ECONOMETRÍA
APLICADA UTILIZANDO R.

PAPIME PE302513 LIBRO ELECTRÓNICO Y COMPLEMENTOS DIDÁCTICOS EN MEDIOS COMPUTACIONALES, PARA EL FORTALECIMIENTO DE LA ENSEÑANZA DE LA ECONOMETRÍA

Capítulo 15.

Econometría Espacial: Aplicaciones con R

Miguel Ángel Mendoza González

Objetivo

Introducir al alumno a la subdisciplina de la econometría espacial, abordando sus principales tópicos.

Introducción

En la actualidad, cada vez se realiza una mayor difusión e información georeferenciada, la cual, aunada al desarrollo de Sistemas de Información Geográfica (GIS por sus siglas en inglés), han permitido el desarrollo de una novel subdisciplina conocida como econometría espacial.

En este capítulo, se abordarán los principales temas concernientes a la econometría espacial, tales como:

- Vecindad
- Dependencia espacial
- Estadísticos de dependencia espacial
- Regresión espacial
- Selección de modelos espaciales

La econometría espacial fue definida a principios de los años setenta por Jean Paelinck como el creciente cuerpo de la literatura en ciencia regional que trata primordialmente con la estimación y prueba de problemas encontrados en la implantación de modelos econométricos multirregionales.

Luc Anselin (1988) uno de los pioneros y grandes impulsores de la econometría espacial considera que el campo de esta disciplina esta formado por:

"…aquellos métodos y técnicas que, sustentados en una representación formal de la estructura de la dependencia y heterogeneidad espacial, provee el medio para llevar a cabo la adecuada especificación, estimación, prueba de hipótesis y predicción para modelos en la ciencia regional."

Los métodos desarrollados por la econometría espacial permiten atender problemas de violación a los supuestos del modelo de regresión, que no es posible resolverlos en el marco de los modelos econométricos estándar. Estos problemas son típicos en los datos espaciales y se refieren a:

1) Dependencia espacial entre observaciones: Correlación espacial.

2) Heterogeneidad espacial entre observaciones: Heteroscedasticidad espacial.

El caso al que se le ha dedicado mayor atención es al primero, debido a que el segundo ha podido estudiarse en el marco de modelos de panel y otras técnicas similares en donde la heterocedasticidad y el cambio estructural juegan un papel relevante.

Vecindad Espacial

Usualmente cuando el economista maneja series económicas, sociales o ambientales lo hace desde una perspectiva en la cual toma como dadas las coordenadas de localización geográfica de las variables.

Obviar el contexto espacial significa una perdida importante de información, puesto que, los datos generalmente presentan algún tipo de dependencia o autocorrelación espacial, la cual puede definirse como la existencia de una relación funcional entre lo que ocurre en un punto del espacio y lo que sucede en otro lugar, lo cual se explica fundamentalmente por razones de interacción humana con su entorno físico-ambiental.

La dependencia espacial implicaría que al tomar en consideración una variable, para diferentes localidades, esperaríamos características más similares en localidades vecinas, que en aquéllas separadas por grandes distancias.

La dependencia espacial puede ser positiva o negativa, de ser positiva la presencia de un atributo en una localidad se extendería a las regiones vecinas y, en caso de ser negativa, obstaculizaría su presencia en sus vecindades.

Los datos espaciales se pueden clasificar de acuerdo con el objeto espacial al que se refieren y al nivel de medida de las variables. Dicha clasificación puede ilustrarse matricialmente como en la figura 1

Figura 1 Matriz de datos espaciales

$$\begin{array}{cc}
\begin{bmatrix}
z_1(1) & z_2(1) & \dots & z_k(1) & \mathbf{s}(1) \\
z_1(2) & z_2(2) & \dots & z_k(2) & \mathbf{s}(2) \\
. & . & . & . & . \\
. & . & . & . & . \\
. & . & . & . & . \\
z_1(n) & z_2(n) & \dots & z_k(n) & \mathbf{s}(n)
\end{bmatrix}
&
\begin{matrix}
\text{Caso 1} \\
\text{Caso 2} \\
. \\
. \\
. \\
\text{Caso n}
\end{matrix}
\end{array}$$

