Dependencia Espacial: Detección, Validación y Modelación

Gabriel Alberto Agudelo Torres

Trabajo de grado presentado como requisito parcial para optar al título de Magíster en Matemáticas Aplicadas

Asesores: Dr. Hermilson Velásquez Ceballos
Dr. Juan Carlos Duque Cardona

Medellín
Universidad EAFIT
Departamento de Ciencias Básicas
Noviembre de 2010

Tabla de Contenido

1 Preliminares Matemáticos **7**

1.1 Ruido blanco espacial 7

1.2 Transformación de variables aleatorias 7

1.3 Distribución normal multivariada 8

1.4 Método de Mínimos Cuadrados Ordinarios 9

1.5 Método de Máxima Verosimilitud 10

1.6 Estimador insesgado 11

1.7 Estimador eficiente 11

1.8 Convergencia en probabilidad 12

1.9 Estimador consistente 12

1.10 Propiedades de la traza de una matriz 13

2 Modelación con datos espaciales **15**

2.1 Matriz de contigüidades (W) 16

2.2 Matriz W como Operador de Retardo Espacial 19

2.3 Detección de efectos espaciales 21

2.3.1 Estadístico de contactos (τ) 21

2.3.2 I de Moran 23

2.3.3 Test c de Geary 25

2.3.4 Test $G(d)$ de Getis y Ord 26

2.4 Dependencia espacial en un modelo de regresión 27

2.5 Taxonomía de Vayá 27

2.5.1 Dependencia espacial sustantiva 27

2.5.2 Dependencia espacial residual 28

2.6 Taxonomía de Florax y Folmer 30

2.6.1 Modelo General de Regresión Espacial 30

2.6.2 Modelo básico de regresión lineal 31

2.6.3 Modelo mixto autorregresivo de regresión espacial o modelo del retardo espacial . . . 32
2.6.4 Modelo mixto regresivo cruzado de regresión espacial . 32
2.6.5 Modelo mixto autorregresivo regresivo cruzado de regresión espacial . . . 33
2.6.6 Modelo de regresión con dependencia espacial en la perturbación aleatoria o modelo del error espacial . . . 33
2.6.7 Modelo mixto autorregresivo de regresión espacial con perturbación aleatoria espacialmente autorregresiva . . 34
2.6.8 Modelo Mixto Autorregresivo de Regresión Espacial, con perturbaciones que incorporan un esquema de dependencia de media móvil . . . 37
2.7 Taxonomía de Anselin . . . 38
2.7.1 Modelo autorregresivo de regresión espacial de orden 1. SAR(1) . . . 38
2.7.2 Modelo básico de regresión lineal . . . 43
2.7.3 Modelo de regresion con dependencia espacial en la perturbación aleatoria o modelo del error espacial . . . 44
2.7.4 Modelo mixto autorregresivo de regresión espacial o modelo del retardo espacial . . . 44

3 Autocorrelación Espacial Espuria 47
3.1 Desarrollo teórico del algoritmo . . . 49
3.2 Resultados del algoritmo . . . 56

4 Defunción de adultos no pensionados en Colombia. Dependencia Espacial 65
4.1 División politico-económica de Colombia . . . 66
4.2 Análisis de dependencia espacial . . . 68

Algoritmo de Agregación Espacial - MatLab 81

Validación del Algoritmo - MatLab 93

Agradecimientos

En 1996, al mirar mi primer cuaderno de álgebra, mi papá me preguntó: "¿usted va a ser ingeniero? Tiene unos números bonitos" El año pasado, siendo ingeniero financiero y habiendo casi terminado todos los cursos de la Maestría en Matemáticas Aplicadas, me preguntó: "¿y después de esto qué?, ¿sigue el doctorado?". Recibió la misma respuesta de hace 14 años.

Al finalizar esta tesis de maestría estoy dando el último paso de una etapa que sin duda fue un gran reto en mi vida académica. Al hacerlo, quiero agradecer a tantas personas que hicieron posible la materialización de esta tesis y la culminación de la Maestría en Matemáticas Aplicadas.

En primer lugar quiero agradecer a Dios por darme la fortaleza, la paciencia y la capacidad para encarar este reto.

Quiero agradecer también a mis padres, Soledad Torres Lara y Conrado Agudelo Gómez, por la formación y apoyo recibidos durante todos estos años de estudio y por estar siempre dispuestos a ayudar en todo sentido ante cualquier dificultad. A mi hermano Jorge quien desde los años de preescolar ha estado a mi lado, silencioso, pero dispuesto siempre a ayudar con una explicación, con la solución de alguna duda o con unas palabras de aliento para seguir adelante. A mi tía Elvira quiero agradecer su cuidado y sus pacientes explicaciones en los primeros años de mi vida académica, años que sin duda marcarían desde esos días el rumbo que seguiría tiempo después.

Debo expresar también un especial agradecimiento a *yes*, por su amor, apoyo incondicional, sacrificio, y palabras de aliento ante las dificultades surgidas durante esta maestría. Sin todo esto no hubiera sido posible la realización de este logro.

Finalmente, quiero agradecer al profesor Juan Carlos Duque Cardona por su colaboración en el desarrollo del algoritmo de agregación espacial, al profesor Hermilson Velásquez Ceballos por su orientación al dirigir esta tesis y por sus claras y valiosas explicaciones, y a la Administradora de Fondos de Pensiones y Cesantías Protección por todo el apoyo que recibí a lo largo de estos tres años.

Prólogo

En años recientes el creciente nivel de integración regional en ámbitos tan diversos como los culturales, comerciales, económicos, religiosos, entre otros; han llevado el análisis económico y especialmente a la econometría, un paso más allá en la búsqueda de explicar fenómenos cuyo origen no se puede determinar *a priori*. Estas condiciones fueron las perfectas para el nacimiento y desarrollo de una rama de la econometría denominada *econometría espacial*, la cual brinda al investigador las herramientas necesarias para recoger en los modelos econométricos tradicionales los efectos espaciales existentes en determinada estructura económica.

La econometría espacial es una rama de la econometría en la cual se consideran en la especificación elementos relacionados con el espacio. Esta relación constituye el primer y más importante efecto espacial entre variables y se denomina dependencia espacial; su presencia, traducida en autocorrelación espacial, puede invalidar modelos econométricos tradicionales tal como la autocorrelación temporal lo hace comúnmente en este tipo de especificaciones.

Moreno y Vayá (2000) definieron la dependencia espacial como "la existencia de una relación funcional entre lo que ocurre en un punto determinado del espacio y lo que ocurre en otro lugar". Es decir, la variable endógena presentará altos niveles de autocorrelación espacial cuando los valores de ella dependen no solo de variables exógenas sino también de la localización geográfica de estos valores (tanto de la variable endógena como de las exógenas).

Para constrastar la presencia de autocorrelación espacial se puede utilizar alguno de los estadísticos que se expondrán en el capítulo 2 de esta tesis, siendo el más utilizado el de Moran (1948).

El segundo efecto espacial se denomina heterogeneidad espacial y hace referencia a la falta de estabilidad en las relaciones interregionales que se presentan. Esta inestabilidad se traduce en variaciones en las especificaciones, en los parámetros o en ambos, de región a región de acuerdo a su localización, impidiendo así encontrar una forma funcional espacial que ayude a explicar de forma razonable el comportamiento de la variable endógena.

El desarrollo de la econometría espacial se remonta a 1948, cuando Morán desarrolla su estadístico I y continúa con Geary en 1954 cuando éste desarrolla un estadístico basado en el Durbin – Watson para la detección de autocorrelación serial.

A finales de la década del 70, Paelinck y Klaasen (1979) publican su obra "Spatial Econometrics", en la cual aparte de proponer un adecuado nombre para esta disciplina, hace lo que ellos denominan "un balance provisional" que les permitirá ordenar sus ideas. Y fue esta publicación la que les permitió cuatro años más tarde aplicar al conjunto de regiones europeas un modelo multi e interregional llamado "Factores de Localización en Europa".

Este modelo llevó a Paelinck a la conclusión que la modelación de efectos espaciales requería mucho más que una especificación "espacial" inmersa en la econometría tradicional. La preferencia de Paelinck por hacer un mayor énfasis en la fase de especificación en vez de modelar los residuos en los pasos finales, le permitió definir cinco reglas que de ahí en adelante guiarían las investigaciones sobre modelación espacial. Estos principios son los siguientes:

1. Las variables endógenas deben modelarse como entidades espacialmente interdependientes.
2. Las relaciones sobre el espacio son fundamentalmente asimétricas debido a la diversidad espacial.
3. Las variables exógenas están influidas por la localización física de los datos.
4. Los modelos no lineales serán más apropiados para estimar los parámetros presentes en las ecuaciones espaciales.
5. Los elementos topológicos (tales como coordenadas, vecindades, distancias, densidades, etc.) no pueden estar ausentes de los modelos espaciales. La simple introducción de un índice para indicar la región no basta para especificar apropiadamente relaciones que tienen lugar sobre el espacio.

A partir de estos aportes de Paelinck y Klaasen la econometría espacial tuvo un gran desarrollo en los años ochenta y noventa e inició con los trabajos de Cliff y Ord (1981) y continuó con los textos de Anselin (1980, 1988A), siendo el texto de Anselin (1988A) el que sería considerado según Chasco (2003) "el manual por excelencia de esta disciplina". Esta literatura se fue robusteciendo con varios artículos publicados por autores tales como el propio Anselin, Florax, Fisher, Getis y Ord.

A pesar de los avances logrados por estos autores, existe aun un marcado desequilibrio entre el estudio de la dimensión espacial con respecto a la dimensión temporal. Desde 1970, con la publicación de la célebre obra "Time Series Analysis. Forecasting and Control" por parte de Box y Jenkings, el mundo académico se volcó hacia la dimensión temporal permitiendo así un incremento inusitado en las investigaciones relacionadas con el análisis de series temporales y el desarrollo de software que acompaña estos procesos de investigación y desarrollo. Estos factores contribuyeron al escaso interés en la dimensión espacial, coherente con las necesidades de predicción, pero incoherente si se tiene en cuenta que los agentes económicos objeto del estudio se encuentran dispersos en el espacio y se esfuerzan por disminuir cada vez más las distancias que los separan.

Capítulo 1

Preliminares Matemáticos

En esta sección se presentan algunas definiciones y deducciones matemáticas y estadísticas que constituyen una pieza fundamental tanto para la siguiente sección dedicada al desarrollo de los modelos econométricos espaciales como para las secciones dedicadas a la detección y corrección de la autocorrelación espacial espuria y a su aplicación en las defunciones de personas entre los 15 y los 64 años en las subregiones colombianas.

1.1 Ruido blanco espacial

Definición 1 *Se dice que una variable aleatoria ε_i es ruido blanco espacial si:*

1. $E[\varepsilon_i] = 0$

2. $E[\varepsilon_i, \varepsilon_j] = \begin{cases} \sigma^2 \text{ para } \varepsilon_i = \varepsilon_j \\ 0 \text{ en otro caso} \end{cases}$

1.2 Transformación de variables aleatorias

Teorema 2 *Sea X una variable aleatoria con función de distribución acumulada $F_X(X)$, $Y = g(X)$, $\mathcal{X} = \{x : f_X(X) > 0\}$ y $\mathcal{Y} = \{y : y = g(x) \text{ para algún } x \in \mathcal{X}\}$.*

Si g es una función creciente en $\mathcal{X}$, entonces $F_Y(y) = F_X(g^{-1}(y))$ para $y \in \mathcal{Y}$.

Si g es una función decreciente en X y X es una variable aleatoria contínua, entonces $F_Y(y) = 1 - F_X\left(g^{-1}(y)\right)$ para $y \in Y$.

La demostración formal de este teorema se encuentra en Casella (2002).

Teorema 3 *Sea X una variable aleatoria con función de densidad de probabilidad $f_X(x)$ y sea $Y = g(X)$, donde g es una función monótona, $X = \{x : f_X(x) > 0\}$ y $Y = \{y : y = g(x) \text{ para algún } x \in X\}$. Suponga que $f_X(x)$ es contínua en X y que $g^{-1}(y)$ tiene derivada contínua en Y. Entonces, la función de densidad de probabilidad está dada por:*

$$f_Y(y) = \begin{cases} f_X\left(g^{-1}(y)\right)\left|\frac{d}{dy}g^{-1}(y)\right|, & y \in Y \\ 0 \text{ en otro caso} \end{cases} \tag{1.1}$$

Demostración. Por Teorema 2 y utilizando la regla de la cadena, se tiene que:

$$f_Y(y) = \frac{d}{dy}F_Y(y) = \begin{cases} f_X\left(g^{-1}(y)\right)\frac{d}{dy}g^{-1}(y) & \text{si } g \text{ es creciente} \\ -f_X\left(g^{-1}(y)\right)\frac{d}{dy}g^{-1}(y) & \text{si } g \text{ es decreciente} \end{cases}$$

lo cual puede ser expresado como:

$$f_Y(y) = \begin{cases} f_X\left(g^{-1}(y)\right)\left|\frac{d}{dy}g^{-1}(y)\right|, & y \in Y \\ 0 \text{ en otro caso} \end{cases} \quad \blacksquare$$

1.3 Distribución normal multivariada

Teorema 4 *Sea $X = [X_1, X_2, ..., X_m]$ un vector de m variables aleatorias con vector de medias $\mu = E[X] = \begin{bmatrix} E(X_1) \\ \vdots \\ E(X_m) \end{bmatrix}$, y matriz de varianzas y covarianzas $\Sigma_X = \begin{bmatrix} V(X_1) & Cov(X_1, X_2) & \cdots & Cov(X_1, X_m) \\ Cov(X_2, X_1) & V(X_2) & \cdots & Cov(X_1, X_m) \\ \vdots & \vdots & \ddots & \vdots \\ Cov(X_m, X_1) & Cov(X_m, X_2) & \cdots & V(X_m) \end{bmatrix} = \begin{bmatrix} \sigma_1^2 & \sigma_{12} & \cdots & \sigma_{1m} \\ \sigma_{21} & \sigma_2^2 & \cdots & \sigma_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{m1} & \sigma_{m2} & \cdots & \sigma_m^2 \end{bmatrix}$.*

Se dice que X tiene una distribución normal m-variante si su función de densidad de probabilidad es:

$$f_X(x) = (2\pi)^{-m/2} (\det \Sigma)^{-1/2} e^{(-1/2)(X-\mu)'\Sigma^{-1}(X-\mu)}$$

Donde Σ es definida positiva.

Demostración. Puede escribirse Σ como $\Sigma = CC'$ donde C es una matriz de $m\text{x}m$ no singular.

Sea $X = CU + \mu$

Donde $U_{m\text{x}1}$ es un vector de variables aleatorias independientes que se distribuyen $N(0,1)$, es decir, U se distribuye $N_m(0, I_m)$. La función de densidad de probabilidad conjunta de $U_1, U_2, ..., U_m$ es:

$$f_U(u) = \prod_{i=1}^{m} (2\pi)^{-1/2} e^{(-1/2)u_i^2}$$

$$f_U(u) = (2\pi)^{-m/2} e^{(-1/2)u'u}$$

La transformación inversa es $U = B(X-\mu)$, con $B = C^{-1}$, y el jacobiano de esta transformación es:

$$\det \begin{bmatrix} \frac{\partial u_i}{\partial x_i} & \cdots & \frac{\partial u_i}{\partial x_m} \\ \vdots & \ddots & \vdots \\ \frac{\partial u_m}{\partial x_i} & \cdots & \frac{\partial u_m}{\partial x_m} \end{bmatrix} = \det \begin{bmatrix} b_{11} & \cdots & b_{1m} \\ \vdots & \ddots & \vdots \\ b_{m1} & \cdots & b_{mm} \end{bmatrix}$$

$\det(B) = \det(C^{-1}) = (\det C^{-1}) = (\det(CC'))^{-1/2} = (\det(\Sigma))^{1/2}$

de esta forma $f_X(x) = (2\pi)^{-m/2} (\det \Sigma)^{-1/2} e^{(-1/2)(x-\mu)'C^{-1\prime}C^{-1}(x-\mu)}$

y como $\Sigma^{-1} = C^{-1\prime}C^{-1}$, entonces

$f_X(x) = (2\pi)^{-m/2} (\det \Sigma)^{-1/2} e^{(-1/2)(x-\mu)'\Sigma^{-1}(x-\mu)}$ ■

1.4 Método de Mínimos Cuadrados Ordinarios

Proposición 5 *Considerese una variable aleatoria endógena $Y_i = f(X_{i2}, X_{i3}, ..., X_{iK}, U_i)$ y la especificación $Y_i = \beta_1 + \beta_2 X_{i2} + \beta_3 X_{i3} + ... + \beta_K X_{iK} + U_i$, con $i = 1, 2, ..., N$. Donde $X_{i2}, X_{i3}, ..., X_{iK}$ son variables aleatorias explicativas de la variabilidad de Y_i o exógenas y U_i es una variable aleatoria ruido blanco.*

Para $i = 1, 2, ..., N$ la información del sistema de ecuaciones generado se puede representar en forma matricial como $Y = X\beta + U$, donde:

$$Y_{Nx1} = \begin{bmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_N \end{bmatrix}; \quad X_{NxK} = \begin{bmatrix} 1 & X_{12} & X_{13} & \cdots & X_{1K} \\ 1 & X_{22} & X_{23} & \cdots & X_{2K} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & X_{N2} & X_{N3} & \cdots & X_{NK} \end{bmatrix}; \quad \beta_{Kx1} = \begin{bmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_K \end{bmatrix};$$

$$U_{Nx1} = \begin{bmatrix} U_1 \\ U_2 \\ \vdots \\ U_N \end{bmatrix}$$

Y_{Nx1}: *Vector de observaciones de la variable endógena.*
X_{NxK}: *Matriz de diseño.*
β_{Kx1}: *Vector de parámetros.*
U_{Nx1}: *Valor de perturbación aleatoria.*

Definición 6 *El criterio de Mìnimos Cuadrados Ordinarios está relacionado con la solución del problema de optimización*

$$\min_{\beta} \sum_{i=1}^{N} U_i^2 = \min_{\beta} U^T U$$

el cual permite obtener $\widehat{\beta}_{MCO} = (X'X)^{-1} X'X$

1.5 Método de Máxima Verosimilitud

Definición 7 *Sea* $x_1, x_2, ..., x_N$ *una muestra aleatoria de una población con función de densidad de probabilidad* $f(X;\theta)$.

La función de verosimitud de la muestra es la funciòn de densidad de probabilidad conjunta y al ser $x_1, x_2, ..., x_N$ *muestras aleatorias, esta es el producto de las funciones de densidad de probabilidad marginales.*

$$L(X;\theta) = L(x_1, x_2, ..., x_N;\theta) = f(x_1, x_2, ..., x_N, \theta) = f(x_1, \theta) f(x_2, \theta) ... f(x_N, \theta)$$

El estimador máximo verosímil de θ *se obtiene resolviendo el problema de optimización:*

$$\max_{\theta} L(X;\theta)$$

1.6 Estimador insesgado

Definición 8 *Un estimador $\widehat{\theta}$ de un parámetro θ es insesgado cuando verifica:*
$E\left[\widehat{\theta}\right] = \theta$

Definición 9 *Se llama sesgo de $\widehat{\theta}$ a la diferencia entre $E\left[\widehat{\theta}\right]$ y θ.*
$Sesgo\left(\widehat{\theta}\right) = E\left[\widehat{\theta}\right] - \theta$

Proposición 10 *El Error Cuadrático Medio (ECM) de un estimador insesgado es $V\left[\widehat{\theta}\right]$*

Demostración. Sea $\widehat{\theta}$ un estimador insesgado de θ.

Por definición de estimador insesgado se sabe que $E\left[\widehat{\theta}\right] = \theta$ y por construcción del error cuadrático medio (ECM):
$ECM\left(\widehat{\theta}\right) = E\left[\left(\widehat{\theta} - \theta\right)^2\right]$
$\Rightarrow ECM\left(\widehat{\theta}\right) = E\left[\widehat{\theta} - E\left(\widehat{\theta}\right) + E\left(\widehat{\theta}\right) - \theta\right]^2$
$\Rightarrow ECM\left(\widehat{\theta}\right) = E\left[\left(\widehat{\theta} - E\left(\widehat{\theta}\right)\right)^2 + \left(Sesgo\left(\widehat{\theta}\right)\right)^2 + 2\left(\widehat{\theta} - E\left(\widehat{\theta}\right)\right) Sesgo\left(\widehat{\theta}\right)\right]$
$\Rightarrow ECM\left(\widehat{\theta}\right) = V\left[\widehat{\theta}\right] + \left[Sesgo\left(\widehat{\theta}\right)\right]^2$
Al ser el estimador insesgado, $Sesgo\left(\widehat{\theta}\right) = 0$, por lo tanto
$ECM\left(\widehat{\theta}\right) = V\left[\widehat{\theta}\right]$ ■

1.7 Estimador eficiente

Definición 11 *Dados dos estimadores insesgados, $\widehat{\theta}_1$ y $\widehat{\theta}_2$, de un parámetro θ, con varianzas $V\left[\widehat{\theta}_1\right]$ y $V\left[\widehat{\theta}_2\right]$ respectivamente, la eficiencia relativa de $\widehat{\theta}_1$ con respecto a $\widehat{\theta}_2$ se define como la razón*
$eficiencia = \frac{V\left[\widehat{\theta}_1\right]}{V\left[\widehat{\theta}_2\right]}$

1.8 Convergencia en probabilidad

Definición 12 *La sucesión de variables aleatorias $\{Y_n\}$ converge en probabilidad a la variable aleatoria Y si y solo si*

$$\underset{n\to\infty}{Lim} P\left(|y_n - y| < \varepsilon\right) = 1, \ \ \forall\varepsilon > 0$$

*La convergencia en probabilidad se denotará como $Y_n \xrightarrow{p} Y$ ó $pLim\ Y_n = Y$. La anterior notación significa **límite en probabilidad de** Y_n **es** Y**.***

1.9 Estimador consistente

Definición 13 *El estimador $\widehat{\theta}_n$ es un estimador consistente de θ si para cualquier número positivo ε,*

$$\underset{n\to\infty}{Lim} P\left(\left|\widehat{\theta}_n - \theta\right| \leq \varepsilon\right) = 1$$

o de forma equivalente,

$$\underset{n\to\infty}{Lim} P\left(\left|\widehat{\theta}_n - \theta\right| > \varepsilon\right) = 0$$

Teorema 14 *Un estimador insesgado $\widehat{\theta}_n$ para θ es un estimador consistente de θ si*

$$\underset{n\to\infty}{Lim} V\left(\widehat{\theta}_n\right) = 0$$

Demostración. Si Y es una variable aleatoria con $E(Y) = \mu$ y $V(Y) = \sigma^2 < \infty$ y si k es cualquier constante no negativa, el teorema de Tchebysheff implica que

$$P\left(|Y - \mu| > k\sigma\right) \leq \frac{1}{k^2}$$

Dado que $\widehat{\theta}_n$ es un estimador insesgado de θ, $E\left(\widehat{\theta}_n\right) = \theta$. Si se aplica el teorema de Tchebysheff para la variable aleatoria $\widehat{\theta}_n$ se obtiene:

$$P\left(\left|\widehat{\theta}_n - \theta\right| > k\sqrt{V\left(\widehat{\theta}_n\right)}\right) \leq \frac{1}{k^2}$$

para cualquier número positivo ε,

$$k = \frac{\varepsilon}{\sqrt{V\left(\widehat{\theta}_n\right)}}$$

es un número positivo. La aplicación del teorema de Tchebysheff para esta elección de k muestra que

$$P\left(\left|\widehat{\theta}_n - \theta\right| > \frac{\varepsilon}{\sqrt{V\left(\widehat{\theta}_n\right)}}\sqrt{V\left(\widehat{\theta}_n\right)}\right) \leq \frac{\sqrt{V\left(\widehat{\theta}_n\right)}}{\varepsilon^2}$$

$\Rightarrow P\left(\left|\widehat{\theta}_n - \theta\right| > \varepsilon\right) \leq \frac{\sqrt{V\left(\widehat{\theta}_n\right)}}{\varepsilon^2}$

si $\underset{n\to\infty}{Lim} V\left(\widehat{\theta}_n\right) = 0$ entonces

$\underset{n\to\infty}{Lim} P\left(\left|\widehat{\theta}_n - \theta\right| > \varepsilon\right) = 0$

Lo que significa que $\widehat{\theta}_n$ es un estimador consistente de θ. ■

1.10 Propiedades de la traza de una matriz

Teorema 15 *Si A y B son matrices de m x n y n x m respectivamente, entonces $Tr(AB) = Tr(BA)$.*

Demostración. *Sea $C = AB$. Así*

$c_{ij} = \sum_{k=1}^{n} a_{ik}b_{kj}$

para $i = j$ la expresión anterior queda de la siguiente forma:

$c_{ii} = \sum_{k=1}^{n} a_{ik}b_{ki}$

Entonces,

$Tr(C) = \sum_{i=1}^{n} c_{ii} = \sum_{i=1}^{m}\sum_{k=1}^{n} a_{ik}b_{ki} = \sum_{k=1}^{n}\sum_{i=1}^{m} a_{ik}b_{ki} = \sum_{k=1}^{n}\sum_{i=1}^{m} b_{ki}a_{ik}$

Por otro lado, si $D = BA$, entonces

$d_{ij} = \sum_{k=1}^{n} b_{ik}a_{kj}$

para $i = j$ la expresión anterior queda de la siguiente forma:

$d_{ii} = \sum_{k=1}^{n} b_{ik}a_{ki}$

Entonces,

$Tr(D) = \sum_{i=1}^{n} d_{ii} = \sum_{k=1}^{n}\sum_{i=1}^{m} b_{ik}a_{ki}$

Comparando los resultados obtenidos, se observa que:

$Tr(C) = Tr(AB) = \sum_{k=1}^{n}\sum_{i=1}^{m} b_{ki}a_{ik}$ *y que* $Tr(D) = \sum_{i=1}^{n} d_{ii} = \sum_{k=1}^{n}\sum_{i=1}^{m} b_{ik}a_{ki}$

Se concluye entonces, que cambiando los nombres de los índices i y k se obtiene:

$Tr\,(AB) = Tr\,(BA)$ ■

Teorema 16 *Si A, B y C son matrices de nxn se cumple que:*

$$Tr\,(ABC) = Tr\,(CAB) = Tr\,(BCA) \tag{1.2}$$

Demostración. *Sea $D = AB$. Luego, $Tr\,(ABC) = Tr\,(DC)$.*

Por teorema 10, se sabe que $Tr\,(DC) = Tr\,(CD)$, por lo tanto $Tr\,(ABC) = Tr\,(CAB)$.

