Cheat Sheet Series de Tiempo

Por Marcelo Moreno - Universidad Rey Juan Carlos
Como parte del Econometrics Cheat Sheet Project

Conceptos básicos

Definiciones

Serie temporal - es una sucesión de observaciones cuantitativas de un fenómeno ordenadas en el tiempo.

Hay algunas variaciones de serie temporal:

- **Datos de panel** - consiste en una serie temporal para cada observación de una sección cruzada.
- **Secciones transversales agrupadas** - combina secciones cruzadas de diferentes periodos de tiempo.

Proceso estocástico - es una secuencia de variables aleatorias que están indexadas en el tiempo.

Componentes de una serie temporal

- **Tendencia** - es el movimiento general a l/p de una serie.
- **Variaciones estacionales** - son oscilaciones periódicas que son producidas en un período igual o inferior al año, y pueden ser fácilmente identificadas en diferentes años (usualmente son el resultado de la climatología).
- **Ciclo** - son oscilaciones periódicas que se producen en un periodo mayor al año (son resultado del ciclo económico).
- **Variaciones residuales** - son movimientos que no siguen una oscilación periódica identificable (resultado de fenómenos eventuales no permanentes que pueden afectar a la variable estudiada en un momento dado).

Tipos de modelos de series temporales

- **Modelos estáticos** - la relación entre y y x's es contemporánea. Conceptualmente:
$$y_t = \beta_0 + \beta_1 x_t + u_t$$
- **Modelos de rezagos distribuidos** - la relación entre y y x's no es contemporánea. Conceptualmente:
$$y_t = \beta_0 + \beta_1 x_t + \beta_2 x_{t-1} + ... + \beta_{s+1} x_{t-s} + u_t$$
El efecto acumulado a largo plazo en y cuando Δx es:
$$\beta_1 + \beta_2 + ... + \beta_{s+1}$$
- **Modelos dinámicos** - un rezago de la variable dependiente es parte de las variables independientes (endogeneidad). Conceptualmente:
$$y_t = \beta_0 + \beta_1 y_{t-1} + ... + \beta_s y_{t-s} + u_t$$
- Combinaciones de lo anterior, como modelos de rezagos distribuidos racionales (rezagos distribuidos + dinámicos).

Supuestos y propiedades

Supuestos MCO bajo series temporales

Bajo estos supuestos, los estimadores de los parámetros MCO presentarán buenas propiedades. **Supuestos Gauss-Markov extendidos en series temporales**:

st1. **Linealidad de parámetros y dependencia débil**.
   a. y_t debe ser una función lineal de β's.
   b. El proceso estocástico $\{(x_t, y_t) : t = 1, 2, ..., T\}$ es estacionario y débilmente dependiente.

st2. **No colinealidad perfecta**.
   - No hay variables independientes que sean constantes: $Var(x_j) \neq 0$
   - No hay una relación lineal exacta entre variables independientes.

st3. **Media condicional cero y correlación cero**.
   a. No hay errores sistemáticos: $E(u_t|x_{1t}, ..., x_{kt}) = E(u_t) = 0 \rightarrow$ **exogeneidad fuerte** (a implica b).
   b. No hay variables relevantes no incluidas en el modelo: $Cov(x_{jt}, u_t) = 0$ para cualquier $j = 1, ..., k \rightarrow$ **exogeneidad débil**.

st4. **Homocedasticidad**. Var. de los residuos es igual para cualquier nivel de x: $Var(u_t|x_{1t}, ..., x_{kt}) = \sigma^2$

st5. **No autocorrelación**. Los residuos no contienen información sobre otros residuos: $Corr(u_t, u_s|x) = 0$ para cualquier $t \neq s$.

st6. **Normalidad**. Los residuos son independientes e idénticamente distribuidos (**i.i.d.**): $u \sim N(0, \sigma^2)$

st7. **Tamaño de datos**. El número de observaciones disponibles debe ser mayor a $(k + 1)$ parámetros a estimar. (YA satisfecho bajo situaciones asintóticas)

Propiedades asintóticas de MCO

Bajo los supuestos del modelo econométrico y el Teorema Central del Límite:

- De (1) a (3a): MCO es **insesgado**. $E(\hat{\beta}_j) = \beta_j$
- De (1) a (3): MCO es **consistente**. $plim(\hat{\beta}_j) = \beta_j$ (a (3b) sin (3a), exogeneidad débil, insesg. y consistente).
- De (1) a (5): **normalidad asintótica** de MCO (entonces, (6) es necesariamente satisfecho): $u \sim_a N(0, \sigma^2)$.
- De (1) a (5): **estimador insesgado de** σ^2. $E(\hat{\sigma}^2) = \sigma^2$
- De (1) a (5): MCO es MELI (Mejor Estimador Lineal Insesgado) or **eficiente**.
- De (1) a (6): contrastes de hipótesis e intervalos de confianza son fiables.

