Análisis de Series Temporales

Clase 2 - Modelos Autorregresivos

Rodrigo Del Rosso
RDelRosso-ext@austral.edu.ar

15 de Octubre de 2021

UNIVERSIDAD AUSTRAL | INGENIERÍA

CIENCIA DE DATOS
Maestría en Ciencia de Datos

Agenda

- FAC y Test Q
- Teorema de Descomposición de Wold
- Representación $MA(\infty)$
- Estacionariedad e Invertibilidad
- Modelos Autorregresivos $AR(p)$

Función de Autocorrelación

Función de Autocorrelación

Para estudiar la serie, lo ideal es previamente depurarla de dichos picos, dado que desarrollar el modelo sin depurarlo implica estimar mayor cantidad de coeficientes, lo cual reduce la confianza de una estimación derivada de una muestra de tamaño determinado.

Esta depuración es análoga a los filtros de audio, que cumplen la función de un colador.

- La FAC no es la más apta para estudiar los ciclos.
- La utilidad de la FAC se limita a detectar, sugerir que existen ciclos.
- Para estudiar los ciclos se aplica a la FAC una transformación ortogonal, que da origen al espectro del proceso.
- Con el correlograma se verifica si la memoria de la serie es corta o larga.
- Si es corta se pueden construir modelos de coyuntura y si es larga se puede transformar a costa de perder información.

Función de Autocorrelación

Para determinar si un proceso es $AR(0)$ se debe detectar si todos los coeficientes de autocorrelación $\rho_j(Y)$ son **iguales a cero**, considerando que dichos valores son desconocidos.

A partir de un conjunto finito de observaciones se forma la serie cronológica, de manera que pueda determinarse un estimador muestral y **construir un test de hipótesis**,

Planteo de Hipótesis

$H_0 : \rho_j\,(Y) = 0 \quad (j = 1, 2, \ldots)$

$H_1 : \rho_j\,(Y) \neq 0$

El estimador maximoverosimil de dicho parámetro es,

$$\hat{\gamma}_j\,(Y) = \sum_{t=1}^{n-j} \frac{(Y_t - m(Y))(Y_{t+j} - m(Y))}{n} \qquad (j = 0, 1, 2, \ldots)$$

$$\hat{\rho}_j\,(Y) = \frac{\hat{\gamma}_j(Y)}{\hat{\gamma}_0(Y)} = \frac{\sum_{t=1}^{n-j} (Y_t - m(Y))(Y_{t+j} - m(Y))}{\sum_{t=1}^{n} (Y_t - m(Y))^2}$$

Supone que es estacionaria

Función de Autocorrelación

El inconveniente es que el numerador tiene $n-j$ términos, de modo que, según el valor de j que se pretenda, la información contenida en el estimador es asimilable a la obtenida a partir de una muestra de tamaño $n-j$, y a medida que se incremente j, se asimilará a una muestra cada vez más chica, reduciéndose la confiabilidad de la estimación.

Se considera que dicha confiabilidad se mantiene dentro de niveles aceptables en tanto se cumpla que $j \leq \frac{n}{4}$, aplicable a series cortas, dado que si se tratara de series largas, de por ejemplo 4000 datos, podría garantizarse la confiabilidad hasta $j = 1000$, y sin embargo dicho valor carece de interés práctico.

La varianza del estimador se aproxima por el primer término de la **Fórmula de Bartlett**,

$$\sigma^2\left(\hat{\rho}_j\left(Y\right)\right) \cong \sum_{i=-(j-1)}^{j-2} \frac{\hat{\rho}_i^2(Y)}{n} = \frac{1}{n}\left(1 + 2\sum_{i=1}^{j-1} \hat{\rho}_i^2\left(Y\right)\right)$$

Función de Autocorrelación

La varianza no es constante. Incrementar j implica agregar más términos, todos positivos, cada uno menor al anterior porque la autocorrelación decrece al separarse en el tiempo las variables. Por lo tanto, la varianza del estimador crece con j pero a una tasa decreciente.

Estadístico

Bajo el supuesto de que H_0 es verdadera, se verifica que,

$$\hat{\rho}_j\left(Y\right) \sim N\left(0; \tfrac{1}{n}\right) \Rightarrow \sqrt{n}\hat{\rho}_j\left(Y\right) \sim N\left(0; 1\right)$$

Regla de Decisión

Para un nivel de significación del 5%, se tiene la siguiente regla de decisión,

$$Si \;\; \left|\hat{\rho}_j\left(Y\right)\right| > \tfrac{2}{\sqrt{n}} \Rightarrow \;\; RH_0$$

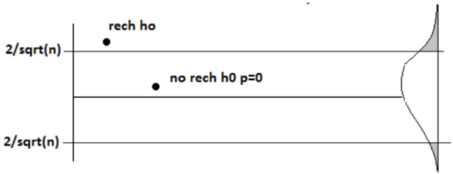
Este test de significatividad individual no es muy recomendable, porque para valores chicos de j se genera una autocorrelación entre los valores estimados, sesgándose el test.

