Spatial Weights and Spatial Lag

SERGIO REY

Geographic Information Analysis
School of Geographical Sciences and Urban Planning
Arizona State University

Outline

Outline

Spatial Autocorrelation Statistic

Structure

- Formal Test of Match between Value Similarity and Locational Similarity
- Statistic Summarizes Both Aspects
- Significance
  - how likely is it (p-value) that the computed statistic would take this (extreme) value in a spatially random pattern

Locational Similarity

- Formalizing the notion of Neighbor
  - when two spatial units a-priori are likely to interact
- Spatial Weights
  - not necessarily geographical
  - many approaches

Why Spatial Weights

Identification Problem

- Spatial covariance $\sigma_{i,j}$
- $n(n-1)/2$ parameters
- We only have n observations in a cross section

Incidental Parameter Problem

- Number of parameters increases $O(n^2)$
- More data (in the cross-section) is not the solution

Weights as a solution

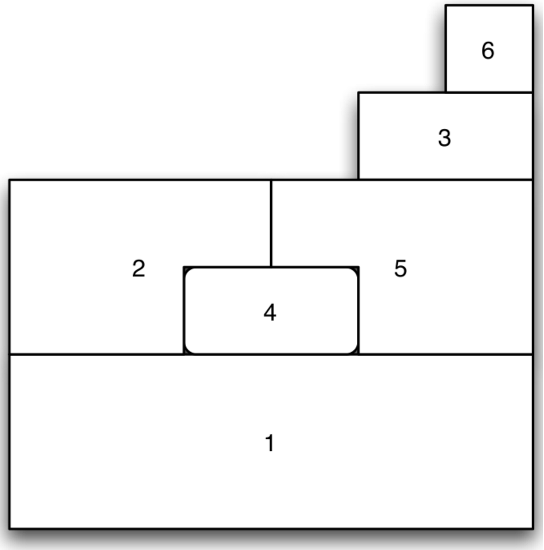
Impose structure on the problem

- set some interactions to zero
- only let “neighbors” interact directly
- constrain the number of neighbors

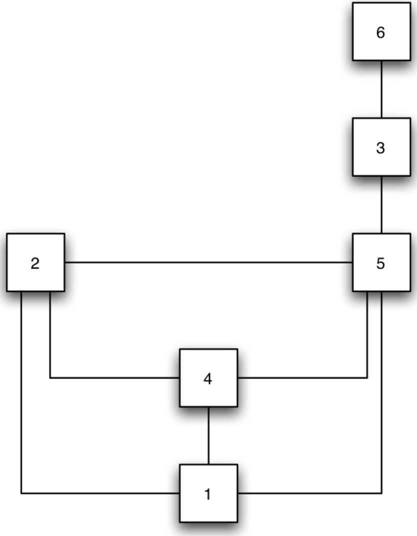
Assume a single parameter

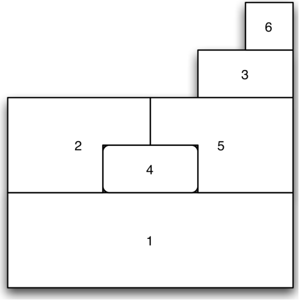
- Spatial autocorrelation coefficient

Irregular Lattice (Polygons)

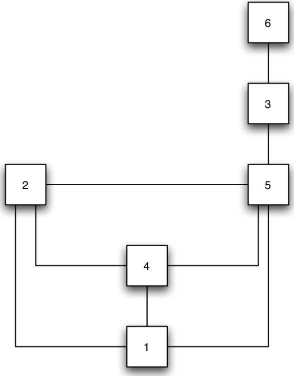


Neighbor Structure as a Graph





Map



Graph

Outline

Spatial Weights Matrix

Definition

- $n \times n$ positive matrix W, elements: $w_{i,j}$
- $w_{i,j} \neq 0$ for neighbor pairs, 0 otherwise
- $w_{i,i} = 0$, no self-neighbors

$$W = \begin{bmatrix} w_{1,1} & w_{1,2} & \dots & w_{1,n} \\ w_{2,1} & w_{2,2} & \dots & w_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ w_{1,1} & w_{1,2} & \dots & w_{1,n} \end{bmatrix} \tag{1}$$

