Revision sheet

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

Part 1. Point referenced data / Geostatistics

Exercise 1. Consider the Gaussian c.f. $c(h) = \sigma^2 \exp\left(-\beta \|h\|_2^2\right)$ for $\sigma^2, \beta > 0$ and $h \in \mathbb{R}^d$. Compute the spectral density from Bochner's theorem

Solution. It is

$$
\begin{aligned}
f(\omega) &= \left(\frac{1}{2\pi}\right)^d \int_{\mathbb{R}^d} \exp\left(-i\omega^\top h\right) \sigma^2 \exp\left(-\beta \|h\|_2^2\right) \mathrm{d}h \\
&= \sigma^2 \left(\frac{1}{2\pi}\right)^d \prod_{j=1}^d \int_{\mathbb{R}} \exp\left(-i\omega_j h_j - \beta h_j^2\right) \mathrm{d}h \\
&= \sigma^2 \left(\frac{1}{2\pi}\right)^d \prod_{j=1}^d \int_{\mathbb{R}} \exp\left(-\beta \left(h_j - \left(-i\omega / (2\beta)\right)\right)^2\right) \mathrm{d}h_j \\
&= \sigma^2 \left(\frac{1}{4\pi\beta}\right)^{d/2} \exp\left(-\|\omega\|_2^2 / (4\beta)\right)
\end{aligned}
$$

Exercise 2. Let $(Z_s)_{s\in\mathcal{S}}$ be a specified statistical model. Assume that process $(Z_s)_{s\in\mathcal{S}}$ has known mean $\mu(s) = \mathrm{E}(Z(s))$ and known covariance function $c(\cdot,\cdot)$. Assume there is available a dataset $\{(s_i, Z_i := Z(s_i))\}_{i=1}^n$. Assume that the matrix C such as $[C]_{i,j} = c(s_i, s_j)$ has an inverse. Consider the "Kriging" estimator μ_{SK} Consider the "Kriging" estimator $Z_{\text{SK}}(s_0)$ of $Z(s_0)$ at an unseen spatial location s_0 as the BLUE (Best Linear Unbiased Estimator)

$$
Z_{\text{SK}}(s_0) = w_{n+1} + \sum_{i=1}^n w_i Z(s_i) = w_{n+1} + w^\top Z,
$$

for some unknown $\{w_i\}$ that we need to learn, and $Z = (Z_1, ..., Z_n)^\top$. Let $w = (w_1, ..., w_n)^\top$.

(1) Find sufficient conditions on $w = (w_1, ..., w_n)^\top$ so that the Kriging estimator $Z_{\text{SK}}(s_0)$ to be unbiased.

(2) Derive the MSE of $Z_{\text{SK}}(s_0)$ as

$$
\mathrm{E}\left(Z_{\text{SK}}(s_0) - Z(s_0)\right)^2 = w^\top C w + c(s_0, s_0) - 2w^\top C_0
$$

where C_0 is a vector such as $[C_0]_i = c(s_0, s_i)$.

(3) Derive the Kriging estimator of $Z(s_0)$ as

$$Z_{\text{SK}}(s_0) = \mu(s_0) + C_0^\top C^{-1}\left[Z - \mu(s_{1:n})\right]$$

where $\mu(s_{1:n})$ is a vector such as $[\mu(s_{1:n})]_i = \mu(s_i)$.

(4) Compute the Kriging standard error $\sigma_{\text{SK}} = \sqrt{\text{E}\left(Z_{\text{SK}}(s_0) - Z(s_0)\right)^2}$.

Solution. The method is called Simple Kriging, and hence we denote it as SK.

(1) It is

$$Z_{\text{SK}}(s_0) = w_{n+1} + \sum_{i=1}^{n} w_i Z(s_i) = w_{n+1} + w^\top Z,$$

where $\{w_i\}$ is a set of unknown weights to be learned.

We assume that assume zero systematic error (unbiasness), hence

$$\text{E}\left(Z_{\text{SK}}(s_0) - Z(s_0)\right) = \text{E}\left(w_{n+1} + \sum_{i=1}^{n} w_i Z(s_i) - Z(s_0)\right) = w_{n+1} + \sum_{i=1}^{n} w_i \mu(s_i) - \mu(s_0)$$

which is satisfied given the assumption

$$w_{n+1} = \mu(s_0) - \sum_{i=1}^{n} w_i \mu(s_i) \iff w_{n+1} = \mu(s_0) - w^\top \mu(s_{1:n})$$

where $w = (w_1, ..., w_n)^\top$.