Esto ocurre porque el numerador y el denominador del estimador poseen $n-j$ términos en común.

Función de Autocorrelación

El porcentaje de términos en común es $\frac{n-j}{n}$, o sea $1 - \frac{j}{n}$.

Cuanto mayor sea j, más tenderá a 0 el porcentaje de términos en común, reduciéndose así el peso de la autocorrelación que se genera.



Este test no es muy recomendado para decir si el fenómeno es $AR(0)$ o no porque los valores de j pequeños tienen mucha autocorrelación y los valores de j grandes no suele importar.

Test Q de Box-Pierce

Es un test de significatividad conjunta más potente que la prueba anterior.

Planteo de Hipótesis

$H_0 : \rho_1 (Y) = \rho_2 (Y) = \ldots = \rho_j (Y) = 0 \quad (j = 1, 2, \ldots)$

$H_1 : \exists \; al \; menos \; \rho_j (Y) \neq 0$

Estadístico

$$Q_j^{BP} = n \sum_{i=1}^{j} \hat{\rho}_i^2 (Y) \sim \chi_j^2$$

Se utiliza $\sigma^2 (\hat{\rho}_j (Y)) \cong \frac{1}{n}$ como estimador de la varianza, el cual produce un sesgo hacia el no rechazo de la hipótesis nula.

Test Q de Ljung-Box

Igual al test anterior de Box Pierce. En vez de emplear a $\sigma^2\left(\hat{\rho}_j\left(Y\right)\right) \cong \frac{1}{n}$ como estimador de la varianza, el cual produce un sesgo hacia el no rechazo de la hipótesis nula, se emplea la siguiente expresión en el estadístico,

$$\sigma^2\left(\hat{\rho}_j\left(Y\right)\right) \cong \frac{n-j}{n(n+2)}$$

Planteo de Hipótesis

$H_0 : \rho_1\left(Y\right) = \rho_2\left(Y\right) = \ldots = \rho_j\left(Y\right) = 0 \quad (j = 1, 2, \ldots)$

$H_1 : \exists \; al \;\; menos \;\; \rho_j\left(Y\right) \neq 0$

Estadístico

$$Q_j^{LB} = n(n+2)\sum_{i=1}^{j}\frac{\hat{\rho}_i^2(Y)}{n-i} \sim \chi_j^2$$

Teorema de Descomposición de Wold

Teorema de Descomposición de Wold

Herman Wold (1948) en su trabajo **"On prediction in stationary time series"** definió el siguiente teorema fundamental para el análisis de las series de tiempo estacionarias,

Si $\{Y_t\}_{t \in T}$ es un proceso estacionario no determinista entonces para todo $t \in T$ se puede expresar al proceso de la forma siguiente,

$$Y_t = \sum_{j=0}^{\infty} \psi_j \epsilon_{t-j} + V_t$$

Donde,

- $\psi_0 = 1$
- $\sum_{j=0}^{\infty} \psi_j^2 < \infty$
- $\{\epsilon_t\}_{t \in T}$ es un proceso Ruido Blanco (WN)
- $\gamma(\epsilon_s, V_t) = 0$ *para todo s y t*
- ϵ_s es el límite de combinaciones lineales de $Y_s \quad (s \leq t)$

Teorema de Descomposición de Wold

El proceso estocástico Y_t está conformado por infinitas variables que afectan su comportamiento de forma autorregresiva.

Cuando $j \to \infty$ se obtiene la siguiente representación,

$$Y_t = \sum_{j=0}^{\infty} \psi_j \epsilon_{t-j}$$

Es decir,

$$Y_t = \epsilon_t + \psi_1 \epsilon_{t-1} + \psi_2 \epsilon_{t-2} + \dots$$

Esta representación se conoce como $MA(\infty)$ **(Moving Average)** y se conforma por la combinación lineal de infinitos ruidos blancos.

$\{\psi_j\}$ es la sucesión de ponderadores que acompañan a las perturbaciones en cada momento del tiempo, por lo tanto ψ_j representa la persistencia de un shock aleatorio de j períodos atrás.

Teorema de Descomposición de Wold

Si se parte de la representación $MA(\infty)$ y se aplica el operador Backward (o Lag) se llega a la siguiente expresión,

$$Y_t = \epsilon_t + \psi_1 \epsilon_{t-1} + \psi_2 \epsilon_{t-2} + \dots$$
$$Y_t = \sum_{j=0}^{\infty} \psi_j \epsilon_{t-j} = \sum_{j=0}^{\infty} \psi_j B^j \epsilon_t = \epsilon_t \sum_{j=0}^{\infty} \psi_j B^j = \psi(B) \epsilon_t$$

Hint: $B^j \epsilon_t = \epsilon_{t-j}$

Donde $\psi(B)$ representa el operador polinómico en B de orden infinito de una representación MA.