Tendencia y estacionalidad

Regresión espuria - es cuando la relación entre y y x es debida a factores que afectan a y y que tienen correlación con x, $Corr(x, u) \neq 0$. Es el **incumplimiento de st3**.

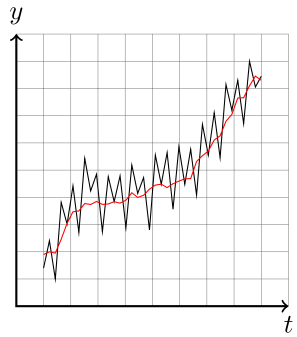
Tendencia

Dos series temporales pueden tener la misma (o contraria) tendencia, lo que lleva a altos niveles de correlación. Esto provoca una falsa apariencia de causalidad, el problema es **regresión espuria**. Dado el modelo:
$$y_t = \beta_0 + \beta_1 x_t + u_t$$
donde:
$$y_t = \alpha_0 + \alpha_1 Tendencia + v_t$$
$$x_t = \gamma_0 + \gamma_1 Tendencia + v_t$$
Añadir una tendencia al modelo puede resolver el problema:
$$y_t = \beta_0 + \beta_1 x_t + \beta_2 Tendencia + u_t$$
Una tendencia puede ser lineal o no lineal (cuadrática, cúbica, exponencial, etc.)

Otra manera, es hacer uso del **filtro Hodrick-Prescott** y extraer la tendencia (suavizado) y el componente cíclico.

Estacionalidad

Una serie temporal puede manifestar estacionalidad. Esto es, que la serie está sujeta a variaciones estacionales o patrones usualmente relacionados al clima. Por ejemplo, el PIB (negro) es usualmente mayor en verano y menor en invierno. Serie ajustada estacionalmente (rojo) en comparación.



- Este problema es **regresión espuria**. Un ajuste estacional puede solucionarlo.

Un **ajuste estacional** sencillo es crear variables estacionales binarias y añadirlas al modelo. Por ejemplo, una serie trimestral (QX_t son variables binarias):
$$y_t = \beta_0 + \beta_1 Q2_t + \beta_2 Q3_t + \beta_3 Q4_t + \beta_4 x_{1t} + ... + \beta_k x_{kt} + u_t$$
Otro método es ajustar estacionalmente (sa) las variables, y entonces, hacer la regresión con las variables ajustadas:
$$z_t = \beta_0 + \beta_1 Q2_t + \beta_2 Q3_t + \beta_3 Q4_t + v_t \rightarrow \hat{v}_t + E(z_t) = \hat{z}_t^{sa}$$
$$\hat{y}_t^{sa} = \beta_0 + \beta_1 \hat{x}_{1t}^{sa} + ... + \beta_k \hat{x}_{kt}^{sa} + u_t$$
Hay métodos mucho mejores y complejos para ajustar estacionalmente, como el **X-13ARIMA-SEATS**.

Autocorrelación

El residuo de cualquier observación, u_t, está correlacionado con el residuo de cualquier otra observación. Las observaciones no son independientes. Es el **incumplimiento de st5**.

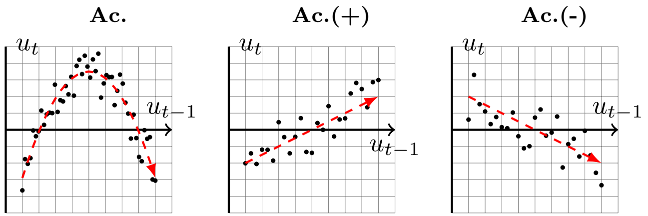
$$Corr(u_t, u_s|x) \neq 0 \text{ para cualquier } t \neq s$$

Consecuencias

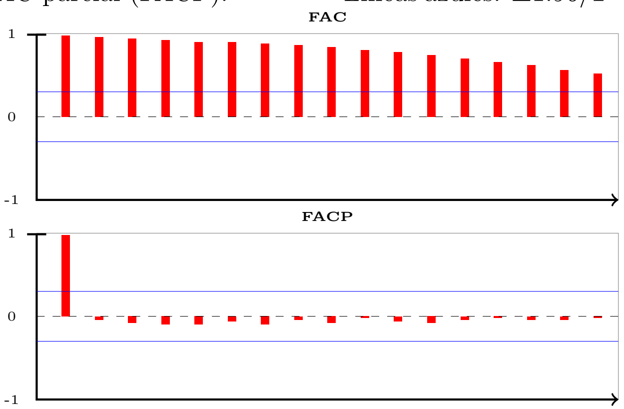
- Estimadores MCO son insesgados.
- Estimadores MCO son consistentes.
- MCO ya **no es eficiente**, pero sigue siendo ELI (Estimador Lineal Insesgado).
- La **estimación de la varianza de los estimadores es sesgada**: la construcción de intervalos de confianza y contraste de hipótesis no son fiables.

