Experimental variogram of the residuals in the universal kriging (UK) model

Nicolas DESASSIS
Didier RENARD

Technical report

R141210NDES

MINES ParisTech
Centre de Géosciences – Equipe Géostatistique
35, rue Saint-Honoré – 77300 Fontainebleau (France)

MINES ParisTech
Centre de Géosciences – Equipe Géostatistique
35, Rue Saint-Honoré
77300 FONTAINEBLEAU (France)

Tél. +33 1 64 69 47 73
Fax. +33 1 64 69 47 05

Equipe	Géostatistique
Visa	Jacques RIVOIRARD

Experimental variogram of the residuals in the universal kriging (UK) model

N. Desassis & D. Renard

1 Model and notations

Suppose that we observe $\mathbf{y} = (y(x_1), \dots, y(x_n))$ the realization of a random function $Y(.)$ at points $x_1, \dots, x_n$ of a domain $\mathcal{D} \subset \mathbb{R}^d$.

We suppose that for all $x \in \mathcal{D}$,

$$Y(x) = \sum_{i=0}^{p} \beta_i f_i(x) + R(x) \tag{1}$$

where R is an intrinsic random function (IRF) with variogram $\gamma^{(R)}$, the f_i are some functions of the coordinates (e.g polynomial drifts or external drifts) known everywhere (at the data locations and at the prediction sites) and the β_i are some real coefficients. Note that the function f_0 is equal to 1 for all $x \in \mathcal{D}$.

The aim of the methodology is to provide some values for the coefficients β_i and a model for the variogram $\gamma^{(R)}$.

We will note

- $\beta = (\beta_0, \beta_1, \dots, \beta_p)'$ the $(p+1)$-vector of drift coefficients,
- $\mathbf{F}$ the $n \times (p+1)$ matrix with $(i,j)^{\text{th}}$ term $f_{j-1}(x_i)$ (the first column of $\mathbf{F}$ is the vector $\mathbf{1}_n$, which has its n components equal to 1),
- $\mathbf{Y} = (Y(x_1), \dots, Y(x_n))$ the random vector of $Y(.)$ at observation sites
- $\mathbf{R} = (R(x_1), \dots, R(x_n))$ the vector of the intrinsic random function at observation sites,

- $\mathbf{r} = (r(x_1), \ldots, r(x_n))$ the current realization of $\mathbf{R}$, that means

$$r(x_j) = y(x_j) - \sum_{i=0}^{p} \beta_i f_i(x_j).$$

With these notations, the model (**??**) can be written

$$\mathbf{Y} = \mathbf{F}\beta + \mathbf{R} \tag{2}$$

2 The usual way

2.1 Algorithm

The traditional way to determine β and γ is

1. Compute some estimated coefficients $\hat{\beta}$
2. Compute the estimated residuals $\hat{\mathbf{r}} = \mathbf{y} - \mathbf{F}\hat{\beta}$
3. Estimate the experimental variogram of the estimated residuals $\hat{\gamma}$ for different lags h
4. Fit a model γ to the experimental variogram $\hat{\gamma}$.

The point 1. is easily done. In this work, we will consider the least-square estimator:

$$\hat{\beta} = (\mathbf{F}'\mathbf{F})^{-1}\mathbf{F}'\mathbf{y}$$

Then the point 2. becomes

$$\hat{\mathbf{r}} = \mathbf{y} - \mathbf{F}\hat{\beta} = \mathbf{P}\mathbf{y}$$

where

$$\mathbf{P} = \mathbf{I} - \mathbf{M}$$

with $\mathbf{I}$ the identity matrix of size n and

$$\mathbf{M} = \mathbf{F}(\mathbf{F}'\mathbf{F})^{-1}\mathbf{F}'.$$

The point 3. is the computation of the variogram of the residuals for different lags h:

$$\hat{\gamma}^{(\hat{r})}(h) = \frac{1}{N(h)} \sum_{(i,j)\in\mathcal{V}(h)} \hat{\gamma}_{ij}^{(\hat{r})}$$

where

$$\hat{\gamma}_{ij}^{(\hat{r})} = \frac{1}{2}(\hat{r}(x_i) - \hat{r}(x_j))^2,$$

$(i,j) \in \mathcal{V}(h)$ is the set of pairs (i,j) such as $x_i - x_j \simeq h$ and $N(h)$ stands for the number of such pairs.

2.2 Bias of the methodology

A major limitation of this methodology comes from the fact that the experimental variogram computed from the estimated residuals $\hat{\mathbf{r}}$ is biased. In other words, if we compute for a given h the expectation of $\hat{\gamma}^{(\hat{R})}(h)$ the random version of $\hat{\gamma}^{(\hat{r})}(h)$ and if we compare with the target $\gamma^{(R)}(h)$ we have:

$$E[\hat{\gamma}^{(\hat{R})}(h)] \leq \gamma^{(R)}(h).$$

To see that, first note that we can write the estimated residuals as follows:

$$\hat{\mathbf{R}} = \mathbf{PY} = \mathbf{P}(\mathbf{F}\beta + \mathbf{R}) = \mathbf{PR} \tag{3}$$

since $\mathbf{PF} = \mathbf{F} - \mathbf{MF} = \mathbf{F} - \mathbf{F}(\mathbf{F}'\mathbf{F})^{-1}\mathbf{F}'\mathbf{F} = 0$.

