A abordagem clássica de Econometria Espacial

Rodrigo de Souza Vieira

VIEIRA, RS. *Crescimento econômico no estado de São Paulo*: uma análise espacial [online]. São Paulo: Editora UNESP; São Paulo: Cultura Acadêmica, 2009. 103 p. ISBN 978-85-7983-013-6. Available from SciELO Books <http://books.scielo.org>.

2
A abordagem clássica de Econometria Espacial

O termo "Econometria Espacial" foi, inicialmente, introduzido por Jean Paelinck no início dos anos 70 para denominar a área do conhecimento que lida com a estimação e teste de modelos econométricos multirregionais.

A existência de uma área da Econometria denominada de Econometria Espacial se justifica, basicamente, por dois aspectos: o primeiro é a importância da questão espacial inerente à ciência regional, em particular, à economia regional. O segundo é que dados distribuídos no espaço podem apresentar dependência ou heterogeneidade em sua estrutura.

Segundo Lesage (1999), a presença de dependência espacial entre as observações, ou heterogeneidade espacial nas relações modeladas ferem os pressupostos básicos de Gauss-Markov, utilizados, de forma tradicional, em modelos de regressão.

Em termos gerais, heterogeneidade espacial significa que o comportamento econômico não é estável através do espaço, e pode gerar padrões espaciais característicos sob a forma de agrupamentos ao longo do *set* de dados, e variar com a unidade. Dessa forma, os parâmetros variam e podem mudar a forma estrutural do modelo, podendo inclusive, gerar heterocedasticidade com possíveis erros de especificação. Entretanto, Anselin (1988) aponta que, na maioria das vezes, os problemas gerados pela heterogeneidade espacial podem ser corrigidos com o uso de instrumentos fornecidos pela econometria padrão. Há casos em que o conhecimento teórico da estrutura espacial dos dados pode levar a procedimentos mais eficientes. Além disso, como afirma Resende (2005), o problema torna-se mais complexo naquelas situações em que a heterogeneidade e a autocorrelação estão presentes ao mesmo tempo. Nessas circunstâncias, as ferramentas da econometria padrão são inadequadas e exigem a utilização de técnicas da econometria espacial.

Por sua vez, a dependência ou autocorrelação espacial surge ao se questionar a independência do conjunto de dados coletados. O pressuposto-base para esse tipo de especificação está diretamente associado à primeira Lei da Geografia, na qual todas as informações são relacionadas entre si, porém informações mais próximas estão mais relacionadas do que informações distantes. Assume-se, desse modo, que a proximidade intensifica o processo de conexões entre as unidades espaciais e gera concentração em determinadas localidades em detrimento de outras.

A noção de proximidade, no entanto, é determinada por meio de uma ideia de espaço relativo, ou distância relativa, uma vez que a proximidade não precisa necessariamente estar relacionada à distância entre as localidades. Critérios distintos àquele do sentido euclidiano estrito podem ser considerados, tal como distâncias econômicas, sociais e políticas. O importante é delimitar as regras para uma potencial interação entre as localidades.

No que se refere à metodologia econométrica tradicional, a presença desses efeitos pode tanto requerer alguma modificação na mesma, como pode até invalidá-la. Em alguns casos, faz-se necessária a criação de novas técnicas para o correto tratamento desses efeitos. Como nota Anselin (1988), geralmente, essas questões são ignoradas pela teoria econométrica tradicional e formam o campo específico da Econometria Espacial.

A econometria espacial é importante não apenas quando faz parte da estrutura do modelo, mas também quando ocorrem erros de especificação nas unidades espaciais, os quais podem surgir da não coincidência entre a unidade espacial considerada e a influência do fenômeno econômico sob consideração, que pode transbordar as fronteiras preestabelecidas. De forma mais específica, a atuação das externalidades pode extrapolar o ambiente de uma cidade, não obedecendo necessariamente seus limites políticos.

Assim como existe a econometria espacial, há também o campo da estatística espacial. Anselin (1988) nota que a distinção entre esses dois campos é sutil, visto que os métodos de uma são amplamente utilizados pela outra. Segundo o autor, mais prático seria deixar a cargo dos próprios pesquisadores referirem seus trabalhos a um ou outro campo.

Em geral, a econometria espacial pauta-se em um modelo ou teoria em particular e tem, como foco, principalmente, a economia regional e urbana, enquanto a estatística espacial trata, de modo primordial, de fenômenos naturais, ligados, principalmente, a campos como a biologia e geologia. A abordagem da econometria espacial consiste basicamente em impor a estrutura do problema por meio da especificação de um modelo *a priori*, ao associá-lo a um teste de especificação com contrapartida em uma hipótese nula. Talvez essa ênfase seja a principal distinção entre a econometria espacial e o campo mais amplo da estatística espacial.

Autocorrelação espacial

Para Anselin & Bera (1998), a autocorrelação espacial pode ser definida como a coincidência entre valores similares e similaridades locacionais. Assim, quando altos ou baixos valores para uma variável aleatória tendem a agrupar-se no espaço, temos o processo de autocorrelação espacial positiva. No entanto, pode acontecer também de as unidades espaciais serem circundadas por unidades com valores significativamente distintos, ou seja, pode ocorrer que altos valores sejam acompanhados por vizinhos com valores baixos, ou vice-versa, processo que se denomina autocorrelação espacial negativa.

Embora os dois processos sejam igualmente importantes e dignos de consideração, a autocorrelação espacial positiva é, sobremaneira, a mais intuitiva, e é encontrada, com maior frequência nos fenômenos econômicos. Na maior parte das vezes, um processo que apresenta autocorrelação espacial negativa é de difícil interpretação.

Em termos práticos, uma amostra de dados espacialmente autocorrelacionada contém menos informação do que sua contrapartida não autocorrelacionada. Em termos de inferência estatística, essa perda de informação precisa ser levada em conta nos testes de estimação e de diagnóstico. Para Anselin & Bera (1998), esta é a essência do problema de autocorrelação espacial em econometria aplicada.

O problema da autocorrelação espacial tem alguma semelhança com a autocorrelação temporal. De fato, se as regiões de um determinado espaço fossem todas "enfileiradas", de tal modo que só existisse o vizinho da "frente" e o de "trás", (ou, em termos estatísticos, só pudessem apresentar dependência unidirecional) como mostra a figura abaixo, recairíamos em uma situação formalmente idêntica a das séries de tempo e, portanto, todo o tratamento econométrico seria idêntico ao das séries de tempo.

1	2	3	4

Figura 1. Espaço com dependência unidirecional.