$\{\psi_j\}$ la sucesión de estos ponderadores representa un desarrollo en serie que puede ser explosivo o convergente para $|B| \leq 1$.

Cuando $B = 1$ la representación es no estacionaria e indica cuánto tarda el sistema en asimilar un shock aleatorio.

Para sea $\psi(B)$ factible, el sistema debe cumplir ciertas condiciones de estabilidad. El operador, encargado de la transformación, se denomina filtro.

El sistema está caracterizado por el operador $\psi(B)$, el cual está definido por un desarrollo en serie de potencias de B el cual puede ser convergente o divergente.

Estacionariedad

La sucesión de ponderaciones va a ser finita cuando se cumpla la condición de "Cuadrado Sumable" en el intervalo $|B| \leq 1$ donde el desarrollo en serie converge.

$$\sum_{j=0}^{\infty} \psi_j^2 < \infty$$

Si la sucesión es convergente y cuadrado sumable entonces es finita, y si se cumple esto el proceso estocástico será estacionario de 2° orden o en las covarianzas.

La condición de estacionariedad está vinculada con la convergencia del operador

La condición de que la sucesión sea de cuadrado sumable es más visible e inobjetable que la autocovarianza ya que significa que el decrecimiento es cuasi-exponencial, lo que es muy rápido.

Estacionariedad

Hasta ahora los únicos indicios de estacionariedad eran,

- que los momentos de primer y segundo orden fueran estables, o sea invariables con respecto al tiempo
- que la función de autocorrelación decreciera exponencialmente en valor absoluto

En el primer indicio, resulta que los momentos son parámetros desconocidos, y sus estimadores satisfacen siempre la condición de invariabilidad respecto del tiempo.

En la otra situación, la condición depende exclusivamente de la lectura de un gráfico, que no siempre suele ser excesivamente prolijo, y por lo tanto la interpretación tiene un cierto grado de subjetividad.

La condición que acabamos de obtener, relacionada con el operador $\psi(B)$ es más clara.

Estacionariedad

Las autocovarianzas de las variables Y_t centradas están dadas en función del argumento j por,

$$\gamma_j\left(Y\right) = E\left(Y_t Y_{t+j}\right) = E\left(\sum_{i=0}^{\infty} \psi_i \epsilon_{t-i} \sum_{h=0}^{\infty} \psi_h \epsilon_{t+j-h}\right)$$

$$\gamma_j\left(Y\right) = E\left(\sum_{i=0}^{\infty} \sum_{h=0}^{\infty} \psi_i \psi_h \epsilon_{t-i} \epsilon_{t+j-h}\right) = \sum_{i=0}^{\infty} \sum_{h=0}^{\infty} \psi_i \psi_h E(\epsilon_{t-i} \epsilon_{t+j-h})$$

Las $E(\epsilon_{t-j}\epsilon_{t+j-h})$ representa las autocovarianzas de las perturbaciones que son Ruido Blanco, y serán todas nulas cuando el orden sea mayor que 0, mientras que serán no nulas cuando se refieran al mismo momento, o sea cuando $t - i = t + j - h$ o equivalentemente $h = i + j$.

$$\gamma_j\left(Y\right) = \sigma_\epsilon^2 \sum_{i=0}^{\infty} \psi_i \psi_{i+j}$$

Para $j = 0$ tenemos la varianza del proceso,

$$\gamma_0\left(Y\right) \equiv \sigma_Y^2 = \sigma_\epsilon^2 \sum_{i=0}^{\infty} \psi_i^2$$

y para asegurar la estabilidad del sistema (y consecuentemente la estacionariedad del proceso Y_t), debió satisfacerse que,

$$\sum_{j=0}^{\infty} \psi_j^2 < \infty$$

De esta forma, se encuentra una condición adicional que garantiza la estacionariedad: La condición necesaria y suficiente para que un proceso sea estacionario, es que su varianza $\sigma_Y^2 < \infty$ sea finita. Sin embargo, dicha varianza es desconocida.

