GeoDa Center
FOR GEOSPATIAL ANALYSIS
AND COMPUTATION

ARIZONA STATE UNIVERSITY

Spatial Autocorrelation

Soe W. Myint

2010

Working Paper Number 27

Spatial Autocorrelation

Spatial autocorrelation is a spatial measure that evaluates how dispersed or clustered points are distributed in space and whether or not this distribution has occurred by chance. Unlike other spatial methods for detecting spatial patterns of point distribution without considering their attributes such as Quadrat analysis and Nearest Neighbor method, spatial autocorrelation also characterizes how similar these points are with respect to their attribute values. There are two typical spatial autocorrelation statistics: Moran's *I* and Geary's *C*. Both Moran's *I* and Geary's *C* measure the proximity of locations and evaluate the similarity of the attributes.

If the statistics show more positive than it can be expected from a randomly distributed pattern, points with similar attribute values are closely distributed in space (Figure 1a) whereas negative spatial autocorrelation statistics indicate that closely associated points are more dissimilar. A chess board is an example of the negative spatial autocorrelation (Figure 1b); every black cell is adjacent to white squares, indicating that the neighbors are not similar. Careful attention should be given when using spatial autocorrelation as it is related to the scale of the data.

FIGURE 1 here.

Figure 1. Hypothetical images with positive spatial autocorrelation (a) and negative spatial autocorrelation (b).

Moran's *I* can be computed as

$$I(d) = \frac{\sum_{i} \sum_{j} {}_{ij} \, w_{ij}(d)(z_i - \bar{z})(z_j - \bar{z})}{s^2 \sum_{i} \sum_{j} w_{ij}(d)}$$

where W_{ij} is the weight at distance d so that $W_{ij} = 1$ if point j is within distance d from point i; otherwise, $W_{ij} = 0$; $\bar{z}$ is the mean attribute value; s^2 is the variance of the attribute values and can be calculated as: $s^2 = \sum_i (z_i - \bar{z})^2 / n$. Moran's I varies from +1.0 for perfect positive correlation to –1.0 for perfect negative correlation.

Geary's C is calculated from the following:

$$C(d) = \frac{\sum_i^n \sum_j^n w_{ij}(d)(z_i - z_j)^2}{2\sum_i^n \sum_j^n w_{ij}(d)s^2}$$

where w_{ij} is the weight; s^2 is the variance of z values and can be computed as: $s^2 = \sum_i (z_i - \bar{z})^2 /(n-1)$. The weight w_{ij} only considers the values 0 and 1 as mentioned above. A typical Geary's C value ranges between 0 and 2. Where a value of less than 1.0 implies positive autocorrelation, 1.0 implies no autocorrelation, and values of greater than 1.0 imply negative autocorrelation.

Similarly, another spatial autocorrelation approach called the Getis statistic (*Gi*) was developed to characterize the presence of hot spots (Figure 2a) and cold spots (Figure 2b), which are a set of clustered points with large or small attribute values respectively.

FIGURE 2 here.

Figure 2. Hypothetical images with hot spot (a) and cold spot (b).

The Getis statistic is computed as:

$$Gi(d) = \frac{\sum_{i}^{n}\sum_{j}^{n} w_{ij}(d) z_i z_j}{\sum_{i}^{n}\sum_{j}^{n} z_i z_j}, \text{ for } i \neq j.$$

The *Gi* is also defined by a distance, *d*, within which areal units can be regarded as neighbors of *i*. The weight w_{ij} (*d*), is 1 if areal unit *j* is within *d* and is 0 otherwise. Some of the z_i, z_j pairs will not be considered in the calculation of numerator if *i* and *j* are more than *d* away from each other.

Soe W. Myint

Further Readings

Getis, A. & J.K. Ord (1992). The analysis of spatial association by use of distance statistics. *Geographical Analysis 24(3)*:1269-1277.

Goodchild, M.F. (1988). *Spatial Autocorrelation. Concepts and Techniques in Modern Geography* (CATMOG) Number 47. Norwich: Geobooks.

Wong, D. W. S. & J. Lee. (2005). *Statistical Analysis and Modeling of Geographic Information with ArcView GIS and ArcGIS*, Hoboken, New Jersey, John Wiley & Sons.

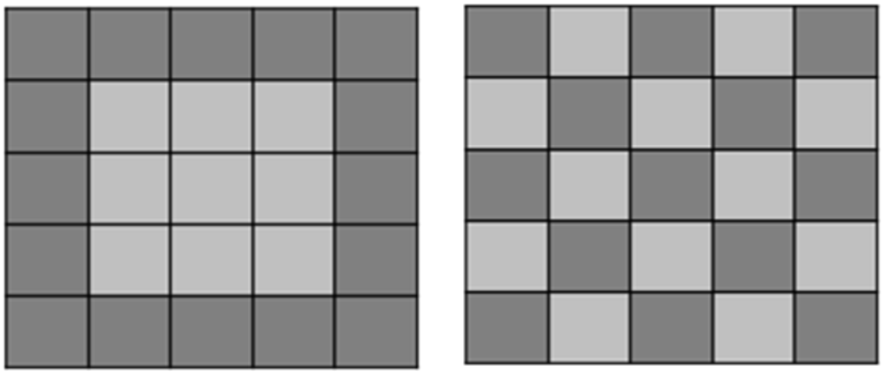

Figure 1 (a) Figure 1 (b)

31	34	32	30	32
30	997	998	997	34
32	998	997	998	31
33	999	997	997	33
31	35	31	33	31

Figure 2 (a)

31	34	32	30	32
30	1	2	1	34
32	2	1	2	31
33	3	1	1	33
31	35	31	33	31

Figure 2 (b)