

MAESTRÍA EN ECONOMÍA ECONOMETRÍA INTERMEDIA ECO743 – MÓDULO 2

Sesión 3 Autocorrelación

Docente: Juan Palomino



- 1 Naturaleza del Problema
- 2 Causas y Consecuencias de la Autocorrelación
- 3 Modelos de Autocorrelación
- 4 Detección de Autocorrelación
- 5 Estimación en presencia de Autocorrelación



1. Naturaleza del Problema



Naturaleza del Problema

❖ Hipótesis de ausencia de autocorrelación

$$cov(\varepsilon_i, \varepsilon_i | X) = 0; \forall i = 1, ..., n; i \neq j$$

Autocorrelación

Este fenómeno ocurre cuando la covarianza entre los términos de perturbación no es cero:

$$cov(\varepsilon_i, \varepsilon_j | X) \neq 0; \forall i = 1, ..., n; i \neq j$$



Naturaleza del Problema

Modelo autocorrelado:

$$y = X\beta + \varepsilon$$

con:

$$Var(\varepsilon|X) = \begin{pmatrix} \sigma^2 & \sigma_{12} & \cdots & \sigma_{1n} \\ \sigma_{21} & \sigma^2 & \cdots & \sigma_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{n1} & \sigma_{n2} & \cdots & \sigma^2 \end{pmatrix}$$

Ocurre frecuentemente en series de tiempo.



Naturaleza del Problema

- ¿Por qué ocurre la autocorrelación?
- ¿Cómo afecta las propiedades estadísticas del estimador de MCO?
- ¿Se alterarán las pruebas de hipótesis en esos casos?



2. Causas y Consecuencias de la Autocorrelación



Causas de la Autocorrelación

- Omisión de variables relevantes, cuyos valores estén autocorrelacionados entre sí.
- Forma funcional incorrecta
- Inercia de los fenómenos económicos
- En datos de corte transversal se pueden dar ciertos efectos de proximidad que provoquen autocorrelación entre las perturbaciones.
- Correlación espacial



Consecuencias de la Autocorrelación

- Los estimadores MCO son insesgados, pero menos eficientes que los MCG (es decir, no tienen varianza mínima).
- Estimación sesgada de la varianza σ^2 .
- Los intervalos de confianza y los estadísticos utilizados para resolver contrastes de hipótesis ya no son fiables. Por tanto, existe la posibilidad de extraer conclusiones erróneas.
- En presencia de autocorrelación es conveniente utilizar los estimadores MCG.



3. Modelos de Autocorrelación



Modelos de Autocorrelación

- Existen diferentes modelos de autocorrelación de acuerdo a su grado de orden:
 - Modelos Autorregresivos (AR)
 - Modelos de promedios móviles (MA)
 - Modelos ARMA



3. Modelos de Autocorrelación

3.1 Modelos Autorregresivos



Un modelo con k variables en donde el término de perturbación presenta correlación serial con las perturbaciones inmediatamente anterior:

$$Y_t = \beta_1 + \beta_2 X_{2t} + \dots + \beta_k X_{kt} + \varepsilon_t$$
$$\varepsilon_t = \phi \varepsilon_{t-1} + \mu_t \qquad |\phi| < 1$$

donde
$$E[\mu_t] = 0$$
, $Var[\mu_t] = \sigma_{\mu}^2$ y $Cov[\mu_t, \mu_{t-s}] = 0$, $\forall s \neq 0$.



