

Homework 1: Geostatistics (building concepts)

Lecturer: Georgios P. Karagiannis

georgios.karagiannis@durham.ac.uk

As formative assessment 1, submit the solutions to all the Exercises

Exercise 1. (★) Let $Z = (Z_s)_{s \in \mathbb{R}^d}$ be an intrinsically stationary stochastic process, and let $\gamma : \mathbb{R}^d \rightarrow \mathbb{R}$ be its semivariogram. Assume $a \in \mathbb{R}^n$ s.t. $\sum_{i=1}^n a_i = 0$.

(1) Let $a \in \mathbb{R}^n$ be a vector of constants. Show that

$$\text{Var} \left(\sum_{i=1}^n a_i Z(s_i) \right) = \sum_{i=1}^n \sum_{j=1}^n a_i a_j c_Y(s_i, s_j)$$

where $c_Y(s, t) = E(Y(s)Y(t))$, and $Y_s = Z_s - Z_0$.

(2) Show that

$$c_Y(s, t) = \gamma(s) + \gamma(t) - \gamma(s - t)$$

(3) Show that for all $\forall \{s_1, \dots, s_n\} \subseteq S$ it is

$$\sum_{i=1}^n \sum_{j=1}^n a_i a_j \gamma(s_i - s_j) \leq 0$$

Solution. Assume $0 \in S$, and a random variable $Z(0)$. Let $Y_s = Z_s - Z_0$.

(1) It is

$$\begin{aligned} \text{Var} \left(\sum_{i=1}^n a_i Z(s_i) \right) &= \text{Var} \left(\sum_{i=1}^n a_i Z(s_i) - \overbrace{\sum_{i=1}^n a_i}^{0=} Z(0) \right) = \text{Var} \left(\sum_{i=1}^n a_i Y(s_i) \right) \\ &= \sum_{i=1}^n \sum_{j=1}^n a_i a_j E(Y(s_i)Y(s_j)) = c_Y(s, t) \end{aligned}$$

(2) For $E(Y(s_i)) = 0$ it is

$$\begin{aligned} \gamma(s - t) &= \frac{1}{2} E(Z(s) - Z(0) + Z(t) - Z(0))^2 \\ &= \frac{1}{2} (2\gamma(s) + 2\gamma(t) - 2c_Y(s, t)) \\ \implies c_Y(s, t) &= \gamma(s) + \gamma(t) - \gamma(s - t) \end{aligned}$$

(3) It is

$$\begin{aligned}
0 &\leq \text{Var} \left(\sum_{i=1}^n a_i Z(s_i) \right) = \sum_{i=1}^n \sum_{j=1}^n a_i a_j c_Y(s_i, s_j) \\
&= \sum_{i=1}^n \sum_{j=1}^n a_i a_j (\gamma(s_i) + \gamma(s_j) - \gamma(s_i - s_j)) \\
&= \sum_{i=1}^n a_i \gamma(s_i) \sum_{j=1}^n a_j + \sum_{j=1}^n a_j \gamma(s_j) \sum_{i=1}^n a_i - \sum_{i=1}^n \sum_{j=1}^n a_i a_j \gamma(s_i - s_j)
\end{aligned}$$

hence

$$\sum_{i=1}^n \sum_{j=1}^n a_i a_j \gamma(s_i - s_j) \leq 0$$

Exercise 2. (★) Consider the zero-mean geostatistical process $Z = (Z_s)_{s \in \mathbb{R}^d}$ with a weakly stationary and isotropic covariance function given by

$$c(h) = \begin{cases} \xi^2 (1 + \rho \|h\|) \exp(-\rho \|h\|), & h > 0 \\ \nu^2 + \xi^2, & h = 0 \end{cases}$$

- (1) Compute the semi-variogram for the geostatistical process (Z_s)
- (2) What are the nugget, sill and partial sill for this covariance model? Justify your answer.
- (3) Would the slightly altered covariance function defined below be a good model for spatial data for $\phi > 0$? Justify your answer.

$$c(h) = \begin{cases} \xi^2 (1 + \rho \|h\|) \exp(-\rho \|h\|) + \phi, & h > 0 \\ \nu^2 + \xi^2 + \phi, & h = 0 \end{cases}$$

Solution.

- (1) For all $h \neq 0$, it is

$$\begin{aligned}
\gamma(h) &= c(0) - c(h), \\
&= \nu^2 + \xi^2 - \xi^2 (1 + \rho \|h\|) \exp(-\rho \|h\|) \\
&= \nu^2 + \xi^2 (1 - (1 + \rho \|h\|) \exp(-\rho \|h\|))
\end{aligned}$$

then

$$\gamma(h) = \begin{cases} \nu^2 + \xi^2 (1 - (1 + \rho \|h\|) \exp(-\rho \|h\|)) & h > 0 \\ 0 & h = 0 \end{cases}$$

- (2)

- The sill is the covariance function at distance 0, that is $c(0) = \nu^2 + \xi^2$. Or since analogously, it is $\lim_{\|h\| \rightarrow \infty} \gamma(h)$. So,

$$\begin{aligned} \lim_{\|h\| \rightarrow \infty} (\|h\| \exp(-\rho \|h\|)) &= \lim_{\|h\| \rightarrow \infty} (\|h\| / \exp(\rho \|h\|)) \\ &= \lim_{\|h\| \rightarrow \infty} (\|h\| / \exp(\rho \|h\|)) = \lim_{\|h\| \rightarrow \infty} (\exp(-\rho \|h\|)) = 0 \end{aligned}$$

then

$$\lim_{\|h\| \rightarrow \infty} \gamma(h) = \nu^2 + \xi^2$$

- The nugget effect is the limiting value of the semi-variogram as $h \rightarrow 0$ from above, hence it is $\gamma(h) \rightarrow \nu^2$ as $h \rightarrow 0^+$.
 - The partial sill is the sill minus the nugget and is hence ξ^2 .
- (3) No, it would be unrealistic because if $\phi > 0$ then the covariance is always positive for infinitely large distances h . In practical terms this means that two points will always be correlated however far apart they are, it would be unrealistic.