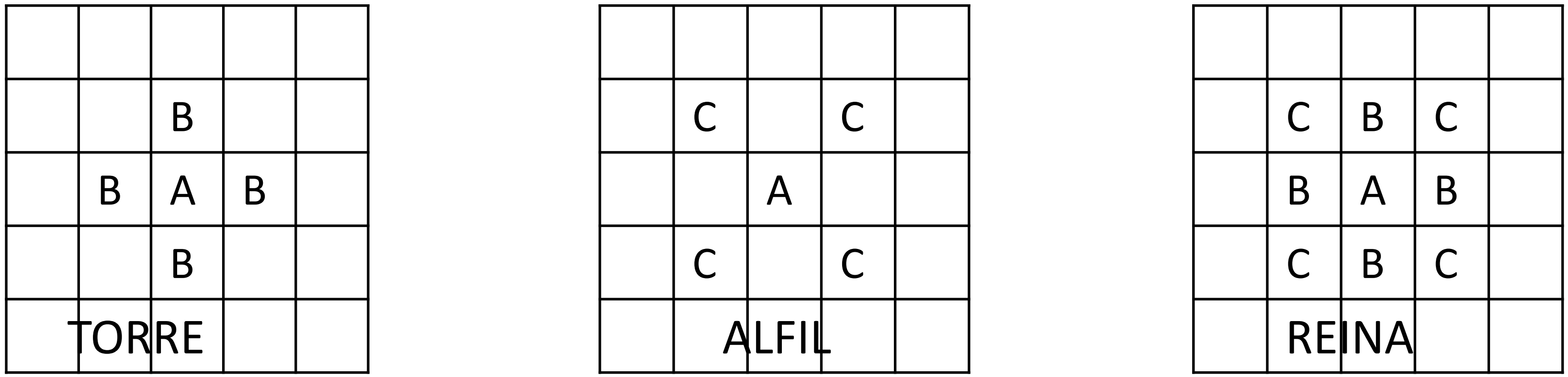
Donde tenemos k variables {z1, z2, …,zk} medidas en la localización s(i) donde i=1,2,..n.

Si incorporamos el factor de temporalidad, podríamos tener una matriz de este tipo para cada período del tiempo. Las relaciones entre las variables y localizaciones clasificadas en la matriz de datos pueden establecerse a través de conectividad o vecindad.

Matriz de vecindad por contigüidad

La noción de vecindad se puede establecer de forma binaria; en tal caso, sí dos unidades espaciales tienen una frontera común se les asigna un uno, en caso contrario se le asigna un cero. Bajo esta sencilla idea, una variable particular podría referenciarse en un mapa, a partir del cual es posible establecer sus fronteras y, en consecuencia, identificar sus vecindades. Luc Anselin (1988) plantea diferentes medidas de vecindad, las cuales se asemejan a un tablero de ajedrez y que podemos apreciar en la figura 2.

Figura 2 Diferentes Vecindades



La vecindad entre puntos también puede ser de orden superior, sí se consideran series de bandas concéntricas alrededor de la localidad bajo consideración.

En la figura 3 y considerando vecindad tipo torre, las celdas C y D son contiguas de segundo orden a la celda A, y son contiguas de primer orden a B.

