Spatial data - general comments

• Why spatial statistics? Noel Cressie ("Statistics for Spatial Data") writes “why, how, when” are not enough. We need to add “where”.

• Today, spatial statistics models appear in areas such as mining, geology, hydrology, ecology, environmental science, medicine, image processing, crop science, epidemiology, forestry, atmospheric science, etc.

• Need to develop models that deal with data collected at spatial locations.

• The basic components are the spatial locations \{s_1, s_2, \ldots, s_n\} and the data observed at these locations denoted as \{Z(s_1), Z(s_2), \ldots, Z(s_n)\}.

• The distance between the observations is important in analyzing spatial data. With distance we mostly mean “Euclidean distance”. However there are other forms of distances (e.g. road miles, travel time, etc.). The latter is modeled through multidimensional scaling. Here we will consider mostly (if not always) Euclidean distances.

• Consider the following example taught in introductory statistics courses (Cressie, 1993): Let the spatial data \(Z(s_1), Z(s_2), \ldots, Z(s_n)\) be an i.i.d. sample from \(N(\mu, \sigma_0)\). The minimum variance unbiased estimator (MVUE) of \(\mu\) is

\[
\bar{Z} = \frac{1}{n} \sum_{i=1}^{n} Z(s_i), \text{ and } var(\bar{Z}) = \frac{\sigma_0^2}{n}.
\]

We know that \(\bar{Z} \sim N(\mu, \frac{\sigma_0^2}{n})\), and therefore we can easily construct a 95% confidence interval for \(\mu\) as follows:

\[
\bar{Z} \pm 1.96 \frac{\sigma_0}{\sqrt{n}}.
\]

• The previous example assumes an i.i.d. sample. This can be too simplistic for spatial data. A more realistic assumption is that the data exhibit some spatial correlation. Suppose this spatial correlation is represented through the covariance function

\[
\text{cov}(Z(s_i), Z(s_j)) = \sigma_0^2 \rho^{|i-j|}, \ 0 < \rho < 1.
\]

In the i.i.d. case \(\text{cov}(Z(s_i), Z(s_j)) = 0\) (independent, therefore the covariance is zero).
• What is the confidence interval for the non-i.i.d. case? Find the variance of $\bar{Z}$ first:

$$\text{var}(Z) = \text{var} \left( \frac{1}{n} \sum_{i=1}^{n} Z(s_i) \right) = \frac{1}{n^2} \sum_{i=1}^{n} \sum_{j=1}^{n} \text{cov}(Z(s_i), Z(s_j))$$

Or after some simplification . . .

$$\text{var}(\bar{Z}) = \frac{\sigma_0^2}{n} \left[ 1 + 2 \left( \frac{\rho}{1-\rho} \right) \left( 1 - \frac{1}{n} \right) - 2 \left( \frac{\rho}{1-\rho} \right)^2 \left( \frac{1-\rho^{n-1}}{n} \right) \right]$$  \hspace{1cm} (1)

• Since $Z(s_1), Z(s_2), \ldots, Z(s_n)$ is multivariate normal the distribution of $\bar{Z}$ is also normal with mean $\mu$ and standard deviation the square root of the above expression.

• Suppose $n = 10$ and $\rho = 0.26$. For this example we get:

$$\text{var}(\bar{Z}) = \frac{\sigma_0^2}{10} (1.608) \Rightarrow \bar{Z} \sim N \left( \mu, \frac{\sigma_0 \sqrt{1.608}}{\sqrt{10}} \right)$$

and a two-sided 95% confidence interval for $\mu$ is

$$\bar{Z} \pm 1.96 \frac{\sigma_0 \sqrt{1.608}}{\sqrt{10}} \quad \text{or} \quad \bar{Z} \pm 2.485 \frac{\sigma_0}{\sqrt{10}}$$

• Conclusion: If we do not realize the presence of spatial correlation in our data and we use $\bar{Z} \pm 1.96 \frac{\sigma_0}{\sqrt{10}}$, we obtain a confidence interval that is too narrow. The actual coverage is 87.8%, not 95%. Why?

