

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. It is

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

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.

- (1) It is $|v| = b - a$

(a) It is

$$\mu(v) = \mu([a, b]) = \frac{1}{|v|} \int_v \mu(s) ds = \frac{1}{|v|} \int_v 1 ds = \frac{1}{|v|} |v| = 1$$

(b) It is

$$\begin{aligned} c(v, s) &= \frac{1}{|v|} \int_v c(t, s) dt = \frac{1}{b-a} \int_a^b \exp(-(t-s)^2) dt \\ &= \frac{1}{b-a} \frac{\sqrt{\pi}}{2} \int_a^b \frac{2}{\sqrt{\pi}} \exp(-(t-s)^2) dt = \frac{1}{b-a} \frac{\sqrt{\pi}}{2} \int_{a-s}^{b-s} \frac{2}{\sqrt{\pi}} \exp(-\xi^2) d\xi \\ &= \frac{1}{b-a} \frac{\sqrt{\pi}}{2} \int_0^{b-s} \frac{2}{\sqrt{\pi}} \exp(-\xi^2) d\xi - \frac{1}{b-a} \frac{\sqrt{\pi}}{2} \int_0^{a-s} \frac{2}{\sqrt{\pi}} \exp(-\xi^2) d\xi \\ &= \frac{1}{b-a} \frac{\sqrt{\pi}}{2} \operatorname{erf}(b-s) - \frac{1}{b-a} \frac{\sqrt{\pi}}{2} \operatorname{erf}(a-s) \end{aligned}$$

(c) It is

$$\begin{aligned}
c(v, v') &= \frac{1}{|v'|} \frac{1}{|v|} \int_{a'}^{b'} \int_a^b c(t, s) dt ds = \frac{1}{|v'|} \frac{1}{|v|} \int_{a'}^{b'} \left[\int_a^b c(t, s) dt \right] ds = \frac{1}{b' - a'} \int_{a'}^{b'} c(v, s) ds \\
&= \frac{1}{b' - a'} \int_{a'}^{b'} \left(\frac{1}{b - a} \frac{\sqrt{\pi}}{2} \operatorname{erf}(b - s) - \frac{1}{b - a} \frac{\sqrt{\pi}}{2} \operatorname{erf}(a - s) \right) ds \\
&= \frac{1}{b' - a'} \frac{1}{b - a} \frac{\sqrt{\pi}}{2} \int_{a'}^{b'} \operatorname{erf}(b - s) ds - \frac{1}{b' - a'} \frac{1}{b - a} \frac{\sqrt{\pi}}{2} \int_{a'}^{b'} \operatorname{erf}(a - s) ds \\
&= \frac{1}{b' - a'} \frac{1}{b - a} \frac{\sqrt{\pi}}{2} (-1) \int_{b-a'}^{b-b'} \operatorname{erf}(\xi) d\xi - \frac{1}{b' - a'} \frac{1}{b - a} \frac{\sqrt{\pi}}{2} (-1) \int_{a-a'}^{a-b'} \operatorname{erf}(\xi) d\xi \\
&= \frac{1}{b' - a'} \frac{1}{b - a} \frac{\sqrt{\pi}}{2} (-1) \left[\xi \operatorname{erf}(\xi) + \frac{\exp(-\xi^2)}{\sqrt{\pi}} \right]_{b-a'}^{b-b'} - \frac{1}{b' - a'} \frac{1}{b - a} \frac{\sqrt{\pi}}{2} (-1) \left[\xi \operatorname{erf}(\xi) + \frac{\exp(-\xi^2)}{\sqrt{\pi}} \right]_{a-a'}^{a-b'} \\
&= -\frac{1}{b' - a'} \frac{1}{b - a} \frac{\sqrt{\pi}}{2} (b - b') \operatorname{erf}(b - b') - \frac{1}{b' - a'} \frac{1}{b - a} \frac{1}{2} (b - b') \exp(-(b - b')^2) \\
&\quad + \frac{1}{b' - a'} \frac{1}{b - a} \frac{\sqrt{\pi}}{2} (b - a') \operatorname{erf}(b - a') + \frac{1}{b' - a'} \frac{1}{b - a} \frac{1}{2} (b - a') \exp(-(b - a')^2) \\
&\quad + \frac{1}{b' - a'} \frac{1}{b - a} \frac{\sqrt{\pi}}{2} (a - b') \operatorname{erf}(a - b') + \frac{1}{b' - a'} \frac{1}{b - a} \frac{1}{2} (a - b') \exp(-(a - b')^2) \\
&\quad - \frac{1}{b' - a'} \frac{1}{b - a} \frac{\sqrt{\pi}}{2} (a - a') \operatorname{erf}(a - a') - \frac{1}{b' - a'} \frac{1}{b - a} \frac{1}{2} (a - a') \exp(-(a - a')^2)
\end{aligned}$$

