Spatio-temporal statistics (MATH4341) Michaelmas term, 2024

Homework 1: Point referenced data (building concepts)

Lecturer: Georgios P. Karagiannis georgios.karagiannis@durham.ac.uk

---

Exercise 1. ($\star$) Let $Z = \left(Z\left(s\right) : s \in \mathbb{R}^d\right)$ be an intrinsic random field with $\mathrm{E}\left(Z\left(s\right) - Z\left(t\right)\right) = 0$ and let $\gamma : \mathbb{R}^d \to \mathbb{R}$ be its semivariogram.

(1) Let $a \in \mathbb{R}^n$ be a vector of constants. Consider sites $\left\{s_1, ..., s_n \subseteq \mathbb{R}^d\right\}$ Show that

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

where $c_I\left(s,t\right) = \gamma\left(s - s_0\right) + \gamma\left(t - s_0\right) - \gamma\left(s - t\right)$ at some additional $s_0 \in \mathbb{R}^d$.

(2) Show that for all $n \in \mathbb{N}$, $\left(a_1, ..., a_n\right) \subseteq \mathbb{R}^n$ s.t. $\sum_{i=1}^{n} a_i = 0$, and for all $\left(s_1, ..., s_n\right) \subseteq S^n$, it is

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

Solution. Assume origin $s_0 \in \mathbb{R}^d$ with random $Z\left(s_0\right)$. .

(1) I use $Z\left(s_0\right)$ at some location let's say s_0. It is

$$\begin{aligned}
\mathrm{Var}\left(\sum_{i=1}^{n} a_i Z\left(s_i\right)\right) =& \mathrm{Var}\left(\sum_{i=1}^{n} a_i Z\left(s_i\right) - \overbrace{\sum_{i=1}^{n} a_i Z\left(s_0\right)}^{0=}\right) = \mathrm{Var}\left(\sum_{i=1}^{n} a_i\left(Z\left(s_i\right) - Z\left(s_0\right)\right)\right) \\
=& \sum_{i=1}^{n}\sum_{j=1}^{n} a_i a_j \mathrm{E}\left(\left(Z\left(s_i\right) - Z\left(s_0\right)\right)\left(Z\left(s_j\right) - Z\left(s_0\right)\right)\right)
\end{aligned}$$

Let $c_I\left(s,t\right) = \mathrm{E}\left(\left(Z\left(s_i\right) - Z\left(s_0\right)\right)\left(Z\left(s_j\right) - Z\left(s_0\right)\right)\right)$.

It is

$$\begin{aligned}
\gamma\left(s - t\right) &= \frac{1}{2}\mathrm{E}\left(Z\left(s\right) - Z\left(s_0\right) + Z\left(t\right) - Z\left(s_0\right)\right)^2 \\
&= \frac{1}{2}\left(2\gamma\left(s - s_0\right) + 2\gamma\left(t - s_0\right) - 2c_I\left(s,t\right)\right) \\
\implies c_I\left(s,t\right) &= \gamma\left(s - s_0\right) + \gamma\left(t - s_0\right) - \gamma\left(s - t\right)
\end{aligned}$$

(2) It is

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

hence

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

Exercise 2. ($\star$) Consider the zero-mean random field $Z=\left(Z(s): s \in \mathbb{R}^d\right)$ with covariogram 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 semivariogram for the random field $\left(Z(s): s \in \mathbb{R}^d\right)$
(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

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

(a)

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

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

  then

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

- The nugget effect is the limiting value of the semicovariogram as $h\to 0$ from above, hence it is $\gamma\left(h\right)\to\nu^2$ as $h\to 0^+$.
- The partial sill is the sill minus the nugget and is hence ξ^2.

(b) 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.