A note on semivariogram
Una nota sul semivariogramma

Giovanni Pistone and Grazia Vicario

Abstract Variograms are usually discussed in the framework of stationary or intrinsically stationary processes. We retell here this piece of theory in the setting of generic Gaussian vectors.
Abstract *Si parla normalmente di variogrammi nell'ambito di processi stazionari o intrinsecamente stazionari. Qui riesponiamo questa parte di teoria nel caso di vettori gaussiani generici.*

Key words: Kriging; Gaussian field; Geostatistics; Variogram.

1 Introduction

In this paper we discuss some preliminary items on variograms as defined by Matheron [5], see the modern exposition in [2, Ch. 2], [1, Ch. 2], [3, Ch.1], [4]. We touched previously into this topic in the course of applied research, see [6, 7]. Our goal now is to rework the basics in order to prepare for a nonparametric Bayes approach to Kriging.

G. Pistone
de Castro Statistics, Collegio Carlo Alberto, Via Real Collegio 30, 10024 Moncalieri, Italy e-mail: giovanni.pistone@carloalberto.org

G. Vicario
DISMA Giuseppe Luigi Lagrange, Politecnico di Torino, Corso Duca degli Abruzzi 24, 10124 Torino, Italy e-mail: grazia.vicario@polito.it

2 Variogram of a normal vector

We consider first generic Gaussian vectors and we intend to specialize our assumptions later on.

Definition 1. Assume $Y \sim \mathrm{N}_n(\mu, \Sigma)$, $\Sigma = [\sigma_{ij}]_{i,j=1}^n$. The *variogram* of Y is the $n \times n$ matrix $\Gamma = [\gamma_{ij}]_{i,j=1\dots n}$ with

$$2\gamma_{ij} = \mathrm{Var}\,(Y_i - Y_j) = (\boldsymbol{e}_i - \boldsymbol{e}_j)'\Sigma(\boldsymbol{e}_i - \boldsymbol{e}_j) = \sigma_{ii} + \sigma_{jj} - 2\sigma_{ij}.$$

The matrix Γ can be written

$$\Gamma = \frac{1}{2}\left(\mathrm{vdiag}\,(\Sigma)\,\mathbf{1}' + \mathbf{1}\,\mathrm{vdiag}\,(\Sigma)'\right) - \Sigma = \frac{1}{2}\left(\mathrm{diag}\,(\Sigma)\,\mathbf{1}\mathbf{1}' + \mathbf{1}\mathbf{1}'\,\mathrm{diag}\,(\Sigma)\right) - \Sigma$$

where $\mathbf{1}$ is the unit column vector and $\mathrm{vdiag}\,(\Sigma) = \mathrm{diag}\,(\Sigma)\,\mathbf{1}$ is the diagonal of Σ as a column vector. Recall that $\frac{1}{n}\mathbf{1}\mathbf{1}'$ is the orthogonal projector on $\mathrm{Span}\,(\mathbf{1})$.

The variogram matrix has the following basic properties.

Proposition 1. *The variogram Γ is symmetric with zero diagonal and it is* conditionally negative definite.

More precisely, the quadratic form of

$$\Sigma = \frac{1}{2}\left(\mathrm{vdiag}\,(\Sigma)\,\mathbf{1}' + \mathbf{1}\,\mathrm{vdiag}\,(\Sigma)'\right) - \Gamma$$

is

$$\boldsymbol{\alpha}'\Sigma\boldsymbol{\alpha} = (\boldsymbol{\alpha}\cdot\mathbf{1})(\boldsymbol{\alpha}\cdot\mathrm{vdiag}\,(\Sigma)) - \boldsymbol{\alpha}'\Gamma\boldsymbol{\alpha},$$

hence $\boldsymbol{\alpha}\cdot\mathbf{1} = 0$ implies $\boldsymbol{\alpha}'\Sigma\boldsymbol{\alpha} = -\boldsymbol{\alpha}'\Gamma\boldsymbol{\alpha}$, in particular Γ is negative definite conditionally to $\sum_j \alpha_j = 0$. A symmetric matrix which has zero diagonal and is conditionally negative definite is called a *variogram matrix*. The variogram matrix carries $n(n-1)/2$ degrees of freedom (df), while the diagonal of Σ carries n df. Together, Λ and Γ form a proper parameterization of Σ.

