Spatial Weights

Luc Anselin

University of Illinois, Urbana-Champaign

http://geog55.geog.uiuc.edu/sa

Outline

- Connectivity in Space
- Spatial Weights
- Practical Issues
- Spatial Lag Operator

Spatial Connectivity

Why Spatial Weights?

- Spatial Correlation
  - $Cov[y_i.y_h] \neq 0$, for $i \neq h$
- Structure of Correlation
  - which i, h interact?
  - N observations to estimate N(N-1)/2 interactions
  - impose structure in terms of what are the "neighbors" for each location

Spatial Arrangement

- Need to Impose Structure on the Extent of Spatial Interaction
- Neighborhood View
  - define neighborhood set N(i) for each location i
  - spatial weights matrix
- Pairs View
  - order pairs of locations i-j in function of separating distance
  - semivariogram (geostatistics)

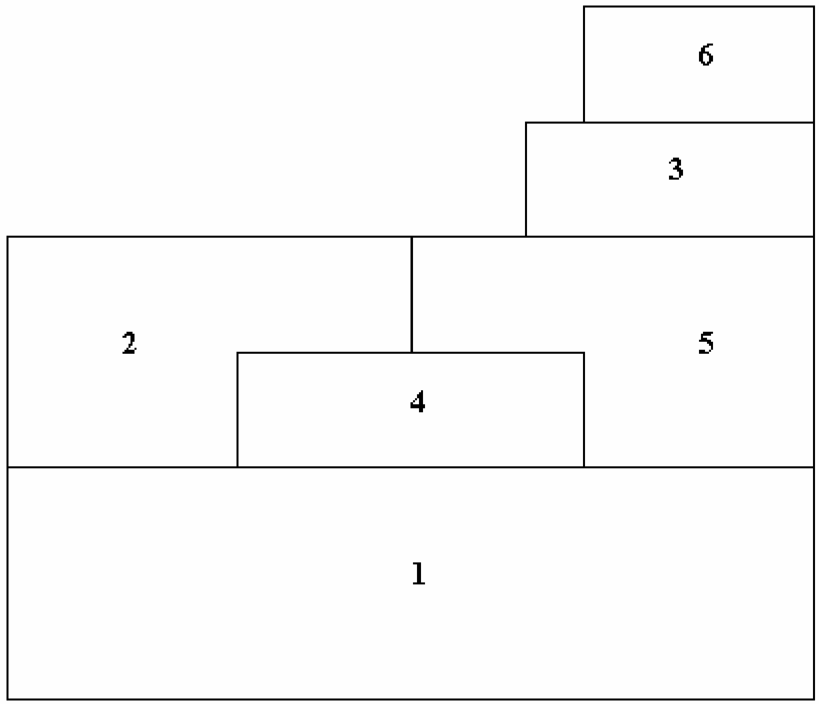
Neighborhood Set

- Geographic-Cartographic Contiguity (GIS)
  - common boundary = contiguity
    - » common border, common vertex
  - distance band = isotropy
  - interaction border length and distance
- Spatial Interaction
  - distance decay, gravity, entropy
  - scale dependent, identification problems

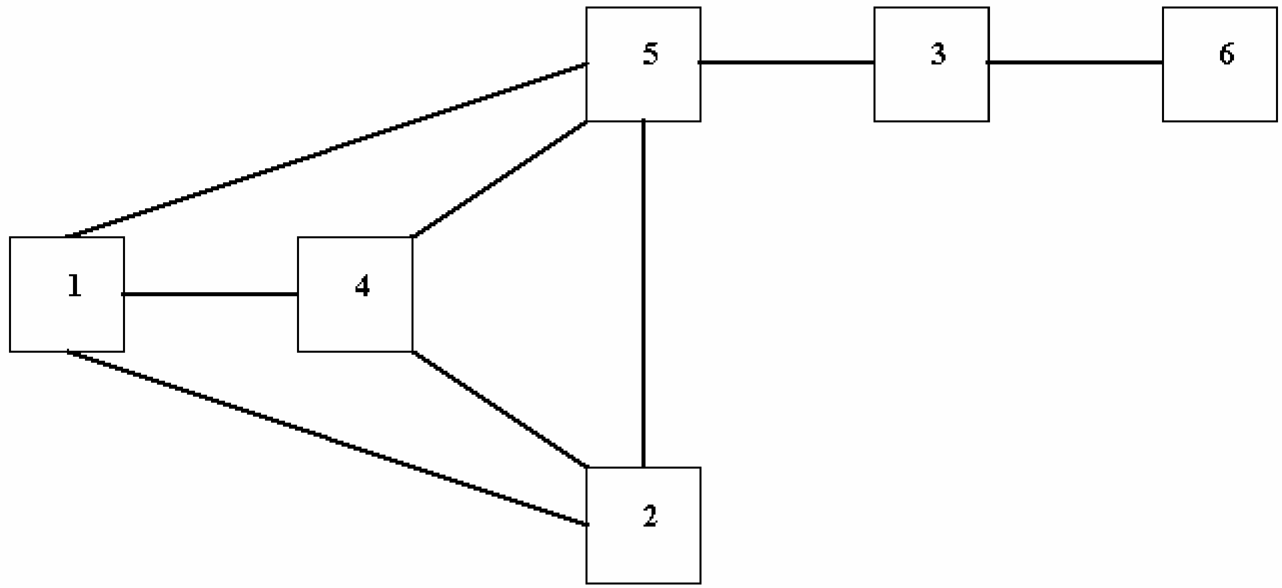
Neighborhood Set (continued)

