Ecole des Mines
de Paris

COURSE ON MULTIVARIATE GEOSTATISTICS

Jacques RIVOIRARD

Cours C-17 3

Février 2003

CENTRE DE GÉOSTATISTIQUE
35 rue Saint-Honoré
77305 FONTAINEBLEAU (France)
Tél. : 33-1 64 69 47 81 Fax : 33-1 64 69 47 05
http://cg.ensmp.fr

Course on multivariate geostatistics

Table of contents

0 Introduction

0.1 Foreword

0.1.1 Content, references

These notes have been written for the course of multivariate geostatistics given at the Centre of Géostatistique of Fontainebleau, within the CFSG (Cycle of Formation Spécialisée en Géostatistique) since 2001 and for the winter "hivernal" professional training 2002. They present the basis of multivariate geostatistics (essentially the linear geostatistics, in the stationary and intrinsic cases). For further developments, see for instance Wackernagel (1993), Wackernagel (1995) and Chilès and Delfiner (1999), where a number of references can be found.

The section on the simplification of cokriging is however largely original, as well as the last one, which treats the particular case of indicators and opens the door to non-linear geostatistics.

0.1.2 Convenient abuses of language

Some abuses of language are very convenient :

Correlated variables and correlated Random Functions

The regionalized variables under study, variables depending on location, will be represented as usual in geostatistics by (a model of) Random Function (RF in short). Although this distinction is fundamental, we will generally and conveniently speak of "variables", even for the RF that represent these, when this does not make confusion.

The explicit reference to RFs will however be made to underline a difference. For instance, uncorrelated RF will represent variables which are not correlated, not only at the same point, but also between different points, that is , variables "without spatial correlation" with each other. The word "theoretical" will refers also to the model of RF, for instance the theoretical variance in the stationary case (sill of variogram model, different from the experimental variance of samples).

In the case of intrinsic RFs, we will speak abusively of variables without spatial correlation, when there is no correlation between the increments of the variables – more exactly, between the increments of the corresponding RFs…

Identical structures, different structures

When there is no confusion, two variables will be said to have the "same" structure, or "identical" structure, when their variograms, for instance, are the same, up to a multiplicative factor, i.e. when they are proportional to each other, like 4 sph (h/100) and 9 sph(h/100).

Similarly the structures of the variables will be said to be different, when they are not proportional, for instance 4 sph(h/100) and 4 nugget(h) + 4 sph(h/100).

Remarks on the word ``intrinsic''

Be careful to the meaning of the adjective « intrinsic », which means: "proper to the object itself, independent of external factors".

At its beginning, the variogram was also called the intrinsic function of dispersion, for unlike the covariance, it describes directly the spatial structure, independently of the means and of the problems posed by the estimation of these means. The model (or hypothesis) of intrinsic RF is precisely characterized by this sole variogram.

0.1.3 Notations and indices

The number of variables and indices is a first difficulty when learning multivariate. To facilitate the approach, we have chosen to put in upper indices (not to be confused with a power) only indices representing components with different scales. For instance, the variogram of the first variable in a set of variables, $Z_1(x)$, may be written : $\gamma_1(h) \, or \, \gamma_{11}(h) = b_{11}^1 \gamma^1(h) + b_{11}^2 \gamma^2(h)$

where the upper index 1 refers for instance to a short ranged component, and the upper index 2 to a large ranged one.

0.2 Why multivariate geostatistics?

Multivariate geostatistics is essentially used to:

- Highlight structural relations between variables;

- Improve the estimation of one variable thanks to other variable(s), that can be sampled at the same points (« isotopic » case) or not (« heterotopic » case) ;

- Improve the consistencies between the estimations of different variables ;

- Simulate jointly several variables.

0.3 Examples of covariables

Here are some examples :

- Elevations of geological horizons:

    top and bottom of a layer (ex: bedrock and overburden), as well as the corresponding thickness

    several layers

    elevation of a layer + data on dip

    elevation of a layer known at wells + seismic known at many points

- grades in different metals, concentrations in different elements

- thickness and metal accumulation of a 2D orebody, as well as their ratio, the 2D grade

- different types of measure, errors

- indicators for different lithofacies

- etc.

1 Reminders on statistics of one variable

In this chapter, we recall some basic notions on the correlation and on the regressions between two variables.

1.1 Scatterplot between 2 variables

The utility of the scatterplot cannot be exaggerated. It allows :

- To distinguish different populations, to easily identify extreme values, or even some artefacts ;
- To visualize the relation between variables.

Beware however of variations in the number of samples throughout the scatterplot, which can make the readability and the interpretation of a scatterplot difficult (a logarithmic transformation, for instance, may facilitate the readability for skewly distributed variables).

1.2 Distribution per class (conditional distributions)

The scatterplot represents the empirical version of a bivariate statistical distribution. Similarly, the distribution of the values of a variable, say Z_2, per class on Z_1, is the empirical version of the distribution of Z_2 knowing (conditional on) $Z_1 = z_1$.

1.3 Mean per class (regression)

The mean of Z_2 per class on Z_1 corresponds to the empirical version of the conditional expectation $E[Z_2|Z_1]$ of Z_2 on Z_1.

1.4 Coefficient of correlation

Consider two variables Z_1 and Z_2 with :

- means :

  $m_1 = E[Z_1]$ and $m_2 = E[Z_2]$

- variances :

  $$\sigma_1^2 = \text{var}\, Z_1 = E[(Z_1 - m_1)^2]$$
  $$= E(Z_1^2) - 2E(Z_1)m_1 + m_1^2 = E(Z_1^2) - m_1^2$$

  and $\sigma_2^2 = \ldots$

- standard deviations :

  $\sigma_1 = \sqrt{\sigma_1^2}$ and σ_2

- covariance :

  $$C_{12} = \text{cov}(Z_1, Z_2) = E[(Z_1 - m_1)(Z_2 - m_2)]$$
  $$= \ldots = E(Z_1 Z_2) - m_1 m_2$$

By definition, the coefficient of correlation is equal to:

$$\rho = \frac{C_{12}}{\sigma_1 \sigma_2}$$

It satisfies: $-1 \leq \rho \leq 1$. (But inconsistencies may appear – a coefficient of correlation exceeding 1, for instance– when both variables are not sampled together and when means, variances and covariance are not computed on the same samples.)

The coefficient of correlation measures the linear dependency between variables. It is zero if the variables are independent, but can also be zero when they are dependent. It is highly sensitive to extreme values.

1.5 Linear regressions between 2 variables

Remember that the linear regression of Z_2 on Z_1 is the best linear estimator of Z_2 from Z_1 :

$$Z_2^* = aZ_1 + b$$

(best in the sense of having no bias and a minimal variance of error).

The committed error can be written:

$$R = Z_2 - Z_2^* = Z_2 - aZ_1 - b$$

The estimator is unbiased if the error has a zero mean, which gives b:

$$E(R) = m_2 - am_1 - b = 0 \Leftrightarrow b = m_2 - am_1$$

so that, in term of centered variables, the estimator can be written :

$$Z_2^* - m_2 = a(Z_1 - m_1)$$

In particular, this shows that the regression line goes through the means of the variable.

The variance of the error can be developed:

$$\begin{aligned} Var(R) &= Var(Z_2) + a^2 Var(Z_1) - 2aCov(Z_1, Z_2) \\ &= \sigma_2^2 + a^2\sigma_1^2 - 2aC_{12} \end{aligned}$$

It is a function of a, and its minimum is obtained by setting to zero its derivative with respect to a. Hence the slope of the regression:

$$a = \frac{Cov(Z_1, Z_2)}{Var(Z_1)} = \frac{C_{12}}{\sigma_1^2} = \rho\frac{\sigma_2}{\sigma_1}$$

With reduced variables, the expression of the regression is easy to memorize :

$$\frac{Z_2^* - m_2}{\sigma_2} = \rho\frac{Z_1 - m_1}{\sigma_1}$$

Residual and known variable are not correlated:

$$Cov(Z_1, R) = Cov(Z_1, Z_2 - aZ_1 - b) = C_{12} - a\sigma_1^2 = 0$$

(Since E(R) = 0, we have also: $E(Z_1R) = 0$: Z_1 and R are said to be "orthogonal")

Then it is possible to decompose Z_2 into two uncorrelated variables:

$$Z_2 = a\, Z_1 + b + R$$

The square of the correlation gives the proportion of variance of one of the variables Z_1 or Z_2, that is explained by its linear dependency to the other variable. We have for instance:

$$\frac{Var\left(Z_2^*\right)}{Var\left(Z_2\right)} = \frac{Var\left(aZ_1 + b\right)}{Var\left(Z_2\right)} = \frac{a^2\sigma_1^2}{\sigma_2^2} = \rho^2$$

The value of the correlation can be misleading. A correlation of 0.7, for instance, explains only 49% of the variance, a correlation of 0.5, 25%, and a correlation of 0.1, only 1%.

Beware: the linear regressions of Z_2 on Z_1 and of Z_1 on Z_2 are two different lines (slopes are not inverse of each other), that cross on the means:

$$Z_2^* - m_2 = \rho\frac{\sigma_2}{\sigma_1}\left(Z_1 - m_1\right)$$

$$Z_1^* - m_1 = \rho\frac{\sigma_1}{\sigma_2}\left(Z_2 - m_2\right)$$

Note also that the regression lines do not represent the “principal directions” of the scatterplot, along which the variability is maximal or minimal.

1.6 Remark

The statistical tools, particularly the simple tools of visualization and exploration, like the scatterplot, are extremely useful.

When dealing with regionalized variables, however, note that statistics such as probability distribution function, coefficient of correlation, regressions, are not “spatial” statistics: they do not depend on the locations of samples in space (permuting the values of variables between 2 points does not change the correlation, for instance).

Moreover, such statistics are not « intrinsic » to the variables, in the sense that they depend on the support (and on the domain) on which the variables are defined. For instance, the coefficient of correlation between two variables measures, in general, the linear dependency between the variables at the support used, not at a different support, which limits its meaning.

Finally, note that the absence of correlation between variables at the same point x (for instance between the residual of a regression and the conditioning variable) does not imply the absence of “spatial” correlation, i.e. between different points x and y.

2 Structural multivariate tools

2.1 Reminder: the monovariate case

Consider a variable (represented by a RF) Z(x). In linear geostatistics, we consider only the moments of order 1 and 2 of the RF (so-called 2nd order RF):

m(x) = E[Z(x)] mean at point x (generally called trend or drift when it depends on x)

E[Z(x)Z(y)] noncentered covariance between Z(x) and Z(y)

Cov[Z(x), Z(y)] = E[Z(x)-m(x)][Z(y)-m(y)] = E[Z(x) Z(y)]-m(x) m(y) covariance (usual, centered) between Z(x) and Z(y).

Then we make some hypotheses of stationarity, i.e. invariance under translation, allowing the inference of such moments from a unique realization.

2.1.1 Model of 2nd order stationary RF

It is defined by:

- a constant mean : E[Z(x)] = m

- a covariance: Cov[Z(x), Z(x +h)] = E[Z(x)-m][Z(x+h)-m] = C(h) function of the distance h between the points

- a constant variance: var Z(x) = E{[Z(x)-m]^2} = C(0)

We have $|C(h)| \leq C(0)$, hence :

$\rho(h) = C(h)/C(0)$, correlogram, or correlation between Z(x) and Z(x+h).

The covariance allows us to compute the variance of any linear combination :

$$Z = \sum_\alpha \lambda_\alpha Z(x_\alpha) = \sum_\alpha \lambda_\alpha Z_\alpha$$

of Z(x) at any points:

$$E(Z) = m\sum_\alpha \lambda_\alpha$$

$$Var(Z) = \sum_\alpha \sum_\beta \lambda_\alpha \lambda_\beta Cov\left(Z_\alpha, Z_\beta\right)$$

$$= \sum_\alpha \sum_\beta \lambda_\alpha \lambda_\beta C\left(x_\beta - x_\alpha\right) \geq 0$$

2.1.2 Model or intrinsic RF

It is defined by increments $Z(x+h) - Z(x)$, having a zero expectation :

E[Z(x+h)-Z(x)] = 0

and a variance depending only on the distance h between the points:

0.5 E{[Z(x+h)-Z(x)]^2} = γ(h) (variogram)

This makes possible the computation of the expectation and variance of any Authorized Linear Combination (ALC), i.e. of any combination :

$$\sum_\alpha \lambda_\alpha Z(x_\alpha) = \sum_\alpha \lambda_\alpha Z_\alpha$$

satisfying to $\sum_\alpha \lambda_\alpha = 0$:

$$E\left[\sum_\alpha \lambda_\alpha Z_\alpha\right] = 0$$

$$Var\left[\sum_\alpha \lambda_\alpha Z_\alpha\right] = -\sum_\alpha \sum_\beta \lambda_\alpha \lambda_\beta \gamma(x_\beta - x_\alpha)$$

A stationary RF is also intrinsic, and we have:

$\gamma(h) = C(0)-C(h)$

that is: $C(h) = C(0) - \gamma(h)$

The sill of the theoretical variogram coincides with the theoretical variance C(0), which represents, in the model, the variance of Z(x) in a large domain (it can be fairly different from the sample variance).

However, conversely, an intrinsic RF is not necessarily stationary (variogram without a sill).

In multivariate (linear) geostatistics, these different characteristics will be complemented by the "cross"-structures (variograms or covariances) between the variables, and possibly by relations between the means of the variables.