(2) It is

$$\begin{aligned}
\text{E}\left(Z_{\text{SK}}(s_0) - Z(s_0)\right)^2 =& \text{Var}\left(Z_{\text{SK}}(s_0) - Z(s_0)\right) = \text{Var}\left(w_{n+1} + w^\top Z - Z(s_0)\right) \\
=& \text{Var}\left(w_{n+1} + w^\top Z\right) + \text{Var}\left(Z(s_0)\right) - 2\text{Cov}\left(w_{n+1} + w^\top Z, Z(s_0)\right) \\
=& w^\top C w + c(s_0, s_0) - 2w^\top \text{Cov}\left(Z, Z(s_0)\right) \\
=& w^\top C w + c(s_0, s_0) - 2w^\top C_0
\end{aligned}$$

where $C_0 = \text{Cov}\left(Z, Z(s_0)\right)$, i.e. $[C_0]_j = c(s_j, s_0)$.

(3) To learn the unknown weights $\{w_i\}$ we need to solve

$$w^{\text{SK}} = \arg\min_{w} \text{E}\left(Z_{\text{SK}}(s_0) - Z(s_0)\right)^2, \quad \text{subject to } \; w_{n+1} = \mu(s_0) - w^\top \mu(s_{1:n})$$

As $\text{E}\left(\mu_{\text{SK}} - Z(s_0)\right)^2$ does not depend on w_{n+1} we minimize

$$\begin{aligned}
0 &= \nabla_w \text{E}\left(Z_{\text{SK}}(s_0) - Z(s_0)\right)^2 = \nabla_w \text{Var}\left(Z_{\text{SK}}(s_0) - Z(s_0)\right) \\
&= 2Cw - 2C_0
\end{aligned}$$

So I get

$$w_{\text{SK}} = C^{-1} C_0$$

So

$$
\begin{aligned}
Z_{\mathrm{SK}}(s_0) &= w_{n+1} + C^{-1}C_0 Z \\
&= \mu(s_0) - \left(C^{-1}C_0\right)^\top \mu(s_{1:n}) + \left(C^{-1}C_0\right)^\top Z \\
&= \mu(s_0) + C_0^\top C^{-1}\left[Z - \mu(s_{1:n})\right]
\end{aligned}
$$

(4) It is

$$
\begin{aligned}
\sigma_{\mathrm{SK}} &= \sqrt{\mathrm{E}\left(Z_{\mathrm{SK}}(s_0) - Z(s_0)\right)^2} \\
&= \sqrt{w_{\mathrm{SK}}^\top C w_{\mathrm{SK}} + c(s_0, s_0) - 2w_{\mathrm{SK}}^\top C_0} \\
&= \sqrt{c(s_0, s_0) - C_0^\top C^{-1} C_0}
\end{aligned}
$$

Part 2. Aerial unit data / spatial data on lattices

Exercise 3. Show that the conditionals $x|y \sim \mathrm{N}\left(a+by, \sigma^2+\tau^2y^2\right)$ and $y|x \sim \mathrm{N}\left(c+dx, \tilde{\sigma}^2+\tilde{\tau}^2x^2\right)$ are compatible if $\tau^2 = \tilde{\tau}^2 = 0$, $d/\tilde{\sigma}^2 = b/\sigma^2$, and $|db| < 1$. In particular see what happens if $x|y \sim \mathrm{N}\left(y, \sigma^2\right)$ and $y|x \sim \mathrm{N}\left(x, \sigma^2\right)$ namely if $\tau^2 = \tilde{\tau}^2 = 0$, $d/\tilde{\sigma}^2 = b/\sigma^2$, $\tilde{\sigma}^2 = \sigma^2$ and $d = b = 1$.