(d) It is

$$\begin{bmatrix} Z(v) \\ Z(v') \\ Z \end{bmatrix} \sim N \left(\begin{bmatrix} \mu(v) \\ \mu(v') \\ \mu(S) \end{bmatrix}, \begin{bmatrix} c(v, v) & c(v, v') & c(v, S) \\ c(v', v) & c(v', v') & c(v', S) \\ c(S, v) & c(S, v') & c(S, S) \end{bmatrix} \right)$$

(e) Taking a better look at part 1, I can see

$$\begin{bmatrix} \begin{bmatrix} Z(v_1) \\ Z(v_2) \\ [Z] \end{bmatrix} \end{bmatrix} \sim N \left(\begin{bmatrix} \begin{bmatrix} \mu(v) \\ \mu(v_2) \\ [\mu(S)] \end{bmatrix} \end{bmatrix}, \begin{bmatrix} \begin{bmatrix} c(v_1, v_1) & c(v_1, v_2) \\ c(v_2, v_1) & c(v_2, v_2) \\ c(S, v_1) & c(S, v_2) \end{bmatrix} & \begin{bmatrix} c(v_1, S) \\ c(v_2, S) \\ c(S, S) \end{bmatrix} \end{bmatrix} \right)$$

From the hint I can see, I can see that

$$\begin{bmatrix} Z(v_1) \\ Z(v_2) \end{bmatrix} | Z \sim N(\mu^\dagger, C^\dagger)$$

with

$$C^\dagger = \begin{bmatrix} C_{11}^\dagger & C_{12}^\dagger \\ C_{21}^\dagger & C_{22}^\dagger \end{bmatrix} = \begin{bmatrix} c(v_1, v_1) & c(v_1, v_2) \\ c(v_2, v_1) & c(v_2, v_2) \end{bmatrix} - \begin{bmatrix} c(v_1, S) \\ c(v_2, S) \end{bmatrix} [c(S, S)]^{-1} \begin{bmatrix} c(S, v_1) & c(S, v_2) \end{bmatrix}$$

and

$$\mu^\dagger = \mu(v_1) + c(v_1, S) [c(S, S)]^{-1} (Z - \mu(S))$$

As this is consistent for any vector of blocks with any size, not only $V = \{v_1, v_2\}$, but also $V = \{v_1, v_2, \dots, v_q\}$ then the predictive stochastic process is a Gaussian Process

$$Z(\cdot) | Z \sim \text{GP}(\mu^*(\cdot), c^*(\cdot, \cdot))$$

with mean function at block v

$$\mu^*(v) = \mu(v) + c(v, S) [c(S, S)]^{-1} (Z - \mu(S))$$

by looking at μ^\dagger , and with covariance function at any pair of blocks v and v'

$$c^*(v, v') = c(v, v') - c(v, S) [c(S, S)]^{-1} c(S, v')$$

looking at any off-diagonal element of C^\dagger e.g. the $(1, 2)$ element marked in red.
