

## Problem class sheet 2

Lecturer &amp; author: Georgios P. Karagiannis

georgios.karagiannis@durham.ac.uk

**Exercise 1.** Let  $(Z_s)_{s \in \mathcal{S}}$  be a specified statistical model. Assume that  $(Z_s)_{s \in \mathcal{S}}$  is weakly stationary with unknown constant mean  $\mu = E(Z(s))$  and known covariogram  $c(\cdot)$ . Assume there is available a dataset  $\{(s_i, Z_i := Z(s_i))\}_{i=1}^n$  and assume they are realizations of  $(Z_s)_{s \in \mathcal{S}}$ . Assume that the matrix  $C$  such as  $[C]_{i,j} = c(\|s_i - s_j\|)$  has an inverse. Consider the “Kriging” estimator  $\mu_{\text{KM}}$  of  $\mu$  as the BLUE (Best Linear Unbiased Estimator)

$$\mu_{\text{KM}} = \sum_{i=1}^n w_i Z(s_i) = w^\top Z,$$

for some unknown  $\{w_i\}$  that we need to learn.

- (1) Find sufficient conditions on  $w = (w_1, \dots, w_n)$  so that the Kriging estimator  $\mu_{\text{KM}}$  to be unbiased.
- (2) Assume  $C$  is invertible. Compute the MSE of  $\mu_{\text{KM}}$  as a function of  $w = (w_1, \dots, w_n)$  and  $C$
- (3) Derive the Kriging estimator  $\mu_{\text{KM}}$  of  $\mu$  as a function of  $C$
- (4) Derive the Kriging standard error as  $\sigma_{\text{KM}} = \sqrt{E(\mu_{\text{KM}} - \mu)^2}$  as a function of  $C$

**Solution 2.** The method is called Kriging the Mean, and hence we denote it as KM.

- (1) It is

$$\mu_{\text{KM}} = \sum_{i=1}^n w_i Z(s_i) = w^\top Z,$$

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

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

$$E(\mu_{\text{KM}} - \mu) = E\left(\sum_{i=1}^n w_i Z(s_i) - \mu\right) = \sum_{i=1}^n w_i \overset{=1}{E(Z(s_i))} - \overset{=\mu}{\mu}$$

which is satisfied given the assumption

$$\sum_{i=1}^n w_i = 1 \iff 1^\top w = 1 \quad (\text{ASSUMPTION})$$

- (2) It is

$$\begin{aligned}
E(\mu_{\text{KM}} - \mu)^2 &= E(\mu_{\text{KM}}^2 + \mu^2 - 2\mu_{\text{KM}}\mu) = \sum_{i=1}^n \sum_{j=1}^n w_i w_j E(Z(s_i) Z(s_j)) - \sum_{i=1}^n w_i \sum_{j=1}^n w_j \mu \\
&= \sum_{i=1}^n \sum_{j=1}^n w_i w_j (c(s_i - s_j) - \mu) = \sum_{i=1}^n \sum_{j=1}^n w_i w_j c(s_i - s_j) = w^\top C w
\end{aligned}$$

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

$$w^{\text{KM}} = \arg \min_w E(\mu_{\text{KM}} - \mu)^2, \text{ subject to } \sum_{i=1}^n w_i = 1$$

The Lagrange function is

$$\begin{aligned}
\mathfrak{L}(w, \lambda) &= \sum_{i=1}^n \sum_{j=1}^n w_i w_j c(s_i - s_j) - 2\lambda \left( \sum_{i=1}^n w_i - 1 \right) \\
&= w^\top C w - 2\lambda (1^\top w - 1)
\end{aligned}$$

The Kriging to mean equations are  $0 = \nabla_{w, \lambda} \mathfrak{L}(w, \lambda)$  producing

$$\begin{cases} 0 = 2 \sum_{j=1}^n w_j^{\text{KM}} c(s_i - s_j) - 2\lambda & \forall i = 1, \dots, n \\ 1 = \sum_{i=1}^n w_i^{\text{KM}} \end{cases}$$

$$\begin{cases} 2Cw^{\text{KM}} - 2\lambda 1 = 0 \\ 1^\top w^{\text{KM}} = 1 \end{cases}$$

Given that  $C^{-1}$  exists, I multiply by  $1^\top C^{-1}$  and I get

$$21^\top C^{-1} C w^{\text{KM}} - 21^\top C^{-1} \lambda 1 = 0$$

so

$$\lambda = \frac{1}{1^\top C^{-1} 1}$$

I substitute and I get

$$w^{\text{KM}} = \frac{C^{-1} 1}{1^\top C^{-1} 1}$$

So

$$\mu_{\text{KM}} = \left( \frac{C^{-1} 1}{1^\top C^{-1} 1} \right)^\top Z$$

(4) It is

$$\sigma_{\text{KM}} = \sqrt{E(\mu_{\text{KM}} - \mu)^2} = \sqrt{\left( \frac{C^{-1} 1}{1^\top C^{-1} 1} \right)^\top C \frac{C^{-1} 1}{1^\top C^{-1} 1}} = \frac{1}{\sqrt{1^\top C^{-1} 1}}$$