- Socio-Economic Distance
  - multidimensional distance based on socio-economic indicators
    - » Euclidean, Mahalanobis
    - » example: income, ethnicity, industrial structure, trade flows, migration flows
  - problem with endogeneity
    - » variables for distance same as in model
  - zero distances
    - » $1/(z_i - z_j)$ when $z_i = z_j$

Spatial Weights



Example: N=6
contiguity = common boundary



contiguity as a graph
link between nodes = contiguity

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

First Order Contiguity Matrix

Spatial Weights Matrix

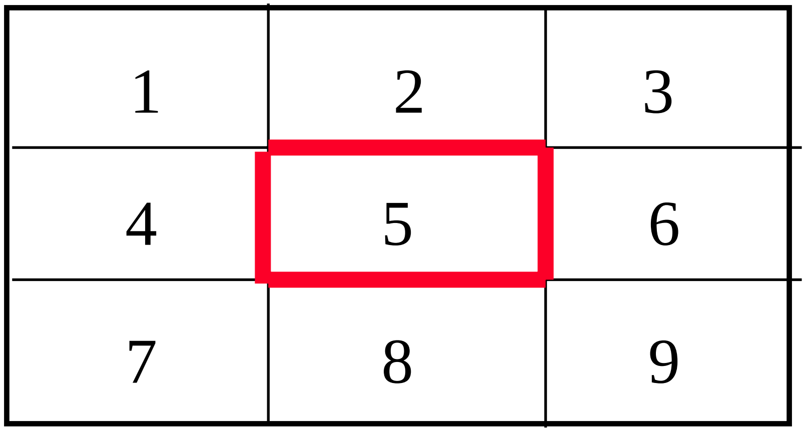
- Definition
  - N by N positive matrix W, with elements w_{ij}
- Simplest Form: Binary Contiguity
  - » $w_{ij} = 1$ for i and j “neighbors” (e.g. $d_{ij} < $ critical distance)
  - » $w_{ij} = 0$ otherwise,
  - » $w_{ii} = 0$ by convention

How to Define Weights

- Contiguity
  - common boundary
- Distance
  - distance band
  - k-nearest neighbors
- General
  - social distance
  - complex distance decay functions

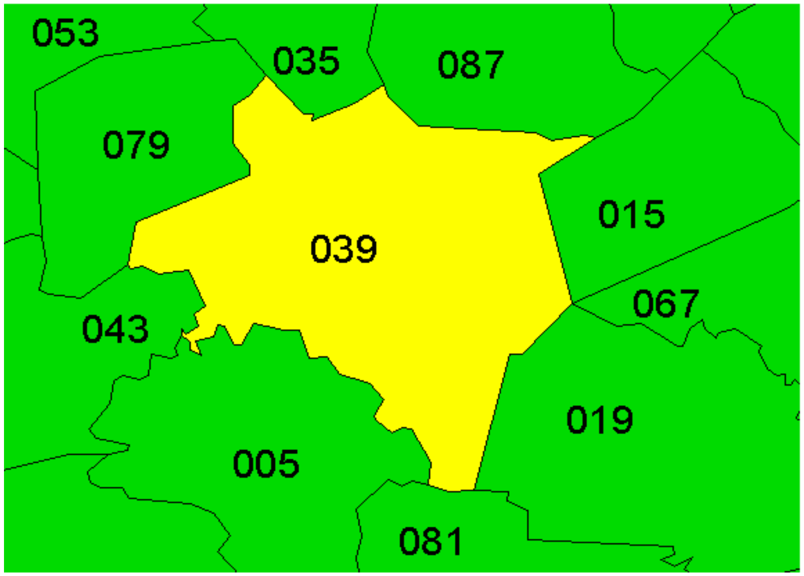
Contiguity – Regular Grid

- Regular Grid
  - rook
    - » 2, 4, 6, 8
  - bishop
    - » 1, 3, 7, 9
  - queen
    - » both