En forma análoga se desarrolla la demostración de la otra igualdad. ■

Capítulo 2

Modelación con datos espaciales

La dependencia espacial es el fenómeno que se presenta cuando los valores observados en una región dependen de los valores de las regiones vecinas. Más adelante se expondrá una situación en la cual el valor tomado por una variable en una región no solo depende del valor de la misma variable en las regiones vecinas, sino también de los valores de otras variables en estas mismas regiones.

El proceso generador de datos del primer caso puede entonces tomar la siguiente forma:

$$
\begin{aligned}
Y_i &= \alpha_i Y_j + X_i \beta + \varepsilon_i \\
Y_j &= \alpha_j Y_i + X_j \beta + \varepsilon_j
\end{aligned}
$$

con $\varepsilon_{i,} \varepsilon_j \sim N(0, \sigma^2)$.

Por simplicidad se asumen términos de error con media cero y varianza σ^2.

Para el caso de tres regiones sería entonces:

$$
\begin{aligned}
Y_i &= \alpha_i Y_j + \lambda_i Y_k + X_i \beta + \varepsilon_i \\
Y_j &= \alpha_j Y_i + \lambda_j Y_k + X_j \beta + \varepsilon_j \\
Y_k &= \alpha_k Y_i + \lambda_k Y_j + X_k \beta + \varepsilon_k
\end{aligned}
$$

con $\varepsilon_{i,}\varepsilon_j, \varepsilon_k \sim N(0, \sigma^2)$.

En esta situación ya se observa claramente que este esquema no es práctico pues supondría hallar estimaciones para un número de parámetros superior al número de observaciones disponibles.

Ord (1975) propuso una especificación diferente para modelar la dependencia espacial, la cual se caracteriza por incluir un solo parámetro que recoge la relación existente entre las agregaciones de regiones vecinas y las propias regiones.

Para desarrollar esta especificación, Ord utilizó la primera ley de la geografía formulada por Tobler (1979), la cual indica que "todo tiene que ver con todo, pero las cosas cercanas están más relacionadas entre sí que las cosas lejanas". Ord materializó este principio incluyendo en sus modelos econométricos una "matriz de retardos espaciales" o "matriz de contigüidades" W cuya forma y utilidad se exponen a continuación.

2.1 Matriz de contigüidades (W)

La matriz de contigüidades representa la materialización de la primera ley de Tobler en términos de proximidad o lejanía. La matriz resultante es entonces del tipo binario, donde 1 representa la proximidad entre la región correspondiente a la fila y 0 el caso contrario (la diagonal principal estará compuesta por ceros al no tener sentido el concepto de proximidad consigo mismo).

Al ser difícil la obtención de información más precisa que denote proximidad o lejanía entre un par de regiones, es comúnmente utilizada la vecindad como expresión de proximidad. Sin embargo, esta vecindad debe ser definida con mayor detalle para lograr una correcta especificación del modelo a utilizar.

Acevedo y Velásquez (2008) recopilan algunas de las formas para definir la presencia o ausencia de contigüidad. Algunas de ellas son:

Contigüidad tipo torre:

	B	
B	A	B
	B	

En esta definición de contigüidad se tienen en cuenta como vecinas regiones que tienen una frontera común con una longitud positiva.

Contigüidad tipo alfil:

B		B
	A	
B		B

En el esquema de contigüidad tipo alfil, se consideran vecinas solo las regiones con las cuales existe una frontera determinada solo por un punto.

Contigüidad tipo reina:

B	B	B
B	A	B
B	B	B

La contigüidad tipo reina se caracteriza por no dar relevancia al tamaño de la frontera entre las regiones, basta con que ésta sea positiva o cercana a cero para que la región sea considerada contigua.

Un ejemplo ilustrativo sobre la construcción de la matriz W a través de un esquema de contigüidad tipo torre puede ser dado a partir de un mapa con 9 regiones de la siguiente forma:

1	2	3
4	5	6
7	8	9

La matriz de contigüidades es entonces la siguiente:

$$W = \begin{array}{c} \\ 1 \\ 2 \\ 3 \\ 4 \\ 5 \\ 6 \\ 7 \\ 8 \\ 9 \end{array} \begin{array}{c} \begin{array}{ccccccccc} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 \end{array} \\ \begin{bmatrix} 0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 1 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 1 & 0 \end{bmatrix} \end{array}$$

Donde el 1 correspondiente a la fila 1 – columna 2 indica contigüidad entre la región 1 y la región 2. Los ceros en la diagonal principal indican la imposibilidad de hablar de una región contigua consigo misma.

Claramente la matriz W, por construcción, es simétrica. Esta característica tomará relevancia más adelante al demostrar características importantes de los modelos econométrico – espaciales.

La matriz de contigüidades así definida puede denominarse "de primer orden"(W_1), pues capta la contigüidad en regiones entre las cuales un individuo tendría que cruzar solo una frontera para desplazarse de la una a la otra.

Utilizando la definición de Mur y Angulo (2008) puede decirse que una matriz W de segundo orden (W_2) capta las relaciones entre dos regiones en las que el camino más corto entre ambas supone cruzar dos fronteras. Volviendo al ejemplo de las nueve regiones, la matriz de contigüidad de segundo orden sería entonces:

$$W = \begin{bmatrix} 0&0&1&0&1&0&1&0&0 \\ 0&0&0&1&0&1&0&1&0 \\ 1&0&0&0&1&0&0&0&1 \\ 0&1&0&0&0&1&0&1&0 \\ 1&0&1&0&0&0&1&0&1 \\ 0&1&0&1&0&0&0&1&0 \\ 1&0&0&0&1&0&0&0&1 \\ 0&1&0&1&0&1&0&0&0 \\ 0&0&1&0&1&0&1&0&0 \end{bmatrix}$$

Mur y Angulo (2008) presentan también un algoritmo que permite obtener una matriz de contigüidad de cualquier orden partiendo de la matriz de orden 1 sin necesidad de hacer un análisis cartográfico.

2.2 Matriz W como Operador de Retardo Espacial

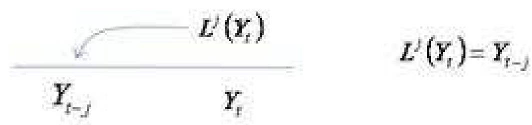
Como se mencionó anteriormente las series de tiempo tienen una característica fundamental y es su unidimensionalidad. Chasco (2003) establece tres características básicas que tienen los datos espaciales: el primero es su naturaleza georreferenciada, es decir, que su posición en el espacio contiene información valiosa para el investigador; la segunda característica es la multidireccionalidad de las relaciones que se establecen en el espacio y la tercera es la multidimensionalidad, puesto que no nos movemos en una sola dimensión sino a través de un espacio geográfico compuesto por tantas dimensiones como regiones existan.

En este sentido, la matriz de contigüidades W es también conocida como "matriz de retardos espaciales" pues se asemeja al operador de retardos utilizado en series temporales. En el contexto temporal, el operador de retardos provoca un desplazamiento en el eje del tiempo equivalente a la potencia j.

Análogamente, la matriz de contactos provoca desplazamientos en el espacio hacia el nivel de contigüidad asociado con la matriz.

Para ilustrar la naturaleza unidimensional de la serie de tiempo y multidireccional de la serie espacial y el efecto causado por los retardos temporales y espaciales pueden utilizarse los siguientes diagramas:

1. Efecto del operador de retardo en una serie temporal:



Este operador de retardo al ser aplicado a una variable, la desplaza en la dimensión temporal tantos períodos como lo indique su potencia.

2. Efecto del operador de retardo en una serie espacial:

$$W_1 Y = \begin{bmatrix} 0&1&0&1&0&0&0&0&0 \\ 1&0&1&0&1&0&0&0&0 \\ 0&1&0&0&0&1&0&0&0 \\ 1&0&0&0&1&0&1&0&0 \\ 0&1&0&1&0&1&0&1&0 \\ 0&0&1&0&1&0&0&0&1 \\ 0&0&0&1&0&0&0&1&0 \\ 0&0&0&0&1&0&1&0&1 \\ 0&0&0&0&0&1&0&1&0 \end{bmatrix} \begin{bmatrix} Y_1 \\ Y_2 \\ Y_3 \\ Y_4 \\ Y_5 \\ Y_6 \\ Y_7 \\ Y_8 \\ Y_9 \end{bmatrix} = \begin{bmatrix} Y_2 + Y_4 \\ Y_1 + Y_3 + Y_5 \\ Y_2 + Y_6 \\ Y_1 + Y_5 + Y_7 \\ Y_2 + Y_4 + Y_6 + Y_8 \\ Y_3 + Y_5 + Y_9 \\ Y_4 + Y_8 \\ Y_5 + Y_7 + Y_9 \\ Y_6 + Y_9 \end{bmatrix}$$

El operador de retardo espacial desplaza y agrega las masas de las regiones contiguas a la región analizada.

Este resultado crea un problema de interpretación con la dimensión de la variable resultado. Si se desea preservar la dimensión, basta con imponer la condición de que la suma de los elementos de cada fila sea 1. Se puede entonces tomar cada fila de W y dividir cada componente entre la suma de todas las componentes de la fila respectiva; este proceso da lugar a la “matriz de contactos normalizada” y supone, no la agregación de masas en el espacio, sino hallar el promedio de estas.

$$W_1Y = \begin{bmatrix} 0 & \frac{1}{2} & 0 & \frac{1}{2} & 0 & 0 & 0 & 0 & 0 \\ \frac{1}{3} & 0 & \frac{1}{3} & 0 & \frac{1}{3} & 0 & 0 & 0 & 0 \\ 0 & \frac{1}{2} & 0 & 0 & 0 & \frac{1}{2} & 0 & 0 & 0 \\ \frac{1}{3} & 0 & 0 & 0 & \frac{1}{3} & 0 & \frac{1}{3} & 0 & 0 \\ 0 & \frac{1}{4} & 0 & \frac{1}{4} & 0 & \frac{1}{4} & 0 & \frac{1}{4} & 0 \\ 0 & 0 & \frac{1}{3} & 0 & \frac{1}{3} & 0 & 0 & 0 & \frac{1}{3} \\ 0 & 0 & 0 & \frac{1}{2} & 0 & 0 & 0 & \frac{1}{2} & 0 \\ 0 & 0 & 0 & 0 & \frac{1}{3} & 0 & \frac{1}{3} & 0 & \frac{1}{3} \\ 0 & 0 & 0 & 0 & 0 & \frac{1}{2} & 0 & \frac{1}{2} & 0 \end{bmatrix} \begin{bmatrix} Y_1 \\ Y_2 \\ Y_3 \\ Y_4 \\ Y_5 \\ Y_6 \\ Y_7 \\ Y_8 \\ Y_9 \end{bmatrix} = \begin{bmatrix} \frac{1}{2}(Y_2 + Y_4) \\ \frac{1}{3}(Y_1 + Y_3 + Y_5) \\ \frac{1}{2}(Y_2 + Y_6) \\ \frac{1}{3}(Y_1 + Y_5 + Y_7) \\ \frac{1}{4}(Y_2 + Y_4 + Y_6 + Y_8) \\ \frac{1}{3}(Y_3 + Y_5 + Y_9) \\ \frac{1}{2}(Y_4 + Y_8) \\ \frac{1}{3}(Y_5 + Y_7 + Y_9) \\ \frac{1}{2}(Y_6 + Y_9) \end{bmatrix}$$

2.3 Detección de efectos espaciales

2.3.1 Estadístico de contactos (τ)

El análisis de la dependencia espacial se ha desarrollado a través del **estadístico de contactos** que denotamos por τ. Para calcularlo se debe especificar primero W.

$$\tau = \sum_{r=1}^{R} \sum_{\substack{s=1 \\ s \neq r}}^{R} W_{sr} Y_{sr}$$

Donde Y_{sr} es una medida de similaridad entre las características de los puntos s y r. Una medida de similaridad podría ser, por ejemplo:

$Y_{sr} = \left[Y_r - \overline{Y}\right]\left[Y_s - \overline{Y}\right]$

De acuerdo a este estadístico, se puede definir una autocorrelación espacial positiva y una negativa de acuerdo a las siguientes reglas:

1. Autocorrelación espacial positiva: Cuando τ es positivo y alto, las regiones con características similares y entre las cuales existe dependencia espacial tienden a situarse geográficamente próximas unas de otras.

○	○	○	○	○	
○	○	○	○		
○	○	○			
○	○				
○					

2. Autocorrelación espacial negativa: Cuando τ es negativo y alto, las regiones con características similares y entre las cuales existe dependencia espacial tienden a alternarse con otras diferentes, en forma regular sobre el espacio.

○		○		○	
	○		○		○
○		○		○	
	○		○		○
○		○		○	
	○		○		○

En ambas situaciones (autocorrelación espacial positiva y negativa) apuntan a la existencia de una estructura sistemática en la serie.

Suponiendo que los datos fueron generados de forma aleatoria sobre el espacio $Y_r \sim N(\mu_Y, \sigma^2)$ se obtiene que:

$$E(\tau) = E(Y_{sr}) \sum_{\substack{r=1 \\ s=1}}^{R} W_{rs} = E_{r \neq s}\left[(Y_r - Y)(Y_s - Y)\right] S_0 = -\tfrac{S_0}{R}\sigma^2$$

$$V(\tau) = \tfrac{\sigma^4}{R^2}\left(2S_0^2 + S_1 R^2 - R S_1\right)$$

donde $S_0 = \sum_{\substack{r=1 \\ s=1}}^{R} W_{rs}$, $S_1 = \frac{1}{2}\sum_{r=1}^{R}\sum_{s=1}^{R}(W_{rs} + W_{sr})^2$ y $S_2 = \sum_{s=1}^{R}(W_{s.} + W_{.s})^2$

siendo $W_{s.} = \sum_{r=1}^{R} W_{rs}$

Cuando la matriz de contactos es binaria tipificada (las filas suman 1) $S_0 = R$, entonces:

$$E(\tau) = -\sigma^2$$
$$V(\tau) = E(\tau^2) - \left[E(\tau)\right]^2 = \sigma^4\left(S_1 + 2 - \tfrac{S_2}{R}\right)$$

2.3.2 I de Moran

La I de Moran se basa en el estadístico τ visto anteriormente y se define como:

$$I = \frac{N}{S_0} \frac{\sum_{i=1}^{N}\sum_{j=1}^{N} W_{ij}\left(Y_i - \overline{Y}\right)\left(Y_j - \overline{Y}\right)}{\sum_{i=1}^{N}\left(Y_i - \overline{Y}\right)^2}$$

Donde:

W_{ij}: elemento correspondiente a la posición (i, j) de la matriz de contigüidades.

S_0: Suma de los pesos espaciales. $\sum_{i=1}^{N}\sum_{j=1}^{N} W_{ij}$

$\overline{Y}$: valor esperado de la variable Y.

N: Número de regiones.

Cuando se utiliza la matriz de contigüidades estandarizada, la especificación de la I de Moran se convierte en:

$$I = \frac{\sum_{i=1}^{N}\sum_{j=1}^{N} W_{ij}\left(Y_i - \overline{Y}\right)\left(Y_j - \overline{Y}\right)}{\sum_{i=1}^{N}\left(Y_i - \overline{Y}\right)^2}$$

Pues la suma de los elementos de cada fila dará como resultado la unidad, por tanto, el resultado de $\sum_{i=1}^{N}\sum_{j=1}^{N} W_{ij}$ sera N.

La I de Moran se basa en el coeficiente de autocorrelación temporal $\rho(\tau)$, donde se tiene que $\rho(\tau) = \frac{\gamma(\tau)}{V(Y_t)}$; siendo $\gamma(\tau)$ la función de autocovarianza evaluada en el rezago τ de la variable Y_t. La I de Moran utiliza de forma similar este concepto, pero ponderando las autocovarianzas espaciales por el término correspondiente en la matriz de contigüidades estandarizada, para posteriormente dividir el resultado por la varianza de Y.

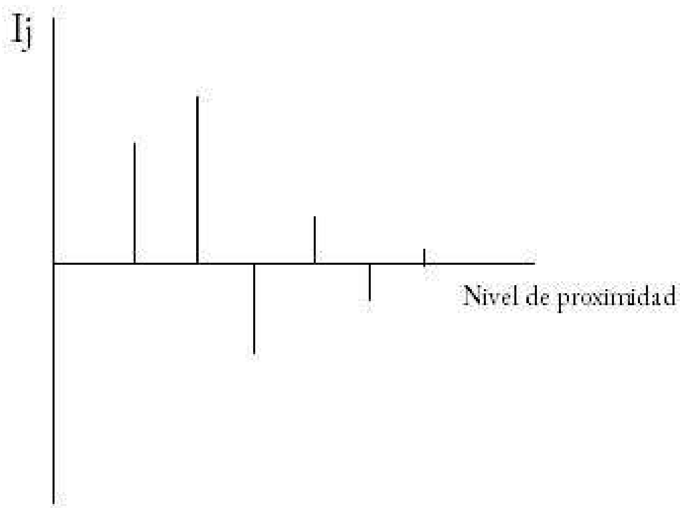
La interpretación de los valores estadísticamente significativos de la I de Moran tipificada sería la siguiente:

1. Valores no significativos del test I estandarizado conducirán a aceptar la hipótesis nula de no autocorrelación espacial.

2. Valores significativos del test I estandarizado y mayores que cero son indicativos de autocorrelación espacial positiva.

3. Valores significativos del test I estandarizado y menores que cero son indicativos de autocorrelación espacial negativa.

Con la colección de valores del estadístico I puede construirse el "correlograma espacial".



Una de las debilidades de este correlograma espacial es que no cuenta con una acotación precisa para los coeficientes, por lo que se parte de la experiencia del investigador y del conocimiento de la región geográfica para interpretarlo.

2.3.3 Test c de Geary

A diferencia de la I de Moran, el test c de Geary no es función de la distancia existente entre cada punto y la media, sino del cuadrado de las distancias entre ellos.

$$c = \frac{(N-1)}{2S_0} \frac{\sum_{i=1}^{N}\sum_{j=1}^{N} W_{ij}(Y_i - Y_j)^2}{\sum_{i=1}^{N}\left(Y_i - \overline{Y}\right)^2}$$

De la forma en que la I de Moran está relacionada con el coeficiente de autocorrelación temporal, la c de Geary está basada en el estadístico d de Durbin y Watson para verificar o descartar la presencia de procesos autorregresivos de primer orden en los residuos de una regresión.

La interpretación de los valores estadísticamente significativos de la c de Geary tipificada sería la siguiente:

1. Valores no significativos del test c estandarizado conducirán a aceptar la hipótesis nula de no autocorrelación espacial.

2. Valores significativos del test c estandarizado y menores que cero son indicativos de autocorrelación espacial positiva.

3. Valores significativos del test c estandarizado y mayores que cero son indicativos de autocorrelación espacial negativa.

2.3.4 Test $G(d)$ de Getis y Ord

Para el cálculo de este estadístico se debe redefinir la matriz de contigüidades W de tal modo que dos regiones serán vecinas si y solo si la distancia entre ellas es menor a una distancia d determinada.

$$G(d) = \frac{\sum_{i=1}^{N}\sum_{j=1}^{N} W_{ij}(d) Y_i Y_j}{\sum_{i=1}^{N}\sum_{j=1}^{N} Y_i Y_j}; \ \text{para } i \neq j$$

La hipótesis nula asociada al estadístico $G(d)$ estandarizado será la ausencia de autocorrelación espacial, mientras que un valor significativo (positivo o negativo) indicará la tendencia a la concentración de valores similares (grandes o pequeños respectivamente).

El test $G(d)$ tiene dos restricciones asociadas a su construcción misma:

1. Solo puede aplicarse a variables positivas, por lo que no puede utilizarse por ejemplo para constatar la presencia de autocorrelación espacial en los residuos de una regresión.

2. La matriz de contigüidades W debe ser simétrica por lo que no puede realizarse ningún tipo de estandarización por filas en ella.

2.4 Dependencia espacial en un modelo de regresión

Para determinar adecuadamente el modelo a utilizar debe analizarse el tipo de dependencia espacial existente en los datos. En caso de omitir erróneamente un retardo espacial en la variable endógena y/o en las exógenas, la dependencia espacial se trasladará directamente al término de perturbación, el cual pasaría a estar correlacionado espacialmente. En este caso la autocorrelación resultante es llamada “autocorrelación espacial sustantiva” y su solución es incluir en el modelo un retardo espacial de la variable sistemática correlacionada espacialmente.

Cuando la dependencia espacial en los residuos no es ocasionada por una omisión errónea del retardo en las variables endógena y/o exógena se considera que existe “autocorrelación espacial residual”.