2.2 The cross-covariance

By definition, the cross-covariance between 2 RF $Z_1(x)$ and $Z_2(x)$ is:

$$Cov\left[Z_1(x), Z_2(y)\right]$$

of any pair of points (x, y).

It is important to clearly make a distinction between this and the covariance at the same point :

$$Cov\left[Z_1(x), Z_2(x)\right]$$

If the cross-covariance is zero for all x and y, we will say that the RF are not correlated, or equivalently that the variables are not spatially correlated, a stronger property than the absence of correlation at the same point.

In some cases (notably for indicators or in a non-stationary case), it is advantageous to consider the non-centered cross-covariance :

$$E\left[Z_1(x)Z_2(y)\right]$$

2.3 Stationary cross-covariance

In this case, the cross-covariance is invariant under translation, i.e. does not depend on the distance h between the points x and x+h. Hence:

- stationary non-centered cross-covariance:

$$E\left[Z_1(x)Z_2(x+h)\right] = K_{12}(h)$$

which can be estimated without bias from data Z_1 and Z_2 distant by h,

- or stationary cross-covariance (usual, centered):

$$Cov\left[Z_1(x), Z_2(x+h)\right] = C_{12}(h)$$

the estimation of which is less easy.

As a matter of fact, this can be written :

$$E\left[(Z_1(x) - E[Z_1(x)])(Z_2(x+h) - E[Z_2(x+h)])\right] = E\left[Z_1(x)Z_2(x+h)\right] - E[Z_1(x)]E[Z_2(x+h)]$$

Its estimation from data necessitates the means $E[Z_1(x)]$ and $E[Z_2(x+h)]$ be known. If these are to be estimated, they will be (in the simplest cases) supposed stationary : $E[Z_1(x)] = m_1$ and $E[Z_2(x+h)] = m_2$ (see the stationary multivariate model further). Then one can use a common global estimation of m_1 and m_2 for all distances h (so-called ergodic covariance) or an estimation based on the pairs $(Z_1(x), Z_2(x+h))$ used for distance h (so-called non-ergodic covariance). In the first case, and even more in the second one, serious problems of bias appear (due to the difficulty of separating, within the variables, the respective parts of mean and of deviation to the mean).

It is also possible to define the cross-correlation, or cross-correlogram :

$$\rho_{12}(h) = corr\left[Z_1(x), Z_2(x+h)\right] = \frac{C_{12}(h)}{\sigma_1\sigma_2}$$

which requires means and variances of $Z_1(x)$ and $Z_2(x+h)$. Here too, one can distinguish between ergodic and non-ergodic correlation: this last one (based on the means and standard deviations from pairs at distance h) ensures by construction the inequality $|\rho_{12}(h)| \leq 1$, but with even more serious problems of bias.

The cross-covariance generalizes the simple covariance (take $Z_2(x) = Z_1(x)$). However it has no reason to be positive, maximal or minimal for h = 0, and is not necessarily even, since $C_{12}(h) = C_{21}(-h)$ can be different from $C_{21}(h) = C_{12}(-h)$. In particular, it can show a correlation maximal for a non-zero distance vector, and for instance indicates a delay, between the variables, in this direction. For instance, the cross-covariance between a stationary RF $Z_1(x)$ and $Z_2(x) = Z_1(x-t)$ is maximal for the distance equal to the delay between the variables :

$$C_{12}(h) = Cov[Z_1(x), Z_1(x+h-t)] = C_1(h-t)$$

2.4 The order-2 stationary multivariate model

A set of RF $(Z_1(x), \ldots, Z_i(x), \ldots, Z_p(x))$ is order-2 stationary if all 1st and 2nd order moments are invariant under translation :

- constant means :

$$E[Z_i(x)] = m_i$$

- covariances depending only on the distance vector between points, for instance for centered covariances :

$$Cov\left[Z_i(x), Z_j(x+h) \right] = C_{ij}(h)$$

This implies the order-2 stationarity of each RF (means and simple covariances $C_i(h) = C_{ii}(h)$ for j = i), and must be complemented by stationarity of cross-covariances (i ≠ j).

In such a model, any linear combination of the ($Z_1(x)$, …, $Z_i(x)$, …, $Z_p(x)$), at any points:

$$Z = \sum_i \sum_\alpha \lambda_{i\alpha} Z_i(x_\alpha) = \sum_i \sum_\alpha \lambda_{i\alpha} Z_{i\alpha}$$

has the mean and variance :

$$E(Z) = \sum_i \sum_\alpha \lambda_{i\alpha} m_i$$

$$\begin{aligned} Var(Z) &= \sum_i \sum_j \sum_\alpha \sum_\beta \lambda_{i\alpha} \lambda_{j\beta} Cov\left(Z_{i\alpha}, Z_{j\beta} \right) \\ &= \sum_i \sum_j \sum_\alpha \sum_\beta \lambda_{i\alpha} \lambda_{j\beta} C_{ij}\left(x_\beta - x_\alpha \right) \geq 0 \end{aligned}$$

Note that the variance uses jointly the simple and cross-covariances of the variables.

2.4.1 Exercise: differentiate

Let Z(x) be a differentiable stationary RF with covariance C(h).

Show that the cross-covariance between Z and Z' is odd, equal to 0 for h = 0:

$$Cov\left[Z(x), Z'(x+h) \right] = C'(h)$$

and that the covariance of Z'(x) can be written:

$$Cov\left[Z'(x), Z'(x+h) \right] = -C''(h)$$

2.5 Cross-variogram

2.5.1 Definition

Let $Z_1(x)$ and $Z_2(x)$ be 2 intrinsic RF (increments $Z_1(x+h) - Z_1(x)$ and $Z_2(x+h) - Z_2(x)$ with a zero expectation and a stationary semi-variance, equal to the variogram of each variable). The cross-variogram is defined as the semi-covariance of increments between x and x+h, supposed to be stationary (invariant under translation):

$$\gamma_{12}(h) = \frac{1}{2} E\left[Z_1(x+h) - Z_1(x) \right] \left[Z_2(x+h) - Z_2(x) \right]$$

Its estimation requires knowing both variables at pairs of points distant by h (up to the usual tolerance).

It follows that:

$$\frac{\gamma_{12}(h)}{\sqrt{\gamma_1(h)\gamma_2(h)}}$$

represents the correlation between increments of both variables.

On the other hand, the ratio :

$$\frac{\gamma_{12}(h)}{\gamma_1(h)}$$

represents the slope of the linear regression :

$$[\,Z_2(x+h)-Z_2(x)\,]^* = a\,[\,Z_1(x+h)-Z_1(x)\,]$$

of $Z_2(x+h)-Z_2(x)$ knowing $Z_1(x+h)-Z_1(x)$.

The cross-variogram generalizes the simple variogram (take $Z_2(x) = Z_1(x)$), but can present much more various shapes. It can be negative (e.g. substitution of a metal or element by another one). It is 0 for h = 0, and is symmetrical:

$$\gamma_{12}(0) = 0$$
$$\gamma_{12}(h) = \gamma_{12}(-h)$$

2.5.2 Relation between cross-variogram and cross-covariance

In the stationary multivariate RF model, we have :

$$\gamma_{12}(h) = K_{12}(0) - \frac{K_{12}(h)+K_{12}(-h)}{2}$$

that is:

$$\gamma_{12}(h) = C_{12}(0) - \frac{C_{12}(h)+C_{12}(-h)}{2}$$

In other words, the cross-variogram sees only the even part of the covariance :

$$C_{12}(h) = \frac{C_{12}(h)+C_{12}(-h)}{2} + \frac{C_{12}(h)-C_{12}(-h)}{2}$$

= even part + odd part

The cross-covariance is a more powerful tool than the cross-variogram. However it requires stronger hypotheses.

Note finally :

Sill of the cross-variogram: $\gamma_{12}(\infty) = C_{12}(0) = Cov\left[Z_1(x), Z_2(x)\right]$

$C_{12}(h) \equiv 0 \quad \forall h \quad \Rightarrow \quad \gamma_{12}(h) \equiv 0 \quad \forall h$

2.6 Intrinsic multivariate model

This section is more difficult. We will begin by a general definition of the model, then we will finish by examining the usual simple characterization of the model, which is limited to the symmetrical case.

2.6.1 Definition

Generally speaking, a set of RF ($Z_1(x)$, …, $Z_i(x)$, …, $Z_p(x)$) can be said to be intrinsic if all increments of the RF have a zero expectation and if all their order-2 moments (simple and cross) are invariant under translation (stationary). We then must have :

$$E\left[Z_i(x+h)-Z_i(x)\right]=0 \qquad \text{for all h and i}$$

and $$Cov[Z_i(x+h)-Z_i(x), Z_j(x+l+h')-Z_j(x+l)] \qquad [1]$$

that is $$E\{[Z_i(x+h)-Z_i(x)][Z_j(x+l+h')-Z_j(x+l)]\}$$

depending only on h, h', and l, not on x.

In particular, we have :

- for l = 0, h = h', i = j, the variogram:

$$\gamma_i(h)=\gamma_{ii}(h)=\frac{1}{2}E\left(\left[Z_i(x+h)-Z_i(x)\right]^2\right)$$

so that each RF separately is intrinsic ;

- for l = 0, h = h', i ≠ j, the cross-variogram:

$$\gamma_{ij}(h)=\frac{1}{2}E\left[Z_i(x+h)-Z_i(x)\right]\left[Z_j(x+h)-Z_j(x)\right]$$

However, in general simple and cross-variograms do not suffice to compute quantities like [1]. For instance, a stationary multivariate RF model is also a multivariate intrinsic model, and [1], which is equal to :

$$C_{ij}(l+h'-h)+C_{ij}(l)-C_{ij}(l+h')-C_{ij}(l-h)$$

can be written in function of $\gamma_{ij}(h)$, if $C_{ij}(h)$ is even:

$$\gamma_{ij}(l+h')+\gamma_{ij}(l-h)-\gamma_{ij}(l+h'-h)-\gamma_{ij}(l)$$

but cannot, in general, if $C_{ij}(h)$ is not even.

2.6.2 Variance of Authorized Linear Combinations

Similarly, only under some conditions do simple and cross-variograms allow us to compute the variance of multivariate ALC (Authorized Linear Combinations), i.e. the combinations :

$$Z=\sum_i\sum_\alpha \lambda_{i\alpha}Z_{i\alpha}$$

satisfying the conditions:

$$\sum_{\alpha} \lambda_{i\alpha} = 0 \quad \forall i$$

(By the way, this implies E(Z)=0. Note also that each variable needs not be considered in each of points x_α: suffices to take $\lambda_{i\alpha} = 0$).

As a matter of fact, let us write, introducing an arbitrary point x_0:

$$Z = \sum_i \sum_{\alpha} \lambda_{i\alpha} Z_{i\alpha} = \sum_i \sum_{\alpha} \lambda_{i\alpha} \left(Z_{i\alpha} - Z_{i0} \right)$$

hence:

$$Var[Z] = \sum_i \sum_j \sum_{\alpha} \sum_{\beta} \lambda_{i\alpha} \lambda_{j\beta} Cov(Z_{i\alpha} - Z_{i0,} Z_{j\beta} - Z_{j0}) \qquad [2]$$

which outputs the terms :

$$Cov(Z_{i\alpha} - Z_{i0,} Z_{j\beta} - Z_{j0})$$

But we have:

$$E[(Z_{i\beta} - Z_{i\alpha})(Z_{j\beta} - Z_{j\alpha})] = E[(Z_{i\beta} - Z_{i0}) - (Z_{i\alpha} - Z_{i0})][(Z_{j\beta} - Z_{j0}) - (Z_{j\alpha} - Z_{j0})]$$

i.e.

$$2\gamma_{ij\alpha\beta} = 2\gamma_{ij_0\beta} + 2\gamma_{ij_0\alpha} - Cov(Z_{i\alpha} - Z_{i0,} Z_{j\beta} - Z_{j0}) - Cov(Z_{i\beta} - Z_{i0,} Z_{j\alpha} - Z_{j0})$$

else :

$$Cov(Z_{i\alpha} - Z_{i0,} Z_{j\beta} - Z_{j0}) + Cov(Z_{i\beta} - Z_{i0,} Z_{j\alpha} - Z_{j0}) = 2\gamma_{ij_0\beta} + 2\gamma_{ij_0\alpha} - 2\gamma_{ij\alpha\beta}$$

If the $\lambda_{i\alpha} \lambda_{j\beta} Cov(Z_{i\alpha} - Z_{i0,} Z_{j\beta} - Z_{j0})$ do not change when permuting i and j, or when permuting α and β, we will have:

$$Var[Z] = -\sum_i \sum_j \sum_{\alpha} \sum_{\beta} \lambda_{i\alpha} \lambda_{j\beta} \gamma_{ij} (x_\beta - x_\alpha)$$

(same formula as in the stationary case, replacing C_{ij} by $-\gamma_{ij}$)

This is the case notably :

- if $\lambda_{i\alpha} = k_i \lambda_\alpha$ (weights being similar for all variables) so that: $\lambda_{i\alpha} \lambda_{j\beta} = \lambda_{i\beta} \lambda_{j\alpha}$

- or if $Cov(Z_{i\alpha} - Z_{i0,} Z_{j\beta} - Z_{j0}) = Cov(Z_{i\beta} - Z_{i0,} Z_{j\alpha} - Z_{j0})$ [3]

(then being equal to $\gamma_{ij_0\beta} + \gamma_{ij_0\alpha} - \gamma_{ij\alpha\beta}$).

