Spatio-temporal statistics (MATH4341) Michaelmas term, 2023

Homework 3: Geostatistics (Change of support)

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

---

Exercise 1. ($\star$) Suppose a large volume V is partitioned into n smaller units v of equal size. Show that the dispersion variance $\sigma^2(v|V) = \frac{1}{n}\sum_{j=1}^{n}\sigma_E^2(v_j, V)$ can be written in term of variogram integrals

$$\bar{\gamma}(v, V) = \frac{1}{|v|\,|V|}\int_{s\in v}\int_{s\in V}\gamma\left(s-s'\right)\mathrm{d}s\mathrm{d}s'$$

as

$$\sigma^2(v|V) = \bar{\gamma}(V,V) - \bar{\gamma}(v,v)$$

Solution.

$$\begin{aligned}
\sigma^2(v|V) =& \frac{1}{n}\sum_{j=1}^{n}\sigma_E^2(v_j, V)\\
=& \frac{1}{n}\sum_{j=1}^{n}\left[2\bar{\gamma}(v_j,V) - \bar{\gamma}(v_j,v_j) - \bar{\gamma}(V,V)\right]\\
=& \frac{2}{n}\sum_{j=1}^{n}\bar{\gamma}(v_j,V) - \frac{1}{n}\sum_{j=1}^{n}\bar{\gamma}(v_j,v_j) - \frac{1}{n}\sum_{j=1}^{n}\bar{\gamma}(V,V)\\
=& \frac{2}{n}\sum_{j=1}^{n}\frac{1}{|v_j|\,|V|}\int_{s\in v_j}\int_{s\in V}\gamma\left(s-s'\right)\mathrm{d}s\mathrm{d}s'\\
&- \frac{1}{n}\sum_{j=1}^{n}\bar{\gamma}(v_j,v_j) - \frac{1}{n}\sum_{j=1}^{n}\bar{\gamma}(V,V) \ \text{(but all } v_j \text{ are of the same size as } v\text{)}\\
=& 2\frac{1}{n\,|v|\,|V|}\sum_{j=1}^{n}\int_{s\in v_j}\int_{s\in V}\gamma\left(s-s'\right)\mathrm{d}s\mathrm{d}s' - \bar{\gamma}(v,v) - \bar{\gamma}(V,V)\\
=& 2\frac{1}{\underbrace{n\,|v|}_{=|V|}\,|V|}\underbrace{\sum_{j=1}^{n}\int_{s\in v_j}}_{\int_{s\in V}}\int_{s\in V}\gamma\left(s-s'\right)\mathrm{d}s\mathrm{d}s' - \bar{\gamma}(v,v) - \bar{\gamma}(V,V)\\
=& 2\frac{1}{n}\sum_{j=1}^{n}\bar{\gamma}(V,V) - \bar{\gamma}(v,v) - \bar{\gamma}(V,V) = \quad \bar{\gamma}(V,V) - \bar{\gamma}(v,v)
\end{aligned}$$

Exercise 2. ($\star$) Consider a statistical model which is a stochastic process $(Z_s)_{s\in\mathbb{R}}$ (so s has dimension 1), where $Z(\cdot) \sim \mathrm{GP}(\mu(\cdot), c(\cdot,\cdot))$ with mean function $\mu(s) = 1$ and covariance function

$c(s,t)=\exp\left(-(s-t)^2\right)$ 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 $Z=(Z_1,...,Z_n)^{\top}$, and $S=\{s_1,...,s_n\}$. 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

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

then it is

$$x_2|x_1\sim \mathrm{N}_{d_2}\left(\mu_{2|1},\Sigma_{2|1}\right)$$

where

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

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

Solution.

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

(2) It is

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

(3) It is

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

(4) It is

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

(5) It is

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

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

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

From the hint I can see, I can see that

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

with

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

and

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

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

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

with mean function at block v

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

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

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

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