**Exercise 3.** Let  $(Z_s)_{s \in \mathcal{S}}$  be a specified statistical model. Assume that  $(Z_s)_{s \in \mathcal{S}}$  is an intrinsic stationary process with unknown constant mean  $\mu(s) = E(Z(s))$  and known semi-variogram  $\gamma(\cdot)$ . Assume there is available a dataset  $\{(s_i, Z_i := Z(s_i))\}_{i=1}^n$ . Consider the “Kriging” estimator  $Z_{\text{OK}}(s_0)$  of  $Z(s_0)$  at any unseen spatial location  $s_0$  as the BLUE (Best Linear Unbiased Estimator)

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

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

- (1) Find sufficient conditions on  $w = (w_1, \dots, w_n)$  so that the Kriging estimator  $Z_{\text{OK}}(s_0)$  to be unbiased.
- (2) Derive the MSE of  $Z_{\text{OK}}(s_0)$  as

$$E(Z_{\text{OK}}(s_0) - Z(s_0))^2 = -w^\top \Gamma w + 2w^\top \gamma_0$$

where  $\gamma_0 = (\gamma(s_0 - s_1), \dots, \gamma(s_0 - s_n))^\top$  and  $\Gamma$  with  $[\Gamma]_{i,j} = \gamma(s_i - s_j)$

- (3) Assume  $\Gamma$  is invertible matrix. Derive the Kriging estimator of  $Z(s_0)$  as

$$Z_{\text{OK}}(s_0) = \Gamma^{-1} \left( \gamma_0 + \frac{1 - 1^\top \Gamma^{-1} \gamma_0}{1^\top \Gamma^{-1} 1} 1 \right) Z$$

- (4) Derive the Kriging standard error of  $Z_{\text{OK}}(s_0)$  as

$$\sigma_{\text{SK}} = \sqrt{\gamma_0^\top \Gamma^{-1} \gamma_0 - \frac{(1 - 1^\top \Gamma^{-1} \gamma_0)^2}{1^\top \Gamma^{-1} 1}}$$

**Solution.** The method is called Ordinary Kriging, and hence we denote it as OK.

- (1) It is

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

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

$$E(Z_{\text{OK}}(s_0)) = w_{n+1} + \sum_{i=1}^n w_i E(Z(s_i)) \Leftrightarrow \mu = w_{n+1} + \mu \sum_{i=1}^n w_i$$

Unbiasness is satisfied given the assumption  $w_{n+1} = 0$ , and

$$\sum_{i=1}^n w_i = 1 \Leftrightarrow 1^\top w = 1 \quad (\text{ASSUMPTION})$$

(2) The MSE of  $Z_{\text{OK}}(s_0)$  is

$$\begin{aligned}
\text{MSE}(Z_{\text{OK}}(s_0)) &= \mathbb{E}(Z_{\text{OK}}(s_0) - Z(s_0))^2 = \mathbb{E}\left(\sum_{i=1}^n w_i Z(s_i) - \underbrace{\sum_{i=1}^n w_i}_{=1} Z(s_0)\right)^2 \\
&= \mathbb{E}\left(\sum_{i=1}^n w_i (Z(s_i) - Z(s_0))\right)^2 \\
&= -\mathbb{E}\left(\frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n w_i w_j (Z(s_i) - Z(s_j))^2 - 2 \sum_{i=1}^n \frac{1}{2} w_i (Z(s_i) - Z(s_0))^2\right) \\
&= -\sum_{i=1}^n w_i \sum_{j=1}^n w_j \frac{1}{2} \mathbb{E}(Z(s_i) - Z(s_j))^2 + 2 \sum_{i=1}^n w_i \frac{1}{2} \mathbb{E}(Z(s_i) - Z(s_0))^2 \\
&= -\sum_{i=1}^n w_i \sum_{j=1}^n w_j \gamma(s_i - s_j) + 2 \sum_{i=1}^n w_i \gamma(s_i - s_0) \\
&= -w^\top \mathbf{\Gamma} w + 2w^\top \boldsymbol{\gamma}_0
\end{aligned}$$

where  $w = (w_1, \dots, w_n)^\top$ ,  $\boldsymbol{\gamma}_0 = (\gamma(s_0 - s_1), \dots, \gamma(s_0 - s_n))^\top$ , and  $[\mathbf{\Gamma}]_{i,j} = \gamma(s_i - s_j)$ .

