Chapter 4
Spatial preprocessing

4. Preprocessing

4.1 Introduction

Once we acquire spatial data, we think that we can start our study immediately. Unfortunately, this is not always the case.

Why?

Data format, georeferencing system, map projection, data resolution, date of data acquisition, and spatial data unit used for reporting attribute data are different among spatial data.

Preprocessing is a computational operation that helps us to use various spatial datasets together in our analysis. It includes conversion of data format, conversion of georeferencing system, and spatial interpolation.

If you are interested in a spatial phenomenon at a local scale, Cartesian coordinate systems are better than longitude-latitude system.

If the data are obtained only at sample points, you may have to interpolate them.

If you want to know the population in a small region and the data are aggregated by prefectures, you have to estimate the population from the data.

The problems discussed in this chapter may seem trivial and may not be so interesting.

If you use only one database, you may not be troubled with such problems.

However, when you use several datasets simultaneously, overlaying them and performing calculations among datasets, you will face those intricate but important problems.

Spatial preprocessing

1. Conversion of data format
2. Conversion of georeferencing system
3. Map transformation (including conversion of map projection)
4. Spatial interpolation
5. Spatial smoothing
6. Raster-vector conversion
7. Areal interpolation
8. Spatial data fusion
4. Preprocessing

4.2 Conversion of data format

1. ArcView (ESRI) shape
2. ArcInfo (ESRI) coverage
3. Mapinfo (Mapinfo) MIF
4. SIS (Informatix) SIS
5. AtlasGIS (?)
6. TIGER (USGS)
7. KIWI (DD)
8. DXF, DIL, PIX
9. JPEG
...

You can convert the data format by writing a program in C or Pascal if the data format is open.

However, you had better use existing GIS software because they can at least read various data formats; in some cases they can even save them in different formats.

For example, ArcView can read spatial data in shape, coverage, MIF, ...

4.3 Conversion of georeferencing system

1. Address system
2. Longitude-latitude system
3. Cartesian coordinate system

Today’s GIS software has a conversion program as an internal function. You can convert georeferencing system by using GIS; you do not have to write a computer program.

Longitude-latitude system ↔ Cartesian coordinate system
Address system ↔ Longitude-latitude system

Geocoding

4.4 Map transformation

1. Map projection
2. Affine transformation
3. Rubber sheeting

4.4.1 Map projections
4. Preprocessing

- References – map projection


There are numerous methods of map projection. Map projection is a process of projecting the earth’s surface on a plane. Therefore, map projections are classified by the location and direction of source light and the location and shape of projection plane.

**orthographic**

**stereographic**

**gnomonic**

Figure: Location and direction of source light

**Azimuthal projection**  **Conic projection**  **Cylindrical projection**

Figure: Location and shape of projection plane

**Mercator’s projection**

Source light: gnomonic projection  
Projection plane: (modified) cylindrical projection

As well as data format, map projection can be converted easily by today’s GIS software.

Mercator’s projection  ↔  Mollweide’s projection  ↔  Goode’s projection
4. Preprocessing

4.4.2 Affine transformation

Suppose a map whose projection or georeferencing system is unknown. If we want to convert it into digital data, we have to fit it to another map whose projection and georeferencing system are known.

To do this we scan the maps and specify spatial objects recorded on both maps in GIS. After that, we transform the former map to fit the latter one.

4.4.3 Rubber sheeting

In some cases it is enough to use Affine transformation to fit a map to another map.

However, if a map is seriously distorted, it is necessary to apply more flexible transformation to obtain good result.

The ‘rubber sheeting’ operation permits flexible transformation of spatial data. It is often used for treating sketch maps, historical maps and cognitive maps in GIS.

We perform rubber sheeting operation by only giving a set of corresponding point pairs.

GIS system then calculates transformation function, not limited to Affine transformation, from the coordinates of the points and apply it to one of the maps.
4. Preprocessing

4.5 Spatial interpolation

Spatial interpolation is a mathematical process of estimating unknown value of a (usually scalar) function at a given point from a set of known values at sample points.

Basic idea of spatial interpolation

Suppose two points and their function values.

If they are closely located, their values are expected to be close.

If they are distant, it is probable that their values are quite different.

4.5.1 Discrete interpolation:

Nearest neighborhood method

This method substitutes the function value of a point by that of its nearest sample point.

To find the nearest sample point from a certain point, we use the Voronoi diagram defined for a set of points.

The Voronoi diagram shows the assignment of points to their nearest sample points, so it gives a spatial tessellation.
4. Preprocessing

4.5.2 One-dimensional continuous interpolation

Though discrete interpolation is theoretically natural and easy to perform, the result is not satisfactory because it yields discontinuity in scalar functions.

It is obviously better to interpolate scalar functions by using continuous functions.

In usual, spatial data are two-dimensional, so interpolation of spatial data requires two-dimensional interpolation methods.

