SERIES DE TIEMPO
LICENCIATURA CIENCIA DE DATOS
IIMAS-UNAM 1

RAMSÉS H. MENA

CLASE 1: PRESENTACIÓN Y CONCEPTOS BÁSICOS

1. Presentación

La teoría de ***series de tiempo*** o ***series temporales*** es la rama de la estadística que se dedica al estudio de la dependencia entre observaciones de un fenómeno aleatorio a través de un conjunto de índices, comúnmente identificados como el " tiempo ". Desde un punto de vista tradicional existen dos enfoques para el análisis de series de tiempo denominados el enfoque en el ***dominio de frecuencias*** y el enfoque en el ***dominio en el tiempo***. El primer enfoque, más vinculado a la Ingeniería y la Física, se basa principalmente en la teoría del análisis de Fourier y en el supuesto de estacionariedad, que abordaremos más adelante. Por otro lado, el enfoque en el dominio del tiempo, más común en disciplinas como Biología, Finanzas y Medicina, incorpora elementos de la teoría de la probabilidad y la estadística, y ofrece mayor flexibilidad en el análisis de procesos no estacionarios. Claramente, ninguno de los estos enfoques es exclusivo de un área dada.

A lo largo de este curso, adoptaremos principalmente el enfoque en el dominio del tiempo, en ocasiones centrado en la metodología Box-Jenkins. Este enfoque permite modelar y predecir series de tiempo mediante modelos autoregresivos integrados de medias móviles (ARIMA) y se basa en la identificación, estimación y diagnóstico del modelo, garantizando que los residuos sean ruido blanco.

En general, el estudio de series de tiempo se puede realizar utilizando la teoría de procesos estocásticos, en particular algunas de sus sub-disciplinas, tales como lo son el estudio de los ***procesos de Markov***, los ***procesos Gaussianos***, o ***procesos débilmente estacionarios***. Sin embargo, para algunas aplicaciones, la mayoría de estas sub-disciplinas son inadecuadas y/o muy generales. Como veremos más adelante el estudio de series temporales se puede ver como una rama de la teoría de procesos estocásticos cuya finalidad es la de simplificar el estudio de observaciones longitudinales.

Dado que parte de los conceptos necesarios para el estudio de series de tiempo, provienen de la teoría general de procesos estocásticos, razón por la cual la primera parte del curso está enfocada a una breve revisión de las nociones relevantes de dicha teoría.

Entonces, el ***objetivo*** principal de la teoría de las series temporales es el estudio, descripción y análisis de la dinámica que gobierna una sucesión de datos aleatorios dados a través del tiempo, ya sea a través de una ley de probabilidad o una ecuación estocástica. Este " análisis ", tiene a su ves diversas finalidades, como lo pueden ser el pronóstico de valores futuros o simplemente la explicación de un fenómeno observado.

En la actualidad existe una gran variedad de modelos que se pueden ajustar a una aplicación en particular, tales como los son los ***Autoregresivos*** (AR) y los modelos de ***promedios móviles*** (MA) o alguna de sus variantes no lineales. Parte de este curso estará concentrado al aprendizaje y aplicación de los modelos básicos lineales, no-lineales, y para datos con estructuras discretas o complejas.

2. Bibliografía recomendada

- Brockwell, P. and Davis, R. (2002). Introduction to Time Series and Forecasting. Springer.
- Shumway, R. and Stoffer, D. (2000). Time Series Analysis and Its Applications. Springer.
- Chatfield, C. (2003). The Analysis of Time Series: An Introduction. Chapman and Hall.

3. Software

La teoría de series de tiempo esta altamente ligada a aplicaciones reales, razón por la cual su aplicación es un factor esencial. Con esta finalidad en mente se utilizara el programa `R` (http://www.r-project.org) como herramienta computacional. Sin embargo, otros lenguajes de programación pueden ser utilizados.

4. Algunos conceptos básicos de procesos estocásticos

Definición 1. Sea $\mathcal{T}$ un conjunto de índices. Un ***proceso estocástico*** se define como un conjunto de variables aleatorias, $\{X_t\}_{t\in\mathcal{T}}$ definidas en un espacio de probabilidad $(\Omega, \mathcal{F}, \mathcal{P})$ y con valores en ***un espacio de estados*** denotado por $\mathcal{E}$.