2.5 Taxonomía de Vayá

Vayá (2000) clasifica las formas funcionales de los modelos espaciales de acuerdo al tipo de dependencia espacial y a su orden de la siguiente manera:

2.5.1 Dependencia espacial sustantiva

De primer orden

1. $Y = \rho W_1 Y + X\beta_1 + W_2 R\beta_2 + \varepsilon$

   $\varepsilon = \lambda W_3 \varepsilon + \mu$

   $\mu \sim N(0, \Omega)\,;\;\; \Omega_{ii} = h_i(Z\alpha)\,;\;\; h_i > 0$

2. $Y = \rho W_1 Y + X\beta_1 + \varepsilon$

   $\mu \sim N(0, \sigma^2 I)$

3. $Y = \rho W_1 Y + X\beta_1 + \varepsilon$

$\varepsilon = \lambda W_3 \varepsilon + \mu$

$\mu \sim N(0, \sigma^2 I)$

4. $Y = \rho W_1 Y + X\beta_1 + W_2 R\beta_2 + \varepsilon$

$\mu \sim N(0, \sigma^2 I)$

5. $Y = X\beta_1 + W_2 R\beta_2 + \varepsilon$

$\mu \sim N(0, \sigma^2 I)$

6. $Y = \rho W_1 Y + X\beta_1 + \varepsilon$

$\varepsilon = \theta W_3 \mu + \mu$

$\mu \sim N(0, \sigma^2 I)$

De ordenes superiores

$$\begin{aligned} Y &= \rho_1 W_1 Y + \rho_2 W_2 Y + ... + \rho_p W_p Y + X\beta_1 + \varepsilon \\ \varepsilon &= \theta_1 W_1 \mu + \theta_2 W_2 \mu + ... + \theta_q W_q \mu + \mu \\ \mu &\sim N(0, \sigma^2 I) \end{aligned}$$

2.5.2 Dependencia espacial residual

De primer orden

Estructura autorregresiva SAR

1. $Y = \rho W_1 Y + X\beta_1 + W_2 R\beta_2 + \varepsilon$

$\varepsilon = \lambda W_3 \varepsilon + \mu$

$\mu \sim N(0, \Omega); \;\; \Omega_{ii} = h_i(Z\alpha); \;\; h_i > 0$

2. $Y = X\beta_1 + \varepsilon$

$\varepsilon = \lambda W_3 \varepsilon + \mu$

$\mu \sim N(0, \sigma^2 I)$

3. $Y = \rho W_1 Y + X\beta_1 + \varepsilon$

$\varepsilon = \lambda W_3 \varepsilon + \mu$

$\mu \sim N(0, \sigma^2 I)$

Estructura de media móvil SMA

1. $Y = X\beta_1 + \varepsilon$

$\varepsilon = \theta W_3 \mu + \mu$

$\mu \sim N(0, \sigma^2 I)$

2. $Y = \rho W_1 Y + X\beta_1 + \varepsilon$

$\varepsilon = \theta W_3 \mu + \mu$

$\mu \sim N(0, \sigma^2 I)$

Modelo de componentes del error espacial

$$Y = X\beta + \varepsilon$$

$$\varepsilon = W\upsilon + \psi$$

$$\psi \sim N(0, \sigma^2 I)$$

De ordenes superiores

Estructura autorregresiva SAR

$$Y = X\beta_1 + \varepsilon$$

$$\varepsilon = \lambda_1 W_1 \varepsilon + \lambda_2 W_2 \varepsilon + ... + \lambda_p W_p \varepsilon + \mu$$

$$\mu \sim N(0, \sigma^2 I)$$

Estructura de media móvil SMA

$$Y = \rho_1 W_1 Y + \rho_2 W_2 Y + ... + \rho_p W_p Y + X\beta_1 + \varepsilon$$

$$\varepsilon = \theta_1 W_1 \mu + \theta_2 W_2 \mu + ... + \theta_q W_q \mu + \mu$$

$\mu \sim N(0, \sigma^2 I)$ Así mismo, Chasco (2003) hace un recuento de las taxonomías propuestas por Anselin (2001) y Florax y Folmer (1992). En el siguiente apartado se desarrollarán aspectos importantes relativos a algunos modelos, los cuales pueden hacerse extensivos a la totalidad de ellos, conservando, claro está, las particularidades asociadas a cada uno de ellos y que constituyen la base fundamental de la aplicabilidad de los modelos de dependencia espacial.

2.6 Taxonomía de Florax y Folmer

Florax y Folmer (1992) clasificaron las diferentes variantes de los modelos en 8 modelos básicos, los cuales parten del llamado Modelo General de Regresión Espacial.

2.6.1 Modelo General de Regresión Espacial

También es llamado Modelo Mixto Regresivo de Regresión Espacial, con Perturbaciones Aleatorias Autorregresivas y Heteroscedásticas.

$$Y = \rho W_1 Y + X\beta_1 + W_2 R\beta_2 + U$$
$$U = \lambda W_3 U + \varepsilon$$
$$\varepsilon \sim N(0, \Omega)\,;\;\; \Omega_{ii} = h_i(z\alpha) = \sigma^2 + z_{i1}\alpha_1 + ... + z_{ip}\alpha_p \;\text{ con }\; h_i > 0$$

Donde:

Y: Vector de (Nx1) observaciones de la variable endógena.
W_1: Matriz de NxN de pesos espaciales de la variable endógena.

ρ: Coeficiente autorregresivo espacial (escalar), que recoge la intensidad de las interdependencias entre las observaciones muestrales.

X: Matriz de (K_1xN). K_1 variables exógenas y N observaciones.

R: Matriz de (K_2xN). K_2 variables exogenas espacialmente retardadas, que pueden o no coincidir con las variables incluidas en X.

W_2: Matriz de NxN de pesos espaciales correspondiente a las variables exógenas espacialmente retardadas.

β_1, β_2: Vectores (K_1x1 y K_2x1) de parámetros de las variables exogenas.

U: Vector (Nx1) de perturbaciones aleatorias autorregresivas (de primer orden) y heteroscedásticas, siendo los elementos de la diagonal principal de la matriz de covarianzas (Ω) funcion $P+1$ de variables exogenas de Z.

α: Vector (Px1) asociado a los terminos no constantes de la matriz Z.

W_3: Matriz de pesos espaciales de la perturbación aleatoria U.

λ: Parámetro autorregresivo (escalar) asociado al retardo espacial W_3U.

ε: Vector de perturbaciones aleatorias, ruido blanco.

Luego, el vector de parámetro es:

$$\Theta = [\rho, \beta_1, \beta_2, \theta, \lambda, \alpha]$$

2.6.2 Modelo básico de regresión lineal

$$Y = X\beta_1 + U$$
$$U \sim N(0, \sigma^2 I)$$

Algunos autores consejan anular el efecto de dependencia espacial a traves de métodos de filtrado espacial, de forma que un modelo con problemas de dependencia espacial pueda ser transformado en el modelo basico de regresion lineal. Getis (1995,2002) propone una transformación de las variables con problemas de dependencia espacial apartir del estadistico $G_i(d)$ visto al principio de esta sección, de forma que se obtenga, por ejemplo, para la variable endogena (Y) espacialmente autocorrelacionada, otra variable (Y^*) sin tendencia espacial, del modo siguiente:

$$Y_i^* = \frac{Y_i\left(\frac{W_i}{N-1}\right)}{G_i(d)};\ \ G_i\left(d\right) = \frac{\sum\limits_{j=1}^{N} W_{ij}(d) Y_j}{\sum\limits_{j=1}^{N} Y_j} \ \ \text{para } j \neq i$$

Donde $W_i = \sum\limits_j W_{ij}\left(d\right)$, por lo que el coeficiente $\frac{W_i}{N-1}$ es la media o esperanza matemática del estadístico G_i. La diferencia $Y_i - Y_i^*$ sería una nueva variable que contiene los efectos espaciales presentes en Y_i. Esta transformación, aplicada a las variables afectadas daría lugar a nuevas variables (Y^*, X_k^*) sin problemas de autocorrelacion espacial que podrían formar parte del modelo básico de regresión líneal.

2.6.3 Modelo mixto autorregresivo de regresión espacial o modelo del retardo espacial

$Y = \rho W_1 Y + X\beta_1 + U$

$U \sim N\left(0, \sigma^2 I\right)$

Esta especificación es adecuada en casos en los que el valor que adopta una variable en un punto o región depende no solo de condiciones internas (valores de otras variables explicativas en el mismo punto región), sino también del valor de esa misma variable en otras regiones vecinas, incumpliéndose así el principio de independencia entre las observaciones muestrales.

2.6.4 Modelo mixto regresivo cruzado de regresión espacial

$Y = X\beta_1 + W_2 R\beta_2 + U$
$U \sim N\left(0, \sigma^2 I\right)$

En este modelo, el efecto de dependencia espacial es también sustantivo, dado que se encuentra presente, en forma de retardo espacial, en una o varias variables exógenas del modelo (no en la dependiente, como en el modelo del retardo espacial).

2.6.5 Modelo mixto autorregresivo regresivo cruzado de regresión espacial

$$Y = \rho W_1 Y + X\beta_1 + W_2 R\beta_2 + U$$
$$U \sim N\left(0, \sigma^2 I\right)$$

Esta especificación es producto de la combinación del modelo del retardo espacial y del modelo mixto regresivo cruzado de regresión espacial.

Un caso particular de esta especificación es el llamado "Modelo Durbin Espacial", en el que se plantea la siguiente restricción:

$$H_0 : \beta_2 = \rho\beta_1$$

2.6.6 Modelo de regresión con dependencia espacial en la perturbación aleatoria o modelo del error espacial

$$Y = X\beta_1 + U$$
$$U = \lambda W_3 U + \varepsilon$$
$$\varepsilon \sim N\left(0, \sigma^2 I\right)$$

El fenómeno de autocorrelación espacial está únicamente presente en la perturbación aleatoria, normalmente en forma autorregresiva de orden 1, SAR(1).

La consideración explícita de un esquema de dependencia espacial en el término de la perturbación aleatoria, resulta adecuada cuando se omiten en un modelo variables que se hallen correlacionadas espacialmente, o también cuando se producen errores de medida.

De esta expresión podrían derivarse otras, por ejemplo un SAR(2):

$$Y = X\beta_1 + U$$
$$U = \lambda_1 W_3^1 U + \lambda_2 W_3^2 U + \varepsilon$$
$$\varepsilon \sim N\left(0, \sigma^2 I\right)$$

Donde W_3^1 y W_3^2 son las matrices de interacción espacial, correspondientes a la perturbación aleatoria de orden 1 y 2 respectivamente.

Podrían además derivarse procesos de medias móviles en la perturbación aleatoria, SMA(q). Por ejemplo:

SMA(1):

$Y = X\beta_1 + U$
$U = \theta W_3 \varepsilon + \varepsilon$
$\varepsilon \sim N(0, \sigma^2 I)$

2.6.7 Modelo mixto autorregresivo de regresión espacial con perturbación aleatoria espacialmente autorregresiva

$$Y = \rho W_1 Y + X\beta + U \tag{2.1}$$

$$U = \lambda W_2 U + \varepsilon \tag{2.2}$$

$$\varepsilon \sim N(0, \Omega)$$

Los elementos de la diagonal de la matriz de covarianzas $\Sigma_U = \Omega$ son:

$$\Omega_{ii} = h_i(z\alpha) = \sigma^2 + z_{i1}\alpha_1 + ... + z_{ip}\alpha_p \; con \; h_i > 0 \tag{2.3}$$

Este modelo surge de la combinación de los modelos del retardo espacial y del error espacial, siendo su desarrollo importante para ilustrar la forma general para hallar los estimadores propios de estos modelo. El desarrollo de los estimadores es el siguiente:

De (2.1)

$$(I - \rho W_1) Y = X\beta + U \tag{2.4}$$

$\Rightarrow AY = X\beta + U$

donde $A = (I - \rho W_1)$

De (2.2)

$$(I - \lambda W_2)\, U = \varepsilon$$

$$\Rightarrow BU = \varepsilon \tag{2.5}$$

donde $B = (I - \lambda W_2)$.

reemplazando (2.5) en (2.4) se obtiene:

$(I - \rho W_1)\, Y = X\beta + B^{-1}\varepsilon$
$\Rightarrow (I - \rho W_1)\, Y = X\beta + (I - \lambda W_2)^{-1}\, \varepsilon$
$\Rightarrow (I - \lambda W_2)\,(I - \rho W_1)\, Y - (I - \lambda W_2)\, X\beta = \varepsilon$
$\Rightarrow BAY - BX\beta = \varepsilon$

Por (1.1) se sabe que $f(Y) = f(\varepsilon)\left|\frac{\partial\varepsilon}{\partial Y}\right|$. Se sabe además que $\frac{\partial\varepsilon}{\partial Y} = BA$, pues $\varepsilon = BAY = BX\beta$.

Como $\varepsilon \sim N(0, \sigma^2 I)$, entonces la función de densidad de probabilidad de ε es:

$f(\varepsilon) = \frac{1}{|\Omega|^{1/2}(2\pi)^{N/2}} e^{-(1/2)\varepsilon'\Omega^{-1}\varepsilon}$

Luego,

$f(Y) = \frac{1}{|\Omega|^{1/2}(2\pi)^{N/2}} e^{-(1/2)\varepsilon'\Omega^{-1}\varepsilon}\, |BA|$

Para hallar los estimadores de los parámetros por el método de Máxima Verosimilitud es conveniente transformar esta función de densidad de probabilidad en otra forma más funcional.

$f(Y) = \frac{1}{|\Omega|^{1/2}(2\pi)^{N/2}} e^{-(1/2)\varepsilon'\Omega^{-1}\varepsilon}\, |BA|$
$\Rightarrow f(Y) = |\Omega|^{-1/2}\,(2\pi)^{-N/2}\, e^{-(1/2)\varepsilon'\Omega^{-1}\varepsilon}\, |BA|$
$\Rightarrow f(Y) = |\Omega|^{-1/2}\,(2\pi)^{-N/2}\, e^{-(1/2)[B(AY-X\beta)]'\Omega^{-1}[B(AY-X\beta)]}\, |B|\,|A|$
$\Rightarrow f(Y) = |\Omega|^{-1/2}\,(2\pi)^{-N/2}\, e^{-(1/2)[B(AY-X\beta)]'\Omega^{-1/2}\Omega^{-1/2}[B(AY-X\beta)]}\, |B|\,|A|$
$\Rightarrow f(Y) = |\Omega|^{-1/2}\,(2\pi)^{-N/2}\, e^{-(1/2)\left[\Omega'^{-1/2}B(AY-X\beta)\right]'\Omega^{-1/2}B(AY-X\beta)}\, |B|\,|A|$

$\Rightarrow f(Y) = |\Omega|^{-1/2} (2\pi)^{-N/2} e^{-(1/2)\left[\Omega^{-1/2}B(AY-X\beta)\right]'\Omega^{-1/2}B(AY-X\beta)} |B| |A|$

Sea $V = \Omega^{-1/2}B(AY - X\beta)$, entonces:

$f(Y) = |\Omega|^{-1/2} (2\pi)^{-N/2} e^{-(1/2)V'V} |B| |A|$

La función de log-verosimilitud es entonces:

$$L = \ln[f(Y)] = -\frac{N}{2}\ln|\pi| - \frac{1}{2}\ln|\Omega| + \ln|B|\ln|A| - \frac{1}{2}V'V - \frac{N}{2}\ln|2| \quad (2.6)$$

y el problema de optimización asociado es:

$$\max_{\beta,\beta',\rho,\lambda,\alpha_p} (L)$$

y las condiciones de primer orden para la optimización son:

$\frac{\partial L}{\partial \beta} = 0$

$\frac{\partial L}{\partial \beta'} = 0$

$\frac{\partial L}{\partial \rho} = 0$

$\frac{\partial L}{\partial \lambda} = 0$

$\frac{\partial L}{\partial \alpha_p} = 0$

siendo α_p el término que representa cada uno de los coeficientes de la ecuación (2.3) asociada a la diagonal de la matriz de covarianzas Ω.

Acevedo y Gómez (2008) muestran el desarrollo algebraico que permite llegar a:

$\frac{\partial L}{\partial \beta} = V'\Omega^{-1/2}BX$

$\frac{\partial L}{\partial \beta'} = \Omega^{-1/2} BXV$

$\frac{\partial L}{\partial \rho} = -Tr\left(A^{-1}\right) W_1 + V'\Omega^{-1/2} BW_1 Y$

$\frac{\partial L}{\partial \lambda} = -Tr\left(B^{-1}\right) W_2 + V'\Omega^{-1/2} W_2 \left(AY - X\beta\right)$

$\frac{\partial L}{\partial \alpha_p} = -\frac{1}{2} Tr\left(\Omega^{-1}\right) H_p + \frac{1}{2} V'\Omega^{-3/2} H_p B \left(AY - X\beta\right)$

donde $\frac{\partial \Omega}{\partial \alpha_p} = H_p$

Resolviendo el anterior sistema de ecuaciones para cada caso particular se obtienen los estimadores del caso.

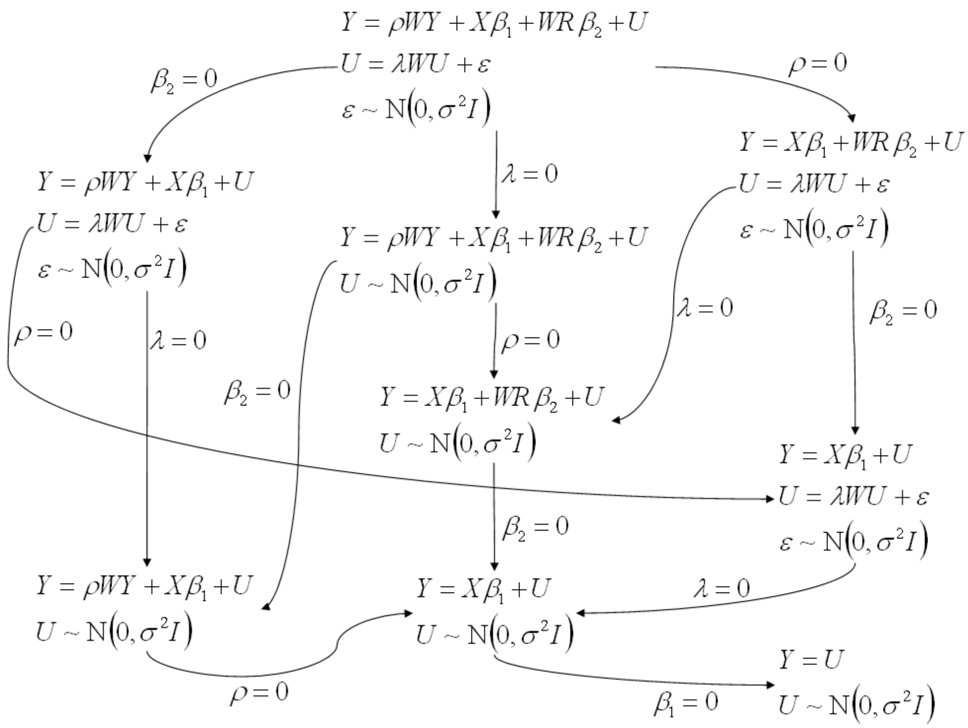
2.6.8 Modelo Mixto Autorregresivo de Regresión Espacial, con perturbaciones que incorporan un esquema de dependencia de media móvil

$Y = \rho_1 W_1^1 Y + \rho_2 W_1^2 Y + ... + \rho_p W_p^p Y + X\beta_1 + \varepsilon + \theta_1 W_3^1 Y + \theta_2 W_3^2 Y + ... + \theta_q W_3^q Y$
$\varepsilon \sim N\left(0, \sigma^2 I\right)$

Esta especificación se denomina $SARMA(p,q)$. Algunos autores lo denominan $SARMAX$ debido a la presencia de variables exógenas.

Las matrices de interacciones espaciales W_i^j llevan expresado como superíndice el orden de contigüidad. En estos casos debe ponerse especial cuidado en evitar el problema de circularidad y redundancia propio de las matrices de ordenes superiores a uno (1).

El paso de una especificación a otra puede resumirse mediante el siguiente diagrama:



2.7 Taxonomía de Anselin

2.7.1 Modelo autorregresivo de regresión espacial de orden 1. SAR(1)

$$Y = \rho WY + \varepsilon$$
$$\varepsilon \sim N\left(0, \sigma^2 I\right)$$

Es la especificación de dependencia espacial más sencilla y hace referencia a una situación en la cual los valores que adopta una variable están determinados por la localización geográfica de las mismas. Esta especificación es importante en la medida que servirá para desarrollar el algoritmo de agregación espacial presentado en el capítulo 3 y para hacer inferencia sobre la tasa de defunción en adultos no pensionados presentada en el capítulo 4.

Esta especificación, por su aparente simplicidad, se utilizará para demostrar uno de los problemas que conlleva utilizar el método de Mínimos Cuadrados Ordinarios para hallar estimadores insesgados y consistentes de ρ.

Este modelo implica que:

$\varepsilon = (I - \rho W) Y$

Dado $\varepsilon \sim N(0, \sigma^2 I)$, la función de log-verosimilitud es:

$\ell(\rho, \sigma^2) = -\frac{N}{2} \ln(2\pi\sigma^2) - \frac{1}{2}\sigma^2 Y'A'AY + \ln|A|$

Demostración. Por especificación del modelo se sabe que $AY = \varepsilon$ y por (1.1) que $f(Y) = f(\varepsilon)\left|\frac{\partial v}{\partial Y}\right|$. Utilizando la misma deducción realizada para encontrar la ecuación (2.6) se obtiene que

$f(Y) = (\sigma^2)^{-N/2} (2\pi)^{-N/2} e^{-(1/2)\varepsilon'\sigma^2\varepsilon} |A|$

$\Rightarrow \ln[f(Y)] = \ell(\rho, \sigma^2) = -\frac{N}{2} \ln(2\pi\sigma^2) - \frac{1}{2}\sigma^2 Y'A'AY + \ln|A| \quad \blacksquare$

Esta función de log-verosimilitud será utilizada únicamente para encontrar un estimador de σ^2, pues la estimación de ρ se hará por el método de Mínimos Cuadrados Ordinarios.