Um espaço como o da figura acima é, com evidência, raro de se obter. O caso mais geral é ilustrado pela Figura 2 (embora, não necessariamente, com a mesma regularidade), onde os dados, regiões, estão dispostos em uma superfície bidimensional, e apresentam dependência bidirecional. Assim, a principal diferença entre a dependência temporal e a dependência espacial situa-se, principalmente, na natureza bidimensional e multidimensional da dependência no espaço.

1	2	3
4	5	6

Figura 2. Espaço com dependência multidimensional.

A autocorrelação ou dependência espacial pode ocorrer, basicamente, de duas formas: na variável dependente, ou nos erros. Formalmente, a existência de autocorrelação espacial pode ser expressa pela seguinte condição de momento:

$$\mathrm{Cov}(y_i, y_j) = \mathrm{E}(y_i, y_j) - \mathrm{E}(y_i).\mathrm{E}(y_j) \neq 0 \quad \text{para i} \neq \text{j} \tag{3.1.1}$$

Em que y_i e y_j são observações de uma variável aleatória nas localizações *i* e *j* respectivamente. *i* e *j* podem ser pontos, tais como localização de estabelecimentos ou áreas metropolitanas – medidas em latitudes e longitudes – ou unidades de área, tal como países, estados ou municípios (Anselin & Bera, 1998). É evidente que a condição estabelecida por (3.1.1) não é suficiente para que haja um processo de autocorrelação espacial, pois para tal é necessário que a correlação existente entre as observações siga um padrão intuitivo lógico em termos de estrutura espacial.

As consequências da autocorrelação espacial são, em princípio, os mesmos da autocorrelação temporal. Em um modelo de regressão, se os erros são correlacionados entre si (temporal ou espacialmente), os estimadores de mínimos quadrados ordinários são ineficientes, e os estimadores das variâncias serão viesados, o que invalida os testes de significância. Por um lado, para o caso de autocorrelação na variável dependente, as estimativas de MQO são viesadas e inconsistentes, por outro lado, quando a correlação está presente no termo de erro, não há viés, nem inconsistência, mas o estimador de MQO deixa de ser o mais eficiente.

Os processos de autocorrelação espacial guardam analogia com os de séries de tempo, de modo que a situação de autocorrelação serial de ordem 1 pode ser representada da seguinte forma:

$$z_t = \ {}_t + \rho\, z_{t-1}, \tag{3.1.2}$$

em que μ_t é um ruído branco e ρ é o coeficiente de correlação. Em contrapartida, a autocorrelação espacial, também de ordem 1, é mostrada abaixo:

$$z = \mu + \rho\, \mathbf{W}_1\, z \tag{3.1.3}$$

No caso, z é um vetor *n* por 1 de observações sobre a variável dependente, W_1z é um vetor *n* por *n* de defasagens espaciais para a variável dependente, ρ é o coeficiente autorregressivo espacial, e μ é um vetor *n* por 1 de termos de erro distribuídos aleatoriamente, ou seja, $\mu \sim (0,\sigma^2 I)$. Esse processo é conhecido como SAR (*spatial autoregressive*), onde $\mathbf{W}_1$ é a matriz de conectividade que, em geral, contém relações de contiguidade de 1ª ordem ou funções de distância.[1] Em linhas gerais, $\mathbf{W}_1$ é montada de modo a captar a influência dos vizinhos na variável em consideração. Esse é, portanto, um SAR (1).

Mais genericamente, pode-se ter também um SARMA (*spatial autoregressive moving average*). Segue abaixo um SARMA(1,1).

$$z = \mu + \rho \mathbf{W}_1 z + \theta \mathbf{W}_1 \mu \qquad (3.1.4)$$

Que pode facilmente incluir ordens superiores, e basta, para tal, incluir as respectivas matrizes de conectividade. Por exemplo, o processo abaixo seria um SAR(2).

$$z = \mu + \rho_1 \mathbf{W}_1 z + \rho_2 \mathbf{W}_2 z \qquad (3.1.5)$$

O índice global de Moran (I) é, segundo Anselin & Florax (1995), uma das formas mais amplamente utilizadas de se medir a autocorrelação espacial. Essa estatística varia entre –1 e 1, fornecendo uma medida geral da associação linear (espacial) entre os vetores Z_t no tempo *t* e a média ponderada dos valores da vizinhança, ou *lags* espaciais (WZ_t). Valores próximos de zero indicam inexistência de autocorrelação espacial significativa: quanto mais próximo do valor unitário, mais autocorrelacionado estará. Se o valor dessa estatística for positivo (negativo), a autocorrelação será positiva (negativa). Esse indicador é uma forma de detectar similaridade entre as áreas e é dado por:

$$I = \left(\frac{n}{S_0}\right)\left(\frac{Z'WZ}{Z'Z}\right) \qquad (3.1.6)$$

1 Uma discussão detalhada sobre a matriz de conectividade será realizada no tópico A matriz de pesos espaciais (capítulo 2).

Onde Z é o vetor de n observações para o desvio em relação à média, e S_0 é um escalar igual à soma de todos os elementos de W.

Sendo o valor esperado:

$$E(I) = -\frac{1}{n-1} \qquad (3.1.7)$$

Quando a matriz de pesos espaciais é normalizada na linha, ou seja, quando a soma dos elementos de cada linha for igual a um, a expressão poderá ser reescrita, como segue:

$$I = \frac{Z'WZ}{Z'Z} \qquad (3.1.8)$$

A estatística I de Moran fornece uma indicação formal do grau de associação linear entre os valores do vetor Z e o vetor espacialmente defasado WZ. Valores maiores do que aqueles esperados, E(I), indicam autocorrelação espacial positiva; negativa, caso contrário.

O diagrama de dispersão de Moran compara os valores normalizados do atributo em uma área com a média normalizada dos vizinhos, o que deriva um gráfico bidimensional de Z(valores normalizados) por WZ (média dos vizinhos). É uma forma de visualizar a dependência espacial e indicar os diferentes padrões espaciais presentes nos dados. O gráfico abaixo representa quatro quadrantes Q_1, Q_2, Q_3 e Q_4 que irão corresponder a quatro padrões de associação local espacial entre as regiões e seus vizinhos.

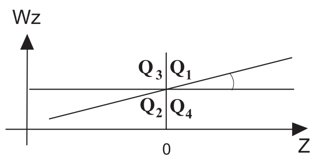

Figura 3. Diagrama de Moran.

O coeficiente I de Moran será a inclinação da curva de regressão de WZ contra Z e indicará o grau de ajustamento. O primeiro quadrante, Q_1, conhecido como alto-alto (AA), ou *high-high* – (HH), mostra regiões com altos valores para a variável, valores acima da média, assim como seus vizinhos. O terceiro quadrante, Q_2, geralmente chamado de baixo-baixo (BB) ou *low-low* – (LL), expressa localidades com baixos valores em relação aos atributos analisados, acompanhados por vizinhos que também apresentam baixos valores. O segundo quadrante, Q_3, classificado como baixo-alto (BA) ou *low-high* – (LH), é constituído por baixos valores dos atributos na região estudada, cercada por vizinhos com altos valores. O último quadrante, Q_4, é formado por regiões com altos valores para as variáveis estudadas cercadas por regiões com baixos valores. Este é o quadrante alto-baixo (AB) ou *high-low* (HL).

