Spatial Econometrics

Urban Economics

Ignacio Sarmiento-Barbieri

Universidad de los Andes

February 11, 2022

Agenda

Intro to Spatial Econometrics

- Applied work in urban economics and regional science relies heavily on sample data that is collected with reference to locations
- What distinguishes spatial econometrics from traditional econometrics?
  - Spatial dependence between the observations and
  - Spatial heterogeneity in the relationships we are modeling.

Spatial Dependence

Cross-sectional iid non-spatial data

- Standard Cross-sectional models

$$y_i = X_i\beta + \epsilon_i \tag{1}$$

$$i = 1, \ldots, n \tag{2}$$

- Independent or statistically independent observations imply

$$E(\epsilon_i \epsilon_j) = E(\epsilon_i)E(\epsilon_j) = 0 \tag{3}$$

Spatial Dependence

Spatial data

- Spatial dependence reflects a situation where values observed at one location or region, say observation i, depend on the values of neighboring observations at nearby locations.

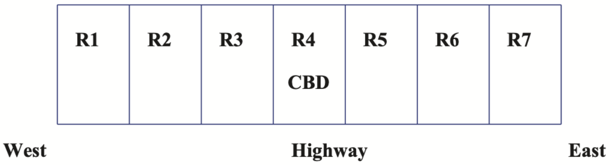
$$y_i = \alpha_i y_j + X_i\beta + \epsilon_i \tag{4}$$

$$y_j = \alpha_j y_i + X_j\beta + \epsilon_j \tag{5}$$

- This situation suggests a simultaneous data generating process, where the value taken by y_i depends on that of y_j and vice versa.

Spatial Dependence
Example

Figure 1: Regions east and west of the CBD



Fuente: LeSage & Pace (2009)

Spatial heterogeneity

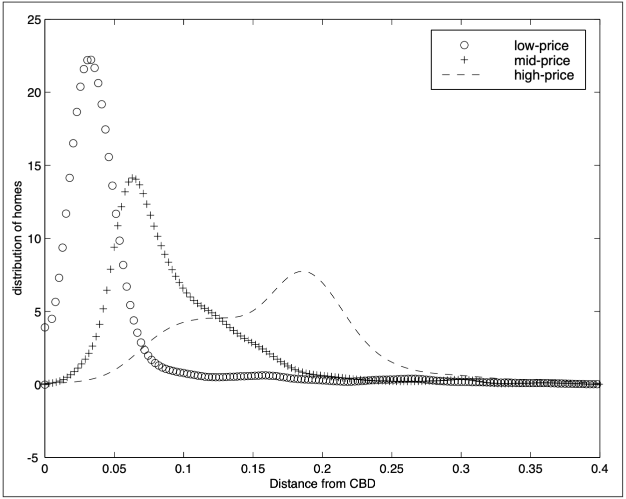
- The term spatial heterogeneity refers to variation in relationships over space.
- In the most general case we might expect a different relationship to hold for every point in space.

$$y_i = X_i\beta_i + \epsilon_i \tag{6}$$

$$i = 1, \ldots, n \tag{7}$$

Spatial heterogeneity

Figure 2: Distribution of low, medium and high priced homes versus distance



The spatial autoregressive process

- The solution to the over-parameterization problem that arises when we allow each dependence relation to have relation-specific parameters is to impose structure on the spatial dependence relations.
- Ord (1975) proposed a parsimonious parameterization for the dependence relations (which built on early work by Whittle (1954)).
- The Spatial autoregressive process.

$$y_i = \lambda \sum_{j=1}^{n} W_{ij} y_j + \epsilon_i \tag{8}$$

$$i = 1, \ldots, n \tag{9}$$

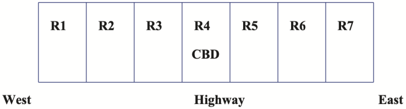
Weights Matrix

- *Weights matrix*:

$$W = \begin{pmatrix} w_{11} & \dots & \dots & w_{n1} \\ \vdots & w_{ij} & & \vdots \\ \vdots & & \ddots & \vdots \\ w_{n_1} & \dots & \dots & w_{nn} \end{pmatrix}_{n \times n} \tag{10}$$

- with generic element: $w_{ij} = \begin{cases} 1 & if\; j \in N(i) \\ 0 & o.w \end{cases}$
- $N(i)$ being the set of neighbors of location j. By convention, the diagonal elements are set to zero, i.e. $w_{ii} = 0$.