Detección

- **Gráficos de dispersión -** buscar patrones de dispersión en u_{t-1} vs. u_t.



- **Correlograma** - compuesto de la función de autocorrelación (FAC) y el FAC parcial (FACP).
  - Eje Y: correlación [-1,1].
  - Eje X: número de retardo.
  - Líneas azules: $\pm 1.96/T^{0.5}$



Conclusiones difieren entre procesos de autocorrelación.

- **Proceso MA(q).** **FAC**: sólo los primeros q coeficientes son significativos. El resto se anulan bruscamente. **FACP**: decrecimiento rápido exponencial atenuado u ondas sinusoidales.
- **Proceso AR(p).** **FAC**: decrecimiento rápido exponencial atenuado u ondas sinusoidales. **FACP**: sólo los primeros p coeficientes son significativos. El resto se anulan bruscamente.
- **Proceso ARMA(p,q).** **FAC**: los coeficientes no se anulan bruscamente y presentan un decrecimiento rápido. **FACP**: los coeficientes no se anulan bruscamente y presentan un decrecimiento rápido.

Si los coeficientes de la FAC no decaen rápidamente, hay claro indicio de falta de estacionariedad en media, lo que llevaría a tomar primeras diferencias en la serie original.

- **Formal tests -** H_0: No auto-correlation.

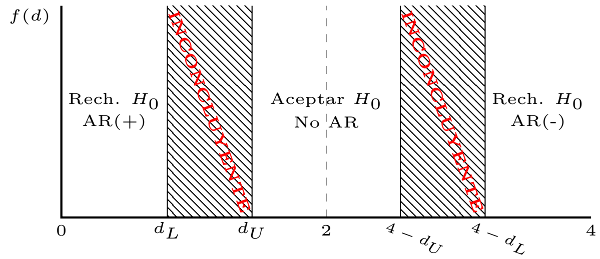
Suponiendo que u_t sigue un proceso AR(1):

$$u_t = \rho_1 u_{t-1} + \varepsilon_t$$

donde ε_t es ruido blanco.

- **Prueba t AR(1)** (regresores exógenos):
  $$t = \frac{\hat{\rho}_1}{ee(\hat{\rho}_1)} \sim t_{T-k-1,\alpha/2}$$
  - $*$ H_1: Autocorrelación de orden uno, AR(1).
- **Estadístico Durbin-Watson** (regresores exógenos y normalidad de residuos):
  $$d = \frac{\sum_{t=2}^{n}(\hat{u}_t - \hat{u}_{t-1})^2}{\sum_{t=1}^{n}\hat{u}_t^2} \approx 2(1-\hat{\rho}_1) \text{ , } 0 \leq d \leq 4$$
  - $*$ H_1: Autocorrelación de orden uno, AR(1).

$d =$	0	2	4
$\rho \approx$	-1	0	1



- **h de Durbin** (regresores endógenos):
  $$h = \hat{\rho}\sqrt{\frac{T}{1-T\times v}}$$
  donde v es la varianza estimada del coeficiente asociado a la variable endógena.
  - $*$ H_1: Autocorrelación de orden uno, AR(1).
- **Prueba Breusch-Godfrey** (regresores endógenos): puede detectar procesos MA(q) y AR(p) (ε_t ruido b.):
  - $*$ MA(q): $u_t = \varepsilon_t - \theta_1 u_{t-1} - ... - \theta_q u_{t-q}$
  - $*$ AR(p): $u_t = \rho_1 u_{t-1} + ... + \rho_p u_{t-p} + \varepsilon_t$

  Bajo H_0: No autocorrelación:
  $$T \times R^2_{\hat{u}_t} \sim_a \chi^2_q \qquad \textbf{or} \qquad T \times R^2_{\hat{u}_t} \sim_a \chi^2_p$$
  - $*$ H_1: Autocorrelación de orden q (ó p).
- **Prueba Ljung-Box Q**:
  - $*$ H_1: Existe autocorrelación.