Desde un punto de vista probabilístico lo ideal sería conocer la ***ley*** o ***distribución*** que rige dicho conjunto de variables aleatorias, en notación no rigurosa

$$\mathcal{P}(X_{t_1}, X_{t_2}, \dots, X_{t_n}), \tag{1}$$

para cualquier sub-conjunto de índices $t_1, t_2, \dots, t_n \in \mathcal{T}$. El estudio de procesos estocáticos desde la perspectiva distribucional, se fundamenta en el supuesto de proyectividad de sus ***distribuciones finito dimensionales*** (1). Sin embargo, suponer una ley para un conjunto de variables aleatorias, las cuales describen alguna característica de interés de un fenómeno dado, es una tarea compleja, en ocasiones restrictiva y poco realista. Tomando elementos de probabilidad y estadística, se pueden establecer un conjunto de técnicas que faciliten la elección de un modelo adecuado.

Un ejemplo de estas técnicas se da en el enfoque clásico de inferencia estadística, en donde un conjunto de observaciones se modelan a través de variables aleatorias independientes e idénticamente distribuidas (***iid***). En este caso la ley de un conjunto de variables aleatorias esta dada mediante el producto de las leyes marginales de variables aleatorias que modelan el fenómeno de interés, ignorando la dependencia temporal, es decir

$$\mathcal{P}(X_{t_1}, X_{t_2}, \ldots, X_{t_n}) = \prod_{i=1}^{n} \mathcal{P}(X_{t_i}).$$

En estadística, la distribución asociada generalmente está indexada por un conjunto de parámetros $\theta \in \Theta$, que deben ser estimados o acotados a partir de observaciones. Sin embargo, el supuesto de independencia entre variables aleatorias y de igualdad en sus distribuciones limita las aplicaciones de este enfoque, especialmente en fenómenos aleatorios observados a lo largo del tiempo, donde la dependencia entre observaciones es crucial. Muchos métodos estadísticos, como el análisis multivariado, la regresión, el diseño de experimentos y la estadística Bayesiana, se enfocan en estudiar la dependencia entre variables aleatorias. En la teoría de procesos estocásticos, varias disciplinas imponen estructuras de dependencia. Por ejemplo, la teoría de los procesos de Markov establece que el futuro y el pasado son independientes, dado el presente.

$$\mathcal{P}(X_{t_n} \mid X_{t_{n-1}}, \ldots, X_{t_1}) = \mathcal{P}(X_{t_n} \mid X_{t_{n-1}}), \tag{2}$$

donde $t_1 < t_2 < \cdots < t_n$.

Por otro lado, dentro de la misma teoría de procesos estocásticos existen, al menos, dos maneras de inducir dependencia entre variables aleatorias: una es mediante suposiciones distribucionales acerca de la dinámica que regula el proceso, e.g. asumir una forma especifica para $\mathcal{P}$ en (2), y la otra mediante una relación o ***ecuación estocástica*** que regula el comportamiento de un proceso a través del tiempo. Un ejemplo de esta última son las ***ecuaciones diferenciales estocásticas***, las cuales juegan un papel muy importante en modelos financieros. Por ejemplo el modelo de Cox-Ingersoll-Ross (CIR) para tasas de interés dado mediante la ecuación

$$dX_t = (\alpha + \beta\, X_t)\, dt + \gamma \sqrt{X_t}\, dW_t, \;\; X_0 = x \geq 0. \tag{3}$$

En la práctica, el estudio de procesos estocásticos se divide, esencialmente, por la naturaleza del conjunto $\mathcal{T}$ y el espacio de estados $\mathcal{E}$ en donde las variables aleatorias en cuestión toman valores.

<table>
<tr><td></td><td></td><td colspan="2">$\mathcal{T}$</td></tr>
<tr><td></td><td></td><td>C</td><td>D</td></tr>
<tr><td rowspan="2">E</td><td>C</td><td>Movimiento Browniano
Procesos de Difusión
Procesos de Lévy</td><td>Procesos de Markov con
espacio de estados gral.
Series de Tiempo</td></tr>
<tr><td>D</td><td colspan="2">Cadenas de Markov
Procesos Poisson</td></tr>
</table>

Cuadro 1. Algunas de las sub-teorías de procesos estocásticos, clasificadas en cuanto al " tiempo ", $\mathcal{T}$, y su espacio de estados, E.