As regiões de *clusters* com valores similares ocorrem nos quadrantes Q_1 e Q_2 – AA e BB – e apresentam autocorrelação espacial positiva. As regiões identificadas pelos quadrantes Q_3 e Q_4 – BA e AB – apresentam, por sua vez, autocorrelação espacial negativa, ou seja, *clusters* com valores diferentes.

Adicionalmente, a estatística I tem sido usada como um teste para a presença de autocorrelação espacial residual, em linha com a estatística de Durbin-Watson para séries de tempo. Nesse caso, o teste I de Moran é aplicado sobre as estimativas dos erros de uma regressão feita por MQO, com a estatística I observada, comparada com uma distribuição aleatória aproximada por seus momentos, sob a hipótese nula de nenhuma correlação residual. Tiefelsdorf & Boots (1995) fornecem os momentos exatos.

Além da estatística I de Moran, aplicada aos resíduos de uma regressão linear, a presença de algum grau de dependência espacial pode ser verificada por meio de alguns testes específicos, entre eles, o teste de Wald, Razão de Verossimilhança (*Likelihood Ratio – LR*) e através de uma família de testes baseada no Multiplicador de Lagrange (*Lagrange Multiplier – LM*).

Os testes de Multiplicador de Lagrange (LM)[2] são, inclusive, os mais indicados por Anselin (2003) para a escolha da especificação

2 Burridge (1980).

mais adequada. Maiores detalhes sobre testes de especificação e escolha dos modelos serão tratados no tópico Testes de especificação dos modelos espaciais (capítulo 2).

Modelos de regressão com dependência espacial

Segundo Lesage (1999), um modelo autorregressivo espacial mais geral (*spatial autoregressive model* – SAC) pode ser representado da seguinte forma:

$$y = \rho W_1 y + X\beta + \varepsilon,$$

$$\text{Com } \varepsilon = \lambda W_2\, \varepsilon + \mu$$

$$\mu \sim N(0,\sigma^2 I_n) \qquad (3.2.1)$$

No modelo acima, y é o vetor nx1 de variáveis dependentes, X é uma matriz nxk de variáveis explicativas, e ε é o termo de erro aleatório normalmente distribuído. W_1 e W_2 são as matrizes nxn de pesos espaciais. Seguindo a definição de contiguidade binária, uma matriz de contiguidade de primeira ordem possui zeros em sua diagonal principal, suas linhas são preenchidas com 0 (zero) nas posições referentes a unidades regionais não contíguas e com 1 (um) naquelas posições vizinhas à unidade que está sendo estudada.[3] ρ, β e λ são parâmetros. É fácil ver que o modelo pode ser reescrito na forma abaixo:

$$(I - \rho W)\, Y = X\beta + (I - \lambda W)^{-1}\, \mu \qquad (3.2.2)$$

O modelo (3.2.1), ou mesmo, sua versão reduzida, em (3.2.2), indica que a dependência espacial se manifesta tanto nas variáveis controladas pelo modelo quanto nas variáveis não controladas. Uma representação esquemática pode ser ilustrada a partir da Figura 4 abaixo.

3 No tópico A matriz de pesos espaciais (capítulo 2) são fornecidos alguns exemplos de especificação para a matriz de pesos espaciais.

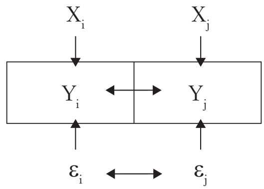

Figura 4. Representação esquemática do modelo SAC.[4]

A Figura 4 ilustra a influência das variáveis explicativas e do termo de erro sobre a variável dependente, sendo que a influência do comportamento dos vizinhos também está presente, tanto na própria variável dependente quanto no termo de erro.

A função logaritmo da verossimilhança (L) para o modelo acima é dada por:

$$
\begin{aligned}
&\mathrm{L} = C-(n/2)\ln(\sigma^2)+\ln(|A|)+\ln(|B|)-(1/2\sigma^2)(e'B'Be)\\
&e=(Ay-X\beta)\\
&A=(I_n-\rho W_1)\\
&B=(I_n-\lambda W_2)
\end{aligned}
\tag{3.2.3}
$$

Para a estimação dos parâmetros do modelo SAC, faz-se necessária a otimização do logaritmo da função de verossimilhança. Dessa forma, os estimadores de máxima verossimilhança para ρ e λ requerem que se encontrem os valores dos parâmetros que maximizam o logaritmo da função dada em (3.2.3). Todavia, no sentido de simplificar o problema de maximização, pode-se obter o logaritmo da função concentrada. É possível concentrar a função usando as seguintes expressões para β e σ^2 (Lesage, 1999):

$$
\begin{aligned}
&\beta=(X'A'AX)^{-1}(X'A'ABy)\\
&e=By-x\beta\\
&\sigma^2=(e'e)/n
\end{aligned}
\tag{3.2.4}
$$

4 Extraído de Almeida (2007).

Dadas as expressões em (3.2.4), é possível calcular o logaritmo da verossimilhança com os valores de ρ e λ. Os valores dos parâmetros β e σ^2 podem ser calculados como uma função de ρ e λ, e com os dados amostrais de y e X.

Do modelo mais geral, SAC, podem-se derivar modelos distintos ao impor-se restrições sobre os parâmetros. Por exemplo, estabelecendo $X = 0$ e $W_2 = 0$ tem-se um modelo espacial autorregressivo na forma:

$$y = \rho W_1 y + \varepsilon,$$

$$\varepsilon \sim N(0,\sigma 2 I_n) \qquad (3.2.5)$$

Aqui, o vetor de variáveis y é expresso em termos de desvio da média no intuito de eliminar o termo de intercepto do modelo. O modelo (3.2.5) busca explicar a variação em y como uma combinação linear das unidades vizinhas, sem qualquer outra variável explicativa.

Todavia, dois casos particulares do modelo geral chamam mais a atenção, a saber: quando $W_1 = 0$ ou quando $W_2 = 0$, cada um com problemas econométricos específicos. Nota-se que, se ambas forem iguais a zero, então, o modelo recai no modelo clássico de regressão linear.