Figura 3 Vecindad de Orden Superior

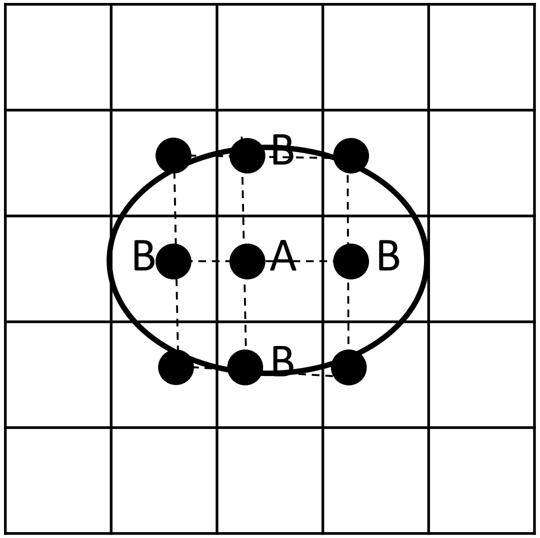
		D		
	C	B	C	
D	B	A	B	D
	C	B	C	
		D		

Matriz de vecindad por distancia

La matriz de vecindades binarias es limitada, ya que únicamente considera la vecindad física, por lo cual no contabiliza la posibilidad de interacción entre regiones alejadas. Por ello, han sido propuestas otras medidas de vecindad alternativas, sustentadas en distancias de diferente tipo y cuya matriz, **W**, es conocida como la matriz de pesos o contactos espaciales.

Anselin plantea que, en caso de que la unidad espacial sea un sistema urbano, la vecindad puede ser obtenida de la trayectoria más corta en una red o gráfica formada por una conexión de puntos. Por ejemplo, en la figura 4, la distancia más corta entre los puntos es representada por la línea punteada y la vecindad por el círculo que conecta los puntos y tiene como centroide a la localidad A.

Figura 4 Vecindad por Distancia más Corta



Considerando los centroides como punto de referencia para medir las distancias geográficas, Fotheringham, Brunsdon y Charlton (2000) proponen las siguientes medidas de distancias:

I. Localización en el plano cartesiano

En un sistema cartesiano, la distancia se mide por el teorema de Pitágoras y la localización es por medio de las coordenadas geográficas: *latitud* y *longitud*.

- *Distancia Euclidiana*

Con base a las coordenadas de latitud (*x)* y la *longitud* (*y*), la distancia entre los centroides de las localidades *i* y *j*:

$$d_{i,j} = \sqrt{\left(x_i - x_j\right)^2 (y_i - y_j)^2}$$

La distancia euclidiana entre dos localidades *i* y *j* con coordenadas ($x_{i,1}$, $x_{i,2}$), ($x_{j,1}$, $x_{j,2}$), se puede escribir también como:

$$d_E(i,j) = \left[\sum_{k=1}^{2} \left(x_{ik} - x_{jk}\right)^2\right]^{1/2}$$

La distancia puede ser generalizada a *m* dimensiones.

$$d_E(i,j) = \left[\sum_{k=1}^{m} \left(x_{ik} - x_{jk}\right)^2\right]^{1/2}$$

- *Métrica de Minkowski*

En el caso de que *p=2* es la distancia euclidiana, si *p=1* es la distancia conocida como *Manhattan* o distancia *taxicab*.

$$d_E(i,j) = \left[\sum_{k=1}^{m} \left|x_{ik} - x_{jk}\right|^p\right]^{\frac{1}{p}}$$

II. Localización en el globo o superficie de la tierra

En el caso de considerar la superficie de la tierra en lugar del plano cartesiano, se necesita de los cálculos geométricos:

- *Trigonometría esférica* (curvatura de la tierra)

$$S_{ij} = R.arcos[\cos(90^0 - \Phi i)\cos(90^0 - \Phi j) + \mathrm{sen}(90^0 - \Phi i)\mathrm{sen}(90^0 - \Phi j)\cos(\lambda_j - \lambda_i)]$$

donde R es el radio de la tierra, arcoseno (arcos), coseno (cos), seno (sen), la latitud y longitud de la locación i son ($\Phi i, \lambda_i$)

- *Mercator* (proyección a una forma cilíndrica)

$x = R\lambda$

$y = R\ln(\tan(\pi/4 + \Phi/2)$

donde *R* es el radio de la tierra, *ln* es el logaritmo natural, tangente (tan), *Φ* es la latitud y *λ* es la longitud.

