SPACE Workshop

- NSF →NCGIA →CSISS →UCGIS →SDSU

- Aldstadt, Getis, Jankowski, Rey, Weeks → SDSU
- F. Goodchild, M. Goodchild, Janelle, Rebich → UCSB

- August 2-8, 2004
- San Diego State University

Some Examples of Spatial Interests in the Social Sciences

- **Sociology**: Behavior in Space; Ethnic Patterns; Spatial Patterns of Criminal Activities; Spatial Manifestation of Demographic Trends
- **Political Science**: Spatial Patterns of Voting; Redistricting; Diffusion of Political Movements
- **Anthropology and Archaeology**: Patterns of Human Activities (usually local in scale); Re-creation of Past Settlement Patterns
- **Economics**: Spatial Aspects of Economic Variables, Trends, Location Patterns, Economic Concentrations, Trade Patterns, Value of Place
- **Urban Planning and Geography**: Human Spatial Interaction Patterns, Distance Decline, Land Use Patterns, Transportation Systems
- **Transportation**: Movement, Forecasting Demand, Accidents
- **History**: Patterns of Change in Socio-economic Patterns Over Time
- **Public Health**: Disease Diffusion; Patterns of Care; Clustering of Disease

Where does *Spatial Understanding* come from?

- Location (basic question -- site and situation)
- Spatial Interaction (communication, information, movement, the role of distance)
- Land Use Patterns (spatial associations)
- Spatial Point Patterns (significant associations, clustering, direction)
- Pattern Processes (evolution of associations)

Where does *Spatial Understanding* come from?

- Transportation Routing (connecting space, direction)
- Inter-Industry Relations (economies of agglomeration, urbanization)
- Urban Growth (spatial growth and form of cities, urban decline, too)
- Irregular Surfaces (why the heterogeneity?)
- Population density (where and intensity of human agglomerations

Where does *Spatial Understanding* come from?

- Population density (where and intensity of human agglomerations)
- Size of Cities (nested hierarchies of cities spatially patterned)
- Scale Effects (at what spatial scale is our understanding best?)
- Diffusion (the spatial spread of ideas, culture, disease, etc.)
- Spatial Autocorrelation (the quantitative relationship of associated sites)

Methods for Learning about Space

• Location	Regression, Linear Programming
• Spatial Interaction	Gravity Analogs, Population Potential Models
• Land Use Patterns	Classification, Factor Analysis
• Spatial Clustering	Poisson Distribution as Randomness
• Pattern Processes	Probability Distributions (two-stage)

Methods for Learning about Space

• Routing	Least Cost Algorithms, Operations Research
• Inter-Industry	Input-Output, Multipliers
• Urban Growth	Basic-Service (Non-basic) Ratios
• Irregular Surfaces	Map Transformations
• Population Density	Negative Exponential, Distance Decay

Methods for Learning about Space

• Size of Cities	Rank-size, Hierarchies
• Scale Effects	Correlation, Variance
• Diffusion	Monte Carlo Simulations
• Spatial Auto-correlation	Correlation, Cross Products

Ways to Study Space

- Map Spatial Data (GIS)
- Explore Properties of Geographical Space (ESDA-GIS/GEODA)
- Visualization techniques (GEODA)
- Location Modeling
- Spatial Statistics (Global/Local)
- Geostatistics
- Spatial Econometrics
- Spatial Choice

Spatial Pattern Identification

- ESDA and GIS
- Data Manipulation
- Geostatistics
- Local and Global Statistics
- Data Mining
- Spatial Autocorrelation Analysis
- Segregation Indices
- Tesselations (Voronoi Polygons)

ESDA/GEODA

- GIS Functionality (buffers, distances, etc)
- Map Pattern Measures
- Histograms, Box Plots
- Multiple Scatter Diagrams, Surfaces
- Residuals from Regression
- Visualization

Temporal vs. Spatial Data

Temporal

- 1 dimension
- Units: day,
  week, month
- Lag: t, t-1, t-2
- Durbin-Watson
- Differencing

Spatial

- 2-3 dimension
- Units: county,
  mile, region
- Lag: near neighbor,
  networks (?)
- Moran's I (W)
- Maps (distortions)