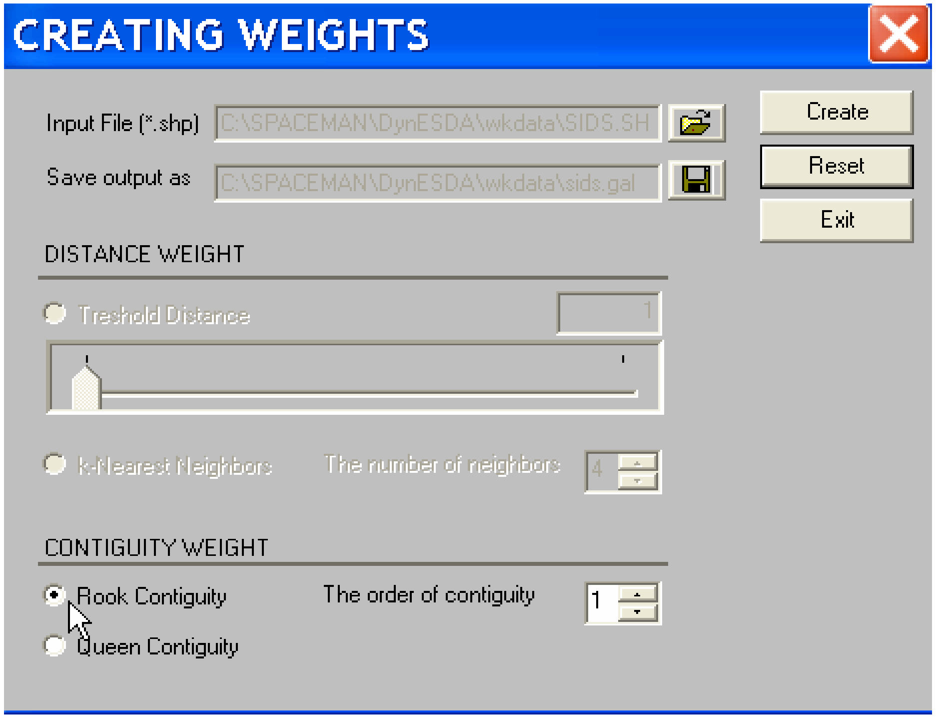


Contiguity – Irregular Units

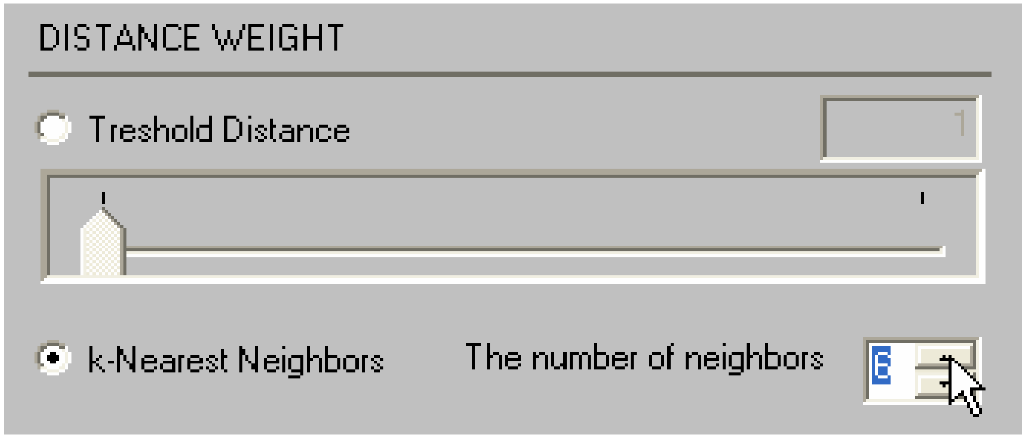
- Irregular Units
  - common border
    - » rook
  - common vertex
    - » 039 and 067
    - » queen