Corrección

- Usar MCO con un estimador de la **matriz de varianzas-covarianzas robusto a la autocorrelación**, por ejemplo, la propuesta de **Newey-West**.
- Usar **Mínimos Cuadrados Generalizados**. Suponiendo $y_t = \beta_0 + \beta_1 x_t + u_t$, con $u_t = \rho u_{t-1} + \varepsilon_t$, donde $|\rho| < 1$ y ε_t es ruido blanco.
  - Si ρ **es conocido, crear un modelo cuasi-diferenciado**:
    $$y_t - \rho y_{t-1} = \beta_0(1-\rho) + \beta_1(x_t - \rho x_{t-1}) + u_t - \rho u_{t-1}$$
    $$y_t^* = \beta_0^* + \beta_1 x_t^* + u_t^*$$
    donde u_t^* es ruido blanco, y estimarlo por MCO.
  - Si ρ **es desconocido, estimarlo** -por ejemplo- **el método de Cochrane-Orcutt** (el método de Prais-Winsten también es bueno):
    1. Obtener $\hat{u}_t$ del modelo original.
    2. Estimar $\hat{u}_t = \rho\hat{u}_{t-1} + \varepsilon_t$ y obtener $\hat{\rho}$.
    3. Crear un modelo cuasi-diferenciado:
       $$y_t - \hat{\rho} y_{t-1} = \beta_0(1-\hat{\rho}) + \beta_1(x_t - \hat{\rho} x_{t-1}) + u_t - \hat{\rho} u_{t-1}$$
       $$y_t^* = \beta_0^* + \beta_1 x_t^* + u_t^*$$
       donde u_t^* es ruido blanco, y estimarlo por MCO.
    4. Obtener $\hat{u}_t^*$ y repetir desde el paso 2.
    5. El método termina cuando los parámetros estimados varían muy poco entre iteraciones.
- Si **no se arregla, buscar fuerte depend.** en la serie.

Estacionariedad y dependencia débil

Estacionariedad es estabilidad de las distribuciones conjuntas de probabilidad de un proceso a medida que este progresa el tiempo. Permite identificar correctamente las relaciones -inalteradas en el tiempo- entre variables.

Procesos estacionarios y no estacionarios

- **Proceso estacionario** (estacionariedad fuerte) - la dist. de prob. es estable en el tiempo: si se toma cualquier colección de variables aleatorias, y se mueven h periodos, la dist. conjunta de prob. debe permanecer inalterada.

- **Proceso no estacionario** - por ejemplo, una serie con tendencia, donde al menos la media cambia con el tiempo.
- **Proceso estacionario en covarianza** - es una forma más débil de estacionariedad:
  - $E(x_t)$ es constante.
  - $Var(x_t)$ es constante.
  - Para cualquier t, $h \geq 1$, la $Cov(x_t, x_{t+h})$ depende sólo de h, no de t.

Series temporales de dependencia débil

Es importante porque reemplaza el requisito de muestreo aleatorio, dando por supuesto la validez del Teorema Central del Límite (requiere estacionariedad y una forma de dependencia débil). Los procesos débilmente dependientes **también se conocen como, I(0)**.

- **Dependencia débil** - restringe cuán cercana la relación entre x_t y x_{t+h} puede ser a medida que la distancia temporal entre las series aumenta (h).

Un **proceso estacionario** $\{x_t : t = 1, 2, ..., T\}$ es débilmente dependiente cuando x_t y x_{t+h} son casi independientes a medida que h aumenta sin límite.
Un **proceso estacionario en covarianza** es débilmente dependiente si la correlación entre x_t y x_{t+h} tiende a 0 lo suficientemente rápido cuando $h \to \infty$ (no están asintóticamente correlacionados).
Algunos ejemplos de series estacionarias y débilmente dependientes:

- **Media móvil** - $\{x_t\}$ es una media móvil de orden uno MA($q = 1$):
$$x_t = e_t + \theta_1 e_{t-1}$$
donde $\{e_t : t = 0, 1, ..., T\}$ es una secuencia *i.i.d.* con media cero y varianza σ_e^2.
- **Proceso autorregresivo** - $\{x_t\}$ es un proceso autorregresivo de orden uno AR($p = 1$):
$$x_t = \rho_1 x_{t-1} + e_t$$
donde $\{e_t : t = 0, 1, ..., T\}$ es una secuencia *i.i.d.* con media cero y varianza σ_e^2.
Si $|\rho_1| < 1$, entonces $\{x_t\}$ es un proceso AR(1) que es débilmente dependiente. Es estacionario en covarianza, $Corr(x_t, x_{t-1}) = \rho_1$.
- **Proceso ARMA** - es una combinación de los dos anteriores. $\{x_t\}$ es un ARMA($p = 1$,$q = 1$):
$$x_t = e_t + \theta_1 e_{t-1} + \rho_1 x_{t-1}$$

Una serie con tendencia no puede ser estacionaria, pero puede ser débilmente dependiente (y estacionaria si la serie es filtrada de tendencia).