Invertibilidad

La siguiente es una representación $AR(\infty)$ donde Y_t son variables centradas y ϵ_t Ruido Blanco,

$$Y_t = -\pi_1 Y_{t-1} - \pi_2 Y_{t-2} - \ldots + \epsilon_t$$

$$Y_t + \pi_1 Y_{t-1} + \pi_2 Y_{t-2} + \ldots = \epsilon_t$$

$$Y_t + \pi_1 B Y_t + \pi_2 B^2 Y_t + \ldots = \epsilon_t$$

$$\left(1 + \pi_1 B + \pi_2 B^2 + \ldots\right) Y_t = \epsilon_t$$

$$\pi(B) Y_t = \epsilon_t$$

Donde $\pi(B)$ es el operador $AR(\infty)$ y esta definido por la serie de potencias de B. La última expresión es legítima si el sistema es estable, es decir si para un entorno de B $|B| \leq 1$ el operador $AR(\infty)$ es convergente $(\pi(B) < \infty)$, lo cual puede garantizarse si el sistema de ponderaciones $\{\pi_j\}$ es absolutamente sumable, es decir si,

$$\sum_{j=0}^{\infty} |\pi_j| < \infty$$

Estacionariedad e Invertibilidad

La condición de sumabilidad absoluta implica la de cuadrados sumables, garantizándose la convergencia de la serie de potencias, y consecuentemente la estabilidad del sistema, con lo cual el proceso Y_t es invertible. Formalmente,

$$\sum_{j=0}^{\infty} |\pi_j| < \infty \Rightarrow \sum_{j=0}^{\infty} \psi_j^2 < \infty$$

La condición de invertibilidad implica la convergencia del operador $AR(\infty)$

Hay procesos que admiten ciertas representaciones AR, y hay procesos que admiten ciertas representaciones MA.

Sin embargo como en las representaciones MA el proceso queda expresado en función de variables aleatorias no observables, si se pretende operar con ese proceso, debe inevitablemente transformarse en una representación AR, dado que en ella el Y_t queda expresado en función de sus realizaciones anteriores, que son observables.

Por lo tanto, en esta idea de invertibilidad, el proceso debe ser tal que al pasar de la forma MA a la forma AR, la forma AR que se obtenga sea estacionaria.

Estacionariedad e Invertibilidad

- Las representaciones **MA son siempre estacionarias**

$$Y_t = \sum_{j=0}^{\infty} \psi_j \epsilon_{t-j} \Rightarrow E(Y_t) = \sum_{j=0}^{\infty} \psi_j E(\epsilon_{t-j})$$

Como $E(\epsilon_{t-j}) = 0$ por ser Ruido Blanco, entonces el **valor esperado del proceso** $\mathbf{E(Y_t) = 0}$.

El momento de primer orden es independiente del tiempo, por lo tanto Y_t es estacionario de primer orden. Respecto a la varianza,

$$\sigma^2(Y_t) = \sum_{j=0}^{\infty} \psi_j^2 \sigma^2(\epsilon_{t-j}) = \sigma_\epsilon^2 \sum_{j=0}^{\infty} \psi_j^2$$

El momento de segundo orden es independiente del tiempo. Luego, el proceso es **estacionario de segundo orden**. Dadas las representaciones *AR* y *MA* si se aplica en cualquiera de ellas un operador miembro a miembro, la igualdad no se altera.

Estacionariedad e Invertibilidad

En el caso particular de aplicar $\pi(B)$ a MA,

$$\pi(B)Y_t = \epsilon_t$$

Se aplica a ambos lados de la igualdad el operador $\psi(B)$

$$\psi(B)\ \pi(B)Y_t = \psi(B)\epsilon_t$$

$$Y_t = \psi(B)\epsilon_t$$

El producto entre los dos operadores $\psi(B)\ \pi(B)$ cumplen con una identidad. En términos formales,

$$\psi\left(B\right)\pi\left(B\right) = 1 \Rightarrow \pi\left(B\right) = \frac{1}{\psi(B)} = \left[\psi\left(B\right)\right]^{-1}$$

Es posible pasar de una representación a la otra, y viceversa, invirtiendo el operador. Con lo cual, las dos representaciones son dos formas de expresar el mismo concepto.

Modelos Autorregresivos

Modelos Autorregresivos

La representación es teórica y el interés radica en usufructuarla en la práctica.

Esto implicaría plantear un modelo que intente explicar al fenómeno Y_t a partir de infinitos shocks aleatorios pasados que generan una determinada influencia ponderada sobre el fenómeno Y_t, lo que requiere estimar cada uno de los respectivos coeficientes de ponderación.

Dado que se dispone de información finita, resulta imposible estimar los infinitos coeficientes.

Se pretende pasar de una representación de orden infinito a una de orden finito

Para que se verifique la estabilidad del sistema, el sistema de ponderaciones π_j debe ser absolutamente sumable $\sum_{j=0}^{\infty} |\pi_j| < \infty$ lo cual implica que las ponderaciones deben decrecer rápidamente en valor absoluto, es decir deben decrecer en forma exponencial (podría demostrarse analíticamente).

Modelos Autorregresivos

Con lo cual, a medida que j crece, los π_j se aproximan rápidamente a 0, de modo que, en términos de un modelo, podrían despreciarse, o sea,

$$Y_t = -\pi_1 Y_{t-1} - \pi_2 Y_{t-2} - \ldots + \epsilon_t$$

$$-\pi_1 = \phi_1$$

$$-\pi_2 = \phi_2$$

$$\vdots$$

$$-\pi_p = \phi_p$$

$$-\pi_j = 0 \quad (j = p+1, p+2, \ldots)$$

Entonces

$$Y_t = \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \ldots + \phi_p Y_{t-p} + \epsilon_t$$

Modelos Autorregresivos

Así, la representación $AR(\infty)$ de orden infinito se transforma en una de orden finito $AR(p)$.