Derivando $\ell(\rho, \sigma^2)$ con respecto a σ^2 se obtiene:

$\frac{\partial \ell}{\partial \sigma^2} = \frac{\partial\left(-\frac{N}{2}\ln\left(2\pi\sigma^2\right) - \frac{1}{2}\sigma^2 Y'A'AY + \ln|A|\right)}{\partial \sigma^2} = -\frac{N}{2}\left(\frac{1}{\sigma^2}\right) + \frac{1}{2\sigma^2} Y'A'AY$

Igualando a cero $\sigma^2 = N^{-1} Y'A'AY$

Proposición 17 *Un estimador de ρ es la solución de la siguiente ecuación:*
$0 = Y'WY - \rho Y'W'WY = Y'(I - \rho W') WY = e'WY$
donde $e = (I - \rho W) Y$

Demostración. Por la especificación del modelo se sabe que $\varepsilon = (I - \rho W)\, Y$

$\Rightarrow \varepsilon' = Y'\,(I - \rho W')$

entonces

$\varepsilon'\varepsilon = Y''\,(I - \rho W')'\,(I - \rho W)\, Y = Y'\,(I - \rho W - \rho W' + \rho^2 W'W)\, Y$

$\Rightarrow \varepsilon'\varepsilon = (Y - \rho Y'W - \rho Y'W' + \rho^2 Y'W'W)\, Y$

$\Rightarrow \varepsilon'\varepsilon = Y'Y - \rho Y'WY - \rho Y'W'Y + \rho^2 Y'W'WY$

Derivando parcialmente $\varepsilon'\varepsilon$ con respecto a Y para hallar el mínimo se obtiene:

$\frac{\partial(\varepsilon'\varepsilon)}{\partial\rho} = -YWY - Y'W'Y + 2\rho Y'W'WY = 0$

$\Rightarrow -2Y'WY + 2\rho Y'W'WY = 0$

$$\Rightarrow Y'WY + \rho Y'W'WY = 0 \tag{2.7}$$

■

Proposición 18 *El estimador de ρ hallado por el Método de Mínimos Cuadrados Ordinarios será insesgado si $E\,(\varepsilon'WY) = 0$.*

Demostración. Despejando ρ de (2.7) se obtiene:

$\widehat{\rho} = \frac{Y'WY}{Y'W'WY}$

$\Rightarrow \widehat{\rho} = \frac{(\rho WY+\varepsilon)'WY}{Y'W'WY} = \frac{(Y'W'\rho+\varepsilon)WY}{Y'W'WY} = \frac{\rho Y'W'WY}{Y'W'WY} + \frac{\varepsilon'WY}{Y'W'WY}$

$\Rightarrow \widehat{\rho} = \rho + \frac{\varepsilon'W(I-\rho W)^{-1}\varepsilon}{Y'W'WY}$

y tomando esperanzas a ambos lados;
$E\,(\widehat{\rho}) = E\left(\rho + \frac{\varepsilon'W(I-\rho W)^{-1}\varepsilon}{Y'W'WY}\right) = \rho + \frac{E\left(\varepsilon'W(I-\rho W)^{-1}\varepsilon\right)}{Y'W'WY} = \rho + \frac{E(\varepsilon'WY)}{Y'W'WY}$

Por lo tanto, para que $\widehat{\rho}$ sea insesgado, es necesario que $E\,(\varepsilon'WY) = 0$ ■

Para analizar la consistencia del estimador se utilizará la siguiente proposición:

Proposición 19 $E(\varepsilon' WY) = Tr\left(W'(I-\rho W)^{-1}\right) = 0$

Demostración. Se sabe que $E(\varepsilon' WY) = 0$

Al tratarse de números reales, es posible tomar la traza a ambos lados de la ecuación de la siguiente forma:

$Tr\left[E(\varepsilon' WY)\right] = 0$

$\Rightarrow E\left[Tr(\varepsilon' WY)\right] = 0$

por (1.2) puede afirmarse que

$E\left[Tr(\varepsilon' WY)\right] = E\left[Tr(WY\varepsilon')\right]$

Entonces, $E\left[Tr(WY\varepsilon')\right] = 0$

$\Rightarrow E\left[Tr(WY\varepsilon')\right] = 0$

$\Rightarrow E\left[Tr\left(W'(I-\rho W)^{-1}\varepsilon\varepsilon'\right)\right] = 0$

$\Rightarrow Tr\left[E\left(W'(I-\rho W)^{-1}\varepsilon\varepsilon'\right)\right] = 0$

$$\Rightarrow Tr\left[W'(I-\rho W)^{-1}E(\varepsilon\varepsilon')\right] = 0 \tag{2.8}$$

Como $E(\varepsilon\varepsilon') = 0$, por ser ε ruido blanco[1], entonces

$$E(\varepsilon' WY) = Tr\left[W'(I-\rho W)^{-1}\right] = 0 \tag{2.9}$$

■

Para garantizar entonces la consistencia del estimador se requiere que: $\underset{n\to\infty}{pLim}\frac{\varepsilon' WY}{Y'W'WY} = 0$

esto porque

[1] $E\left[(\varepsilon - E(\varepsilon))(\varepsilon - E(\varepsilon))'\right] = 0$

$\widehat{\rho} = \rho + \frac{\varepsilon' WY}{Y'W'WY}$

entonces

$V\left(\widehat{\rho}\right) = V\left(\rho\right) + V\left(\frac{\varepsilon' WY}{Y'W'WY}\right)$
$\Rightarrow V\left(\widehat{\rho}\right) = V\left(\frac{\varepsilon' WY}{Y'W'WY}\right)$

Por definición de consistencia $\underset{n\to\infty}{Lim} V\left(\widehat{\rho}\right) = 0$, entonces se tiene que:

$$\underset{n\to\infty}{Lim} V\left(\frac{\varepsilon' WY}{Y'W'WY}\right) = 0$$

$$\Rightarrow \underset{n\to\infty}{pLim} \frac{\varepsilon' WY}{Y'W'WY} = 0$$

El que $\underset{n\to\infty}{pLim} \frac{\varepsilon' WY}{Y'W'WY} = 0$ (para que haya consistencia) también implica, por ecuación (2.9), que:

$$\underset{n\to\infty}{Lim} \frac{Tr\left[W'(I-\rho W)^{-1}\right]}{Tr\left[(I-\rho W)^{-1} W'W (I-\rho W)^{-1}\right]} = 0$$

Esto porque,

$$\underset{n\to\infty}{pLim} \frac{\varepsilon' WY}{Y'W'WY} = \underset{n\to\infty}{Lim} \frac{E(\varepsilon' WY)}{E(Y'W'WY)} = \underset{n\to\infty}{Lim} \frac{Tr\left[W'(I-\rho W)^{-1} E(\varepsilon\varepsilon')\right]}{E\left[\varepsilon'(I-\rho W)^{-1} W'W (I-\rho W)^{-1} \varepsilon\right]}$$

Utilizando la misma deducción de (2.8) se obtiene:

$$\underset{n\to\infty}{pLim} \frac{\varepsilon' WY}{Y'W'WY} = \underset{n\to\infty}{Lim} \frac{Tr\left[W'(I-\rho W)^{-1}\right] E(\varepsilon\varepsilon')}{Tr\left[(I-\rho W)^{-1} W'W (I-\rho W)^{-1}\right] E(\varepsilon\varepsilon')}$$

$$\Rightarrow \underset{n\to\infty}{pLim} \frac{\varepsilon' WY}{Y'W'WY} = \underset{n\to\infty}{Lim} \frac{Tr\left[W'(I-\rho W)^{-1}\right]}{Tr\left[(I-\rho W)^{-1} W'W (I-\rho W)^{-1}\right]}$$

Bajo el supuesto de consistencia, $\underset{n\to\infty}{pLim} \frac{\varepsilon' WY}{Y'W'WY} = 0$, entonces se debe cumplir que:

$$\underset{n\to\infty}{Lim} \frac{Tr\left[W'(I-\rho W)^{-1}\right]}{Tr\left[(I-\rho W)^{-1} W'W (I-\rho W)^{-1}\right]} = 0$$

Ord (1975) plantea que para remover la inconsistencia se debe encontrar una matriz G tal que:

$$E\left[\varepsilon' W' G Y\right] = 0$$

Y que se puede demostrar que G debe ser una función de ρ y que la elección mas simple es $G = I - \rho W = A$

El estimador $\widehat{\rho}$ por mínimos cuadrados ordinarios es una solución de la forma cuadrática en ρ, $Y'AWAY = 0$, porque:

$$Y'WY - \rho Y'W'WY = 0$$
$$Y'\left(W - \rho W'W\right)Y = 0$$
$$Y'\left(I - \rho W'\right)WY = 0$$
$$Y'A'WY = 0$$

Incluyendo a $G = I - \rho W = A$ se obtiene

$$Y'A'WAY = 0$$

De esta forma puede hallarse un estimador consistente de ρ.

2.7.2 Modelo básico de regresión lineal

$$Y = X\beta + U$$
$$U \sim N\left(0, \sigma^2 I\right)$$

Esta especificación será correcta para modelar una variable con dependencia espacial, solo en el caso en el que este efecto esté explicado en su totalidad por las variables exógenas del modelo. La significancia de las variables exógenas explicativas de la dependencia espacial implica la no presencia de autocorrelación espacial residual.

2.7.3 Modelo de regresion con dependencia espacial en la perturbación aleatoria o modelo del error espacial

1. $Y = X\beta + U$

   $U = \lambda WU + \varepsilon$; con $\varepsilon \sim N\left(0, \sigma^2 I\right)$

2. $Y = X\beta + U$

   $U = \theta W\varepsilon + \varepsilon$; con $\varepsilon \sim N\left(0, \sigma^2 I\right)$

Esta es la especificación más utilizada cuando el modelo básico de regresión lineal resulta ineficaz para explicar el efecto espacial presente en la variable endógena, trasladando entonces este efecto hacia los términos de error.

Anteriormente se explicaron las razones por las cuales no es apropiado estimar los parámetros del modelo por Mínimos Cuadrados Ordinarios. Utilizaremos ahora el método de máxima verosimilitud para hallar los estimadores necesarios.

2.7.4 Modelo mixto autorregresivo de regresión espacial o modelo del retardo espacial

$Y = \rho WY + X\beta + U$

$U \sim N\left(0, \sigma^2 I\right)$

El modelo del retardo espacial también puede ser conveniente para explicar la dependencia espacial de la variable endógena; basta con rezagar espacialmente esta variable y agregarla como explicativa del modelo junto con otras variables exógenas que puedan explicar apropiadamente la variabilidad de la endógena.

Para hallar un estimador de ρ se puede utilizar el Método de Máxima Veromilitud y posteriormente, si se quiere, reemplazarlo en la especificación y estimar β por Mínimos Cuadrados Ordinarios. Esto es posible porque teniendo un estimador de ρ puede reemplazarse este último en la especificación y convertir así el modelo espacial en un modelo lineal general.

Para demostrarlo se hace $Z = (I - \rho W)\, Y$ con lo que el modelo se convierte en el lineal general:

$$Z = X\beta + U$$
$$U \sim N\left(0, \sigma^2 I\right)$$

Estimando ρ por el Método de Máxima Verosimilitud y reemplazando en Z se obtiene $\widehat{Z}$

Se sabe que el estimador de β tanto por Mínimos Cuadrados Ordinarios como por Máxima Verosimilitud para el modelo lineal general es:

$$\widehat{\beta} = (X'X)^{-1}\, X'\widehat{Z}$$
$$\widehat{\beta} = (X'X)^{-1}\, X'\,(I - \widehat{\rho}W)\, Y$$

Con esto demostramos que el estimador $\widehat{\beta}$ es insesgado.

Capítulo 3

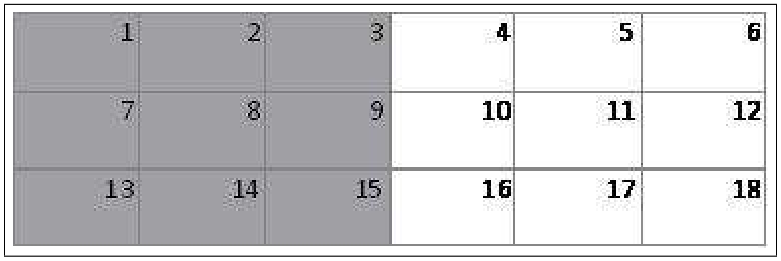
Autocorrelación Espacial Espuria

Uno de los aspectos relevantes en la econometría de series de tiempo está relacionado con las regresiones espurias, fenómeno que se puede originar cuando se consideran regresiones entre series de tiempo no estacionarias, procedimiento cuyo fundamento formal ha sido desarrollado y divulgado ampliamente.

La *autocorrelación espacial espuria* puede ser definida como el fenómeno en el cual existe una aparente dependencia espacial entre las regiones, pero que en realidad es producida por la consideración de lo que en la práctica es una sola región como dos o más regiones independientes. Este fenómeno no ha recibido un tratamiento especial por parte de la comunidad científica.

El fenómeno de autocorrelación espacial espuria es factible que se presente en aquellos espacios en los cuales se tenga una alta densidad de regiones y las cuales, en la práctica, no representan cada una a una región claramente diferenciada de las demás; sino que por el contrario la unión de algunas de ellas conforman en realidad una sola región.

Para ilustrar el problema bajo consideración, se construyó un espacio dividido en dos regiones con procesos generadores de datos muy diferentes: Y_{i1} con $i = 1, 2, 3, 7, 8, 9, 13, 14, 15$ y Y_{i2} con $i = 4, 5, 6, 10, 11, 12, 16, 17, 18$.



El hecho de tener 9 regiones representando las mismas características hace que los test, desde los más complejos hasta el más sencillo, indiquen la presencia de dependencia espacial aunque, en la práctica, la autocorrelación espacial sea espuria al no depender el comportamiento de una región de sus vecinas, sino al ser ella y sus vecinas en realidad una sola región con un mismo proceso generador de datos.

Para detectar y corregir la presencia de autocorrelación espacial espuria se desarrolló un algoritmo que permite hallar la mejor forma de realizar un número de agregaciones dado, con el fin de atacar la hipótesis alternativa de presencia de este fenómeno[1].

El algoritmo que se propone parte de la especificación del modelo SAR(1), analizado en la sección 2.7.1.

$$Y = \rho WY + \varepsilon$$

Cuya prueba de hipótesis para el parámetro de ρ es:

$$H_0 : \rho = 0$$
$$H_a : \rho \neq 0$$

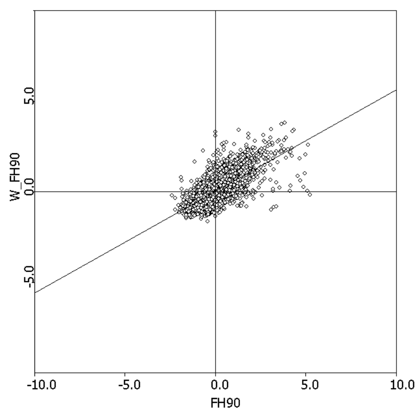
Utilizando el Test de Multiplicadores de Lagrange se obtiene el siguiente estadístico para realizar la prueba:

$$LM - LAG = \frac{\left[\frac{Y'W\widehat{\varepsilon}}{\widehat{\sigma}^2}\right]}{V[\widehat{\rho}]} \sim \chi^2\,(1)$$

Al ser el parámetro ρ la pendiente de la recta, puede inferirse que $\widehat{\rho} \approx 0$ indicará la ausencia de dependencia espacial. Este principio se utilizará como

[1] Para efectos de este trabajo de grado se considera el término "agregación" como la suma de dos o más valores de la variable aleatoria objeto de estudio.

la medida de efectividad de la agregación, pues el objetivo de descartar la dependencia espacial se traduce en reducir tanto como sea posible la pendiente de la recta.



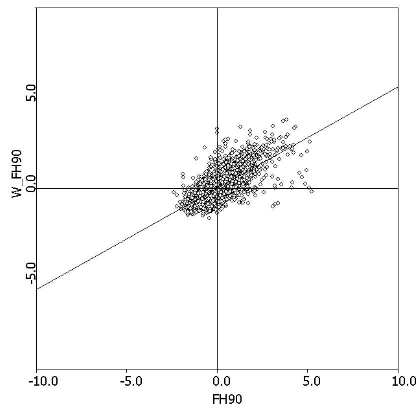
La agregación espacial debe contar con dos reglas básicas:

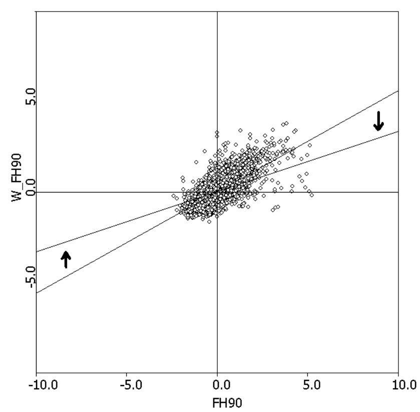
1. Las regiones a agregar deben ser contigüas: Esta restricción se impone pues no tiene sentido plantear que las regiones a agregar pertenecen en la práctica a una sola región, mientras estas estén separadas por dos o más fronteras.

2. Deben tener la misma medida de posición: Las regiones a agregar deben pertenecer al mismo cuartil, decil u otra medida de posición. Esta restricción permite que se conserve al máximo la información de la población y por consiguiente la verdadera dependencia espacial que pueda llegar a existir, pues impide la agregación de regiones con características muy diferentes.

3.1 Desarrollo teórico del algoritmo

Como se analizó anteriormente, detectar y eliminar la posible autocorrelación espacial espuria implica agregar una cantidad dada de regiones, con las restricciones asociadas, tal que la pendiente de la recta definida por $Y = \rho WY$ sea lo más cercana a cero posible pero rotando la recta de regresión en sentido dextrógiro, esto para lograr que algoritmo ataque con más fuerza la hipótesis alternativa de $\rho \neq 0$ al requerir menos esfuerzo para llevar la pendiente de la recta a un valor cercano a cero.





En la segunda sección de este trabajo de grado se detalló el efecto que produce en la variable endógena el operador de retardo espacial. Este efecto es importante porque permite inferir las características que deberán tener las dos regiones a agregar en cuanto a la "calidad" de la agregación espacial que producirían sus vecinos, haciendo claro que un movimiento dextrógiro en la pendiente de la recta se dará con una mayor probabilidad de ocurrencia cuando los vecinos agregados espacialmente sean tan diferentes que produzcan un valor de la variable retardada cercana a cero.

El nuevo punto en el scatter plot producido por la agregación espacial será entonces $(y_1 + y_2, w'y')$, con:

$$w'y' = w_1 y_1 + w_2 y_2 - y_1 - y_2$$

donde $w_i y_i$ es el retardo espacial de la región y_i. El punto $w'y'$ se define de esta forma pues luego de realizar la agregación espacial de las regiones vecinas $(w_1 y_1 + w_2 y_2)$ debe restarse las propias regiones y_1 y y_2 pues son vecinas entre sí.

Las regiones que al momento de agregarse espacialmente producirán puntos más cercanos al eje de las abscisas serán aquellas que teniendo la misma medida de posición estén rodeadas de regiones tan diferentes entre sí, que produzcan un valor cercano a cero en la suma de sus retardos espaciales, es decir que $w_1 y_1 + w_2 y_2 \to 0$[2]

Partiendo de lo anterior es lógico suponer qué las regiones a agregar se ubicarán en las fronteras de las pseudoregiones que agrupan regiones contiguas con la misma medida de posición, pero que cuentan con un grupo de regiones lo más diferentes a ellas posible al otro lado de la frontera.

1	2	3	4	5	6	7	8	9	10
11	12	13	14	15	16	17	18	19	20
21	22	23	24	25	26	27	28	29	30
31	32	33	34	35	36	37	38	39	40
41	42	43	44	45	46	47	48	49	50
51	52	53	54	55	56	57	58	59	60
61	62	63	64	65	66	67	68	69	70
71	72	73	74	75	76	77	78	79	80
81	82	83	84	85	86	87	88	89	90
91	92	93	94	95	96	97	98	99	100

[2] El término $w'y'$ podría tender a cero con una mayor intensidad escogiendo regiones pertenecientes a diferentes medidas de posición cuyos valores individuales sean los suficientemente diferentes como para producir un valor cercano a cero; sin embargo, escoger este tipo de regiones violaría la segunda regla básica de agregación, eliminando no solo la dependencia espacial espuria, sino también la real.

6	6	6	8	8	8	7	7	7	7
6	6	6	8	8	8	7	7	7	7
6	6	6	8	8	8	7	7	7	7
9	9	4	4	1	1	1	10	10	10
4	9	4	4	1	1	1	10	10	10
4	4	4	4	1	1	1	10	10	10
3	3	3	5	5	5	2	2	2	2
3	3	3	5	5	5	2	2	2	2
3	3	3	5	5	5	2	2	2	2
3	3	3	5	5	5	2	2	2	2

Para ilustrar el proceso de construcción se consideran 100 regiones y se dividieron en deciles. Las regiones que se analizarán en este ejemplo serán entonces las que se encuentran a lado y lado de la frontera que separa los deciles 1 y 10 puesto que no existen regiones contiguas con una mayor diferencia en sus medidas de posición, lo que indica que el retardo espacial de las regiones agregadas sería lo más cercano a cero que se podría encontrar en este espacio con subdivisiones del mismo por deciles.

El primer paso del algoritmo será entonces hallar las regiones a analizar, es decir determinar cuáles regiones se encuentran a lado y lado de una frontera que divida dos grupos de regiones los más diferentes entre sí que sea posible.

Para determinar cuáles regiones pertenecen a las grandes regiones a analizar se toma la matriz de contigüidades (W) y se reemplazan los unos (1) por la expresión (3.1); esta conversión dará lugar a una matriz de diferencias W^d. En adelante, por facilidad en la exposición del algoritmo, se asumirá como medida de posición el decil y se denotará como $d(\cdot)$.

$$w_{ij}^{d} = \begin{cases} |d(j) - d(i)| \text{ si } d(j) \neq d(i) \\ \xi \text{ en otro caso; con } \xi \in \mathbb{R}^{-} \end{cases} \tag{3.1}$$

ξ deberá ser cualquier número real negativo, pues de lo contrario podría confundirse en la siguiente fase del algoritmo con los generados por $|d(j) - d(i)|$ cuando $d(j) \neq d(i)$.

Esta expresión w_{ij}^d garantizará la identificación de los vecinos con la máxima diferencia entre sus deciles sin perder la funcionalidad como matriz de contigüidades, pues las entradas iguales a cero (0) continuarán representando regiones sin frontera común.

La máxima diferencia entre deciles será entonces:

$$M^d = \max\left(w_{ij}^d\right)$$

Y las grandes regiones a analizar serán las correspondientes a las filas (o columnas, por simetría de esta matriz) que tengan por lo menos una entrada igual a M^d y sus respectivos vecinos.

Por ejemplo, una matriz

$$W^d = \begin{bmatrix} 0 & 9 & 1 & 0 & 0 & 3 & 7 & 0 & 0 & 0 \\ 9 & 0 & 0 & 5 & 0 & 2 & 0 & 0 & 2 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 4 & 0 & 0 & 0 \\ 0 & 5 & 0 & 0 & 0 & 0 & 1 & 9 & 0 & \xi \\ 0 & 0 & 0 & 0 & 0 & 0 & 6 & 0 & 3 & 0 \\ 3 & 2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 7 & 0 & 4 & 1 & 6 & 0 & 0 & 0 & \xi & 0 \\ 0 & 0 & 0 & 9 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 & 3 & 0 & \xi & 0 & 0 & 0 \\ 0 & 0 & 0 & \xi & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

Dará como resultado una diferencia máxima entre deciles de

$$M^d = 9$$

Por lo que las grandes regiones a analizar serán las representadas por las filas que contengan por lo menos una entrada igual a 9; es decir, las regiones $\{1, 2, 4, 8\}$ serán consideradas como regiones con diferencia máxima y $\{1, 2, 3, 4, 6, 7, 8, 9\}$ como vecinas de estas regiones de acuerdo a la forma de construcción de la matriz de contigüidades.

El análisis se centrará entonces en determinar el valor del retardo espacial de las nuevas regiones, las cuales serán todas las combinaciones de pares de regiones que se puedan producir entre las regiones a analizar.

Como se dijo anteriormente, este retardo espacial estará determinado por la siguiente expresión:

$w'y' = w_1y_1 + w_2y_2 - y_1 - y_2$

donde w_iy_i es el retardo espacial de la región y_i.

Los valores de los retardos espaciales y los correspondientes a las regiones pueden ser obtenidos de los vectores WY y Y respectivamente, por lo que se puede construir una matriz cuadrada compuesta por los retardos espaciales de las nuevas regiones, conformadas a partir de la agregación de dos regiones originales: la correspondiente a la fila y la correspondiente a la columna.