No caso de W_2 ser igual a zero, tem-se o modelo com defasagens espaciais SAR (*mixed regressive-spatial autorregressive model*),[5] dado por:

$$y = \rho W_1 y + X\beta + \varepsilon \qquad (3.2.6)$$

O modelo apresenta uma variável explicativa, $W_1 y$, que é o valor médio da variável dependente nos vizinhos. Nesse caso, cada localidade é vizinha de seus vizinhos, tal que o efeito dos vizinhos precisa ser tratado como endógeno. É fácil perceber a similaridade do modelo SAR com o modelo de variáveis dependentes defasadas das séries de tempo. Neste, o período de tempo mais próximo importa, enquanto, naquele, os lugares mais próximos possuem maior relevância. O pa-

5 Anselin (1988) denominou esse modelo como "modelo misto regressivo-autorregressivo espacial" (*mixed regressive-spatial autorregressive model*) porque combina o modelo de regressão-padrão com uma variável espacialmente defasada.

râmetro ρ do modelo (3.2.6) mede a influência média das observações vizinhas sobre as observações do vetor *y*, o que quer dizer que, para o caso de ρ significativo, uma parcela da variação total de *y* é explicada pela dependência de cada observação de seus vizinhos.

Nota-se que a presença de um termo para a defasagem espacial do lado direito da equação induz a uma correlação dos erros diferente de zero. Além disso, a defasagem espacial para uma dada observação *i* não é apenas correlacionada com o termo de erro em *i*, mas também com os termos de erro em todas as outras localidades. Dado que a simultaneidade incorporada no termo W_1y deve ser explicitamente levada em consideração, a estimativa por MQO será viesada e inconsistente, quando se deve utilizar a função de verossimilhança para estimação. Anselin (1988) fornece um método de Máxima Verossimilhança (MV) para estimar os parâmetros desse modelo.

A figura abaixo ilustra a interação presente no modelo SAR.

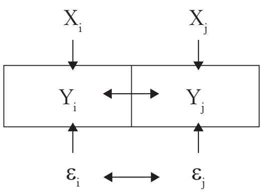

Figura 5. Representação esquemática do modelo SAR.[6]

Como pode ser observado na Figura 5, há uma influência mútua da variável dependente com os seus vizinhos.

Quando uma variável dependente defasada é omitida do modelo de regressão, mas se faz presente no processo gerador dos dados, o problema resultante é similar àquele observado para variáveis omitidas no modelo de regressão linear clássico. Uma alternativa ao método da máxima verossimilhança, nesse caso, seria o uso de variáveis instrumentais, o qual não requer uma suposição de normalidade.

De outra forma, uma maneira de introduzir-se a autocorrelação espacial no modelo de regressão linear é a especificação de uma

6 Extraído de Almeida (2007).

estrutura espacial para o termo de erro. Tal procedimento é necessário quando W_1 é igual a zero. Nesse caso, ocorre um problema de autocorrelação espacial que, do ponto de vista econométrico, tem as mesmas consequências do tradicional problema da autocorrelação temporal: os estimadores de MQO serão ineficientes.

O modelo SEM (*spatial error model*) com autocorrelação espacial no termo de erro é apresentado da seguinte forma:

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

$$\text{com } \varepsilon = \lambda W_2 \varepsilon + \mu$$

$$\mu \sim N(0,\sigma 2 I_n) \tag{3.2.7}$$

y é um vetor nx1 de variáveis dependentes, X representa a usual matriz nxk de variáveis explicativas, e W_2 consiste em uma matriz de pesos espaciais previamente definida. λ e β são parâmetros.

Ao recorrer-se à forma reduzida de (3.2.8), segue-se que:

$$y = X\beta + (I - \lambda W_2)^{-1}\mu \tag{3.2.8}$$

Neste modelo, a covariância dos erros toma a forma:

$$\in[\varepsilon\varepsilon'] = \sigma^2(I - \lambda W_2)^{-1}(I - \lambda W_2')^{-1} =$$

$$= \sigma^2[(I - \lambda W_2)'(I - \lambda W_2]^{-1} \tag{3.2.9}$$

Na estrutura da matriz de variância-covariância de (3.2.9), cada localidade é correlacionada com todas as outras localidades do sistema, mas de forma mais intensa com aquelas mais próximas, seguindo a já mencionada Lei de Tobler. O parâmetro de erro espacial, λ, quando significativo, reflete a autocorrelação espacial nos erros ou nas variáveis que foram omitidas do modelo.

Da mesma forma que o processo gerador do modelo de defasagens espaciais, o modelo autorregressivo de erro conduz a uma covariância dos erros diferente de zero para cada par de observações, mas decrescente à medida que aumenta a ordem da contiguidade. Nesse caso, também se deve recorrer à função de verossimilhança.

Segundo Rey & Montouri (1999), quando $\lambda \neq 0$, um choque ocorrido em uma unidade geográfica não se espalha apenas entre seus vizinhos imediatos, mas sim por todas as outras unidades. Uma alternativa é o uso de estimadores do método dos momentos generalizados (GMM), apresentado por Conley (1999).

Segue o quadro esquemático representativo do modelo SEM.

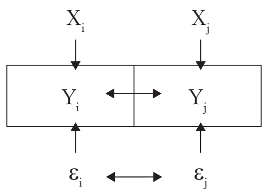

Figura 6. Representação esquemática do modelo SEM.

No caso do modelo SEM, a influência espacial encontra-se nas variáveis omitidas do modelo, como pode ser observado na Figura 6.

Nota-se que, a partir do modelo geral, é possível utilizar a variável W_3X, isto é, a defasagem espacial das variáveis explicativas. Nesse caso, como X é, em princípio, uma matriz de variáveis exógenas, então não há inconveniente sob o ponto de vista econométrico. Esse modelo é conhecido como Modelo Espacial de Durbin (*Spatial Durbin Model* – SDM) (Anselin & Bera, 1998) e assume a forma:

$$y = \rho W_1 y + X\beta - \theta W_3 X\beta + \mu \tag{3.2.10}$$

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

Sendo que ρW_1 e θW_3 representam as respectivas matrizes de pesos espaciais associadas a seus parâmetros. Nota-se que, em (3.2.10), existe defasagem espacial tanto na variável dependente quanto nas variáveis explicativas. O modelo de *Durbin* pode também ser expresso em termos de variáveis espacialmente filtradas, na forma:

$$(I - \rho W_1)y = (I - \theta W_3)X\beta + \mu \tag{3.2.11}$$

Este é um modelo de regressão com variáveis dependentes e explicativas espacialmente filtradas e com um termo de erro não autocorrelacionado.

Da mesma forma como foi representado para os outros modelos, segue a ilustração do processo gerado pelo modelo de *Durbin*.

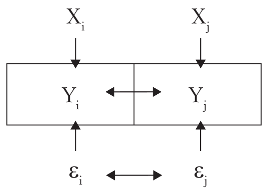

Figura 7. Representação esquemática do modelo SDM.

