ESTRUCTURA DE AUTOCORRELACIÓN ESPACIAL DE LA ACTIVIDAD COMERCIAL EN LOS MUNICIPIOS DE LA REGIÓN DE MURCIA

Juana María Vivo Molina[1], Jorge Chica Olmo[2]
[1]*Universidad de Murcia*
[2]*Universidad de Granada*

RESUMEN

En este trabajo, se realiza una aproximación al estudio de la dependencia entre variable económica y localización espacial aplicando la metodología de la Teoría de las Variables Regionalizadas (TVR). En particular, nos centraremos en la localización de actividades económico-comerciales de los municipios de la Comunidad Autónoma de la Región de Murcia (CARM).

Aplicando métodos estadísticos para la reducción de datos que nos permiten eliminar información redundante, detectamos estructuras latentes en el desarrollo económico-comercial dentro de los municipios de la CARM.

La metodología TVR nos permite constatar la presencia de correlación espacial, considerando los municipios de la Región de Murcia georeferenciados dentro del plano topográfico, en las estructuras latentes detectadas.

1. INTRODUCCIÓN

La aplicación de la Geoestadística en el estudio del desarrollo económico es la clave de este trabajo; en el que llevaremos a cabo un análisis de la estructura espacial de autocorrelación, considerando la Teoría de las Variables Regionalizadas (TVR), en los factores latentes de desarrollo económico-comercial observados en los municipios de la Región de Murcia. Dichos factores los obtendremos aplicando la técnica multivariante: Análisis Factorial y es requisito imprescindible que estos factores puedan ser considerados variables regionalizadas, con el fin de aplicar la metodología TVR.

En particular nos detendremos en el análisis de la autocorrelación espacial de la actividad comercial-industrial en los municipios de la Comunidad Autónoma de la Región de Murcia.

2. TEORÍA DE LAS VARIABLES REGIONALIZADAS[1]

En el análisis económico espacial el concepto de variable económica espacial implica la dependencia entre localización espacial y variable económica, en este caso diremos que dicha variable es una variable económica regionalizada. La principal característica de la variable regionalizada es la presencia de autocorrelación espacial, en nuestro caso estudiaremos la presencia de autocorrelación espacial en la actividad comercial de los municipios de la Región de Murcia.

La aplicación de la TVR, nos permitirá identificar la estructura de autocorrelación espacial de la actividad comercial en los municipios y, en su caso, se podría llegar a realizar la inferencia espacial. La localización espacial se refiere a la posición relativa de cada municipio respecto de los restantes municipios, éstos aparecen georeferenciados dentro del plano topográfico de la CARM, donde se localiza la actividad comercial.

En este caso, el dominio o campo geométrico en el que se realiza el estudio de la actividad comercial es la Región de Murcia, se denota por D y el soporte estaría constituido por los municipios.

Para llevar a cabo, el tratamiento adecuado de una variable regionalizada se ha de considerar el doble aspecto que presenta: por un lado, aleatoriedad (los datos varían

[1] Este apartado representa un resumen de la obra del autor Chica Olmo (1994)

irregular e imprevisiblemente) y por otro lado, estructuración (los datos no son totalmente independientes de su localización, se dice que en puntos próximos presentan autocorrelación).

Este doble carácter de las variables regionalizadas, se puede abordar desde la interpretación probabilística de estas variables a partir de las funciones aleatorias (Matheron, 1970). La función aleatoria $Z(x)$ es función de la localización $x \in D$. Consideramos $z(x_0)$ el valor de la variable en un punto cualquiera $\boldsymbol{x}_0$ de la muestra de una variable aleatoria $Z(x_0)$. En la práctica, estudiaremos la función aleatoria (f.a.) en unas ciertas localizaciones, puntos de la muestra, esto es, analizaremos las variables aleatorias: $Z(x_1), Z(x_2), ..., Z(x_n)$; al conjunto $Z(x)$ de todas estas v. a. se le denomina función aleatoria. Una realización u observación de una función aleatoria es una función, que es la interpretación, desde el punto de vista de la Teoría de las Funciones Aleatorias, del concepto de variable regionalizada.