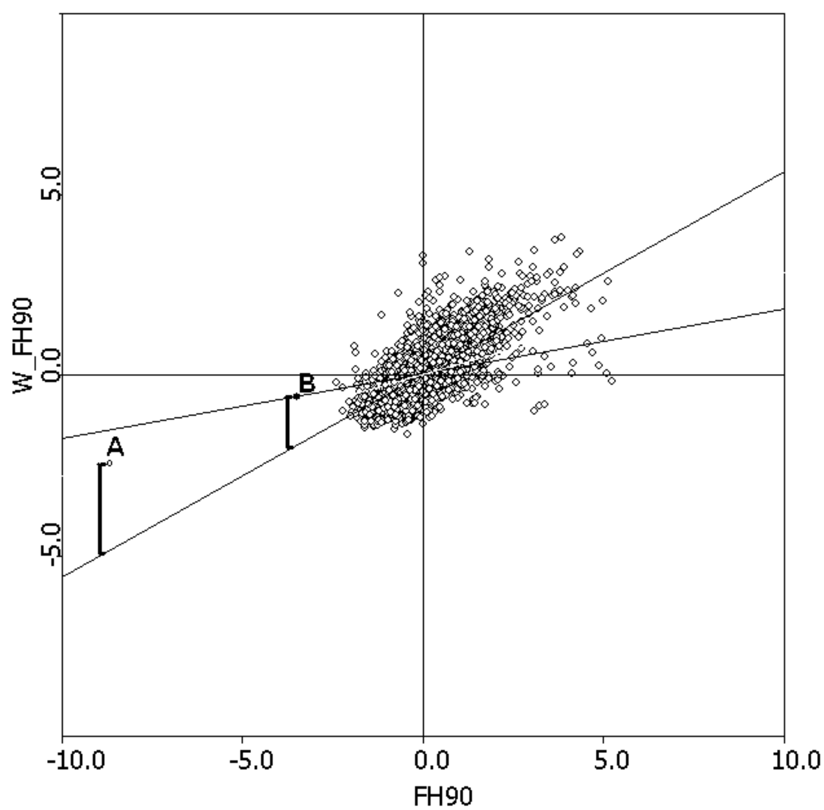
Se puede entonces conformar una matriz de contigüidades especial (W^{ε}) que incluirá, solo para las regiones a analizar, el valor del retardo espacial para cada par de regiones (fila – columna) bajo el supuesto de ser agregadas.

Cada uno de estos puntos se graficaron en un scatter plot que incluye el nuevo punto, producto de la agregación de dos regiones, y los puntos correspondientes a las regiones no agregadas. Este nuevo punto es el que produce una modificación en la pendiente de la recta de regresión y para escoger el mejor par de puntos a agregar se deben analizar cuál es el que produce el mayor efecto sobre esta.

Un indicador que puede establecerse como regla de decisión es el valor absoluto de la diferencia entre la componente vertical del punto agregado y el valor de la variable endógena WY, obtenida reemplazando en la ecuación de la recta la variable exógena Y por el valor de la componente horizontal del punto agregado. Sin embargo, esta medida no tiene en cuenta otra variable a considerar y es la cercanía o lejanía del punto con respecto a al eje de las ordenadas; tener en cuenta estos dos aspectos se traduce en impactar en fuertemente la pendiente de la recta de regresión.

$\upsilon = \left|WY - \overline{WY}\right|$

En el siguiente gráfico podemos notar cómo la mayor distancia del punto A a la recta de regresión original no implica una mayor reducción en la pendiente de esta, pues el punto B contando con una menor distancia la impacta con mayor fuerza al estar su componente horizontal más cerca del origen.



Esta deducción nos lleva a otro indicador que será adoptado como regla de decisión:

$\psi = \frac{\left|WY - \overline{WY}\right|}{|Y|}$

ψ relaciona la distancia vertical del punto agregado y la recta de regresión original con la componente horizontal del primero, para producir una medida de distancia relativa que penalizará la lejanía del nuevo punto con respecto al eje de las ordenadas. Así, entre más lejano, más distancia vertical necesitará el punto para establecerse como agregación óptima.

3.2 Resultados del algoritmo

Para validar la efectividad del algoritmo se construyó un programa en MATLAB que permite la agregación de tantas regiones como se quiera y arroja como resultado las pendientes de cada una de las agregaciones y su variación con respecto a la pendiente original.

Posteriormente se desarrolló otro programa en la misma herramienta que permite realizar agregaciones aleatorias con una sola restricción, la contigüidad entre las regiones agregadas.

Los códigos de ambos programas pueden ser consultados en los apéndices de este trabajo de grado.

Para determinar la capacidad del algoritmo para encontrar las regiones óptimas a agregar se construyó un mapa con 100 regiones con valores aleatorios provenientes de una distribución uniforme. Las regiones fueron agrupadas de acuerdo a su cercanía y a cada grupo se le asignó un rango para la generación de los números aleatorios; esto permitió formar deciles como se muestra a continuación.

1930	2890	2910	7610	9378	7185	-101	-117	-1886	-1268
1406	1204	2203	7898	8734	8527	-338	-1504	918	-1435
2095	2172	2340	8445	7393	7186	-1708	-802	-1170	-1631
-7801	-5514	-5927	-7886	4935	3528	4311	-4655	-2551	-2275
-3904	-6727	-5476	-3342	2408	2315	2257	-4553	-3763	-4416
-5732	-5514	-7751	-6658	3035	4820	3103	-2414	-2783	-3430
302	252	757	8792	8422	8769	3830	4977	4460	4737
471	1931	1271	7492	7416	8383	4464	4708	4949	5937
1260	1525	686	7559	7823	8103	5198	5037	5712	3026
1874	356	675	7550	8918	7993	3134	3927	3652	4720

Decil
1
2
3
4
5
6
7
8
9
10

1	2	3	4	5	6	7	8	9	10
11	12	13	14	15	16	17	18	19	20
21	22	23	24	25	26	27	28	29	30
31	32	33	34	35	36	37	38	39	40
41	42	43	44	45	46	47	48	49	50
51	52	53	54	55	56	57	58	59	60
61	62	63	64	65	66	67	68	69	70
71	72	73	74	75	76	77	78	79	80
81	82	83	84	85	86	87	88	89	90
91	92	93	94	95	96	97	98	99	100

5	6	6	9	10	8	4	4	3	3
5	5	6	9	10	10	3	3	4	3
5	5	6	10	9	9	3	3	3	3
1	1	1	1	8	7	7	2	2	3
2	1	1	2	6	6	6	2	2	2
1	1	1	1	6	8	6	2	2	2
4	4	4	10	10	10	7	8	7	8
4	5	5	9	9	10	7	7	8	8
5	5	4	9	9	10	8	8	8	6
5	4	4	9	10	9	7	7	7	7

Para facilitar su tratamiento, las regiones fueron nombradas con códigos del 1 al 100.

El programa de MATLAB fue construido de tal forma que nombre la región producto de la agregación espacial con el máximo código más uno (1). De tal forma que para el mapa construido (de 100 regiones) la primera región producto de la agregación será codificada como 101. A partir de ahí las nuevas regiones serán nombradas como 102, 103, etc.

El resultado del algoritmo encontró que la primera agregación debe involucrar las regiones 33 y 34. Esta agregación cumple con las dos reglas básicas de la agregación espacial: i) son regiones contiguas y ii) pertenecen a la misma medida de posición (primer decil).

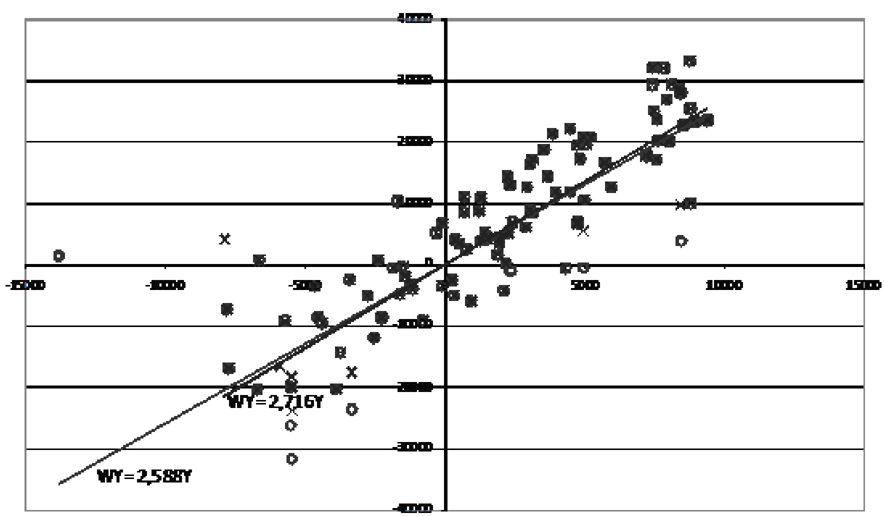
Esta agregación permite disminuir la pendiente de la recta de regresión en 4,71% al pasar de 2,7164 a 2,5880.

El nuevo mapa quedará entonces de la siguiente manera:

1	2	3	4	5	6	7	8	9	10
11	12	13	14	15	16	17	18	19	20
21	22	23	24	25	26	27	28	29	30
31	32	101		35	36	37	38	39	40
41	42	43	44	45	46	47	48	49	50
51	52	53	54	55	56	57	58	59	60
61	62	63	64	65	66	67	68	69	70
71	72	73	74	75	76	77	78	79	80
81	82	83	84	85	86	87	88	89	90
91	92	93	94	95	96	97	98	99	100

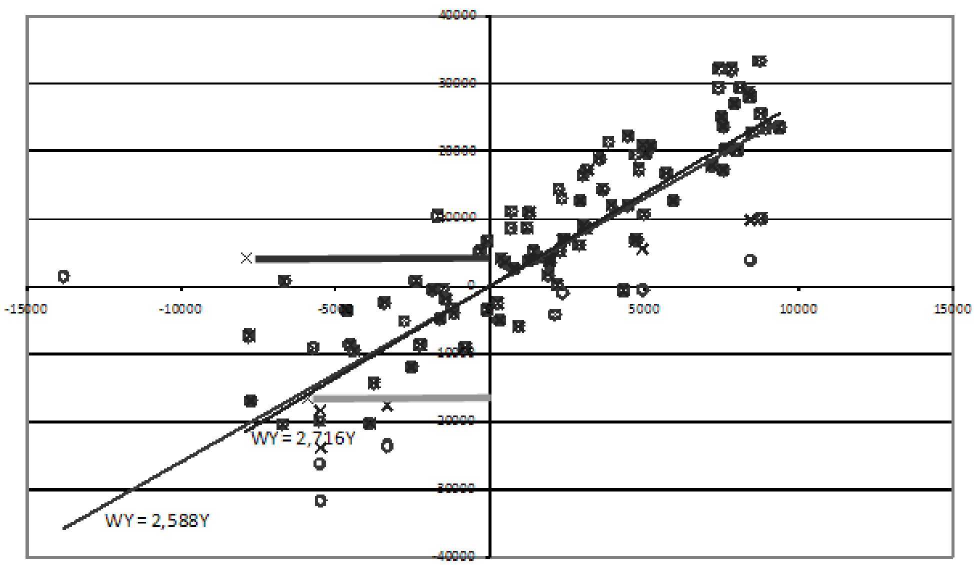
1930	2890	2910	7610	9378	7185	-101	-117	-1886	-1268
1406	1204	2203	7898	8734	8527	-338	-1504	918	-1435
2095	2172	2340	8445	7393	7186	-1708	-802	-1170	-1631
-7801	-5514	-13813		4935	3528	4311	-4655	-2551	-2275
-3904	-6727	-5476	-3342	2408	2315	2257	-4553	-3763	-4416
-5732	-5514	-7751	-6658	3035	4820	3103	-2414	-2783	-3430
302	252	757	8792	8422	8769	3830	4977	4460	4737
471	1931	1271	7492	7416	8383	4464	4708	4949	5937
1260	1525	686	7559	7823	8103	5198	5037	5712	3026
1874	356	675	7550	8918	7993	3134	3927	3652	4720

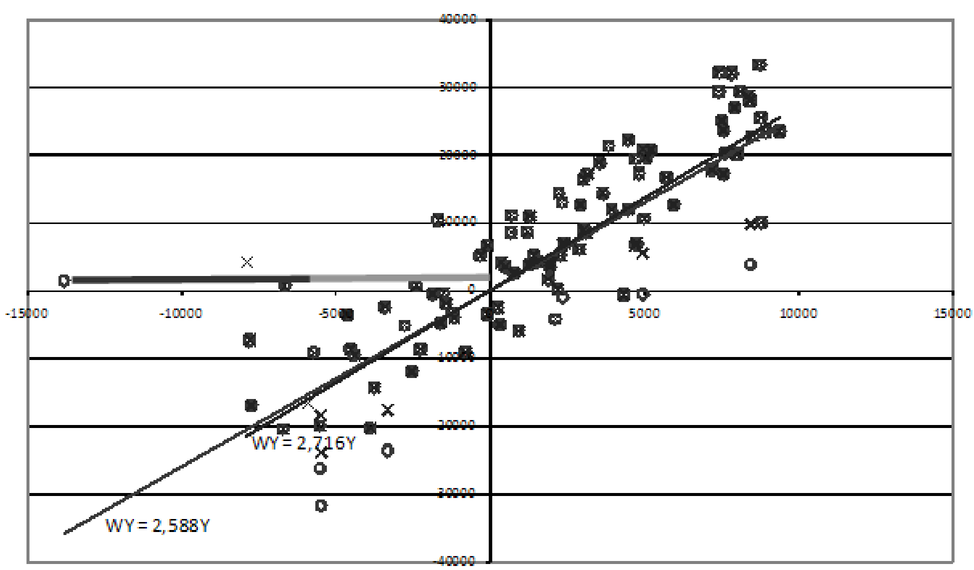
Y el scatter plot con los puntos originales y los puntos producidos por la agregación espacial es el siguiente:



En este scatter plot pueden observarse dos efectos de la agregación espacial:

El primero de ellos es que la agregación producirá una componente horizontal igual a la suma de las componentes horizontales de las regiones agregadas





El segundo efecto es que la agregación no solo afecta a las dos regiones que se fusionan sino también a las regiones vecinas de cada una, pues ahora contarán con un nuevo vecino, más grande, que reemplazará a las dos regiones agregadas, por lo que su retardo espacial se verá lógicamente impactado.

Es en este segundo efecto donde radica la dificultad en la disminución de la pendiente de la recta de regresión; al incluir la restricción de no agregación de regiones con diferente medida de posición, se hace muy probable la fusión de regiones con valores de la variable endógena ambos positivos o ambos negativos. Esta igualdad de signos producirá un efecto lógico en la componente vertical de cada uno de los puntos correspondientes a regiones vecinas: valores negativos de las regiones agregadas producirán disminuciones en los retardos espaciales de sus vecinos y valores positivos producirán el efecto contrario. La presencia de este efecto hará contrapeso al movimiento de la recta de regresión haciendo más difícil la disminución de la pendiente.

Producto del cambio en la componente vertical de los puntos correspondientes a las regiones vecinas a las fusionadas, se produce una segunda consecuencia: agregar repetidas veces no implica una constante diminución en la pendiente con respecto a la pendiente original. A fin de detectar y corregir la autocorrelación espacial debe establecerse previamente el número de agregaciones a realizar, siendo este el número de observaciones a perder, y elegirse el número de agregaciones óptimo para minimizar la pendiente.

Los resultados obtenidos para 5 agregaciones se muestran en la siguiente tabla:

Reg. 1	Reg. 2	Reg. creada	Pend.	Pend. orig.	Reducción pend.
33	34	101	2,5880	2,7164	-0,0473
32	101	102	2,4847	2,7164	-0,0853
31	102	103	2,1483	2,7164	0,2091
47	57	104	2,1473	2,7164	-0,2095
58	59	105	2,1272	2,7164	-0,2169

Los mapas generados por cada agregación son los siguientes:

1	2	3	4	5	6	7	8	9	10
11	12	13	14	15	16	17	18	19	20
21	22	23	24	25	26	27	28	29	30
31	32	101		35	36	37	38	39	40
41	42	43	44	45	46	47	48	49	50
51	52	53	54	55	56	57	58	59	60
61	62	63	64	65	66	67	68	69	70
71	72	73	74	75	76	77	78	79	80
81	82	83	84	85	86	87	88	89	90
91	92	93	94	95	96	97	98	99	100

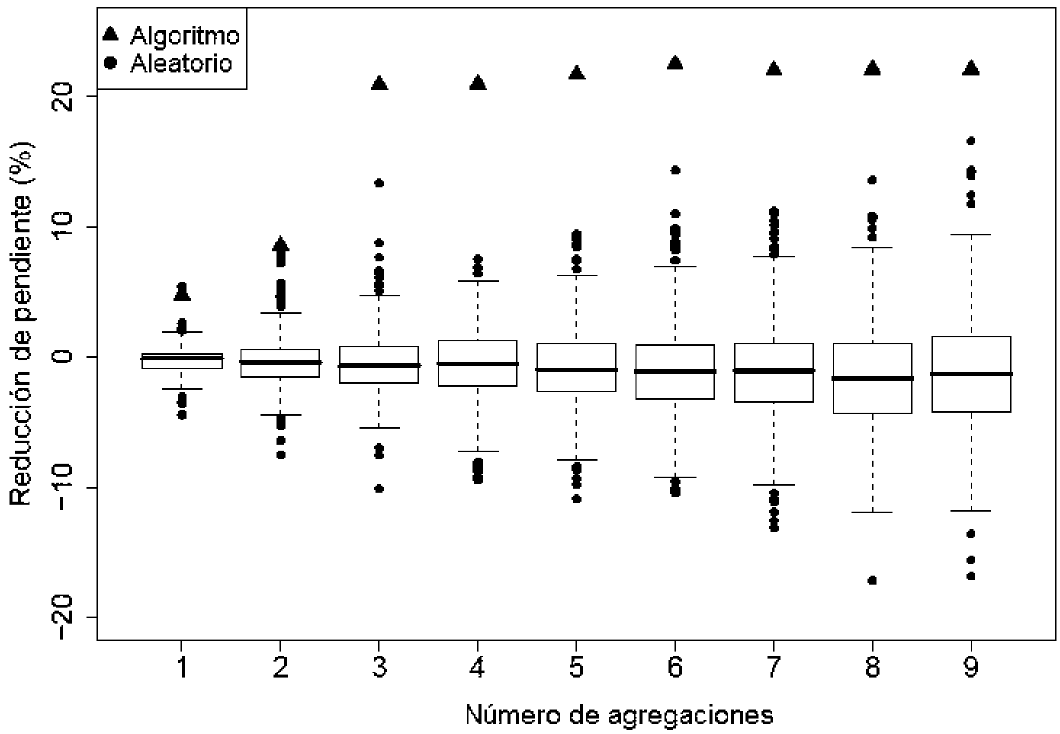
1	2	3	4	5	6	7	8	9	10
11	12	13	14	15	16	17	18	19	20
21	22	23	24	25	26	27	28	29	30
31	102			35	36	37	38	39	40
41	42	43	44	45	46	47	48	49	50
51	52	53	54	55	56	57	58	59	60
61	62	63	64	65	66	67	68	69	70
71	72	73	74	75	76	77	78	79	80
81	82	83	84	85	86	87	88	89	90
91	92	93	94	95	96	97	98	99	100

1	2	3	4	5	6	7	8	9	10
11	12	13	14	15	16	17	18	19	20
21	22	23	24	25	26	27	28	29	30
103				35	36	37	38	39	40
41	42	43	44	45	46	47	48	49	50
51	52	53	54	55	56	57	58	59	60
61	62	63	64	65	66	67	68	69	70
71	72	73	74	75	76	77	78	79	80
81	82	83	84	85	86	87	88	89	90
91	92	93	94	95	96	97	98	99	100

<table>
<tr><td>1</td><td>2</td><td>3</td><td>4</td><td>5</td><td>6</td><td>7</td><td>8</td><td>9</td><td>10</td></tr>
<tr><td>11</td><td>12</td><td>13</td><td>14</td><td>15</td><td>16</td><td>17</td><td>18</td><td>19</td><td>20</td></tr>
<tr><td>21</td><td>22</td><td>23</td><td>24</td><td>25</td><td>26</td><td>27</td><td>28</td><td>29</td><td>30</td></tr>
<tr><td colspan="4">103</td><td>35</td><td>36</td><td>37</td><td>38</td><td>39</td><td>40</td></tr>
<tr><td>41</td><td>42</td><td>43</td><td>44</td><td>45</td><td>46</td><td rowspan="2">104</td><td>48</td><td>49</td><td>50</td></tr>
<tr><td>51</td><td>52</td><td>53</td><td>54</td><td>55</td><td>56</td><td>58</td><td>59</td><td>60</td></tr>
<tr><td>61</td><td>62</td><td>63</td><td>64</td><td>65</td><td>66</td><td>67</td><td>68</td><td>69</td><td>70</td></tr>
<tr><td>71</td><td>72</td><td>73</td><td>74</td><td>75</td><td>76</td><td>77</td><td>78</td><td>79</td><td>80</td></tr>
<tr><td>81</td><td>82</td><td>83</td><td>84</td><td>85</td><td>86</td><td>87</td><td>88</td><td>89</td><td>90</td></tr>
<tr><td>91</td><td>92</td><td>93</td><td>94</td><td>95</td><td>96</td><td>97</td><td>98</td><td>99</td><td>100</td></tr>
</table>

<table>
<tr><td>1</td><td>2</td><td>3</td><td>4</td><td>5</td><td>6</td><td>7</td><td>8</td><td>9</td><td>10</td></tr>
<tr><td>11</td><td>12</td><td>13</td><td>14</td><td>15</td><td>16</td><td>17</td><td>18</td><td>19</td><td>20</td></tr>
<tr><td>21</td><td>22</td><td>23</td><td>24</td><td>25</td><td>26</td><td>27</td><td>28</td><td>29</td><td>30</td></tr>
<tr><td colspan="4">103</td><td>35</td><td>36</td><td>37</td><td>38</td><td>39</td><td>40</td></tr>
<tr><td>41</td><td>42</td><td>43</td><td>44</td><td>45</td><td>46</td><td rowspan="2">104</td><td>48</td><td>49</td><td>50</td></tr>
<tr><td>51</td><td>52</td><td>53</td><td>54</td><td>55</td><td>56</td><td colspan="2">105</td><td>60</td></tr>
<tr><td>61</td><td>62</td><td>63</td><td>64</td><td>65</td><td>66</td><td>67</td><td>68</td><td>69</td><td>70</td></tr>
<tr><td>71</td><td>72</td><td>73</td><td>74</td><td>75</td><td>76</td><td>77</td><td>78</td><td>79</td><td>80</td></tr>
<tr><td>81</td><td>82</td><td>83</td><td>84</td><td>85</td><td>86</td><td>87</td><td>88</td><td>89</td><td>90</td></tr>
<tr><td>91</td><td>92</td><td>93</td><td>94</td><td>95</td><td>96</td><td>97</td><td>98</td><td>99</td><td>100</td></tr>
</table>

Para contrastar la efectividad del algoritmo de agregación, se realizó una prueba empírica desarrollando 500 agregaciones aleatorias sucesivas con una única restricción: solo se permite la agregación de regiones vecinas. Posteriormente se realizó el análisis de la reducción de la pendiente de la recta de regresión para verificar si el algoritmo desarrollado está en capacidad de encontrar mejores soluciones que un proceso aleatorio. Los resultados se presentan en el siguiente gráfico:



Estos resultados son concluyentes en el sentido que permiten afirmar que el algoritmo está en capacidad de encontrar las regiones óptimas a agregar, con las restricciones dadas, de tal suerte que la pendiente de la recta de regresión se haga mínima. Solo en el primer Box – Plot el algoritmo fue superado por una agregación aleatoria; el amplio número de ensayos realizados (500) permite afirmar que, en este espacio, una agregación aleatoria de dos regiones tiene una probabilidad cercana a 0,002 de obtener una reducción en la pendiente mayor a la alcanzada por el algoritmo.

