.1.2. Procesos autorregresivos (AR(1)).

Una clase muy importante de procesos estacionarios son los procesos autorregresivos, que resultan de imponer una dependencia entre las variables del proceso, similar a la de una ecuación de regresión.[Peña Sanchez, 1995].
La forma de dependencia más simple es relacionar z_t con z_{t-1} mediante la ecuación de *"autorregresion"*

$$z_t = c + \phi z_{t-1} + a_t, \quad |\phi| < 1 \tag{2.1}$$

donde c y ϕ son constantes a determinar y a_t es un proceso de ruido blanco, independiente de z_{t-k} para todo k positivo. El proceso (2.1) se denomina proceso autorregresivo de primer orden o , brevemente, $AR(1)$.

Las propiedades principales de este proceso son:

1. Si el proceso comienza con $z_0 = h$, siendo h un valor cualquiera fijo, $z_1 = c + \phi h + a_1$ tendrá una distribución normal y lo mismo le ocurrirá a z_t para cualquier t además, sustituyendo sucesivamente y suponiendo c=0.

$$
\begin{aligned}
z_1 &= \phi h + a_1 \\
z_2 &= \phi^2 h + \phi a_1 + a_2 \\
z_3 &= \phi^3 h + \phi^2 a_1 + \phi a_2 + a_3 \\
&. \\
&. \\
&. \\
z_t &= \phi^t h + \sum_{i=0}^{t-1} \phi^i a_{t-i}.
\end{aligned}
$$

   tomando esperanzas, como $E[a_t] = 0$

$$E[z_t] = \phi^t h$$

   el proceso sólo puede ser estacionario si $|\phi| < 1$. Entonces, después de un periodo transitorio incial, cuando $t \to \infty$, ϕ^t tenderá a cero.

2. *Función de medias.* Tomando esperanzas en 2.1 suponiendo que $|\phi| < 1$, de manera que $E[z_t] = E[z_{t-1}] = \mu$.

$$
\begin{aligned}
z_t &= c + \phi z_{t-1} + a_t \\
E[z_t] &= E[c + \phi z_{t-1} + a_t] \\
\mu &= E[c] + E[\phi z_{t-1}] + E[a_t] \\
\mu &= c + \phi E[z_{t-1}] + 0 \\
\mu &= c + \phi \mu \\
\mu &= \frac{c}{1 - \phi}.
\end{aligned}
$$

Llamando $\overline{z}_t = z_t - \mu$, el proceso puede escribirse en desviaciones a su media:

$$\overline{z}_t = \phi \overline{z}_{t-1} + a_t \tag{2.2}$$

que es la expresión más usada del $AR(1)$.

3. *Función de varianzas.* Como z_{t-1} y a_t son independientes, llamando σ_z^2 a la varianza del proceso:

$$\begin{aligned}
z_t &= c + \phi z_{t-1} + a_t \\
V[z_t] &= V[c + \phi z_{t-1} + a_t] \\
\sigma_z^2 &= V[c] + V[\phi z_{t-1}] + V[a_t] \\
\sigma_z^2 &= 0 + \phi^2 V[z_{t-1}] + \sigma_a^2 \\
\sigma_z^2 &= \phi^2 \sigma_z^2 + \sigma_a^2 \\
\sigma_z^2 &= \frac{\sigma_a^2}{1 - \phi^2}
\end{aligned}$$

De nuevo apareciendo la condición $|\phi| < 1$ que es necesaria ahora para que σ_z^2 sea finita y positiva.

4. *Función de autocovarianzas.* Utilizando (2.2) y escribiendo:

$$z_t - \mu = \phi(z_{t-1} - \mu) + a_t$$

multiplicando por z_{t-k}, tomando esperanzas y llamando:

$$\begin{aligned}
\gamma_k &= E[(z_{t-k} - \mu)(z_t - \mu)] \\
&= E[z_{t-k} z_t - z_{t-k}\mu - \mu z_t + \mu^2] \\
&= E[z_{t-k} z_t - z_{t-k}\mu - \mu^2 + \mu^2] \\
&= E[z_{t-k}(z_t - \mu)]
\end{aligned}$$

se obtiene

$$\gamma_k = \phi \gamma_{k-1} \quad k = 1, 2, 3, \ldots ; \gamma_o = \sigma_z^2$$

como $|\phi| < 1$ la dependencia entre observaciones se amortigua al aumentar el retardo en particular usando la formula encontrada en la función de varianzas:

$$\gamma_1 = \frac{\phi\sigma_a^2}{1-\phi^2}. \tag{2.3}$$

5. *Función de autocorrelación*. Se verifica:

$$\varrho_k = \frac{\gamma_k}{\gamma_0} = \frac{\phi\gamma_{k-1}}{\gamma_0} = \phi\varrho_{k-1}$$

como, según la formula encontrada para la función de varianzas y (2.3), $\varrho_1 = \phi$, esta función será:

$$\varrho_k = \phi^k \tag{2.4}$$

y cuando k es grande, ϱ_k tiende a cero con rapidez que depende de ϕ.

