GEOG/Math574 Introduction to Geostatistics

Spatial Stochastic Process

Daoqin Tong

School of Geography & Development
408 Harvill Building
Email: daoqin@email.arizona.edu

Potentials of spatial data

- Distance
- Adjacency
- Interaction
- Neighborhood

Distance

- Distance can be calculated with spatial data

$$d_{ij} = \sqrt{(x_i - x_j)^2 + (y_i - y_j)^2}$$

  - Euclidean distance
  - Network distance
  - Other distance measures (travel time)

Adjacency

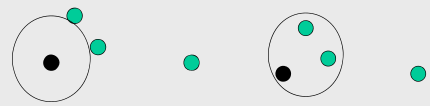
- Two spatial entities are either adjacent or not.
- Define adjacency
  - Two entities are adjacent if they share a common boundary (e.g. California and Oregon)
  - Two entities are adjacent if they are within a specified distance
- Important for network analysis, spatial autocorrelation, and spatial interpolation

Interaction

- Refinement of distance and adjacency
- Tries to quantify the strength of some relationship between objects
  - Inverse distance $\omega_{ij} \propto \frac{1}{d^k}$
    - ⇒ IDW
  - Weighed inverse distance $\omega_{ij} \propto \frac{p_i p_j}{d^k}$
    - ⇒ Spatial interaction model

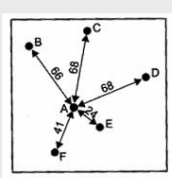
Neighborhood

- Define a neighborhood
  - It could be a region maybe around some object
  - It could be a collection of objects considered to be neighbors of some object



This is where matrix becomes useful …

- Distance, adjacency, interaction, neighborhood uses some value to describe a spatial relationship between two objects
- A powerful way to record this is in a matrix

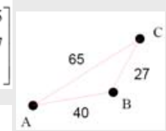


$$\begin{array}{c|c} & \begin{array}{cccccc} \mathbf{A} & \mathbf{B} & \mathbf{C} & \mathbf{D} & \mathbf{E} & \mathbf{F} \end{array} \\ \begin{array}{c} \mathbf{A} \\ \mathbf{B} \\ \mathbf{C} \\ \mathbf{D} \\ \mathbf{E} \\ \mathbf{F} \end{array} & \begin{bmatrix} 0 & 66 & 68 & 68 & 24 & 41 \\ 66 & 0 & 51 & 110 & 99 & 101 \\ 68 & 51 & 0 & 67 & 91 & 116 \\ 68 & 110 & 67 & 0 & 60 & 108 \\ 24 & 99 & 91 & 60 & 0 & 45 \\ 41 & 101 & 116 & 108 & 45 & 0 \end{bmatrix} \end{array}$$

Different matrices used

- Distance matrix
  - Symmetric

$$\mathbf{D} = \begin{bmatrix} 0 & 40 & 65 \\ 40 & 0 & 27 \\ 65 & 27 & 0 \end{bmatrix}$$



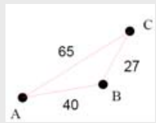
- Adjacency matrix
  - Depending on adjacency rule not always symmetric

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

- Interaction (weights) matrix
  - Using some inverse distance rule (1/d)
  - Frequently used for autocorrelation, point pattern analysis, interpolation

$$\mathbf{W} = \begin{bmatrix} \infty & 0.025 & 0.015 \\ 0.025 & \infty & 0.037 \\ 0.015 & 0.037 & \infty \end{bmatrix}$$

Which city has the greatest interaction with the region?



$$
\begin{array}{c} \\ W = \begin{bmatrix} \infty & 0.025 & 0.015 \\ 0.025 & \infty & 0.037 \\ 0.015 & 0.037 & \infty \end{bmatrix} \end{array}
\begin{array}{l} \textbf{•row totals} \\ \textbf{•0.04} \\ \textbf{•0.062} \\ \textbf{•0.052} \end{array}
$$

$$
W = \begin{bmatrix} \infty & 0.625 & 0.375 \\ 0.403 & \infty & 0.597 \\ 0.288 & 0.712 & \infty \end{bmatrix}
$$

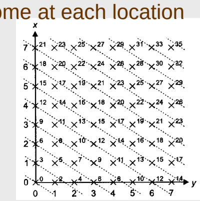
column totals **0.691** **1.337** **0.972**

Spatial process

- **Spatial process**
  - Spatial process is a description of how a spatial pattern might be generated.
- **Deterministic process**
  - It always produces same outcome at each location

Z = 2x + 3y

where x and y are two spatial coordinates
z is the numerical value for a variable