Tenemos que:

Entonces:

$$\varepsilon_t = \phi^t \varepsilon_0 + \sum_{i=0}^{t-1} \phi^i \mu_{t-i}$$



Si $t \to \infty$:

$$\varepsilon_t = \sum_{i=0}^{t-1} \phi^i \mu_{t-i}$$

Entonces, la esperanza es:

$$E(\varepsilon_t) = E(\sum_{i=0}^{t-1} \phi^i \mu_{t-i})$$
$$= \sum_{i=0}^{t-1} \phi^i E(\mu_{t-i})$$
$$= 0$$

Para hallar la varianza:

$$\gamma_{0} = Var(\varepsilon_{t}) = E(\varepsilon_{t}^{2}) \qquad dado \ que \ \varepsilon_{t} = \phi^{t} \varepsilon_{0} + \sum_{i=0}^{t-1} \phi^{i} \mu_{t-i}$$

$$= Var(\phi^{t} \varepsilon_{0} + \sum_{i=0}^{t-1} \phi^{i} \mu_{t-i})$$

$$= (\phi^{t})^{2} Var(\varepsilon_{0}) + \sum_{i=0}^{t-1} (\phi^{i})^{2} Var(\mu_{t-i})$$

$$= (\phi^{t})^{2} Var(\varepsilon_{0}) + \sigma_{\mu}^{2} \sum_{i=0}^{t-1} (\phi^{i})^{2}$$

Tenemos que:

$$\sum_{i=0}^{t-1} \phi^{2i} = \phi^0 + \phi^2 + \phi^4 + \phi^6 + \dots + \phi^{2t-4} + \phi^{2t-2}$$



Luego:

$$z = 1 + \phi^{2} + \phi^{4} + \phi^{6} + \dots + \phi^{2t-4} + \phi^{2t-2}$$

$$z = 1 + \phi^{2} [1 + \phi^{2} + \dots + \phi^{2t-6} + \phi^{2t-4}]$$

$$z = 1 + \phi^{2} [1 + \phi^{2} + \dots + \phi^{2t-6} + \phi^{2t-4} + \phi^{2t-2}] - \phi^{2t-2}]$$

$$z = 1 + \phi^{2} [z] - \phi^{2t-2}]$$

$$z = 1 + \phi^{2} z - \phi^{2t}$$

$$(1 - \phi^{2})z = 1 - \phi^{2t}$$

$$z = \frac{1 - \phi^{2t}}{1 - \phi^{2}}$$



Entonces, para la función de autocovarianza:

$$\gamma_0 = (\phi^t)^2 var(\varepsilon_0) + \sigma_\mu^2 (\frac{1 - \phi^{2t}}{1 - \phi^2})$$

Si
$$t \to \infty$$

$$\gamma_0 = \frac{\sigma_\mu^2}{1 - \phi^2}$$



La autocovarianza es:

$$\begin{split} \gamma_1 &= E[\varepsilon_t \varepsilon_{t-1}] \\ &= E[(\sum_{i=0}^{\infty} \phi^i \, \mu_{t-i}) (\sum_{i=0}^{\infty} \phi^i \, \mu_{(t-1)-i})] \\ &= E[(\mu_t + \phi \mu_{t-1} + \phi^2 \mu_{t-2} + \cdots) (\mu_{t-1} + \phi \mu_{t-2} + \phi^2 \mu_{t-3} + \cdots)] \\ &= E(\phi \mu_{t-1}^2 + \phi^3 \mu_{t-2}^2 + \phi^5 \mu_{t-3}^2 + \cdots) \\ &= \phi E(\mu_{t-1}^2) + \phi^3 E(\mu_{t-2}^2) + \phi^5 E(\mu_{t-3}^2) + \cdots \\ &= \phi \sigma_{\mu}^2 + \phi^3 \sigma_{\mu}^2 + \phi^5 \sigma_{\mu}^2 + \cdots \\ &= \phi \sigma_{\mu}^2 [1 + \phi^2 + \phi^4 + \cdots] \\ \gamma_1 &= \frac{\phi}{1 - \phi^2} \sigma_{\mu}^2 \end{split}$$

La función de autocovarianza de orden superior tienen el mismo procedimiento:

$$cov(\varepsilon_t, \varepsilon_{t-s}) = \gamma_s = \frac{\phi^s}{1 - \phi^2} \sigma_{\mu}^2$$



