GIS Resource Document 2-06 (GIS_RD_02-06) November, 2002

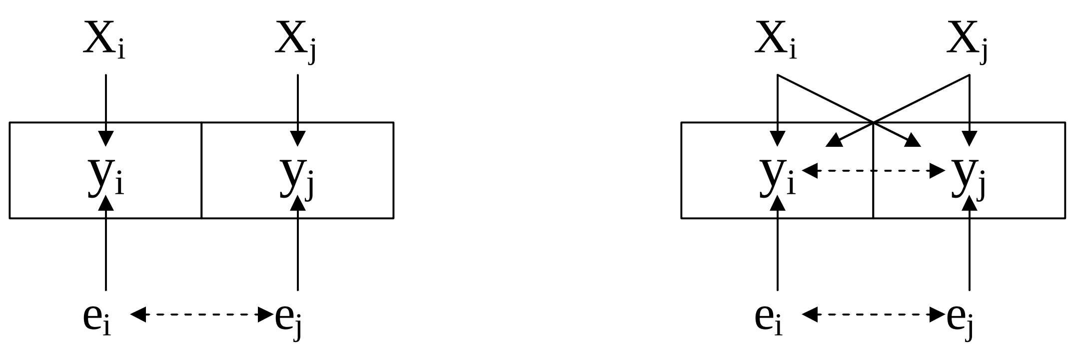
Spatial Regression Analysis

Jason M. Smith

Introduction

Spatial autocorrelation is the idea that a value observed at a location depends on the values observed at neighboring locations. In simpler terms, the value for a variable at a given location depends on the values of that variable at "neighboring" locations. There are two primary types of spatial autocorrelation: spatial error (correlation across space in the error term), and spatial lag (the dependent variable in space "i" is affected by the independent variables in space "i" as well as those in space "j".) The following figures may help:

Spatial Error | Spatial Lag



(from Baller, et al., 2001)

With spatial error in OLS Regression, we violate the assumption that our error terms are uncorrelated, and we get inefficient coefficients. With spatial lag, we violate the assumption that our observations are independent (as well as uncorrelated error terms), thus we get both biased and inefficient coefficients. Therefore we must accommodate these conditions in our models.

SpaceStat, created by Luc Anselin (now at U. of Illinois), provides the user with the widest range of functionality for performing specification tests and estimation of OLS and spatial regression models. (S+ also provides the ability to run some spatial regression models.) In order to use SpaceStat, the easiest path is to have a dataset for use in ArcView 3.X GIS software, and export the dataset into a SpaceStat-ready format. This is done through the SpaceStat extension in ArcView. The user also needs to create a spatial weights matrix – also done through the SpaceStat extension in ArcView – that defines a given location's neighbors. There are two possibilities for defining the weights matrix from ArcView: rook-based and queen-based. A rook-based weights matrix will define as neighbors any locations that share a common border, while a queen-based matrix defines neighbors as locations that share either a border or a vertex in their boundaries. Other definitions of neighbors are possible, and can be created through the Tools Module in SpaceStat. These include distance-based neighbors (i.e., any location within so many miles, kilometers, feet, etc.) and k-nearest neighbors.

The process for estimating spatial regression models begins with estimating an OLS model in SpaceStat, including a spatial weights matrix. Included in the output are indicators of spatial lag and spatial error. Inspection of these "tests for spatial dependence" then leads the user to choose either a spatial lag model or spatial error model. These different specifications can then be estimated using a variety of techniques, including Maximum Likelihood, 2-stage Least Squares, Instrumental Variables, and General Method of Moments Estimation. Output from these procedures include coefficients, standard errors, and significance tests for the variables in the model, which now include parameters for either a spatial lag operator or spatial error operator (depending on the type of model estimated.) Also in the output are tests for the un-modeled type of spatial dependence. That is, in a spatial lag model, the user gets a test for remaining spatial error in the data, and vice-versa. Usually, when the correct type of model is run (lag vs. error), the other type of spatial dependence no longer exists in the data, and substantive interpretation of the model can be made. Statistics on model fit, heteroskedasticity, and multicollinearity are also provided by SpaceStat.

References

Extensive resources are available either through the GIA Core (8th floor of Oswald Tower, www.pop.psu.edu/gia-core) or on the web. Other excellent places to start are the following websites:

www.spacestat.com The home on the web for all things SpaceStat. AND

www.csiss.org Center for Spatially Integrated Social Science

Between these two you can find SpaceStat manuals, tutorials, exercises, and data, all available for downloading at no cost. Workshops, seminars, and classes are also advertised here.

Readings *(all of these are available through the GIA Core, and online search engines)*

Anselin, L. (forthcoming). Under the hood. Issues in the specification and interpretation of spatial regression models. *Agricultural Economics.* This article discusses both theory-driven and data-driven motivations for using spatial regression, and the assumptions and inferential framework that comes with using it.

Anselin, L. 2000. Computing environments for spatial data analysis. Journal of Geographical Systems, 2: 201-220. This article provides an overview of the functionality of SpaceStat, the Spacestat Extension for ArcView, and the DynESDA Extension for ArcView.

Anselin, L., & A. Bera. 1998. Spatial dependence in linear regression models with an introduction to spatial econometrics. In A. Ullah and D. Giles (eds.), *Handbook of Applied Economic Statistics*, pp. 237-289. Marcel Dekker, New York. This chapter goes through the technical derivations and definitions of spatial regression. It includes discussions of Maximum Likelihood, Generalized Methods of Moments, & Instrumental Variable estimation techniques, specification tests, and other tests for spatial dependence available in SpaceStat.

Baller, R., L. Anselin, S. Messner, G. Deane and D. Hawkins. 2001. Structural covariates of US county homicide rates: incorporating spatial effects. *Criminology* 39, 561-590. *This article includes a discussion of spatial regression and incorporating spatial approaches to criminological research.*

Goodchild, M., L. Anselin, R. Appelbaum, and B. Harthorn. 2000. Toward Spatially Integrated Social Science. *International Regional Science Review* 23, 139-159. This article provides philosophical and practical motivations and justifications for employing a spatial approach across social science disciplines, both to further the individual disciplines, as well as to unite them and further science in general.

Special Issue on Spatial Analysis. Journal of Geographical Systems. 2000. No. 2, pp. 1-110. Includes several short (6-10 page) articles on various aspects of spatial analysis.