Matrix Representation

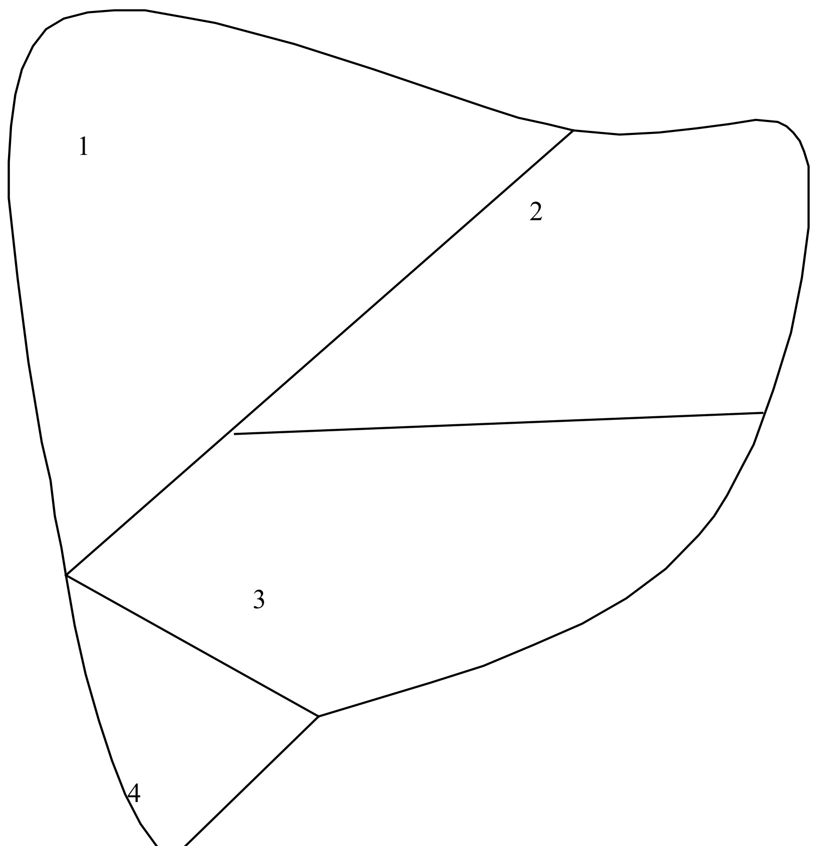
- **W**
- The Spatial Weights Matrix
- The Spatial Association of All Sites to All Other Sites
- d, d^2, 1/0, 1/d

- **Y**
- The Attribute Association Matrix
- The Association of the Attributes at Each Site to the Attributes at All Other Sites
- +,-,/,x

The Spatial Weights Matrix

W is the formal expression of the spatial association between objects

(it is the pair-wise geometry of objects being studied).



0	1	1	0
1	0	1	0
1	1	0	1
0	0	1	0

Typical W

- Spatially contiguous neighbors (rook, queen: one/zero)
- Inverse distances raised to a power: ($1/d$, $1/d^2$, $1/d^5$)
- Geostatistics functions (spherical, gaussian, exponential)
- Lengths of shared borders (perimeters)
- All centroids within distance *d*
- n^{th} nearest neighbor distance
- Links (number of)

Attribute Relationships

Y

- **Types of Relationships**

Additive association (clustering): $(Y_i + Y_j)$

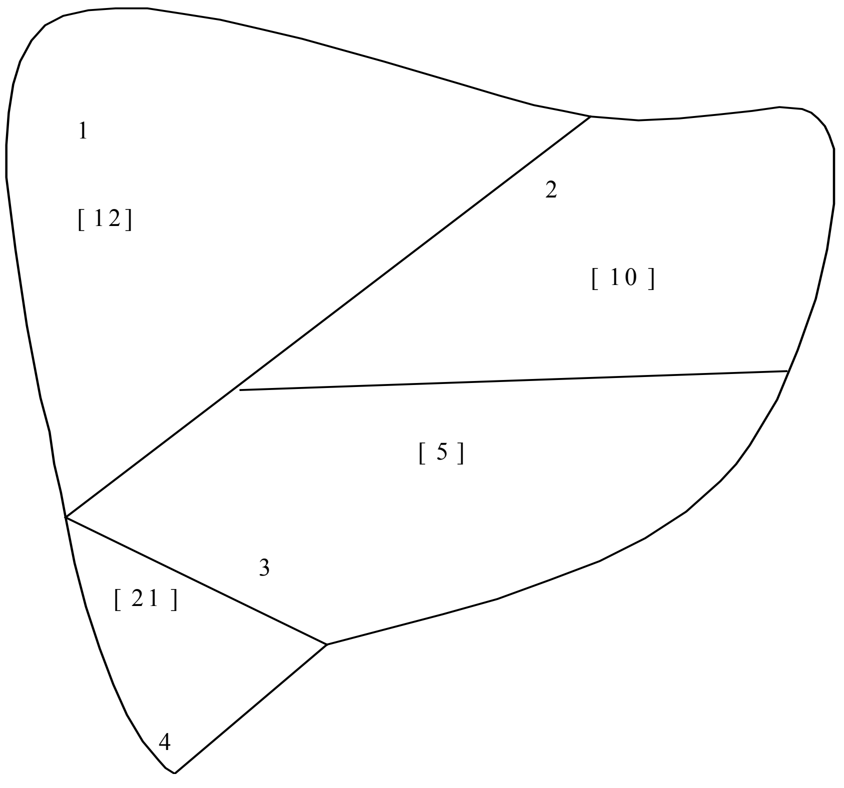
Multiplicative association (product): $(Y_i Y_j)$

Covariation (correlation): $(Y_i - Ybar)(Y_j - Ybar)$

Differences (homogeneity/heterogeneity): $(Y_i - Y_j)$

Inverse (relativity): (Y_i/Y_j)

- All Relationships Subject to Mathematical Manipulation (power, logs, abs, etc.)