La función de autocorrelación es:

$$corr(\varepsilon_t, \varepsilon_{t-s}) = \rho_s$$

$$\rho_{s} = \underbrace{\frac{\gamma_{s}}{\sqrt{var(\varepsilon_{t})}} \underbrace{\sqrt{var(\varepsilon_{t-s})}}_{\sqrt{\gamma_{0}}} = \frac{\gamma_{s}}{\gamma_{0}}}_{}$$

$$\rho_{s} = \frac{\gamma_{s}}{\gamma_{0}} = \frac{\frac{\phi^{s}}{1 - \phi^{2}} \sigma_{\mu}^{2}}{\frac{1}{1 - \phi^{2}} \sigma_{\mu}^{2}} = \phi^{s}$$



La matriz de varianzas-covarianzas:

$$E(\varepsilon\varepsilon') = \begin{bmatrix} \gamma_0 & \gamma_1 & \dots & \gamma_{t-1} \\ \gamma_1 & \gamma_0 & \dots & \vdots \\ \vdots & \vdots & \ddots & \vdots \\ \gamma_{t-1} & \dots & \dots & \gamma_0 \end{bmatrix} = \begin{bmatrix} \frac{\sigma_{\mu}^2}{1-\phi^2} & \frac{\phi}{1-\phi^2}\sigma_{\mu}^2 & \dots & \frac{\phi^{t-1}}{1-\phi^2}\sigma_{\mu}^2 \\ \frac{\phi}{1-\phi^2}\sigma_{\mu}^2 & \frac{\sigma_{\mu}^2}{1-\phi^2} & \dots & \frac{\phi^{t-2}}{1-\phi^2}\sigma_{\mu}^2 \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\phi^{t-1}}{1-\phi^2}\sigma_{\mu}^2 & \dots & \dots & \frac{\sigma_{\mu}^2}{1-\phi^2} \end{bmatrix}$$

$$= \frac{\sigma_{\mu}^{2}}{1 - \phi^{2}} \begin{bmatrix} 1 & \phi & \dots & \phi^{t-1} \\ \phi & 1 & \dots & \phi^{t-2} \\ \vdots & \vdots & \ddots & \vdots \\ \phi^{t-1} & \dots & \dots & 1 \end{bmatrix}$$

La matriz de varianzas-covarianzas:

$$E(\varepsilon\varepsilon') = \frac{\sigma_{\mu}^{2}}{1 - \phi^{2}} \begin{bmatrix} 1 & \rho_{1} & \dots & \rho^{t-1} \\ \rho_{1} & 1 & \dots & \rho^{t-2} \\ \vdots & \vdots & \ddots & \vdots \\ \rho^{t-1} & \rho^{t-2} & \dots & 1 \end{bmatrix}$$

$$E(\varepsilon\varepsilon') = \sigma^2 \begin{bmatrix} 1 & \rho_1 & \dots & \rho^{t-1} \\ \rho_1 & 1 & \dots & \rho^{t-2} \\ \vdots & \vdots & \ddots & \vdots \\ \rho^{t-1} & \rho^{t-2} & \dots & 1 \end{bmatrix}$$

Donde
$$\sigma^2 = \frac{\sigma_\mu^2}{1-\phi^2}$$



Sabiendo que $\rho_s = \phi^s$

$$E(\varepsilon\varepsilon') = \sigma^2 \begin{bmatrix} 1 & \phi & \dots & \phi^{t-1} \\ \phi_1 & 1 & \dots & \phi^{t-2} \\ \vdots & \vdots & \ddots & \vdots \\ \phi^{t-1} & \phi^{t-2} & \dots & 1 \end{bmatrix}$$

Donde
$$\sigma^2 = \frac{\sigma_\mu^2}{1-\phi^2}$$

3. Modelos de Autocorrelación

3.2 Modelos de Promedios Móviles



Considerar un modelo con cambio en la estructura de autocorrelación de los errores:

$$Y_t = \beta_1 + \beta_2 X_{2t} + \dots + \beta_k X_{kt} + \varepsilon_t$$
$$\varepsilon_t = \mu_t + \theta \mu_{t-1}$$

donde el parámetro θ es el peso del shock pasado sobre el futuro.