Una característica de la función aleatoria es la continuidad. Tenemos que una f.a. es continua en media si:

$$\lim_{x \to x_0} E[Z(x) - Z(x_0)] = 0\text{,}$$

y es continua en media cuadrática cuando:

$$\lim_{x \to x_0} E\left\{[Z(x) - Z(x_0)]^2\right\} = 0\text{.}$$

Cuando la continuidad en media cuadrática no se cumple se dice que existe "efecto pepita" y se denota por C_0:

$$\lim_{x \to x_0} E\left\{[Z(x) - Z(x_0)]^2\right\} = C_0 \neq 0$$

y si la f.a. presenta direcciones particulares de variabilidad, diremos que tiene un comportamiento anisótropo.

Es posible inferir, usando la Teoría de las Funciones Aleatorias, si se conoce la distribución de probabilidad conjunta de las v.a. $Z(x_1), Z(x_2), ..., Z(x_n)$, o por lo menos, sus primeros momentos que no se pueden obtener a partir de una única realización $z(x_1), z(x_2), ..., z(x_n)$. Por ello, para realizar inferencia estadística se introducen ciertas hipótesis restrictivas sobre $Z(x)$.

1. Hipótesis Estacionaria

Una función aleatorias se dice estacionaria estrictamente si la ley de distribución conjunta es invariante por traslación. Sea h el vector de traslación:

$$F(Z(x_1),Z(x_2),...,Z(x_n)) = F(Z(x_1+h),Z(x_2+h),...,Z(x_n+h))$$

Esta hipótesis puede relajarse requiriendo que sólo los dos primeros momentos, la media y la varianza, permanezcan invariantes mediante la traslación. Nos referiremos a ella, en adelante, con el nombre de estacionario.

Hipótesis estacionaria de 2º orden:

a) La función aleatoria es estacionaria de segundo orden si:

$$E[Z(x)] = m(x) = m = E[Z(x+h)], \forall x$$

b) Para cada par de v.a. $Z(x)$ y $Z(x+h)$ la covarianza existe y sólo depende del vector h en magnitud y dirección:

$$C(h) = E\{[Z(x)-m][Z(x+h)-m]\} = E\{Z(x)Z(x+h)\} - m^2$$

En particular, si $h = 0$:

$$C(0) = E[Z(x)-m]^2 = Var[Z(x)], \forall x$$

Así pues, bajo esta hipótesis se asume la existencia de varianza finita, generalmente desconocida y que estimaremos con el valor de la varianza experimental.

En el estudio de fenómenos aleatorios no se suele cumplir esta hipótesis, basta que presenten una capacidad de dispersión ilimitada ($C(0)$ no exista) o haya una cierta tendencia o deriva (no se puede asumir media constante, $E[Z(x)] = m(x)$). Por ello, es necesario considerar una hipótesis menos restrictiva que la anterior.

2. Hipótesis Intrínseca

Esta hipótesis considera que los incrementos de primer orden son estacionarios, es decir, la media y la varianza de los incrementos $Z(x+h) - Z(x)$, existen y son independientes del punto de apoyo x.

Una función aleatoria $Z(x)$ se dice intrínseca de orden 0 si:

a) La esperanza de los incrementos de primer orden es nula:

$$E[Z(x+h) - Z(x)] = 0$$

b) La varianza de los incrementos de primer orden existe y no depende de los puntos x y $x+h$.

$$Var[Z(x+h)-Z(x)] = 2\gamma(h)$$

$$\gamma(h) = \frac{1}{2}Var[Z(x+h)-Z(x)] = \frac{1}{2}E\{[Z(x+h)-Z(x)]^2\}$$

La función $\gamma(h)$ se denomina "semi-variograma", en general, simplemente variograma. Es un instrumento básico para la interpretación de la estructura del fenómeno, así como para su estimación.

Una función aleatoria de 2º orden es siempre intrínseca, el recíproco no es necesariamente cierto. Si el fenómeno verifica alguna de estas hipótesis es posible la estimación de los momentos de la función aleatoria a partir de una única realización.

Si se restringe el cumplimiento de las condiciones anteriores a un cierto entorno y para distancias limitadas, se dice que la f.a. es cuasiestacionaria o cuasiintrínseca.

2.1. Función Variograma

Si $Z(x)$ es estacionaria de 2º orden, se tiene la relación: $\gamma(h) = C(0) - C(h)$.