Otro resultado importante de este ejercicio es que existe un punto donde agregaciones adicionales no permiten disminuir más la pendiente de la recta, esto sumado a la creciente dispersión en los resultados producidos por las agregaciones aleatorias ocasionan que con cada nueva agregación la probabilidad de obtener la misma pendiente por parte del procedimiento aleatorio aumente. Este resultado, sin embargo, se lograría sacrificando información al no tener la restricción de agregación de regiones similares.

Los resultados empíricos son concluyentes al demostrar que mediante el algoritmo desarrollado se obtiene una reducción de la pendiente que permite atacar la hipótesis alternativa de presencia de autocorrelación espacial con mayor eficiencia y menor pérdida de información que con un procedimiento de agregación aleatorio. Este desarrollo permitirá al investigador corroborar la presencia de autocorrelación espacial, corregir la espuria o descartar de plano algún tipo de dependencia espacial en el espacio analizado.

Capítulo 4

Defunción de no pensionados en Colombia

El sistema pensional colombiano se caracteriza por tener dos regímenes iguales en cuanto a la tasa de cotización en cada uno de ellos y los tipos de riesgo cubiertos (vejez, invalidez y sobrevivencia), pero diferentes en las condiciones para acceder a la prestación al entrar a la tercera edad.

El régimen de prima media, administrado casi en su totalidad por el Instituto de Seguros Sociales (ISS) es un sistema de reparto en el cual las cotizaciones de todos sus afiliados van a una bolsa común de donde sufragan los gastos de funcionamiento y el pago de las mesadas pensionales a cientos de miles de colombianos que ya adquirieron el derecho a recibir la prestación, sea la indemnización sustitutiva para aquellas personas que no cumplieron con los requisitos de edad y semanas cotizadas para acceder a la pensión de vejez, o la mesada pensional para aquellos que sí las cumplieron.

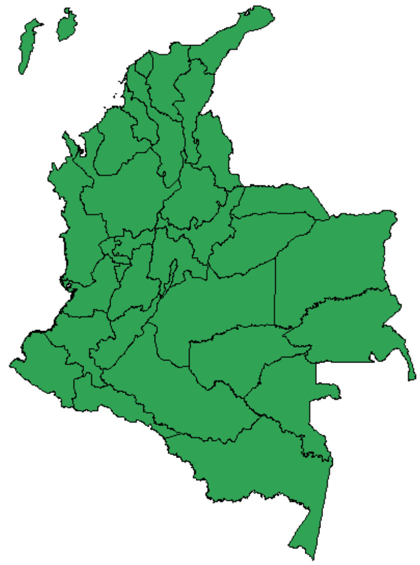
El régimen de ahorro individual, gestionado por las Administradoras de Fondos de Pensiones y Cesantías (AFP) es un sistema en el cual cada afiliado es titular de su propia cuenta y propietario de los recursos que mes a mes son cotizados al sistema en cabeza suya. La prestación por lo tanto, será o la devolución de la totalidad del saldo de la cuenta para aquellas personas que no acumularon el suficiente saldo como para acceder a su pensión de vejez y no desean seguir cotizando, o la mesada pensional para aquellas cuyo saldo les alcanzaría para acceder a una pensión con un monto superior un (1) salario mínimo mensual legal vigente (SMMLV) de acuerdo al cálculo actuarial realizado al momento de solicitar la prestación.

Uno de los aspectos en los cuales no existe diferencia alguna entre los dos regímenes pensionales es el de la *pensión de sobrevivencia.* Con este término se hace referencia a aquella prestación que recibe el conyuge o compañero permanente de un pensionado o de un afiliado que habiendo fallecido cumplía con el requisito que la ley 100 de 1993 define para acceder a tal derecho: haber cotizado al Sistema General de Pensiones 50 semanas en los últimos 3 años. Es este tipo de pensión el que plantea un reto poco exporado en el pais, tanto por parte del ISS como de las AFP; la estimación de la tasa de defunción de personas adultas no pensionadas es importante por cuanto permitiría a las administradoras de ambos regímenes, acercarse aun más a una proyección de gastos operativos y desembolsos de recursos que le permitan planear sus flujos de efectivo de manera más acertada. La estimación de la tasa de defunción de personas pensionadas se encuentra bien documentada en las tablas de mortalidad expedidas recientemente por la Superintendencia Financiera de Colombia por lo que su estudio no presenta tanta relevancia.

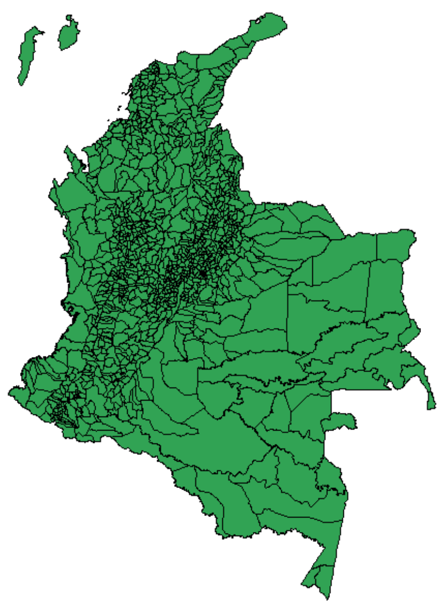
4.1 División politico-económica de Colombia

Colombia es un pais caracterizado por tener cultural, política y económicamente muchas más regiones que cualquier otro país de América; es así como se distinguen regiones como la de la costa caribe, la andina, la cundiboyacence, la de la costa pacífico entre muchas otras. Esta diversidad de regiones grandes y fuertes a permitido evitar riesgos como el de permitir dictaduras o de caídas estrepitosas en la producción nacional, pero también ha limitado la unión nacional para encarar retos comunes.

Políticamente, la República de Colombia está dividida en 32 departamentos y un Distrito Capital. Estos 32 departamentos a su vez se dividen en 1101 municipios, algunos de ellos tan pequeños que en otros países serían catalogados como poblados.



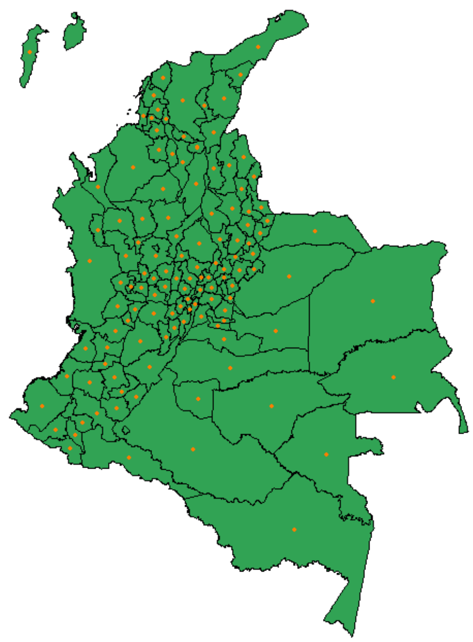
Departamentos de Colombia



Municipios de Colombia

En cuanto al análisis espacial, debe tenerse especial cuidado tal como lo sugieren algunos autores, al trabajar con divisiones políticas de los territorios pues estas no siempre tienen en cuenta las diferentes características que marcan diferencias entre poblaciones, sino que corresponden a procesos políticos y administrativos.

Para el caso de Colombia y debido a factores culturales, climáticos, económicos, raciales, etc., las grandes divisiones que son los departamentos se han ido dividiendo en subregiones bien definidas desde la Constitución de 1991, hasta llegar en la actualidad a 124. El objetivo de estas divisiones agrupar municipios con características similares permitiendo una mejor y más sencilla administración por parte de las gobernaciones departamentales. La similitud de características intrarregionales permite pensar que es la subdivisión más adecuada para estudiar una variable vinculada al espacio en Colombia.

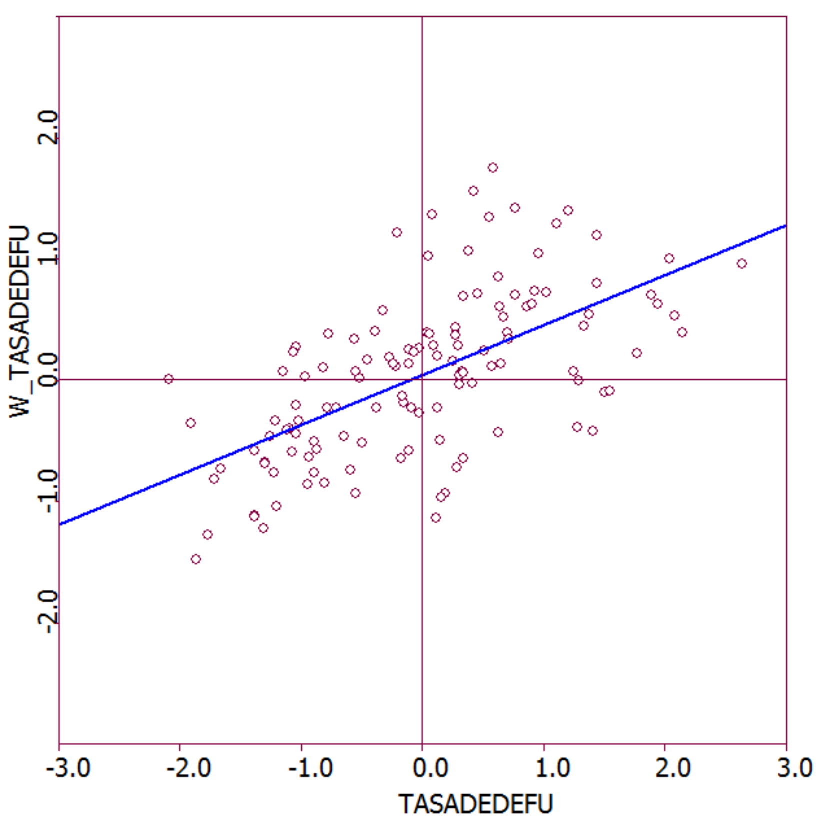


4.2 Análisis de dependencia espacial

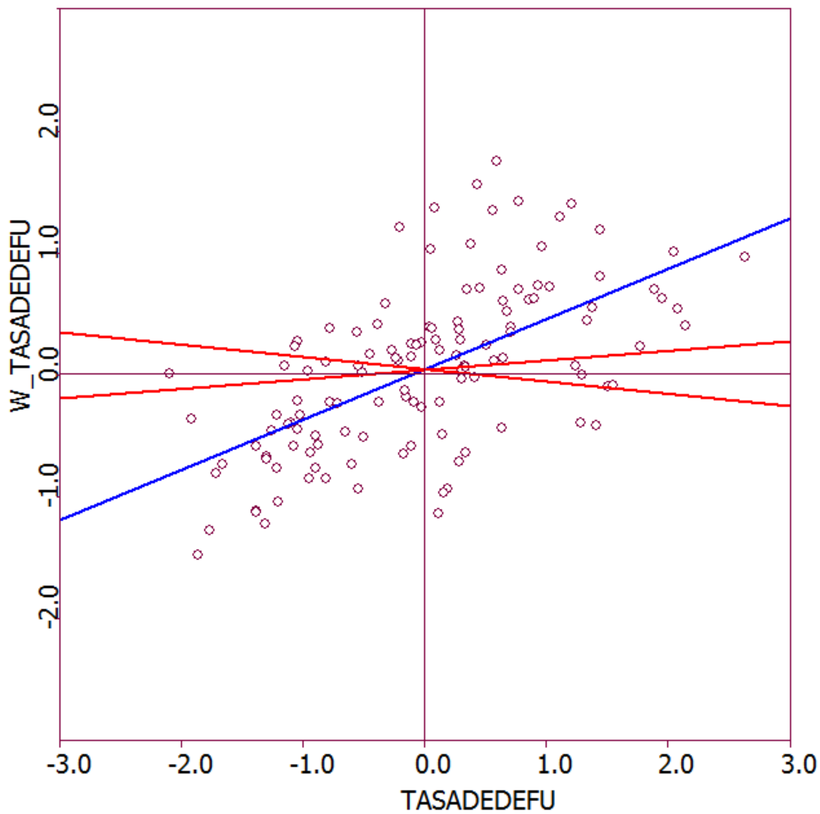
Para el análisis de la tasa de defunción de adultos no pensionados en las subregiones colombianas se asumió que corresponde a la misma tasa de defunción de personas entre los 15 y los 64 años, datos proveídos por el Departamento Nacional de Estadística (DANE) a través de sus estadísticas de defunciones no fetales correspondientes al año 2008.

Se define entonces a Y_i como la tasa de defunción de adultos no pensionados en la región i por cada 100.000 habitantes con estas mismas características. $i = 1, 2, ..., 124$.

Como es apenas lógico pensar, la tasa de defunciones por cada 100.000 habitantes debería estar autocorrelacionada espacialmente, pues condiciones económicas, climáticas, sociales, sanitarias y de violencia, entre muchas otras, permiten inferir que el espacio físico es relevante al intentar explicar su variabilidad. Para comprobar esta hipótesis se generó, mediante el software GeoDa, un scatter plot de esta tasa de defunción contra la misma variable rezagada, el cual se presenta a continuación.



A través de este software se puede también generar las bandas por fuera de las cuales no se puede descartar la presencia de autocorrelación espacial.



Como se observa, la recta de regresión que relaciona la tasa de defunción por cada 100.000 habitantes con la misma variable rezagada espacialmente está por fuera de las bandas que indican ausencia de autocorrelación espacial, por lo que se infiere que la localización geográfica de los datos es importante para explicar su variabilidad.

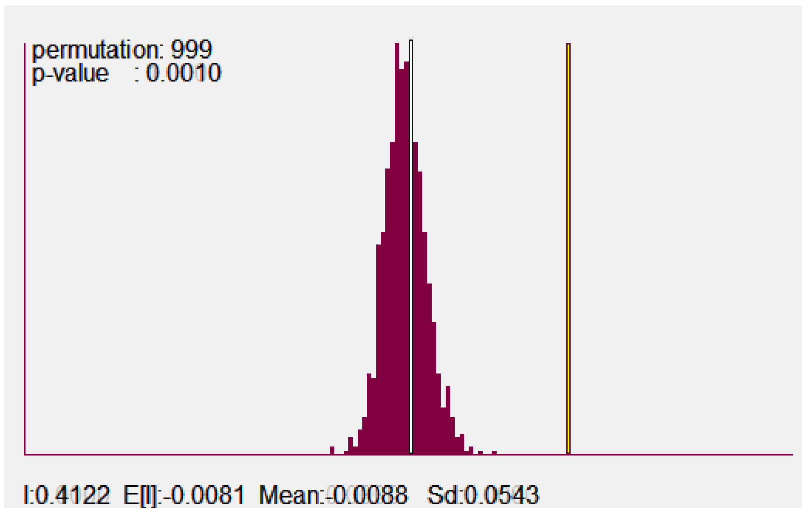
Estos métodos gráficos pueden ser validados a través de las pruebas de autocorrelación espacial analizadas en el capítulo 2. El software GeoDa permite también realizar la prueba de hipótesis de presencia de dependencia espacial a través de la I de Moran, donde las hipótesis planteadas son:

$H_0 : \rho = 0$
$H_a : \rho \neq 0$

.
Y el estadístico de prueba:

$$Z(I) = \frac{I - E(I)}{\sqrt{V(I)}}\ , \quad \text{donde } I = \frac{N}{S_0} \frac{\sum_{i=1}^{N}\sum_{j=1}^{N} W_{ij}\left(Y_i - \overline{Y}\right)\left(Y_j - \overline{Y}\right)}{\sum_{i=1}^{N}\left(Y_i - \overline{Y}\right)^2}$$

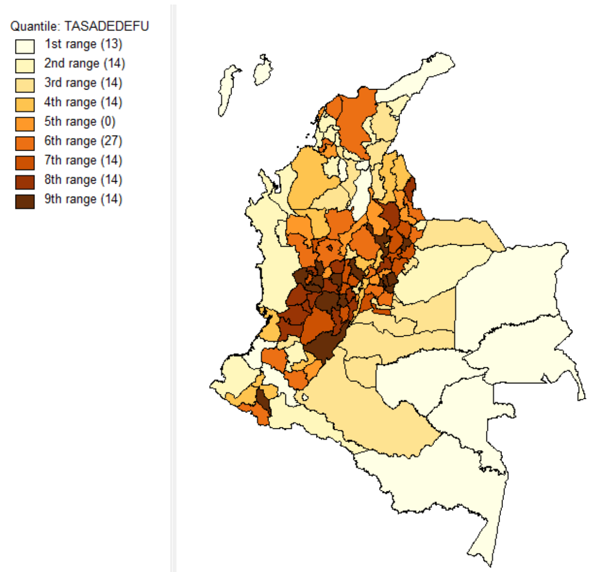
Los resultados obtenidos fueron los siguientes:



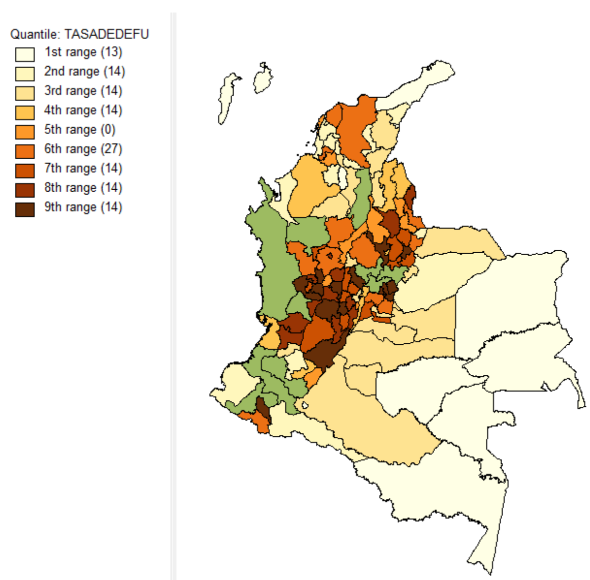
El valor p indica que se rechaza la hipótesis nula de no dependencia espacial, por lo que es de esperarse que la variable rezagada espacialmente WY sea significativa para explicar la variabilidad de Y.

Sin embargo, antes de iniciarse el proceso de modelación de la variable debe descartarse la presencia de autocorrelación espacial espuria, máxime si se tiene en cuenta que estas subregiones fueron creadas tomando como base los departamentos ya existentes y criterios posiblemente más políticos que técnicos.

Para hacerlo se utilizó el algoritmo desarrollado en el capítulo anterior y mediante el software MatLab se implementó realizando las 12 mejores agregaciones espaciales para tratar de descartar la hipótesis de presencia de autocorrelación espacial. Estas 12 agregaciones implican perder un máximo de 12 observaciones de la muestra, que corresponden a aproximadamente el 10% de la misma. Se obtuvieron los siguientes resultados:



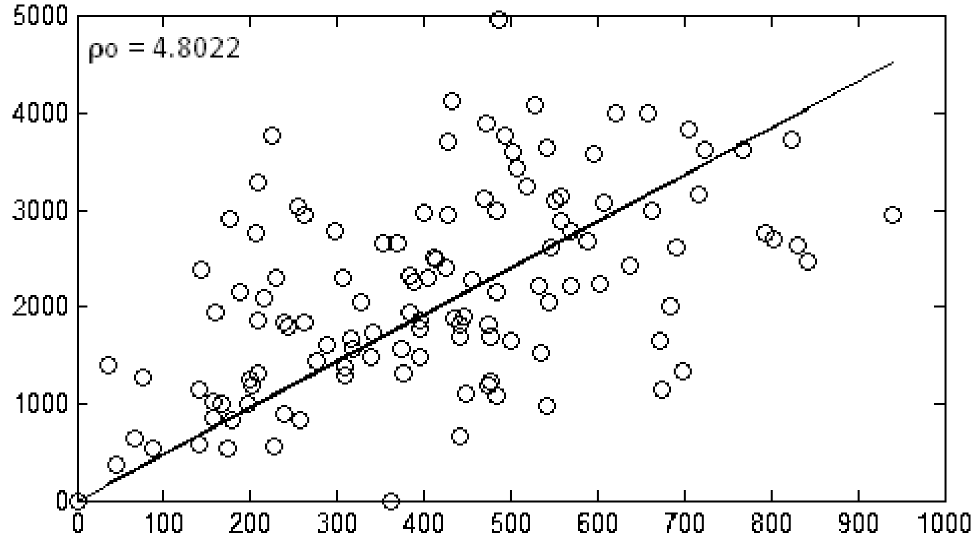
Subregiones colombianas

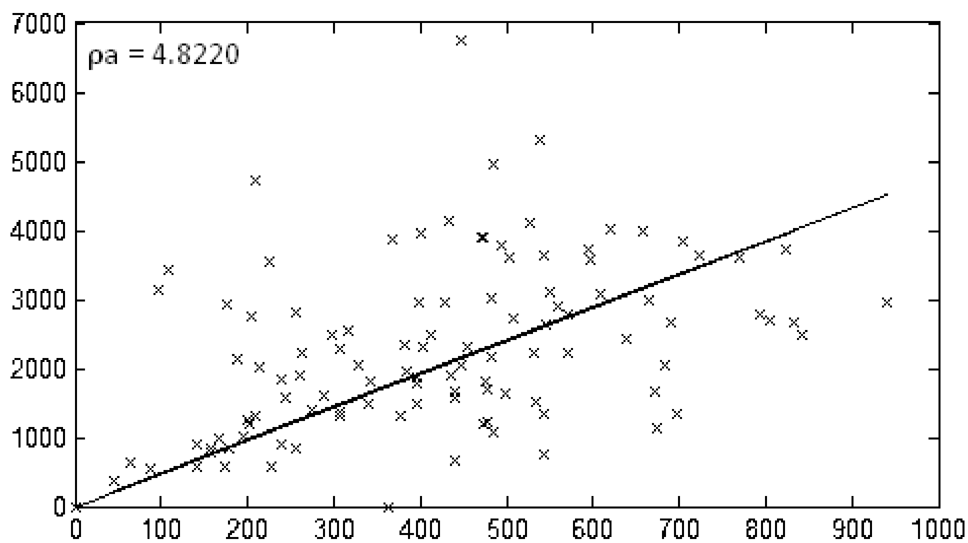


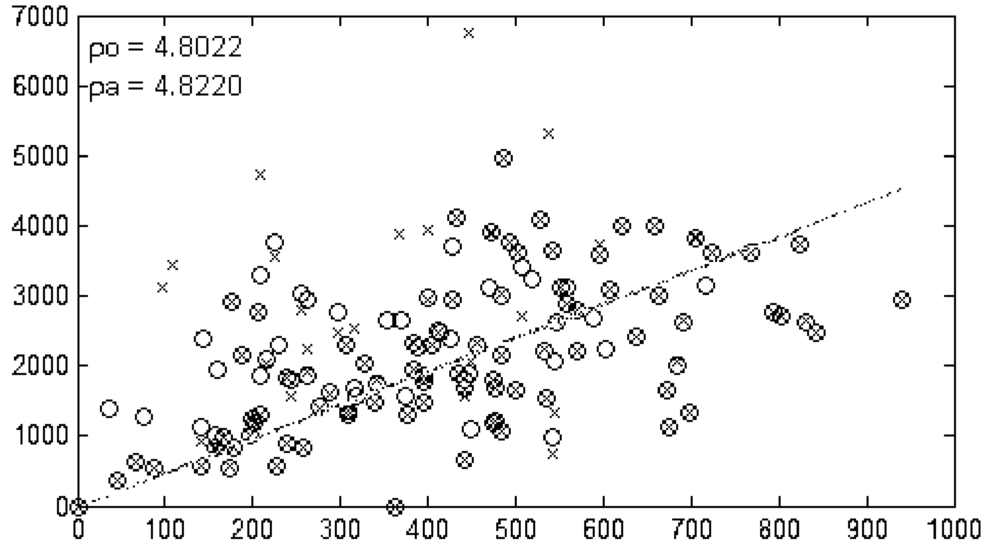
Subregiones colombianas luego de 12 agregaciones espaciales

Reg. 1	Reg. 2	Reg. Creada	Pend.	Pend. Orig.	Variación Pend.
31	104	125	4.8237	4.8022	0,4477%
79	107	126	4.8233	4.8022	0.4394%
51	55	127	4.8220	4.8022	0.4123%
73	35	128	4.8329	4.8022	0.6393%
20	57	129	4.8526	4.8022	1.0495%
88	76	130	4.8924	4.8022	1.8783%
30	105	131	4.9267	4.8022	2.5926%
129	119	132	4.9544	4.8022	3.1694%
78	92	133	4.9697	4.8022	3.4880%
7	120	134	4.8812	4.8022	1.6451%
67	77	135	4.9003	4.8022	2.0428%
9	133	136	4.9178	4.8022	2.4072%

La tabla anterior indica que la menor pendiente lograda es de 4.8220, alcanzada luego de 3 agregaciones sucesivas; sin embargo, esta resulta ser superior a la pendiente inicial, es decir sin agregaciones. En las siguientes figuras se observan las pendientes original y producto de la agregación y los puntos originales (marcador en forma de círculo) y los puntos desplazados producto de la agregación espacial (marcador en forma de x).


