Spatial Analysis Using Grids

Learning Objectives

• Continuous surfaces or spatial fields representation of geographical information
• Grid data structure for representing numerical and categorical data
• Map algebra raster calculations
• Interpolation
• Calculate slope on a raster using
  – ArcGIS method based in finite differences
  – D8 steepest single flow direction
Readings - Theory

- ArcMap help: What is raster data?

- ArcMap help: Fundamentals of raster data from Cell size of raster data to Raster dataset attribute tables
Readings – ArcGIS Spatial Analyst Tools

• An overview of the Spatial Analyst Toolbox

Read this to get a general sense of the tools available. Tools we will use now or later include, from the Surface toolset: Slope, Aspect, Contour; from the Zonal toolset: Zonal Statistics; from Map Algebra: Raster Calculator; from Interpolation: Spline and Kriging; and many tools in the Hydrology toolset.
Readings – Slope and Aspect

• How Slope Works

• How Aspect Works
Two fundamental ways of representing geography are **discrete objects** and **fields**.

The **discrete object view** represents the real world as objects with well defined boundaries in empty space.

The **field view** represents the real world as a finite number of variables, each one defined at each possible position.
Numerical representation of a spatial surface (field)

Grid or Raster

TIN

Contour and flowline
Discrete (vector) and continuous (raster) data

Raster and Vector Data

Raster data are described by a cell grid, one value per cell

*Point*

*Line*

*Polygon*

**Vector** → **Raster**

Zone of cells
Line as a Sequence of Cells
Polygon as Zone of Grid Cells

From: http://resources.arcgis.com/en/help/main/10.2/index.html#/How_features_are_represented_in_a_raster/009t00000006000000/
Raster and Vector are two methods of representing geographic data in GIS

- Both represent different ways to encode and generalize geographic phenomena
- Both can be used to code both fields and discrete objects
- In practice a strong association between raster and fields and vector and discrete objects
A **grid** defines geographic space as a mesh of identically-sized square cells. Each cell holds a numeric value that measures a geographic attribute (like elevation) for that unit of space.
The grid data structure

- Grid size is defined by **extent**, **spacing** and **no data value** information
  - Number of rows, number of column
  - Cell sizes (X and Y)
  - Top, left, bottom and right coordinates
- Grid values
  - Real (floating decimal point)
  - Integer (may have associated attribute table)
NODATA Cells
Cell Networks
Floating Point Grids

Continuous data surfaces using floating point or decimal numbers
Integer valued grids to represent zones
Value attribute table for categorical (integer) grid data

Attributes of grid zones

<table>
<thead>
<tr>
<th>Value</th>
<th>Count</th>
<th>Type</th>
<th>Code</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>6</td>
<td>Maple</td>
<td>200</td>
</tr>
<tr>
<td>2</td>
<td>23</td>
<td>Oak</td>
<td>400</td>
</tr>
<tr>
<td>3</td>
<td>8</td>
<td>Pine</td>
<td>300</td>
</tr>
</tbody>
</table>
Raster Sampling

(1) values are averages for cells

(2) values are samples at cell centers

(3) values are samples at the grid nodes

Cell size of raster data

- Smaller cell size
- Higher resolution
- Higher feature spatial accuracy
- Slower display
- Slower processing
- Larger file size

- Larger cell size
- Lower resolution
- Lower feature spatial accuracy
- Faster display
- Faster processing
- Smaller file size

Cell by cell evaluation of mathematical functions

Example

Precipitation - Losses (Evaporation, Infiltration) = Runoff
Runoff generation processes

- Infiltration excess overland flow
  aka Horton overland flow

- Partial area infiltration excess overland flow

- Saturation excess overland flow
  Soils with high infiltration capacities
Runoff generation at a point depends on

- Rainfall intensity or amount
- Antecedent conditions
- Soils and vegetation
- Depth to water table (topography)
- Time scale of interest

These vary spatially which suggests a spatial geographic approach to runoff estimation
Cell based discharge mapping flow
accumulation of generated runoff

- Radar Precipitation grid
- Soil and land use grid
- Runoff grid from raster calculator operations implementing runoff generation formula’s
- Accumulation of runoff within watersheds
Raster calculation – some subtleties

Resampling or interpolation (and reprojection) of inputs to target extent, cell size, and projection within region defined by analysis mask.

Analysis extent

Analysis cell size

Analysis mask
Spatial Snowmelt Raster Calculation Example

The grids below depict initial snow depth and average temperature over a day for an area.