Se dice que Y_t admite una representación $AR(p)$, o sea $Y_t : AR(p)$. En términos formales,

$$
\begin{gathered}
Y_t = \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \ldots + \phi_p Y_{t-p} + \epsilon_t \\
Y_t - \phi_1 Y_{t-1} - \phi_2 Y_{t-2} - \ldots - \phi_p Y_{t-p} = \epsilon_t \\
Y_t - \phi_1 B Y_t - \phi_2 B^2 Y_t - \ldots - \phi_p B^p Y_t = \epsilon_t \\
\left(1 - \phi_1 B - \phi_2 B^2 - \ldots - \phi_p B^p\right) Y_t = \epsilon_t \\
\Phi(B) Y_t = \epsilon_t
\end{gathered}
$$

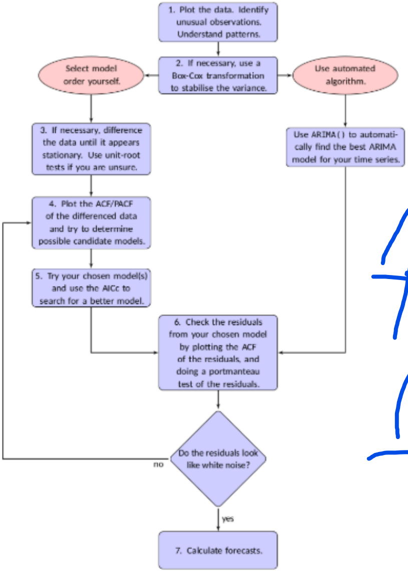
Donde $\Phi(B)$ es el operador $AR(p)$ y está definido por el polinomio de orden p.

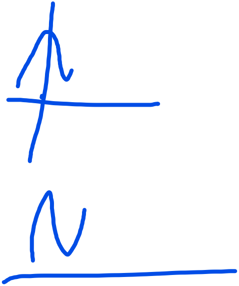
Metodología de Box-Jenkinks

Existe la Metodología de Box Jenkins nombrada así en honor a los estadísticos **George E. P. Box** y **Gwilym Jenkins**, se aplica a los modelos autorregresivos de media móvil **ARMA** o a los modelos autorregresivos integrados de media móvil **(ARIMA)** para encontrar el mejor ajuste de una serie temporal de valores, a fin de que los pronósticos sean más acertados

- Identificación del Modelo (Model Identification)
- Estimación de Parámetros (Parameter Estimation)
- Diagnóstico comparativo (Comparative Diagnosis)
- Pronóstico (Forecast)

Metodología de Box-Jenkinks





Invertibilidad y Estacionariedad de un $AR(p)$

Invertibilidad y Estacionariedad de un $AR(p)$

Condiciones de estacionariedad e invertibilidad de la $AR(p)$

La condición de invertibilidad está vinculada a la convergencia del operador $AR(\infty)$.

Los π_j fueron reemplazados por los ϕ_j determinando la siguiente sucesión finita de términos:

$$1 - \phi_1 B - \phi_2 B^2 - ... - \phi_p B^p$$

La cual es convergente porque posee una cantidad infinita de términos.

- La invertibilidad no está vinculada con los procesos $AR(p)$.
- La invertibilidad está vinculada con la convergencia del operador $MA(\infty)$.

Para analizar dicha condición, se requiere hallar la representación MA de este $AR(p)$.

$$\Phi(B) Y_t = \epsilon_t$$

Invertibilidad y Estacionariedad de un $AR(p)$

Las raíces surgen del Polinomio Característico del $AR(p)$,

$$1 - \phi_1 B - \phi_2 B^2 - ... - \phi_p B^p = 0$$

Esta ecuación produce p raíces $G_i \quad (i = 1, 2, ..., p)$, sobre las cuales se impone la siguiente **Condición de Estacionariedad**,

$$\left| G_j^{-1} \right| < 1 \qquad (j = 1, 2, \ldots, p)$$

Esta condición es equivalente a pedir que,

$$|G_j| > 1 \qquad (j = 1, 2, \ldots, p)$$

Ambas condiciones en forma alternativa garantizan la estacionariedad del proceso.

En la práctica ocurrirá que no se conocerán los valores de las raíces, y deberán estimarse, teniendo que decidir, a partir de los valores estimados, si las verdaderas raíces satisfacen la condición.

No se puede pasar en forma general de condiciones en las raíces a condiciones en los coeficientes, sino que debe plantearse caso por caso para cada valor de p.