• Intuitive understanding of the effect of spatial correlation. Write (1) as $\text{var}(\bar{Z}) = \frac{\sigma_0^2}{n'}$ where,

$$n' = \frac{n}{1 + 2 \left( \frac{\rho}{1-\rho} \right) \left( 1 - \frac{1}{n} \right) - 2 \left( \frac{\rho}{1-\rho} \right)^2 \left( \frac{1-\rho^{n-1}}{n} \right)}$$

If $n = 10$ and $\rho = 0.26$ then $n' \approx 6$. What if $n$ is large?

• Linear models with spatially dependent error term:

The classical regression model in matrix form when the error terms are i.i.d. random variables is given by

$$Z = X\beta + \epsilon, \text{ with } \text{var}(\epsilon) = \sigma^2 I$$

and the estimation of $\beta$ is obtained through ordinary least squares (OLS)

$$\hat{\beta}_{ols} = (X'X)^{-1}X'Z$$
• When the error terms are spatially correlated the above model can be written as

\[ Z = X\beta + \delta, \text{ with } var(\delta) = \Sigma, \]

where \( \Sigma \) is the \( n \times n \) variance-covariance matrix. The form of \( \Sigma \) usually is not known. If \( \Sigma \) is known the estimation of \( \beta \) could be obtained through generalized least squares

\[ \hat{\beta}_{gls} = (X'\Sigma^{-1}X)^{-1}X'\Sigma^{-1}Z \]

• If \( \Sigma \) is unknown it can be estimated to obtain estimated generalized least squares:

\[ \hat{\beta}_{vgl} = (X'\hat{\Sigma}^{-1}X)^{-1}X'\hat{\Sigma}^{-1}Z \]
Geostatistics

• Matheron (1963) develop a comprehensive theory mainly for spatial prediction for the estimation of ore reserves.
• Geostatistics can be applied to data sets not only pertaining the earth sciences.
• The data are thought as a realization of

\[ Z(s) : s \in D \]

• The data are denoted with \( \{Z(s_1), Z(s_2), \ldots, Z(s_n)\} \) and are collected at locations \( \{s_1, s_2, \ldots, s_n\} \).
• The cornerstone of geostatistics is the variogram which describes the spatial correlation.
• First let’s discuss the so called “h-scatterplots”.

An h-scatterplot shows all possible pairs of data values whose locations are separated by a certain distance in a particular direction. For example, the figure below shows a 10 \( \times \) 10 grid (100 data points - with distance between two consecutive data points 1 m in north-south and east-west direction).

And on the next page we see the value at each location.
In this first example we will plot each value (we place it on the $x$-axis) against a value that is 1 $m$ apart on the south-north direction (we place it on the $y$-axis). How many pairs are there? There are $9 \times 10 = 90$ pairs. The scatterplot shows a cloud of points distributed around the 45-degree line. We observe some similarity between nearby data points. If we increase the distance from 1 $m$ to 2 $m$, 3 $m$, and 4 $m$ we see that the cloud becomes “fatter” indicating that the values separated by longer distance are not as close as with the 1 $m$ case. These plots are shown below:
• Similarly, if we move on southwest-northeast direction we will pair all values that are $\sqrt{2} m$ apart. In this case there are 81 pairs.

• As with many graphical displays, we need a quantitative summary of the information contained on an $h$-scatterplot. In the previous figure we can measure the “fatness” of the points by the correlation coefficient. The correlation coefficient decreases as the separation distance increases. The relation between the correlation coefficient and $h$ (the separation distance) it is called correlogram. We can also plot the covariance against $h$ to get the so called covariogram.

• How do we compute the covariance and the correlation coefficient for data points separated by $h$?

$$C(h) = \frac{1}{N(h)} \sum_{h_{ij}=h} (Z_i - \bar{Z}_s)(Z_j - \bar{Z}_{s+h})$$

$$\rho(h) = \frac{C(h)}{\sigma_{z_s} \sigma_{z_{s+h}}}$$

where $\sigma_{z_s}, \sigma_{z_{s+h}}$ are the standard deviations of the variable $Z_s$ and the variable $Z_{s+h}$. 
• The two figures below show the covariogram and correlogram for these data.