Binary Contiguity

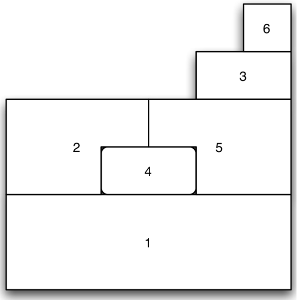
Binary

- $w_{i,j} = 1 \textit{ or } 0$

Contiguity

- $w_{i,j} = 1$ if i and j are contiguous
- Share an edge or vertex
- $w_{i,j} = 0$ if i and j are noncontiguous

Binary Contiguity Example



Map

$$W = \begin{bmatrix} 0 & 1 & 0 & 1 & 1 & 0 \\ 1 & 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 \\ 1 & 1 & 0 & 0 & 1 & 0 \\ 1 & 1 & 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \end{bmatrix}$$

How to Define W

- Geographic Weights
  - contiguity
  - distance
  - general
  - graph-based
- Socioeconomic Weights

Outline

Contiguity Weights

- Contiguity
  - sharing a common boundary of non-zero length
- What is a non-zero boundary?
- Three Views of Contiguity
  - rook
  - queen
  - bishop

Example: Regular Lattice ($n = 9$)

1	2	3
4	5	6
7	8	9

Rook Contiguity

1	2	3
4	5	6
7	8	9

Map

$$W = \begin{bmatrix} 0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 1 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 1 & 0 \end{bmatrix}$$

Neighbors to center cell 5

Bishop Contiguity

1	2	3
4	5	6
7	8	9

Map

$$W = \begin{bmatrix} 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 1 & 0 \end{bmatrix}$$

Neighbors to center cell 5

Queen Contiguity

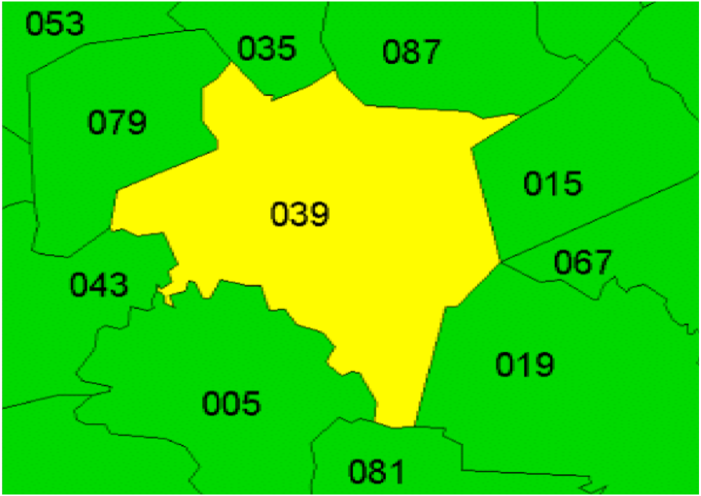
1	2	3
4	5	6
7	8	9

Map

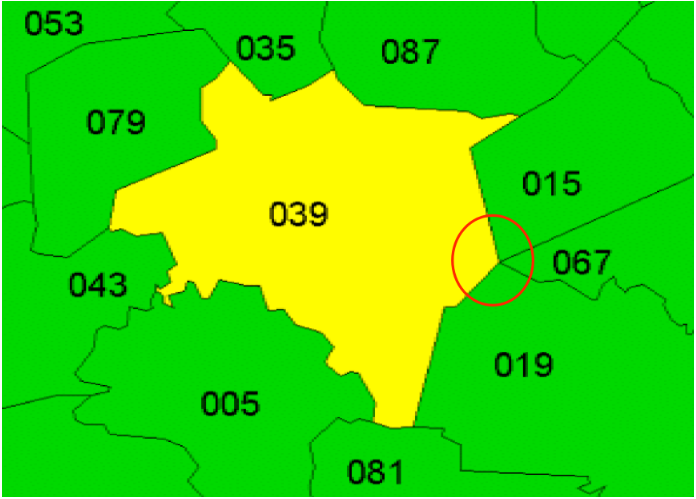
$$W = \begin{bmatrix} 0 & 1 & 0 & 1 & 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 1 & 1 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 & 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 1 & 0 & 1 & 1 & 0 \\ 1 & 1 & 1 & 1 & 0 & 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 0 & 1 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 & 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 1 & 1 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 & 1 & 0 & 1 & 0 \end{bmatrix}$$