In the general case, it is only under this symmetry relation between any points, that simple and cross-variograms allow to compute the variance of a ALC.

2.6.3 Remark: the symmetry relation

This symmetry relation on 3 points is equivalent to the following symmetry relation on 4 points :

$$Cov[Z_i(x+h)-Z_i(x), Z_j(x+l+h')-Z_j(x+l)]$$

$$= Cov[Z_j(x+h)-Z_j(x), Z_i(x+l+h')-Z_i(x+l)] \quad [4]$$

These covariances, which characterize the multivariate intrinsic model, are then equal to :

$$Cov[Z_i(x+h)-Z_i(x), Z_j(x+l+h')-Z_j(x)+Z_j(x)-Z_j(x+l)]$$

$$= Cov[Z_i(x+h)-Z_i(x), Z_j(x+l+h')-Z_j(x)]$$

$$- Cov[Z_i(x+h)-Z_i(x), Z_j(x+l)-Z_j(x)]$$

that is: $\gamma_{ij}(l+h')+\gamma_{ij}(l-h)-\gamma_{ij}(l+h'-h)-\gamma_{ij}(l)$

As seen before, such a symmetry relation is necessarily satisfied by a stationary multivariate model when cross-covariances are even, not when they are not symmetrical. The cross-variogram is unable to account for a dissymmetry in the cross-structure.

We will retain that simple and cross-variograms describe a multivariate intrinsic model only in the symmetrical case, hence the following characterization. (In the more general non-stationary case, another tool is required, the generalized cross-covariance, difficult to capture, and outside of the present course.)

2.6.4 Characterization of the intrinsic model in the symmetrical case

Let ($Z_1(x)$, …, $Z_i(x)$, …, $Z_p(x)$) be a set of intrinsic RF, satisfying the symmetry relations [3] or [4] for all points. This model is entirely characterized by :

- expectations of increments equal to 0:

$$E\left[Z_i(x+h)-Z_i(x)\right]=0$$

- simple variograms:

$$\gamma_i(h)=\gamma_{ii}(h)=\frac{1}{2}E\left(\left[Z_i(x+h)-Z_i(x)\right]^2\right)$$

- cross-variograms:

$$\gamma_{ij}(h)=\frac{1}{2}E\left[Z_i(x+h)-Z_i(x)\right]\left[Z_j(x+h)-Z_j(x)\right]$$

Simple and cross-variograms then allow to compute the variance of any Authorized Linear Combination, according to [2].

2.7 The pseudo cross-variogram

Definition:

$$\psi_{12}(h) = \frac{1}{2} E\left[Z_2(x+h) - Z_1(x)\right]^2$$

It can be computed between variables known at different points (unlike $\gamma_{12}(h)$).

However it requires:

- that the difference between variables makes sense (units in particular), for instance a single variable but measured in different conditions ;

- that $Z_2(x) - Z_1(x)$ (or more exactly the expectation of its square) is stationary, for :

$$\psi_{12}(0) = \frac{1}{2} E\left[Z_2(x) - Z_1(x)\right]^2$$

does not depend on x.

It is possible to have : $\psi_{12}(h) \neq \psi_{12}(-h)$, its even part being related to the cross-variogram :

$$\frac{\psi_{12}(h) + \psi_{12}(-h)}{2} = \frac{1}{2} E\left[Z_2(x) - Z_1(x)\right]^2 + \gamma_{12}(h)$$

In the stationary case, we have :

$$\psi_{12}(h) = \frac{1}{2}\left[K_1(0) + K_2(0)\right] - K_{12}(h)$$

where $K_{12}(h) = E\left[Z_1(x) Z_2(x+h)\right]$ is the non-centered cross-covariance.

Remark : the pseudo cross-variogram, through not symmetrical in general, does not correspond to the generalized cross-covariance evoked at the end of section 2.6.3 (Künsch and al. 1997).

2.8 Exercise: compute a bivariate structure

Let Z_1 and Z_2 the grades in two metals (in %), every meter along a hole:

Z_1	Z_2
0	0
1	0
0	0
0	1
0	1
0	0
0	0
1	0
0	1
0	1

Mean, mean square, variance, standard deviation, for each variable, correlation coefficient?

Compute and plot, up to h = 4:

- the simple variogram for each variable;

- the cross-variogram;

- the "non-centered" covariance for each variable;

- the covariance for each variable.

Compute and plot, from h = - 4 up to h = 4:

- the non-centered cross-covariance;

- the cross-covariance;

- the cross-correlation.

How are the cross-structures modified when the last value of Z_2 is equal to 3 instead of 1 ? when it is unknown?

Remark : in the heterotopic case (variables known at possibly different points), inconsistencies may appear between experimental simple and cross-structures, when these are computed on different sets of points.

2.9 Support

The formula of regularization and variances in the monovariate case can be generalized to the multivariate case, using the cross-covariances and cross-variograms (stationary and intrinsic cases, respectively).

2.9.1 Regularization

Cross-covariance regularized on support v :

$$C_{ij,v}(h) = Cov\left[Z_i(v), Z_j(v_h)\right] = \frac{1}{|v|^2}\int_v \int_{v_h} C_{ij}(x-y)dxdy = \overline{C_{ij}}\left(v, v_h\right)$$

Cross-variogram regularized on support v :

$$\gamma_{ij,v}(h) = \overline{\gamma_{ij}}(v, v_h) - \overline{\gamma_{ij}}(v, v)$$

2.9.2 Co-dispersion

Dispersion variance of v within V :

$$Var\left[Z_i(\underline{v}) - Z_i(V)\right] = \overline{\gamma}_i(V,V) - \overline{\gamma}_i(v,v)$$

$\underline{v}$ describing uniformly the different volumes with support v that partition the domain V (i.e. $\overline{C}_i(v,v)$ for a large domain in the stationary case).

Dispersion covariance of v within V :

$$Cov\left[Z_i(\underline{v}) - Z_i(V), Z_j(\underline{v}) - Z_j(V)\right] = \overline{\gamma}_{ij}(V,V) - \overline{\gamma}_{ij}(v,v)$$

(i.e. $\overline{C}_{ij}(v,v)$ for a large domain in the stationary case).

Correlation of v within V :

$$\rho_{ij}(v \mid V) = \frac{\overline{\gamma}_{ij}(V,V) - \overline{\gamma}_{ij}(v,v)}{\sqrt{\overline{\gamma}_i(V,V) - \overline{\gamma}_i(v,v)}\sqrt{\overline{\gamma}_j(V,V) - \overline{\gamma}_j(v,v)}}$$

(i.e. $\dfrac{\overline{C}_{ij}(v,v)}{\sqrt{C_i(v,v)C_j(v,v)}}$ for a large domain in the stationary case).

The linear regression :

$$\left[Z_j(\underline{v}) - Z_j(V) \right]^* = a\left[Z_i(\underline{v}) - Z_i(V) \right]$$

of $Z_j(v)$ on $Z_i(v)$, for v describing V, has the slope :

$$\frac{\overline{\gamma}_{ij}(V,V) - \overline{\gamma}_{ij}(v,v)}{\overline{\gamma}_i(V,V) - \overline{\gamma}_i(v,v)}$$

(i.e. $\dfrac{\overline{C}_{ij}(v,v)}{\overline{C}_i(v,v)}$ for a large domain in the stationary case).

2.9.3 Exercise : correlation at different supports

Consider a lead-zinc-silver deposit, where the quasi-punctual sample grades $Z_1(x)$, $Z_2(x)$ and $Z_3(x)$ have the means :

$m_1 = 13.2$ (%)

$m_2 = 10.3$ (%)

$m_3 = 3.72$ oz/ton (1 once per ton $\approx$ 30 g/t)

and the simple and cross-structures (where Pep is the nugget, or pepitic, variogram function with sill 1, and Sph the spherical variogram function with sill and ranges equal to 1) :

$\gamma_1(h) = 11\ Pep(h) + 39\ Sph(h/60)$

$\gamma_2(h) = 9\ Pep(h) + 15\ Sph(h/60)$

$\gamma_3(h) = 1.1\ Pep(h) + 1.8\ Sph(h/60)$

$\gamma_{12}(h) = \quad 14.5\ Sph(h/60)$

$\gamma_{13}(h) = \quad 5\ Sph(h/60)$

$\gamma_{23}(h) = \quad 3.8\ Sph(h/60)$

(after P. Dowd, cited in Mining geostatistics, Journel and Huijbregts, Academic Press, 1978)

What are the values, in the model, of the correlation coefficients between the different sample grades ? (The domain is supposed large compared to the range.)

What are approximately the values of the coefficients of correlation between the different grades for small blocks ? (small compared to the range)

2.10 Structure of linear combinations of the variables

Knowing simple and cross covariances or variograms makes it possible to compute the simple structure of a new variable obtained by combining linearly the variables, or the cross-structure between such combinations.

Example :

Consider: $H(x) = Z_2(x) - Z_1(x)$ the thickness of a layer, between top Z_2 and bottom Z_1 (with mean $m_H = m_2 - m_1$ in the stationary case).

The structure of H can be written, in covariance:

$$C_H(h) = Cov[H(x), H(x+h)] = Cov[Z_2(x) - Z_1(x), Z_2(x+h) - Z_1(x+h)]$$
$$= C_1(h) + C_2(h) - C_{12}(h) - C_{12}(-h)$$

or in variogram:

$$\gamma_H(h) = \frac{1}{2} E\left[H(x+h) - H(x)\right]^2$$
$$= \gamma_1(h) + \gamma_2(h) - 2\gamma_{12}(h)$$

The cross-structure between H and Z_1 , for instance, is :

$$Cov[H(x), Z_1(x+h)] = C_{12}(-h) - C_1(h)$$

that is, in variogram:

$$\gamma_{12}(h) - \gamma_1(h)$$

Other example :

Let $Z(x) = m + Z_1(x) + Z_2(x)$. We have:

$$m_Z = m + m_1 + m_2$$
$$C_Z(h) = C_1(h) + C_2(h) + C_{12}(h) + C_{12}(-h)$$
$$\gamma_Z(h) = \gamma_1(h) + \gamma_2(h) + 2\gamma_{12}(h)$$

If Z_1 and Z_2 have a zero mean and are not correlated with each other, we have:

$$m_Z = m$$
$$C_Z(h) = C_1(h) + C_2(h)$$
$$\gamma_Z(h) = \gamma_1(h) + \gamma_2(h)$$

Inversely, in a monovariate case, a nested structure (superimposition of several structural components, corresponding for instance to different scales) can be interpreted as a decomposition of the variable itself into components at different scales. This example is the seed of the linear model of coregionalization studied further.

2.11 Closing relation

Consider the following closing relation (sum of variables constant) :

$$Z(x) = \sum_i Z_i(x) = 1$$

with:

$$\sum_i m_i = 1$$

Cross $Z_i(x)$ with the constant Z(x). With covariances, we obtain :

$$\sum_{j \neq i} C_{ij}(h) = -C_i(h)$$

which is negative for h = 0 (minus the variance $C_i(0)$). With variograms :

$$\sum_{j \neq i} \gamma_{ij}(h) = -\gamma_i(h) \leq 0$$

So the closing relation is responsible for necessary negative values in at least some of the cross covariances or variograms.

3 Some common models

The cross-structures are able to present much more various shapes than the simple structures, that they generalize (for instance, linear combinations of basic structures with positive or negative coefficients will represent cross-structures in the linear model of coregionalization). However the cross-structures between variables are not independent of (and must not be modelled independently of) the simple structures of these variables.

As a matter of fact, we have seen that the variances of linear combinations could be developed in function of simple and cross covariances in the stationary case, and (under some conditions) in function of simple and cross variograms in the intrinsic case. As such variances must be positive or null, cross- structures must be modelled jointly with simple structures, using an authorized multivariate model. Here are some current models. Some other models will be described in the Section 5 on the simplification of cokriging.

3.1 The model of intrinsic correlation

3.1.1 Definition

This is the simplest of all models : all structures, simple or cross, are the same (proportional) :

$$\gamma_{ij}(h) = b_{ij}\gamma(h) \qquad \forall i, j$$

which is convenient to write matricially :

$$(\gamma_{ij}(h)) = (b_{ij})\,\gamma(h)$$

that is :

$$\begin{bmatrix} \gamma_{11}(h) & \gamma_{12}(h) & \cdots & \gamma_{1p}(h) \\ \gamma_{21}(h) & \gamma_{22}(h) & \cdots & \gamma_{2p}(h) \\ \cdots & \cdots & \ddots & \vdots \\ \gamma_{p1}(h) & \gamma_{p2}(h) & \cdots & \gamma_{pp}(h) \end{bmatrix} = \begin{bmatrix} b_{11} & b_{12} & \cdots & b_{1p} \\ b_{21} & b_{22} & \cdots & b_{2p} \\ \cdots & \cdots & \ddots & \vdots \\ b_{p1} & b_{p2} & \cdots & b_{pp} \end{bmatrix} \gamma(h)$$

It follows immediately that the correlation between increments is constant (independent of h):

$$\frac{\gamma_{ij}(h)}{\sqrt{\gamma_i(h)\gamma_j(h)}} = \frac{b_{ij}}{\sqrt{b_{ii}b_{jj}}}$$

In this model the correlation between $Z_i(v)$ and $Z_j(v)$ within a domain V (section 2.9.2) is :

$$\rho_{ij}(v \mid V) = \frac{b_{ij}}{\sqrt{b_{ii}b_{jj}}} \text{ constant}$$

It is said to be « intrinsic » for it does not depend on the support v, nor on the domain V (hence the name of the model).