- *Lambert* (proyección a un área cilíndrica)

$x = R\lambda$

$y = R sen\Phi$

Estadísticos de Dependencia Espacial

Para la medición de dependencia espacial se han propuesto numerosos estadísticos, uno de los más utilizados es el índice de Moran (1948), que se define en la fórmula siguiente:

$$\boldsymbol{I} = \frac{R}{\sum_i \sum_j w_{ij}} \frac{\sum_i \sum_j w_{ij}(x_i - \bar{x})(x_j - \bar{x})}{\sum_i (x_i - \bar{x})^2}$$

Donde x_i es la variable cuantitativa en la región i, $\bar{x}$ es su media muestral, w_{ij} son los pesos de la matriz **W**, R es el tamaño de muestra (Regiones); y

$$E(I) = \frac{-1}{R-1}$$

$$V(\boldsymbol{I}) = \frac{RS_4 - S_3 S_1(1-2R)}{(R-1)(R-2)(R-3)(\sum_i \sum_j w_{ij})^2}$$

$$S_1 = \frac{1}{2}\sum_i \sum_j (w_{ij} + w_{ji})^2$$

$$S_2 = \sum_i \left(\sum_j w_{ij} + \sum_j w_{ji}\right)^2$$

$$S_3 = \frac{R^1 \sum_i (x_i - \bar{x})^4}{(R^1 \sum_i (x_i - \bar{x})^2)^2}$$

$$S_4 = (R^2 - R + 3)S_1 - RS_2 + 3\left(\sum_i \sum_j w_{ij}\right)^2$$

Bajo la hipótesis nula de no autocorrelación, el estadístico de Moran es asintóticamente normal:

$$I^* = \frac{I - E(I)}{\sqrt{V(I)}}$$

Un valor positivo (negativo) significativo del índice Z(I) llevará al rechazo de la hipótesis nula de no autocorrelación espacial y a la aceptación de autocorrelación espacial positiva (negativa).

Es posible graficar la información del índice en un diagrama de dispersión de Moran. Dicho diagrama, presenta en el eje horizontal a la variable x normalizada y en el eje vertical a la variable multiplicada por la matriz de pesos W, lo cual da lugar al retardo espacial de dicha variable. La visualización de un patrón aleatorio en la gráfica brinda evidencia de la ausencia de autocorrelación espacial.

Dependencia Espacial

La dependencia espacial es multidireccional, es decir, una región puede estar afectada no solamente por otra región contigua o vecina sino por otras que la rodean, al igual que ella puede afectar a las otras.

Por tal motivo, es posible utilizar la matriz W de efectos espaciales como operador de rezago espacial, que se puede leer como una media ponderada de los valores vecinos y se define como:

$$\boldsymbol{W}Y = \sum_{j=1}^{N} w_{ij}\, y_j$$

donde y_j es el valor que toma el atributo medido en la vecindad j, w_{ij} es un ponderador cuya suma es la unidad.

Indicador Local de Asociación Espacial (LISA)

*En pr*ocesos en los cuales existen patrones de agrupación local o clúster, el índice de Moran no los puede detectar, dado que sólo evalúa la dependencia global de todas las regiones. Como alternativa se han propuesto estadísticos locales, tal es el caso del índice local de Moran que se calcula en cada región o localidad y su definición es la siguiente:

$$\boldsymbol{I}_i = \frac{z_i}{\sum_i z_i^2/N_j}\sum_i w_{ij} z_j$$

donde z_j es el valor de la variable correspondiente en la región i, N_j es el conjunto de regiones vecinas a i. Un valor elevado, positivo (negativo) y significativo del estadístico da lugar a la existencia de un clúster alrededor de la región i de valores similares elevados (bajos). Con base en el índice local, $\boldsymbol{I}_i$, es posible encontrar su contribución al índice global, $\boldsymbol{I}$, y detectar sus valores extremos lo cual lo convierte en un LISA.

Modelos Espaciales

En el dado caso de que los estadísticos de dependencia espacial indiquen asociación de este tipo, será necesario un modelo de regresión espacial que tome en cuenta dicha dependencia.

