Spatial Regression

1. Introduction and Review

Luc Anselin

http://spatial.uchicago.edu
matrix algebra basics

spatial econometrics - definitions

pitfalls of spatial analysis

spatial autocorrelation

spatial weights
Matrix Algebra Basics
• Review (on your own)

  matrix multiplication, quadratic forms
  matrix inverse
  determinant
  Cholesky decomposition
  eigenvalues, eigenvalue decomposition
  optional: matrix partial derivates
Spatial Econometrics
Definitions
When is Analysis Spatial?

geo-spatial data:
location + value (attribute)

“non-spatial” analysis:
location does NOT matter = locational invariance

spatial analysis:
when the location changes, the information content of the data changes
Spatial Distribution
A-Spatial Distribution
Histogram
Spatial Analysis
Global Spatial Autocorrelation
Moran Scatter Plot
• **Definition of Spatial Econometrics**
  (e.g., Anselin 1988, 2006)

  a subset of econometric methods concerned with spatial aspects present in cross-sectional and space-time observations

  explicit treatment of location, distance and arrangement in model specification, estimation, diagnostics and prediction

  spatial effects: dependence, heterogeneity
• Spatial Effects

  spatial heterogeneity

  structural change, varying coefficients

  standard methods apply

  spatial dependence

  two-dimensional and multidirectional cross-sectional dependence

  not a direct extension of time series methods
Four dimensions of spatial econometrics

- specifying the structure of spatial dependence/heterogeneity
- testing for the presence of spatial effects
- estimating models with spatial effects
- spatial prediction
Pitfalls of Spatial Analysis
• Ecological Fallacy

  individual behavior cannot be explained at the aggregate level

  issue of interpretation

    e.g., county homicide rates do not explain individual criminal behavior

  model aggregate dependent variables with aggregate explanatory variables

  alternative: multilevel modeling
• Modifiable Areal Unit Problem (MAUP)

what is the proper spatial scale of analysis?

a million spatial autocorrelation coefficients (Openshaw)

spatial heterogeneity - different processes at different locations/scales

both size and spatial arrangement of spatial units matter
(a) Hypothetical country with 5 regions

(b) Scaling problem: Change of aggregation level

(c) Zoning problem: Change of boundaries

(d) MAUP-free distance based approach

Source: Scholl and Brenner (2012)
Change of Support Problem (COSP)

- variables measured at different spatial scales
  - nested, hierarchical structures
  - non-nested, overlapping

solutions

- aggregate up to a common scale
- interpolate/impute - Bayesian approach
Spatial Autocorrelation
• The Null Hypothesis

spatial randomness is absence of any pattern

spatial randomness is not very interesting

if rejected, then there is evidence of spatial structure
• Tobler’s First Law of Geography

  everything depends on everything else, but closer things more so

  structures spatial dependence

  importance of distance decay
• Positive Spatial Autocorrelation

impression of clustering

clumps of like values

like values can be either high (hot spots) or low (cold spots)

difficult to rely on human perception
< positive s.a.

random >
• Negative Spatial Autocorrelation

  checkerboard pattern

  hard to distinguish from spatial randomness
Quantile: ZARN09

< negative s.a.

random >
• Spatial Autocorrelation Statistic

captures both attribute similarity and locational similarity

how to construct an index from the data that captures both attribute similarity and locational similarity (i.e., neighbors are alike)
Attribute Similarity

summary of the similarity (or dissimilarity) of observations for a variable at different locations

variable y

locations i, j

how to construct $f(y_i, y_j)$

cross product, $y_i \cdot y_j$

squared difference, $(y_i - y_j)^2$

absolute difference, $|y_i - y_j|$
Locational Similarity

formalizing the notion of neighbors
= spatial weights \((w_{ij})\)

when are two spatial units \(i\) and \(j\) a priori likely to interact

not necessarily a geographical notion, can be based on social network concepts or general distance concepts (distance in multivariate space)
• General Spatial Autocorrelation Statistic

general form

sum over all observations of an attribute similarity measure with the neighbors

\[ f(x_i x_j) \text{ is attribute similarity between } i \text{ and } j \text{ for } x \]

\( w_{ij} \) is a spatial weight between \( i \) and \( j \)

\[ \text{statistic} = \sum_{ij} f(x_i x_j) \cdot w_{ij} \]
• Moran’s I

the most commonly used of many spatial autocorrelation statistics

\[
l = \frac{\left[ \sum_i \sum_j w_{ij} z_i z_j / S_0 \right]}{\left[ \sum_i z_i^2 / N \right]}
\]

with \( z_i = y_i - m_x \) : deviations from mean

cross product statistic \((z_i . z_j)\) similar to a correlation coefficient

value depends on weights \((w_{ij})\)
Spatial Weights
Basic Concepts
Why Spatial Weights

formal expression of locational similarity

spatial autocorrelation is about interaction

n x (n - 1)/2 pairwise interactions but only n observations in a cross-section

insufficient information to extract pattern of interaction from cross-section

example: North Carolina has 100 counties 5,000 pairwise interactions, 100 observations
Solution