Teniendo en cuenta las siguientes propiedades:

PROPIEDADES DE LA COVARIANZA:	PROPIEDADES DEL VARIOGRAMA:
$C(0) = Var[Z(x)] \geq 0$	$\gamma(h) \geq 0; \gamma(0) = 0$
$C(h) = C(-h)$	$\gamma(h) = \gamma(-h)$
$\lvert C(h)\rvert \leq C(0)$	

Tabla 1 Propiedades de la covarianza y del variograma

se puede representar gráficamente la relación entre ambas. Se denomina alcance a la distancia a que indica la zona de influencia de la variable regionalizada, tal que $C(h) = 0, \forall h \geq a$.

Por otro lado, el valor de $\gamma(h)$ cuando $h \geq a$ recibe el nombre de meseta. Así pues, si se verifica la hipótesis estacionaria de 2º orden, el variograma será limitado, diremos que presenta un comportamiento estacionario y bajo la hipótesis intrínseca, el variograma será ilimitado, diremos que presenta un comportamiento no estacionario o deriva.

2.1.1. Variograma experimental

El estudio de la autocorrelación espacial de la variable actividad económica lo realizaremos a través del cálculo del variograma experimental.

El estimador insesgado del variograma en $\mathbb{R}^2$ a partir de la definición de $\gamma(h)$ viene dada por:

$$\gamma^*(h_\alpha) = \frac{1}{2NP(h_\alpha)} \sum_{i=1}^{NP(h_\alpha)} \left[Z(x_i + h) - Z(h)\right]^2$$

donde:

$Z(x_i)$, $Z(x_i + h)$: representan a la variable Z en las localizaciones x_i y $x_i + h$.

h_α: distancia entre pares de datos en la dirección α del plano, denominada paso del variograma. En general, se fija α para una de las cuatro direcciones del plano.

$NP(h_\alpha)$: número de pares de datos separados por una distancia h_α

En la práctica, para calcular el variograma experimental en una determinada dirección α, se selecciona una cierta tolerancia en la dirección ($\delta\alpha$, tolerancia angular) y en la distancia (εh, tolerancia de la distancia) y se usa cada dato $Z(x_i)$, junto con todos los puntos situados en el ángulo $\alpha \pm \delta\alpha$ y a una distancia $h \pm \varepsilon h$.

2.1.2. Variograma *teórico*

Una vez obtenido el variograma experimental, será necesario ajustarle un modelo de variograma teórico que cumpla ciertas condiciones de tipo matemático.

Para los variogramas teóricos es aconsejable elegir modelos conocidos o combinaciones lineales de éstos. Los modelos básicos se obtienen construyendo matemáticamente una función aleatoria y calculando su variograma teóricamente, imponiéndole la condición de que $\gamma(h)$ debe ser una función condicionalmente definida negativa (o equivalentemente $-\gamma(h)$ ha de ser condicionalmente definida positiva).

Los modelos de variogramas pueden agruparse en dos tipos: con meseta (f.a. limitados o estacionarios con crecimiento rápido en el origen que se estabiliza en torno a la meseta) y sin meseta (f.a. intrínseca sin varianza finita, que presenta una capacidad de crecimiento ilimitada).

Los modelos con meseta más habituales son el esférico o de Matheron, el modelo exponencial y el Gausiano; y entre los más utilizados sin meseta tenemos el modelo potencial.

3. APLICACIÓN A LA ACTIVIDAD COMERCIAL EN LA REGIÓN DE MURCIA

Este estudio se centra en las siguientes características económico-comerciales observadas en los municipios de la CARM: PARO[2], NIVEL ECONÓMICO, ACTIVIDAD INDUSTRIAL[2], ACTIVIDAD COMERCIAL MAYORISTA[2], ACTIVIDAD COMERCIAL MINORISTA[2], TELÉFONO[2] Y AUTOMÓVIL[2].

La base de datos, con la que trabajamos, procede del Anuario Económico de La Caixa, del que hemos seleccionado las variables mencionadas. Este anuario considera los municipios de más de 1.000 habitantes individualmente, motivo por el que no se reflejan las características económico-comerciales de Ojós y Ulea.

