12 AUG 2013

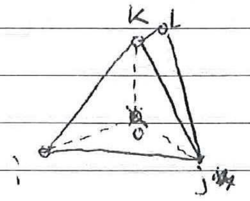
TU Delft
Technische Universiteit Delft

naam / *name*	Lars Keuris
studienummer / *student number*	
vak / *course*	Data Analyse & Geostatics
code / *code*	TA 2060
datum / *date*	
opleiding / *program*	
aantal ingeleverde vellen / *total number of sheets*	1 of 3
opgave nummer / *question number*	

A. Deterministische interpolatie

1,2 Triangular interpolation makes use of the following formula:

$$h_0 = \frac{A_{0jk}}{A_{ijk}} h_i + \frac{A_{i0k}}{A_{ijk}} h_j + \frac{A_{ij0}}{A_{ijk}} h_k$$



h_0 : prediction location height
h_i : ~~[illegible]~~ height of vertex i
$\frac{A_{0jk}}{A_{ijk}}$: weight of location i

Since the surfaces of the three sub triangles are almost equal, $\frac{1}{3}$ of the full triangle, the weights of the sub triangles are also equal.
Vertex l does not have a weight for prediction location "0", because it is not a vertex of a triangle in which "0" is present.

$$h_0 = \frac{1}{3}\cdot 2 + \frac{1}{3}\cdot 4 + \frac{1}{3}\cdot 20 = 8\frac{2}{3}$$

3,4 Inverse distance interpolation

$$h_0 = \frac{1}{\sum_{i=1}^{n} \frac{1}{d_i^p}} \left(\frac{h_1}{d_1^p} + \dots + \frac{h_n}{d_n^p} \right)$$

$$= \frac{1}{\frac{1}{1,1^2} + \frac{1}{1,1^2} + \frac{1}{1,1^2} + \frac{1}{1,2^2}} \left(\frac{2}{1,1^2} + \frac{4}{1,1^2} + \frac{20}{1,1^2} + \frac{21}{1,2^2} \right)$$

$$= 0,32 \cdot 36,[illegible] = 11,4$$

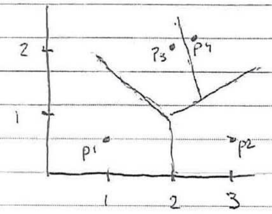
One could also say that the distances are all nearly equal. Whereas there are 4 points, so the weights are $\frac{1}{4}$ for each point.

$$h_0 = \frac{1}{4} \cdot 2 + \frac{1}{4} \cdot 4 + \frac{1}{4} \cdot 20 + \frac{1}{4} \cdot 21 = 12{,}50$$

Which is in this case a good assumption.

5 Increasing the power will result in a faster decrease in weight when you get further away from your prediction location. But only if your distance is larger than one.

6



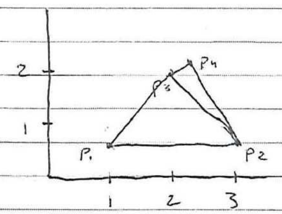
7 Nearest neighbour or Nearest Centroid uses the Voronoi diagrams to determine the value of the prediction location.

Using: if: $p_0 \in V(p_i)$: then $h_0 = h_i$

Saying that if a prediction location lies within the Voronoi cell of a certain p_i. Meaning that the nearest known location is p_i. Then the height, or other data, ~~is to~~ of the prediction location is to be same as the nearest known location.

So now we can deduce the height of (3,2).
→ (3,2) is closest to and lies within the Voronoi cell of, p_4 ~~thus~~ thus the height will be $h_0 = h_4 = 21$

8



You could draw the triangulation in two ways. Using p_1, p_2, p_3 and p_3, p_2, p_4 OR p_1, p_4, p_2 and p_1, p_3, p_4.

However we use the Delaunay triangulation, so we maximize the minimum angle overall triangulations. Meaning that the triangles are nicely distributed over all points.

9 The convex hull is the polygon ~~e~~ enclosing all points, so the line follows: p_1, p_2, p_4, p_3 back to p_1

10 There is no triangle enclosing (3,2), so my guess is the ~~the~~ triangular interpolation will say it's value is zero.