ρo = 4.8022
5000
4000
3000
2000
1000
0
0
100
200
300
400
500
600
700
800
900
1000


ρa = 4.8220
7000
6000
5000
4000
3000
2000
1000
0
0
100
200
300
400
500
600
700
800
900
1000



De la ubicación de los nuevos puntos sobre el scatter plot se infiere que el efecto en la disminución en la pendiente provocado por la agregación de las dos mejores regiones para hacerlo, se ve contrarrestado e incluso superado por el movimiento en los puntos correspondientes a las regiones vecinas.

La lejanía de la pendiente de la recta con respecto a cero permite descartar la hipótesis nula de no autocorrelación espacial, permitiendo al mismo tiempo afirmar que existe un fenómeno de dependencia espacial genuino. A partir de esta conclusión puede entonces comenzarse con la modelación de la variable teniendo en cuenta el efecto espacial que ha sido validado a través del algoritmo, para ello se utilizará la especificación $SAR(1)$ descrita en el capítulo 2.

Los resultados de la modelación espacial fueron los siguientes:

Variable	Coefficient	Std.Error	t-Statistic	Probability
CONSTANT	192.6316	35.94294	5.359373	0.0000004
WTASADEDEF	0.1019209	0.014826	6.874473	0.0000000

Como se esperaba, la variable rezagada $WTASADEDEF$ fue significativa para explicar la variabilidad de la tasa de defunción de adultos no pensionados en las subregiones colombianas.

Los test de Multiplicadores de Lagrange para la dependencia espacial sustantiva y para la residual permiten confirmar por un lado la presencia de la primera y descartar por otro la presencia de la segunda.

Test de Multiplicadores de Lagrange para la dependencia espacial sustantiva

H_0 : Existe dependencia espacial sustantiva
H_a : No existe dependencia espacial sustantiva
Multiplicador de Lagrange
Grados de libertad: 1
Valor: 16,181842
Valor p: 0,000057
Resultado: No se rechaza la hipótesis nula

Test de Multiplicadores de Lagrange para la dependencia espacial residual

H_0 : Existe dependencia espacial residual
H_a : No existe dependencia espacial residual
Multiplicador de Lagrange
Grados de libertad: 1
Valor: 1,091880
Valor p: 0,296055
Resultado: Se rechaza la hipótesis nula

Los resultados encontrados permiten concluir que la tasa de defunciones en adultos no pensionados en Colombia depende del valor de la misma variable en localizaciones vecinas. Este resultado fue detectado a través del análisis de la pendiente de la recta de regresión realizada entre WY y Y y a través de la I de Moran; así mismo, fue validada esta dependencia espacial y descartada la autocorrelación espacial espuria utilizando el algoritmo de agregación espacial presentado en el capítulo 3.

La estimación del modelo $SAR(1)$ permitió contrastar la significatividad de la relación entre la variable y la variable rezagada, permitiendo así concluir que es posible explicar la variabilidad en la tasa de defunción de adultos no pensionados en las subregiones colombianas a través de esa misma variable en localizaciones cercanas.

Los resultados obtenidos en este capítulo permitirán ajustar las mediciones de riesgo tanto de las Administradoras de Fondos de Pensiones y Cesantías como de las aseguradoras con las cuales contratan sus seguros previsionales, utilizados para cubrir el capital faltante en las cuentas individuales de sus afiliados, ante eventuales derechos pensionales de estos por los riesgos de invalidez o sobrevivencia. Así mismo, permitirá a las compañías aseguradoras mejorar sus estimaciones de riesgo para establecer con una mayor precisión las reservas matemáticas necesarias para su operación en seguros de vida y pensiones.

Conclusiones

Antes de los trabajos de Paelinck y Klaassen (1979) la inclusión de efectos espaciales se realizaba a través de indices que se incluían en los modelos lineales tradicionales; a partir de ese momento la econometría espacial ha tenido un desarrollo lento con respecto a la econometría financiera o de series de tiempo, aunque no por ello autores como Anselin, Florax, Folmer, Ord entre otros y mas recientemente Coro Chasco y LeSage han dejado de lado el análisis espacial y la ciencia regional. Sus valiosas contribuciones han proporcionado una herramienta de modelación de efectos espaciales confiable y robusta a cada vez más investigadores al rededor del mundo, permitiendo explicar el comportamiento de variables vinculadas al espacio con mayor precisión.

A lo largo de esta tesis se han visto los fundamentos conceptuales y matemáticos que conlleva una correcta modelación espacial. El análisis de conceptos relevantes como dependencia y heterogeneidad espacial y la exposición de los problemas al intentar modelar estos efectos espaciales a través de modelos lineales generales, permiten al lector comprender la importancia de los desarrollos aquí explicados y adentrarse posteriormente con mayor facilidad en las taxonomías planteadas por Vayá, Florax y Folmer, y Anselin, las cuales permiten tener una visión completa de las familias de modelos espaciales. Simultáneamente, el análisis de cada uno de ellos permitió exponer el método general de estimación por Máxima Verosimilitud y plantear las restricciones a tener en cuenta al hallar estimadores por Mínimos Cuadrados Ordinarios.

En el capítulo 3 se hizo una descripción completa de uno de los problemas menos estudiados en la econometría espacial: la autocorrelación espacial espuria. Intentar modelar un efecto espacial inexistente puede llevar al investigador fácilmente a conclusiones erradas y a tomar decisiones incorrectas.

En este capítulo se realizó el planteamiento del problema y se construyó un algoritmo que permite su detección y corrección a través de sucesivas agregaciones espaciales que permiten atacar con fuerza al hipótesis de autocorrelación espacial real; así mismo se contrastó contra agregaciones aleatorias encontrando una marcada eficiencia del algoritmo con respecto a este último método. El desarrollo de este algoritmo permitirá en un futuro el d de un test de autocorrelación espacial espuria que podrá ser utilizad investigadores antes de comenzar un proceso de modelación espacial.

Así mismo, se incluyó en esta tesis un análisis de dependencia espacial de la tasa de defunción de adultos no pensionados en las subregiones colombianas. Este análisis fue realizardo mediante un método gráfico, un test de autocorrelación espacial y el algoritmo construído en el capítulo 3 y permitió establecer que el fenómeno de dependencia espacial en esta variable es genuino y significativo; para comprobarlo, se realizaron 12 agregaciones espaciales (permitiendo así la pérdida de hasta el 10% de la muestra) luego de las cuales la pendiente de la recta de regresión entre la variable y la misma variable rezagada espacialmente no solo no disminuyó, sino que aumentó con respecto a su valor inicial; además se estimó un modelo $SAR(1)$ que permitió establecer que el valor de la variable en regiones vecinas es significativo al momento de explicar la variabilidad en la tasa de defunción de adultos no pensionados. Este resultado permite concluir que esta tasa de fallecimientos depende de la localización geográfica de la observación, por lo que los modelos econométricos utilizados para reemplazar las tablas de mortalidad colombiana, de uso frecuente para población pensionada, deben incluir este efecto espacial; omitirlo podría significar sobrecostos a las Administradoras de Fondos de Pensiones y Cesantías a la hora de contratar sus seguros previsionales y pérdidas a las aseguradoras por falta de competitividad y de métodos adecuados para estimar sus reservas matemáticas.

Apéndice

Algoritmo de Agregación Espacial - MatLab

```
function AutocorrelacionEspacialEspuria(Y)
    clc
    % INICIO DE PARAMETRIZACIÓN ————————————————
    % Se requiere:
    % Y: Vector de valores de la variable dependiente
    % R: Vector con la numeración de cada región (va de 1 a R)
    % W: Matriz de contigüidades
    % NumeroDeAgregaciones: Número de agregaciones que se quieren hacer
    % FIN DE PARAMETRIZACIÓN ————————————————
    % INICIO DEL PROGRAMA PARA AGREGACIÓN OPTIMA DE
    % REGIONES ————————————————
    % Preparativos para iteraciones
    % Copia de datos iniciales
    YOriginal=Y;
    WOriginal=W;
    ROriginal=R;
    DimensionYOriginal=size(YOriginal);
    ResumenRegionesAAgregar=[];
    ResumenBeta=[];
    ResumenDiferenciaPrimerBeta=[];
    Resumen=[];
    % ————————————————
    for Iteracion=1:NumeroDeAgregaciones
```

```
% Determina los deciles a los que pertenece cada región
D= ceil(10*tiedrank(Y)/length(Y));
%——————————————————————————
% Crea una matriz con número de la región, valor de la región y decil al
% que pertenece
RYD=[R Y D];
% ——————————————————————————
% Crea una matriz con la diferencia entre deciles de regiones contiguas
DimensionY=size(Y);
WDecil=zeros(DimensionY(1,1));
DimensionYFilas=DimensionY(1,1);
for A=1:DimensionYFilas
for B=1:DimensionYFilas
if W(A,B)==1
for C=1:DimensionYFilas
if RYD(C,1)==A
DecilFila=RYD(C,3);
end
end
for D=1:DimensionYFilas
if RYD(D,1)==B
DecilColumna=RYD(D,3);
end
end
if DecilFila-DecilColumna~=0
WDecil(A,B)=abs(DecilFila-DecilColumna);
WDecil(B,A)=abs(DecilFila-DecilColumna);
else
WDecil(A,B)=-99;
WDecil(B,A)=-99;
end
end
end
end
% ——————————————————————————
% Determina la máxima diferencia que existe entre todas las regiones
MaximaDiferencia=0;
for E=1:DimensionY
```

```
for F=1:DimensionY
if WDecil(E,F)>= MaximaDiferencia
MaximaDiferencia=WDecil(E,F);
end
end
end
% ————————————————————————————
% Agrega las regiones a WDecil
WDecilConRegiones1=[R';WDecil];
RegionesConCero=[0;R];
WDecilConRegiones=[RegionesConCero WDecilConRegiones1];
% ————————————————————————————
% Determina las regiones con diferencia igual al máximo con alguna otra
% región
RegionesMaximaDiferencia1=[];
for G=2:DimensionY
for H=2:DimensionY
if WDecilConRegiones(G,H)==MaximaDiferencia
RegionesMaximaDiferencia1=[RegionesMaximaDiferencia1;_
WDecilConRegiones(G,:)];
break
end
end
end
RegionesMaximaDiferencia=[RegionesConCero';_
RegionesMaximaDiferencia1];
% ————————————————————————————
%Ajuste para hallar regiones a analizar (numero de la región en la
%diagonal)
DimensionRegionesMaximaDiferencia=size(RegionesMaximaDiferencia);
for I=2:DimensionRegionesMaximaDiferencia
RegionesMaximaDiferencia(I,RegionesMaximaDiferencia(I,1)+1)=_
RegionesMaximaDiferencia(I,1);
end
% ————————————————————————————
% Agrega las regiones a W
size(R');
DimensionWIteracionMayorA1=size (W);
```

```
WRegiones1=[R';W];
RegionesConCero=[0;R];
WConRegiones=[RegionesConCero WRegiones1];
% ———————————————————————————
% Halla las regiones a analizar
RegionesAAnalizar=[];
SumaColumna=0;
for J=2:DimensionRegionesMaximaDiferencia(1,2)
SumaColumna=0;
for K=2:DimensionRegionesMaximaDiferencia(1,1)
SumaColumna=SumaColumna+RegionesMaximaDiferencia(K,J);
end
if SumaColumna~=0
RegionesAAnalizar=[RegionesAAnalizar;_
RegionesMaximaDiferencia(1,J)];
end
end
%RegionesAAnalizar
% ————————————————————————————-
% Elimina las filas de W que no sean las que se van a analizar
DimensionRegionesAAnalizar=size(RegionesAAnalizar);
WConRegionesAAnalizar=[];
for L=1:DimensionRegionesAAnalizar
for M=2:DimensionY
if WConRegiones(M,1)==RegionesAAnalizar(L,1)
WConRegionesAAnalizar=[WConRegionesAAnalizar;_
WConRegiones(M,:)];
end
end
end
WConRegionesAAnalizar=[RegionesConCero';_
WConRegionesAAnalizar];
DimensionWConRegionesAAnalizar=size(WConRegionesAAnalizar);
% ———————————————————————————
% Reemplaza cada entrada de WConRegionesAAnalizar con el valor
% de la región de la columna
for N=2:DimensionWConRegionesAAnalizar(1,1)
for O=2:DimensionY+1
```

```
if WConRegionesAAnalizar(N,O)==1
for P=1:DimensionY
if RYD(P,1)==WConRegionesAAnalizar(1,O)
WConRegionesAAnalizar(N,O)=RYD(P,2);
end
end
end
end
end
% ————————————————————————————————
% Suma las filas de WConRegionesAAnalizar
VecinosAgregados=[];
SumaFila=0;
for Q=2:DimensionWConRegionesAAnalizar(1,1)
for Ra=2:DimensionY+1
SumaFila=SumaFila+WConRegionesAAnalizar(Q,Ra);
end
VecinosAgregados=[VecinosAgregados;SumaFila];
SumaFila=0;
end
% ————————————————————————————————
% Agrega las regiones a VecinosAgregados
VecinosAgregados=[0;VecinosAgregados];
VecinosAgregados=[WConRegionesAAnalizar(:,1) _
VecinosAgregados];
VecinosAgregados(1,:)=[];
DimensionVecinosAgregados=size(VecinosAgregados);
% ————————————————————————————————
% Elimina las columnas de WConRegionesAAnalizar que no pertenezcan
% a las regiones a analizar y crea WSoloConRegionesAAnalizar
WSoloConRegionesAAnalizar=[WConRegionesAAnalizar(:,1)];
for S=1:DimensionVecinosAgregados(1,1)
for T=2:DimensionY+1
if VecinosAgregados(S,1)==WConRegionesAAnalizar(1,T)
WSoloConRegionesAAnalizar=[WSoloConRegionesAAnalizar _
WConRegionesAAnalizar(:,T)];
end
end
```

```
end
% ————————————————————————————————
% Reemplaza en WSoloConRegionesAAnalizar cualquier número
% diferente de 0 por la suma de las regiones vecinas de la region de la
% fila y la columna menos el valor de la region de la fila menos el valor
% de la columna
ValorFila=0;
ValorColumna=0;
ValorFila1=0;
ValorColumna1=0;
for U=2:DimensionRegionesAAnalizar
for V=2:DimensionRegionesAAnalizar
if WSoloConRegionesAAnalizar(U,V)~=0
for Wa=1:DimensionVecinosAgregados(1,1)
if VecinosAgregados(Wa,1)==WSoloConRegionesAAnalizar(U,1)
ValorFila=VecinosAgregados(Wa,2);
end
end
for X=1:DimensionVecinosAgregados(1,1)
if VecinosAgregados(X,1)==WSoloConRegionesAAnalizar(1,V)
ValorColumna=VecinosAgregados(X,2);
end
end
for Ya=1:DimensionY
if RYD(Ya,1)==WSoloConRegionesAAnalizar(U,1)
ValorFila1=RYD(Ya,2);
end
end
for Z=1:DimensionY
if RYD(Z,1)==WSoloConRegionesAAnalizar(1,V)
ValorColumna1=RYD(Z,2);
end
end
WSoloConRegionesAAnalizar(U,V)= ValorFila + ValorColumna - _
ValorFila1 - ValorColumna1;
ValorFila=0;
ValorColumna=0;
ValorFila1=0;
```

```
ValorColumna1=0;
end
end
end
% ———————————————————————————
% Determina la pendiente de la recta inicial
Beta=(1/(Y'*Y))*Y'*W*Y;
if Iteracion==1
BetaOriginal=Beta;
end
% ———————————————————————————
% Evalua la distancia relativa entre el nuevo punto y el valor en la recta.
% Esto se divide entre el valor del eje de las abscisas.
WConRegionesAAnalizar=zeros(DimensionRegionesAAnalizar(1,1)+1);
for AA=2:DimensionRegionesAAnalizar
for AB=2:DimensionRegionesAAnalizar
if WSoloConRegionesAAnalizar(AA,AB)~=0
for AC=1:DimensionY
if RYD(AC,1)==WSoloConRegionesAAnalizar(AA,1)
ValorRegionFila=RYD(AC,2);
end
end
for AD=1:DimensionY
if RYD(AD,1)==WSoloConRegionesAAnalizar(1,AB)
ValorRegionColumna=RYD(AD,2);
end
end
WConRegionesAAnalizar(AA,AB)=_
abs(WSoloConRegionesAAnalizar(AA,AB)-Beta*(ValorRegionFila+_
ValorRegionColumna))/abs(ValorRegionFila+ValorRegionColumna);
end
end
end
WConRegionesAAnalizar(:,1)=[];
WConRegionesAAnalizar(1,:)=[];
WConRegionesAAnalizar=[RegionesAAnalizar';WConRegionesAAnalizar];
RegionesConCero=[0;RegionesAAnalizar];
WConRegionesAAnalizar=[RegionesConCero WConRegionesAAnalizar];
```

```
% ————————————————————————————
% Corrección de la matriz WConRegionesAAnalizar para garantizar
% agregaciones en el mismo decil. Crea
% WConRegionesAAnalizarMismosDeciles y WFinal
WConRegionesAAnalizarMismosDeciles=_
zeros(DimensionRegionesAAnalizar+1);
for AC=2:DimensionRegionesAAnalizar+1
for AD=2:DimensionRegionesAAnalizar+1
for AE=1:DimensionY
if RYD(AE,1)== WConRegionesAAnalizar(AC,1)
DecilDeLaFila=RYD(AE,3);
end
end
for AF=1:DimensionY
if RYD(AF,1)== WConRegionesAAnalizar(1,AD)
DecilDeLaColumna=RYD(AF,3);
end
end
if DecilDeLaFila==DecilDeLaColumna
WConRegionesAAnalizarMismosDeciles(AC,AD)=1;
end
end
end
PrimeraColumna=[0;RegionesAAnalizar];
WConRegionesAAnalizarMismosDeciles(:,1)=PrimeraColumna(:,1);
WConRegionesAAnalizarMismosDeciles(1,:)=PrimeraColumna(:,1)';
WFinal=zeros(DimensionRegionesAAnalizar+1);
for AG=2:DimensionRegionesAAnalizar+1
for AH=2:DimensionRegionesAAnalizar+1
WFinal(AG,AH)=WConRegionesAAnalizar(AG,AH)*_
WConRegionesAAnalizarMismosDeciles(AG,AH);
end
end
WFinal(:,1)=PrimeraColumna(:,1);
WFinal(1,:)=PrimeraColumna(:,1)';
% ————————————————————————————
% Determina las regiones a agregar. Crea RegionesAAgregar.
Maximo=0;
```

```
for AI=2:DimensionRegionesAAnalizar+1
for AJ=2:DimensionRegionesAAnalizar+1
if WFinal(AI,AJ)>Maximo
Maximo=WFinal(AI,AJ);
RegionAAgregar1=WFinal(AI,1);
RegionAAgregar2=WFinal(1,AJ);
end
end
end
RegionesAAgregar=[RegionAAgregar1 RegionAAgregar2]
% ————————————————————————————————————
% Agregación de datos
RY=[R Y];
RConCero=[0;R];
W=[R W];
W=[RConCero';W];
for AL=1:DimensionY(1,1)
if RY(AL,1)==RegionAAgregar2
ValorRegion2=RY(AL,2);
FilaAEliminar=AL;
end
end
for AK=1:DimensionY(1,1)
if RY(AK,1)==RegionAAgregar1
ValorRegion1=RY(AK,2);
RY(AK,2)=ValorRegion1+ValorRegion2;
RY(AK,1)= DimensionYOriginal(1,1)+Iteracion;
end
end
RY(FilaAEliminar,:)=[];
R=RY(:,1);
Y=RY(:,2);
for AM=2:DimensionY(1,1)+1
if W(AM,1)==RegionAAgregar2
FilaWAAgregar2=W(AM,:);
FilaWAEliminar=AM;
end
end
```

```
for AN=2:DimensionY(1,1)+1
if W(AN,1)==RegionAAgregar1
FilaWAAgregar1=W(AN,:);
W(AN,:)=FilaWAAgregar1+FilaWAAgregar2;
W(AN,1)= DimensionYOriginal(1,1)+Iteracion;
end
end
W(FilaWAEliminar,:)=[];
for AO=2:DimensionY(1,1)+1
if W(1,AO)==RegionAAgregar2
ColumnaWAAgregar2=W(:,AO);
ColumnaWAEliminar=AO;
end
end
for AP=2:DimensionY(1,1)+1
if W(1,AP)==RegionAAgregar1
ColumnaWAAgregar1=W(:,AP);
W(:,AP)=ColumnaWAAgregar1+ColumnaWAAgregar2;
W(1,AP)= DimensionYOriginal(1,1)+Iteracion;
end
end
W(:,ColumnaWAEliminar)=[];
for AQ=2:DimensionY(1,1)+1
W(AQ,AQ)=0;
end
W(1,:)=[];
W(:,1)=[];
DimensionWCorregida = size(W);
FilaWACorregir=DimensionWCorregida(1,1);
W(FilaWACorregir,:)=[];
W(:,FilaWACorregir)=[];
% ------------------------------------------------
% Resumen
ResumenRegionesAAgregar=_
[ResumenRegionesAAgregar;RegionesAAgregar];
ResumenBeta=[ResumenBeta;Beta];
ResumenDiferenciaPrimerBeta=[ResumenDiferenciaPrimerBeta;_
(Beta - BetaOriginal)/BetaOriginal];
```

```
Resumen=[ResumenRegionesAAgregar ResumenBeta _
ResumenDiferenciaPrimerBeta];
% ————————————————————————————
end
Resumen=[ResumenRegionesAAgregar ResumenBeta _
ResumenDiferenciaPrimerBeta]
%plot(Y,W*Y,'o',Y,BetaOriginal*Y,Y,min(ResumenBeta)*Y)
subplot(2,2,1)
plot(YOriginal,WOriginal*YOriginal,'o',YOriginal,BetaOriginal*_
YOriginal,YOriginal,min(ResumenBeta)*YOriginal)
subplot(2,2,2)
plot(Y,W*Y,'x',YOriginal,BetaOriginal*YOriginal,YOriginal,_
min(ResumenBeta)*YOriginal)
subplot(2,2,3)
plot(YOriginal,WOriginal*YOriginal,'o',Y,W*Y,'x',_
YOriginal,BetaOriginal*YOriginal,YOriginal,min(ResumenBeta)*YOriginal)
subplot(2,2,4)
plot(Resumen(:,3))
end
% FIN DEL PROGRAMA PARA AGREGACIÓN OPTIMA DE
% REGIONES ————————————————————————
```