Contiguity Weights in DynESDA2



Distance Based Weights



General Spatial Weights

- Cliff-Ord Weights
  - w_{ij} to reflect potential spatial interaction between i and j
  - $w_{ij} = [d_{ij}]^{-a}.[b_{ij}]^{b}$
    - » with
      d_{ij} as distance between i and j
      b_{ij} as share of common boundary between i and j in perimeter of i

General Spatial Weights
(continued)

- Weights May Contain Parameters
  - inverse distance weights
    - » $w_{ij} = 1 / d_{ij}^{\alpha}$
  - estimated from data or chosen a priori
    - » in practice: second power (gravity model)
  - identification problems in nonlinear weights
    - » interaction is multiplicative: $\rho \, . \, w_{ij} = \rho \, (1 / d_{ij}^{\alpha})$
    - » parameters ρ and α not separately identified

General Spatial Weights

(continued)

- K-Nearest Neighbor Weights
  - k neighbors, irrespective of actual distance
  - "warps" space
- Economic Weights (Case)
  - block structure, state effect
    - » $w_{ij} = 1$ for all i, j in "block"
  - economic distance $|r_i - r_j|$, weight = $1/|r_i - r_j|$
    - » e.g., r = total employment

Row-Standardized Spatial Weights

- Motivation
  - averaging of neighboring values = form of spatial smoothing
  - spatial parameters comparable
- $w^s_{ij} = w_{ij} / \Sigma_j w_{ij}$
- $\Sigma_j w^s_{ij} = 1$

Practical Issues

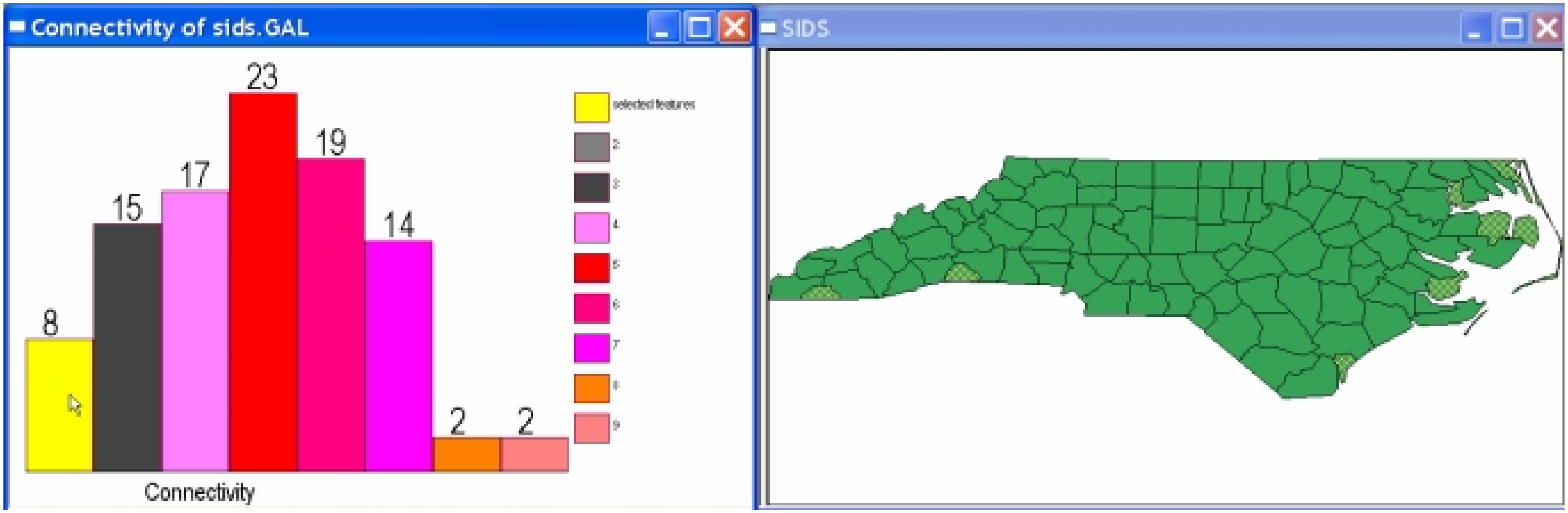
Characteristics of Spatial Weights

- Measures of Overall Connectedness
  - percent nonzero weights (sparseness)
  - average weight
  - average number of links
  - principal eigenvalue
- Location-Specific Measures
  - most/least connected observations
  - unconnected observations = islands

```
Connectivity characteristics of weights matrix wvhouse
Weights matrix stored in GAL format
Weights are row-standardized
Dimension:                          55
# nonzero links:                   248
% nonzero weights:                8.35017
Average weight:                   0.221774
Average # links:                  4.50909
With observations referred to by indicator variable FIPSNO
 1 most connected observation(s) - with  8 link(s):
 54039
 2 least connected observation(s) - with  1 link(s):
 54029  54037
```