Weights Matrix



Fuente: LeSage & Pace (2009)

$$W = \begin{array}{c} \\ R1 \\ R2 \\ R3 \\ R4 \\ R5 \\ R6 \\ R7 \end{array} \begin{array}{c} \begin{array}{ccccccc} R1 & R2 & R3 & R4 & R5 & R6 & R7 \end{array} \\ \begin{pmatrix} 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 \end{pmatrix} \end{array} \tag{11}$$

Weights Matrix

- Quite often the W matrices are standardized to sum to one in each row $w^*_{ij} = \frac{w_{ij}}{\sum_{j=1}^{n} w_{ij}}$

$$
W = \begin{array}{c} \\ R1 \\ R2 \\ R3 \\ R4 \\ R5 \\ R6 \\ R7 \end{array}
\begin{array}{ccccccc} R1 & R2 & R3 & R4 & R5 & R6 & R7 \end{array}
\left(\begin{array}{ccccccc}
0 & 1 & 0 & 0 & 0 & 0 & 0 \\
1/2 & 0 & 1/2 & 0 & 0 & 0 & 0 \\
0 & 1/2 & 0 & 1/2 & 0 & 0 & 0 \\
0 & 0 & 1/2 & 0 & 1/2 & 0 & 0 \\
0 & 0 & 0 & 1/2 & 0 & 1/2 & 0 \\
0 & 0 & 0 & 0 & 1/2 & 0 & 1/2 \\
0 & 0 & 0 & 0 & 0 & 1 & 0
\end{array} \right) \tag{12}
$$

Spatial Lag Model

Spatial Autoregressive (SAR) Models

Let's consider the following model:

$$y = \lambda W y + X\beta + u$$

we assume that W is exogenous

If W is row standardized:

- Guarantees $|\lambda| < 1$ (Anselin, 1982)
- $[0, 1]$ Weights
- Wy Average of neighboring values
- W is no longer symmetric $\sum_j w_{ij} \neq \sum_i w_{ji}$ (complicates computation)

Note that we can write

$$(I - \lambda W)y = X\beta + u$$

Spatial Lag Model

In this case, endogeneity emerges because the spatially lagged value of y is correlated with the stochastic disturbance.

$$E((Wy)u') \neq 0 \tag{13}$$

- One solution is to add an extra assumption

$$u \sim_{iid} N(0, \sigma^2 I) \tag{14}$$

- if

$$y = (I - \lambda W)^{-1} X\beta + (I - \lambda W)^{-1} u$$

- Then

$$y \sim_{iid} N(E(y), V(y)) \tag{15}$$

Spatial Lag Model
Maximum Likelihood Estimator

- The associated likelihood function is then

$$\mathcal{L}\left(\sigma^2, \lambda, y\right) = \left(\frac{1}{\sqrt{2\pi}}\right)^n |\sigma^2\Omega|^{-\frac{1}{2}} exp\left\{-\frac{1}{2\sigma^2}(y-(I-\lambda W)^{-1}X\beta)'\Omega^{-1}(y-(I-\lambda W)^{-1}X$$

- the log likelihood

$$l\left(\sigma^2, \lambda, y\right) = constant - \frac{1}{2}ln|\sigma^2\Omega| - \frac{1}{2\sigma^2}(y-(I-\lambda W)^{-1}X\beta)'\Omega^{-1}(y-(I-\lambda W)^{-1}X\beta)$$

Spatial Lag Model

Maximum Likelihood Estimator

- The determinant $|(I - \lambda W)|$ is quite complicated because in contrast to the time series, where it is a triangular matrix, here it is a full matrix.
- However, Ord (1975) showed that it can be expressed as a function of the eigenvalues ω_i

$$|(I - \lambda W)| = \Pi_{i=1}^{n}(1 - \lambda\omega_i)$$

So the log likelihood is simplified to

$$\begin{aligned} l\left(\sigma^2, \lambda, y\right) = {} & constant - \frac{n}{2} ln\left(\sigma^2\right) \\ & - \frac{1}{2\sigma^2}((I - \lambda W)y - X\beta)'((I - \lambda W) - X\beta) \\ & + \sum ln(1 - \lambda\omega_i) \end{aligned} \tag{16}$$

Spatial Lag Model

Maximum Likelihood Estimator

Applying FOC, the ML estimates for β and σ^2 are:

$$\hat{\beta}_{MLE} = (X'X)^{-1}X'(I - \lambda W)y$$

$$\hat{\sigma}^2_{MLE} = \frac{1}{n}(y - \lambda Xy - X\hat{\beta}_{MLE})'(y - \lambda Xy - X\hat{\beta}_{MLE})$$

- Conditional on λ these estimates are simply OLS applied to the spatially filtered dependent variable and explanatory variables X.