Neighbors to center cell 5

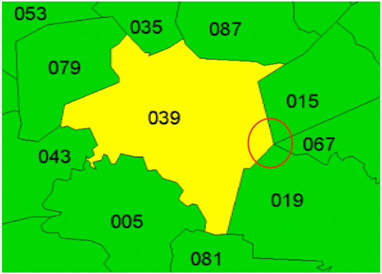
Irregular Lattice Contiguity



Irregular Lattice Contiguity

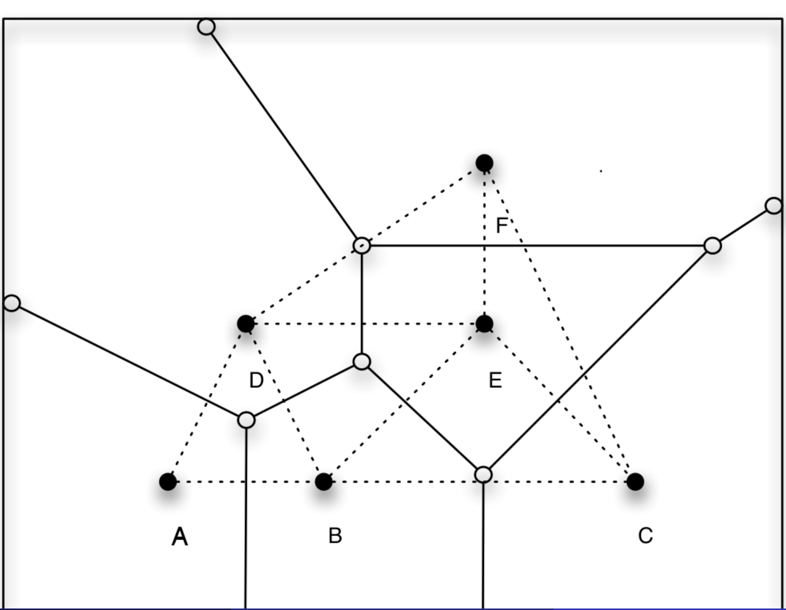


Irregular Lattice Contiguity

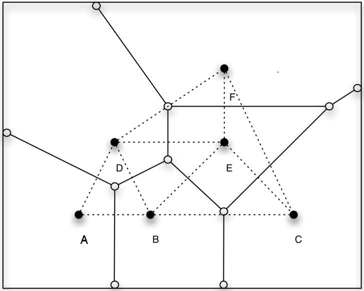


- Rook: 039 is not a neighbor of 067
- Bishop: 039 is a neighbor of 067
- Queen: 039 is a neighbor of 067

Point Contiguity Weights



Point Contiguity Weights



$$W = \begin{bmatrix} 0 & 1 & 0 & 1 & 0 & 0 \\ 1 & 0 & 1 & 1 & 1 & 0 \\ 0 & 1 & 0 & 0 & 1 & 1 \\ 1 & 1 & 0 & 0 & 1 & 1 \\ 0 & 1 & 1 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 & 1 & 0 \end{bmatrix}$$

Outline

Lattice Distances

Distance Measures

- Point: (x_i, y_i)
- Interpoint Distances
- Metric
  - Euclidean: $d^e_{i,j} = \sqrt{(x_i - x_j)^2 + (y_i - y_j)^2}$
  - Manhattan: $d^m_{i,j} = |x_i - x_y| + |y_i - y_j|$
  - Minkowski: $d^p_{i,j} = \left(|x_i - x_y|^p + |y_i - y_j|^p\right)^{(1/p)}$