2.1.3. Proceso autorregresivo general (AR(p)).

Un proceso z_t es autorregresivo de orden p si

$$\overline{z}_t = \phi_1\overline{z}_{t_1} + ... + \phi_p\overline{z}_{t-p} + a_t \tag{2.5}$$

donde $\overline{z}_t = z_t - \mu$ y a_t es un proceso de ruido blanco independiente de z_{t-h}, para todo $h \geq 1$. Para manejar de forma más cómoda estos procesos se usara el *operador de desplazamiento hacia atrás*, B, definido en la sección anterior, de tal forma que la ecuación de un proceso $AR(P)$ es:

$$\overline{z}_t = (\phi B + ... + \phi_p B^p)\overline{z}_t + a_t$$

es decir:

$$(1 - \phi_1 B - \phi_2 B^2 - ... - \phi_p B^p)\overline{z}_t = a_t. \tag{2.6}$$

y renombrando a $\phi_p(B) = 1 - \phi_1 B - \phi_2 B^2 - ... - \phi_p B^p$ como un polinomio de grado p en el operador de retardo, cuyo primer término es la unidad, obtenemos,

$$\phi_p(B)\overline{z}_t = a_t \tag{2.7}$$

que es la expresión general de un proceso autorregresivo. Y se nombrará *ecuación característica* del proceso a: $\phi_p(B) = 0$, que estará en función de B. Esta ecuación tendrá p raìces $G_1^{-1}, ..., G_p^{-1}$, en general distintas escribiendo,

$$\phi_p(B) = \prod_{i=1}^{p}(1 - G_i B)$$

Finalizaremos, comprobando que el proceso será estacionario si $|G_i| < 1$, para todo i.
La modelización de una serie a través de un modelo AR exige que el modelo sea estacionario en media y varianza. La condición de estacionariedad en media exige que la $E(Y_t)$ no sea función del tiempo y la $E(Y_t)$ debe ser finita y determinada, observaremos qué condiciones deben verificar los parámetros $\phi_1, ...\phi_p$ para que el proceso sea estacionario.

1. *Función de medias* Tomando esperanzas en (2.5) e imponiendo que la media sea constante se obtiene:

$$\begin{aligned}
z_t &= c + \phi_1 z_{t-1} + ... + \phi_p z_{t-p} + a_t \\
E[z_t] &= E[c + \phi_1 z_{t-1} + ... + \phi_p z_{t-p} + a_t] \\
\mu &= E[c] + E[\phi_1 z_{t-1}] + ... + E[\phi_p z_{t-p}] + E[a_t] \\
\mu &= c + \phi_1 E[z_{t-1}] + ... + E[\phi_p z_{t-p}] + 0 \\
\mu &= c + \phi_1 \mu + ... + \phi_p \mu \\
\mu &= \frac{c}{1 - \phi_1 - ... - \phi_p}.
\end{aligned}$$

2. *Función de varianzas* Una forma alternativa de comprobar que el modelo autorregresivo de orden p es estacionario en varianza es comprobando que las raíces del polinomio característico, en módulo, sean menores que la unidad. Con el fin de comprobar si el modelo es estacionario en varianza se van a calcular las raíces del polinomio característico del modelo, para ello se iguala a cero la parte autorregresiva del modelo:

$$z_t - \phi_1 z_{t-1} - ... - \phi_p z_{t-p} = 0$$

sustituye z_t por λ^t y se obtiene,

$$\lambda^t - \phi_1 \lambda^{t-1} - ... - \phi_p \lambda^{t-p} = 0$$

dividiendo por λ^{t-p} se tiene,

$$\lambda^p - \phi_1 \lambda^{p-1} - ... - \phi_p = 0$$

La solución de la ecuación o raíces del polinomio de grado p se puede escribir como se hizo anteriormente,

$$\phi_p(B) = \prod_{i=1}^{p} (1 - G_i B)$$

así pues, si el modelo especificado para representar la serie (2.5) cumple las condiciones $|G_i| < 1$ y $\sum_{i=1}^{p} G_i^k \neq 1$ el modelo será estacionario en media y varianza.