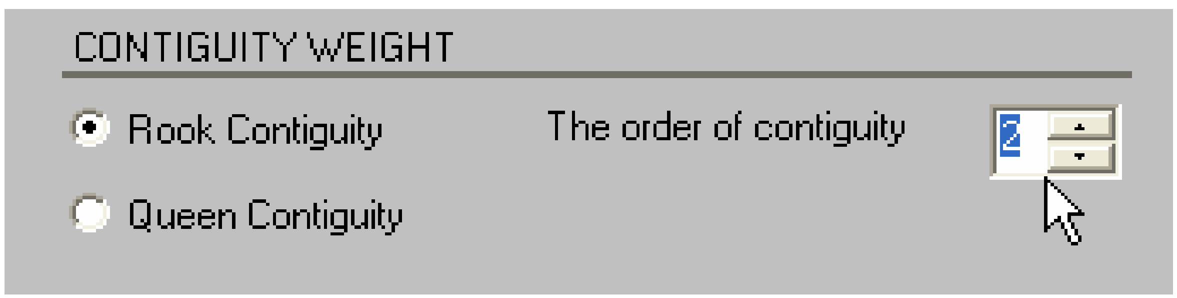
```
Frequency Count for Weights File wvhouse
Connections     Frequency

          1           2
          2           6
          3           7
          4          10
          5          14
          6          10
          7           5
          8           1
```

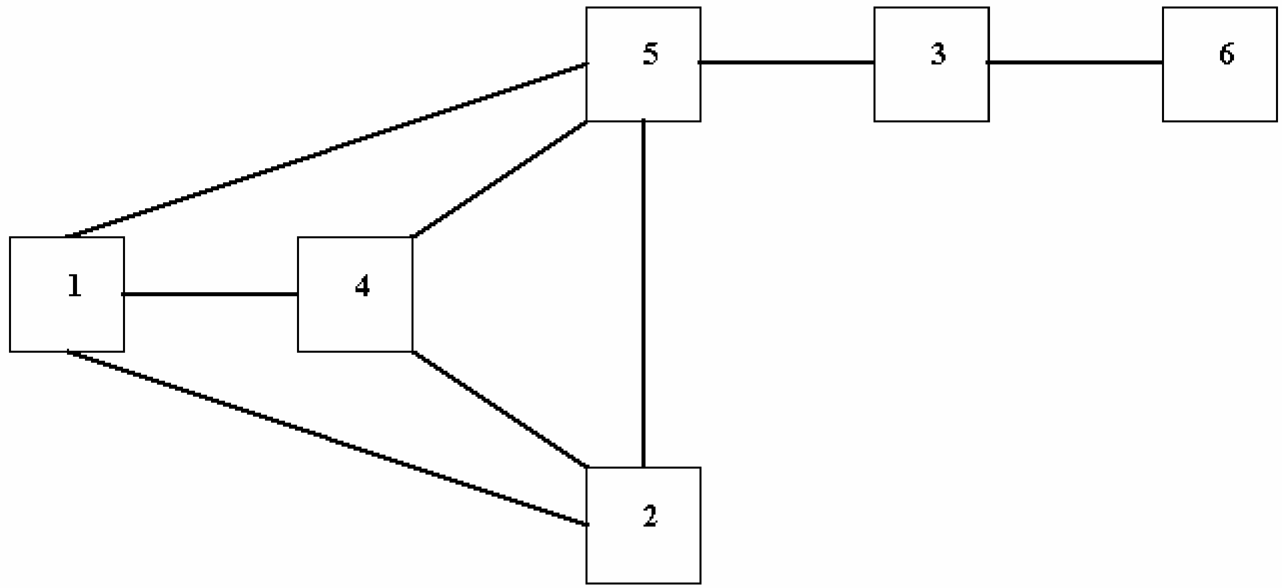
Weights Characteristics in DynESDA2



Higher Order Contiguity

- Recursive Definition
  - contiguous of order k is first order contiguous to order k-1
    - » 2^{nd} is first order contiguous to first
- Remove Circularity and Redundancy





contiguity as a graph
link between nodes = contiguity

Circularity and Redundancy in Higher Order Weights

- Powering of the Weights Matrix
  - standard approach invalid for contiguity
- Removing Circularity and Redundancy
  - sparse network representation of weights
  - modified Dijkstra algorithm to identify number of steps between nearest neighbors (Anselin and Smirnov)
  - number of steps = order of contiguity