El modelo general planteado es:

$$y_i = \rho W_1 y_i + \beta X_i + \theta W_2 X_i + \varepsilon_i$$

$$\varepsilon_i = \lambda W_3 \varepsilon_i + u_i$$

con $u_i \sim N(0, \Omega)$ siendo los elementos diagonales de $\Omega_{ij} = h_i(z\alpha)$ con $h_i > 0$.
donde y_i es el vector de la variable endógena, X_i es una matriz de variables exógenas y el término de error ε_i que incorpora una estructura de dependencia espacial autorregresiva, W_1, W_2 y W_3 son matrices de pesos espaciales.

1) Modelo de regresión clásico sin efectos espaciales: $\rho = 0, \lambda = 0, \theta = 0$

$$y_i = \beta X_i + \varepsilon_i$$
$$\varepsilon_i = u_i$$

2) Modelo autorregresivo: $\rho \neq 0, \lambda = 0, \theta = 0$

$$y_i = \rho W_1 y_i + \beta X_i + \varepsilon_i$$
$$\varepsilon_i = u_i$$

3) Modelo de error espacial autorregresivo: $\rho = 0, \lambda \neq 0, \theta = 0$

$$y_i = \beta X_i + \varepsilon_i$$
$$\varepsilon_i = \lambda W_3 \varepsilon_i + u_i$$

4) Modelo Durbin Espacial: $\rho \neq 0, \lambda = 0, \theta \neq 0$

La estrategia de Durbin sobre el factor común se aplica al modelo de Rezago Espacial, como:

$$y_i = \rho W_1 y_i + \beta X_i + \theta W_1 X_i + u_i$$

5) Modelo mixto autorregresivo espacial con errores espaciales autorregresivos (SARMA): $\rho \neq 0, \lambda \neq 0, \theta = 0$

$$y_i = \rho W_1 y_i + \beta X_i + (I - \lambda W_3)^{-1} u_i$$

6) Modelo Error Durbin Espacial: $\rho = 0, \lambda \neq 0, \theta \neq 0$

La estrategia de Durbin sobre el factor común se aplica al modelo de Error Espacial, con los siguientes pasos:

a) De la primera ecuación despejar los errores y sustituir en la segunda

$$y_i - \beta X_i = \lambda W_3(y_i - \beta X_i) + u_i$$

b) Al despejar y_i, se obtiene:

$$y_i = \lambda W_3 y_i + \beta X_i + \theta W_3 X_i + u_i$$

donde: $\theta = -\beta\lambda$

Métodos de Estimación

La estimación del modelo espacial se realiza a través del método de máxima verosimilitud en concordancia con el modelo espacial específico que se seleccione.

De acuerdo a Lesage y Pace (2009) la estrategia de estimación de los modelos Durbin Espacial (SDM) y Rezago Espacial (SAR) por sus siglas en inglés, es la siguiente:

<u>*El modelo SDM*</u>

$$y = \rho W y + \alpha i_n + X\beta + WX\theta + \varepsilon$$
$$\varepsilon \sim N(0, \sigma^2 I_n)$$

donde 0 representa un vector de ceros de n×1 y i_n un vector de unos n×1 asociados con el término de la constante α. Este modelo puede ser escrito de forma compacta con $Z = [i_n \;\; X \;\; WX]$ y $\delta = [\alpha\ \beta\ \theta]'$ para definir el caso del modelo SAR cuando $Z = [i_n \;\; X \;\;]$ y $\delta = [\alpha\ \beta\ ]'$

- *El modelo SAR*

$$y = \rho W y + Z\delta + \varepsilon$$
$$\varepsilon \sim N(0, \sigma^2 I_n)$$

Si el valor del parámetro rho (ρ) fuera conocido por decir ρ^*, el modelo se puede escribir como:

$$y - \rho^* W y = Z\delta + \varepsilon$$

Por lo que se puede resolver el problema de estimación de δ como:

$$(I_n - \rho^* W) y = Z\delta + \varepsilon$$
$$\hat{\delta} = (Z'Z)^{-1} Z' (I_n - \rho^* W) y$$

También se encuentra la estimación de la varianza

$$\hat{\sigma}^2 = n^{-1}e(\rho^*)'e(\rho^*) \text{ donde } e(\rho^*) = y - \rho^* W y - Z\hat{\delta}$$

donde e son los errores de estimación.