In the stationary case, we can write :

$$C_{ij}(h) = b_{ij}C(h)$$

or more conveniently (which will be supposed from now on):

$$C_{ij}(h) = b_{ij}\rho(h)$$

with $\rho(0) = 1$ (by the way, note that in this model the cross-covariances are necessarily even).

3.1.2 Matrix of coefficients

We have (stationary case):

$$b_{ij} = C_{ij}(0) = \text{cov}[Z_i(x), Z_j(x)]$$

and $\quad b_{ii} = C_{ii}(0) = \text{var}[Z_i(x)]$

so that the matrix of sills (symmetrical since $b_{ij} = b_{ji}$):

$$\left(b_{ij}\right) = \begin{bmatrix} b_{11} & b_{12} & \cdots & b_{1p} \\ b_{21} & b_{22} & \cdots & b_{2p} \\ \cdots & \cdots & \ddots & \vdots \\ b_{p1} & b_{p2} & \cdots & b_{pp} \end{bmatrix}$$

represents, in the model, the variances-covariances matrix of the variables at the same point (like in the monovariate case, the sills of the model can differ from the experimental variances-covariances).

Such a matrix cannot be arbitrary : it must be (semi-) positive definite, which means that the variances of any linear combination of the variables at the same point is positive, possibly zero :

$$\text{var}(\sum_i \lambda_i Z_i(x)) = \sum_{ij} \lambda_i \lambda_j Cov[Z(x_i), Z(x_j)]$$

$$= \sum_{ij} \lambda_i \lambda_j b_{ij} \geq 0$$

Remark : in the purely intrinsic case, (b_{ij}) must also be semi-positive definite. As a matter of fact, 2 (b_{ij}) $\gamma(h)$ represents the variances-covariances matrix of the increments $Z_i(x+h)- Z_i(x)$) between two points, hence:

$$\text{var}(\sum_i \lambda_i [Z_i(x+h) - Z_i(x)]) = 2\sum_{ij} \lambda_i \lambda_j \gamma_{ij}(h)$$

$$= 2(\sum_{ij} \lambda_i \lambda_j b_{ij})\,\gamma(h) \geq 0$$

Come back to the stationary case, easier to present. The matrix (b_{ij}) being (semi-) positive definite, we have necessarily :

$$b_{ii} = \text{var}[Z_i(x)] \geq 0 \text{ for all i}$$

and $b_{ii} b_{jj} - b_{ij}^{\,2} \geq 0$ for all pair (i, j)

which corresponds (if their variances are not zero) to a correlation coefficient $\rho_{ij} = \frac{b_{ij}}{\sqrt{b_{ii} b_{jj}}}$

between the variables $Z_i(x)$ and $Z_j(x)$ that lies between–1 and +1 as it must be.

These conditions, which must necessarily be satisfied for all i and j, are however not sufficient to ensure the positive definiteness of the matrix (b_{ij}) with more than two variables. This property will be automatically ensured if it is possible to associate to the matrix a factorization of the variables.

3.1.3 Factorization

First let us give two properties, elementary though important, of the model of intrinsic correlation :

- It is easy to see that the intrinsic correlation between variables can be extended to variables formed by linear combinations : in effect, simple and cross variograms of all linear combinations are proportional to $\gamma(h)$.

- On the other hand :

$$C_{ij}(0) = 0 \;\Rightarrow\; b_{ij} = 0 \;\Rightarrow\; C_{ij}(h) = 0$$

So, in the model of intrinsic correlation, the absence of correlation between two variables (or two linear combinations) at the same point implies the absence of spatial correlation.

Suppose now that the matrix (b_{ij}) can be associated with a factorization of the variables, i.e. a decomposition of the variables $Z_i(x)$ into « factors » $Y_q(x)$ not correlated with each other at the same point, for instance in the stationary case, and with $E[Y_q(x)] = 0$:

$$Z_i(x) = \sum_q a_{iq} Y_q(x) + m_i$$

and suppose that conversely the $Y_q(x)$ can be expressed as a function of the $Z_i(x)$. Such a decomposition guarantees the positive definiteness of the matrix (b_{ij}) : in effect, any linear combination of the $Z_i(x)$ at the same point x is also a linear combination of the uncorrelated $Y_q(x)$, and so has a variance positive or null :

$$\mathrm{var}\left(\sum_i \lambda_i Z_i(x) \right) = \sum_q \left(\sum_i \lambda_i a_{iq} \right)^2 \mathrm{var}(Y_q) \geq 0$$

Developing variances and covariances of the variables gives the following relation (if the $Y_q(x)$ share the same variogram $\gamma(h)$, including sill or vertical scale):

$$b_{ij} = \sum_q a_{iq} a_{jq}$$

In the model of intrinsic correlation, such factors $Y_q(x)$, being not correlated at the same point, are moreover not spatially correlated with each other. So any statistical factorization (by successive residuals, or principal components for instance) gives a spatial factorization of the model. The number of useful factors is in general equal to the number of variables (but can be

lower when variables are linked). Though they depend on the chosen method for factorization, these factors are objective, in the sense that their values at a point x can be determined when the values of the Z_i at this point are known.

Finally, the model of intrinsic correlation is entirely specified by its basic structure $\gamma(h)$, which must be authorized, and by the variances-covariances matrix (b_{ij}), which must be (semi-) positive definite, which is ensured when a factorization can be made. Thus the variables can be interpreted as linear combinations of factors having the same variogram and being without correlation with each other. In practice, once the basic structure is chosen, it is possible to call for an automatic fitting procedure that will give the b_{ij} and a_{iq}.

3.1.4 Examples

a) Consider two variables Z_1 and Z_2 with structures :

$\gamma_1(h) = 100\ \mathrm{sph}(h/20)$

$\gamma_2(h) = 25\ \mathrm{sph}(h/20)$

$\gamma_{12}(h) = -40\ \mathrm{sph}(h/20)$

This is a model of intrinsic correlation, for we can write $\gamma_{ij}(h) = b_{ij}\gamma(h)$ with:

- the basic structure $\gamma(h) = \mathrm{sph}(h/20)$
- the variances-covariance matrix : $\left(b_{ij}\right) = \begin{pmatrix} 100 & -40 \\ -40 & 25 \end{pmatrix}$

This matrix is positive definite for it corresponds to two variables with positive variances and having a correlation –0.8. But this can also be seen by producing a possible factorization of the model. Let us write for instance, with Y_1 and Y_2 with unit variances and uncorrelated (leaving the means aside), first:

$Z_1 = 10\ Y_1$

then:

$Z_2 = -4Y_1 + 3\ Y_2$

(here Y_1 is but the variable Z_1 reduced, and $3Y_2$ represents the residual of the linear regression of Z_2 knowing Z_1 or Y_1). Hence inversely the factors:

$Y_1 = 0.1\ Z_1$

$Y_2 = (4/30)\ Z_1 + (1/3)\ Z_2$

b) similarly the model for three variables $\gamma_{ij}(h) = b_{ij}\gamma(h)$ with $\gamma(h) = \mathrm{sph}(h/20)$ and

$\left(b_{ij}\right) = \begin{pmatrix} 100 & -40 & -30 \\ -40 & 25 & 6 \\ -30 & 6 & 17 \end{pmatrix}$ is also a model of intrinsic correlation.

The matrix is positive definite, for it is possible to factorize the model. For instance, introduce Y_1, Y_2, and Y_3, with unit variances and uncorrelated. To honour successively b_{11}, b_{12}, b_{22}, b_{13}, b_{23}, b_{33}, we can write:

$Z_1 = 10\ Y_1$

$Z_2 = -4Y_1 + 3\ Y_2$

$Z_3 = -3Y_1 - 2\ Y_2 + 2\ Y_3$

Then the factors are conversely given by:

$Y_1 = 0.1\ Z_1$

$Y_2 = (4/30)\ Z_1 + (1/3)\ Z_2$

$Y_3 = (17/60)\ Z_1 + (1/3)\ Z_2 + (1/2)\ Z_3$

As an exercise, produce other possible factorizations: for instance by the same method of successive residuals, but taking the variables in a different order; by analysis in principal components; or else by analysis in principal components on normalized variables.

(One could check that if we had $b_{33} = 13$, two factors would suffice to ensure positive definiteness, attesting a linear relation between the three variables; but with $b_{33} = 10$, factorization is not possible: the model is not authorized, although all correlation coefficients lie between –1 and +1.)

c) Suppose that, within model (b), a nugget component is added to the simple variogram of the 3rd variable. Then only the two variables Z_1 and Z_2 remain intrinsically correlated.

3.1.5 Exercise : examples of models of intrinsic correlation

Are the following models, models of intrinsic correlation?

$\gamma_1(h) = 64\ \mathrm{sph}(h/20)$

$\gamma_2(h) = 64\ \mathrm{sph}(h/20)$

$\gamma_{12}(h) = 38.4\ \mathrm{sph}(h/20)$

$\gamma_1(h) = 36\ \mathrm{pep}(h) + 64\ \mathrm{sph}(h/20)$

$\gamma_2(h) = 36\ \mathrm{pep}(h) + 64\ \mathrm{sph}(h/20)$

$\gamma_{12}(h) = 38.4\ \mathrm{sph}(h/20)$

$\gamma_1(h) = 36\ \mathrm{pep}(h) + 64\ \mathrm{sph}(h/20)$

$\gamma_2(h) = 36\ \mathrm{pep}(h) + 64\ \mathrm{sph}(h/20)$

$\gamma_{12}(h) = 21.6\ \mathrm{pep}(h) + 38.4\ \mathrm{sph}(h/20)$

$\gamma_1(h) = 36\ \mathrm{pep}(h) + 64\ \mathrm{sph}(h/20)$

$\gamma_2(h) = 36\ \mathrm{pep}(h) + 64\ \mathrm{sph}(h/20)$

$\gamma_{12}(h) = 18\ pep(h) + 38.4\ sph(h/20)$

$\gamma_1(h) = 64\ sph(h/20)$

$\gamma_2(h) = 16\ sph(h/20)$

$\gamma_{12}(h) = 38.4\ sph(h/20)$

3.2 The linear model of coregionalization

In this model, all simple and cross structures are linear combinations of the same basic structural components, and this can be interpreted by a decomposition of the variables themselves into linear combinations of corresponding components. This can be seen as a generalization to the multivariate case, of the nested structure for one variable.

3.2.1 Nested structure for one variable: the linear model of regionalization

Consider the example (see section 2.9.3) of lead grade Z(x), with mean 13.2 (%) and with a variogram made of 2 nested structural components, one being a nugget of 11, the other being spherical with sill 39 and range 60 :

$$\gamma(h) = 11\ Pep(h) + 39\ Sph(h/60)$$

that is: $\gamma(h) = b^1\gamma^1(h) + b^2\gamma^2(h)$

This variable can be interpreted as the sum of two components $Z^1(x)$ and $Z^2(x)$ not correlated with each other, with zero means and having respectively the structures $b^1\gamma^1(h) = 11\,Pep(h)$ and $b^2\gamma^2(h) = 39\,Sph(h/60)$:

$$Z(x) = Z^1(x) + Z^2(x) + m$$

It is often convenient to used normalized components (variance 1), here $Y^1(x)$ and $Y^2(x)$, with respective variograms $\gamma^1(h) = Pep(h)$ and $\gamma^2(h) = Sph(h/60)$, hence finally:

$$Z(x) = 3.32\ Y^1(x) + 6.24\ Y^2(x) + 13.2$$

that is: $Z(x) = a^1\ Y^1(x) + a^2\ Y^2(x) + m$

with $a^1 = \sqrt{b^1} = \sqrt{11} = 3.32$ and $a^2 = \sqrt{b^2} = \sqrt{39} = 6.24$ (beware : exponents are indices, not powers).

More generally, a variable with nested structure :

$$\gamma(h) = \sum_k b^k \gamma^k(h)$$

can be interpreted as a decomposition :

$$Z(x) = \sum_k Z^k(h) + m$$

or else, using normalized components:

$$Z(x) = \sum_k a^k Y^k(h) + m$$

with $b^k = \left(a^k\right)^2$ (factors are supposed here to have a mean of 0, or a mean of increments of 0, the mean m disappearing in the purely intrinsic case, i.e. not stationary).

3.2.2 Multivariate case: the linear model of coregionalization

Decomposition of structures

To introduce the notations, start from the example of the two variables that are the grades in lead and zinc :

$$\gamma_1(h) = 11\ Pep(h) + 39\ Sph(h/60)$$
$$\gamma_2(h) = 9\ Pep(h) + 15\ Sph(h/60)$$
$$\gamma_{12}(h) = \quad 14.5\ Sph(h/60)$$

which can be noted:

$$\gamma_1(h) = b_{11}^1 \gamma^1(h) + b_{11}^2 \gamma^2(h)$$
$$\gamma_2(h) = b_{22}^1 \gamma^1(h) + b_{22}^2 \gamma^2(h)$$
$$\gamma_{12}(h) = b_{12}^1 \gamma^1(h) + b_{12}^2 \gamma^2(h)$$

that is, matricially :

$$[\gamma_{ij}(h)] = [b_{ij}^1] \gamma^1(h) + [b_{ij}^2] \gamma^2(h)$$

with $\gamma^1(h) = Pep(h)$, $\gamma^2(h) = Sph(h/60)$, and matrices $b^1 = \left(b_{ij}^1\right) = \begin{pmatrix} 11 & 0 \\ 0 & 9 \end{pmatrix}$ and $b^2 = \left(b_{ij}^2\right) = \begin{pmatrix} 39 & 14.5 \\ 14.5 & 15 \end{pmatrix}$.