**Covariogram:**

![Covariogram Graph](image)

**Correlogram:**

![Correlogram Graph](image)

<table>
<thead>
<tr>
<th>$h$</th>
<th>$C(h)$</th>
<th>$\rho(h)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>453.4</td>
<td>0.742</td>
</tr>
<tr>
<td>2</td>
<td>345.1</td>
<td>0.590</td>
</tr>
<tr>
<td>3</td>
<td>327.8</td>
<td>0.560</td>
</tr>
<tr>
<td>4</td>
<td>295.2</td>
<td>0.476</td>
</tr>
</tbody>
</table>
You can verify the result for $h = 1$ using the 90 paired values below:

<table>
<thead>
<tr>
<th>(Z(s))</th>
<th>(Z(s+1))</th>
<th>(Z(s))</th>
<th>(Z(s+1))</th>
</tr>
</thead>
<tbody>
<tr>
<td>87</td>
<td>77</td>
<td>109</td>
<td>143</td>
</tr>
<tr>
<td>77</td>
<td>75</td>
<td>143</td>
<td>99</td>
</tr>
<tr>
<td>75</td>
<td>74</td>
<td>99</td>
<td>101</td>
</tr>
<tr>
<td>74</td>
<td>77</td>
<td>101</td>
<td>113</td>
</tr>
<tr>
<td>77</td>
<td>89</td>
<td>113</td>
<td>108</td>
</tr>
<tr>
<td>89</td>
<td>88</td>
<td>108</td>
<td>64</td>
</tr>
<tr>
<td>88</td>
<td>82</td>
<td>64</td>
<td>91</td>
</tr>
<tr>
<td>82</td>
<td>82</td>
<td>91</td>
<td>77</td>
</tr>
<tr>
<td>82</td>
<td>81</td>
<td>77</td>
<td>19</td>
</tr>
<tr>
<td>100</td>
<td>84</td>
<td>0</td>
<td>91</td>
</tr>
<tr>
<td>84</td>
<td>80</td>
<td>91</td>
<td>95</td>
</tr>
<tr>
<td>80</td>
<td>80</td>
<td>95</td>
<td>96</td>
</tr>
<tr>
<td>80</td>
<td>82</td>
<td>96</td>
<td>79</td>
</tr>
<tr>
<td>82</td>
<td>88</td>
<td>79</td>
<td>73</td>
</tr>
<tr>
<td>88</td>
<td>70</td>
<td>73</td>
<td>84</td>
</tr>
<tr>
<td>70</td>
<td>74</td>
<td>84</td>
<td>73</td>
</tr>
<tr>
<td>74</td>
<td>61</td>
<td>73</td>
<td>52</td>
</tr>
<tr>
<td>61</td>
<td>77</td>
<td>52</td>
<td>40</td>
</tr>
<tr>
<td>47</td>
<td>74</td>
<td>98</td>
<td>52</td>
</tr>
<tr>
<td>74</td>
<td>83</td>
<td>52</td>
<td>48</td>
</tr>
<tr>
<td>83</td>
<td>85</td>
<td>48</td>
<td>72</td>
</tr>
<tr>
<td>85</td>
<td>88</td>
<td>72</td>
<td>102</td>
</tr>
<tr>
<td>88</td>
<td>94</td>
<td>102</td>
<td>107</td>
</tr>
<tr>
<td>94</td>
<td>103</td>
<td>107</td>
<td>105</td>
</tr>
<tr>
<td>103</td>
<td>97</td>
<td>105</td>
<td>115</td>
</tr>
<tr>
<td>97</td>
<td>110</td>
<td>115</td>
<td>111</td>
</tr>
<tr>
<td>110</td>
<td>103</td>
<td>111</td>
<td>111</td>
</tr>
<tr>
<td>111</td>
<td>108</td>
<td>134</td>
<td>136</td>
</tr>
<tr>
<td>108</td>
<td>87</td>
<td>136</td>
<td>139</td>
</tr>
<tr>
<td>87</td>
<td>90</td>
<td>139</td>
<td>128</td>
</tr>
<tr>
<td>90</td>
<td>101</td>
<td>128</td>
<td>120</td>
</tr>
<tr>
<td>101</td>
<td>110</td>
<td>120</td>
<td>118</td>
</tr>
<tr>
<td>110</td>
<td>111</td>
<td>118</td>
<td>113</td>
</tr>
<tr>
<td>111</td>
<td>105</td>
<td>113</td>
<td>118</td>
</tr>
<tr>
<td>105</td>
<td>121</td>
<td>118</td>
<td>117</td>
</tr>
<tr>
<td>121</td>
<td>112</td>
<td>117</td>
<td>114</td>
</tr>
<tr>
<td>124</td>
<td>121</td>
<td>144</td>
<td>144</td>
</tr>
<tr>
<td>121</td>
<td>94</td>
<td>144</td>
<td>145</td>
</tr>
<tr>
<td>94</td>
<td>97</td>
<td>145</td>
<td>130</td>
</tr>
<tr>
<td>97</td>
<td>109</td>
<td>130</td>
<td>121</td>
</tr>
<tr>
<td>109</td>
<td>116</td>
<td>121</td>
<td>127</td>
</tr>
<tr>
<td>116</td>
<td>122</td>
<td>127</td>
<td>123</td>
</tr>
<tr>
<td>122</td>
<td>112</td>
<td>123</td>
<td>129</td>
</tr>
<tr>
<td>112</td>
<td>119</td>
<td>129</td>
<td>124</td>
</tr>
<tr>
<td>119</td>
<td>123</td>
<td>124</td>
<td>120</td>
</tr>
</tbody>
</table>
Constructing \textit{h}-scatterplots using the library \texttt{gstat}:

Let’s use only the first three rows of the data set above. Again the distance on the north-south and east-west direction is 1 \textit{m}.

Here are the data with their coordinates:

\begin{verbatim}
   x  y  z
1  0  0  87
2  1  0 100
3  2  0  47
4  3  0 111
5  4  0 124
6  5  0 109
7  6  0  0
8  7  0  98
9  8  0 134
10 9  0 144
11 0  1  77
12 1  1  84
13 2  1  74
14 3  1 108
15 4  1 121
16 5  1 143
17 6  1  91
18 7  1  52
19 8  1 136
20 9  1 144
21 0  2  75
22 1  2  80
23 2  2  83
24 3  2  87
25 4  2  94
26 5  2  99
27 6  2  95
28 7  2  48
29 8  2 139
30 9  2 145
\end{verbatim}
Here are all the pairs separated by distance $h = 1$ north-south and east-west:

\[
\begin{array}{ccc}
  z(s) & z(s+1) \\
  1 & 87 & 77 \\
  2 & 77 & 75 \\
  3 & 100 & 84 \\
  4 & 84 & 80 \\
  5 & 47 & 74 \\
  6 & 74 & 83 \\
  7 & 111 & 108 \\
  8 & 108 & 87 \\
  9 & 124 & 121 \\
  10 & 121 & 94 \\
  11 & 109 & 143 \\
  12 & 143 & 99 \\
  13 & 0 & 91 \\
  14 & 91 & 95 \\
  15 & 98 & 52 \\
  16 & 52 & 48 \\
  17 & 134 & 136 \\
  18 & 136 & 139 \\
  19 & 144 & 144 \\
  20 & 144 & 145 \\
  21 & 87 & 100 \\
  22 & 100 & 47 \\
  23 & 47 & 111 \\
  24 & 111 & 124 \\
  25 & 124 & 109 \\
  26 & 109 & 0 \\
  27 & 0 & 98 \\
  28 & 98 & 134 \\
  29 & 134 & 144 \\
  30 & 77 & 84 \\
  31 & 84 & 74 \\
  32 & 74 & 108 \\
  33 & 108 & 121 \\
  34 & 121 & 143 \\
  35 & 143 & 91 \\
  36 & 91 & 52 \\
  37 & 52 & 136 \\
  38 & 136 & 144 \\
  39 & 75 & 80 \\
  40 & 80 & 83 \\
  41 & 83 & 87 \\
  42 & 87 & 94 \\
  43 & 94 & 99 \\
  44 & 99 & 95 \\
  45 & 95 & 48 \\
  46 & 48 & 139 \\
  47 & 139 & 145 \\
\end{array}
\]

The $h$–scatterplot of $z(s)$ against $z(s + 1)$ is shown below:
Using the library \texttt{gstat} you can produce $h$–scatterplots as follows:

\begin{verbatim}
a <- read.table("h_scatter_1.txt", header=T) #Read the data (page 9 above) coordinates(a) <- ~x+y #Convert the data into spatial data hscat(z~1, a, c(0,1)) #This will produce the $z(s)$ against $z(s+1)$ plot. hscat(z~1, a, c(0,1,1.5, 2,3)) #This will produce 4 different #h-scatterplots with $h=1, 2^{0.5}, 2, 3$.
\end{verbatim}