No último caso, tanto as variáveis explicativas quanto a variável dependente apresentam uma dada estrutura para as unidades espaciais.

Testes de especificação dos modelos espaciais

Como foi visto no tópico anterior, os componentes espaciais do modelo podem aparecer, basicamente, por meio de três formas: (1) na forma de defasagem espacial na variável dependente (Wy), (2) na forma de defasagem nas variáveis explicativas (Wx), ou então (3) como defasagem no termo de erro (Wμ). Tais componentes podem aparecer de forma isolada ou em conjunto. Os testes para modelos espaciais, geralmente, tomam como base a estimação por MV ou por MQO.

Anselin & Bera (1998) enfatizam que, assim como na literatura econométrica clássica, os estágios iniciais da abordagem de econometria espacial foram marcados pela ênfase nas técnicas de estimação. Nesse sentido, Cliff & Ord (1973) desenvolveram a estimação por máxima verossimilhança. Na econometria tradicional, Durbin & Watson (1950, 1951) introduziram a estatística para correlação em modelos de séries de tempo, a qual consistiu no primeiro teste de especificação aplicado a modelos de regressão. No entanto, outros testes, como: homocedasticidade, normalidade, exogeneidade e forma funcional não tiveram a merecida atenção antes dos anos 80. A estatística descoberta por Rao (1947), que ficou conhecida na literatura como teste do Multiplicador de Lagrange (LM), foi uma exceção e tornou-se amplamente utilizada em função de sua facilidade operacional. Outros testes de natureza "assintótica" também foram desenvolvidos, como o teste de Razão de Verossimilhança (LR) e o teste de Wald.

Apesar de o caminho percorrido pela econometria espacial, em termos de testes de especificação, ter sido muito parecido com o caminho da econometria tradicional, a implementação desses testes se mostrou bastante distinta entre os dois campos de pesquisa. Anselin & Bera (1998) chamam atenção para o fato de que os testes para os modelos de econometria espacial não seguem a forma padrão da maioria dos testes da econometria tradicional, na forma "NR^2" – em que N é o tamanho da amostra, e R^2 é o coeficiente de determinação. Além disso, a possibilidade de defasagem espacial tanto na variável dependente quanto no termo de erro tornam os testes dos modelos espaciais mais complexos.

Conforme já mencionado, a estatística I de Moran surgiu como uma analogia bidimensional ao teste de Durbin-Watson para séries de tempo e, desde então, é a técnica mais utilizada para diagnosticar autocorrelação espacial em modelos de regressão.

A estatística de Moran possui como hipótese nula a inexistência de qualquer forma de dependência espacial, mas não apresenta uma correspondência direta com uma hipótese alternativa particular. Assim, apesar de ser um bom identificador de correlação espacial, o teste não é capaz de distinguir qual estrutura de dependência espacial está presente no modelo.

Recentemente, uma variedade de testes alternativos à estatística I tem sido desenvolvida.[7] Assim como o teste de Moran, outros testes também são baseados nos resultados de uma regressão de

7 Para maiores detalhes ver Anselin & Florax (1995).

MQO clássica, ao apresentarem como hipótese nula a ausência de autocorrelação espacial.

Um modelo mais geral de dependência espacial é o modelo SARMA, demonstrado em (3.1.4). Aqui, acrescenta-se ao modelo uma componente com variáveis explicativas exógenas, conforme Anselin & Florax (1995).

$$y = \rho W_1 y + X\beta + \theta_1 W_2 \mu \qquad (3.3.1)$$

Sendo que as notações permanecem as mesmas das equações anteriores. O primeiro termo $\rho W_1 y$ representa a variável dependente espacialmente defasada, com um parâmetro espacial autorregressivo ρ. O segundo termo do lado direito da equação, $X\beta$, representa a matriz de variáveis explicativas exógenas mais o vetor de parâmetros β. O último termo $\theta_1 W_1 \mu$ refere-se à defasagem no termo de erro, mais o parâmetro θ_1.

Do modelo geral, segue-se que os testes baseados nas estimativas de MQO são aplicados somente a um tipo de dependência, sendo assumida, de forma condicional, a ausência do outro tipo. Assim, a hipótese nula para testar a presença de um processo autorregressivo espacial é H_0: $\rho = 0$, condicionado a $\theta_1 = 0$. Anselin & Florax (1995) chamam atenção para quando essas condições não são satisfeitas, ou seja, quando a presença de uma outra forma de dependência espacial está presente no modelo. Nesse caso, os testes não podem mais ser baseados nos resultados da regressão de MQO, e devem ser levados a cabo por meio das estimativas de MV do modelo espacial apropriado; ou ainda, podem-se utilizar testes robustos que considerem a presença da outra forma de dependência espacial.

No caso deste trabalho, além do I de Moran, dois testes familiares de dependência espacial em modelos de regressão linear são investigados, LM-ERR e LM-LAG. Assim como a estatística de Moran, a família de testes LM utiliza apenas os resultados das estimativas por MQO, sob a luz de uma H_0 de nenhuma dependência espacial. A estatística LM-ERR foi sugerida por Burridge (1980) e é, basicamente, um coeficiente de Moran em escala quadrática. A estatística para o teste apresenta-se da seguinte forma:

$$\text{LM-ERR} = \frac{\left(e'W_1 e / s^2\right)^2}{T_1} \tag{3.3.2}$$

Em que $s^2 = e'e/n$ e $T_1 = \text{tr}\left(W'_1 W_1 + W_1^2\right)$, com tr como um operador traço da matriz. A estatística LM-ERR segue uma distribuição χ^2 com 1 (um) grau de liberdade e possui, como hipótese alternativa, a presença de dependência espacial no termo de erro.

O teste LM para a presença de dependência espacial na variável dependente é dado por:

$$\text{LM-LAG} = \left(\frac{e'W_1 y}{s^2}\right)^2 \frac{1}{\left(nJ_{\rho-\beta}\right)} \tag{3.3.3}$$

Com $J_{\rho\text{-}\beta} = [T_1 + (W_1 X\beta)' M (W_1 X\beta) / s^2]$ e $M = I - X(X'X)^{-1}X'$ que é a matriz de projeção usual. A estatística LM-LAG também segue uma distribuição χ^2 com 1 grau de liberdade.

Bera & Yoon (1993) fornecem as versões robustas dos testes LM-ERR e LM-LAG, as quais consideram o efeito da dependência espacial que não é captado pelo teste. O teste LM-EL é o teste LM para dependência espacial no termo de erro, robusto à dependência espacial na variável dependente. O teste é computado da seguinte forma:

$$\text{LM-EL} = \frac{[e'W_1 e / s^2 - T_1 (nJ_{\rho-\beta})^{-1} (e'W_1 y / s^2)]^2}{[T_1 - T_1^2 (nJ_{\rho-\beta})^{-1}]} \tag{3.3.4}$$

E a notação permanece a mesma das anteriores, e a distribuição de LM-EL permanece uma χ^2 com 1 grau de liberdade.