Calculamos la varianza:

$$Var(\varepsilon_t) = Var(\mu_t + \theta \mu_{t-1})$$

$$= Var(\mu_t) + Var(\theta \mu_{t-1}) + 2\theta cov(\mu_t \mu_{t-1})$$

$$= \sigma_\mu^2 + \theta^2 \sigma_\mu^2 + 0$$

$$= (1 + \theta^2)\sigma_\mu^2 = \sigma^2$$

Calculamos la covarianza:

$$\begin{split} \gamma_1 &= Cov(\varepsilon_t, \varepsilon_{t-1}) \\ &= E[(\mu_t + \theta \mu_{t-1})(\mu_{t-1} + \theta \mu_{t-2})] \\ &= E[(\mu_t \mu_{t-1} + \theta \mu_{t-1}^2 + \theta \mu_t \mu_{t-2} + \theta^2 \mu_{t-1} \mu_{t-2})] \\ &= E[\mu_t \mu_{t-1}] + \theta E[\mu_{t-1}^2] + \theta E[\mu_t \mu_{t-2}] + \theta^2 E[\mu_{t-1} \mu_{t-2}] \\ &= \theta \sigma_{\mu}^2 \end{split}$$

Calculamos la correlación de orden 1:

$$\rho_{1} = corr(\varepsilon_{t}, \varepsilon_{t-1})$$

$$\rho_{1} = \frac{\gamma_{1}}{\sqrt{var(\varepsilon_{t})} \sqrt{var(\varepsilon_{t-1})}} \sqrt{\frac{1+\theta^{2}\sigma_{\mu}^{2}}{\sqrt{1+\theta^{2}\sigma_{\mu}^{2}}}} \sqrt{\frac{1+\theta^{2}\sigma_{\mu}^{2}}{1+\theta^{2}\sigma_{\mu}^{2}}}$$

$$\rho_{1} = \frac{\theta \sigma_{\mu}^{2}}{1+\theta^{2}}$$

$$\rho_{1} = \frac{\theta}{1+\theta^{2}}$$

Calculamos la covarianza de orden 2:

La correlación de orden 2:

$$\gamma_2 = Cov(\varepsilon_t, \varepsilon_{t-2})$$

$$= E[(\mu_t + \theta \mu_{t-1})(\mu_{t-2} + \theta \mu_{t-3})]$$

$$= 0$$

$$\rho_2 = corr(\varepsilon_t, \varepsilon_{t-2}) = 0$$

Para otras covarianzas mayor igual al orden 2, todas serán cero. La matriz de varianza-covarianza es:

$$Var(\varepsilon) = \begin{bmatrix} \sigma^2 & \frac{\theta}{1+\theta^2} & 0 & \dots & 0\\ \frac{\theta}{1+\theta^2} & \sigma^2 & \frac{\theta}{1+\theta^2} & \dots & 0\\ 0 & \frac{\theta}{1+\theta^2} & \sigma^2 & \dots & 0\\ \vdots & \vdots & \vdots & \ddots & \vdots\\ 0 & 0 & 0 & \dots & \sigma^2 \end{bmatrix}$$

4. Detección de Autocorrelación



Detección de la Autocorrelación

- Disponemos de varias herramientas de diagnóstico que nos pueden ayudar a detectar la autocorrelación.
- Se clasifican en:
 - Procedimientos informales
 - Procedimientos formales