Proposition 2. *The mapping from a positive definite Σ to a positive diagonal Λ and a variogram matrix Γ defined by*

$$\Sigma \mapsto \left(\mathrm{diag}\,(\Sigma), \frac{1}{2}(\mathrm{vdiag}\,(\Sigma)\,\mathbf{1}' + \mathbf{1}\,\mathrm{vdiag}\,(\Sigma)') - \Sigma\right) = (\Lambda, \Gamma) \tag{1}$$

is injective and its inverse

$$(\Lambda, \Gamma) \mapsto \frac{1}{2}(\Lambda\mathbf{1}\mathbf{1}' + \mathbf{1}\mathbf{1}'\Lambda) - \Gamma = \frac{1}{2}(\mathbf{vec}\,(\Lambda)\,\mathbf{1}' + \mathbf{1}\,\mathbf{vec}\,(\Lambda)') - \Gamma. \tag{2}$$

is defined on all Λ, Γ positive diagonal and conditionally negative definite, respectively, and satisfying

$$\boldsymbol{\alpha}\cdot\mathbf{1} = 1 \Rightarrow \boldsymbol{\alpha}\cdot\mathbf{vec}\,(\Lambda) \geq \boldsymbol{\alpha}'\Gamma\boldsymbol{\alpha}. \tag{3}$$

Proof. Let Λ and Γ be generic positive diagonal and conditionally negative definite, respectively. Then for a generic $\boldsymbol{\alpha} = \boldsymbol{\alpha}_0 + \bar{\alpha}\mathbf{1}$, with $\boldsymbol{\alpha}_0 \cdot \mathbf{1} = 0$ and $\bar{\alpha} = \frac{1}{n}\boldsymbol{\alpha} \cdot \mathbf{1}$, we have

$$\boldsymbol{\alpha}' \left[\frac{1}{2}(\Lambda \mathbf{1}\mathbf{1}' + \mathbf{1}\mathbf{1}'\Lambda) - \Gamma\right] \boldsymbol{\alpha} = n\bar{\alpha}\ \boldsymbol{\alpha} \cdot \mathbf{vec}\,(\Lambda) - \boldsymbol{\alpha}'\Gamma\boldsymbol{\alpha}$$
$$= \begin{cases} -\boldsymbol{\alpha}_0\Gamma\boldsymbol{\alpha}_0 \geq 0 & \text{if } \bar{\alpha} = 0, \\ \boldsymbol{\alpha} \cdot \mathbf{vec}\,(\Lambda) - \boldsymbol{\alpha}'\Gamma\boldsymbol{\alpha} & \text{if } n\bar{\alpha} = 1. \end{cases}$$

□

If $\det(\Sigma) \neq 0$, similar formulæ are obtained by considering the correlation matrix

$$R = \operatorname{diag}(\Sigma)^{-1/2}\,\Sigma\,\operatorname{diag}(\Sigma)^{-1/2},$$

viz

$$\Gamma = \frac{1}{2}(\mathbf{vec}\,(\Sigma)\,\mathbf{1}' + \mathbf{1}\,\mathbf{vec}\,(\Sigma)') - \operatorname{diag}(\Sigma)^{1/2} R \operatorname{diag}(\Sigma)^{1/2}.$$

This formula is sometimes preferred because both Γ and R carry the same number of degrees of freedom.

Given a variogram matrix Γ, Eq. (3) is a quadratic programming problem. Here, we do not discuss it in full generality, but we move to consider the case where the variance is constant, as it is common in many applications.

Proposition 3. *If the variance is constant,* $\operatorname{diag}(\Sigma) = \lambda I_n$*, Eq.s* (1) *and* (2) *become*

$$\Sigma \mapsto \left(\lambda, \lambda\mathbf{1}\mathbf{1}' - \Sigma\right) = (\lambda, \Gamma), \quad (\lambda, \Gamma) \mapsto \lambda\mathbf{1}\mathbf{1}' - \Gamma,$$

respectively, and the existence condition becomes

$$\lambda \geq \max\left\{\boldsymbol{\alpha}'\Gamma\boldsymbol{\alpha}\middle|\boldsymbol{\alpha}\cdot\mathbf{1} = 1\right\}. \tag{4}$$

The knowledge of the support of the parameterization with λ and Γ is crucial in the choice of a coherent apriori distribution.