Lo anterior indica que el método de estimación se concentra en el log de verosimilitud con respecto a los parámetros de β y σ^2 y por tanto la maximización de la verosimilitud se convierte a un problema de optimización univariante en el parámetro ρ.

Propuesta para estimar al mismo tiempo todo:

- Estimar la función de log-verosimilitud concentrada con respecto a los parámetros $\beta \; y \; \sigma^2$, para obtener soluciones muy cercanas a las condiciones de primer orden junto con rho.

- Sustituir las estimaciones de $\beta \; y \; \sigma^2$, por lo que la función de log-verosimilitud depende de la muestra de datos y el parámetro desconocido rho.

- En este punto la función de log-verosimilitud esta concentrada con respecto rho, por lo que se usa para encontrar la estimación de máxima verosimilitud $\hat{\rho}$ que será usada a su vez en la estimación de $\hat{\beta}(\hat{\rho}) \; y \; \hat{\sigma}^2(\hat{\rho})$ en la siguiente vuelta.

La función de verosimilitud para SDM y SAR toma la forma siguiente

$$lnL = -\left(\frac{n}{2}\right)\ln(\pi\sigma^2) + ln|I_n - \rho W| - \frac{e'e}{2\sigma^2}$$
$$e = y - \rho W y - Z\hat{\delta}$$
$$\rho \in (\min(\omega)^{-1}, \max(\omega)^{-1})$$

donde ω es el vector de $n \times 1$ raíces características de la matriz W. Dado que la matriz siempre esta construida para tener raíces máximas de 1, entonces $\rho \in (\min(\omega)^{-1}, 1)$ el cual es un subconjunto del empleado en la práctica $\rho \in [0,1)$.

La función de log-verosimilitud concentrada en los valores de ln L(ρ) se escribe como

$$lnL(\rho)=\kappa+ln|I\_n-\rho W|-(n/2)\ln(S(\rho))$$

$$S(\rho) = e(\rho)'e(\rho) = e_0'e_0 - 2\rho e_0'e_d + \rho^2 e_d'e_d$$
$$e(\rho) = e_0 - \rho e_d$$
$$e_0 = y - Z\delta_0$$
$$e_d = Wy - Z\delta_d$$
$$\delta_0 = (Z'Z)^{-1}Z'y$$
$$\delta_d = (Z'Z)^{-1}Z'Wy$$

La optimización es con respecto al parámetro rho y una vez estimado $\hat{\rho}$ con máxima verosimilitud se llega a la estimación con máxima verosimilitud de $\hat{\delta}$ y $\boldsymbol{\hat{\sigma}^2}$

$$\hat{\delta} = \delta_0 - \hat{\rho}\delta_d$$
$$\hat{\sigma}^2 = \mathrm{n}^{-1}S(\hat{\rho})$$
$$\widehat{\Omega} = \widehat{\sigma}^2[(I_n - \hat{\rho}W)'(I_n - \hat{\rho}W)]^{-1}$$

Estimación del modelo de Error Espacial (SEM)

Con una estrategia parecida, se obtiene la solución para *SEM*

$$y = X\beta + u$$
$$u = \lambda W u + \varepsilon$$
$$\varepsilon \sim N(0, \sigma^2 I_n)$$

$$ln\ L = -\left(\frac{n}{2}\right)\ln(\pi\sigma^2) + ln|I_n - \lambda W| - \frac{e'e}{2\sigma^2}$$
$$e = (I_n - \lambda W)(y - X\beta)$$