3. *Función de autocovarianzas* Para esto multiplicamos (2.5) por $\bar{z}_{t-k}(k > 0)$, tomando esperanza, resulta

$$\gamma_k = \phi_1 \gamma_{k-1} + ... + \phi_p \gamma_{k-p} \quad k \geq 1 \tag{2.8}$$

además,

$$\gamma_0 = \phi_1 \gamma_1 + \phi_2 \gamma_2 + ... + \phi_p \gamma_p + \sigma_a^2$$

4. *Función de autocorrelación* Si dividimos (2.8) por γ_0 se obtienen las relaciones entre correlaciones

$$\varrho_k = \phi_1 \varrho_{k-1} + ... + \phi_p \varrho_{k-p} \qquad k > 0$$

y las correlaciones satisfacen la misma ecuación que el proceso

$$\phi_p(B) \varrho_k = 0 \quad k > 0 \tag{2.9}$$

La solución general de la ecuación es

$$\varrho_k = \sum_{i=1}^{p} A_i G_i^k \tag{2.10}$$

donde los A_i son constantes a determinar a partir de las condiciones iniciales y G_i es una solución de la ecuación característica.
Comprobar la condición de $|G_i| < 1$, es observar que la condición $|\varrho| < 1$ exige que no exista en (2.10) ningún G_i mayor que la unidad, ya que entonces cuando k aumenta el término G_i^k crecería sin limite, y no puede existir una raíz unitaria ya que entonces su componente G_i^k no decrece y los coeficientes ϱ_k no tienden a cero para ningún retardo [Peña Sanchez, 1995].

CAPÍTULO 3

Aplicación

Una de las características definitivas de las series temporales es que se trata de una lista de observaciones en las que el orden es importante ya que ordenar es muy importante porque hay dependencia y cambiar el orden podría cambiar el significado de los datos.

3.1. Análisis series de tiempo

Algunas preguntas importantes a tener en cuenta al mirar una serie de tiempo son [of Science, 2016]:

- ¿Existe una tendencia , lo que significa que, en promedio, las mediciones tienden a aumentar (o disminuir) con el tiempo?
- ¿Hay estacionalidad , lo que significa que hay un patrón repetitivo de altos y bajos relacionados con el tiempo del calendario, tales como estaciones, cuartos, meses, días de la semana, y así sucesivamente?

- ¿Son sus valores atípicos ? En la regresión, los valores extremos están muy lejos de su línea. Con datos de series de tiempo, sus valores extremos están muy lejos de otros datos.
- ¿Existe un ciclo o período de largo plazo sin relación con los factores de estacionalidad?
- ¿Existe varianza constante en el tiempo, o es la varianza no constante?
- ¿Hay cambios abruptos en el nivel de la serie o en la varianza?

3.1.1. Ejemplo

La siguiente tabla muestra las muertes accidentales mensuales en los Estados Unidos, 1973-1978. La siguiente gráfica es una gráfica de series de tiempo del número de muertes

	1973	1974	1975	1976	1977	1978
Jan	9007	7750	8162	7717	7792	7836
Feb	8106	6981	7306	7461	6957	6892
Mar	8928	8038	8124	7776	7726	7791
Apr	9137	8422	7870	7925	8106	8129
May	10017	8714	9387	8634	8890	9115
Jun	10826	9512	9556	8945	9299	9434
Jul	11317	10120	10093	10078	10625	10484
Aug	10744	9823	9620	9179	9302	9827
Sep	9713	8743	8285	8037	8314	9110
Oct	9938	9129	8433	8488	8850	9070
Nov	9161	8710	8160	7874	8265	8633
Dec	8927	8680	8034	8647	8796	9240

Cuadro 3.1: Muertes accidentales mensuales en U.S.A

accidentales mensuales en U.S.A., durante 6 años consecutivos. Mediante un gráfico de series de tiempo, simplemente queremos decir que la variable se representa en función del tiempo.