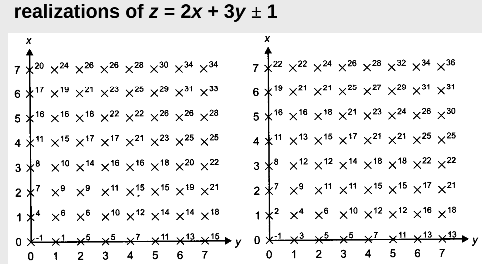


Spatial stochastic process

- **Stochastic process**
  - A process which is subject to uncertainty, also called random process
  - Governed by the laws of probability
    - ⇒Random variables
    - ⇒Probability distribution
    - ⇒Time series
    - ⇒White noise: random components are independent of each other
- **Spatial stochastic process**
  - The outcome at each location is subject to uncertainty
  - A spatial pattern is one of the realizations of the underlying stochastic process

Pattern is a realization of process…

For a spatial stochastic process z=2x + 3y ±1, there are different realizations. The following are two realizations of *z* = 2*x* + 3*y* ± 1



Complete Spatial Randomness (CSR)

- This is probably the most commonly used ‘standard’ process
  - It’s also called the *independent random process* (IRP)
- Formally, CSR postulates two conditions
  - Equal probability: an event has equal probability of occurring anywhere in A
  - Independent: event locations are independent
- CSR is not realistic, otherwise geography would have little meaning and most GIS operations would be pointless

Specifying Stochastic Process

- Simple cases
  - Theoretic knowledge
  - Example, throw of a fair die, flip a coin
- Real situation
  - Complex
  - Experiment
  - Observation
  - Data-> Model

Spatial Stochastic Process

- Spatial stochastic process
  - Statistical model which specifies a probability distribution for the random variable (or variables) representing a spatially referenced stochastic phenomenon

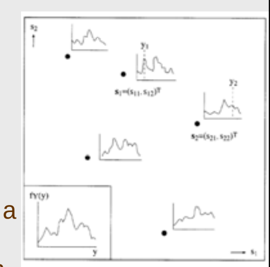


Fig. 1.8 Illustration of a spatial stochastic process

Spatial Stochastic Process

- Spatial stochastic process is typically represented by the joint distribution of a set of (non-independent) random variables Y indexed by location vector $s=(s_1,s_2)^T$ on study region R or by sub-region A of R

  $\{Y(s), s \in \Re\} \quad or \quad \{Y(A), A \subseteq \Re\}$
- Theoretically, to fully specify a spatial stochastic process, we need to model the joint distributions of every possible configurations of locations or sub-regions

Specifying Spatial Stochastic Process

- In reality, we only have one realization from the joint distribution observed, which does not give much information about the distribution
- We have to make some "reasonable" assumptions about the nature of the spatial phenomena so that the model can be specified using both the observed data and the assumptions
- By doing so, we obtain a general mathematical form of the probability distribution with some parameters whose values are left unspecified

Specifying Spatial Stochastic Process

- The general form is refined, or fitted to the observed data, i.e. the values of unknown parameters are estimated from observed data
- Then, fitted model can be evaluated -> modified assumptions -> adjusted or refined model and so on

An Example

- We make the following assumptions
  - The random variables $\{Y(s), s \in \Re\}$ are independent
  - Each *Y(s)* has a normal distribution with constant variance σ^2 and mean values as a linear combination of location

$$E(Y(s)) = \beta_0 + \beta_1 s_1 + \beta_2 s_2$$

- Then we have the probability model for each *Y(s)*

$$Y(s) \sim N(\beta_0 + \beta_1 s_1 + \beta_2 s_2, \sigma^2)$$

- The independence assumption implies that the joint probability is just the product of the distributions at each site. Model specification reduces to the estimation of unknown parameters from the data

Likelihood Function

- Given n observed data

$$(y(s_1), y(s_2), \ldots, y(s_n)) \text{ or simply as } (y_1, y_2, \ldots, y_n)$$

- Under the previous assumptions, the joint distribution is

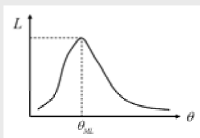
$$f(y_1, y_2, \ldots, y_n; \theta) = \prod_{i=1}^{n} \left(\frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(y_i - \beta_0 - \beta_1 s_{i,1} - \beta_2 s_{i,2})^2}{2\sigma^2}} \right)$$

- $f(y_1, y_2, \ldots, y_n; \theta)$ represents the probability (probability density) of observing the data under the proposed model (or given the unknown parameters θ)
- We call it the likelihood of the data and denote it as

$$L(y_1, y_2, \ldots, y_n; \theta)$$

Maximum Likelihood Estimation

- ML solution $\theta_{ML} = \arg\{\max_{\theta} L(y_1, y_2, \dots, y_n; \theta)\}$