Para un valor dado de λ,

$$\beta(\lambda) = \left(X(\lambda)'X(\lambda)\right)^{-1}X(\lambda)'y(\lambda), \text{ donde}$$
$$X(\lambda) = (X - \lambda WX)$$
$$y(\lambda) = (y - \lambda Wy)$$
$$\sigma^2(\lambda) = e(\lambda)'e(\lambda)/n$$
$$e(\lambda) = y(\lambda) - X(\lambda)\beta(\lambda)$$

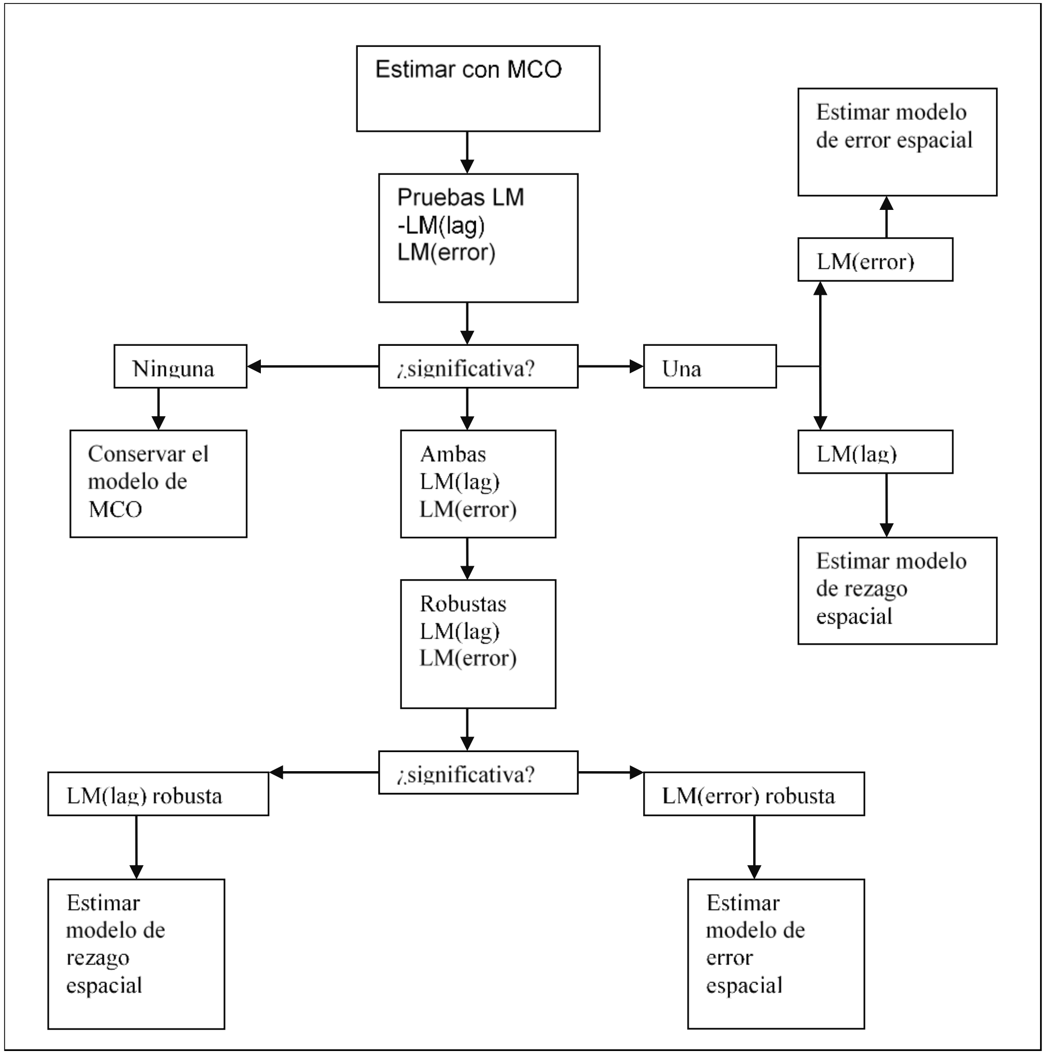
La función de log-verosimilitud concentrada en los parámetros β y σ^2