O teste robusto para LM-LAG consiste em um teste para defasagem espacial que considera a influência da dependência espacial no erro. O teste LM-LE é definido formalmente, como segue:

$$\text{LM-LE} = \frac{\left(e'W_1 y / s^2 - e'W_1 e / s^2\right)^2}{nJ_{\rho-\beta} - T_1} \sim \chi^2_{(1)} \tag{3.3.5}$$

Florax et al. (2003) afirmam que os testes robustos do multiplicador de lagrange possuem um poder maior em apontar a alternativa correta para a especificação do modelo, ao invés de serem adotados os testes LM tradicionais. Adicionalmente, os autores fornecem uma estratégia de especificação híbrida que combina a taxonomia clássica de especificação com o emprego dos testes robustos, como segue:

1. estima-se o modelo inicial $y = X\beta + \varepsilon$ através de MQO;
2. testa-se a hipótese de nenhuma dependência espacial em função de uma defasagem espacial omitida, ou em função de erros espacialmente autorregressivos, utilizando LM-LAG e LM-ERR, respectivamente;
3. se ambos os testes forem não significativos, as estimativas iniciais do passo (1) devem ser usadas como a especificação final; caso contrário, procede-se como sugerido em (4);
4. se ambos os testes são significativos, estima-se a especificação apontada por aquele mais significativo dos dois testes robustos. Por exemplo, se LM-LE > LM-EL, então, estima-se (2) usando LM-LAG. Se LM-EL > LM-LE, então, estima-se (2) usando LM-ERR. De outra forma, procede-se como em (5);
5. se LM-LAG é significante, mas LM-ERR não o é, estima-se (2) utilizando LM-LAG. Caso contrário, procede-se como sugerido em (6);
6. estima-se (2) usando LM-ERR.

Lesage (1999) apresenta um outro teste com base no multiplicador de lagrange que possibilita analisar se a presença do termo de defasagem espacial elimina a dependência espacial presente nos resíduos do modelo de MQO. Essa estatística testa a presença da dependência espacial nos resíduos, condicionada à existência de um parâmetro ρ para a defasagem espacial diferente de zero.

O teste é baseado no seguinte modelo (Lesage, 1999):

$$y = \rho Cy + X\beta + \mu$$

$$\mu = \lambda W \mu + \varepsilon$$

$$\varepsilon \sim N(0,\sigma^2 I_n) \qquad (3.3.6)$$

Sendo que o foco do teste é sobre o parâmetro λ.

A estatística para o teste apresenta a seguinte forma:

$$\left(e'We/\sigma^2\right)\left[T_{22}-\left(T_{21}\right)^2\mathrm{var}\left(\rho\right)\right]^{-1}\sim\chi^2_{(1)} \tag{3.3.7}$$

Com $T_{22}=\mathrm{tr}\left(W.*W+W'W\right)$

$$T_{21}=\mathrm{tr}\left(W.*CA^{-1}+W'CA^{-1}\right)$$

Tem-se que W é a matriz de pesos espaciais escolhida, $A=(I_n-\rho C)$, var(ρ) é a estimativa de MV para a variância do parâmetro ρ no modelo e . ∗ simboliza a operação de multiplicação da matriz elemento por elemento.

Por sua vez, o teste assintótico de Wald não é baseado nos resultados de uma regressão de MQO, mas sim no cômputo das estimativas de máxima verossimilhança do modelo espacial apropriado. A estatística de Wald pode ser utilizada tanto para averiguar a presença de dependência espacial na variável dependente quanto no termo de erro. Contudo, é mais comum encontrar o teste aplicado ao modelo de erro espacial, com H_0: λ=0, sendo a hipótese alternativa o modelo SEM. O teste aplicado ao modelo de erro espacial é definido como:

$$W=\lambda^2\left[t_2+t_3-\left(1/n\right)\left(t_1^2\right)\right]\sim\chi^2_{(1)}$$

$$t_1=tr(W.*B^{-1})$$

$$t_2=tr(WB^{-1})^2$$

$$t_3=tr(WB^{-1})'(WB^{-1}) \tag{3.3.8}$$

Em que $B=(I_n-\lambda W)$, com λ sendo a estimativa de MV.

Por fim, o teste de Razão de Verossimilhança (LR) é baseado na diferença entre o logaritmo (log) da verossimilhança do modelo SEM e o log da verossimilhança do modelo de MQO. Dessa forma, Anselin (1988) define o teste como:

$$LR=2[L(\theta)-L(\theta_R)] \tag{3.3.9}$$

Sendo que $L(\theta)$ corresponde ao log da verossimilhança do modelo não restrito – modelo SEM – e $L(\theta_R)$ corresponde ao log da verossimilhança do modelo restrito, ou seja, o modelo de MQO. O teste *LR* é distribuído assintoticamente como uma χ^2 com *q* graus de liberdade.

A matriz de pesos espaciais

Considerando um modelo de regressão linear familiar da forma:

$$y = X\beta + \varepsilon \qquad (3.4.1)$$

A matriz de variância-covariância dos erros, cov[ε ε'], expressa uma covariância espacial quando os elementos fora da diagonal principal são diferentes de zero e seguem uma dada estrutura ou ordenamento espacial (Anselin 2003), que especifica os pares de localidades i-j (com $i \neq j$) cuja covariância será diferente de zero, ou $\in[\varepsilon_i \, \varepsilon_j] \neq 0$.

Há duas maneiras de encontrar o padrão espacial dessa estrutura. A primeira maneira consiste em especificar, diretamente, a covariância como uma função da distância que separa quaisquer dois pares de localidades. Essa abordagem é comumente empregada em geoestatística, onde as superfícies espaciais são contínuas. Segundo Anselin (2003) tal abordagem requer uma função decrescente para a distância e um parâmetro espacial que assegurem uma matriz de variância-covariância definida positiva.

A segunda forma de encontrar o ordenamento espacial – mais adequada para pontos de observação discretos no espaço – requer a especificação de um processo estocástico que relacione o valor de uma variável aleatória em uma localidade aos valores dessa variável em localidades vizinhas. Assim, em vez de ligar todos os pares por meio de uma função de decaimento da distância, os vizinhos de cada localidade são especificados por meio da chamada matriz de pesos espaciais, W (Anselin, 2003). Dessa forma, para cada ponto do espaço, é definido um conjunto de vizinhança relevante que, potencialmente, interage com ele.