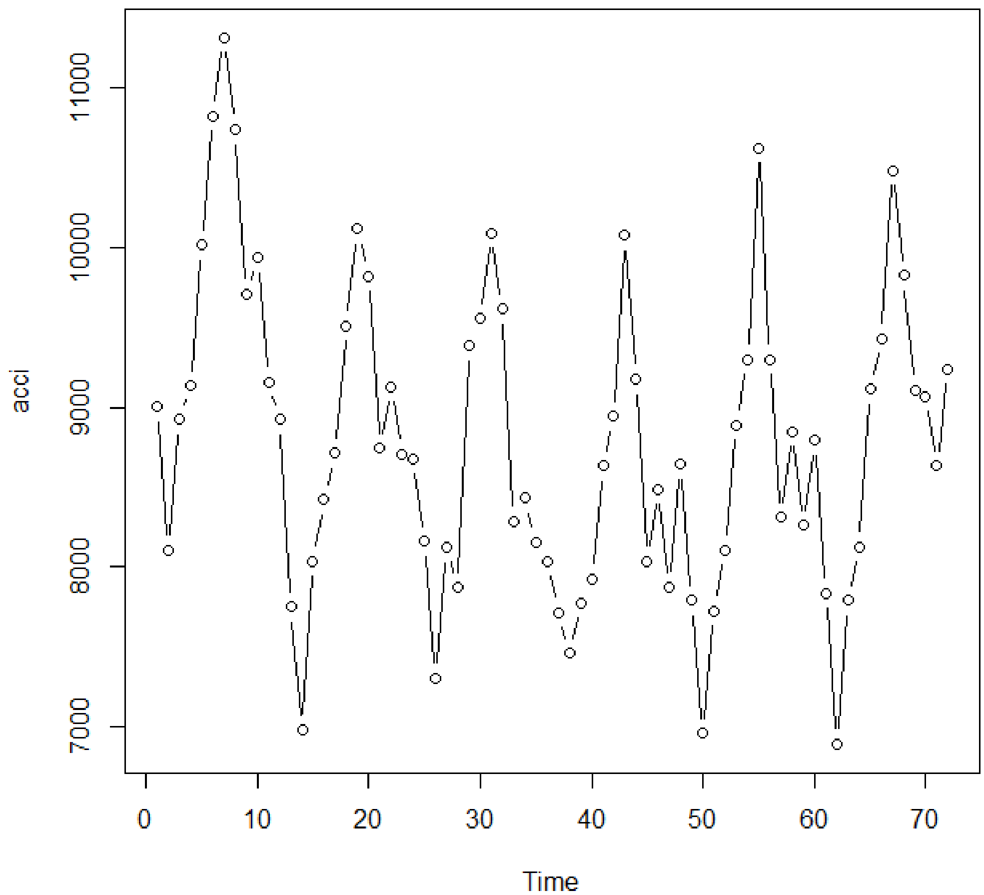

Figura 3.1: Gráfica muertes accidentales mensuales en U.S.A., $1973 - 1978$

De la gráfica anterior podemos decir, que no hay una tendencia consistente (ascendente o descendente) sobre todo el intervalo de tiempo. La serie parece vagar lentamente hacia arriba y hacia abajo. Casi por definición, se podría decir que existe una componente estacional pronunciada con periodo 12, no hay valores atípicos obvios y se podría juzgar si la varianza es constante o no.
El código usado para la figura 3.1, se muestra a continuación.

```
months<-c(1:72)
acci<-c
    (9007,8106,8928,9137,10017,10826,11317,10744,9713,9938,9161,8927,\\
```

```
6981,8038,8422,8714,9512,10120,9823,8743,9129,8710,8680,8162,7306,\\
8124,7870,9387,9556,10093,9620,8285,8433,8160,8034,7717,7461,7776,\\
7925,8634,8945,10078,9179,8037,8488,7874,8647,7792,6957,7726,8106,\\
8890,9299,10625,9302,8314,8850,8265,8796,7836,6892,7791,8129,9115,\\
9434,10484,9827,9110,9070,8633,9240)
names(acci)<-months
acci<-ts(acci)
plot(acci,type="b")
library(astsa)
```

El modelo lineal que se usará para predecir el valor en el tiempo presente usando un valor en un tiempo pasado será el llamado $AR(1)$.El orden del modelo expresa el número de observaciones retrasadas de la serie temporal analizada que intervienen en el modelo. Una de las formas de evaluar un modelo $AR(1)$ es gráficar los datos de la serie contra los valores retrasados un período de tiempo.