Here they are:
The variogram

Suppose that the vector $Z$ of the observed values at spatial locations is a realization of a random process $Z(s) : s \in D$. Then intrinsic stationarity is defined as follows:

$$E(Z(s + h) - Z(s)) = 0$$

and

$$\text{Var}(Z(s + h) - Z(s)) = 2\gamma(h)$$

The quantity $2\gamma(h)$ is known as the variogram and is very crucial in geostatistics. The variogram says that differences of variables lagged $h$-apart vary in a way that depends only on $h$ through the length of $h$. This is called isotropic variogram as opposed to anisotropic variogram which depends not only on the length $h$ but also the direction. Because of the assumption of constant mean (no trend) we have $E(Z(s)) = \mu$ and we can write

$$\text{Var}(Z(s + h) - Z(s)) = E((Z(s + h) - Z(s))^2) = 2\gamma(h)$$

Therefore we can use the method of moments estimator for the variogram (also called the classical estimator):

$$2\hat{\gamma}(h) = \frac{1}{N(h)} \sum_{N(h)} (Z(s_i) - Z(s_j))^2,$$

where the sum is over $N(h)$ such that $s_i - s_j = h$.

- The following is the so called semivariogram.

$$\hat{\gamma}(h) = \frac{1}{2N(h)} \sum_{N(h)} (Z(s_i) - Z(s_j))^2,$$

- Important: The values of $C(h), \rho(h), \gamma(h)$ are unaffected if we switch all he $i, j$ subscripts. For example, south-north or north-south. Therefore:

$$C(h) = C(-h), \quad \rho(h) = \rho(-h), \quad \gamma(h) = \gamma(-h)$$

- The table below gives the values of the semivariogram for the values of $h = 1, 2, 3, 4$ meters on the south-north direction.

<table>
<thead>
<tr>
<th>$h$</th>
<th>Variogram</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>312.6</td>
</tr>
<tr>
<td>2</td>
<td>478.6</td>
</tr>
<tr>
<td>3</td>
<td>521.6</td>
</tr>
<tr>
<td>4</td>
<td>654.1</td>
</tr>
</tbody>
</table>
• For our data the variogram is shown below. We observe that it increases as \( h \) increases.

Variogram:

\[
\hat{\gamma}(h) = \frac{1}{2N(\approx h)} \sum_{N(h)} (Z(s_i) - Z(s_j))^2,
\]

In this formula we choose to sum over the pairs that are approximately \( h \) apart.

- The semivariogram now is calculated as follows:

- Our goal is to predict every single point on a dense grid of the area of interest to produce a raster map (reconstruct the surface of the variable of interest). Therefore we want to calculate the variogram for many values of the lagged distance \( (h) \). In our example, what if we want to pair data points which are 1.3 \( m \) apart? Based on the geometry of our data there isn’t such a distance. This is solved if we can specify tolerances for both \( h \) and direction. We can use a tolerance of \( \pm 0.5 \ m \) and \( \pm 20 \) degrees. Using these tolerances we will be able to produce enough pairs.

- The semivariogram now is calculated as follows:

- Same applies for irregular sampling. To compute the variogram when \( h = 100 \ m \) with a tolerance of 20 \( m \) we will pair all the observations that fall between 80 – 120 \( m \).

- It is a good practice to calculate directional variograms (unless we are sure that there is no anisotropy in our data).
Another example:
Using the data below, calculate and plot the semivariogram for the following directions and separation distances \((h)\):

- North-south \((h = 100, 200, 300, 400, 500)\).
- East-west \((h = 100, 200, 300, 400, 500)\).
- Northeast-southwest \((h = 141.42, 282.84, 424.26)\).
- Northwest-southeast \((h = 141.42, 282.84, 424.26)\).