Estacionariedad $AR(1)$ y $AR(2)$

Caso $AR(1)$

El Polinomio Característico es $\Phi(B) = 1 - \phi_1 B = 0$.

La raíz es $G_1 = \frac{1}{\phi_1} \Rightarrow \; G_1 = \left|\frac{1}{\phi_1}\right| > 1 \Rightarrow |\phi_1| < 1$

Caso $AR(2)$

El Polinomio Característico es $\Phi(B) = 1 - \phi_1 B - \phi_2 B^2 = 0$.

Las raíces G_1 y G_2 cumplirán con la condición si $|G_i| > 1$.

Mediante las **condiciones de Liapunov** deberá cumplirse que los coeficientes caigan dentro del círculo unitario,

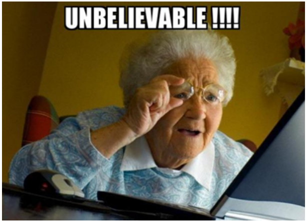
$$\phi_2 + \phi_1 < 1$$

$$\phi_2 - \phi_1 < 1$$

$$|\phi_2| < 1$$

Para otros valores de p el planteo genérico se complica cada vez más.

Estacionariedad e Invertibilidad



En resumen,

Todo modelo autorregresivo es estacionario si y solo si el valor absoluto de todas las raíces del PC de la ecuación que lo describe es mayor a uno

Todo proceso autorregresivo estacionario es invertible en un modelo $MA(\infty)$ que siempre es estacionario

Especificación del $AR(p)$

- A partir de la FAC se logra la especificación
- Para cada valor de p, el proceso tiene una familia de FAC típica asociada.
- Si se identifica que tiene determinada FAC típica asociada, entonces es un punto de partida para intentar conocer el valor p.

Para construir la FAC, deben primero determinarse las autocovarianzas (FAS).

Para ello tomamos la representación $AR(p)$ y miembro a miembro se multiplica por Y_{t-j} y se aplica el operador Esperanza Matemática.

$$Y_t = \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \ldots + \phi_p Y_{t-p} + \epsilon_t$$

La FAS se expresa como $\gamma_j(Y) = E\left(Y_t Y_{t-j}\right)$

La cual resulta ser para un $AR(p)$

$$\gamma_j(Y) = \phi_1 E(Y_{t-1} Y_{t-j}) + \phi_2 E(Y_{t-2} Y_{t-j}) + \ldots + \phi_p E(Y_{t-p} Y_{t-j}) + E(\epsilon_t Y_{t-j})$$

Función de Autocovarianzas (FAS)

Para $j = 0$ se obtiene la autocovarianza de orden 0, que es la varianza de Y,

$$\gamma_0(Y) = E(Y_t Y_t) = \phi_1\gamma_1(Y) + \phi_2\gamma_2(Y) + \ldots + \phi_p\gamma_p(Y) + E(\epsilon_t Y_t)$$

El último sumando de la ecuación anterior es siempre la varianza del Ruido Blanco,

$$E(\epsilon_t Y_t) = \gamma_0(\epsilon) = \sigma_\epsilon^2$$

$$\sigma_Y^2 = \phi_1\gamma_1(Y) + \phi_2\gamma_2(Y) + \ldots + \phi_p\gamma_p(Y) + \sigma_\epsilon^2$$

La Función de Autocovarianzas (FAS) del $AR(p)$ es infinita; es decir, existe y es no nula para cualquier valor de j.

$$\gamma_j(Y) = \phi_1\gamma_{|j-1|}(Y) + \phi_2\gamma_{|j-2|}(Y) + \ldots + \phi_p\gamma_{|j-p|}(Y) \qquad (j = 1, 2, \ldots)$$

Función de Autocorrelación

La FAC hereda las propiedades de la función de autocovarianzas, con lo cual la FAC del $AR(p)$ es infinita.

$$\rho_j(Y) = \frac{\gamma_j(Y)}{\gamma_0(Y)}$$

$$\rho_j(Y) = \frac{\phi_1\gamma_{j-1}(Y)+\phi_2\gamma_{j-2}(Y)+\ldots+\phi_p\gamma_{j-p}(Y)}{\gamma_0(Y)}$$

$$\rho_j(Y) = \phi_1\rho_{j-1}(Y) + \phi_2\rho_{j-2}(Y) + \ldots + \phi_p\rho_{j-p}(Y) \qquad (j = 1, 2, \ldots)$$

En el caso práctico de un proceso específico, se plantea la FAC. **Si es infinita** entonces todo **parece indicar que se podría representar como un AR**.

Una vez asignada la familia, debe determinarse el orden p, por lo tanto se define la **Función de Autocorrelaciones Parcial (FACP)**.