With the three variables, lead-zinc-silver, we will have :

$$\gamma_1(h) = 11\ Pep(h) + 39\ Sph(h/60)$$
$$\gamma_2(h) = 9\ Pep(h) + 15\ Sph(h/60)$$
$$\gamma_3(h) = 1.1\ Pep(h) + 1.8\ Sph(h/60)$$
$$\gamma_{12}(h) = \quad 14.5\ Sph(h/60)$$
$$\gamma_{13}(h) = \quad 5\ Sph(h/60)$$
$$\gamma_{23}(h) = \quad 3.8\ Sph(h/60)$$

now with $b^1 = \left(b_{ij}^1\right) = \begin{pmatrix} 11 & 0 & 0 \\ 0 & 9 & 0 \\ 0 & 0 & 1.1 \end{pmatrix}$ and $b^2 = \left(b_{ij}^2\right) = \begin{pmatrix} 39 & 14.5 & 5 \\ 14.5 & 15 & 3.8 \\ 5 & 3.8 & 1.8 \end{pmatrix}$

Generally, a linear model of coregionalization can be written:

- in term of variograms:

$$\gamma_{ij}(h) = \sum_k b_{ij}^k \gamma^k(h)$$

that is, matricially:

$$[\gamma_{ij}(h)] = \sum_k [b_{ij}^k] \gamma^k(h)$$

- or in covariances, when stationary:

$$C_{ij}(h) = \sum_k b_{ij}^k C^k(h)$$

$$C_{ij}(h) = \sum_k b_{ij}^k \rho^k(h)$$

with

$$\rho^k(0) = 1$$

where the $\gamma^k(h)$ represent for instance basic structures with different scales (hence $\gamma^1(h) = Pep(h)$, $\gamma^2(h) = Sph(h/60)$, …, but taking related structures, like adding Pep(h)+Sph(h/60) to the two previous structures, will be avoided).

Note that in such a linear model of coregionalization, the cross-covariances are necessarily even.

Decomposition of the variables into different scales

In such a model, the variables can be split in the following way (for instance, in the stationary case, with m_i , mean of Z_i and factors with mean 0):

$$Z_i(x) = \sum_k Z_i^k(x) + m_i$$

The components with scale k, $Z_i^k(x)$, having the same structure proportional to $\gamma^k(h)$ or $\rho^k(h)$, are intrinsically correlated, and $\left(b_{ij}^k\right)$ represents their variances-covariances matrix. Each matrix $b^k = \left(b_{ij}^k\right)$ must then be positive definite. In particular, we must have :

$$\left|b_{ij}^k\right| \le \sqrt{b_{ii}^k b_{jj}^k}$$

So a structural component (nugget, small ranged, or even long ranged) cannot be present in a cross-structure without being present in each of the two corresponding simple structures. However, a component can be present in two simple structures (or only one of these) without being present in the corresponding cross-structure.

The cross-variogram, which is a linear combination, with positive or negative coefficients, of the basic structures, lies necessarily within the envelop :

$$-\sum_k \sqrt{b_{ii}^k b_{jj}^k}\, \gamma^k(h) \le \gamma_{ij}(h) \le \sum_k \sqrt{b_{ii}^k b_{jj}^k}\, \gamma^k(h)$$

Factorization of each scale

Going further, it is possible to factorize the intrinsically correlated components for each scale k (see section 3.1.3). Then we obtain the following decomposition into factors :

$$Z_i(x) = \sum_k \sum_q a_{iq}^k Y_q^k(x) + m_i$$

where the $Y_q^k(x)$ with same k share the structure $\gamma^k(h)$ and with $b_{ij}^k = \sum_q a_{iq}^k a_{jq}^k$. Such a decomposition guarantees the (semi-) positive definiteness of the matrices b_{ij}^k.

In practice, once the basic structures are chosen, it is possible and very convenient to call for an automatic procedure that will be give the b_{ij}^k and a_{iq}^k.

3.2.3 Example

Consider the following linear model of coregionalization:

$$\begin{aligned}
\gamma_{11}(h) &= 25\gamma^1(h) + 100\gamma^2(h) + 16\gamma^3(h) \\
\gamma_{12}(h) &= \quad\quad\quad -40\gamma^2(h) + 12\gamma^3(h) \\
\gamma_{22}(h) &= 16\gamma^1(h) + 25\gamma^2(h) + 13\gamma^3(h) \\
\gamma_{13}(h) &= \quad\quad\quad -30\gamma^2(h) \\
\gamma_{23}(h) &= \quad\quad\quad 6\gamma^2(h) \\
\gamma_{33}(h) &= 9\gamma^1(h) + 17\gamma^2(h)
\end{aligned}$$

that is, matricially:

$$\begin{pmatrix} \gamma_{11}(h) & \gamma_{12}(h) & \gamma_{13}(h) \\ \gamma_{12}(h) & \gamma_{22}(h) & \gamma_{23}(h) \\ \gamma_{13}(h) & \gamma_{23}(h) & \gamma_{33}(h) \end{pmatrix}$$

$$= \begin{pmatrix} 25 & 0 & 0 \\ 0 & 16 & 0 \\ 0 & 0 & 9 \end{pmatrix} \gamma^1(h) + \begin{pmatrix} 100 & -40 & -30 \\ -40 & 25 & 6 \\ -30 & 6 & 17 \end{pmatrix} \gamma^2(h) + \begin{pmatrix} 16 & 12 & 0 \\ 12 & 13 & 0 \\ 0 & 0 & 0 \end{pmatrix} \gamma^3(h)$$

where $\gamma^1(h)$ = Pep(h), $\gamma^2(h)$ = Sph(h/20), and $\gamma^3(h)$ is large ranged (or even intrinsic, linear for instance). This is an authorized model, for we can write for instance (up to the means):

$$\begin{aligned}
Z_1 &= 5Y_1^1 + 10Y_1^2 + 4Y_1^3 \\
Z_2 &= 4Y_2^1 - 4Y_1^2 + 3Y_2^2 + 3Y_1^3 + 2Y_2^3 \\
Z_3 &= 3Y_3^1 - 3Y_1^2 - 2Y_2^2 + 3Y_3^2
\end{aligned}$$

where the Y_q^1 have the structure $\gamma^1(h)$ = Pep(h), the Y_q^2 the structure $\gamma^2(h)$ = Sph(h/20), and the Y_q^3 the large ranged structure. In this example, 3 factors are required to factorize each of the first two components, but 2 suffice for the third one (which is not present in Z_3).

Exercise: propose other possible factorizations of this model.

As a general rule, it is possible to factorize a linear model of coregionalization by taking a number of factors equal, at maximum, to the number of variables times the number of components.

3.2.4 Exercise : examples of linear models of coregionalization

Are the following models (authorized) linear models of coregionalization?

$\gamma_1(h) = 16\ \mathrm{sph}(h/20) + 4\ \mathrm{sph}(h/100)$

$\gamma_2(h) = 25\ \mathrm{sph}(h/20)$

$\gamma_{12}(h) = 12\ \mathrm{sph}(h/20)$

$\gamma_1(h) = 4\ \mathrm{sph}(h/100)$

$\gamma_2(h) = 25\ \mathrm{sph}(h/20)$

$\gamma_{12}(h) = 12\ \mathrm{sph}(h/20)$

$\gamma_1(h) = 4\ \mathrm{sph}(h/20)$

$\gamma_2(h) = 25\ \mathrm{sph}(h/20)$

$\gamma_{12}(h) = 12\ \mathrm{sph}(h/20)$

$\gamma_1(h) = 16\ \mathrm{sph}(h/20) + 4\ \mathrm{sph}(h/100)$

$\gamma_2(h) = 25\ \mathrm{sph}(h/20) + 9\ \mathrm{sph}(h/100)$

$\gamma_{12}(h) = 12\ \mathrm{sph}(h/20) - 3\ \mathrm{sph}(h/100)$

3.3 Exercise : model for error

Let $Z_1(x)$ (mean m_1, covariance $C_1(h)$, variogram $\gamma_1(h)$) known at some points and :

$$Z_2(x) = Z_1(x) + R(x)$$

a measure of $Z_1(x)$ (possibly at different points), affected by an error $R(x)$ with no spatial correlation with $Z_1(x)$ and being pure nugget.

Show the following relations:

Means :

$$m_2 = m_1 + m_R$$

Variogram of Z_2 :

$$\gamma_2(h) = \gamma_1(h) + \mathrm{var}(R)\ \mathrm{Pep}(h)$$

Cross-variogram of Z_1 and Z_2 :

$$\gamma_{12}(h) = \gamma_1(h)$$

Pseudo cross-variogram of Z_1 and Z_2 :

$\psi_{12}(h) = \gamma_1(h) + 0.5\ E[R^2]$

3.4 Exercise : multi-support

One single variable, but studied simultaneously on different supports, can be considered as a particular multivariate case. Although linear geostatistics can theoretically make the structural link between the supports, it is often more convenient to use the following approximate model (« discrete » or « discretized » model).

One chooses a reference support (say, a block v), with known variogram $\gamma_v(h)$, supposing that:

- a larger support can be considered as a support multiple of v (classical formula of regularization);
- the domain itself is partitioned ("discretized") into units of support v;
- a support finer than v (for instance, quasi-punctual x) is localized randomly, uniformly and independently in its support v (then let us denote it by $\underline{x}$).

Show (in the stationary case) that the linear regression of $Z(\underline{x})$, knowing its block Z(v), coincides with Z(v) (this would not be the case for a point x fixed in v), hence:

$$E[Z(\underline{x}) \mid Z(v)] = Z(v)$$

We have the decomposition:

$$Z(\underline{x}) = Z(v) + [Z(\underline{x}) - Z(v)]$$

Using for instance the hypothesis that, conditionally on its block Z(v), $Z(\underline{x})$ is independent of all other blocks or point samples, show that we have (with $\underline{x} \in v$, $\underline{y} \in v_h$, and v_h denoting v translated by h):

- point-block cross-structure:

$$\frac{1}{2} E\{[Z(\underline{y}) - Z(\underline{x})][Z(v_h) - Z(v)]\} = \gamma_v(h)$$

- discretized point structure:

$$\frac{1}{2} E\{[Z(\underline{y}) - Z(\underline{x})]^2\} = \frac{1}{2} E\{[(Z(\underline{y}) - Z(v_h)) + (Z(v_h) - Z(v)) + (Z(v) - Z(\underline{x}))]^2\}$$

$$= \gamma_v(h) + \mathrm{var}[Z(\underline{x}) - Z(v)]$$

(included h = 0, for $\underline{x} \neq \underline{y} \in v$)

So the model is defined solely by the variogram for the reference support $\gamma_v(h)$ and by $\mathrm{var}[Z(\underline{x}) - Z(v)]$. Note the analogy with the previous exercise: the term $Z(\underline{x}) - Z(v)$ is but the estimation error of Z(v) by $Z(\underline{x})$.

4 Cokriging

4.1 Generalities on the cokriging of one or several variables

To simplify the writing, and without loss of generality, we will consider the case of 2 variables $Z_1(x)$ and $Z_2(x)$, informed respectively on two sets of points S_1 and S_2, either identical (isotopic case) or not (heterotopic case). Cokriging allows us to account for information on both variables by estimating at best, linearly on Z_1 and Z_2 data : either $Z_1(x_0)$, or $Z_2(x_0)$, or each of these two variables, or these variables regularized, or else a linear combination like $Z_2(x_0) - Z_1(x_0)$. Cokriging ensures the consistency between estimations :

$$[Z_2(x_0) - Z_1(x_0)]^{CK} = Z_2(x_0)^{CK} - Z_1(x_0)^{CK}$$

which is not the case of kriging in general (see exercises 4.2.1, 4.2.2 and 4.3.1).

In the following we will begin by developing calculus considering the estimation of a target variable denoted $Z(x_0)$ (which can be $Z_1(x_0)$, $Z_2(x_0)$, or for instance $Z_2(x_0) - Z_1(x_0)\ldots$) in the most general way. Let us consider an estimator made by a linear combination of the data:

$$Z^*(x_0) = \sum_{S_1} \lambda_{1\alpha} Z_1(x_\alpha) + \sum_{S_2} \lambda_{2\alpha} Z_2(x_\alpha) + \lambda_0$$

which corresponds to the estimation error :

$$\varepsilon = Z_0 - Z_0^* = Z_0 - \sum_{S_1} \lambda_{1\alpha} Z_1(x_\alpha) - \sum_{S_2} \lambda_{2\alpha} Z_2(x_\alpha) - \lambda_0$$

We will then look for the weights such that the estimation error is an authorized linear combination (which allows to compute its mean and variance), has a zero expectation (unbiased estimator) and finally has a minimal variance (optimality). In terms of covariance, and with shortened notations, the variance to be minimized can be written :

$$\begin{aligned} Var(\varepsilon) \quad &= C(Z_0, Z_0) - 2\sum_{S_1} \lambda_{1\alpha} C(Z_0, Z_{1\alpha}) - 2\sum_{S_2} \lambda_{2\alpha} C(Z_0, Z_{2\alpha}) \\ &+ 2\sum_{S_1}\sum_{S_2} \lambda_{1\alpha}\lambda_{2\beta} C(Z_{1\alpha}, Z_{2\beta}) \\ &+ \sum_{S_1}\sum_{S_1} \lambda_{1\alpha}\lambda_{1\beta} C(Z_{1\alpha}, Z_{1\beta}) \\ &+ \sum_{S_2}\sum_{S_2} \lambda_{2\alpha}\lambda_{2\beta} C(Z_{2\alpha}, Z_{2\beta}) \end{aligned}$$

Clearly, if $Z(x_0)$ is present in the data, the cokriging of $Z(x_0)$, best linear estimator, coincides with itself: like kriging, cokriging is an exact interpolator.