• impose structure

limit the number of parameters to be estimated

incidental parameter problem = number of parameters grows with sample size

for spatial interaction, number of parameters grows with $n^2$
• **Spatial Weights**

  exclude some interactions

  constrain the number of neighbors, e.g., only those locations that share a border

  single parameter = spatial autocorrelation coefficient

• strength of interaction = combined effect of coefficient and weights

  small coefficient with large weights

  large coefficient with small weights
six polygons - neighbors share common border
neighbor structure as a graph

node

link
Spatial Weights Matrix Definition

N by N positive matrix $W$ with elements $w_{ij}$

- $w_{ij}$ non-zero for neighbors
- $w_{ij} = 0$, i and j are not neighbors
- $w_{ii} = 0$, no self-similarity
Geography-Based Spatial Weights
• Binary Contiguity Weights

contiguity = common border

i and j share a border, then $w_{ij} = 1$

i and j are not neighbors, then $w_{ij} = 0$

weights are 0 or 1, hence binary
$$W = \begin{bmatrix}
0 & 1 & 0 & 1 & 1 & 0 \\
1 & 0 & 0 & 1 & 1 & 0 \\
0 & 0 & 0 & 0 & 1 & 1 \\
1 & 1 & 0 & 0 & 1 & 0 \\
1 & 1 & 1 & 1 & 0 & 0 \\
0 & 0 & 1 & 0 & 0 & 0 \\
\end{bmatrix}.$$ 

binary contiguity weights matrix for six-region example
contiguity on a regular grid - different definitions
Rook contiguity

Neighbors for 5: 2, 4, 6, 8

Common border

rook contiguity - edges only
2, 4, 6, 8 are neighbors of 5
bishop contiguity - corners only
1, 3, 7, 9 are neighbors of 5
queen contiguity - edges and corners
5 has eight neighbors
Solano county, CA contiguity

rook

queen
• **Distance-Based Weights**

  distance between points

  distance between polygon centroids or central points

  in general, can be any function of distance that implies distance decay, e.g., inverse distance

  in practice, mostly based on a notion of contiguity defined by distance
• Distance-Band Weights

\[ w_{ij} \text{ nonzero for } d_{ij} < d \]
less than a critical distance \( d \)

potential problem: isolates = no neighbors

make sure critical distance is max-min, i.e., the largest of the nearest neighbor distance for each observation
• k-Nearest Neighbor Weights

k nearest observations, irrespective of distance

fixes isolates problem for distance bands

same number of neighbors for all observations

in practice, potential problem with ties

needs tie-breaking rule (random, include all)
Spatial Weights Transformations
• Row-Standardized Weights

rescale weights such that $\sum_j w_{ij} = 1$

$w_{ij}^* = w_{ij} / \sum_j w_{ij}$

constrains parameter space

makes analyses comparable

spatial lag = average of the neighbors
\[ W^* = \begin{bmatrix}
0 & 1/3 & 0 & 1/3 & 1/3 & 0 \\
1/3 & 0 & 0 & 1/3 & 1/3 & 0 \\
0 & 0 & 0 & 0 & 1/2 & 1/2 \\
1/3 & 1/3 & 0 & 0 & 1/3 & 0 \\
1/4 & 1/4 & 1/4 & 1/4 & 0 & 0 \\
0 & 0 & 1 & 0 & 0 & 0 \\
\end{bmatrix} . \]

row-standardized weights matrix
Higher Order Weights
second order contiguity: neighbor of neighbor
redundancy in higher order contiguity
paths of length 2 between 1 and other cells
• Higher Order Weights

  recursive definition

  • k-th order neighbor is first order neighbor of (k-1)th order neighbor

  avoid duplication, only unique neighbors of a given order (not both first and second order)

  pure contiguity or cumulative contiguity, i.e., lower order neighbors included in weights
exclusive of first order

inclusive of first order

Solano county, CA, second order contiguity
Properties of Weights
• Connectivity Histogram

histogram of number of neighbors

neighbor cardinality

diagnostic for “isolates” or neighborless units

assess characteristics of the distribution
• Things to Watch for

isolates

need to be removed for proper spatial analysis

do not need to be removed for standard analysis

very large number of neighbors

bimodal distribution