Solution. It is

$$
\begin{aligned}
\frac{g(x|y)}{q(y|x)} =& \frac{\mathrm{N}\left(x|a+by, \sigma^2+\tau^2y^2\right)}{\mathrm{N}\left(y|c+dx, \tilde{\sigma}^2+\tilde{\tau}^2x^2\right)} \\
=& \frac{\sqrt{\tilde{\sigma}^2+\tilde{\tau}^2x^2}}{\sqrt{\sigma^2+\tau^2y^2}} \exp\left(-\frac{1}{2}\left(\frac{(x-a-by)^2}{\sigma^2+\tau^2y^2} - \frac{(y-c-dx)^2}{\tilde{\sigma}^2+\tilde{\tau}^2x^2}\right)\right) \text{ (set } \tau^2 = \tilde{\tau}^2 = 0\text{)} \\
=& \frac{\sqrt{\tilde{\sigma}^2}}{\sqrt{\sigma^2}} \exp\left(-\frac{1}{2}\left(\frac{(x-a-by)^2}{\sigma^2} - \frac{(y-c-dx)^2}{\tilde{\sigma}^2}\right)\right) \\
=& \frac{\sqrt{\tilde{\sigma}^2}}{\sqrt{\sigma^2}} \exp\left(-\frac{1}{2}\left(\frac{x^2}{\sigma^2} + \frac{a^2}{\sigma^2} + \frac{b^2y^2}{\sigma^2} - \frac{2xa}{\sigma^2} - 2\frac{xby}{\sigma^2} + \frac{2aby}{\tilde{\sigma}^2}\right.\right. \\
& \left.\left. - \frac{y^2}{\tilde{\sigma}^2} - \frac{c^2}{\tilde{\sigma}^2} - \frac{d^2x^2}{\tilde{\sigma}^2} + \frac{2yc}{\tilde{\sigma}^2} + 2\frac{ydx}{\tilde{\sigma}^2} - \frac{2cdx}{\tilde{\sigma}^2}\right)\right)
\end{aligned}
$$

If $d/\tilde{\sigma}^2 = b/\sigma^2$ (and $\tau^2 = \tilde{\tau}^2 = 0$)

$$\frac{g(x|y)}{q(y|x)} \propto \underbrace{\exp\left(-\frac{1}{2}\left(\left(\frac{1}{\sigma^2} - \frac{d^2}{\tilde{\sigma}^2}\right)x^2 - 2\left(\frac{a}{\sigma^2} + \frac{cd}{\tilde{\sigma}^2}\right)x\right)\right)}_{u(x)}$$

$$\times \underbrace{\exp\left(+\frac{1}{2}\left(\left(\frac{1}{\tilde{\sigma}^2} - \frac{b^2}{\sigma^2}\right)y^2 - 2\left(\frac{c}{\tilde{\sigma}^2} + \frac{ab}{\sigma^2}\right)x\right)\right)}_{v(y)}$$

for $N_g = N_q = N = \mathbb{R}$. Also it is

$$\int u(x)\,\mathrm{d}x =$$

$$\int_{\mathbb{R}} \exp\left(-\frac{1}{2}\left(\left(\frac{1}{\sigma^2} - \frac{d^2}{\tilde{\sigma}^2}\right)x^2 - 2\left(\frac{a}{\sigma^2} + \frac{cd}{\tilde{\sigma}^2}\right)x\right)\right)\mathrm{d}x < \infty$$

when $|db| < 1$.

If $\tau^2 = \tilde{\tau}^2 = 0$, $d/\tilde{\sigma}^2 = b/\sigma^2$, $\tilde{\sigma}^2 = \sigma^2$ and $d = b = 1$, then

$$\frac{g(x|y)}{q(y|x)} \propto \exp\left(-\frac{1}{2}\left(\left(\frac{1}{\sigma^2} - \frac{1}{\sigma^2}\right)x^2\right)\right)\exp\left(-\frac{1}{2}\left(\left(\frac{1}{\sigma^2} - \frac{1}{\sigma^2}\right)y^2\right)\right) \propto \text{const}$$

that is $u(x)$ is constant and hence $\int u(x)\,\mathrm{d}x = \infty$ implying that they are not compatible.