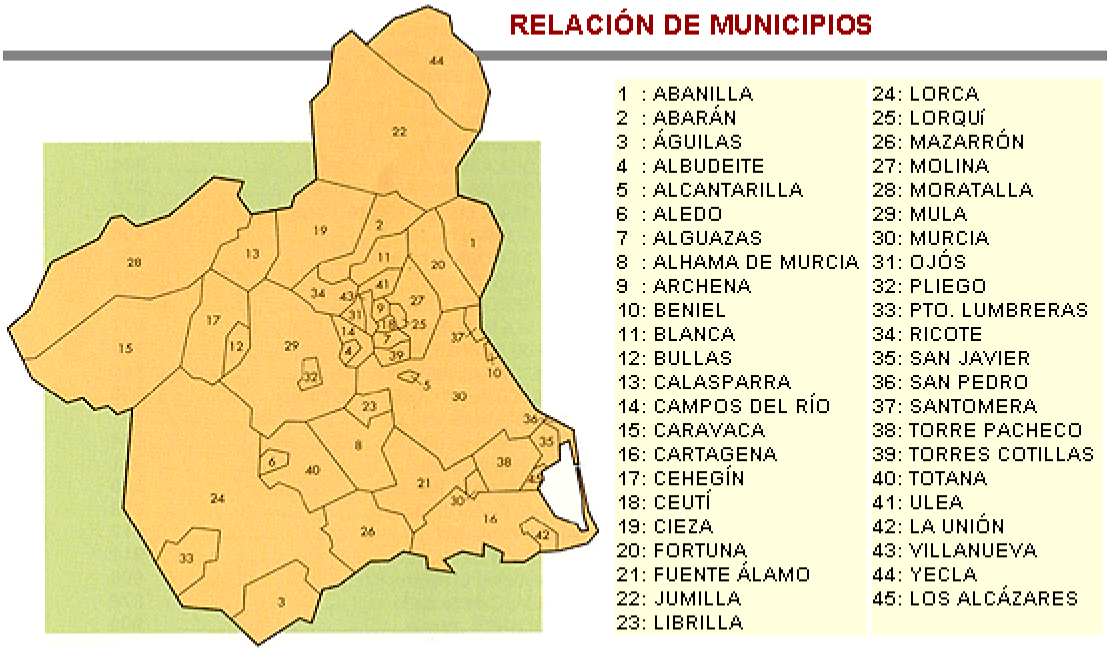

Figura 1: Mapa de municipios de la Región de Murcia

En el conjunto de datos con el que trabajamos, consideramos las coordenadas geográficas x e y de los 43 municipios. Estas coordenadas corresponden al huso en el que se encuentra la península ibérica. El origen del eje x (primera coordenada) estaría colocado en el extremo más occidental del huso y el origen del eje y en el ecuador. De esta manera, todas las coordenadas son del primer cuadrante.

[2] Estas variables han sido relativizadas por la población del Padrón de Habitantes de 1/01/2002.

3.1. Resultados del Análisis Factorial

Aplicamos un análisis de Componentes Principales a las variables económico-comercial mencionadas.

Considerando las componentes con mayor información, la técnica extrae las dos primeras, que conjuntamente explican el 65’05% de la variabilidad total de los datos.

Matriz de componentes[a]

	Componente	
	1	2
TELÉFONO	*,822*	*-,433*
NIVEL ECONÓMICO	*,804*	*-,318*
PARO	*-,781*	*-,147*
ACTIVIDAD COMERCIAL MINORISTA	*,699*	*-,390*
AUTOMÓVIL	*,677*	*,315*
ACTIVIDAD COMERCIAL MAYORISTA	*,575*	*,551*
ACTIVIDAD INDUSTRIAL	*,507*	*,472*

Método de extracción: Análisis de componentes principale

a. 2 componentes extraídos

Tabla 2 Matriz factorial sin rotar.

La primera componente (49’52%) contiene a todas las variables objeto de nuestro estudio y por ello, la interpretamos como el desarrollo económico-comercial de los municipios de la Región de Murcia en general. En lo sucesivo, nos referiremos a dicha componente como *desarrollo económico-comercial*, ver Tabla 2.

La segunda componente (15'54%) contrapone desarrollo de la actividad mayorista e industrial (cargas factoriales positivas) a desarrollo de la actividad minorista (cargas factoriales negativas). Denominaremos a esta componente *Tipología de municipios*, ver Tabla 2.

Dado que se verifican las condiciones para realizar un análisis Factorial, KMO próximo a 0’7 (aceptable) y el p-valor asociado a la Prueba de esfericidad de Bartlett es menor que 0’01; rotamos la solución inicial obtenida, mediante el método ortogonal Varimax, con el fin de formar factores independientes entre sí. y simplificar las columnas de la matriz de factores. El 36’04% de la variabilidad de los datos viene explicada por el primer factor y el 29’02% por el segundo.