Now, let's compute $E[\hat{\gamma}_{ij}^{(\hat{R})}]$ where $\hat{\gamma}_{ij}^{(\hat{R})}$ is the random version of $\hat{\gamma}_{ij}^{(\hat{r})}$.

$$\begin{aligned} E[\hat{\gamma}_{ij}^{(\hat{R})}] &= \frac{1}{2}E[(\hat{R}(x_i) - \hat{R}(x_j))^2] \\ &= \frac{1}{2}\mathrm{Var}(\hat{R}(x_i) - \hat{R}(x_j)) \end{aligned}$$

since $E[\hat{R}(x_i)] = 0$.

Now, let's first remark that

$$\hat{R}(x_i) - \hat{R}(x_j) = (P_{i.} - P_{j.})\mathbf{R},$$

where $P_{i.}$ and $P_{j.}$ are respectively the row vectors of the i^{th} and the j^{th} line of P.

It follows that $\hat{R}(x_i) - \hat{R}(x_j)$ can be seen as an authorized linear combination (ALC) of the elements of $\mathbf{R}$. Indeed, let's compute $S(P_{i.})$, the sum of the components of $P_{i.}$.

$$\begin{aligned} S(P_{i.}) &= 1 - S(M_{i.}) \\ &= 1 - M_{i.}\mathbf{1}_n \\ &= 0 \end{aligned}$$

since $\mathbf{MF} = \mathbf{F}$ and the first column of $\mathbf{F}$ is equal to $\mathbf{1}_n$.
Therefore, we can compute the variance of the ALC.
First, note that (**??**) implies:

$$\begin{aligned} \hat{R}(x_i) - \hat{R}(x_j) &= R(x_i) - M_{i.}\mathbf{R} - (R(x_j) - M_{j.}\mathbf{R}) \\ &= R(x_i) - R(x_j) - \sum_{k=1}^{n} M_{ik}R(x_i) + \sum_{k=1}^{n} M_{jk}R(x_j) \end{aligned}$$

where M_{ij} stands for the $(i,j)^{\text{th}}$ term of the matrix $\mathbf{M}$.
Therefore, we have

$$E[\hat{\gamma}_{ij}^{(\hat{R})}] = \frac{1}{2}\text{Var}(R(x_i) - R(x_j)) + \mathbf{b}_{ij}$$

where $\mathbf{b}_{ij}$, the bias associated to the pair (i,j), can be written:

$$\begin{aligned} \mathbf{b}_{ij} = &\sum_{k=1}^{n}(M_{ik} - M_{jk})(\gamma^{(R)}(h_{ik}) - \gamma^{(R)}(h_{jk})) \\ &+ \sum_{k=1}^{n}\sum_{l=1}^{n}\left(M_{ik}M_{jl} - \frac{M_{ik}M_{il} + M_{jk}M_{jl}}{2}\right)\gamma^{(R)}(h_{kl}) \end{aligned} \tag{4}$$

where h_{kl} stands for the separation vector $x_k - x_l$.
The total bias for the lag h is:

$$\mathbf{B}(h) = \frac{1}{N(h)}\sum_{(i,j)\in\mathcal{V}(h)}\mathbf{b}_{ij}. \tag{5}$$

To compute the total bias associated to the experimental variogram at lag h, we need to know the value of the variogram at all the distances which is exactly the quantity we are trying to compute. To circumvent this difficulty, we use the iterative procedure described in the next section.

3 Bias correction

To perform the bias correction, we propose the following algorithm (inspired from the one of Beckers and Bogaert, 1998):

Algorithm to compute $\hat{\gamma}^{(r)}(h)$

Initialisation:

At step 0 do:

(a) Compute $\hat{\gamma}^{(\hat{r})}(h)$, the experimental variogram of the estimated residuals.

(b) Set $\hat{\gamma}^{(r)}(h) = \hat{\gamma}^{(\hat{r})}(h)$.

Iterations:

At step n do:

(a) Fit a model $\gamma^{(n)}$ on $\hat{\gamma}^{(r)}(h)$.

(b) Compute the bias $\mathbf{B}^{(n)}(h)$ by using equations (**??**) and (**??**) with

$$\gamma^{(R)} = \gamma^{(n)}.$$

(c) Compute the corrected experimental variogram

$$\hat{\gamma}^{(r)} = \hat{\gamma}^{(\hat{r})} - \mathbf{B}^{(n)}(h)$$

Note that step (a) at step n can be performed with the algorithm of Desassis and Renard (2012).

References

Beckers, F. and Bogaert, P. (1998) Nonstationarity of the mean and unbiased variogram estimation: extension of the weighted least-square method. Mathematical Geology 30 (2)

Desassis N, Renard D (2013) Automatic Variogram Modeling by Iterative Least Squares: Univariate and Multivariate Cases. Mathematical Geosciences 45:453–470