El código usado para la figura 3.2 se muestra a continuación,

```
> xlag1=lag(acci,-1)
> y=cbind(acci,xlag1)
> head(y)
      acci  xlag1
[1,]  9007     NA
[2,]  8106   9007
[3,]  8928   8106
[4,]  9137   8928
[5,] 10017   9137
[6,] 10826  10017
> lag1.plot(acci,1)
```

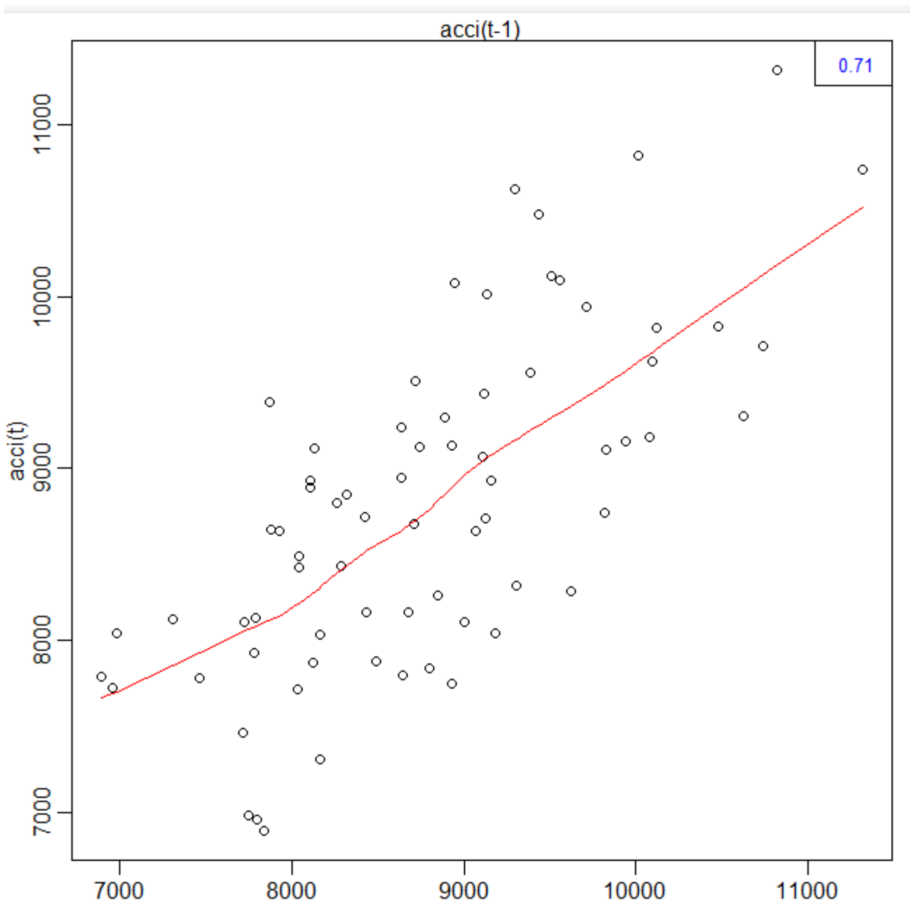

Figura 3.2: Gráfica x_t v.s. x_{t-1}

Por tanto,en este modelo, el valor de x en el tiempo t es una función lineal del valor de x en el tiempo $t-1$.Y el modelo escogido sera optímo para la predicción a realizar. La figura 3.3 es la función de autocorrelación donde, si $\phi_1 > 0$,la ACF decrece exponencialmente a 0 cuando $lag-h$ crece, o si $\phi_1 < 0$, la ACF decrece exponencialmente a 0 cuando $lag-h$ cree, pero el signo de las correlaciones se alterna entre positivo y negativo.El código de la figura 3.3 se muestra a continuación,

```
ar1fit=lm(y[,1]~ y[, 2])
> summary(ar1fit)
> plot(ar1fit$fit,ar1fit$residuals)
> cor<-acf(acci,type="covariance")
```