Distance Measures

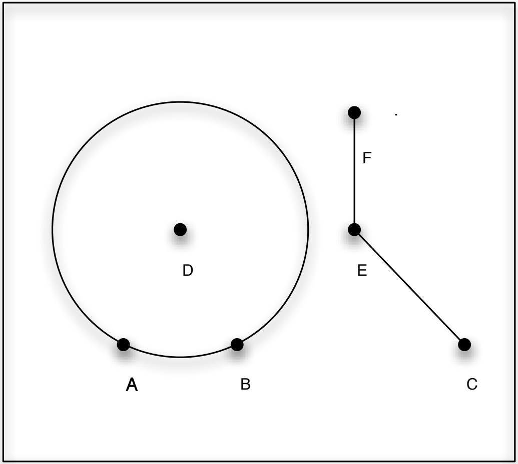
Other

- Road network
- Actual travel time

Unprojected coordinates

- Straight line distance measure inappropriate
- Use great circle distance

Interpoint Distance



The coordinates of the points are (in arbitrary units): A (10, 10), B (20, 10), C (40, 10), D (15, 20), E (30, 20), and F (30, 30).

Interpoint Distance

Table: Interpoint Euclidean Distance for Point Layout

	B	C	D	E	F
A	10.0	30.0	11.2	22.4	28.3
B		20.0	11.2	14.1	22.4
C			26.9	14.1	22.4
D				15.0	18.0
E					10.0

Distance Bands

- $w_{i,j} = 1$ when $d_{i,j} \leq \delta$
- $w_{i,j} = 0$ if $d_{i,j} > \delta$
- δ is a preset critical distance cutoff

Distance Bands

- $w_{i,j} = 1$ when $d_{i,j} \leq \delta$
- $w_{i,j} = 0$ if $d_{i,j} > \delta$
- $\delta = 11.2$

$$\mathbf{W} = \begin{bmatrix} 0 & 1 & 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 & 0 \end{bmatrix}$$

Distance Bands

- $w_{i,j} = 1$ when $d_{i,j} \leq \delta$
- $w_{i,j} = 0$ if $d_{i,j} > \delta$
- $\delta = 11.2$

$$\mathbf{W} = \begin{bmatrix} 0 & 1 & 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 & 0 \end{bmatrix}$$

Observation C is an *island* or *isolate*.

Solution to Islands

Maximum Nearest Neighbor Distance

10 (A-B), 10 (B-A), 14.1 (C-E), 11.2 (D-A and D-B), 10 (E-F), and 10 (F-E)

Threshold δ

- $\delta = MAX(nnd)$
- $\delta = 14.1$

Distance Bands

- $w_{i,j} = 1$ when $d_{i,j} \leq \delta$
- $w_{i,j} = 0$ if $d_{i,j} > \delta$
- $\delta = 14.1$

$$\mathbf{W} = \begin{bmatrix} 0 & 1 & 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 & 0 \end{bmatrix}$$

Max nn-distance Band Weights

Properties

- Symmetric Matrix
- Avoids Islands

Problems

- Can create too many neighbors for clustered locations
- Variance in the connectivity cardinalities
- Driven by maximum nn-distance

K Nearest Neighbors Weights

Properties

- Avoids islands
- Same number of neighbors for all locations
- e.g., $k = 3$

$$\mathbf{W} = \begin{bmatrix} 0 & 1 & 0 & 1 & 1 & 0 \\ 1 & 0 & 0 & 1 & 1 & 0 \\ 0 & 1 & 0 & 0 & 1 & 1 \\ 1 & 1 & 0 & 0 & 1 & 0 \\ 0 & 1 & 1 & 0 & 0 & 1 \\ 0 & 1 & 1 & 1 & 1 & 0 \end{bmatrix}$$

K Nearest Neighbors Weights

Issues

- How to deal with ties?
- Not necessarily symmetric
  - Only symmetric if nn graph is identical to the mutual nn graph
  - No ties broken randomly
- Problems for certain statistical methods