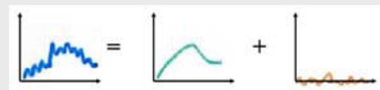


- Log likelihood

$$l(y_1, y_2, \dots, y_n; \theta) = \log[L(y_1, y_2, \dots, y_n; \theta)]$$

Structure of Spatial Phenomena

- Spatial process can be decomposed into different components as:



Spatial data = large-scale variation + small-scale variation

Spatial data = the first order effect + the second order effect

First Oder Effect

- Variation of the mean value in space, also called global or large scale trend
- Often resulting from large scale trends in space
- The trend component measures spatial heterogeneity
- Modeling usually employs conventional statistical methods
- Covariates are often used to estimate the localized mean

Second Order Effect

- Correlation in the deviation of values of the process from the mean
- Resulting from spatial autocorrelation structure-local or small scale effects; this is the core of spatial statistics
- Often deviations from mean ‘follow’ each other in neighboring sites, positive effects, but they may also be negative effects due to competition

Some Remarks

- The distribution of first and second order effects is artifacts of the modeler and not the reality
- Often it depends on the scale and purpose of the study, which effects should modeled as first and which as second order
- The combined first and second order effects violate the independence assumption in conventional statistics
- An alternative assumption that incorporates covariance structure to accommodate the second order effects is needed

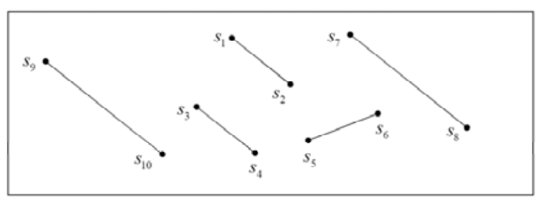
Stationarity

- Spatial data usually represent a single realization of a random process, so some degree of stationarity must be assumed to make inferences about the data
- The second order effect is often modeled as a stationary spatial process
- Stationarity is the quality of a process in which the statistical parameters (mean and variance) of the process do not change with space or time

Stationarity

- A spatial process $\{Y(s), s \in \Re\}$ is said to be stationary if its statistical properties are independent of the absolute location in $\Re$
- Specifically, stationarity is defined as
  - The mean *E(Y(s))* and the variance Var(Y(s)) are constant in $\Re$ and do not depend on location *s*
  - The covariance Cov($Y(s_i)$,$Y(s_j)$) between any two locations s_i and s_j depends only on the relative location of the two sites in distance and direction (i.e. the difference vector $h=s_i-s_j$) and not the absolution location in $\Re$

Stationarity



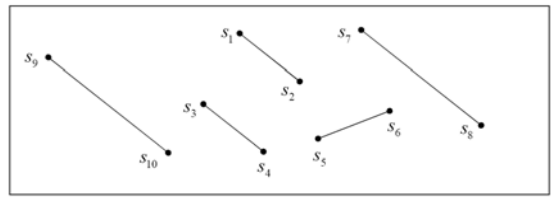
$$E(Y(s_1)) = E(Y(s_2)) = \cdots = E(Y(s_{10}))$$

$$s_1 - s_2 = s_3 - s_4 \Rightarrow Cov(Y(s_1), Y(s_2)) = Cov(Y(s_3), Y(s_4))$$
$$s_1 - s_2 \neq s_5 - s_6 \Rightarrow Cov(Y(s_1), Y(s_2)) \neq Cov(Y(s_5), Y(s_6))$$
$$s_7 - s_8 = s_9 - s_{10} \Rightarrow Cov(Y(s_7), Y(s_8)) = Cov(Y(s_9), Y(s_{10}))$$

Isotropic Process

- Related to the properties of a spatial process with respect to different directions
- A process is called isotropic if in addition to stationarity, the covariance depends only on distance between locations and not the direction in which they are separated
- A process is called anisotropic if the correlation and covariance differs with direction
- Most spatial statistical methods assume spatial correlation is isotropic

Isotropy



$$E(Y(s_1)) = E(Y(s_2)) = \cdots = E(Y(s_{10}))$$

$$|s_1 - s_2| = |s_3 - s_4| = |s_5 - s_6| \Rightarrow$$
$$Cov(Y(s_1), Y(s_2)) = Cov(Y(s_3), Y(s_4)) = Cov(Y(s_5), Y(s_6))$$

Non-stationarity

- Non-stationarity or heterogeneity refers to “drifts” in statistical properties of a spatial process over
  - Mean, variance, covariance
- Non-stationarity in covariance makes the model complex and hard to estimate parameters
- General strategy is to assume heterogeneity in the mean and stationarity in the second order effects