Observando las cargas factoriales, de la Tabla 3, tenemos que el primer factor explicaría la actividad comercial minorista, así como las características económicas de la población. Por ello, de manera genérica lo vamos a denominar *minorista*. La actividad comercial mayorista y la actividad industrial, quedarían explicadas por el segundo factor extraído del análisis y lo vamos a denotar por *mayorista*.

Matriz de componentes rotados[a]

	Componente	
	1	**2**
TELÉFONO	*,911*	*,181*
NIVEL ECONÓMICO	*,825*	*,259*
ACTIVIDAD COMERCIAL MINORISTA	*,788*	*,137*
ACTIVIDAD COMERCIAL MAYORISTA	*,099*	*,790*
ACTIVIDAD INDUSTRIAL	*,097*	*,686*
AUTOMÓVIL	*,328*	*,671*
PARO	*-,514*	*-,606*

Método de extracción: Análisis de componentes principale

Método de rotación: Normalización Varimax con Kaiser.

a. La rotación ha convergido en 3 iteraciones.

Tabla 3 Matriz factorial rotada.

En el siguiente apartado vamos a estudiar los factores obtenidos, mediante el Análisis de Componentes Principales y el Análisis Factorial, desde la óptica de la TVR.

3.2. Análisis de la autocorrelación espacial

Para estudiar la presencia de autocorrelación espacial en las variables obtenidas en el apartado anterior, esto es, los factores, utilizaremos la función variograma. Así, para cada una de ellos se han obtenido los variogramas experimentales. La forma que adopta el variograma nos indica el tipo de estructura de autocorrelación espacial que presenta la variable analizada.

En nuestro caso se observa que en el factor 1 sin rotar (Figura 2.a) la variable no presenta una clara correlación espacial. Sin embargo, la Figura 2.b, muestra que el factor 2 si está correlacionado espacialmente, ya que a medida que aumenta la distancia entre los municipios la variabilidad se incrementa y por lo tanto la correlación disminuye.

Por otra parte, al rotar los factores, se observa que el primer factor rotado, presenta una clara correlación espacial de tipo potencial (Figura 2.c). Igual ocurre en el caso del segundo factor rotado (Figura 2.d), sólo que en este caso la estructura espacial es de tipo estacionario ya que a medida que aumenta la distancia entre los municipios, la correlación de este factor se incrementa hasta los 10 ó 15 km. aproximadamente y a partir de esta distancia, la correlación se estabiliza.