$$lnL(\lambda) = \kappa + ln|I_n - \lambda W| - (n/2)\ln(S(\lambda))$$

$S(\lambda) = e(\lambda)'e(\lambda)$ no es cuadrático, se necesita todo un proceso simultáneo

$$\hat{\beta} = \beta(\hat{\lambda})$$
$$\hat{\sigma}^{\mathbf{2}} = \mathrm{n}^{-1}S(\hat{\lambda})$$
$$\widehat{\Omega} = \widehat{\sigma}^2\left[\left(I_n - \hat{\lambda}W\right)'\left(I_n - \hat{\lambda}W\right)\right]^{-1}$$

Figura 5. Estrategia de Selección de Modelos



Fuente: Anselin, Luc (2005**)**

Contrastes de autocorrelación espacial

Estos contrastes se aplican después de estimar el modelo clásico para analizar la presencia de algún tipo de dependencia espacial. La hipótesis nula es que el tipo de dependencia espacial es igual a cero, contra la hipótesis alternativa de que es diferente de cero.

1) Test I de Moran

Mide el efecto de autocorrelación espacial en los residuos e_i en un modelo no-espacial o clásico, sin distinguir estructuras de Rezago o Error Espacial:

$$I = \frac{N}{S_0} \frac{\sum_{(2)} w_{ij} e_i e_j}{\sum_{i=1}^{N} e_i^2} = \frac{N}{S_0} \frac{e'We}{e'e}$$

La inferencia se hace con el valor z estandarizado. El primer y segundo momento

$$E[I] = \frac{N}{S_0} \frac{tr(MW)}{N-K} \;;\; E[I]^2 = \frac{\left(\frac{N}{S_0}\right)^2 tr(MWMW') + tr(MW)^2 + [tr(MW)]^2}{(N-K)(N-K+2)}$$

Se distribuye como una χ^2 con un grado de libertad

2) Test LM-LAG: Rezago Espacial

Por rezago espaciales de la variable endógena (Anselin, 1988):

$$LM - LAG = \frac{\left[\frac{e'Wy}{s^2}\right]^2}{\left(R\tilde{J}_{\rho-\beta}\right)}$$

Se distribuye como una χ^2 con un grado de libertad

3) Test LM-LE: Rezago Espacial (Robusto)

El estadístico es robusto ante la presencia de dependencia local del error espacial (Anselin, 1988):

$$LM - LE = \frac{\left[\frac{e'W_1y}{s^2} - \frac{e'W_1e}{s^2}\right]^2}{\left(R\tilde{J}_{\rho-\beta} - T_1\right)}$$

Se distribuye como una χ^2 con un grado de libertad

3) Test SARMA: Rezago y Error Espacial

Es robusto ante la presencia de dependencia local y del error espacial (Anselin, 1988):

$$SARMA = \frac{\left[\frac{e'Wy}{s^2} - \frac{e'We}{s^2}\right]^2}{\left(R\tilde{J}_{\rho-\beta} - T_1\right)} + \frac{\left[\frac{e'We}{s^2}\right]^2}{T_1}$$

Se distribuye como una χ^2 con dos grados de libertad.

Referencias

Anselin, L. (1988) *Spatial Econometrics: Methods and Models*. Kluwer Academic Publishers, Dordrecht, The Netherlands.

Anselin, L. (2012) *OpenGeoDa 1.2 User's Guide*. Spatial Analysis Laboratory (SAL). Department of Agricultural and Consumer Economics, University of Illinois, Urbana-Champaign, IL.

Fotheringham, Brunsdon y Charlton (2000) *Quantitative Geography: Perspectives on Spatial Data Analisys*.

Haining, Robert (2003) *Spatial Data Analysis: Theory and Practice*, 1st edition, Cambridge University Press

LeSage, J. y Pace, K. (2009) *Introduction of Spatial Econometrics*, Taylor & Francis Group, LLC.