4.2 Simple cokriging (or cokriging with known means)

We consider here the stationary case. In this case there is no condition to impose to be able to compute mean and variance of the estimation error. The non-bias can be written:

$$E(\varepsilon) = m - m_1 \sum_{S_1} \lambda_{1\alpha} - m_2 \sum_{S_2} \lambda_{2\alpha} - \lambda_0 = 0$$

hence :

$$\lambda_0 = m - m_1 \sum_{S_1} \lambda_{1\alpha} - m_2 \sum_{S_2} \lambda_{2\alpha}$$

which amounts to considering the centered variables:

$$Z^*(x_0) - m_0 = \sum_{S_1} \lambda_{1\alpha}[Z_1(x_\alpha) - m_1] + \sum_{S_2} \lambda_{2\alpha}[Z_2(x_\alpha) - m_2]$$

To minimize the variance of the error, we set its derivative to zero with respect to the weight of each data. Each data gives an equation, making the linear system:

(system of simple cokriging)

$$\begin{cases} \sum_{\beta \in S_1} \lambda_{1\beta} C(Z_{1\alpha}, Z_{1\beta}) + \sum_{\beta \in S_2} \lambda_{2\beta} C(Z_{1\alpha}, Z_{2\beta}) = C(Z_0, Z_{1\alpha}) \quad \forall x_\alpha \in S_1 \\ \sum_{\beta \in S_1} \lambda_{1\beta} C(Z_{1\beta}, Z_{2\alpha}) + \sum_{\beta \in S_2} \lambda_{2\beta} C(Z_{2\alpha}, Z_{2\beta}) = C(Z_0, Z_{2\alpha}) \quad \forall x_\alpha \in S_2 \end{cases}$$

that is, matricially :

$$\begin{bmatrix} \underline{C_{11}} & \underline{C_{12}} \\ \underline{C_{12}}^t & \underline{C_{22}} \end{bmatrix} \times \begin{bmatrix} \underline{\lambda_1} \\ \underline{\lambda_2} \end{bmatrix} = \begin{bmatrix} \underline{C_{01}} \\ \underline{C_{02}} \end{bmatrix}$$

covariance between data — covariance target-data

Solving the system leads to the optimal weights $\lambda_{1\alpha}$ and $\lambda_{2\alpha}$ (weights of simple cokriging).

The variance of the error (cokriging variance) can then be written:

$$\sigma_{CKS}^2 = \mathrm{var}(\varepsilon) = C(Z_0, Z_0) - \sum_{S_1} \lambda_{1\alpha} C(Z_0, Z_{1\alpha}) - \sum_{S_2} \lambda_{2\alpha} C(Z_0, Z_{2\alpha})$$

that is, matricially :

$$\sigma_{CKS}^2 = \mathrm{var}(\varepsilon) = C(Z_0, Z_0) - \begin{bmatrix} \underline{\lambda_1} \\ \underline{\lambda_2} \end{bmatrix}^t \begin{bmatrix} \underline{C_{01}} \\ \underline{C_{02}} \end{bmatrix}$$

4.2.1 Exercise: consistency between kriged values

Let Z_1 and Z_2 be two variables with simple and cross variograms $\gamma_1(h)$, $\gamma_2(h)$ and $\gamma_{12}(h)$, known at the same points. Show that, if $\gamma_2(h)$ and $\gamma_{12}(h)$ are proportional to $\gamma_1(h)$, the kriging of the difference between the two variables is equal to the difference between their kriging (which is not the case in general).

4.2.2 Exercise: consistency between cokriged values

Show that the cokriging of $Z_2(x_0) - Z_1(x_0)$ (or more generally of a linear combination of $Z_1(x_0)$ and $Z_2(x_0)$) is equal to the difference between the cokriging of $Z_1(x_0)$ and $Z_2(x_0)$.

(Use the linearity of cokriging weights as a function of the second member of the cokriging system.)

4.2.3 Exercise: simple cokriging

Let Z_1 and Z_2 be stationary with means m_1 and m_2. We know $Z_1(A)$ and $Z_2(B)$ at two points A and B. Simple cokriging of $Z_1(O)$ at point O ?

4.3 Ordinary cokriging (or cokriging with unknown means)

Consider the stationary case. We are able to compute the mean and variance of the estimation error. However, the means are supposed unknown here. Then we will specify the variable to be estimated, for instance $Z_1(x_0)$ (the general case will be examined in the exercise 4.3.1). The non-bias can then be written:

$$E(\varepsilon) = m_1\left(1 - \sum_{S_1} \lambda_{1\alpha}\right) - m_2 \sum_{S_2} \lambda_{2\alpha} - \lambda_0 = 0$$

To be satisfied whatever the values of m_1 and m_2 (and supposing these are not linked), we impose the following conditions:

$\lambda_0 = 0$ which definitely disappears from the equations

$\sum_{S_1} \lambda_{1\alpha} = 1$ (sum of weights for the target variable)

$\sum_{S_2} \lambda_{2\alpha} = 0$ (sum of weights for the complementary variable, or accordingly for each of the complementary variables)

We then look for the estimator:

$$Z_1^*(x_0) = \sum_{S_1} \lambda_{1\alpha} Z_1(x_\alpha) + \sum_{S_2} \lambda_{2\alpha} Z_2(x_\alpha)$$

for which the estimation variance is minimum, given the two above conditions. This leads to the following linear system of equations, where μ_1 and μ_2 are the Lagrange parameters introduced to satisfy these conditions :

(system of ordinary cokriging)

$$\begin{cases} \sum_{\beta\in S_1}\lambda_{1\beta}C\left(Z_{1\alpha},Z_{1\beta}\right)+\sum_{\beta\in S_2}\lambda_{2\beta}C\left(Z_{1\alpha},Z_{2\beta}\right)+\mu_1=C\left(Z_0,Z_{1\alpha}\right) & \forall x_\alpha\in S_1 \\ \sum_{\beta\in S_1}\lambda_{1\beta}C\left(Z_{1\beta},Z_{2\alpha}\right)+\sum_{\beta\in S_2}\lambda_{2\beta}C\left(Z_{2\alpha},Z_{2\beta}\right)+\mu_2=C\left(Z_0,Z_{2\alpha}\right) & \forall x_\alpha\in S_2 \\ \sum_{\alpha\in S_1}\lambda_{1\alpha}=1 & \\ \sum_{\alpha\in S_2}\lambda_{2\alpha}=0 & \end{cases}$$

that is, matricially:

$$\begin{bmatrix} \underline{C_{11}} & \underline{C_{12}} & 1 & 0 \\ \underline{C_{12}}^t & \underline{C_{22}} & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix} \times \begin{bmatrix} \underline{\lambda_1} \\ \underline{\lambda_2} \\ \mu_1 \\ \mu_2 \end{bmatrix} = \begin{bmatrix} \underline{C_{01}} \\ \underline{C_{02}} \\ 1 \\ 0 \end{bmatrix}$$

The variance of the error (cokriging variance) can be written:

$$\sigma_{CKO}{}^2 = \mathrm{var}(\varepsilon) = C\left(Z_0,Z_0\right)-\sum_{S_1}\lambda_{1\alpha}C\left(Z_0,Z_{1\alpha}\right)-\sum_{S_2}\lambda_{2\alpha}C\left(Z_0,Z_{2\alpha}\right)-\mu_1$$

that is, matricially:

$$\sigma_{CKO}{}^2 = \mathrm{var}(\varepsilon) = C\left(Z_0,Z_0\right)-\begin{bmatrix} \underline{\lambda_1} \\ \underline{\lambda_2} \\ \mu_1 \\ \mu_2 \end{bmatrix}^t \begin{bmatrix} \underline{C_{01}} \\ \underline{C_{02}} \\ 1 \\ 0 \end{bmatrix}$$

When cross-covariances are even, it can be checked that, thanks to the conditions on the weights (and unlike simple cokriging), the system can be written nearly identically in variogram (suffices to change each covariance term, simple or cross, by the opposite of corresponding variogram : the system is then identical, but with changed signs for Lagrange parameters). In fact, as will now be seen, ordinary cokriging then coincides with an intrinsic cokriging.

4.3.1 Exercise: consistency between cokriged values

Suppose we want to estimate, by ordinary cokriging, a linear combination of $Z_1(x_0)$ and $Z_2(x_0)$, for instance $Z_2(x_0)-Z_1(x_0)$, from data on $Z_1(x)$ and $Z_2(x)$.

- Which are the non-bias conditions to be introduced?
- Establish the cokriging system.
- Check that the cokriging of $Z_2(x_0)-Z_1(x_0)$ is the difference between the cokriging of $Z_2(x_0)$ and $Z_1(x_0)$.

-

4.4 Intrinsic cokriging

As a matter of fact, in the intrinsic case, the conditions :

$$\sum_{S_1} \lambda_{1\alpha} = 1, \quad \sum_{S_2} \lambda_{2\alpha} = 0$$

are precisely those for which the estimation error :

$$\begin{aligned} \varepsilon &= Z_1(x_0) - Z_1^*(x_0) \\ &= Z_1(x_0) - \sum_{S_1} \lambda_{1\alpha} Z_1(x_\alpha) - \sum_{S_2} \lambda_{2\alpha} Z_2(x_\alpha) \end{aligned}$$

is an authorized linear combination. Its mean is zero (this ensures no-bias). Its variance can be developed using the simple and cross variograms, and leads to a cokriging system identical to the previous one, but written in variograms (signs being changed for Lagrange parameters).

4.5 Cokriging with unknown, but linked, means

Cokriging will change if the means, supposed to be unknown, are linked. Take the example of two variables with means that are unknown but equal: $m_1 = m_2$.

The non-bias:

$$E(\varepsilon) = m_1 \left(1 - \sum_{S_1} \lambda_{1\alpha} - \sum_{S_2} \lambda_{2\alpha} \right) - \lambda_0 = 0$$

must be respected whatever the mean, so we have:

$$\begin{aligned} &\lambda_0 = 0 \\ &\sum_{S_1} \lambda_{1\alpha} + \sum_{S_2} \lambda_{2\alpha} = 1 \end{aligned}$$

Show that this leads to the following cokriging system :

$$\begin{cases} \sum_{\beta \in S_1} \lambda_{1\beta} C(Z_{1\alpha}, Z_{1\beta}) + \sum_{\beta \in S_2} \lambda_{2\beta} C(Z_{1\alpha}, Z_{2\beta}) + \mu = C(Z_0, Z_{1\alpha}) & \forall x_\alpha \in S_1 \\ \sum_{\beta \in S_1} \lambda_{1\beta} C(Z_{1\beta}, Z_{2\alpha}) + \sum_{\beta \in S_2} \lambda_{2\beta} C(Z_{2\alpha}, Z_{2\beta}) + \mu = C(Z_0, Z_{2\alpha}) & \forall x_\alpha \in S_2 \\ \sum_{\alpha \in S_1} \lambda_{1\alpha} + \sum_{\alpha \in S_2} \lambda_{2\alpha} = 1 \end{cases}$$

4.5.1 Exercise: filtering errors

We consider here the model of error seen in exercise 3.3.

Errors are supposed to be systematic $E[R] \neq 0$ and the means are supposed unknown: kriging and kriging system of $Z_1(x_0)$?

Errors are supposed to be non-systematic E[R]=0 and m_1 unknown: kriging and kriging system of $Z_1(x_0)$?

4.6 Remarks

4.6.1 Cokriging variance

Adding data from Z_2, to estimate the variable Z_1, can only increase the precision : the cokriging variance can only be less or equal to the kriging variance from the same Z_1 data (as a matter of fact, kriging amounts to giving a weight of zero, generally not optimal, to Z_2 data).

4.6.2 Cokriging weights

Beware to the amplitude of weights : the weights of the target variable Z_1 have no unit, but those of a co-variable Z_2 are in (unit Z_1 /unit Z_2)

In ordinary cokriging, the weights of a co-variable Z_2 sum to 0. The negative weights, if associated to high Z_2 values, can easily make negative the estimation of a positive Z_1 variable.

5 Simplification of cokriging

Cokriging is a heavy operation when the number of variables, or of sampled points, is high. Hence the interest of using models allowing simplifications, when this is compatible with the structure and the localization of data.

5.1 Preliminary remarks

5.1.1 Consistency of the estimations

As it has been seen, cokriging ensures the consistency between the estimations of different variables (the cokriging of a linear combination being the linear combination of the cokrigings). Thus the cokriging of Z_2 - Z_1 (from data on Z_1 and Z_2 for instance) is equal to the difference between the cokrigings of Z_2 and Z_1.

