

Introducción a Series de Tiempo

Leonardo A. Caravaggio

Abril 2022

¿Qué es una serie de tiempo?

Una serie de tiempo es un conjunto de datos indexado en el tiempo.

Constituyen la principal fuente de información utilizada en la estimación de Modelos Econométricos.

Componentes

La relación entre estas variables puede escribirse de la siguiente manera:

Observación=Componente sistémica o determinística + Perturbación aleatoria

Es decir:

$$y_t = x_{kt} + u_t$$

donde u_t representa un término de error, es decir, una diferencia aleatoria entre las variables.

Componentes

El procesamiento tradicional para la modelización de las series de tiempo consiste en la descomposición, sea en forma aditiva o multiplicativa, de sus componentes elementales: Tendencia (T), Variación Estacional (S), Variación Cíclica (C) y componente aleatoria (e). Si se adopta el esquema aditivo, la representación sería:

$$Y_t = T_t + S_t + C_t + e_t$$

Estacionariedad

No confundir **estacionariedad** con estacionalidad

Un proceso estocástico es **estacionario en sentido estricto** si todas las variables aleatorias que componen el proceso están idénticamente distribuidas, independientemente del momento del tiempo en que se estudie el proceso.

La **estacionariedad fuerte** es de difícil constatación empírica, por lo que en general se utiliza el criterio de estacionariedad débil. Se dice que un proceso es débilmente estacionario si se verifican simultáneamente las condiciones de estacionariedad en media y en varianza.

Autocorrelación

La **función de autocorrelación** (FAS) y la **función de autocorrelación parcial** (FAP) son medidas de asociación entre valores de series actuales y pasadas e indican cuáles son los valores más útiles para predecir valores futuros.

Práctica

Dejemos por un rato la teoría para ver algo de práctica.

Vamos a bajar algunas series de tiempo y observarlas usando Google Colab y la API de Alphacast.

Ver Ejercicio 1 en el Repositorio

Test de Estacionariedad

El test más conocido para chequear si una serie es no estacionaria (es decir, si tiene raíz unitaria) es el Test de Dicky Fuller aumentado.

Existen alternativas, como por ejemplo el test de Phillips–Perron (PP) o el test de ADF-GLS.

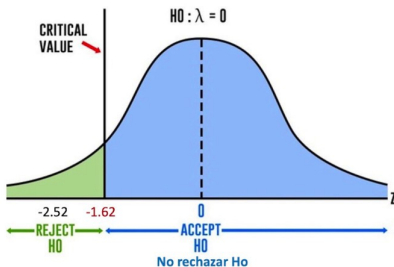
Dicky Fuller

Se trata de un test de hipótesis donde:

H0: Existe raíz unitaria (no es estacionario)

H1: No existe raíz unitaria (es estacionario)

Es decir que si el p-valor es menor al umbral seleccionado, podemos quedarnos tranquilos que la serie es no estacionaria.



Análisis Econométrico de las Series de Tiempo

La mayoría de las series empíricas pertenecientes al ámbito de la economía tienen un comportamiento **no estacionario**.

Esto quiere decir, que para evitar modelizar una correlación espuria, será necesario tornarlos estacionarios. Esto puede hacerse calculando el operador diferencia, aplicando una escala logarítmica, una escala inversa, etc.

Modelo sencillo

Podemos pensar un modelo sencillo de estimación del EMAE en base a dos estimadores adelantados de la actividad, el IPI y el ISAC.

$$EMAE_t = \beta_0 + \beta_1 IPI_t + \beta_2 ISAC_t + u_t$$

Modelo ARIMA

Es un modelo que combina una parte autoregresiva, una integración y una parte de medias móviles. Por ejemplo, un modelo ARIMA(p,d,q):

$$Y_t = -(\Delta^d Y_t - Y_t) + \phi_0 + \sum_{i=1}^p \phi_i \Delta^d Y_{t-i} - \sum_{i=1}^q \theta_i \varepsilon_{t-i} + \varepsilon_t$$

Más práctica

Podemos probar ahora los test de estacionariedad usando el paquete Statsmodel, y armar el modelo sencillo y un ARIMA bajando los datos con la API de Alphacast.

Ver Ejercicio 2 en el repositorio.

Modelos VAR

Un problema del modelo sencillo planteado anteriormente es que le imponen una dirección de causalidad a las variables.

Para evitar esto, cuando dos series de tiempo están interrelacionadas, se puede usar un modelo de Vectores Autoregresivos (VAR).

Modelos VAR

El nombre del modelo es bastante intuitivo respecto de su estructura. Se trata de un modelo autoregresivo, pero multivariado, permitiendo así la interacción entre las variables y los rezagos de las mismas.

$$Y_{1t} = \alpha_1 + \beta_{11} Y_{1,t-1} + \beta_{12} Y_{2,t-1} + u_{1t}$$

$$Y_{2t} = \alpha_2 + \beta_{21} Y_{1,t-1} + \beta_{22} Y_{2,t-1} + u_{2t}$$

Este sería un modelo VAR(1) porque cada variable tiene un rezago.

Test de Granger

Para identificar si una serie adelanta en información a la otra puede realizarse un test de Granger

Cointegración

Dos series están cointegradas si existe una combinación lineal que sea estacionaria.

¡Esto puede suceder aunque las dos series sean no estacionarias!

Test de Johanson

Una de las maneras de chequear si dos series están cointegradas es realizar el test de Johanson.

Test de Durbin Watson

Finalmente para chequear que no quede ninguna relación no identificada entre las series corresponderá chequear que no existan relaciones lineales entre los residuos. Esto puede hacerse con un test de Durbin Watson

Práctica de Modelos VAR

Vamos ahora a armar un modelo VAR en Python usando los datos de Alphacast. Previamente corresponderá también correr los test de Johanson y Granger.

Ver Ejercicio 3 en el repositorio.

Modelos ECM

Otra posibilidad para analizar series de tiempo cointegradas son los modelos de Corrección de Error (ECM).

Modelos VEC

También sería posible incorporar las ventajas de los modelos de vectores al modelo de corrección de errores, de manera de evitar la imposición de un sentido de causalidad. Este tipo de modelos se llama VEC (Corrección de errores vectorial).

Práctica Modelos ECM y VEC

Veamos para finalizar la implementación de estos modelos en Python.

Estos modelos se desarrollan en los Ejercicios 4 y 5 del repositorio.