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

First Order Contiguity Matrix

3	2	1	2	2	0
2	3	1	2	2	0
1	1	2	1	0	0
2	2	1	3	2	0
2	2	0	2	4	1
0	0	0	0	1	1

Second Power of First Order Contiguity

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

Correct Second Order Contiguity Matrix

-1	1	2	1	1	3
1	-1	2	1	1	3
2	2	-1	2	1	1
1	1	2	-1	1	3
1	1	1	1	-1	2
3	3	1	3	2	-1

Matrix Constructed by Bottom-Up Algorithm

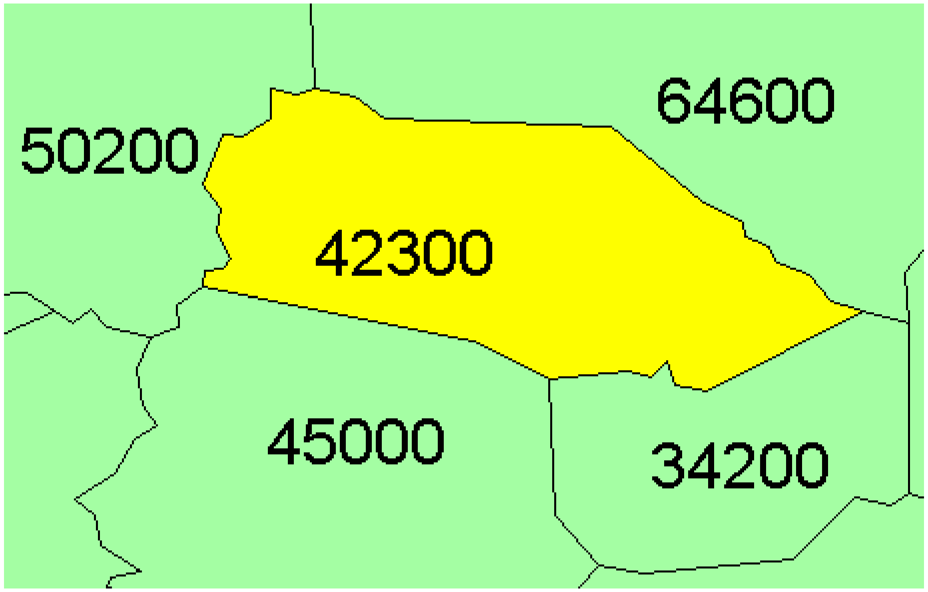
Spatial Lag Operator

Spatial Shift

- No Direct Counterpart to Time Series Shift Operator
  - time series: $L_k = y_{t-k}$
  - spatial series: which h are shifted by "k" from location i?
    - » on regular lattice: east, west, north, south
    - » $(i - 1, j)$ $(i + 1, j)$ $(i, j - 1)$ $(i, j + 1)$
  - arbitrary for irregular lattice
    - » different number of neighbors by observation

Spatial Lag Operator

- Distributed Lag
  - row-standardized weights $\Sigma_j w_{ij} = 1$
  - spatial lag is weighted average of neighboring values
    - » $\Sigma_j w_{ij}.y_j$, for each i
    - » vector Wy
    - » spatial lag does not contain y_i
  - spatial lag is a smoother
    - » not a window average



value y_i = $42,300

4 neighbors

values for neighbors: $50,200, $64,600, $45,000, $34,200

spatial lag = (1/4)$50,200 + (1/4)$64,600 + (1/4)$45,000 + (1/4)$34,200 = $48,500

Interpretation of Spatial Lag

- Linear Association = Spatial Autocorrelation
  - comparison of value of y at i to average of values at neighboring locations
    - » y_i and $(Wy)_i$ similar = positive spatial autocorrelation (high-high, low-low)
    - » y_i and $(Wy)_i$ dissimilar = negative spatial autocorrelation (low-high, high-low)