A segunda abordagem aproxima-se mais da realidade dos dados econômicos, uma vez que tal perspectiva é uma extensão do caso tradicional para as séries de tempo, no entanto, para um ordenamento em um espaço bidimensional. De fato, a Econometria Espacial propriamente dita está relacionada aos dados em treliças, pontos discretos no espaço. Uma "treliça" de locações vem da ideia de pontos espaçados (regiões) ligados a seus "vizinhos", como o exemplo da Figura abaixo.

1	2	3
4	5	6

Figura 8. Dados em treliças.

O ordenamento das informações ao longo do espaço pode ser feito de diversas maneiras. Uma delas é o critério de contiguidade, que reflete a posição de uma unidade em relação às demais unidades no espaço. Medidas de contiguidade necessitam de informações a respeito do tamanho e forma das unidades regionais. Quanto à dependência espacial, pressupõe-se que regiões vizinhas, contíguas, apresentem um grau maior de dependência do que as demais.

Por exemplo, seguindo o critério "Rainha" (*Queen*) de contiguidade,[8] na Figura 8, a região 1 é vizinha das regiões 2, 4 e 5, enquanto a região 5 é vizinha de todas as demais. O critério rainha estabelece que essas relações de vizinhança podem ser representadas pela matriz de conectividade **W** abaixo:

$$\mathbf{W} = \begin{pmatrix} 0 & 1 & 0 & 1 & 1 & 0 \\ 1 & 0 & 1 & 1 & 1 & 1 \\ 0 & 1 & 0 & 0 & 1 & 1 \\ 1 & 1 & 0 & 0 & 1 & 0 \\ 1 & 1 & 1 & 1 & 0 & 1 \\ 0 & 1 & 1 & 0 & 1 & 0 \end{pmatrix}$$

Obviamente, a matriz W é simétrica, e evidencia que, assim como a região 1 é vizinha da região 2, a região 2 é vizinha da região

8 O critério Rainha considera como vizinhas as unidades que possuem fronteiras ou vértices comuns, em que a unidade vizinha é definida da forma $w_{ij} = 1$, enquanto o elemento, que não possui relação de vizinhança, é definido $w_{ij} = 0$.

1. Além disso, por convenção, a matriz sempre tem zeros em sua diagonal principal.

Pode-se também construir a matriz **w**, que seria a matriz **W** normalizada pelas linhas, isto é, alterada de tal modo que a soma em cada linha seja exatamente igual a 1. Isso é feito simplesmente dividindo o valor de cada elemento da matriz pelo total das linhas.[9] Dessa forma, a soma das influências dos vizinhos é igual para cada unidade em consideração, o que torna possível a comparabilidade. Além disso, com a normalização da matriz **W**, a amplitude de possibilidades dos parâmetros é restrita ao intervalo de -1 a 1.

$$\mathbf{w} = \begin{pmatrix} 0 & 1/3 & 0 & 1/3 & 1/3 & 0 \\ 1/5 & 0 & 1/5 & 1/5 & 1/5 & 1/5 \\ 0 & 1/3 & 0 & 0 & 1/3 & 1/3 \\ 1/3 & 1/3 & 0 & 0 & 1/3 & 0 \\ 1/5 & 1/5 & 1/5 & 1/5 & 0 & 1/5 \\ 0 & 1/3 & 1/3 & 0 & 1/3 & 0 \end{pmatrix}$$

A motivação para a normalização da matriz W foi ilustrada por Lesage (1999), da seguinte forma:[10] em primeiro lugar, considera-se a matriz de multiplicação **w** e um vetor de observações de alguma variável associada com as seis regiões a que se chama de y.

$$\begin{pmatrix} y_1^* \\ y_2^* \\ y_3^* \\ y_4^* \\ y_5^* \\ y_6^* \end{pmatrix} = \begin{pmatrix} 0 & 1/3 & 0 & 1/3 & 1/3 & 0 \\ 1/5 & 0 & 1/5 & 1/5 & 1/5 & 1/5 \\ 0 & 1/3 & 0 & 0 & 1/3 & 1/3 \\ 1/3 & 1/3 & 0 & 0 & 1/3 & 0 \\ 1/5 & 1/5 & 1/5 & 1/5 & 0 & 1/5 \\ 0 & 1/3 & 1/3 & 0 & 1/3 & 0 \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \\ y_5 \\ y_6 \end{pmatrix}$$

A matriz produto $y^* = \mathbf{w}y$ representa uma nova variável igual à média das observações das regiões contíguas.

9 Cf. Anselin (1988).

10 Para o caso deste trabalho, adaptou-se a representação esquemática de Lesage ao mapa fornecido pela Figura 7.

$$\begin{pmatrix} y_1^* \\ y_2^* \\ y_3^* \\ y_4^* \\ y_5^* \\ y_6^* \end{pmatrix} = \begin{pmatrix} 1/3y_2 + 1/3y_4 + 1/3y_5 \\ 1/5y_1 + 1/5y_3 + 1/5y_4 + 1/5y_5 + 1/5y_6 \\ 1/3y_2 + 1/3y_5 + 1/3y_6 \\ 1/3y_1 + 1/3y_2 + 1/3y_5 \\ 1/5y_1 + 1/5y_2 + 1/5y_3 + 1/5y_4 + 1/5y_6 \\ 1/3y_2 + 1/3y_3 + 1/3y_5 \end{pmatrix}$$

Uma das maneiras de quantificar a relação $y_i = f(y_j)$, $j \neq i$, é por meio da matriz de conectividade binária. Considera-se, em ambos os casos, vizinhança de ordem 1. A região 1 não é vizinha (de ordem 1) da região 3, mas elas são vizinhas de segunda ordem, pois a região 1 é vizinha da região 2 que, por sua vez, é vizinha da região 3. Para relações de ordem 2 ou superiores, são necessárias, portanto, diferentes matrizes de conectividade. A matriz **W** de ordem 0 ($\mathbf{W_0}$) é a própria matriz identidade.

Existem outras formas de montar a matriz W. Basta que a forma escolhida considere algum tipo de medida que estabeleça a participação dos vizinhos. Nesse sentido, busca-se aqui testar uma variedade de matrizes de pesos a fim de se identificar aquela que mais se aproxima da verdadeira correlação espacial apresentada pelos dados. Para tal, algumas matrizes W serão testadas, entre elas, a matriz de contiguidade binária.

Como já foi mencionado, a matriz W tem o intuito de captar a estrutura de correlação espacial apresentada pelos dados. Assume-se, dessa forma, uma estrutura específica para o erro, sendo que a literatura de econometria espacial admite uma certa arbitrariedade na seleção da matriz de pesos e permite escolhas *ad hoc* por parte do pesquisador. Conforme Anselin (1988), a escolha apropriada da matriz de pesos espaciais é uma das questões metodológicas atuais mais controversas em econometria espacial.