4. Detección de Autocorrelación

4.1 Procedimientos Informales



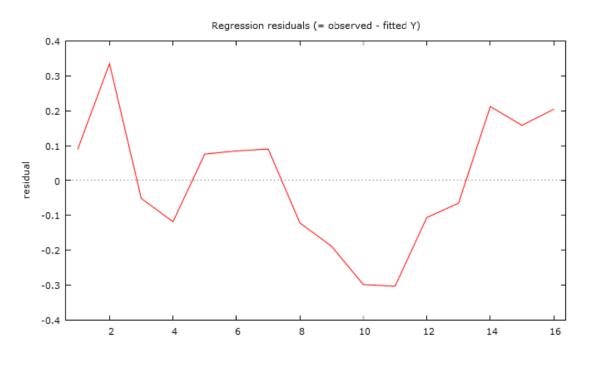
Procedimientos Informales

Son subjetivos y se basan en la representación gráfica de los residuos:

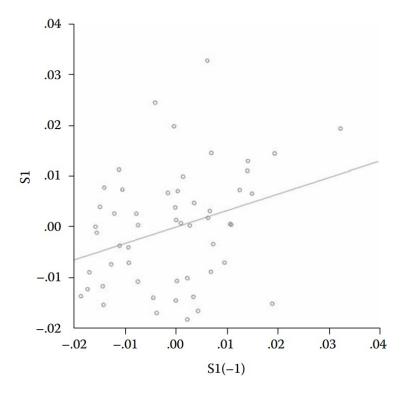
- Se calculan los residuos estimados el modelo por MCO y
 - 1. Representarlos como una serie temporal, esto es, los $\hat{\varepsilon}_t$ en función de t, o bien
 - 2. Representarlos frente a los residuos retardados en el orden de autocorrelación del que se sospecha. Por ejemplo, si creemos que puede haber autocorrelación de orden 1, representamos los $\hat{\mu}_t$ frente a los $\hat{\mu}_{t-1}$.
 - 3. Si el comportamiento es aleatorio, no hay evidencia de existencia de autocorrelación. Si por el contrario, se aprecia un patrón sistemático, nos hará sospechar que hay autocorrelación.



Procedimientos Informales



(a) Residuos frente al tiempo



(b) Residuos presente vs rezagados

Figura 1. Procedimiento Informal



Procedimientos Informales

Se pueden utilizar gráficos descriptivos de series temporales para las autocorrelaciones y las autocorrelaciones parciales, de modo que si los valores salen fuera de los límites de confianza podemos sospechar que hay autocorrelación. Estos gráficos además indican el orden en que la autocorrelación se presenta.



Procedimientos Informales

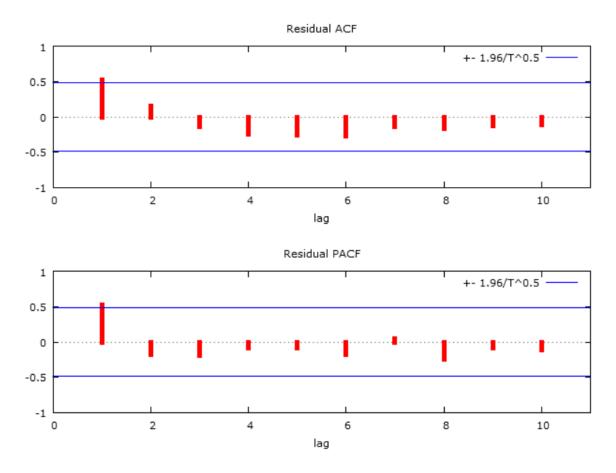


Figura 2. Gráficos PAC y FACP



4. Detección de Autocorrelación

4.2 Procedimientos Formales



Procedimientos Formales

- Son objetivos ya que se trata de contrastes de hipótesis.
- Existe diversos test de autocorrelación, cada uno de ellos con un propósito distinto:
 - Test de Durbin-Watson
 - Test de Breusch-Godfrey



Test de Durbin y Watson

• Supongamos un modelo donde se sospecha que ε_t sigue un proceso AR(1)

$$H_0$$
: $\phi = 0$ (No Autocorrelación)
 H_1 : $\phi \neq 0$

• El test se basa en el cálculo del estadístico de Durbin-Watson (DW) a partir de los residuos de la regresión del modelo estimado por MCO:

$$DW = \frac{\sum_{t=2}^{n} (\hat{\varepsilon}_{t} - \hat{\varepsilon}_{t-1})^{2}}{\sum_{t=1}^{n} \hat{\varepsilon}_{t}^{2}}$$