On the other hand, the cokriging of a quantity being the best linear combination of the data, it is possible to replace these data by any equivalent (generating the same space) family of linear combinations. Thus it is theoretically equivalent to make a cokriging (of Z_1, Z_2 or Z = Z_2 - Z_1, etc) from equivalent data on Z_1 and Z_2, on Z_1 and Z, on Z_2 and Z, or else on Z and $Z_1 + Z_2$. In practice some choices can be preferable for the simplifications they bring to cokriging.

5.1.2 Redundant data

Let us first examine redundancy in a very simple example of kriging.

The simple kriging of Z(x) from duplicated values at point x+h (two data values equal to Z(x+h)) can be written :

$$Z(x)^K = \lambda Z(x+h) + \lambda' Z(x+h) + (1-\lambda-\lambda')m$$

The kriging system is made out of two equations identical to:

$$\lambda C(0) + \lambda' C(0) = C(h)$$

hence:

$$\lambda + \lambda' = \frac{C(h)}{C(0)}$$

The duplicates being consistent (same value), the kriging is perfectly determined:

$$Z(x)^K = (\lambda + \lambda')Z(x+h) + (1-\lambda-\lambda')m$$

and coincides with the kriging from only one of these data values (in case of inconsistency - two different values at the same data point – kriging would not be defined).

In practice however, redundancies will be avoided, for the solution of the kriging system is theoretically not unique (indetermination) and can happen to be unstable.

It is the same for cokriging. One will avoid for instance to simultaneously use as data the values of Z_1, Z_2, and of $Z = Z_2 - Z_1$ at the same point. Similarly, when having p variables with constant sum, known at the same points, only p-1 will be retained. In theory, cokriging will not depend on the choice of the selected variables, but some choices may bring simplification to cokriging – which is useful in practice – and then be more judicious.

5.2 Variables without spatial correlation

Let us consider a set of variables:

$$Z_1, Z_2, ..., Z_p$$

without spatial correlation:

$$C_{ij}(h) = Cov\left[Z_i(x), Z_j(x+h) \right] = 0 \quad \forall h$$

or

$$\gamma_{ij}(h) = 0 \quad \forall h$$

Then (simple or ordinary) cokriging of one of the variables, for instance Z_1, is reduced to its own kriging:

$$Z_i(x_0)^{CK} = Z_i(x_0)^K$$

One can check that the kriging weights are the solution of the cokriging system. However this supposes that the data (possibly heterotopic) are made of the values of these variables. Moreover the simplification disappears if the means are unknown but linked. See the following exercises for these exceptions.

5.2.1 Exercise: composite variable

Let $Z_1(x)$ and $Z_2(x)$ be two variables without spatial correlation, and $Z(x) = Z_2(x) - Z_1(x)$.

We know $Z(A)$ and $Z_2(B)$.

Simple cokriging of $Z_1(O)$?

Ordinary cokriging of $Z_1(O)$?

5.2.2 Exercise: linked means (or linked drifts)

Let $Z_1(x)$ and $Z_2(x)$ be two variables without spatial correlation, but having the same unknown mean m.

We know $Z_1(A)$ and $Z_2(B)$.

Cokriging of $Z_1(O)$?

5.3 Factorized model of « isofactorial » type

In these models, it is possible to go from the variables $Z_1, Z_2, ..., Z_p$ to "factors" Y^k without spatial correlation, and inversely:

$$Y^k(x) = \sum_i e_i^k Z_i(x)$$

$$Z_i(x) = \sum_k a_i^k Y^k(x)$$

the number of the Y^k not exceeding that of the Z_i.

We have:

$$\gamma_{ij}(h) = \sum_k a_i^k a_j^k \gamma^k(h)$$

If the structure of data locations allows it (for instance in the isotopic case, where the Z_i, hence the Y^k, are known at the same points), we are back to the previous case of variables without spatial correlation (the Y^k), from where the cokriging of the Z_i : can be deduced by linear combination:

$$[Z_i(x)]^{CK} = \sum_k a_i^k [Y^k(x)]^K$$

Isofactorial models are very much used in non-linear geostatistics (disjunctive kriging).

5.3.1 Search of factors

It is important to note that the factors coming from a statistical method of factorization (principal components, for instance), uncorrelated at the same point x, have a priori no reason to be spatially uncorrelated. Moreover any multivariate coregionalization model cannot necessarily be decomposed into factors with no spatial correlation. However it is possible to determine a family of factors with no correlation at distance 0 (i.e. at same point) as well as at another previously chosen distance, say h_0 : this is what Desbarats and Dimitrakopoulos (2000) propose, referring to Switzer and Green (1984), by searching the factors explaining a spatial continuity (correlogram) less and less marked for this distance (« min/max autocorrelation factors »). After this, whether the spatial correlation between factors can be neglected at other distances must be checked. This method is attractive when the number of variables is high, for it allows us to retain the best spatially structured factors, by opposition to the quasi-nugget factors (note that the best structured factors do not necessarily explain the most of statistical variability, which is an essential difference with the principal components).

5.4 Self-krigeable variables

5.4.1 Preliminary exercise: screen effect in cokriging

Simple cokriging of $Z_1(x)$ knowing $Z_1(x+h)$ and $Z_2(x+h)$?

Condition for the weight of Z_2 to be zero?

5.4.2 Self-krigeability

Let $Z_1(x)$, …, $Z_i(x)$, …, $Z_p(x)$ be a set of variables. One variable, for instance Z_1, is said to be self-krigeable if, in the isotopic case, its cokriging coincides with its own kriging.

A condition for this is that its cross-structure with the other variables is identical (more exactly, proportional) to its own structure:

$$C_{1j}(h) \equiv C_1(h)$$

or

$$\gamma_{1j}(h) \equiv \gamma_1(h)$$

for all j. (From the preliminary exercise, one can show that this is a necessary and sufficient condition in the stationary case with known means; the intrinsic case is more complex.)

Suppose that there are two self-krigeable variables, for instance Z_1 and Z_2. Then:

- either they have the same structure, in which case they are intrinsically correlated as well as their linear combinations (see section 3.1.3) :

$$\gamma_1(h) \equiv \gamma_{12}(h) \equiv \gamma_2(h)$$

- or they have different structures, in which case they are spatially uncorrelated:

$$\gamma_{12}(h) = 0 \quad \forall h$$

It can be deduced more generally that the self-krigeable variables (or linear combinations of these) can be dispatched into groups, without spatial correlation between groups, each group being made up of intrinsically correlated variables with a different structure. (Mathematically: self-krigeable RFs belong to orthogonal subspaces, each of these being characterized by its own structure.)

Two cases are particularly interesting:

- it is possible to go, by linear transformation, from the set of original variables to new variables, spatially uncorrelated and with structures generally different from each other: those are the factorized models seen above;
- all variables (as well as their linear combinations) are self-krigeable and belong to the same group: this is the model of intrinsic correlation seen above, characterized by identical simple and cross structures.

5.5 Model with residual between 2 variables

This model (sometimes called Markov model) is a bivariate factorized model, where one variable is self-krigeable.

5.5.1 Characterization

Consider two variables $Z_1(x)$ and $Z_2(x)$ with a cross-structure identical to the simple structure of one of these, say $Z_1(x)$:

$$C_{12}(h) = a\ C_1(h)$$

or:

$$\gamma_{12}(h) = a\ \gamma_1(h)$$

(In other words, $Z_1(x)$ is self-krigeable.) Then the residual of the linear regression of $Z_2(x)$ on $Z_1(x)$:

$$R(x) = Z_2(x) - a\ Z_1(x) - b$$

(by construction with mean zero and uncorrelated to $Z_1(x)$ at same point) has no spatial correlation with $Z_1(x)$:

$$Cov\left[Z_1(x), R(x+h)\right] = C_{12}(h) - aC_1(h) = 0$$

The model can then be factorized with factors $Z_1(x)$ and $R(x)$:

$$Z_2(x) = a\ Z_1(x) + b + R(x)$$
$$\gamma_2(h) = a^2\gamma_1(h) + \gamma_R(h)$$

Note that knowing only the structure of the variable $Z_1(x)$ gives no information on the residual structure. (Moreover, in principle, nothing prevents the structure of the residual from being more continuous than that of Z_1.)

Remark: in the pure intrinsic case (variogram but no stationary covariance), we must think in increments: b disappears, and the increments R(x+h)-R(x) = [Z_2(x+h) - Z_2(x)] - a [Z_1(x+h) - Z_1(x)] are uncorrelated to [Z_1(x+h) - Z_1(x)].

5.5.2 Simplification of cokriging

Suppose that $Z_2(x)$ is known on the same set (isotopic case), or on a subset, of the datapoints of $Z_1(x)$. Then knowing Z_1 and Z_2 on their sets of points is equivalent to knowing Z_1 and R on the same sets, respectively. Thanks to the factorization, we then have:

$$\begin{cases} Z_1^{\,CK} = Z_1^{\,K} \\ Z_2^{\,CK} = a\ Z_1^{\,K} + b + R^K \end{cases}$$

This can be simple cokriging (from simple krigings of Z_1 and of R), or ordinary cokriging. Replacing R by its definition, one can check that, in the ordinary cokriging of Z_2, the sum of weights for Z_2 is 1, and the sum of weights for Z_1, 0.

5.5.3 Kriging of residual or collocated cokriging

Consider the heterotopic case where $Z_1(x)$ is densely known, i.e. known at any desired point. Since kriging (or cokriging) is an exact interpolator, its value at a point x_0 coincides with the true value $Z_1(x_0)$. The cokriging of $Z_2(x_0)$ corresponds to the estimation obtained by kriging the residual of $Z_2(x_0)$ on $Z_1(x_0)$, which can be written :

$$\begin{aligned} Z_2^{\,CK}(x_0) &= a\ Z_1(x_0) + b + R^K(x_0) \\ &= a\ Z_1(x_0) + b + \sum_\alpha \lambda_\alpha \left[Z_{2\alpha} - aZ_{1\alpha} - b \right] \\ &= \sum_\alpha \lambda_\alpha Z_{2\alpha} + a\left(Z_1(x_0) - \sum_\alpha \lambda_\alpha Z_{1\alpha} \right) + b\left(1 - \sum_\alpha \lambda_\alpha \right) \end{aligned}$$

This is a collocated cokriging, for it makes use of Z_1 only at x_0 and where Z_2 is known : the other points where Z_1 is available receive a zero weight and do not appear.

This can be a simple cokriging (corresponding to simple kriging of the residual), or an ordinary cokriging (corresponding to the ordinary cokriging of the residual, with the condition $\sum \lambda_\alpha = 1$ which filters b out).

5.5.4 Exercise : « dislocated cokriging »

Consider, in the stationary case, a model with residual, say:

$Z_2(x) = Z_1(x) + R(x)$

where R(x) is spatially uncorrelated with $Z_1(x)$ and has a nugget structure. For instance $Z_2(x)$ can represent a measure of $Z_1(x)$ spoiled with an error R(x) (see exercise 3.3). This is a model with residual with $\gamma_{12}(h) = \gamma_1(h)$ and $\gamma_2(h) = \gamma_1(h) + \gamma_R(h)$.

Consider an arbitrary heterotopic configuration, where Z_1 and Z_2 are known at possibly different points.

Show that :

- the simple cokriging of $Z_1(x_0)$ depends on $Z_2(x)$ only at points where $Z_1(x)$ is unknown (it is a dislocated cokriging, opposite collocated cokriging) ;
- the simple cokriging of $Z_2(x_0)$, where it is unknown, depends on $Z_2(x)$ also only at points where $Z_1(x)$ is unknown.

5.5.5 Exercise: Model with residual and intrinsic correlation

Let $Z_1(x)$ and $Z_2(x)$ be two variables with $Z_1(x)$ self-krigeable. In other words, we are in a model with residual.

Show that the variables are intrinsically correlated if and only if the structure of the residual is identical to this of Z_1.

Z_1, Z_2, or any linear combination, is then self-krigeable.

5.5.6 Exercise: Examples of models with residual

Are those models with residuals? Decomposition of the variables? Structure of the residual?

$\gamma_1(h) = 16\ \mathrm{sph}(h/20)$

$\gamma_2(h) = 25\ \mathrm{sph}(h/20) + 25\ \mathrm{sph}(h/100)$

$\gamma_{12}(h) = 16\ \mathrm{sph}(h/20)$

$\gamma_1(h) = 16\ \mathrm{sph}(h/20) + 16\ \mathrm{sph}(h/100)$

$\gamma_2(h) = 25\ \mathrm{sph}(h/20) + 25\ \mathrm{sph}(h/100)$

$\gamma_{12}(h) = 16\ \mathrm{sph}(h/20)$

$\gamma_1(h) = 16\ \mathrm{sph}(h/20) + 16\ \mathrm{sph}(h/100)$

$\gamma_2(h) = 25\ \mathrm{sph}(h/20)$

$\gamma_{12}(h) = 16\ \mathrm{sph}(h/20)$

5.5.7 Passage to a model with drift

Consider a model with residual, and suppose, like for the kriging of the residual, that Z_2 is densely known. Suppose moreover that Z_1 and R are, not only spatially uncorrelated, but independent RFs. By conditioning the model on the values z_1 of Z_1, this variable gets deterministic and we have:

$$Z_2(x) = a\ z_1(x) + b + R(x) = E\left[Z_2(x)\right] + R(x)$$

which allows us to consider $a\ z_1(x) + b$ as the drift of $Z_2(x)$:

$$E\left[Z_2(x)\right] = a\ z_1(x) + b$$

defined externally from $z_1(x)$.