Función de Autocorrelaciones Parcial (FACP)

Se construye el **Sistema de Ecuaciones de Yule-Walker** llegando a la siguiente expresión,

$$\rho_1(Y) = \phi_{1p} + \phi_{2p}\rho_1(Y) + \ldots + \phi_{pp}\rho_{p-1}(Y)$$

$$\rho_2(Y) = \phi_{1p}\rho_1(Y) + \phi_2 + \ldots + \phi_{pp}\rho_{p-2}(Y)$$

$$\vdots$$

$$\rho_p(Y) = \phi_{1p}\rho_{p-1}(Y) + \phi_{2p}\rho_{p-2}(Y) + \ldots + \phi_{pp}$$

Si $p = 1$ implica un proceso $AR(1)$, con lo cual todas las autocovarianzas (y consecuentemente las autocorrelaciones) de orden mayor que 1 serían nulas. Además, se tomaría solamente una ecuación para el **Sistema de Yule-Walker**,

$$\boxed{\phi_{11}(Y) = \rho_1(Y)}$$

- La $ACP(1)$ es igual a la $AC(1)$

Si $p = 2$ implica un proceso $AR(2)$,

$$\boxed{\phi_{22} = \frac{\rho_2(Y) - \rho_1^2(Y)}{1 - \rho_1^2(Y)}}$$

A medida que p crece, se complica la resolución del sistema. Existe un algoritmo iterativo para computadoras: Durbin-Levinson (Ver Landro).

$AR(1)$

$$Y_t = \phi_1 Y_{t-1} + \epsilon_t$$

$$\rho_j\,(Y) = \phi_1 \rho_{j-1}\,(Y)$$

$$\rho_j\,(Y) = \phi_1^j = \rho_1^j$$

$$\phi_{11} = \rho_1\,(Y)$$

$$\phi_{22} = \frac{\rho_2(Y) - \rho_1^2(Y)}{1 - \rho_1^2(Y)} = \frac{\phi_1^2 - \phi_1^2}{1 - \phi_1^2} = 0$$

$$\vdots$$

$$\boxed{\phi_{pp} = 0}$$

Se cumple que ϕ_{jj} toma valores significativos para j $\leq$ p y se anula para j $>$ p.

Este criterio representa una pauta débil para sugerir el orden p de autorregresividad, a partir del análisis de la FACP.

Test de Significatividad Individual

Los verdaderos valores de los coeficientes ϕ_{jj} son desconocidos, y se dispone de sus estimadores muestrales y de un test de significatividad individual tal que:

Planteo de Hipótesis

$H_0 : \phi_{jj} = 0$

$H_1 : \phi_{jj} \neq 0$

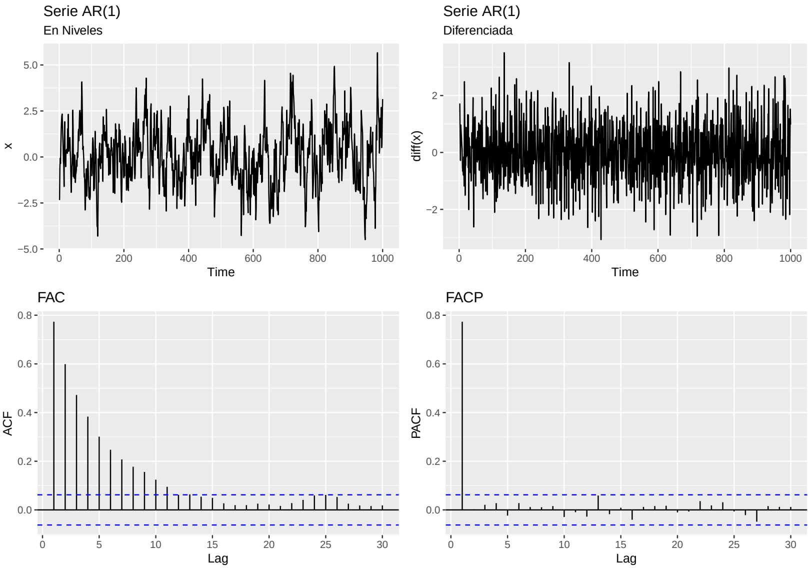
Se puede demostrar, bajo el supuesto de que H_0 es verdadera, que,

$$\hat{\phi}_{jj} \sim \; N\left(0; \tfrac{1}{n}\right) \Rightarrow \sqrt{n}\hat{\phi}_{jj} \sim N\left(0; 1\right)$$

Con un nivel de significación del 5%, H_0 se rechaza cuando,

$$\left|\hat{\phi}_{jj}\right| > \tfrac{2}{\sqrt{n}} \qquad \left(j = 1, 2, \ldots, \tfrac{n}{4}\right)$$

Gráficos


Serie AR(1)
En Niveles
x
Time
Serie AR(1)
Diferenciada
diff(x)
Time
FAC
ACF
Lag
FACP
PACF
Lag

Estimación de Parámetros $AR(p)$

Los coeficientes de la representación pueden estimarse por 3 (tres) métodos distintos,

- Yule-Walker
- Mínimos Cuadrados
- Máxima Verosimilitud

En la práctica se suele emplear *MCO* para estimar los parámetros de la representación autorregresiva.