Se puede comprobar que DW es:

$$DW \approx 2\left(1 - \frac{\sum_{t=2}^{n} \hat{\varepsilon}_{t} \hat{\varepsilon}_{t-1}}{\sum_{t=1}^{n} \hat{\varepsilon}_{t}^{2}}\right)$$
$$\approx 2\left(1 - \hat{\phi}\right)$$

Donde $\hat{\phi}$ es aproximadamente el coeficiente estimado por MCO de la regresión sin intercepto entre $\hat{\varepsilon}_t$ y $\hat{\varepsilon}_{t-1}$.

Test de Durbin y Watson

Tenemos que:

- Si $\hat{\phi} \approx 0 \Rightarrow DW \approx 2$. No hay autocorrelación
- Si $\hat{\phi} \approx 1 \Rightarrow DW \approx 0$. Existe autocorrelación positiva
- Si $\hat{\phi} \approx -1 \Longrightarrow DW \approx 4$. Existe autocorrelación negativa

El valor calculado de DW se compara con los valores críticos de las tablas de Durbin y Watson, los cuales dependen del número de observaciones y del número de variables en el modelo.



Test de Durbin y Watson

Limitaciones:

- Solo se prueba en autocorrelación del tipo AR(1).
- No es útil para probar hipótesis de autocorrelación de orden superior



Test de Breusch-Godfrey

Este test permite probar si existe autocorrelación de cualquier orden.

*H*₀: *No Autocorrelación*

 H_1 : Autocorrelación AR(q) en los errores



Test de Breusch-Godfrey

❖ Pasos:

- Estimar el modelo $y = X\beta + \varepsilon$ por MCO ignorando el problema de autocorrelación y calcular los residuos.
- Hacer una regresión de los residuos de MCO contra sus valores rezagados q periodos atrás y contra las variables X sin intercepto:

$$\hat{\varepsilon}_t = \phi_1 \hat{\varepsilon}_{t-1} + \phi_2 \hat{\varepsilon}_{t-2} + \dots + \phi_q \hat{\varepsilon}_{t-q} + \alpha_2 X_{2t} + \dots + \alpha_k X_{kt} + \mu_t$$

- Calcular $(N-q)R_{aux}^2$ donde R_{aux}^2 es del paso 2 y N es el número total de observaciones del paso 1.
- Si $(N-q)R_{aux}^2 > \chi_{1-\alpha}^2(q)$ se rechaza la hipótesis nula de no autocorrelación.



5. Estimación en presencia de Autocorrelación

5.1 Estimación de β por MCG



1. Estimación de β por MCG

Todo dependerá de cómo se define a la matriz Ω . Tenemos que los errores son del tipo AR(1):

$$\varepsilon_t = \phi \varepsilon_{t-1} + \mu_t \qquad \mu_t \sim iid(0, \sigma_\mu^2)$$

Entonces, podemos obtener a la matriz de transformación C que garantiza que $C'C = \Omega^{-1}$ en:

$$\hat{\beta}_{MCG} = (X'\Omega^{-1}X)^{-1}X'\Omega^{-1}y$$

Siendo esta matriz:

$$C = \begin{bmatrix} \sqrt{(1 - \phi^2)} & 0 & 0 & \dots & 0 & 0 \\ -\phi & 1 & 0 & \dots & 0 & 0 \\ 0 & -\phi & 1 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & -\phi & 1 \end{bmatrix}$$

Esta es la matriz de transformación de Prais-Winsten por la corrección $\sqrt{(1-\phi^2)}$ que se hace en la primera observación.



