

Thinking about, and then Making Spatial Weights.pdf 2017 BurkeyAcademy

This handout (and videos, etc.) available at <http://spatial.burkeyacademy.com>

Intro Video on Contiguity and Spatial Correlation: <https://www.youtube.com/watch?v=cdWH5frHg5E>

YouTube Playlist of all videos: https://www.youtube.com/playlist?list=PLInEW8MeJ4z6Du_cbY6o08KsU6hNDkt4k

Preliminaries:

- 1) You should think about using weights that make sense.
- 2) You cannot have “Islands”: Areas with no neighbors.
- 3) At the opposite extreme, some say that weights matrixes should be rather *sparse*. I.E., most entries should be zeros. Some use distance-based weights with no cutoff, so that **all** are neighbors, to some degree.
- 4) **Lesage and Pace (2014)** “Biggest Myth” applies for SAR and SDM, but not SLX and perhaps not in other circumstances <http://www.mdpi.com/2225-1146/2/4/217>

Summary: Slope coefficients are sensitive to choice of W, but marginal effects are not (so much)

- 5) Weights need to be standardized.
 - i. For Queen, rook, KNN, or “everybody within x distance is a neighbor”
 - a. Row Standardization (weights between 0 and 1, interpretation as average of neighbors)
 $y = \rho W y + X \beta + W X \theta + u, u = \lambda W u + \varepsilon$ (*Manski Model*)
 - ii. For inverse distance or other “strength-based” weights (*could row standardize*)
 - b. Divide by largest element of matrix (keeps all between zero and 1)
 - c. Divide by largest eigenvalue (see *eigenW* in *spdep*) (*Boots and Royle, 1991?*)

iii. Cliff-Ord Weights: $w_{ij} = \frac{b_{ij}^{\beta}}{d_{ij}^{\alpha}}$ b_{ij} = border share amount, d_{ij} = distance

α and β are parameters, $\alpha=2$ is gravity decay model, should probably standardize...

- 6) Create your Spatial Weights in GeoDa, then import them into R, Matlab, Stata, python, ... (e.g. *readgal*)
Why?
 - i. knn (k nearest neighbors) based on centroids is what many people do in R, Matlab, Stata, Julia, etc.
because it is EASY, not because it is right: but should not be your default choice
 - ii. You can visualize your spatial areas AND your weights to make sure they seem reasonable.
- 6) Defining Non-spatial neighbors *E.g., countries’ trading partners, teenagers friend network, etc.*

BE CAREFUL! Weights must be EXOGENOUS.

For more: Lesage and Pace (2011) Review of Regional Studies <http://journal.srsa.org/ojs/index.php/RRS/article/view/39/205>

More on Spatial Weights: Getis and Aldstadt (2010) Geographical Analysis <http://onlinelibrary.wiley.com/doi/10.1111/j.1538-4632.2004.tb01127.x/pdf>

GeoDa Weights Files:

GAL file: Contain only neighbor information

GWT File: Contain neighbors AND distance information

Queen, Rook, knn GAL file

Distance-based weights GWT

Let’s Experiment in GeoDa: Open our NC and VA data, try some different weights!

Our Dataset we’ve been working with:

<http://spatial.burkeyacademy.com/home/files/2.1%20NCVA%20CO%20Variables%20Shapefile.zip>

Spatial Correlation of Per Capita Liquor Sales:

| | | |
|-------------------------|--------------------|-----------------|
| Queen: | Moran's I: -0.0139 | pseudo p: 0.44 |
| Rook: | Moran's I: -0.0130 | pseudo p: 0.44 |
| Distance-based 50mi: | Moran's I: -0.0118 | pseudo p: 0.22 |
| Distance-based 200mi: | Moran's I: 0.0165 | pseudo p: 0.004 |
| Distance-based MAX : | Moran's I: -0.0159 | pseudo p: 0.001 |
| Knn: 2 nearest n'bers: | Moran's I: -0.0536 | pseudo p: 0.22 |
| Knn: 10 nearest n'bers: | Moran's I: -0.0078 | pseudo p: 0.476 |
| Knn: 30 nearest n'bers: | Moran's I: 0.0365 | pseudo p: 0.008 |