KIT-HYD / **geostatistics_lecture**

Code | Issues 1 | Pull requests | Actions | Projects

master | geostatistics_lecture / exercises / **variogram** / ...

mmaelicke Added decomposition ... ... ✓ on 27 Apr 2020 **History**

..	
PREC_GAUGE_COOR.csv	12 months ago
README.md	12 months ago
artificial.csv	12 months ago
cressie.m	12 months ago
matheron.m	12 months ago
test_cressie.m	12 months ago
test_matheron.m	12 months ago

README.md

Variograms

In the fourth lecture will take the step from univariate and simple multivariate statistics to **geostatistics**. You will need the code of the last lecture. Remeber task 5 which asked you to calculate a functional relationship between the *proximity in distance* and the *proximity in observation value* for all observation point pairs. We will basically substitute the mean as an *estimator* for proximity by an estimator called **semi-variance** and end up with a function that is usually referred to as the **experimental variogram**.

Task 1

The by far most common semi-variance estimator is used so heavily, that it is often referred to as **the semi-variance**: the Matheron semi-variance, introduced by Geoges Matheron in 1963. This semi-variance is defined as:

$$\gamma(h) = \frac{1}{2N} \sum_{i=1}^{N} (x_i - (x_i + h))^2$$

Implement this formula as the `matheron.m` function. Note that `matheron.m` should accept the **array of differences**, not the array of values. The differences are the part ![matheron] (http://latex.codecogs.com/svg.latex?x_{h}=\{x_i-(x_i+h)\} for a given value of h.

You can use the `test_matheron.m` to test your code.

Task 2

Calculate the experimental variogram of the given observations in `PREC_GAUGE_COOR.csv` as you did in task 5 of the last exercise and plot the two functions next to each other.

Task 3

In the folder of this lecture, you can find a file `artificial.csv` containing a sample drawn from an artificial dataset. Is there an apparent spatial structure? How did you come up with your semivariogram?

Task 4

You learned about the three variogram parameters in class:

- nugget
- sill
- effective range

The *nugget* is the y-axis interaction and specifies the amount of variance present at no distance. In other words this is the intrinsic uncertainty in the data at the observation points themselves.

The *sill* is the overall variance in the data set, that is approached with increasing distances. The distance at with this happens is called *effective range*. This is the distance at which the sample becomes statistically independed as the point pairs do not bear any predictive information about one another.

Estimate the three variogram parameters for both variograms from task 3 visually.

Task 5

As stated in the introduction to this exercise, the Matheron semi-variance is often referred to as **the** semi-variance, but in fact it is only one estimator among others. Another estimator used is the Cressie-Hawkins estimator defines as:

$$\gamma(h) = \frac{1}{2} \left(\frac{1}{N(h)} \sum_{i,j} \sqrt{|z(x_i) - z(x_j)|} \right)^4 \left(0.457 + \frac{0.494}{N(h)} + \frac{0.045}{N^2(h)} \right)^{-1}$$

Implement this estimator into `cressie.m` and use `test_cressie.m` to test your code.

Task 6

Create a figure of a Matheron semi-variogram and a Cressie-Hawkins semi-variogram next to each other for both data sets. Describe the differences. How do they change if you change the binning?