Let us discuss first the case $n = 2$. We have

$$\begin{bmatrix} \sigma & \sigma_{1,2} \\ \sigma_{1,2} & \sigma \end{bmatrix} = \begin{bmatrix} \lambda & \lambda - \gamma \\ \lambda - \gamma & \lambda \end{bmatrix}, \quad \gamma \geq 0,$$

and we need the sign of

$$\det \begin{bmatrix} \lambda & \lambda - \gamma \\ \lambda - \gamma & \lambda \end{bmatrix} = \lambda^2 - (\lambda - \gamma)^2 = \gamma(2\lambda - \gamma),$$

which is positive if $\lambda \geq \gamma/2$. This shows existence and shows that there is a restriction on λ which is worthwhile to investigate further. The condition in (4) involves the lower bound $\max_\alpha 2\alpha(1-\alpha)\gamma = \gamma/2$.

Assume now $n = 3$, that is

$$\begin{bmatrix} \sigma & \sigma_{1,2} & \sigma_{1,3} \\ \sigma_{1,2} & \sigma & \sigma_{2,3} \\ \sigma_{1,3} & \sigma_{2,3} & \sigma \end{bmatrix} = \begin{bmatrix} \lambda & \lambda-\gamma_{12} & \lambda-\gamma_{13} \\ \lambda-\gamma_{12} & \lambda & \lambda-\gamma_{23} \\ \lambda-\gamma_{13} & \lambda-\gamma_{23} & \lambda \end{bmatrix},$$

with Γ conditionally negative definite. We have to assume $\lambda > \gamma_{12}/2$ and moreover we need the sign of

$$\det \begin{bmatrix} \lambda & \lambda-\gamma_{12} & \lambda-\gamma_{13} \\ \lambda-\gamma_{12} & \lambda & \lambda-\gamma_{23} \\ \lambda-\gamma_{13} & \lambda-\gamma_{23} & \lambda, \end{bmatrix} = \\ -2\gamma_{12}\gamma_{13}\gamma_{23} + \lambda\left(-\gamma_{12}^2 + 2\gamma_{12}\gamma_{13} - \gamma_{13}^2 + 2\gamma_{12}\gamma_{23} + 2\gamma_{13}\gamma_{23} - \gamma_{23}^2\right) \geq 0.$$

The solution of such algebraic inequalities is difficult in general, but we see that the admissible values of λ form a semi-infinite interval. In this and other similar cases, we can use a symbolic software such as Sage [9] to help with the algebra.

We now change our point of view to consider the same problem from a different angle. We can associate the variogram with the state space description of the Gaussian vector. This is of use, for example, when a simulation is required.

Proposition 4. *Given a variogram matrix Γ, the matrix*

$$\Sigma_0 = -(I - \frac{1}{n}\mathbf{1}\mathbf{1}')'\Gamma(I - \frac{1}{n}\mathbf{1}\mathbf{1}')$$

is symmetric and positive definite. If $Y_0 \sim \mathrm{N}_n(0, \Sigma_0)$, then its variogram is Γ and it is supported by $\mathrm{Span}\,(\mathbf{1})^\perp$.

Proof. The matrix Σ_0 is symmetric and positive definite because for a generic vector $\boldsymbol{\alpha}$ the vector $(I - \frac{1}{n}\mathbf{1}\mathbf{1}')\boldsymbol{\alpha}$ is orthogonal to $\mathbf{1}$, hence

$$\boldsymbol{\alpha}'\Sigma_0\boldsymbol{\alpha} = -\boldsymbol{\alpha}'(I - \frac{1}{n}\mathbf{1}\mathbf{1}')'(-\Gamma)(I - \frac{1}{n}\mathbf{1}\mathbf{1}')\boldsymbol{\alpha} \geq 0.$$

Also $\boldsymbol{e}_i - \boldsymbol{e}_j \in \mathrm{Span}\,(\mathbf{1})^\perp$, so that

$$(\boldsymbol{e}_i - \boldsymbol{e}_j)'\Sigma_0(\boldsymbol{e}_i - \boldsymbol{e}_j) = \\ (\boldsymbol{e}_i - \boldsymbol{e}_j)'(I - \frac{1}{n}\mathbf{1}\mathbf{1}')'(-\Gamma)(I - \frac{1}{n}\mathbf{1}\mathbf{1}')(\boldsymbol{e}_i - \boldsymbol{e}_j) = \\ -\gamma_{ii} - \gamma_{jj} + 2\gamma_{ij} = 2\gamma_{ij}.$$