Apéndice

Validación del Algoritmo - MatLab

```
function AutocorrelacionEspacialEspuria(Y)
    clc
    % INICIO DE PARAMETRIZACIÓN ————————————————
————
    % Se requiere:
    % Y: Vector de valores de la variable dependiente
    % R: Vector con la numeración de cada región (va de 1 a R)
    % W: Matriz de contigüidades
    % NumeroDeAgregaciones: Número de agregaciones que se quieren hacer
    % FIN DE PARAMETRIZACIÓN ————————————————————
    % INICIO DEL PROGRAMA PARA AGREGACIÓN OPTIMA DE
    % REGIONES ——————————————————————-
    % Preparativos para iteraciones
    % Copia de datos iniciales
    YOriginal=Y;
    WOriginal=W;
    ROriginal=R;
    DimensionYOriginal=size(YOriginal);
    ResumenRegionesAAgregar=[];
    ResumenBeta=[];
    ResumenDiferenciaPrimerBeta=[];
    Resumen=[];
```

```
% ———————————————————————————————————————————-
for Iteracion=1:NumeroDeAgregaciones
% Determina los deciles a los que pertenece cada región
D= ceil(10*tiedrank(Y)/length(Y));
%———————————————————————————————————————————
% Crea una matriz con número de la región, valor de la región y decil al
% que pertenece
RYD=[R Y D];
% ———————————————————————————————————————————-
% Crea una matriz con la diferencia entre deciles de regiones contiguas
DimensionY=size(Y);
WDecil=zeros(DimensionY(1,1));
DimensionYFilas=DimensionY(1,1);
for A=1:DimensionYFilas
for B=1:DimensionYFilas
if W(A,B)==1
for C=1:DimensionYFilas
if RYD(C,1)==A
DecilFila=RYD(C,3);
end
end
for D=1:DimensionYFilas
if RYD(D,1)==B
DecilColumna=RYD(D,3);
end
end
if DecilFila-DecilColumna~=0
WDecil(A,B)=abs(DecilFila-DecilColumna);
WDecil(B,A)=abs(DecilFila-DecilColumna);
else
WDecil(A,B)=-99;
WDecil(B,A)=-99;
end
end
end
end
end
% ———————————————————————————————————————————-
% Determina la máxima diferencia que existe entre todas las regiones
```

```
MaximaDiferencia=0;
for E=1:DimensionY
for F=1:DimensionY
if WDecil(E,F)>= MaximaDiferencia
MaximaDiferencia=WDecil(E,F);
end
end
end
% ------------------------------------------------------------
% Agrega las regiones a WDecil
WDecilConRegiones1=[R';WDecil];
RegionesConCero=[0;R];
WDecilConRegiones=[RegionesConCero WDecilConRegiones1];
% ------------------------------------------------------------
% Determina las regiones con diferencia igual al máximo con alguna otra
% región
RegionesMaximaDiferencia1=[];
for G=2:DimensionY
for H=2:DimensionY
if WDecilConRegiones(G,H)==MaximaDiferencia
RegionesMaximaDiferencia1=[RegionesMaximaDiferencia1;WDecilConRegiones(G,:)];
break
end
end
end
RegionesMaximaDiferencia=[RegionesConCero';_
RegionesMaximaDiferencia1];
% ------------------------------------------------------------
%Ajuste para hallar regiones a analizar (numero de la región en la
%diagonal)
DimensionRegionesMaximaDiferencia=size(RegionesMaximaDiferencia);
for I=2:DimensionRegionesMaximaDiferencia
RegionesMaximaDiferencia(I,RegionesMaximaDiferencia(I,1)+1)_
=RegionesMaximaDiferencia(I,1);
end
% ------------------------------------------------------------
% Agrega las regiones a W
size(R');
```

```
DimensionWIteracionMayorA1=size (W);
WRegiones1=[R';W];
RegionesConCero=[0;R];
WConRegiones=[RegionesConCero WRegiones1];
% ---------------------------------------------------------
% Halla las regiones a analizar
RegionesAAnalizar=[];
SumaColumna=0;
for J=2:DimensionRegionesMaximaDiferencia(1,2)
SumaColumna=0;
for K=2:DimensionRegionesMaximaDiferencia(1,1)
SumaColumna=SumaColumna+RegionesMaximaDiferencia(K,J);
end
if SumaColumna~=0
RegionesAAnalizar=[RegionesAAnalizar;_
RegionesMaximaDiferencia(1,J)];
end
end
%RegionesAAnalizar
% ---------------------------------------------------------
% Elimina las filas de W que no sean las que se van a analizar
DimensionRegionesAAnalizar=size(RegionesAAnalizar);
WConRegionesAAnalizar=[];
for L=1:DimensionRegionesAAnalizar
for M=2:DimensionY
if WConRegiones(M,1)==RegionesAAnalizar(L,1)
WConRegionesAAnalizar=[WConRegionesAAnalizar;_
WConRegiones(M,:)];
end
end
end
WConRegionesAAnalizar=[RegionesConCero';_
WConRegionesAAnalizar];
DimensionWConRegionesAAnalizar=size(WConRegionesAAnalizar);
% ---------------------------------------------------------
% Reemplaza cada entrada de WConRegionesAAnalizar con el valor de
la
% región de la columna
```

```
for N=2:DimensionWConRegionesAAnalizar(1,1)
for O=2:DimensionY+1
if WConRegionesAAnalizar(N,O)==1
for P=1:DimensionY
if RYD(P,1)==WConRegionesAAnalizar(1,O)
WConRegionesAAnalizar(N,O)=RYD(P,2);
end
end
end
end
end
% ------------------------------------------------------------
% Suma las filas de WConRegionesAAnalizar
VecinosAgregados=[];
SumaFila=0;
for Q=2:DimensionWConRegionesAAnalizar(1,1)
for Ra=2:DimensionY+1
SumaFila=SumaFila+WConRegionesAAnalizar(Q,Ra);
end
VecinosAgregados=[VecinosAgregados;SumaFila];
SumaFila=0;
end
% ------------------------------------------------------------
% Agrega las regiones a VecinosAgregados
VecinosAgregados=[0;VecinosAgregados];
VecinosAgregados=[WConRegionesAAnalizar(:,1) VecinosAgregados];
VecinosAgregados(1,:)=[];
DimensionVecinosAgregados=size(VecinosAgregados);
% ------------------------------------------------------------
% Elimina las columnas de WConRegionesAAnalizar que no pertenezcan
% a las regiones a analizar y crea WSoloConRegionesAAnalizar
WSoloConRegionesAAnalizar=[WConRegionesAAnalizar(:,1)];
for S=1:DimensionVecinosAgregados(1,1)
for T=2:DimensionY+1
if VecinosAgregados(S,1)==WConRegionesAAnalizar(1,T)
WSoloConRegionesAAnalizar=[WSoloConRegionesAAnalizar _
WConRegionesAAnalizar(:,T)];
end
```

```
end
end
% ————————————————————————————————————————-
% Reemplaza en WSoloConRegionesAAnalizar cualquier número
% diferente de 0 por la suma de las regiones vecinas de la region de la fila
% y la columna menos el valor de la region de la fila menos el valor de
% la columna
ValorFila=0;
ValorColumna=0;
ValorFila1=0;
ValorColumna1=0;
for U=2:DimensionRegionesAAnalizar
for V=2:DimensionRegionesAAnalizar
if WSoloConRegionesAAnalizar(U,V)~=0
for Wa=1:DimensionVecinosAgregados(1,1)
if VecinosAgregados(Wa,1)==WSoloConRegionesAAnalizar(U,1)
ValorFila=VecinosAgregados(Wa,2);
end
end
for X=1:DimensionVecinosAgregados(1,1)
if VecinosAgregados(X,1)==WSoloConRegionesAAnalizar(1,V)
ValorColumna=VecinosAgregados(X,2);
end
end
for Ya=1:DimensionY
if RYD(Ya,1)==WSoloConRegionesAAnalizar(U,1)
ValorFila1=RYD(Ya,2);
end
end
for Z=1:DimensionY
if RYD(Z,1)==WSoloConRegionesAAnalizar(1,V)
ValorColumna1=RYD(Z,2);
end
end
WSoloConRegionesAAnalizar(U,V)= ValorFila + ValorColumna - _
ValorFila1 - ValorColumna1;
ValorFila=0;
ValorColumna=0;
```

```
ValorFila1=0;
ValorColumna1=0;
end
end
end
% ——————————————————————————-
% Determina la pendiente de la recta inicial
Beta=(1/(Y'*Y))*Y'*W*Y;
if Iteracion==1
BetaOriginal=Beta;
end
% ——————————————————————————-
% Evalua la distancia relativa entre el nuevo punto y el valor en la recta.
% Esto se divide entre el valor del eje de las abscisas.
WConRegionesAAnalizar=zeros(DimensionRegionesAAnalizar(1,1)+1);
for AA=2:DimensionRegionesAAnalizar
for AB=2:DimensionRegionesAAnalizar
if WSoloConRegionesAAnalizar(AA,AB)~=0
for AC=1:DimensionY
if RYD(AC,1)==WSoloConRegionesAAnalizar(AA,1)
ValorRegionFila=RYD(AC,2);
end
end
for AD=1:DimensionY
if RYD(AD,1)==WSoloConRegionesAAnalizar(1,AB)
ValorRegionColumna=RYD(AD,2);
end
end
WConRegionesAAnalizar(AA,AB)=_
abs(WSoloConRegionesAAnalizar(AA,AB)-Beta*(ValorRegionFila+_
ValorRegionColumna))/abs(ValorRegionFila+ValorRegionColumna);
end
end
end
WConRegionesAAnalizar(:,1)=[];
WConRegionesAAnalizar(1,:)=[];
WConRegionesAAnalizar=[RegionesAAnalizar';_
WConRegionesAAnalizar];
```

```
RegionesConCero=[0;RegionesAAnalizar];
WConRegionesAAnalizar=[RegionesConCero _
WConRegionesAAnalizar];
% ——————————————————————————————————————-
% Corrección de la matriz WConRegionesAAnalizar para garantizar
% agregaciones en el mismo decil. Crea
% WConRegionesAAnalizarMismosDeciles y WFinal
WConRegionesAAnalizarMismosDeciles=_
zeros(DimensionRegionesAAnalizar+1);
for AC=2:DimensionRegionesAAnalizar+1
for AD=2:DimensionRegionesAAnalizar+1
for AE=1:DimensionY
if RYD(AE,1)== WConRegionesAAnalizar(AC,1)
DecilDeLaFila=RYD(AE,3);
end
end
for AF=1:DimensionY
if RYD(AF,1)== WConRegionesAAnalizar(1,AD)
DecilDeLaColumna=RYD(AF,3);
end
end
if DecilDeLaFila==DecilDeLaColumna
WConRegionesAAnalizarMismosDeciles(AC,AD)=1;
end
end
end
PrimeraColumna=[0;RegionesAAnalizar];
WConRegionesAAnalizarMismosDeciles(:,1)=PrimeraColumna(:,1);
WConRegionesAAnalizarMismosDeciles(1,:)=PrimeraColumna(:,1)';
WFinal=zeros(DimensionRegionesAAnalizar+1);
for AG=2:DimensionRegionesAAnalizar+1
for AH=2:DimensionRegionesAAnalizar+1
WFinal(AG,AH)=WConRegionesAAnalizar(AG,AH)*_
WConRegionesAAnalizarMismosDeciles(AG,AH);
end
end
WFinal(:,1)=PrimeraColumna(:,1);
WFinal(1,:)=PrimeraColumna(:,1)';
```

```
% ----------------------------------------------------------------
% Determina las regiones a agregar. Crea RegionesAAgregar.
Maximo=0;
for AI=2:DimensionRegionesAAnalizar+1
for AJ=2:DimensionRegionesAAnalizar+1
if WFinal(AI,AJ)>Maximo
Maximo=WFinal(AI,AJ);
RegionAAgregar1=WFinal(AI,1);
RegionAAgregar2=WFinal(1,AJ);
end
end
end
RegionesAAgregar=[RegionAAgregar1 RegionAAgregar2]
% ----------------------------------------------------------------
% Agregación de datos
RY=[R Y];
RConCero=[0;R];
W=[R W];
W=[RConCero';W];
for AL=1:DimensionY(1,1)
if RY(AL,1)==RegionAAgregar2
ValorRegion2=RY(AL,2);
FilaAEliminar=AL;
end
end
for AK=1:DimensionY(1,1)
if RY(AK,1)==RegionAAgregar1
ValorRegion1=RY(AK,2);
RY(AK,2)=ValorRegion1+ValorRegion2;
RY(AK,1)= DimensionYOriginal(1,1)+Iteracion;
end
end
RY(FilaAEliminar,:)=[];
R=RY(:,1);
Y=RY(:,2);
for AM=2:DimensionY(1,1)+1
if W(AM,1)==RegionAAgregar2
FilaWAAgregar2=W(AM,:);
```

```
FilaWAEliminar=AM;
end
end
for AN=2:DimensionY(1,1)+1
if W(AN,1)==RegionAAgregar1
FilaWAAgregar1=W(AN,:);
W(AN,:)=FilaWAAgregar1+FilaWAAgregar2;
W(AN,1)= DimensionYOriginal(1,1)+Iteracion;
end
end
W(FilaWAEliminar,:)=[];
for AO=2:DimensionY(1,1)+1
if W(1,AO)==RegionAAgregar2
ColumnaWAAgregar2=W(:,AO);
ColumnaWAEliminar=AO;
end
end
for AP=2:DimensionY(1,1)+1
if W(1,AP)==RegionAAgregar1
ColumnaWAAgregar1=W(:,AP);
W(:,AP)=ColumnaWAAgregar1+ColumnaWAAgregar2;
W(1,AP)= DimensionYOriginal(1,1)+Iteracion;
end
end
W(:,ColumnaWAEliminar)=[];
for AQ=2:DimensionY(1,1)+1
W(AQ,AQ)=0;
end
W(1,:)=[];
W(:,1)=[];
DimensionWCorregida = size(W);
FilaWACorregir=DimensionWCorregida(1,1);
W(FilaWACorregir,:)=[];
W(:,FilaWACorregir)=[];
% ——————————————————————————-
% Resumen
ResumenRegionesAAgregar=[ResumenRegionesAAgregar;_
RegionesAAgregar];
```

```
ResumenBeta=[ResumenBeta;Beta];
ResumenDiferenciaPrimerBeta=[ResumenDiferenciaPrimerBeta;(Beta - _
BetaOriginal)/BetaOriginal];
Resumen=[ResumenRegionesAAgregar ResumenBeta _
ResumenDiferenciaPrimerBeta];
% ---------------------------------------------------------------------
end
Resumen=[ResumenRegionesAAgregar ResumenBeta _
ResumenDiferenciaPrimerBeta]
%plot(Y,W*Y,'o',Y,BetaOriginal*Y,Y,min(ResumenBeta)*Y)
subplot(2,2,1)
plot(YOriginal,WOriginal*YOriginal,'o',YOriginal,BetaOriginal*_
YOriginal,YOriginal,min(ResumenBeta)*YOriginal)
subplot(2,2,2)
plot(Y,W*Y,'x',YOriginal,BetaOriginal*YOriginal,_
YOriginal,min(ResumenBeta)*YOriginal)
subplot(2,2,3)
plot(YOriginal,WOriginal*YOriginal,'o',Y,W*Y,'x',_
YOriginal,BetaOriginal*YOriginal,YOriginal,min(ResumenBeta)*_
YOriginal)
subplot(2,2,4)
plot(Resumen(:,3))
% FIN DEL PROGRAMA PARA AGREGACIÓN OPTIMA DE
% REGIONES ------------------------------------------------------------
ResumenProcesosAleatorios=[];
for LAMBDA=1:NumeroDeProcesosAleatorios
% INICIO DE EVALUACIÓN DEL ALGORITMO ----------------------------------
W=WOriginal;
R=ROriginal;
Y=YOriginal;
WConRegiones=[R W];
RConCero=[0;R];
WConRegiones=[RConCero';WConRegiones];
for ALFA=1:NumeroDeAgregaciones
DimensionW=size(W);
% Selecciona una región en forma aleatoria: RegionAAgregar1
RegionAleatoria1=fix(rand*DimensionW(1,1))+1;
```

```
if RegionAleatoria1==1
RegionAleatoria1=fix(rand*DimensionW(1,1))+1;
end
for Gamma=2:DimensionW(1,1)+1
if WConRegiones(1,Gamma)==RegionAleatoria1
RegionAAgregar1=WConRegiones(Gamma,1);
FilaAAgregar=Gamma;
end
end
% ----------------------------------------------------------
% Selecciona una región contigua a RegionAAgregar1 en forma
% aleatoria: RegionAleatoriaAAgregar2
SumaFila=0;
for BETA=2:DimensionW(1,1)+1
SumaFila=SumaFila+WConRegiones(FilaAAgregar,BETA);
end
IndicativoVecinoAAgregar=fix(unifrnd(1,SumaFila));
SumaColumna=0;
for Delta=2:DimensionW(1,1)+1
if WConRegiones(FilaAAgregar,Delta)==1
SumaColumna=SumaColumna+1;
if SumaColumna==IndicativoVecinoAAgregar
RegionAAgregar2=WConRegiones(1,Delta);
end
end
end
% ----------------------------------------------------------
% Muestra las dos regiones aleatorias a agregar: _
RegionesAleatoriasAAgregar
RegionesAleatoriasAAgregar=[RegionAAgregar1 RegionAAgregar2];
% ----------------------------------------------------------
% Agregación de datos
RY=[R Y];
RConCero=[0;R];
W=[R W];
W=[RConCero';W];
DimensionY=size(Y);
for AL=1:DimensionY(1,1)
```

```
if RY(AL,1)==RegionAAgregar2
ValorRegion2=RY(AL,2);
FilaAEliminar=AL;
end
end
for AK=1:DimensionY(1,1)
if RY(AK,1)==RegionAAgregar1
ValorRegion1=RY(AK,2);
RY(AK,2)=ValorRegion1+ValorRegion2;
RY(AK,1)= DimensionYOriginal(1,1)+ALFA;
end
end
RY(FilaAEliminar,:)=[];
R=RY(:,1);
Y=RY(:,2);
for AM=2:DimensionY(1,1)+1
if W(AM,1)==RegionAAgregar2
FilaWAAgregar2=W(AM,:);
FilaWAEliminar=AM;
end
end
for AN=2:DimensionY(1,1)+1
if W(AN,1)==RegionAAgregar1
FilaWAAgregar1=W(AN,:);
W(AN,:)=FilaWAAgregar1+FilaWAAgregar2;
W(AN,1)= DimensionYOriginal(1,1)+ALFA;
end
end
W(FilaWAEliminar,:)=[];
for AO=2:DimensionY(1,1)+1
if W(1,AO)==RegionAAgregar2
ColumnaWAAgregar2=W(:,AO);
ColumnaWAEliminar=AO;
end
end
for AP=2:DimensionY(1,1)+1
if W(1,AP)==RegionAAgregar1
ColumnaWAAgregar1=W(:,AP);
```

```
W(:,AP)=ColumnaWAAgregar1+ColumnaWAAgregar2;
W(1,AP)= DimensionYOriginal(1,1)+ALFA;
end
end
W(:,ColumnaWAEliminar)=[];
for AQ=2:DimensionY(1,1)+1
W(AQ,AQ)=0;
end
W(1,:)=[];
W(:,1)=[];
DimensionWCorregida = size(W);
FilaWACorregir=DimensionWCorregida(1,1);
W(FilaWACorregir,:)=[];
W(:,FilaWACorregir)=[];
RConCero=[0;R];
WConRegiones=[R W];
WConRegiones=[RConCero';WConRegiones];
% Calcula el BetaAleatorio
if ALFA==NumeroDeAgregaciones
Beta;
BetaAleatorio=(1/(Y'*Y))*Y'*W*Y;
ResumenProcesosAleatorios=[ResumenProcesosAleatorios;_
BetaAleatorio];
end
end
end
ResumenProcesosAleatorios;
Resumen;
```

Bibliografía

Abdi, H. Least-squares. M. Lewis-Beck, A. Bryman, T. Futing (2003). *Encyclopedia for research methods for the social sciences.* Thousand Oaks (CA): Sage. pp. 792-795.

Acevedo Bohórquez, Ingrid, Velásquez Ceballos, Ermilson (2008). *Algunos conceptos de la econometría espacial y el análisis exploratorio de datos espaciales.* Colombia : Universidad EAFIT, Octubre.

Anselin, Luc. (1980). *Estimation methods for spatial autoregressive structures.* Regional Science Dissertation and Monograph Series (Ithaca, NY).

Anselin, Luc. (1988A). *Spatial econometrics: methods and models.* Kluwer Academic Publishers.

Anselin, Luc. (2001). *Spatial externalities, spatial multipliers and spatial econometrics.* Discussion Paper del Regional Economics Applications Laboratory, REAL 01-T-11

Arbia, Giuseppe (2009). *Spatial Econometrics, Methods and Applications.* Studies in Empirical Economics.

Casella, G. and Berger, R.L. (2002). *Statistical Inference.* 2nd edition. Duxbury Press.

Chasco, Coro. (2003). *Econometría espacial aplicada a la predicción-extrapolación de datos espaciales.* Ed. Comunidad de Madrid. Madrid.

Cliff, A.D., and J.K. Ord. (1981) *Spatial processes: models and applications.* Taylor& Francis.

Florax, R. y H. Folmer (1992). *Specification and estimation of spatial linear regression models: Monte Carlo evaluation of pre-test estimators.* Regional Science and Urban Economics, 22; pp. 405-32

James Le Sage, Robert K. Pace (2009). *Introduction to spatial econometrics.* CRC Press Inc., 354 pp

Jesús Mur, Ana Angulo (2008). *Procesos Estocásticos Espaciales.* Facultad de Ciencias Económicas y Empresariales, Universidad de Zaragoza.

Mendenhall W, Wackerly D, Sheaffer R. (1994). *Estadística Matemática con Aplicaciones.* Grupo Editorial ibero América.

Moreno, R. and E. Vayá (2000). *Técnicas econométricas para el tratamiento de datos espaciales: La Econometría Espacial.* Edicions Universitat de Barcelona, n. 44. (Econometric tools when dealing with spatial data: Spatial Econometrics).

Muirhead, Robb J. (2008). *The Multivariate Normal and Related Distributions, in Aspects of Multivariate Statistical Theory.* John Wiley & Sons, Inc., Hoboken, NJ, USA.

Ord, J.K. (1975). *Estimation methods for models of spatial interaction.* Journal of the American Statistical Association, 70; pp. 120-26.

Paelinck, J.H.P. and Klaassen, L.H., (1979). *Spatial Econometrics.* Saxon House Farnborough.

Ron C. Mittelhammer (1996). *Mathematical Statistics for Economics and Business.* Published by Springer-Verlag .

Tobler, W. (1979), *Cellular Geography.* Philosophy in Geography, ed. S. Gale y G. Olsson, Dordrecht: Reidel; pp. 379-86.