(a) Initial snow depth (cm)

(b) Temperature (°C)

One way to calculate decrease in snow depth due to melt is to use a temperature index model that uses the formula

\[ D_{\text{new}} = D_{\text{old}} - m \cdot T \]

Here \( D_{\text{old}} \) and \( D_{\text{new}} \) give the snow depth at the beginning and end of a time step, \( T \) gives the temperature and \( m \) is a melt factor. Assume melt factor \( m = 0.5 \, \text{cm/°C/day} \).

Calculate the snow depth at the end of the day.
Lets Experiment with this in ArcGIS

snow.asc

ncols 3
nrows 3
xllcorner 0
yllcorner 0
cellsze 100
NODATA_value -9999
40 50 55
42 47 43
42 44 41

temp.asc

ncols 2
nrows 2
xllcorner 0
yllcorner 0
cellsze 150
NODATA_value -9999
4 6
2 4
New depth calculation using Raster Calculator

“snow100” - 0.5 * “temp150”
The Result

• Outputs are on 150 m grid.

• How were values obtained?
Nearest Neighbor Resampling with Cellsize Maximum of Inputs

100 m

40 50 55

42 47 43

42 44 41

150 m

4 6

2 4

40 - 0.5 * 4 = 38

55 - 0.5 * 6 = 52

42 - 0.5 * 2 = 41

41 - 0.5 * 4 = 39

38 52

41 39
Inferences

- ArcGIS has used nearest neighbors to align cells.
- ArcGIS has done raster computation at the scale of the coarsest input (largest cell size of all rasters).

- Is this what we want?
- How to control scale?
- What is meant by scale in this context?
Scale issues in interpretation of measurements and modeling results

The scale triplet

a) Extent  
b) Spacing  
c) Support

Fig. 2.3.1 The scale triplet of measurement scale and modelling scale: (a) extent; (b) spacing; (c) support in either space or time.

Fig. 2.3.6 The effect of sampling for measurement scales not commensurate with the process scale. (a) Spacings larger than the process scale cause aliasing in the data; (b) Extents smaller than the process scale cause a trend in the data; (c) Supports larger than the process scale cause excessive smoothing in the data. The process scale is the period.
Interpolation

Estimate values between known values.

A set of spatial analyst functions that predict values for a surface from a limited number of sample points creating a continuous raster.

Apparent improvement in resolution may not be justified
Interpolation methods

- Nearest neighbor
- Inverse distance weight
- Bilinear interpolation
- Kriging (best linear unbiased estimator)
- Spline

\[
z = \sum \frac{1}{r_i} z_i
\]

\[
z = (a + bx)(c + dy)
\]

\[
z = \sum w_i z_i
\]

\[
z = \sum c_i x^{e_i} y^{e_i}
\]
Resample to get consistent cell size

Spacing & Support
Calculation with consistent 100 m cell size grid

“snow100” - 0.5 * “temp100”

• Outputs are on 100 m grid as desired.

• How were these values obtained?
100 m cell size raster calculation

40 - 0.5*4 = 38
50 - 0.5*5 = 47.5
55 - 0.5*6 = 52
42 - 0.5*3 = 40.5
47 - 0.5*4 = 45
43 - 0.5*5 = 40.5
41 - 0.5*4 = 39
42 - 0.5*2 = 41
44 - 0.5*3 = 42.5
Comparison

• 150 cell size

• 100 m cell size
Point to Raster Interpolation

Nearest Neighbor “Thiessen” Polygon Interpolation

Spline Interpolation
Interpolation Comparison

Figure 2.4. Comparison of interpolation methods for a one-dimensional example: (a) Thiessen method; (b) inverse distance squared; (c) example of overfitting using a sixth-order polynomial; (d) thin plate splines with different tension parameters; (e) kriging (zero nugget, large range); (f) kriging (solid line: large nugget, large range; dashed line: zero nugget, very short range).
Further Reading


Chapter 2. Spatial Observations and Interpolation

Full text online at:

Spatial Surfaces used in Hydrology

Elevation Surface — the ground surface elevation at each point
3-D detail of the Tongue river at the WY/Mont border from LIDAR.

Roberto Gutierrez
University of Texas at Austin
Topographic Slope

- Defined or represented by one of the following
  - Surface derivative $\nabla z$  ($dz/dx, dz/dy$)
  - Vector with x and y components ($S_x, S_y$)
  - Vector with magnitude (slope) and direction (aspect) ($S, \alpha$)

Slope and Aspect

Degree of slope = $\theta$

\[
\frac{\text{rise}}{\text{run}} = \tan \theta
\]

Percent of slope = \(\frac{\text{rise}}{\text{run}} \times 100\)

Degree of slope = 30
Percent of slope = 58

$\alpha$ = aspect clockwise from North

Large slope!