As $\mathbf{1}'(\boldsymbol{e}_i - \boldsymbol{e}_j) = 0$, then $\mathbf{1}'(I - \frac{1}{n}\mathbf{1}\mathbf{1}')'(-\Gamma)(I - \frac{1}{n}\mathbf{1}\mathbf{1}')\mathbf{1} = 0$, the distribution of Y_0 is supported by the space $\mathbf{1}^\perp$. $\square$

Let us derive some other equivalent expression for Σ_0. The h-th element of $\mathrm{diag}\,(\Sigma_0)$ is

$$\boldsymbol{e}_h'\Sigma_0\boldsymbol{e}_h = -\boldsymbol{e}_h'(I - \frac{1}{n}\mathbf{1}\mathbf{1}')'\Gamma(I - \frac{1}{n}\mathbf{1}\mathbf{1}')\boldsymbol{e}_h = -(\boldsymbol{e}_h - \frac{1}{n}\mathbf{1})'\Gamma(\boldsymbol{e}_h - \frac{1}{n}\mathbf{1})$$

hence

$$\operatorname{diag}(\Sigma_0) = -\sum_h \boldsymbol{e}_h \boldsymbol{e}_h'(I - \frac{1}{n}\mathbf{1}\mathbf{1}')'\Gamma(I - \frac{1}{n}\mathbf{1}\mathbf{1}')\boldsymbol{e}_h \boldsymbol{e}_h'$$

and also

$$\begin{aligned}\operatorname{diag}(\Sigma_0)\,\mathbf{1}\mathbf{1}' &= -\sum_h \boldsymbol{e}_h \boldsymbol{e}_h'(I - \frac{1}{n}\mathbf{1}\mathbf{1}')'\Gamma(I - \frac{1}{n}\mathbf{1}\mathbf{1}')\boldsymbol{e}_h \mathbf{1}' \\ \mathbf{1}\mathbf{1}'\operatorname{diag}(\Sigma_0) &= -\sum_h \mathbf{1}\boldsymbol{e}_h'(I - \frac{1}{n}\mathbf{1}\mathbf{1}')'\Gamma(I - \frac{1}{n}\mathbf{1}\mathbf{1}')\boldsymbol{e}_h \boldsymbol{e}_h'.\end{aligned}$$

Let us compute the (i,j) element.

$$\begin{aligned}\boldsymbol{e}_i'\operatorname{diag}(\Sigma_0)\,\mathbf{1}\mathbf{1}'\boldsymbol{e}_j &= -\boldsymbol{e}_i'(I - \frac{1}{n}\mathbf{1}\mathbf{1}')'\Gamma(I - \frac{1}{n}\mathbf{1}\mathbf{1}')\boldsymbol{e}_j \\ \boldsymbol{e}_i'\mathbf{1}\mathbf{1}'\operatorname{diag}(\Sigma_0)\,\boldsymbol{e}_j &= -\boldsymbol{e}_j'(I - \frac{1}{n}\mathbf{1}\mathbf{1}')'\Gamma(I - \frac{1}{n}\mathbf{1}\mathbf{1}')\boldsymbol{e}_j.\end{aligned}$$

The previous computation are of use in the analysis of the variogram of Σ_0,

$$\frac{1}{2}(\operatorname{diag}(\Sigma_0)\,\mathbf{1}\mathbf{1}' + \mathbf{1}\mathbf{1}'\operatorname{diag}(\Sigma_0)) + (I - \frac{1}{n}\mathbf{1}\mathbf{1}')'\Gamma(I - \frac{1}{n}\mathbf{1}\mathbf{1}')$$

The following is an additive decomposition of a generic Gaussian vector into a term whose variance is that obtained in the previous proposition and a Gaussian vector proportional to $\mathbf{1}$.

Proposition 5. *Let $Y \sim \mathrm{N}_n(\mu, \Sigma)$ with variogram Γ. Let $Y_0 = (I - \frac{1}{n}\mathbf{1}\mathbf{1}')Y$ be the projection of Y onto $\mathbf{1}^\perp$. As $\mathbf{vec}(\Sigma)\,\mathbf{1}'(I - \frac{1}{n}\mathbf{1}\mathbf{1}') = 0$, then the variance of Y_0 is $\Sigma_0 = (I - \frac{1}{n}\mathbf{1}\mathbf{1}')\Sigma(I - \frac{1}{n}\mathbf{1}\mathbf{1}') = -(I - \frac{1}{n}\mathbf{1}\mathbf{1}')\Gamma(I - \frac{1}{n}\mathbf{1}\mathbf{1}')$. We can write $Y = Y_0 + Z$, where each component of Z is $\frac{1}{n}\sum_j Y_j$.*