Series temporales de depend. fuerte

La mayoría del tiempo, las series económicas presentan dependencia fuerte (o fuerte persistencia temporal). Algunos casos especiales de procesos de raíz unitaria, I(1):

- **Paseo aleatorio** - un proceso AR(1) con $\rho_1 = 1$.
$$y_t = y_{t-1} + e_t$$
donde $\{e_t : t = 1, 2, ..., T\}$ es una secuencia *i.i.d.* con media cero y varianza σ_e^2 (la última cambia con el tiempo). El proceso no es estacionario, es persistente.
- **Paseo aleatorio con deriva** - un proceso AR(1) con $\rho_1 = 1$ y una constante.
$$y_t = \alpha + y_{t-1} + e_t$$
donde $\{e_t : t = 1, 2, ..., T\}$ es una secuencia *i.i.d.* con media cero y varianza σ_e^2.
El proceso no es estacionario, es persistente.

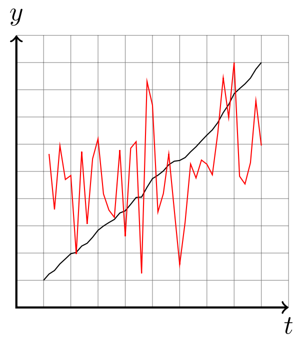
Detección de I(1)

- **Prueba aumentada de Dickey-Fuller (ADF)** - donde H_0: el proceso es raíz unitaria, I(1).
- **Prueba Kwiatkowski–Phillips–Schmidt–Shin (KPSS)** - donde H_0: el proceso no es raíz unitaria, I(0).

Transformar raíz unitaria a depend. débil

Los procesos de raíz unitaria son **integrados de orden uno**, I(1). Esto significa que **la primera diferencia del proceso es débilmente dependiente** ó I(0) (y usualmente, estacionaria). Por ejemplo, un paseo aleatorio:
$$\Delta y_t = y_t - y_{t-1} = e_t$$
donde $\{e_t\} = \{\Delta y_t\}$ es *i.i.d.*

Tomar la primera diferencia de una serie también elimina su tendencia.
Por ejemplo, una serie con tendencia (negro), y su primera diferencia (rojo).



Cuando una serie I(1) es estrictamente positiva, se suele transformar a logaritmos antes de tomar primeras diferencias. Esto es, para obtener el cambio porcentual (aprox.) de la serie:
$$\Delta log(y_t) = log(y_t) - log(y_{t-1}) \approx (y_t - y_{t-1})/y_{t-1}$$

Cointegración

Cuando **dos series son I(1), pero una combinación lineal de estas es I(0)**. Si es el caso, la regresión de una serie sobre la otra no es espuria, sino que expresa algo sobre la relación a largo plazo.
Por ejemplo: $\{x_t\}$ y $\{y_t\}$ son I(1), pero $y_t - \beta x_t = u_t$ donde $\{u_t\}$ es I(0). (β toma el nombre de parámetro cointegrador).

Heterocedasticidad en series temp.

Afecta al supuesto st4, lo que lleva a que **MCO no sea eficiente**.
Algunas pruebas que funcionan pueden ser la de Breusch-Pagan o la de White, donde H_0: No heterocedasticidad. Es **importante que no haya autocorrelación para el correcto funcionamiento de las pruebas** (así que, primero es necesario probar la existencia de autocorrelación).

ARCH

La heterocedasticidad condicional autorregresiva (ARCH), es un modelo para analizar una forma de heterocedasticidad dinámica, donde la varianza del error sigue un proceso AR(p).
Dado el modelo:
$$y_t = \beta_0 + \beta_1 z_t + u_t$$
donde, hay AR(1) y heterocedasticidad:
$$E(u_t^2 | u_{t-1}) = \alpha_0 + \alpha_1 u_{t-1}^2$$

GARCH

La heterocedasticidad condicional autorregresiva general (GARCH), es un modelo similar a ARCH, pero en este caso, la varianza del error sigue un proceso ARMA(p,q).

Predicciones

Dos tipos de predicciones:

- Del valor medio de y para un valor específico de x.
- De un valor individual de y para un valor específico de x.

Si los valores de las variables (x) se aproximan al valor medio ($\overline{x}$), la amplitud del intervalo de confianza de la predicción será menor.