However, two-dimensional interpolation methods are more complicated and difficult to understand than one-dimensional methods. Because of this, before going to two-dimensional case, one-dimensional interpolation methods will be explained.

- References - spline interpolation

4. Preprocessing

**Terminology**

**Sample points**

Sample points are the locations where the value of a scalar function is known.

The elevation is measured at sample points and interpolated to generate the earth’s surface.

**Interpolation region**

Interpolation region is the region where a scalar function is estimated from known values at sample points. This term is not widely used in GIS, but convenient for explaining spatial interpolation.

In usual, interpolation region is bounded by sample points. However, it can also be extended to the area of no sample points. In this case we use the term ‘extrapolation’ instead of interpolation.

**Interpolation region is divided into subregions.**

In one-dimensional case, their boundary points are called knots. In two-dimensional case, they are simply called boundaries.
Knots can be determined arbitrarily. However, in usual, sample points are used as knots.

Outline of the methods

Unlike discrete interpolation, continuous interpolation always gives smooth a function whose value at knots is continuous between subregions.

In each subregion, a continuous function is fitted which is estimated from the function value at sample points.

In usual, polynomial function called spline function is used because of its tractability.

Spline function

$m$: the number of subregions
   ... the number of knots is $m-1$
   ... the number of sample points is $m+1$

$x_i$: the coordinate of the $i$th sample point ($i=0, ..., m$)
$h_i$: the measured value at the $i$th sample point ($i=0, ..., m$)

$n$: the degree of spline function

Spline function for the $i$th ($i=1, ..., m$) subregion:

$$f_i(x) = a_{0i} + a_{1i}x + a_{2i}x^2 + \cdots + a_{ni}x^n$$

We have $m(n+1)$ unknown parameters ($a_{ij}$) to be estimated.
How do we estimate parameter values?

We put two conditions satisfied by spline functions, in order to obtain a continuous scalar continuous.

### Condition C1

At each sample point the value of spline function agrees with the known (measured) value.

\[ f_i(x_{i-1}) = h_{i-1} \]
\[ f_i(x_i) = h_i \quad (i = 1, \ldots, m) \]

\[ f(x) \quad f(x) \quad f(x) \quad f(x) \quad f(x) \]
\[ x_0 \quad x_1 \quad x_2 \quad x_3 \quad \ldots \quad x_m \]

### Condition C2

At each knot two spline functions have the same 1st, 2nd, \ldots, \( n-1 \)th derivative values.

\[ \frac{d^i}{dx^i} f_i(x_i) = \frac{d^i}{dx^i} f_{i+1}(x_i) \quad (i = 1, \ldots, m-1; \quad l = 1, \ldots, n-1) \]

\[ f(x) \quad f(x) \quad f(x) \quad f(x) \quad f(x) \]
\[ x_0 \quad x_1 \quad x_2 \quad x_3 \quad \ldots \quad x_m \]

Therefore, we have

\[ m(n+1) - 2m - (m-1)(n-1) = n - 1 \]

free parameters (degree of freedom).

This indicates that if \( n = 1 \) the spline functions are unique. If \( n > 1 \), additional conditions are necessary to determine spline functions.

### Linear spline

Linear spline uses linear functions where \( n = 1 \).

\[ f_i(x) = a_i + a_i x \]

Degree of freedom: 0

A set of line segments that simply connect the known values of the function at sample points.
4. Preprocessing

**Quadratic spline**

Quadratic spline uses quadratic functions \((n=2)\).

\[ f(x) = a_0 + a_1 x + a_2 x^2 \]

Degree of freedom: 1

To determine spline functions, we either

1. give the first derivative at one end of the interpolation region, or,

2. estimate the spline functions that minimize their total length.

**Cubic spline**

Cubic spline uses quadratic functions \((n=3)\).

\[ f(x) = a_0 + a_1 x + a_2 x^2 + a_3 x^3 \]

Degree of freedom: 2

To determine spline functions, we usually give the first derivative at both ends of the interpolation region.

**Cubic spline** is the most popular among polynomial spline functions, because

1. it is well balanced. Functions interpolated by linear or quadratic spline function are not so smooth. On the other hand, polynomial functions of higher degree are too sensitive to the variation of known values at sample points, that is, they often yield functions that wildly fluctuate between sample points.
2. It minimizes the total curvature among all the functions. This is a fact that was theoretically (mathematically) proved.

3. It gives natural curves found in the real world (it gives the same curve as the curve ruler).

Extensions

Spline function and its extensions are also used for representing smooth curves on a two-dimensional plane.

- Bezier curves
- Post Script
- True type

4.5.3 Two-dimensional continuous interpolation 1: Introduction

Two-dimensional interpolation is a natural extension of one-dimensional spline interpolation.

A crucial difference lies in the way of dividing interpolation region into subregions.
In one-dimensional interpolation, sample points are distributed on a one-dimensional space, that is, a line segment. This leads us to natural definition of subregions and knots used for spatial interpolation – sample points are also used as knots.