0	2	7	-9
-2	0	5	-11
-7	-5	0	-16
9	11	16	0

Spatial Statistics

- Based on Conventional Statistical Theory
- Chance or Non-chance Occurrences
- Measures of Association, Segregation
- All Interevent Distances: *k*-functions
- Clustering Statistics: *I, G, c,* etc.
- Specially Developed Tests on Randomness (or Normal) Hypotheses

Global Statistics

- Nearest Neighbor
- *k*-Function
- Global Autocorrelation Statistics

  Moran's *I*

  Geary's *c*

  Semivariance

Local Statistics

- Focus on a site *i*
- To what extent are values at sites in vicinity of i associated with *i*
- Anselin's LISA: **Moran's**, Geary's local
- Getis and Ord's G_i* (A clustering statistic)

The I_i Statistic (A LISA Statistic): Local Indicator of Spatial Association

- The I_i statistic is local, that is, it is focused on a site and is normally distributed. It is designed to yield a measure of pattern that can be translated into standard normal variates.
- A covariance statistic. A measure of similarity (differences). Well-suited for study of residuals from regression
- Indicates the extent to which a location (site) is surrounded to a distance *d* (or by contiguity) by a cluster of high or low values (H-H, L-L, H-L).
- Can identify clusters of H-H, L-L, H-L.

Spatial Econometrics

- Regression Models with One or More Spatial Parameters (that describes the effect of distance -- in conjunction with certain variables -- on dependent variable)
- Development of Spatial Association Matrices (describes hypothesized spatial effects)
- Parameter Estimation and Testing (model identification)
- Spatial Filtering (removing spatial effects)
- Study of Model Assumptions Including Spatially Random Residuals

Typology of Spatial Econometric Models

- General Model:

$\mathbf{Y} = \theta \mathbf{W_1 Y} + \mathbf{X}\beta + \varepsilon$

$\varepsilon = \kappa\, \mathbf{W_2}\varepsilon + \lambda$ with λ normal, 0 mean, and constant variance Ω (i.e., variance is the same for every variable and covariance for every combination of variables is always 0)

Typology of Spatial Econometric Models

$\mathbf{Y} = \theta\mathbf{W_1 Y} + \mathbf{X}\beta + \varepsilon$ **and** $\varepsilon = \kappa\, \mathbf{W_2}\varepsilon + \lambda$

set θ=0, κ = 0 ------ RESULT is $\mathbf{Y} = \mathbf{X}\beta + \varepsilon$

set κ = 0 ------- RESULT is $\mathbf{Y} = \theta\mathbf{W_1 Y} + \mathbf{X}\beta + \varepsilon$

set θ=0 ------- RESULT is $\mathbf{Y} = \mathbf{X}\beta + (\mathbf{I} - \kappa\, \mathbf{W_2})^{-1}\lambda$

also ------- $\mathbf{Y} = \theta\mathbf{W_1 Y} + \mathbf{X}\beta + (\mathbf{I} - \kappa\, \mathbf{W_2})^{-1}\lambda$

These are respectively: linear regression model,

spatial lag model,

spatial disturbance model,

spatial lag and disturbance model (spatial Durbin).

Testing the Models

- In any model, if error term is correlated, OLS is inappropriate device to find parameters. Usual tests on parameters and R^2 cannot be used.
- Use maximum likelihood approach (i.e., the parameters most likely to give you your data). Spatial lag and error models.

Wald test on parameters; **Likelihood Ratio** test on the goodness of the model; **La Grange Multiplier** test on residuals (non-spatial); **Moran's I** test on residuals (spatial)

Example

- **Crime in Columbus**
- OLS Diagnostics
- CR = 68.6 - 1.6 IN - 0.3 Ho
- (14.5) (-4.8) (-2.7)
- R^2 = .552 Adj R^2 = .533
- Z(I) = 2.95
- Spatial disturbance
- Therefore, model inappropriate

Spatial Lag Model

- Bisection search for θ parameter yields 0.**4310**.

  CR = 45.1 + **0.43** W_CR - 1.0 IN - 0.3 Ho
- (6.3) (3.7) (-3.4) (-3.0)
- Now, La Grange Test on Residuals (non-spatial) = OK

  Z(I) = 0.65
- Conclusion: Not strong evidence of a spatial disturbance process after introduction of a spatially lagged dependent variable.