Spatial Lag Model

Maximum Likelihood Estimator

- Substituting these in the log likelihood we have a concentrated log-likelihood as a nonlinear function of a single parameter λ

$$l(\lambda) = -\frac{n}{2} ln \left(\frac{1}{n} (e_0 - \lambda e_L)'(e_0 - \lambda e_L) \right) + \sum ln(1 - \lambda \omega_i) \tag{17}$$

- where e_0 are the residuals in a regression of y on X and
- e_L of a regression of Wy on X.
- This expression can be maximized numerically to obtain the estimators for the unknown parameters λ.
- with λ^*, get $\hat{\beta}_{MLE}$ and $\hat{\sigma}^2_{MLE}$

Spatial Lag Model
Maximum Likelihood Estimator

The asymptotic variance follows as the inverse of the information matrix

$$AsyVar\left(\lambda,\beta,\sigma^2\right) = \begin{pmatrix} tr(W_A)^2 + tr(W_A'W_A) + \frac{(W_AX\beta)'(W_AX\beta)}{\sigma^2} & \frac{(X'W_AX\beta)'}{\sigma^2} & \frac{tr(W_A)'}{\sigma^2} \\ \frac{(X'W_AX\beta)'}{\sigma^2} & \frac{(X'X)}{\sigma^2} & 0 \\ \frac{tr(W_A)'}{\sigma^2} & 0 & \frac{n}{2\sigma^4} \end{pmatrix}^{-1} \quad (18)$$

- where $W_A = W(I - \lambda W)^{-1}$.
- Note that
  - the covariance between β and σ^2 is zero, as in the standard regression model,
  - this is not the case for λ and σ^2.

Spatial Lag Model

Two-Stage Least Squares estimators

- An alternative to MLE we can use 2SLS to eliminate endogeneity.
- Key is to identify proper instruments
  - Need to be uncorrelated with the error term
  - Correlated with Wy

Spatial Lag Model

Two-Stage Least Squares estimators

Consider the following

$$E(y) = (I - \lambda W)^{-1} X\beta$$

now, since $|\lambda| < 1$ we can use Neumann series property to expand the inverse matrix as

$$(I - \lambda W)^{-1} = I + \lambda W + \lambda^2 W^2 + \lambda^3 W^3 + ...$$

hence

$$E(y) = (I + \lambda W + \lambda^2 W^2 + \lambda^3 W^3 + ...) X\beta$$

$$= X\beta + \lambda W X\beta + \lambda^2 W^2 X\beta + \lambda^3 W^3 X\beta + ...$$

so we can express $E(y)$ as a function of X, WX, $W^2 X$,...

Spatial Lag Model

Two-Stage Least Squares estimators

We can use the first three elements of the expansion as instruments. Let's define H as the matrix with our instruments

$$H = [X, WX, W^2X]$$

Now,

$$y = \lambda Wy + X\beta + u$$

$$= M\theta + u$$

where $M = [Wy, X]$ and $\theta = [\lambda, \beta]$.

Spatial Lag Model

Two-Stage Least Squares estimators

- The first stage is: $M = H\gamma + \eta$
  - where $\hat{\gamma} = (H'H)^{-1}H'M$
  - and $\hat{M} = H\hat{\gamma} = P_H M$
- The second stage is

$$y = \hat{M}\theta + u \tag{19}$$

and

$$\begin{aligned} \hat{\theta}_{2SLS} &= (\hat{M}'\hat{M})^{-1}\hat{M}'y \\ &= (M'P_H M)^{-1}M'P_H y \end{aligned} \tag{20}$$

Further Readings

- Arbia, G. (2014). A primer for spatial econometrics with applications in R. Palgrave Macmillan. (Chapter 2 and 3)
- Anselin, Luc, & Anil K Bera. 1998. "Spatial Dependence in Linear Regression Models with an Introduction to Spatial Econometrics." Statistics Textbooks and Monographs 155. MARCEL DEKKER AG: 237–90.
- Anselin, L. (1982). A note on small sample properties of estimators in a first-order spatial autoregressive model. Environment and Planning A, 14(8), 1023-1030.
- LeSage, J. (1999). The Theory and Practice of Spatial Econometrics. Mimeo
- LeSage, J. & Pace, R. K. (2009). Introduction to spatial econometrics. Chapman and Hall/CRC.
- Tobler, WR. 1979. "Cellular Geography." In Philosophy in Geography, 379–86. Springer.