Ties in Knn Weights: $d_{F,B} = d_{F,C} = 22.4$

Table: Interpoint Euclidean Distance for Point Layout

	B	C	D	E	F
A	10.0	30.0	11.2	22.4	28.3
B		20.0	11.2	14.1	22.4
C			26.9	14.1	22.4
D				15.0	18.0
E					10.0

$$\mathbf{W} = \begin{bmatrix} 0 & 1 & 0 & 1 & 1 & 0 \\ 1 & 0 & 0 & 1 & 1 & 0 \\ 0 & 1 & 0 & 0 & 1 & 1 \\ 1 & 1 & 0 & 0 & 1 & 0 \\ 0 & 1 & 1 & 0 & 0 & 1 \\ 0 & 1 & 1 & 1 & 1 & 0 \end{bmatrix}$$

Ties K Nearest Neighbors Weights

$$\mathbf{W} = \begin{bmatrix} 0 & 1 & 0 & 1 & 1 & 0 \\ 1 & 0 & 0 & 1 & 1 & 0 \\ 0 & 1 & 0 & 0 & 1 & 1 \\ 1 & 1 & 0 & 0 & 1 & 0 \\ 0 & 1 & 1 & 0 & 0 & 1 \\ 0 & 1 & 1 & 1 & 1 & 0 \end{bmatrix}$$

- Not symmetric
- B is a 3nn to F
- F is not a 3nn to B

Outline

Spatial Lag

$$yl_i = \sum_j w^*_{i,j} y_j$$

i	y_i	yl_i
1	20	?
2	10	?
3	40	?
4	22	?
5	30	?
6	50	?

Standardization

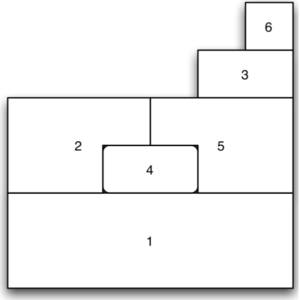
Row Standardization of W

$$w_{i,j}^* = \frac{w_{i,j}}{\sum_j w_{i,j}} \tag{2}$$

Properties

1. $0 \leq w_{i,j}^* \leq 1$
2. $\sum_j w_{i,j}^* = 1 \; \forall i$
3. $\sum_i \sum_j w_{i,j}^* = n$

Example of Row Standardization



Map

$$W = \begin{bmatrix} 0 & 1 & 0 & 1 & 1 & 0 \\ 1 & 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 \\ 1 & 1 & 0 & 0 & 1 & 0 \\ 1 & 1 & 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \end{bmatrix}$$

$$W^* = \begin{bmatrix} 0 & 1/3 & 0 & 1/3 & 1/3 & 0 \\ 1/3 & 0 & 0 & 1/3 & 1/3 & 0 \\ 0 & 0 & 0 & 0 & 1/2 & 1/2 \\ 1/3 & 1/3 & 0 & 0 & 1/3 & 0 \\ 1/4 & 1/4 & 1/4 & 1/4 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \end{bmatrix}$$

Outline

Spatial Lag

$$yl_i = \sum_j w^*_{i,j} y_j$$

i	y_i	yl_i
1	20	?
2	10	?
3	40	?
4	22	?
5	30	?
6	50	?

Spatial Lag

i	y_i	yl_i
1	20	20.67
2	10	24.00
3	40	40.00
4	22	20.00
5	30	23.00
6	50	40.00

$$W^* = \begin{bmatrix} 0 & 1/3 & 0 & 1/3 & 1/3 & 0 \\ 1/3 & 0 & 0 & 1/3 & 1/3 & 0 \\ 0 & 0 & 0 & 0 & 1/2 & 1/2 \\ 1/3 & 1/3 & 0 & 0 & 1/3 & 0 \\ 1/4 & 1/4 & 1/4 & 1/4 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \end{bmatrix}$$

$$\begin{aligned} yl_2 &= \sum_{j=1} w^*_{2,j} y_j \\ &= w^*_{2,1}y_1 + w^*_{2,2}y_2 + w^*_{2,3}y_3 + w^*_{2,4}y_4 + w^*_{2,5}y_5 + w^*_{2,6}y_6 \\ &= 1/3(20) + 0(10) + 0(40) + 1/3(22) + 1/3(30) + 0(50) \\ &= 1/3(20 + 22 + 30) \\ &= 24.00 \end{aligned}$$