



UNIVERSIDAD DE SONORA

DIVISIÓN DE CIENCIAS EXACTAS Y NATURALES

FÍSICA COMPUTACIONAL

Actividad 6: Pronóstico de Series de Tiempo.

Grupo 2.

María José Ramos Murillo

Introducción

En esta actividad comenzamos con los **pronósticos de series de tiempo**. Esto realizando un modelo para predecir valores futuros con base en los datos pasados.

Usamos las bibliotecas:

```
import statsmodels.api as sm
from statsmodels.tsa.stattools import adfuller
from statsmodels.tsa.arima_model import ARIMA
from statsmodels.tsa.arima_model import ARIMAResults
from statsmodels.tsa.stattools import acf, pacf
```

Estas bibliotecas son adicionales a las anteriores que habíamos estado utilizando, pues estas últimas son para el análisis de Series de Tiempo y Modelos de Pronóstico.

Una serie temporal se define como una colección de observaciones de una variable recogidas secuencialmente en el tiempo. Estas observaciones se suelen recoger en instantes de tiempo equiespaciados. Si los datos se recogen en instantes temporales de forma continua, se debe o bien digitalizar la serie, es decir, recoger sólo los valores en instantes de tiempo equiespaciados, o bien acumular los valores sobre intervalos de tiempo.

Cuando se estudia una serie temporal, lo primero que se tiene que considerar:

- Si los datos presentan forma creciente (tendencia).
- Si existe influencia de ciertos periodos de cualquier unidad de tiempo (estacionalidad).
- Si aparecen outliers (observaciones extrañas o discordantes).

Las componentes o fuentes de variación que se consideran habitualmente son las siguientes:

1. **Tendencia:** Se puede definir como un cambio a largo plazo que se produce en relación al nivel medio, o el cambio a largo plazo de la media. La tendencia se identifica con un movimiento suave de la serie a largo plazo.
2. **Efecto Estacional:** Muchas series temporales presentan cierta periodicidad o dicho de otro modo, variación de cierto periodo (anual, mensual ...). Por ejemplo, el paro laboral aumenta en general en invierno y disminuye en verano. Estos tipos de efectos son fáciles de entender y se pueden medir explícitamente o incluso se pueden eliminar del conjunto de los datos, desestacionalizando la serie original.

-
3. **Componente Aleatoria:** Una vez identificados los componentes anteriores y después de haberlos eliminado, persisten unos valores que son aleatorios. Se pretende estudiar qué tipo de comportamiento aleatorio presentan estos residuos, utilizando algún tipo de modelo probabilístico que los describa.

En otras palabras, La estacionaridad significa que sus propiedades estadísticas (promedio, varianza, covarianza) permanezcan constantes en el tiempo. De esta manera se prevee que las propiedades de la serie de tiempo se van a comportar de la misma forma en el futuro.

Para aceptar o rechazar una hipótesis de que una serie de tiempo es no estacionaria, primero hay que hacer una *Prueba Aumentada de Dick-Fuller (ADF)*, la cual se mide con base en el valor $-p$.

Ya que obtuvimos la serie de tiempo estacionaria, ahora podemos conocer sus componentes, para lo cual construimos un modelo de pronóstico con el método estadístico ARIMA (Auto Regresive Integrated Moving Average).

Este modelo depende de los parámetros (p,d,q) , que se describen a continuación:

- p : es el número de términos AR (Auto-Regresivos). Por ej. si $p=3$, el predictor de $y(t)$ serán los términos $y(t-1)$, $y(t-2)$, $y(t-3)$.
- q : es el orden de términos del promedio móvil MA (Moving-Average). Por ej. si $q=3$, los términos para estimar el predictor de $y(t)$ serán los términos $y(t-1)$, $y(t-2)$, $y(t-3)$.
- d : es el número de diferenciaciones para hacer la serie estacionaria.

Los valores de p y q son enteros, y se obtienen a partir de las gráficas de ACF y PACF. Cuando estas se crucen por primera vez con la línea de $+1.96 \sigma$, donde σ es la desviación estándar. El cruce de ACF nos da información sobre el valor de q y el cruce de PACF nos da la información del valor de p .

Luego usamos el Criterio de Información de Akaike (AIC), como indicador de cuál es el mejor modelo, teniendo que el mejor es el que tiene menor AIC.

Primeras Impresiones

Esta actividad me pareció muy interesante porque ya pudimos realizar pronósticos de tiempo utilizando las herramientas de Python. La dificultad de los ejercicios no me pareció muy compleja debido a que hemos estado trabajando con el mismo tema y la secuencia hizo que cada vez se hiciera más sencillo aprender las cosas nuevas de

cada semana. Considero que la carga de trabajo fue adecuada y que es de un grado intermedio-avanzado.