(3) The Lagrange multiplier function to minimize the MSE under the assumption is

$$\begin{aligned}
\mathfrak{L}(w, \lambda) &= -\sum_{i=1}^n w_i w_j \gamma(s_i - s_j) + 2 \sum_{i=1}^n w_i \gamma(s_0 - s_i) - \lambda \left(\sum_{i=1}^n w_i - 1\right) \\
&= -w^\top \mathbf{\Gamma} w + 2w^\top \boldsymbol{\gamma}_0 - \lambda (1^\top w - 1)
\end{aligned}$$

The OK system of equations is  $0 = \nabla_{(\{w_i\}, \lambda)} L(w, \lambda)|_{(w, \lambda)}$  producing

$$\begin{cases} 0 = -2 \sum_{j=1}^n w_j^{\text{OK}} \gamma(s_i - s_j) + 2\gamma(s_0 - s_i) - \lambda, & i = 1, \dots, n \\ 1 = \sum_{i=1}^n w_i^{\text{OK}} \end{cases} \iff$$

$$\begin{cases} 0 = -2\mathbf{\Gamma} w_{\text{OK}} + 2\boldsymbol{\gamma}_0 - \lambda_{\text{OK}} 1 \\ 1 = 1^\top w_{\text{OK}} \end{cases}$$

Assuming  $\mathbf{\Gamma}$  is invertable and multiplying by  $1^\top \mathbf{\Gamma}^{-1}$  it is

$$0 = -2\mathbf{\Gamma} w_{\text{OK}} + 2\boldsymbol{\gamma}_0 - \lambda_{\text{OK}} 1 \iff$$

$$0 = -2 \cancel{1^\top \mathbf{\Gamma}^{-1} \mathbf{\Gamma} w_{\text{OK}}} \overset{=1}{=} 1^\top \mathbf{\Gamma}^{-1} \boldsymbol{\gamma}_0 - 1^\top \mathbf{\Gamma}^{-1} \lambda_{\text{OK}} 1 \iff$$

$$\lambda_{\text{OK}} = 2 \frac{1^\top \mathbf{\Gamma}^{-1} \boldsymbol{\gamma}_0 - 1}{1^\top \mathbf{\Gamma}^{-1} 1}$$

By substitution I get

$$w_{\text{OK}} = \mathbf{\Gamma}^{-1} \left( \gamma_0 + \frac{1 - \mathbf{1}^\top \mathbf{\Gamma}^{-1} \gamma_0}{\mathbf{1}^\top \mathbf{\Gamma}^{-1} \mathbf{1}} \mathbf{1} \right)$$

Hence

$$Z_{\text{OK}}(s_0) = w_{\text{OK}} Z = \mathbf{\Gamma}^{-1} \left( \gamma_0 + \frac{1 - \mathbf{1}^\top \mathbf{\Gamma}^{-1} \gamma_0}{\mathbf{1}^\top \mathbf{\Gamma}^{-1} \mathbf{1}} \mathbf{1} \right) Z$$

(4) It is

$$\begin{aligned} \sigma_{\text{OK}}(s_0) &= \sqrt{\text{MSE}(Z_{\text{OK}}(s_0))} \\ &= \sqrt{-w^\top \mathbf{\Gamma} w + w^\top \gamma_0} \\ &= \sqrt{\gamma_0 \mathbf{\Gamma}^{-1} \gamma_0 - \frac{(1 - \mathbf{1}^\top \mathbf{\Gamma}^{-1} \gamma_0)^2}{\mathbf{1}^\top \mathbf{\Gamma}^{-1} \mathbf{1}}} \end{aligned}$$

Note regarding the calculations in MSE:

$$\begin{aligned}
\left( \sum_{i=1}^n w_i (Z(s_i) - Z(s_0)) \right)^2 &= \left( \sum_{i=1}^n w_i (Z_i - Z_0) \right)^2 \\
&= \sum_{i=1}^n w_i^2 (Z_i - Z_0)^2 + 2 \sum_{1 \leq i < j \leq n} w_i (Z_i - Z_0) w_j (Z_j - Z_0) \\
&= \sum_{i=1}^n \sum_{j=1}^n w_i w_j (Z_i - Z_0) (Z_j - Z_0) \\
&= 2 \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n w_i w_j (Z_i - Z_0) (Z_j - Z_0) \\
&\quad - \frac{1}{2} \sum_{i=1}^n w_i (Z_i - Z_0)^2 - \frac{1}{2} \sum_{j=1}^n w_j (Z_j - Z_0)^2 \\
&\quad + 2 \frac{1}{2} \sum_{i=1}^n w_i (Z_i - Z_0)^2 \\
&= - \frac{1}{2} \left( \sum_{i=1}^n w_i \sum_{j=1}^n w_j [(Z_i - Z_0)^2 + (Z_j - Z_0)^2 - 2w_i w_j (Z_i - Z_0) (Z_j - Z_0)] \right) \\
&\quad + 2 \frac{1}{2} \sum_{i=1}^n w_i (Z_i - Z_0)^2 \\
&= - \frac{1}{2} \left( \sum_{i=1}^n w_i \sum_{j=1}^n w_j [(Z_i - Z_0) - (Z_j - Z_0)]^2 \right) + 2 \frac{1}{2} \sum_{i=1}^n w_i (Z_i - Z_0)^2 \\
&= - \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n w_i w_j (Z_i - Z_j)^2 + 2 \frac{1}{2} \sum_{i=1}^n w_i (Z_i - Z_0)^2 \\
&= - \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n w_i w_j (Z(s_i) - Z(s_j))^2 + 2 \frac{1}{2} \sum_{i=1}^n w_i (Z(s_i) - Z(s_0))^2
\end{aligned}$$


---