

Homework 3: Geostatistics (Change of support)

Lecturer: Georgios P. Karagiannis

georgios.karagiannis@durham.ac.uk

Exercise 1. (★) Assume we wish to estimate the average value in a domain V

$$Z_V = \frac{1}{|V|} \int_V Z(s) \, ds$$

with the average of n sample points $\{s_i; i = 1, \dots, n\}$.

$$\hat{Z} = \frac{1}{n} \sum_{i=1}^n Z(s_i)$$

Show that the estimation variance (or else extension variance)

$$\text{Var}(\hat{Z} - Z_V) = -\frac{1}{n^2} \sum_{i=1}^n \sum_{j=1}^n \gamma(s_i - s_j) + \frac{1}{n|V|} \sum_{i=1}^n \int_V \gamma(s_i - x) \, dx - \frac{1}{|V|^2} \int_{x \in B} \int_{y \in B} \gamma(x - y) \, dx dy$$

Hint:: Consider as known that

$$\text{Cov}(Z(t) - Z(s), Z(v) - Z(u)) = \gamma(t - u) + \gamma(s - v) - \gamma(s - u) - \gamma(t - v)$$

Solution.

Exercise 2. (★) Consider a statistical model which is a stochastic process $(Z_s)_{s \in \mathbb{R}}$ (so s has dimension 1), where $Z(\cdot) \sim \text{GP}(\mu(\cdot), c(\cdot, \cdot))$ with mean function $\mu(s) = 1$ and covariance function $c(s, t) = \exp(-(s - t)^2)$ for any $s \in \mathbb{R}$ and $t \in \mathbb{R}$. Assume there is available a dataset $\{(Z_i, s_i)\}_{i=1}^n$ where $Z_i = Z(s_i)$ and $s_i \in \mathbb{R}$ are point sites.

- (1) Compute the length $|v|$ of the block $v = [a, b] \subset \mathbb{R}$.
- (2) Compute the block mean $\mu(v)$ for some block $v = [a, b] \subset \mathbb{R}$ and point $s \in \mathbb{R}$.
- (3) Compute the block covariance function $c(v, s)$ for some block $v = [a, b] \subset \mathbb{R}$ and point $s \in \mathbb{R}$.
- (4) Compute the block covariance function $c(v, v')$ for some blocks $v = [a, b] \subset \mathbb{R}$ and $v' = [a', b'] \subset \mathbb{R}$.
- (5) Denote $Z = (Z_1, \dots, Z_n)^\top$, and $S = \{s_1, \dots, s_n\}$. Let $v = [a, b] \subset \mathbb{R}$ and $v' = [a', b'] \subset \mathbb{R}$ be two intervals. Compute the joint distribution of $(Z(v), Z(v'), Z)^\top$ as a function of $c(\cdot, \cdot)$, S , v , v' , Z , and $\mu(\cdot)$. What is the name of the distribution and what are the parameter functions defining it?
- (6) (Bayesian Kriging) Compute the predictive stochastic process $[Z(v) | Z]$ at blocks $v = [a, b] \subset \mathbb{R}$ with $|v| > 0$.

Hint-1:: Let $x_1 \in \mathbb{R}^{d_1}$, and $x_2 \in \mathbb{R}^{d_2}$. If

$$\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \sim N_{d_1+d_2} \left(\begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix}, \begin{bmatrix} \Sigma_1 & \Sigma_{21}^\top \\ \Sigma_{21} & \Sigma_2 \end{bmatrix} \right)$$

then it is

$$x_2|x_1 \sim N_{d_2} (\mu_{2|1}, \Sigma_{2|1})$$

where

$$\mu_{2|1} = \mu_2 + \Sigma_{21}\Sigma_1^{-1}(x_1 - \mu_1) \quad \text{and} \quad \Sigma_{2|1} = \Sigma_2 - \Sigma_{21}\Sigma_1^{-1}\Sigma_{21}^\top$$

Hint-2: You can use that $\int \operatorname{erf}(x) \, dx = x \operatorname{erf}(x) + \frac{\exp(-x^2)}{\sqrt{\pi}} + \text{const}$, when $\operatorname{erf}(x) = \frac{2}{\sqrt{\pi}} \int_0^x \exp(-t^2) \, dt$.

Solution.