1. Estimación de β por MCG

Luego, para un modelo bivariado con error tipo AR(1) como:

$$y_t = \beta_1 + \beta_2 X_t + \varepsilon_t$$

$$\varepsilon_t = \phi \varepsilon_{t-1} + \mu_t$$

Rezagando el error tenemos:

$$\varepsilon_{t-1} = y_{t-1} - \beta_1 - \beta_2 X_{t-1}$$

Reemplazando en ε_t , se obtiene el modelo transformado a partir de t=2 en adelante:

$$(y_t - \beta_1 - \beta_2 X_t) = \phi(y_{t-1} - \beta_1 - \beta_2 X_{t-1}) + \mu_t$$

$$(y_t - \phi y_{t-1}) = \beta_1 (1 - \phi) + \beta_2 (X_t - \phi X_{t-1}) + \mu_t$$

En t = 1 es:

$$\sqrt{1 - \phi^2} y_t = \sqrt{1 - \phi^2} \beta_1 + \sqrt{1 - \phi^2} \beta_2 X_t + \mu_t$$

1. Estimación de β por MCG

Una forma de estimar el modelo es mediante el procedimiento de estimación iteractivo de Cochrane-Orcutt, con el cual se omite la primera observación:

- 1. Asegurar un $\hat{\phi}$ inicial: digamos $\hat{\phi}^0 = 1 \frac{DW}{2}$
- 2. Utilizar este valor en (1) y obtener, por MCO, $\hat{\beta}_1^0$ y $\hat{\beta}_2^0$ de la regresión

$$(y_t - \hat{\phi}^0 y_{t-1}) = \beta_1 (1 - \hat{\phi}^0) + \beta_2 (X_t - \hat{\phi}^0 X_{t-1}) + \mu_t$$

3. Usar $\hat{\beta}_1^0$ y $\hat{\beta}_2^0$ del paso 2 en la ecuación:

$$(y_t - \hat{\beta}_1^0 - \hat{\beta}_2^0 X_t) = \phi(y_{t-1} - \hat{\beta}_1^0 - \hat{\beta}_2^0 X_{t-1}) + \xi_t$$

Estimar este modelo por MCO y llame $\hat{\phi}^1$ al estimador.

- 4. Usando $\hat{\phi}^1$, repetir los pasos 2, 3 y 4.
- 5. Continuar hasta que $\hat{\phi}$ converja a un valor.



5. Estimación en presencia de Autocorrelación

5.2 Estimación por MCO con corrección de matriz var-cov



2. Estimación por MCO con corrección en la matriz de var-cov

- Newey-West propone un estimador de las varianzas y covarianzas que es consistente con heterocedasticidad y autocorrelación.
- El estimador de la matriz Ω es:

$$\widehat{\Omega} = \begin{bmatrix} \widehat{\gamma}_0 & \widehat{\gamma}_1 & \widehat{\gamma}_2 & \cdots & \widehat{\gamma}_{n-1} \\ \widehat{\gamma}_1 & \widehat{\gamma}_0 & \widehat{\gamma}_1 & \cdots & \widehat{\gamma}_{n-2} \\ \widehat{\gamma}_2 & \widehat{\gamma}_1 & \widehat{\gamma}_0 & \cdots & \widehat{\gamma}_{n-3} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \widehat{\gamma}_{n-1} & \widehat{\gamma}_{n-2} & \widehat{\gamma}_{n-3} & \cdots & \widehat{\gamma}_0 \end{bmatrix}$$

Donde $\hat{\gamma}_j = (1 - \frac{j}{q+1}) \frac{1}{n} \sum_{t=j+1}^n \hat{\varepsilon}_t \hat{\varepsilon}_{t-j}$ si $0 \le j \le q$ y $\hat{\gamma}_j = 0$ si j > q, donde q son los rezagos que se están utilizando en la estimación y $\hat{\varepsilon}_t$ son los residuos MCO.



2. Estimación por MCO con corrección en la matriz de var-cov

La matriz de varianzas y covarianzas del estimador MCO consistentes con heterocedasticidad asintótica es:

$$AVar(\hat{\beta}_{MCO}) = (X'X)^{-1}X'\widehat{\Omega}X(X'X)^{-1}$$



