Spatial Linear Regression Models

Ignacio Sarmiento-Barbieri

Universidad de los Andes

June 30, 2022

Recap

- Closeness
- Weights matrix
- Examples of weight matrices weights matrices in R
- Example of spatial regression.

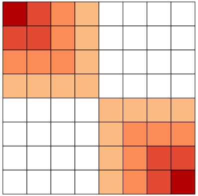
Agenda

Motivation

"Everything is related to everything else, but close things are more related than things that are far apart" (Tobler, 1979).

$$y = X\beta + u$$

- Independence assumption between observation is no longer valid.
- Attributes of observation i may influence the attributes of observation j.
- Spatial dependence introduces a misspecification problem.



Motivation

"Everything is related to everything else, but close things are more related than things that are far apart" (Tobler, 1979).

- One of the major differences between standard econometrics and standard spatial econometrics lies, in the fact that, in order to treat spatial data, we need to use two different sets of information.

  1. Observed values of the variables.
  2. Particular location where those variables are observed and to the various links of proximity between all spatial observations.

Spatial Econometrics: Weights Matrix

- At the heart of traditional spatial econometrics is the definition of the *weights matrix*:

$$W = \begin{pmatrix} w_{11} & \dots & \dots & w_{n1} \\ \vdots & w_{ij} & & \vdots \\ \vdots & & \ddots & \vdots \\ w_{n_1} & \dots & \dots & w_{nn} \end{pmatrix}_{n \times n} \tag{1}$$

- 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. $w_{ii} = 0$.
- Quite often the W matrices are standardized to sum to one in each row $w^*_{ij} = \frac{w_{ij}}{\sum_{j=1}^{n} w_{ij}}$

SARAR(1,1)

Spatial auto-regressive with additional auto-regressive error structure

Let's consider the following model:

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

with

$$u = \lambda W u + \epsilon$$

we assume that W is exogenous

If W is row standardized:

- Guarantees $|\lambda| < 1$ $|\rho| < 1$ (Anselin, 1982)

Interpretation of Parameters

- Consider the following model for the $i-th$ observation

$$y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \cdots + \beta_r x_{ir} + \cdots + \beta_k x_{ik} \;\; i = 1, \ldots, n$$

Interpretation of Parameters

- Lets consider the case of a simple Spatial Lag model with a single regressor

$$y_i = \alpha + \beta x_i + \lambda \sum w_{ij} y_j + \epsilon_i \tag{2}$$

with $|\lambda| < 1$, and

$$\beta \neq \frac{\partial y_i}{\partial x_i}$$

Interpretation of Parameters

More generally consider

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

Then

$$
E(y) = (I - \lambda W)^{-1} X\beta \tag{3}
$$

Interpretation of Parameters

On this basis, LeSage and Pace (2009) suggested the following impact measures that can be calculated for each independent variable X_i included in the model

- *Average Direct Impact*: this measure refers to the impact of changes in the $i-th$ observation of x, which we denote x_i, on y_i. This is the average of all diagonal entries in S

$$
\begin{aligned}
ADI &= \frac{tr(S(W))}{n} \\
&= \frac{1}{n}\sum_{i=1}^{n} S(W)_{ii}
\end{aligned} \tag{4}
$$

Interpretation of Parameters

- *Average Total Impact To* an observation: this measure is related to the impact produced on one single observation y_i. For each observation this is calculated as the sum of the $i-th$ row of matrix S

$$
\begin{aligned}
ATIT_j &= \frac{\iota' S(W)}{n} \\
&= \frac{1}{n}\sum_{i=1}^{n} S(W)_{ij}
\end{aligned}
\tag{5}
$$

Interpretation of Parameters

- *Average Total Impact From* an observation: this measure is related to the total impact on all other observations y_i . For each observation this is calculated as the sum of the $j-th$ column of matrix S

$$\begin{aligned} ATIF_i &= \frac{1}{n} S(W) \iota \\ &= \frac{\sum_{j=1}^{n} S(W)_{ij}}{n} \end{aligned} \tag{6}$$

Interpretation of Parameters

- A Global measure of the average impact obtained from the two previous measures.
- It is simply the average of all entries of matrix S

$$ATI = \frac{1}{n}\iota' S(W)\iota = \frac{1}{n}\sum_{i=1}^{n} ATIT_i = \frac{1}{n}\sum_{j=1}^{n} ATIF_i \tag{7}$$

- The numerical values of the summary measures for the two forms of average total impacts are equal.
- The ATIF relates how changes in a single observation j influences all observations.
- In contrast, the ATIT considers how changes in all observations influence a single observation i.

Interpretation of Parameters

- *Average Indirect Impact* obtained as the difference between ATI and ADI

$$AII = ATI - ADI \tag{8}$$

- It is simply the average of all off-diagonal entries of matrix S

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.
- Tobler, WR. 1979. "Cellular Geography." In Philosophy in Geography, 379–86. Springer.