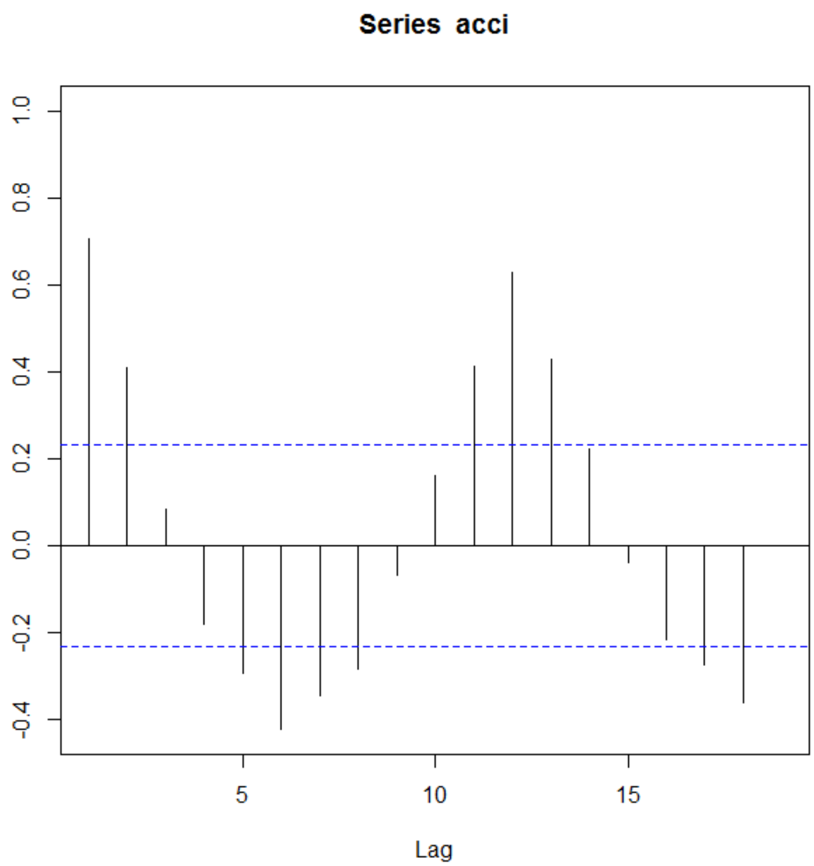

Figura 3.3: Gráfica 1

```
> plot(cor)
```

El principal propósito del correlograma es detectar autocorrelaciones en las series de tiempo luego de haberles removido y estimado la tendencia y la variación estacional.

CAPÍTULO 4

Conclusiones

En el presente trabajo se estudiarón las series de tiempo, profundizando desde los espacios de Hilbert, dada la relación que existe entre la geometría y en particular los conceptos de ortogonalidad y proyecciones que se vieron en teoremas esenciales como el teorema de la proyección que permite la práctica de las series de tiempo, donde pueden mencionarse varios motivos por los cuales es útil analizarlas por ejemplo, para obtener pronósticos de valores futuros de la serie, con el fin de ayudar a tomar decisiones que tienen consecuencias importantes o para entender mejor el mecanismo de generación de datos, que puede no ser claro inicialmente en una investigación. Se implemento un metodología que describió el patrón de la serie temporal y se concluyó que los usos para este modelo son: describir las características importantes del patrón de series temporales, explicar cómo el pasado afecta el futuro o cómo dos series de tiempo pueden interactuar para predecir los valores futuros de la serie y posiblemente servir como un estándar de control para una variable que mide la calidad del producto en algunas situaciones.

Bibliografía

[Bartle, 1995] Bartle, R. (1995). *The elements of integration and Lebesgue measure*. Wiley classics library. Wiley.

[Bosq, 2000] Bosq, D. (2000). *Linear Processes in Function Spaces, Theory and applications*. Springer science.

[Brezis, 2010] Brezis, H. (2010). *Functional Analysis, Sobolev Spaces and Partial Differential Equations*. Universitext. Springer New York.

[Brockwell, 2006] Brockwell, P. D. (2006). *Time series theory and methods*. Springer science.

[Coleman, 1986] Coleman, R. (1986). *Procesos estocásticos:Volumen 14*. Limusa.

[Grigoriu, 2002] Grigoriu, M. (2002). *Stochastic Calculus*. Springer science.

[Kreyszig, 1989] Kreyszig, E. (1989). *Introductory Functional Analysis with Applications*. John Wiley and Sons.

[of Science, 2016] of Science, E. C. (2016). Applied time series analysis. *The Pennsylvania State University*, pages 1–5.

[Peña Sanchez, 1995] Peña Sanchez, D. (1995). *Estadística Modelos y métodos*. Madrid:Alianza.

[Rudin, 1981] Rudin, W. (1981). *Real and Complex Analysis*. Mc Graw-Hill International Edtiions, 3 edition.