

Iowa Social Science Research Center
2018-19 workshop series



Spatial analysis with GeoDa and ArcMap

Seungwon Kim

Geographical and Sustainability Sciences

Feb 21, 2019

GeoDa

- A **FREE** software package
- With GeoDa, you may be able to do..
 - Spatial analysis (okay)
 - Geo-visualization (poor)
 - Spatial modeling (easy & not bad but recommend to use R in the future)
 - Linear Regression, Spatial Lag Model, Spatial Error Model
- Mac, Windows, Linux
- Unstable but evolving slowly
- <http://geodacenter.github.io/>

ArcMap

- A software package (pretty expensive)
- With ArcMap, you may be able to do..
 - Spatial analysis (okay..)
 - Geo-visualization (Very powerful)
 - Spatial modeling (limited)
 - Geographically weighted regression
- Only on Windows system



Simple Linear Regression (SLR)

- y_i is the response variable (a.k.a. Dependent variable)
- x_i is the explanatory variable or predictor variable (a.k.a Independent variable)
- We are using X to explain some or most of the variability of y , in particular the non-random part of y 's variability

Steps for linear regression

- Response variable
 - Type of variable
 - Continuous: Simple or multiple regression
 - Count data: Poisson regression or negative binomial regression
 - Binary: Logistic regression
 - Independent?
 - Study design
 - Ex) School, classroom, students
- Linear relationship between response and predictor(s)
 - Scatter plot

Steps for linear regression (cont.)

- Predictor
 - Check multicollinearity
 - Correlation matrix
 - Variance Inflation Factor (VIF)
- Error
 - Normality
 - Q-Q plot
 - Shapiro-Wilk test
 - H_0 : The population is normally distributed
 - Heteroscedasticity
 - Plot (studentized) residuals vs. predicted y
 - Breusch-Pagan test
 - H_0 : Homoskedasticity
 - Autocorrelation
 - Plot
 - Durbin-Watson test

The first law of geography

- Waldo Tobler
 - “Everything is related to everything else, but near things are more related than distant things.”
- Spatial dependence
 - The co-variation of properties within geographic space
 - Weather map (Kriging)
 - Spatial autocorrelation

Back to regression analysis

- Response variable

- Type of variable

- Continuous: Simple or multiple regression

- Count data: Poisson regression or negative binomial regression

- Binary: Logistic regression

- Independent? (Spatially independent?)

- Study design

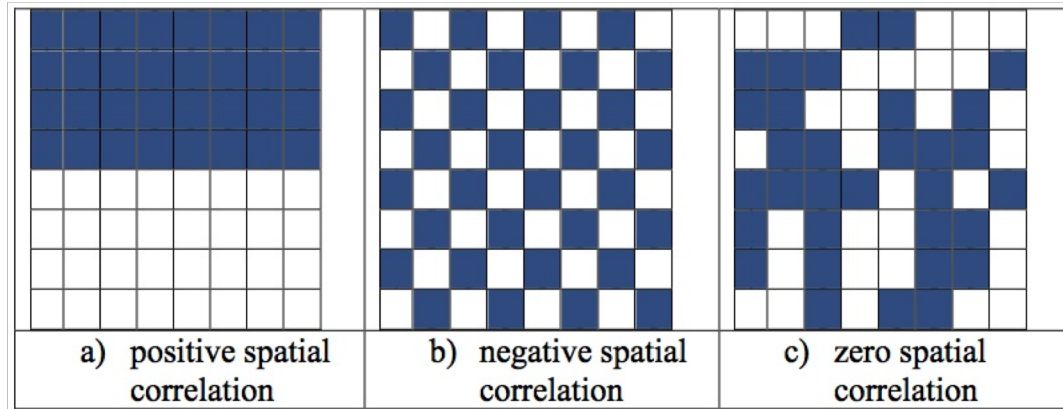
- Ex) School, classroom, students

Back to regression analysis

- Error
 - Normality
 - Q-Q plot
 - Shapiro-Wilk test
 - Heteroscedasticity
 - Plot (studentized) residuals vs. predicted y
 - Breusch-Pagan test
 - Autocorrelation (Spatial autocorrelation)
 - Plot
 - Burbin-Watson test

Check Spatial dependence

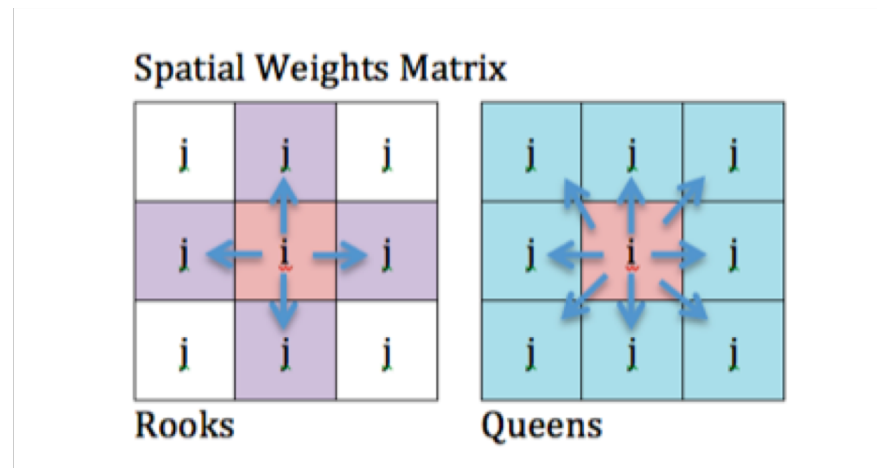
- Moran's I
 - A measure of spatial autocorrelation



- A summary statistic with p-value from bootstrapping

Spatial Weight matrix

- The spatial relationships that exist among the features in your dataset
- NxN matrix
- Two ways to determine neighbors
 - Rooks case
 - Queens case



Spatial regression models

- Spatial Lag model

- Assumes that dependencies exist directly among the levels of the response variable
- Used when we know the structure of spatial dependence
- $Y = \rho WY + X\beta + \varepsilon$

- Spatial Error model

- Error term has dependence
- Used when structure of dependence is unknown
- $Y = X\beta + \lambda W\varepsilon + \xi$

Lab 1. Moran's I

- Go to: <http://geodacenter.github.io/>
 - Download GeoDa if you want to use your own pc or mac
 - Go to data and download 'NYC Education'
- Open *.shp in GeoDa
- Tools > Weights Manager > create > Add ID variable > Ok > Create
- Save
- Space > Univariate Moran's I
- Space > Univariate local Moran's I

Lab 2. MLR vs Spatial Lag model

- MLR
 - Regression > Regression
 - Set Dependent variable and Covariates
 - Check the box next to 'Weights File'
 - Check the box next to 'Pred. Val. And Res.'
 - Run and save to file
- Spatial lag model
 - Same but select 'Spatial lag' instead of 'Classic' in the model
 - Run and save to file

Lab 3. Visualization in ArcMap

- Open ArcMap
- Add *.shp
- Mapping residuals from MLR and Spatial regression models
- Compare