Steering Workshop 2020
STATISTICS AND INNOVATION FOR INDUSTRY 4.0

Poster: Empirical variograms

Giovanni Pistone

de Castro Statistics, Collegio Carlo Alberto, Torino

Florence, February 20–21, 2020

web-page: www.giannidiorestino.it
e-mail: giovanni.pistone@carloalberto.org

- Consider a response surface on a given real domain D (usually a rectangle). A measurement is available at each testing points $x \in D$. We want to assess the conformity of the shape of the response surface to some standard. For example: "is the surface bended in some direction?" Or: "Is there a waviness of a type associate to a specific technology?" These are possible defects that cannot be specified in a parametric way.
- A very popular modeling method relies on the assumption that the surface under study is the realization of a random field, for example, a Gaussian random field $(\zeta_x)_{x \in D}$. In such a case, the observed characteristics of the surface will in fact depend on the auto-covariance of the random field.
- More specifically, under the intrincic stationarity assumpion, the properties of the random surface depend on the random field variogram.
- Grazia Vicario has suggested in to use the variogram as a non-parametric method, without assuming any randomness of the surface, but using instead the idea of a sistematic or random sampling of couples of test points on the given domain.

Definition *Let X and Y be independent random variables whose common distribution μ is supported by the domain $D \in \mathbb{R}^n$ and let d be a distance on D. Let $F\colon D \to \mathbb{R}$ be a response function. The empirical variogram γ is defined by*

$$\gamma(d(X,Y)) = \frac{1}{2}\,\mathsf{E}\left(|F(X) - F(Y)|^2 \middle| d(X,Y)\right) .$$

Here are some immediate properties of the empirical variogram.

1. If F is constant, then $\gamma = 0$.
2. The variogram depends quadratically on the gradient ∇F. In fact,

$$F(v+u) - F(v) = \int_0^1 \nabla F(v + \theta u) \cdot u \; d\theta \; .$$

3. If $D =]0,1[$ and X, Y are uniform,

$$\gamma(t) = \frac{1}{2(1-t)} \int_0^{1-t} (F(s) - F(s+t))^2 \; ds \; .$$

4. If F is linear, $F(x) = a \cdot x$, then $\left|F(v+u) - F(v)\right|^2 = (a \cdot u)^2$ and the defining equation becomes

$$\int \Phi(t)\gamma(t)\ \nu(dt) = \int_D \Phi(\|u\|)\frac{1}{2}(a \cdot u)^2\ dudv\ .$$

5. If F is Lipschitzs, we can derive an upper bound for the variogram. In fact,

$$\left|F(v+u) - F(v)\right|^2 \leq \|F\|_{\mathrm{Lip}}^2 \|u\|^2\ ,$$

implies

$$\gamma(t) \leq \frac{1}{2}\|F\|_{\mathrm{Lip}}^2\, t^2\ .$$