3 Variograms and stationarity: final remarks

Let G be an additive topological locally compact group e.g., $\mathbb{Z}$ or $\mathbb{R}$ with the ordinary sum $x+y$. A centered Gaussian random process $(Y(x))_{x\in G}$ is *stationary* if $\operatorname{Cov}(Y(x), Y(y)) = \operatorname{Cov}(Y(x-y), Y(0)) = C(x-y)$. The autocovariance function C is positive definite, that is $\sum_{ij=0}^n \alpha_i \alpha_j C(x_i - x_j) \geq 0$, $n \in \mathbb{N}$, $x_1, \ldots, x_n \in G$, $\alpha \in \mathbb{R}^n$. The process is *intrinsically stationary* if $\operatorname{Var}(Y(x) - Y(y)) = \operatorname{Var}(Y(x-y) - Y(0)) = 2\gamma(x-y)$. The variogram function γ is conditionally negative definite, i.e. the matrix $\Gamma = [\gamma(x_i - x_j)]_{i,j=1}^n$, $n \in \mathbb{N}$, $x_1, \ldots, x_n \in G$, is conditionally negative definite, as in Prop. 1.

We plan to discuss, in a paper to come, the existence of an intrinsically stationary process Y given a conditionally negative definite function and we want to characterize specific classes of variogram functions e.g., those which are increasing (if an order is available) and bounded as $x \to \infty$. Increasing and bounded variograms are considered especially adapted to Geostatistics. In fact, D.G. Krige himself assumed that the variance of the difference between values measured in two locations is increasing with the distance between the locations, while the covariance vanishes. In the stationary case, these assumptions are still valid; therefore, we can use the results of the previous section, together with a further characterization of variograms, which is based on the following theorem.

Proposition 6 ([8, Th. 6.1.8]). *Let γ: G and $f(0) \geq 0$. Then γ is conditionally negative definite if, and only if, for all finite sequence $x_1, \dots, x_n$, the matrix $A = [\gamma(x_i - x_j) - \gamma(x_i) - \gamma(-x_j)]_{i,j=1}^n$ is negative definite.*

Proof. If the matrix A is negative definite and $\sum_i \alpha_i = 0$, then

$$0 \geq \sum_{i,j=1}^n \alpha_i \alpha_j (\gamma(x_i - x_j) - \gamma(x_i) - \gamma(-x_j)) = \sum_{ij=1}^n \alpha_i \alpha_j \gamma(x_i - x_j)$$

Viceversa, from generic $x_1, \dots, x_n$, $\alpha_1, \dots, \alpha_n$, define $x_{n+1} = 0$ and $\alpha_{n+1} = -\sum_i \alpha_i$, then write the condition for conditional negativity. □

Finally, in this setting one must take advantage of the harmonic representation of positive definite functions.

Acknowledgements G. Pistone is supported by the de Castro Statistics Initiative, Collegio Carlo Alberto, and is a member of GNAMPA–INdAM.

References

1. Chilès, J.P., Delfiner, P.: Geostatistics, second edn. Wiley Series in Probability and Statistics. John Wiley & Sons Inc., Hoboken, NJ (2012).
2. Cressie, N.A.C.: Statistics for spatial data. Wiley Series in Probability and Mathematical Statistics: Applied Probability and Statistics. John Wiley & Sons Inc., New York (1993).
3. Gaetan, C., Guyon, X.: Spatial statistics and modeling. Springer Series in Statistics. Springer, New York (2010).
4. Gneiting, T., Sasvári, Z., Schlather, M.: Analogies and correspondences between variograms and covariance functions. Adv. in Appl. Probab. **33**(3), 617–630 (2001).
5. Matheron, G.: Traité de géostatistique appliqué. No. 14 in Mem. Bur. Rech. Geog. Minieres. Editions Technip (1962)
6. Pistone, G., Vicario, G.: Comparing and generating latin hypercube designs in kriging models. ASTA **94**, 353–366 (2010).
7. Pistone, G., Vicario, G.: Kriging prediction from a circular grid: application to wafer diffusion. Appl. Stochastic Models Bus. Ind. **29**(4), 350–361 (2013).
8. Sasvári, Z.: Positive definite and definitizable functions, *Mathematical Topics*, vol. 2. Akademie Verlag, Berlin (1994)
9. Stein, W., et al.: Sage Mathematics Software (Version 5.9). The Sage Development Team (2013). `http://www.sagemath.org`