B Stochastische interpolatie

11. mean = $\frac{2 + 4 + 20 + 21}{4} = 11\frac{3}{4}$

12. $f(x) = e^{-3x}$, where x is distance

distance $p_1 \rightarrow p_2 = 2$
distance $p_3 \rightarrow p_4 = 0{,}15$

Now you could calculate the covariance between the observations by using the formula.

$f(2) = e^{-6} \approx 0$
$f(0{,}15) = e^{-0{,}45} \approx 0{,}6$

Or you could look in the graph.

13. Kriging the mean ensures that observations that are close to each other, or highly correlated have less influence. Kriging the mean works practically as follows:

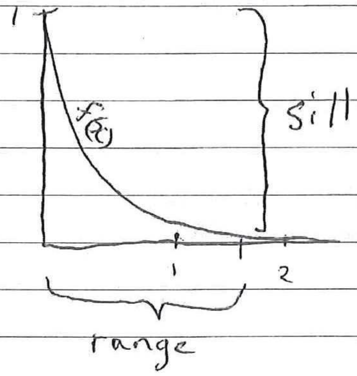
- p_3 and p_4 are highly correlated, merge the two points by making p_5 in between with value 20,5.
- Now we have 3 points; p_1, p_2 and p_5 with more or less the same correlation (because of equal distances). So the weights should be equal as well.

$$\rightarrow \quad \text{mean} = \frac{1}{3} \cdot 2 + \frac{1}{3} \cdot 4 + \frac{1}{3} \cdot 20{,}5 = 8\frac{5}{6}$$

14. If the observations are not correlated, all values deserve the same weight. This will be:

$$\underset{\text{(uncorrelated)}}{\text{mean}} = \frac{1}{4} \cdot 2 + \frac{1}{4} \cdot 4 + \frac{1}{4} \cdot 20 + \frac{1}{4} \cdot 21 = 11\frac{3}{4}$$

15.



range = 1,5
sill = 1

16. Redundancy matrix is a matrix built up out of the correlation of distances of observations among other observations.
On the diagonal are the distances to the observation itself, which is 0. Taking the correlation of that distance: $f(0) = 1$.
So there are ones on the diagonal.
Since the distance from p_1 to p_2 is the same as the distance from p_2 to p_1, the correlation values are diagonally symmetric.

naam *name*	lars keuris
studienummer *student number*	
vak *course*	Data Analysis and Geostatistics
code *code*	TA2060
datum *date*	
opleiding *program*	
aantal ingeleverde vellen *total number of sheets*	2 of 3
opgave nummer *question number*	

TU Delft
Technische Universiteit Delft

So let's calculate distances first, then correlation

$$\begin{bmatrix} C_{11} & C_{12} & C_{13} & C_{14} & 1 \\ C_{21} & C_{22} & C_{23} & C_{24} & 1 \\ C_{31} & C_{32} & C_{33} & C_{34} & 1 \\ C_{41} & C_{42} & C_{43} & C_{44} & 1 \\ 1 & 1 & 1 & 1 & 0 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & C_{12} & C_{13} & C_{14} & 1 \\ C_{12} & 1 & C_{23} & C_{24} & 1 \\ C_{13} & C_{23} & 1 & C_{34} & 1 \\ C_{14} & C_{24} & C_{34} & 1 & 1 \\ 1 & 1 & 1 & 1 & 0 \end{bmatrix}$$

The ones and zero bottom and right are always present.

$d_{12} = 2 \rightarrow C_{12} = f(d_{12}) = f(2) = e^{-6} \approx 0$

$d_{13} = 2 \rightarrow C_{13} = f(d_{13}) = f(2) = e^{-6} \approx 0$

$d_{14} = 2{,}15 \rightarrow C_{14} = f(d_{14}) = f(2{,}15) = e^{-6{,}45} \approx 0$

$d_{23} = 2 \rightarrow C_{23} = f(d_{23}) = f(2) = e^{-6} \approx 0$

$d_{34} = 0{,}15 \rightarrow C_{34} = f(d_{34}) = f(0{,}15) = e^{-0{,}45} \approx 0{,}64$

$d_{24} = 2{,}1 \rightarrow C_{24} = f(d_{24}) = f(2{,}1) = e^{-6{,}3} \approx 0$