In two-dimensional interpolation, on the other hand, sample points are distributed on a two-dimensional plane, which may be a regular pattern or may be an irregular distribution. This makes it difficult to divide the interpolation region naturally into the subregions in which spline functions are estimated. In two-dimensional interpolation, we also have to discuss the way of dividing interpolation region into subregions.

Except the division of interpolation region, two-dimensional interpolation is not substantially different from one-dimensional interpolation.

In the following, two-dimensional interpolation when sample points are regularly located is discussed first, though it rarely happens in the real world, because it is similar to one-dimensional interpolation. After that, the case of irregular sample points is explained.

In each cell, we consider a different spline function.

$n$: the degree of spline function

$$f_0 (x, y) = a_{00} + a_{10} x + a_{01} y + a_{11} xy + a_{20} x^2 + a_{02} y^2 + a_{12} x^2 y + a_{21} x y^2 + a_{22} x^2 y^2 + \cdots$$

We have $m_x m_y (n+1)^2$ unknown parameters to be estimated.
4. Preprocessing

On the other hand, we have constraints to be satisfied at the corner and boundary of cells.

After all, the degree of freedom is $(m_x + m_y + 2)(n - 1) + (n - 1)^2$

Bilinear spline

Bilinear spline uses two-dimensional linear function defined on a two-dimensional plane.

Spline function for the $(i,j)$th $(i=1, \ldots, m_x; j=1, \ldots, m_y)$ subregion:

$$f_{ij}(x, y) = a_{00} + a_{10}x + a_{01}y + a_{11}xy$$

Note that the function does not represent a flat surface but a curved surface (linear function on a one-dimensional space is a straight line).

Degree of freedom: 0 (no additional constraint is necessary)

Figure: A surface generated by the bilinear spline

Biquadratic spline

Biquadratic spline uses two-dimensional quadratic function:

$$f_{ij}(x, y) = a_{00} + a_{10}x + a_{01}y + a_{11}xy + a_{12}x^2 + a_{02}y^2 + a_{22}x^2y + a_{21}xy^2 + a_{20}x^2y^2$$

Degree of freedom: $m_x + m_y + 3$

Figure: A surface generated by the bilinear spline
To determine spline functions, we have to impose $m_x + m_y + 3$ constraints. We usually give the 1st derivatives on the outer boundary of interpolation region.

1. The 1st derivatives by $y$ at the lower end points: $m_y + 1$
2. The 1st derivatives by $x$ at the right end points: $m_x + 1$
3. The 1st derivatives by $y$ and $x$ at the lower right corner point: 1

Figure: A surface generated by the biquadratic spline

Bicubic spline

Bicubic spline uses two-dimensional cubic function:

$$f(x, y) = a_{00} + a_{10}x + a_{01}y + a_{11}xy + a_{20}x^2 + a_{02}y^2 + a_{21}xy^2 + a_{22}x^2y^2$$

Degree of freedom: $2(m_x + m_y + 4)$

To determine spline functions, we have to impose $2(m_x + m_y + 4)$ constraints. In cubic spline, we again give the 1st derivatives on the outer boundary of interpolation region.
1. The 1st derivatives by $y$ at the upper and lower end points: $2(m_x+1)$
2. The 1st derivatives by $x$ at the right and left end points: $2(m_y+1)$
3. The 1st derivatives by $y$ and $x$ at the corner points: $4m_x m_y$

4. Preprocessing

4.5.5 Two-dimensional continuous interpolation 3: Irregular sample points

If the sample points are regularly distributed, we can naturally define rectangular subregions for interpolating continuous function.

However, to measure the function value at regularly distributed points is practically impossible. Sample points may be located in a river, on a cliff, or in a building.

Division of the interpolation region

We use Triangular Irregular Network (TIN) to divide the interpolation region into subregions.

Different spline functions are defined in individual triangular subregions. They are estimated to satisfy constraints similar to those imposed for interpolation by regularly-distributed sample points.

Cubic spline is generally used as the spline function.
There are numerous triangular networks that divide the interpolation region into subregions.

Figure: A point distribution

Figure: Two different triangular networks

Among numerous TINs, in spatial interpolation, Delaunay triangulation is usually chosen, because it does not produce elongated, acute-angled triangles. This makes spline functions smooth, not fluctuated.

We can construct Delaunay triangulation by connecting points whose Voronoi regions are neighboring. Recent GIS provides a function of constructing Delaunay triangulation from a point distribution.

Homework Q.4.1 (10 pts)
Find an application of spatial interpolation, and report its objective, data acquisition method, interpolation method, and so forth.

Homework Q.4.2 (20 pts)
Suppose you create spatial data of the surface elevation of the earth. Then you have to 1) determine the sample points, the locations whose elevation you measure, and 2) interpolate the elevation data measured to obtain the surface data.

Discuss how you should determine the sample points.