MCO para Series Temporales

El método consiste en minimizar la suma del cuadrado de los residuos. En términos formales,

Modelo Estimado

$$\hat{Y}_t = \hat{\phi}_1 Y_{t-1} + \hat{\phi}_2 Y_{t-2} + \ldots + \hat{\phi}_p Y_{t-p}$$

Residuo

$$\hat{\epsilon}_t = Y_t - \hat{Y}_t$$

Problema de Optimización

$$min \sum_{t=p+1}^{n} \hat{\epsilon}_t^2 = \sum_{t=p+1}^{n} (Y_t - \hat{Y}_t)^2$$

$$min \sum_{t=p+1}^{n} (Y_t - \hat{\phi}_1 Y_{t-1} - \hat{\phi}_2 Y_{t-2} - \ldots - \hat{\phi}_p Y_{t-p})^2 = f(\hat{\phi}_1, \hat{\phi}_2, \ldots, \hat{\phi}_p)$$

Condición Necesaria (CPO)

$$\frac{\partial f}{\partial \hat{\phi}_1} = 0 \qquad \frac{\partial f}{\partial \hat{\phi}_2} = 0 \quad \ldots \quad \frac{\partial f}{\partial \hat{\phi}_p} = 0$$

Se demuestra que el extremo será siempre mínimo (Hessiano Positivo). Se puede garantizar que si el Ruido Blanco es Normal, entonces los estimadores minimocuadráticos son asintóticamente eficientes, convergen en distribución a la normal y son consistentes.

Máxima Verosimilitud

El método es eminentemente paramétrico, debiendo definirse una distribución de probabilidad, en este caso la normal porque los Ruido Blanco se distribuyen de dicha forma. Se busca maximizar la siguiente función de probabilidades conjunta para el conjunto de n valores observados. Se sabe que $\epsilon_t \ \sim N(0, \sigma_\epsilon)$.

Por lo tanto la función de densidad marginal es,

$$f_{\epsilon_t}(\epsilon) = \frac{1}{\sqrt{2\pi\sigma_\epsilon^2}} e^{-\frac{1}{2}\left(\frac{\epsilon_t - 0}{\sigma_\epsilon}\right)^2}$$

La Función de Verosimilitud para el conjunto de n valores observados es,

$$\mathcal{L} = \frac{1}{(2\pi\sigma_\epsilon^2)^{\frac{n}{2}}} e^{-\frac{1}{2\sigma_\epsilon^2}\sum_{t=1}^{n}\epsilon_t^2}$$

El problema de optimización resulta ser,

$$\max \mathcal{L} = max \frac{1}{(2\pi\sigma_\epsilon^2)^{\frac{n}{2}}} e^{-\frac{1}{2\sigma_\epsilon^2}\sum_{t=1}^{n}\epsilon_t^2}$$

Al resolver el problema de optimización se obtiene el valor estimado de los parámetros.

Test de Significatividad Individual

Se pretende determinar que regresores pueden permanecer en el modelo y son estadísticamente significativos. A tal fin se construye un estadístico T que converge a una distribución t y se plantea un test de significatividad individual para un modelo con variables centradas.

Planteo de Hipótesis

$H_0 : \phi_j = 0$

$H_1 : \phi_j \neq 0$

Se puede demostrar, bajo el supuesto de que H_0 es verdadera, que,

$$T = \frac{\hat{\phi}_j - \phi_j}{\sigma(\hat{\phi}_j)} \sim \ t_{n-p}$$

Regla de Decisión

$$Si \ \ |T| > 2 \Rightarrow RH_0$$

Si para algún coeficiente se aceptara la hipótesis nula, entonces sería posible eliminar el respectivo término de la representación, sin alterar el orden p del modelo. Si se tratara de un modelo con variables no centradas, se tendría un término independiente, con lo cual habría un coeficiente más a estimar, y la distribución t del estadístico tendría $n - p - 1$ grados de libertad.

Próxima Clase

* Moving Average - $MA(q)$

* Autorregresive Moving Average - $ARMA(p, q)$

Lecturas

Enders. Applied Econometric Time Series. Fourth Edition. Wiley Series. Leer **Capítulos 1 - 2**.

Hyndman. Forecasting: Principles and Practice. Second Edition. Monash University, Australia. Leer **Capítulo 2**.

Uriel. Análisis de Series Temporales. Modelos ARIMA. Leer **Capítulo 2** destinado a Procesos Estocástico

Peña. Análisis de Series Temporales. Alianza Editorial. Leer **Capítulos 1 - 3**

Fin