Filling in:

$$\begin{bmatrix} 1 & e^{-6} & e^{-6} & e^{-6{,}45} & 1 \\ e^{-6} & 1 & e^{-6} & e^{-6{,}3} & 1 \\ e^{-6} & e^{-6} & 1 & e^{-0{,}45} & 1 \\ e^{-6{,}45} & e^{-6{,}3} & e^{-0{,}45} & 1 & 1 \\ 1 & 1 & 1 & 1 & 0 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0{,}64 & 1 \\ 0 & 0 & 0{,}64 & 1 & 1 \\ 1 & 1 & 1 & 1 & 0 \end{bmatrix}$$

17 Proximity vector states the distances from all observations to the prediction location.

$$\begin{bmatrix} C_{1x} \\ C_{2x} \\ C_{3x} \\ C_{4x} \\ 1 \end{bmatrix} \rightarrow \begin{bmatrix} e^{-3{,}3} \\ e^{-3{,}3} \\ e^{-3{,}3} \\ e^{-3{,}6} \\ 1 \end{bmatrix}$$

$d_{1x} = 1{,}1 \rightarrow C_{1x} = f(d_{1x}) = f(1{,}1) = e^{-3{,}3}$

18 The redundancy vector is squared, otherwise it would not be possible.

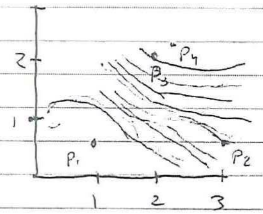
19 OK (Ordinary Kriging) formula:

- redundancy matrix · Weight matrix = proximity vector
- $\Gamma \cdot \lambda = g \quad \rightarrow \quad \lambda = \Gamma^{-1} \cdot g$

Γ is a n+1, n+1 matrix (redundancy)
λ is a n+1, 1 vector (weights)
g is a n+1, 1 vector (proximity)

Use your graphic calculator to calculate the inverse of Γ and to calculate the multiplication.

20



At a distance of one, the covariance is nearly zero. This means that an observation has little to no influence when you're further than one away. The contour lines are less densely grouped when further away from an observation.

C. Landgebruik

21. Hoogtedata
- radar
- gravity
- GPS or actually GNSS

22: GIS raster format is a huge matrix, the location in the matrix is also the location in the image. It ~~most~~ is widely used for continuous data like topography.

23. GIS vector format is used to ~~make~~ create certain points, lines or polygons. Every point has to be saved explicitly and is useful for linear data, like roads and rivers.

24. The UTM coordinate system is divided into zones to reduce the distortion of the map itself. The smaller the zones, the smaller the distortion. UTM preserves its angle, but not surface.

25. WGS 84 is written in latitude and longitude. However it is easy to convert the units and it is done mostly automatically.

26. $y = ax + b$ (2D), $z = ax + by + c$ (3D) ?

27. 3D, 2D whatever...

A	B	C
D	E	F
G	H	I

$S_x = \frac{1}{2}(D-F)$

$S_y = \frac{1}{2}(B-H)$

$S = \frac{1}{2}\sqrt{(D-F)^2 + (B-H)^2}$

28. This can be done using MonteCarlo simulations.

29. Spectral band data is necessary to distinguish specific classes, otherwise classification would not be possible. Geometric data is relevant to determine slopes and heights for different classes. These could be unique.

30. Feature space is the amount of bands available. This is 14, one could get that from question 35. as well.

31. Supervised, because they are using training data.

32. Nearest centroid classification makes use of Voronoi diagrams. ~~Drawing these~~ Of each unknown pixel, the nearest training data is determined. Then the unknown pixel copies the land cover class from this specific training data.

33. ...

34 Commission error: ~~land~~ water was wrongly determined as forest
Omission error: ~~be~~ land cover class that should be determined water was not determined as water but forest :)

35 Highly correlated bands do not bring much new data. While little to no correlated bands should be maintained since this is telling something new. Highly correlated bands could be deleted, while keeping one of each pair of the highly correlated bands.

D Stijgsnelheid.

36. Yes, it does. The GPS signal could've ~~and~~ went to the device via a multipath.

37. $y = ax + b$, 2 parameters

38. $\underline{y} = \begin{bmatrix} 1300 \\ 1640 \\ 1570 \\ 1800 \end{bmatrix}$, look at the graph.