ArcGIS Topographic “Slope” tool

\[
\begin{array}{|c|c|c|}
\hline
a & b & c \\
\hline
d & e & f \\
\hline
g & h & i \\
\hline
\end{array}
\]

\[
\frac{a-c}{2\Delta} + \frac{d-f}{2\Delta} + \frac{d-f}{2\Delta} + \frac{g-i}{2\Delta} = \frac{(a + 2d + g) - (c + 2f + i)}{8\Delta}
\]

Similarly

\[
\frac{dz}{dx} = \frac{(g + 2h + i) - (a + 2b + c)}{8\Delta}
\]
ArcGIS Aspect –
the steepest downslope direction

\[
\begin{align*}
\frac{dz}{dy} & \quad \frac{dz}{dx} \\
\text{atan} \left( \frac{\frac{dz}{dx}}{\frac{dz}{dy}} \right) & 
\end{align*}
\]
Example

\[
\begin{align*}
\frac{dz}{dx} &= \frac{(a + 2d + g) - (c + 2f + i)}{8 \times x_{\text{mesh\_spacing}}} \\
&= \frac{(80 + 2 \times 69 + 60) - (63 + 2 \times 56 + 48)}{8 \times 30} \\
&= \frac{60}{0.229} \\
\frac{dz}{dy} &= \frac{(g + 2h + i) - (a + 2b + c)}{8 \times y_{\text{mesh\_spacing}}} \\
&= \frac{(60 + 2 \times 52 + 48) - (80 + 2 \times 74 + 63)}{8 \times 30} \\
&= -0.329
\end{align*}
\]

Slope = $\sqrt{0.229^2 + 0.329^2}$

= 0.401

atan(0.401) = 21.8°

Aspect = atan\left(\frac{0.229}{-0.329}\right) = -34.8° + 180°

⇒ 145.2°

Aspect is then converted to a compass direction.
Hydrologic Slope (Flow Direction Tool) - Direction of Steepest Descent

Slope: \[ \frac{67 - 48}{30\sqrt{2}} = 0.45 \]

\[ \frac{67 - 52}{30} = 0.50 \]

Hydrologic Slope = Rise/Run...elevation difference/distance from center of cell
Eight Direction Pour Point Model

ESRI Direction encoding
Limitation due to 8 grid directions.
The D∞ Algorithm

Proportion flowing to neighboring grid cell 4 is $\frac{\alpha_1}{\alpha_1 + \alpha_2}$

Steepest direction downslope

Proportion flowing to neighboring grid cell 3 is $\frac{\alpha_2}{\alpha_1 + \alpha_2}$

The $D^\infty$ Algorithm

If $\alpha_1$ does not fit within the triangle the angle is chosen along the steepest edge or diagonal resulting in a slope and direction equivalent to D8.

The diagram illustrates the steepest direction downslope with elevations at each vertex $z_i$.

Let $\alpha_1 = \tan^{-1} \left( \frac{z_1 - z_2}{z_0 - z_1} \right)$.

The slope $S$ is given by:

$$S = \sqrt{\left( \frac{z_1 - z_2}{\Delta} \right)^2 + \left( \frac{z_0 - z_1}{\Delta} \right)^2}$$
D∞ Example

\[ \alpha_1 = \arctan \left( \frac{Z_7 - Z_8}{Z_0 - Z_7} \right) \]

\[ = \arctan \left( \frac{52 - 48}{67 - 52} \right) = 14.9^\circ \]

\[ S = \sqrt{\left( \frac{52 - 48}{30} \right)^2 + \left( \frac{67 - 52}{30} \right)^2} \]

\[ = 0.517 \]

From ArcGIS Pro Help:

- The D-Infinity (DINF) flow method, described by Tarboton (1997), determines flow direction as the steepest downward slope on eight triangular facets formed in a 3x3 cell window centered on the cell of interest. Flow direction output is a floating point raster represented as a single angle in degrees going counter-clockwise from 0 (due east) to 360 (again due east).
Summary Concepts

• **Grid (raster) data structures** represent surfaces as an array of grid cells

• **Raster calculation** involves algebraic like operations on grids

• **Interpolation and Generalization** is an inherent part of the raster data representation
Summary Concepts (2)

• The elevation surface represented by a grid digital elevation model is used to derive slope important for surface flow
• The eight direction pour point model approximates the surface flow using eight discrete grid directions.