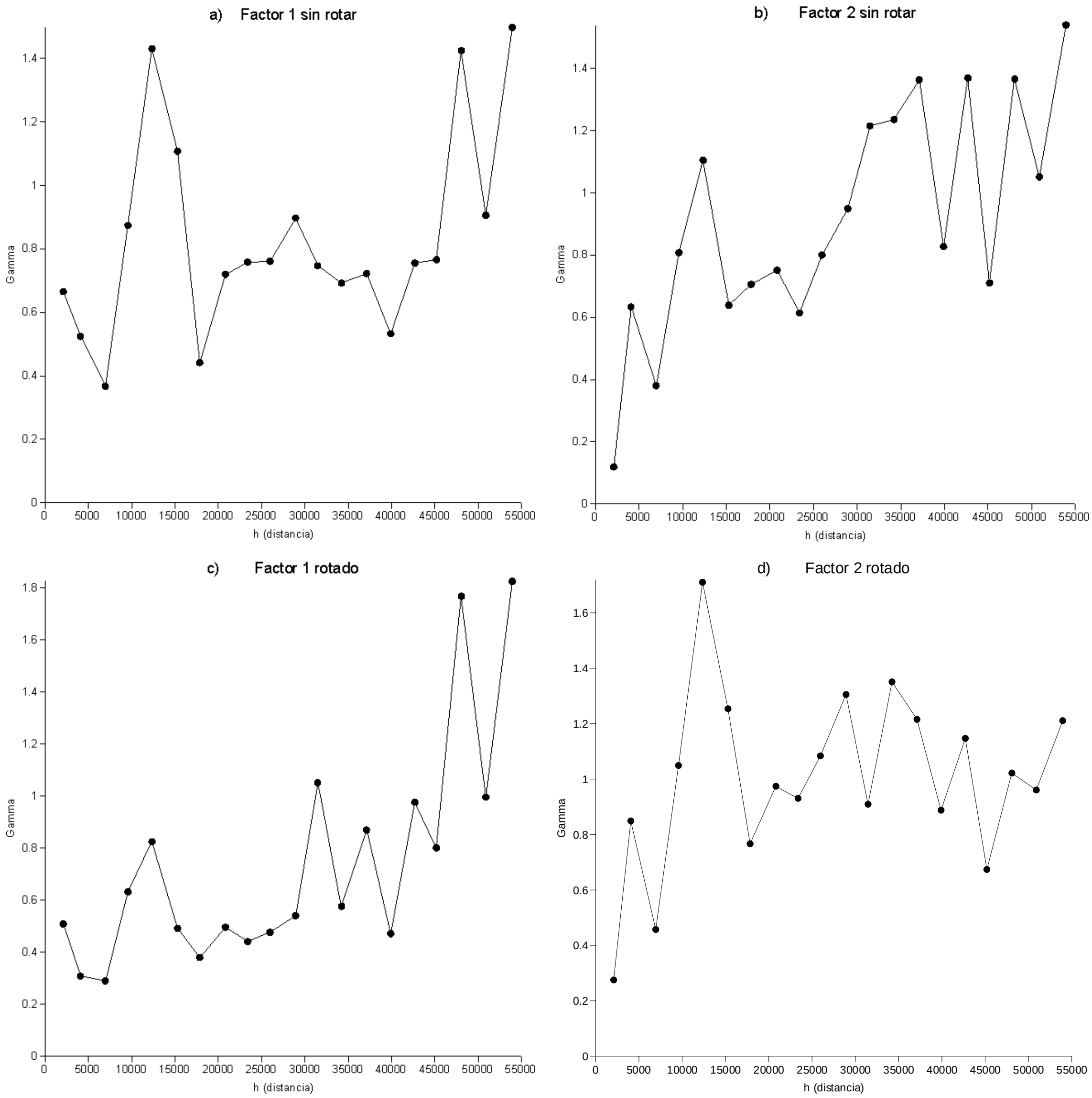

Figura 2: Variogramas de las componentes principales y de los factores rotados.

4. CONCLUSIONES

La autocorrelación espacial de las variables de desarrollo económico de los municipios, viene dada por la dependencia de estas con la localización de los mismos.

Mediante las técnicas multivariantes, Análisis de Componentes Principales y Análisis Factorial, estudiamos dichas variables de desarrollo económico y desde la óptica de la Teoría de las Variables Regionalizadas podemos concluir que:

Municipios próximos tienen *tipología* similar, es decir, desarrollan el mismo tipo de actividades, ya sea actividad comercial minorista o actividad comercial mayorista y actividad industrial.

Asimismo, conforme aumenta la distancia entre los municipios la relación entre sus características, que hemos etiquetado como *minoristas*, disminuye, o lo que es lo mismo su variabilidad incrementa, sin que se aprecie que dicha relación se estabilice. Por otra parte, si analizamos sus características *mayoristas* se aprecia que la relación también disminuye pero sólo lo hace hasta una distancia de unos 15 km. y a partir de dicha distancia, la relación entre ellas se estabiliza.

5. REFERENCIAS BIBLIOGRÁFICAS

- CANO GUERVÓS, R., CHICA OLMO, J. y HERMOSO GUTIÉRREZ, J.A. (2003). "A Geo-Statistical Method to Define Districts within a City". Journal of Real Estate Finance and Economics, 27, 1, pp. 61-85.
- CHICA OLMO, J. (1994). "Teoría de las Variables Regionalizadas. Aplicación en Economía Espacial y Valoración Inmobiliaria". Granada: Biblioteca de Económicas y Empresariales. Servicio de Publicaciones de la Universidad de Granada.
- MATHERON, G. (1970). "La Théorie des Variables Regionasées et ses applications". CMM Fas. 1 ENSMP. París.
- MORAL, F.J. y MARQUES, J.R. (2002). "Empleo de representación gráfica de una variable regionalizada". Actas de XIV Congreso Internacional de Ingeniería Gráfica. Santander.
- SERVICIOS DE ESTUDIOS DE LA CAIXA (2003). "Anuario Económico de España". Caixa.