Dado que as hipóteses sobre a matriz W são feitas *a priori*, a eliminação dos resíduos espacialmente autocorrelacionados não é uma condição suficiente para a eliminação do viés de estimação e pode diferir da verdadeira função. Daí a importância da escolha adequada da matriz de pesos. Florax & Rey (1995) exploram e enfatizam as implicações da má especificação da matriz W.

Lesage (1999) sugere que o princípio direcionador da escolha da matriz mais adequada deve ser a natureza do problema a ser modelado, sendo relevante, também, o emprego de informações adicionais não fornecidas pela amostra. Abreu et al. (2005) atentam para a importância de fundamentar-se a escolha da matriz de pesos, segundo um conjunto de hipóteses teóricas feitas *a priori*.

As matrizes de pesos espaciais mais tradicionais são construídas a partir de atributos físicos e geográficos, como vizinhança, distâncias geográficas e tempo de deslocamento.

De acordo com a distância geográfica, a construção da matriz W está baseada no ordenamento de um espaço cartesiano representado por latitudes e longitudes. Esse tipo de ordenamento permite calcular as distâncias de quaisquer pontos no espaço. Com relação à dependência espacial, pressupõe-se que o grau de dependência é negativamente relacionado com a distância. Em outras palavras, assume-se que a intensidade da dependência espacial declina à medida que a distância entre as unidades aumenta.

Uma matriz de pesos baseada em distâncias geográficas pode ser calculada por meio do inverso da distância euclidiana, na forma:

$$w^*_{ij} = \frac{1}{\sqrt{(x_i - x_j)^2 + \left(y_i - y_j\right)^2}}, \text{ se } i = j$$

$$w^*_{ij} = 0, \text{ se } i = j \qquad (3.4.2)$$

x_i, x_j, y_i *e* y_j são as coordenadas dos centroides das unidades i e j.

Nessa especificação, a construção da matriz de pesos é feita com base em um grande círculo entre as regiões, centroides. Entretanto, o emprego do inverso da distância euclidiana ao quadrado é mais usual, uma vez que maiores distâncias são penalizadas mais rapidamente, ou seja, atribui-se maior peso aos vizinhos mais próximos. Tem-se, portanto:

$$w^*_{ij} = \frac{1}{\left(x_i - x_j\right)^2 + \left(y_i - y_j\right)^2}, \text{ se } i \neq j$$

$$w^*_{ij} = 0, \text{ se } i = j \tag{3.4.3}$$

A notação permanece a mesma dos conjuntos de equações anteriores.

Uma especificação adicional consiste em construir a matriz de pesos espaciais por meio de uma distância limite, mas com um número fixo, k, de vizinhos mais próximos. A matriz W(k) usual é definida, como se segue:

$$w^*_{ij}(k) = 0 \text{ se } i = j, \forall k$$

$$w^*_{ij}(k) = 1 \text{ se } d_{ij} \leq d_i(k)$$

$$w^*_{ij}(k) = 0 \text{ se } d_{ij} > d_i(k) \tag{3.4.4}$$

Em que $d_i(k)$ é a distância do vizinho de ordem k. Segundo Ertur & Gallo (2003), existem diversas vantagens para a preferência dessa matriz em contrapartida à matriz de contiguidade simples. Em primeiro lugar, ela tem vantagem quando ilhas importantes fazem parte da amostra de dados. Um exemplo é o caso da Grã-Bretanha, que seria relegada, no caso de uma análise espacial dos países europeus, porque não possui vizinhos. Em segundo lugar, de acordo com os referidos autores, ao escolher um número fixo de vizinhos, evita-se uma série de problemas metodológicos que surgem quando se permite a variação nesse número.

Tysler (2006) chama atenção para uma matriz de pesos que utiliza um número fixo de vizinhos e que não é construída de forma binária, mas sim através da distância entre os centroides. Essa matriz W é construída da seguinte forma:

$$w^*_{ij}(k) = 0, \text{ se } i = j$$

$$w^*_{ij}(k) = \frac{1}{d_{ij}^{\ 2}}, \text{ se } d_{ij} \leq d_i(k)$$

$$w^*_{ij}(k) = 0, \text{ se } d_{ij} > d_i(k) \tag{3.4.5}$$

Nesse caso, $d_i(k)$ é a distância de corte, isto é, a distância do vizinho de ordem k.

No caso deste trabalho, maior atenção será dada a algumas matrizes em específico, a saber, a matriz de pesos binária, que leva em consideração as fronteiras e vértices comuns, e a matriz formada pelo inverso da distância geográfica com um número fixo de vizinhos. A última foi escolhida, conforme a recomendação de Ertur & Gallo (2003), que, como foi visto, apontam inúmeras vantagens em se utilizar uma matriz que fixa o número de vizinhos.

Case & Rosen (1993) e Conley & Ligon (2002), entre outros, sugerem o uso de pesos baseados na "distância econômica" entre as regiões. Especificamente, Case & Rosen (1993) sugerem usar pesos (antes da padronização) na forma $w_{ij} = \frac{1}{|x_i - x_j|}$, em que x_i e x_j são observações socioeconômicas características da unidade, tais como renda *per capita* ou percentual da população em determinado grupo ético ou racial.

Esses autores utilizam o conceito de similaridade para pressupor uma conexão maior entre as unidades espaciais, ao invés de unidades próximas. Nesse sentido, um município que lidera, economicamente, uma determinada região pode sofrer mais influência de um município líder da região vizinha do que de municípios mais próximos, mas que não possuem economia similar sua.

Na prática, busca-se captar as diferenças existentes entre os municípios para um mesmo indicador socioeconômico, no qual a distância euclidiana invertida é a mais comumente usada.

Conley & Ligon (2002) utilizam uma matriz de distância econômica que discrimina os custos de transporte do capital físico daqueles observados para o capital humano. Para tal, os autores fazem uso dos custos de transporte de encomendas (*United Parcel Service* – UPS) entre capitais de países selecionados, no sentido de medir os custos de transporte do capital físico e os preços de passagens aéreas, buscando capturar os custos de transportar-se capital humano. Os autores procuram, dessa forma, montar a matriz W. Ao levar em conta tal especificação, a razão da distância econômica entre Brasil e Áustria, por exemplo, e a distância entre Brasil e Austrália é menor do que a mesma razão quando se considera a distância geográfica.

Contudo, os autores admitem a correlação existente entre os custos de transporte propriamente ditos e a distância geográfica, uma vez que, para eles, se o principal impedimento aos *spillovers* consiste no custo de transportar fatores, então parece evidente que a distância geográfica será correlacionada com esses custos. (Conley & Ligon, 2002, p.168)