El Cuadro 1, muestra algunas de las teorías más importantes de procesos estocásticos, subdivididas en cuanto a su espacio de estados y a su naturaleza en el tiempo. Aunque esta clasificación no es excluyente, e.g. se pueden tener series de tiempo con valores en un espacio discreto, nos da una idea de como son estudiadas en la literatura.

Definición 2. Una ***serie de tiempo*** se puede ver como una realización de un proceso estocástico a tiempo discreto, $\{X_t\}_{t=1}^{\infty}$.

En la práctica cuando observamos una serie de tiempo, e.g. el índice diario de precios y cotizaciones, el proceso " verdadero " que rige dicha serie, no se conoce. La idea detrás de la teoría de series de tiempo es el estudio de ***modelos para series de tiempo***, $\{X_t\}_{t=1}^{\infty}$, que nos ayuden a entender las características de dependencia (la dinámica) de la serie observada.

Al igual que en el estudio de procesos estocásticos los modelos para series de tiempo se pueden definir a través de una representación estocástica o mediante suposiciones distribucionales acerca de las variables aleatorias que representan una serie dada.

De manera ilustrativa, considérese el modelo autoregresivo de orden uno (AR(1)), este modelo se puede definir a través de una ecuación dada por

$$X_t = \theta\, X_{t-1} + \varepsilon_t, \tag{4}$$

donde $\varepsilon_t \overset{\text{iid}}{\sim} \mathrm{N}(0, \sigma^2)$. De manera " distribucional " se podría definir el mismo modelo mediante una sucesión de variables aleatorias $\{X_t\}_{t=1}^{\infty}$ con distribuciones condicionales $X_t \mid X_{t-1} \sim \mathrm{N}(\theta X_{t-1}, \sigma^2)$, lo cual, bajo ciertos supuestos implicaría que las distribuciones finito dimensionales que caracterizan el proceso definido por (4) están dadas por

$$f(x_t, x_{t-1}, \ldots, x_1) = \pi(x_1) \prod_{i=1}^{t-1} f(x_{i+1} \mid x_i), \tag{5}$$

donde $f(x_t \mid x_{t-1}) = \mathrm{N}(x_t; \theta\, x_{t-1}, \sigma^2)$ y $\pi(\cdot)$ denota la función de densidad de una distribución inicial.

Notación. La función $\mathrm{D}(x; \theta_1, \theta_2, \ldots)$, denotará la función de densidad (o masa, según sea el caso) correspondiente a una variable aleatoria con distribución $X \sim D(\theta_1, \theta_2, \ldots)$, e.g. $\mathrm{N}(\mu, \sigma)$, $\mathrm{Ga}(\alpha, \beta)$, etc.

Un tipo de procesos estocásticos muy importante en el estudio de series de tiempo son los ***procesos estacionarios***.

Definición 3. Un proceso estocástico $\{X_t\}$ se dice que es ***estrictamente estacionario*** si sus distribuciones finito dimensionales son invariantes ante traslaciones. Es decir,

$$\{X_{t_1}, X_{t_2}, \ldots, X_{t_n}\} \stackrel{\mathrm{d}}{=} \{X_{t_1+h}, X_{t_2+h}, \ldots, X_{t_n+h}\}, \quad (6)$$

para todo conjunto de indices, $h, t_1, t_2, \ldots, t_n \in \mathcal{T}$.

Suponer que un proceso es estrictamente estacionario es muy fuerte, sin embargo, razón por la cuál muchos gran parte de la literatura ha optado por "suavizar" dicho supuesto mediante el condicionamiento de momentos más que el de toda la distribución como se hace en (6). Para entender este punto mejor, primero necesitamos introducir la función de autocovarianza.