The vertical and horizontal distance between two points is 100 \(m\).
Stationarity

- **Stationarity of order 2:**
  A random function is said to be stationary of order 2 if
  
  1. The expected value of \( Z(s) \) does not depend on point \( s \). Therefore,
     \[
     E(Z(s)) = \mu, \quad \forall \ s
     \]
  2. For each pair of random variables \( Z(s) \) and \( Z(s + h) \) the covariance exists and
depends only on the separation distance \( h \):
     \[
     \text{cov}[Z(s + h), Z(s)] = C(h) = E(Z(s) - \mu)(Z(s + h) - \mu) \Rightarrow
     \]
     \[
     C(h) = E[Z(s)Z(s + h)] - \mu^2
     \]
  3. As a result of the above the following is true:
     \[
     \text{Var}(Z(s)) = E[(Z(s) - \mu)^2] = C(0), \quad \forall \ s
     \]
     \[
     \gamma(h) = \frac{1}{2}E[(Z(s + h) - Z(s))^2] = C(0) - C(h), \quad \forall \ s
     \]
     
     It follows that under the second-order stationarity assumption the covariance and
     the variogram are two equivalent tools for describing the correlation between the
     random variables \( Z(s + h) \) and \( Z(s) \) separated by a distance \( h \).
  4. Another tool is the correlogram that follows directly from the definition \( \rho = \frac{\text{Cov}}{\sigma_1 \sigma_2} \):
     \[
     \rho(h) = \frac{C(h)}{C(0)} = \frac{C(0) - \gamma(h)}{C(0)} = 1 - \frac{\gamma(h)}{C(0)}
     \]

- **Intrinsic stationarity:**
  The second-order stationarity assumes that the covariance exists. But there are cases
  where the covariance and the variance do not exist. However in these cases the vari-
  ogram does exist! Therefore, intrinsic stationarity is defined as follows:

  1. The expectation does not depend on \( s \):
     \[
     E(Z(s)) = \mu, \quad \forall \ s
     \]
  2. For all separation distances \( h \) the variance of the difference \( Z(s + h) - Z(s) \) does
not depend on \( s \):
     \[
     \text{Var}[Z(s + h) - Z(s)] = E[(Z(s + h) - Z(s))^2] = 2\gamma(h)
     \]
     
     We observe that second-order stationarity (a stricter form of stationarity) implies
     intrinsic stationarity but not the other way around!
Properties of the covariance and the variogram

1. Covariance:

Let $Z(s_1), Z(s_2), \ldots, Z(s_n)$ be stationary random variables with expectation $\mu$ and covariance $C(h)$. Let $Y$ be a linear combination of these variables, i.e.

$$Y = \sum_{i=1}^{n} \omega_i Z(s_i)$$

Aside note: we will consider such linear combinations later in kriging.

This linear combination is a random variable itself and must have nonnegative variance

$$\text{Var}(Y) = \sum_{i=1}^{n} \sum_{j=1}^{n} \omega_i \omega_j C(s_i - s_j) \geq 0$$

Therefore the covariance function must ensure that the variance of $Y$ is always nonnegative. Such a function it is called “positive definite” and this is a property that must be possessed by the covariance function.

2. Variogram:

Using the expression

$$\gamma(h) = \frac{1}{2} E [(Z(s + h) - Z(s))^2] = C(0) - C(h), \ \forall \ s$$

the variance of $Y$ in terms of the variogram can be expressed as follows:

$$\text{Var}(Y) = C(0) \sum_{i=1}^{n} \omega_i \sum_{j=1}^{n} \omega_j - \sum_{i=1}^{n} \sum_{j=1}^{n} \omega_i \omega_j \gamma(s_i - s_j)$$

In the case where the variance $C(0)$ does not exist the expression above becomes

$$\text{Var}(Y) = - \sum_{i=1}^{n} \sum_{j=1}^{n} \omega_i \omega_j \gamma(s_i - s_j)$$

This is true on the condition that $\sum_{i=1}^{n} \omega_i = 0$. Again, because the variance of $Y$ must be nonnegative, the variogram must satisfy the property of conditional negative definiteness:

$$\sum_{i=1}^{n} \sum_{j=1}^{n} \omega_i \omega_j \gamma(s_i - s_j) \leq 0$$
**Some other properties:**
In practice, the covariance between two random variables $Z(s + h)$ and $Z(s)$ disappears as $h$ becomes too large.

\[ C(h) \to 0, \text{ when } h \to \infty \]

We set $C(h) = 0$ when $h \geq \alpha$ where $\alpha$ is the so called range (see figure below).

For the semi-variogram using

\[ \gamma(h) = C(0) - C(h), \ \forall \ s \]

we observe that:

\[ \gamma(\infty) = C(0) = Var(Z(s)) \]

The semi-variogram stops increasing beyond a certain distance. The variogram reaches a plateau which is called the “sill”. This is also the variance of the random field $C(0)$. 