Exercise 4. Consider that $Z(s)$ represents presence or absence of a characteristic at location $s \in \mathcal{S}$. Mathematically, assume random field Z taking values on a set of indices $\mathcal{S}$ in $\mathcal{Z} = \{0,1\}$ on $\mathcal{S} = \{1,...,n\}$, $n \in \mathbb{N} - \{0\}$. Consider that for a given z_{-i} it is

$$\begin{cases} z_i|z_{-i} & \sim \text{Logit}\left(\theta_i\left(z_{-i}\right)\right), \quad i \in \mathcal{S} \\ \theta_i\left(z_{-i}\right) & = \alpha_i + \sum_{j:j\sim i} \beta_{i,j} z_j \end{cases}$$

(1) Show that the conditionals $z_i|z_{-i}$ are compatible as a Besag's auto-model when $\{\alpha_i\}$ and $\{\beta_{i,j}\}$ satisfy certain conditions, and specify these conditions.

Hint: The PMF of Logistic distribution $x|\theta \sim \text{Logit}(\theta)$ can be written as

$$\Pr(x|\theta) = \frac{\exp(x\theta)}{1 + \exp(\theta)} 1(x \in \{0,1\})$$

(2) Write down the marginal distribution of the associated random field.

(3) What would be the sign of $\{\beta_{i,j}\}$ if you wish to introduce competition between neighboring sites? What would be the sign of $\{\beta_{i,j}\}$ if you wish to introduce similarity between neighboring sites? What does α_i represent when $\beta_{i,j} = 0$?

Solution.

(1) Then the characteristics are

$$\Pr_i (z_i|z_{-i}) = \frac{\exp(z_i\theta_i(z_{-i}))}{1+\exp(\theta_i(z_{-i}))} 1(z_i \in \{0,1\})$$

Now, we have parameterized $\{\theta_i(\cdot)\}$ as

$$\theta_i(z_{-i}) = \alpha_i + \sum_{j:j\sim i} \beta_{i,j} z_j$$

for $\{\alpha_i\}$ and $\{\beta_{i,j}\}$ with $\beta_{i,j} = \beta_{j,i}$. Hence, we've got

$$\log\left(\Pr_i (z_i|z_{-i})\right) = \underbrace{z_i}_{B_i(z_i)} \underbrace{\left(\alpha_i + \sum_{j\sim i} \beta_{i,j} \overbrace{z_j}^{B_i(z_j)}\right)}_{A_i(z_{-i})} + \underbrace{0}_{C_i(z_i)} + \underbrace{\left(-\log\left(1+\exp\left(\alpha_i + \sum_{j:j\sim i} \beta_{i,j} z_j\right)\right)\right)}_{D_i(z_{-i})}$$

Notice that all the conditionals $z_i|z_{-i}$ follow an Exponential family with

$$\begin{aligned} A_i(z_{-i}) =& \alpha_i + \sum_{j\sim i} \beta_{i,j} B_i(z_j) \\ B_i(z_i) =& z_i \\ C_i(z_i) =& 0 \\ D_i(z_{-i}) =& -\log\left(1+\exp\left(\alpha_i + \sum_{j:j\sim i} \beta_{i,j} z_j\right)\right) \end{aligned}$$

I can get $C_i(\zeta) = 0$ and $D_i(\zeta, ..., \zeta) = 0$ by considering a reference point $\zeta = 0$. Hence the conditionals $z_i|z_{-i}$ are compatible as a Besag's auto-model for any $\{\alpha_i\}$ and $\{\beta_{i,j}\}$ with $\beta_{i,j} = \beta_{j,i}$ according to a Theorem discussed in the Lectures.

(2) The Besag auto-model has marginal distribution

$$\Pr_Z(z) \propto \exp\left(\overbrace{\sum_i \underbrace{\alpha_i \underbrace{z_i}_{B_i(z_i)}}_{V_i(z_i)} + \underbrace{\sum_i \sum_{j:j\sim i}}_{\sum_{\{i,j\}:j\sim i}} \beta_{i,j} z_i z_j}^{U(z)=}\right)$$

according to a Theorem discussed in the Lectures.

(3) I observe that: (1.) the model has spatially dependent coefficients $\{\alpha_i\}$ and $\{\beta_{i,j}\}$. (2.) when $\beta_{i,j} = 0$, for all j such as $j \sim i$, it is $\Pr_i(z_i|z_{-i}) = \frac{\exp(\alpha_i)}{1+\exp(\alpha_i)}$ and (3.) Characteristic's present at site i is encouraged in neighboring site j when $\beta_{i,j} > 0$, and discouraged when $\beta_{i,j} < 0$.