Definición 4. Sea $\{X_t\}_{t\in\mathcal{T}}$ un proceso tal que $\mathrm{Var}(X_t) < \infty$ para todo $t \in \mathcal{T}$, entonces la ***función de autocovarianza (ACV)*** $\gamma_X(\cdot,\cdot)$ de $\{X_t\}$ está definida por

$$\begin{aligned} \gamma_X(r,s) &= \mathrm{Cov}(X_r, X_s) \\ &= \mathbb{E}[(X_r - \mu_r)(X_s - \mu_s)], \end{aligned} \quad (7)$$

para toda $s, t \in \mathcal{T}$ y donde $\mu_r := \mathbb{E}[X_r]$.

Esta función es claramente una forma de medir la dependencia entre observaciones de la serie (o proceso) dadas en diferentes puntos del tiempo.

Definición 5. Una serie de tiempo $\{X_t\}_{t\in\mathcal{T}}$ se dice que es ***débilmente estacionaria*** o ***estacionaria de segundo orden*** si para toda $r, s, t \in \mathcal{T}$ se cumplen las siguientes condiciones

(i) $\mathbb{E}[|X_t|^2] < \infty$

(ii) $\mathbb{E}[X_t] = m$

(iii) $\gamma_X(r,s) = \gamma_X(r+t, s+t)$.

Si $\{X_t\}$ es estacionario de segundo orden entonces $\gamma_X(r,s) = \gamma_X(r-s, 0)$. Por lo tanto la función de autocovarianza de un proceso estacionario se puede denotar como

$$\gamma_X(h) := \gamma_X(h, 0). \quad (8)$$

A esta función comúnmente se le conoce como la ***función de autocovarianza con rezago (o "lag")*** h.

De la misma manera se puede definir la ***función de autocorrelación (ACF)*** de un proceso estacionario como mediante

$$\rho_X(h) := \frac{\gamma_X(h)}{\gamma_X(0)}. \quad (9)$$

Otra forma análoga de estudiar la dependencia subyacente a un proceso estacionario $\{X_t\}$ puede ser mediante el análisis de ***momentos mayores a dos***. Esto es mediante funciones del estilo

$$\xi_X^{s,r}(h) := \mathrm{Cov}(X_t^s, X_{t+h}^r). \quad (10)$$

Sin embargo, en la mayoría de los modelos que estudiaremos en este curso dicha cantidad no será necesaria.

5. Cantidades Muestrales

Una forma de estudiar la dependencia en un modelo es a través de los "momentos cruzados". Sin embargo, cuando nos enfrentamos a datos provenientes de una serie de tiempo real dichas cantidades no están directamente disponibles ya que en primera instancia no se ha asumido ningún modelo. Así pues, para inferir el grado de dependencia subyacente a un conjunto de ***datos observados***, $\{x_i\}_{i=1}^n$, se pueden usar las cantidades muestrales dadas a continuación.

- ***Media muestral:*** $\bar{x} = \frac{1}{n}\sum_{t=1}^n x_t$
- ***ACV muestral:*** $\hat{\gamma}(h) = \frac{1}{n}\sum_{t=1}^{n-|h|}(x_{t+|h|} - \bar{x})(x_t - \bar{x})$
- ***ACF muestral:*** $\hat{\rho}(h) = \hat{\gamma}(h)/\hat{\gamma}(0), \quad -n < h < n.$

En la Figura 1, se puede observa una serie de tiempo, la cual, por ejemplo, se podría modelar a través de un modelo como el dado mediante la ecuación (4).

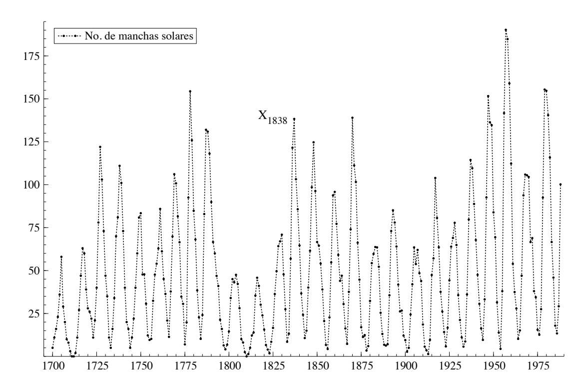

Figura 1: Número de manchas solares.