39.

$A = \begin{bmatrix} 10 & 20 & 30 & 40 \end{bmatrix}$

40. $y = ax + b$

$\underline{y} = A\underline{x} + \underline{e} \rightarrow \underline{e} = \underline{y} - A\underline{x}$

$A^T \cdot \underline{e} = 0$

$A^T(\underline{y} - A\underline{x}) = 0$

$A^T\underline{y} - A^TA\underline{x} = 0$

$A^T\underline{y} = A^TA\underline{x}$

$(A^TA)^{-1}A^T\underline{y} = x$

naam / *name*	lars keuris
studienummer / *student number*	
vak / *course*	Data Analyse en Geostatistiek
code / *code*	TA 2060
datum / *date*	
opleiding / *program*	
aantal ingeleverde vellen / *total number of sheets*	3 van 3
opgave nummer / *question number*	

TU Delft

Technische Universiteit Delft

41 $x = \begin{bmatrix} 16,2 \\ 1129 \end{bmatrix} \rightarrow y = 16,2x + 1129$

16,2 m/min

42/ 43. ~~Original~~

t(min)	10	20	30	40	
h(m) ~~t(min)~~	1300	1640	1570	1800	observation
h(m)	1286	1448	1610	1772	Least squares trend
difference(m)	+14	+192	−40	+28	

Observation 2 has the highest deviatoon.

44. RanSaC
- Select 2 random points
- Fit a line through the points
- Determine residuals between points and line
  a. inliers: points with small residuals
  b. outliers: points with high residuals
- Score of the run: number of inliers
- Repeat the previous steps and choose the line with the most inliers (highest score)

45. Around 50 meters.

18 The redundancy vector is squared, otherwise it would not be possible.

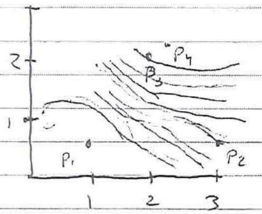
19 OK (Ordinary Kriging) formula:

- redundancy matrix · Weight matrix = proximity vector
- $\Gamma \cdot \lambda = g \rightarrow \lambda = \Gamma^{-1} \cdot g$

Γ is a n+1, n+1 matrix (redundancy)
λ is a n+1, 1 vector (weights)
g is a n+1, 1 vector (proximity)

Use your graphic calculator to calculate the inverse of Γ and to calculate the multiplication.

20



At a distance of one, the covariance is nearly zero. This means that an observation has little to no influence when you're further than one away. The contour lines are less densely grouped when further away from an observation.

C. Landgebruik

21. Hoogtedata
- radar
- gravity
- GPS or actually GNSS

22: GIS raster format is a huge matrix, the location in the matrix is also the location in the image. It ~~most~~ is widely used for continuous data like topography.

23. GIS vector format is used to ~~make~~ create certain points, lines or polygons. Every point has to be saved explicitly and is useful for linear data, like roads and rivers.

24. The UTM coordinate system is divided into zones to reduce the distortion of the map itself. The smaller the zones, the smaller the distortion. UTM preserves its angle, but not surface.

25. WGS84 is written in latitude and longitude. However it is easy to convert the units and it is done mostly automatically.

26. $y = ax + b$ (2D), $z = ax + by + c$ (3D) ?

27. 3D, 2D whatever...

A	B	C
D	E	F
G	H	I

$S_x = \frac{1}{2}(D - F)$
$S_y = \frac{1}{2}(B - H)$
$S = \frac{1}{2}\sqrt{(D-F)^2 + (B-H)^2}$

28. This can be done using MonteCarlo simulations.

29. Spectral band data is necessary to distinguish specific classes, otherwise classification would not be possible. Geometric data is relevant to determine slopes and heights for different classes. These could be unique.

30. Feature space is the amount of bands available. This is 14, one could get that from question 35. as well.

31. Supervised, because they are using training data.

32. Nearest centroid classification makes use of Voronoi diagrams. ~~Drawing these~~ Of each unknown pixel, the nearest training data is determined. Then the unknown pixel copies the land cover class from this specific training data.