Be careful: z_1 being deterministic, the (usual, centered) covariance to be used for Z_2 is now this of the residual:

$$Cov[Z_2(x), Z_2(x+h)] = E\{[Z_2(x) - a\ z_1(x) - b][Z_2(x+h) - a\ z_1(x+h) - b]\}$$
$$= \text{cov}[R(x), R(x+h)]$$

5.5.8 Kriging with external drift

Consider the expression of the cokriging estimator, but now with z_1 deterministic. This does not change the estimation obtained by kriging the residual. If b is unknown, the estimation gets rid of it by introducing the condition $\sum \lambda_\alpha = 1$ (ordinary kriging). Now if both a and b are unknown, the estimation can be achieved by imposing the conditions :

$$\sum_\alpha \lambda_\alpha = 1 \text{ and } \sum_\alpha \lambda_\alpha z_{1\alpha} = z_{10}$$

This is the kriging with external drift. One can show that the weights are the solution of the following linear system:

(system of kriging with external drift)

$$\begin{cases} \sum_\beta \lambda_\beta Cov\left(R_\alpha, R_\beta\right) + \mu_0 + \mu_1 z_{1\alpha} = Cov\left(R_0, R_\alpha\right) \quad \forall x_\alpha \\ \sum_\alpha \lambda_\alpha = 1 \\ \sum_\alpha \lambda_\alpha z_{1\alpha} = z_{10} \end{cases}$$

that is, matricially:

$$\begin{bmatrix} \underline{C} & 1 & \underline{z_1} \\ 1 & 0 & 0 \\ \underline{z_1}^t & 0 & 0 \end{bmatrix} \times \begin{bmatrix} \underline{\lambda} \\ \mu_0 \\ \mu_1 \end{bmatrix} = \begin{bmatrix} \underline{C_{01}} \\ 1 \\ z_{10} \end{bmatrix}$$

The kriging variance can be written:

$$\sigma_{KDE}{}^2 = \text{var}(\varepsilon) = Cov\left(R_0, R_0\right) - \sum_\alpha \lambda_\alpha Cov\left(R_0, R_\alpha\right) - \mu_0 - \mu_1 z_{10}$$

that is, matricially:

$$\sigma_{KDE}{}^2 = \mathrm{var}(\varepsilon) = Cov\left(R_0, R_0\right) - \begin{bmatrix} \underline{\lambda} \\ \mu_0 \\ \mu_1 \end{bmatrix}^t \begin{bmatrix} \underline{C_{01}} \\ 1 \\ z_{10} \end{bmatrix}$$

5.5.9 Remarks on the kriging of residual and on the kriging with external drift

- The variable Z_1, known everywhere, hence not to be estimated, does not really require any hypothesis of stationarity.

- The basic hypothesis is the spatial uncorrelation, or even the independence, between the dense variable and the residual of the target variable on this.

- The estimation depends on the structure of the residual only. The direct determination of this structure is however a problem in external drift (the increments of the residual, and so its variogram, depend a priori on *a*, considered as unknown); cross-validation can then be used to estimate or refine this structure.

It is possible to extend the methods of kriging of the residual and of kriging with external drift to several conditioning variables. Then the structure to be used is this of the residual of the linear regression, at the same point, of the target variable on the conditioning variables. This residual is supposed to be spatially uncorrelated to the conditioning variables.

6 Kriging analysis

Kriging analysis allows us, within a linear model of regionalization or coregionalization, to filter or to map one or several components.

6.1 Stationary case with one variable

Consider a variable Z(x) with mean m and with structure:

$$\gamma(h) = b^1\gamma^1(h) + b^2\gamma^2(h) + b^3\gamma^3(h)$$

where the structural components $\gamma^k(h)$ describe for instance different scales. This model can be interpreted as a decomposition of the variable itself into components at different scales (chosen with means 0):

$$Z(x) = Z^1(x) + Z^2(x) + \cdots + m$$

We have:

$$Cov\left[Z^k(x), Z(x+h)\right] = Cov\left[Z^k(x), Z^k(x+h)\right]$$

which allows us to cokrige each $Z^k(x_0)$ from the sole data $Z_\alpha = Z(x_\alpha)$ on Z.

If the mean m is known, this will be a simple cokriging:

$$\left(Z^k(x_0)\right)^{CKS} = \sum_\alpha \lambda_\alpha \left(Z_\alpha - m\right)$$

As an exercise, one can develop the system (identical to the kriging system of $Z(x_0)$, but changing the right hand side according to the component Z^k to be estimated).

When the mean is unknown, be careful to ensure non-bias. Hence for instance the estimation of $Z^3(x_0)$:

$$\left[Z^3(x_0)\right]^{CK} = \sum \lambda_\alpha Z_\alpha$$
$$\varepsilon = Z_0^3 - \sum \lambda_\alpha Z_\alpha$$
$$E(\varepsilon) = 0 - m\sum \lambda_\alpha = 0 \;\Rightarrow\; \sum \lambda_\alpha = 0$$

and this of $[Z^3(x_0) + m]$:

$$\left[Z^3(x_0) + m\right]^{CK} = \sum \lambda_\alpha Z_\alpha$$
$$\varepsilon = Z^3(x_0) + m - \sum \lambda_\alpha Z_\alpha$$
$$E(\varepsilon) = 0 + m\left(1 - \sum \lambda_\alpha\right) = 0 \;\Rightarrow\; \sum \lambda_\alpha = 1$$

We have the consistency:

$$\left[Z_0^3 + m\right]^{CK} = \left(Z_0^3\right)^{CK} + m^K$$

as well as globally:

$$Z_0^{\,K} = \left(Z_0^1\right)^{CK} + \left(Z_0^2\right)^{CK} + \left(Z_0^3\right)^{CK} + m^K$$

due to the linearity of the weights with respect to the second member of the system, which changes with the quantity to be estimated.

Be careful: while the hidden factors Z^k (depending on the structural decomposition, unknown and non objective) are not spatially correlated, their estimations, since they are made from the same data, are correlated.

6.2 Intrinsic case with one variable

The intrinsic case is more complex. Consider for instance a variable with structure:

$$\gamma(h) = b^1\gamma^1(h) + b^2\gamma^2(h) + b^3\gamma^3(h)$$

This can be interpreted as the sum of 3 intrinsic components without spatial correlation :

$$Z(x) = Z^1(x) + Z^2(x) + Z^3(x)$$

Suppose that we want to estimate $Z^3(x_0)$. The error can be written:

$$\varepsilon = Z_0^3 - \sum \lambda_\alpha Z_\alpha^k = Z_0^3 - \sum \lambda_\alpha (Z_\alpha^1 + Z_\alpha^2 + Z_\alpha^3) = Z_0^3 - \sum \lambda_\alpha Z_\alpha^3 - \sum \lambda_\alpha Z_\alpha^1 - \sum \lambda_\alpha Z_\alpha^2$$

For this to be an authorized linear combination, $\sum \lambda_\alpha$ should be at the same time equal to 1 and to 0, which is impossible. This leads us to consider only one purely intrinsic component (say Z^3), the others (Z^1 and Z^2) being stationary with mean zero. We then have, for instance:

$$\left[Z^2(x_0)\right]^{CK} = \sum \lambda_\alpha Z_\alpha$$
$$E(\varepsilon) = 0 \Rightarrow \sum \lambda_\alpha = 0$$

or

$$\left[Z^3(x_0)\right]^{CK} = \sum \lambda_\alpha Z_\alpha$$
$$E(\varepsilon) = 0 \Rightarrow \sum \lambda_\alpha = 1$$

Note the similarity between the estimation of the intrinsic component Z^3 in this frame, and the estimation of (Z^3+m) in the stationary case with unknown mean:

$$Z(x) = Z^1(x) + Z^2(x) + Z^3(x) \ +\text{m}$$

6.3 Multivariable case

The multivariate case is similar:

$$\gamma_{ij}(h) = \sum_k b_{ij}^k \gamma^k(h)$$

In the stationary case for instance, this can be interpreted as :

$$
\begin{aligned}
Z_i(x) &= \sum_k Z_i^k(x) + m_i \\
&= \sum_k \sum_q a_{iq}^k Y_q^k(x) + m_i
\end{aligned}
$$

Kriging analysis consists of extracting, by cokriging, the Z_i^k (non objective, given by the structural decomposition) or the Y_q^k (depending moreover on the method chosen to factorize the components Z_i^k at scale k).

7 Synthesis and other remarks

7.1 Multivariate models

We have seen:

- the model of intrinsic correlation, where all structures are identical ;
- the linear model of coregionalization, where all structures are linear combinations of a given number of basic structures;
- the factorized models, essentially the particular case of the model with residual, where the factors consist of the self-krigeable variable and the residual (this model, which leads to the estimation by kriging the residual and by kriging with external drift, can represent notably a model for errors or a multi-support model).

It is possible to complement such models, or to imagine many more, according to the situation, for instance :

- models including delays between some components ;
- models made from convolutions (regularizations) or from derivatives ;
- models of indicators resulting from thresholding a RF;
- etc.

7.2 Simulation

The models presented here are particularly suited to simulations in the (multi-)gaussian case (i.e. any linear combination is gaussian), the absence of correlation being then equivalent to independence:

- the simulation of p intrinsically correlated variables can be obtained from independent simulations of the factors resulting from an arbitrary factorization (in number p, in general);
- the linear model of coregionalization, for p variables in function of k basic structures, can be simulated also from independent simulations of factors (in number kp, in general) ;
- the factorized model with p variables can be simulated from independent simulations of each factor (in number p in general, that is, two in the bivariate model with residual), each factor possibly having however a nested structure.

7.3 Simplification of cokriging

If one of the variables, say $Z_1(x)$, has a cross-structure with each other variable that is identical to its own structure, then it is self-krigeable : its cokriging coincides with its own kriging, in the isotopic case, but also in the heterotopic case if Z_1 is known at all sampled points.

When the number of variables is two and in such a configuration, it can be advantageous to replace $Z_2(x)$ by the residual $R(x)$ of its linear regression on the self-krigeable variable $Z_1(x)$ at the same point, this residual being not spatially correlated with $Z_1(x)$. This is the model with residual, where cokriging can be obtained by kriging the factors $Z_1(x)$ and $R(x)$, at least if $Z_1(x)$ is known at every sampled point.

All these remarks apply in particular to the model of intrinsic correlation : the cokriging of Z_1 is identical to its kriging, supposing that Z_1 is known at every data point.

Lastly, more generally in the factorized models, cokriging can be reduced to kriging each factor, providing that the data can be expressed into values taken individually by these factors (this is notably so in the isotopic case).

7.4 Cokriging neighbourhood

In the monovariate case, the choice of the cokriging neighbourhood (by this we mean the neighbouring data that will be retained, for instance the N nearest data per angular sector) depends on :

- the location (number, distribution, in particular geographic density) of the data ;
- the structure of the variable;
- the limit distance of the assumed hypothesis of stationarity (on the variable or its increments).

Different parameters can help for this choice : kriging variance (for the precision), weight of the mean in simple kriging (to see notably the influence of the mean in ordinary kriging), slope of the linear regression true value on estimated value in ordinary kriging (to get closer to conditional unbiasness where this slope is 1).

The same considerations can be also applied in multivariate, but the situation can be far more complex.

Note that the cokriging of a variable is sure to be at least as precise as its own kriging, if the cokriging neighbourhood includes the kriging neighbourhood (which is not guaranteed in practice, in particular because of the number of data), and if the mono- and multi-variate models are consistent.

To ensure consistency between the cokriged values of the different variables, it is necessary to take a neighbourhood identical for the estimation of each variable. In the isotopic case, this suggests choosing the neighbourhood according to the variable that requires the largest neighbourhood (while a smaller neighbourhood could be sufficient for the other variables).

The heterotopic case is far more difficult than the isotopic case, for besides the geographical location of the values for each variable, their mutual arrangement has its role (through the model).

For instance, when estimating a variable Z_1 from data on this variable and from a densely known other variable Z_2, a collocated neighbourhood consists of retaining a neighbourhood that includes, in addition to the data on Z_1, those on Z_2 at target point and at points where Z_1 is informed, but not those on Z_2 at other places. Such a neighbourhood is optimal in the model with residual where the cross-structure between Z_1 and Z_2 is identical to the structure of Z_2 : in effect, cokriging is collocated in this model, and this is also the model for kriging with external drift. However such a neighbourhood can lead to a substantial loss of information in other cases.

For instance, if, in an arbitrary heterotopic stationary case, the cross-structure between Z_1 and Z_2 is identical to this of Z_1 with a nugget residual from $Z_2(x)$ on $Z_1(x)$, simple cokriging of Z_1 at an unsampled point makes use of Z_2 only where Z_1 is unknown: the neighbourhood is dislocated, in opposition to the previous collocated neighbourhood which would then ignore the second variable Z_2, except at target point if available.

7.5 Uses of cokriging

Cokriging allows us to improve the consistency between the estimations of different variables.

In the isotopic case, it can improve the estimation of a variable, if its simple structure contrasts with its cross-structures with the other variables.

In the heterotopic case, cokriging allows us in particular, to account for an auxiliary variable that is better sampled : in the case of a densely sampled auxiliary variable and of the model with residual, notably, cokriging is collocated and coincides with kriging the residual, and kriging with external drift can also be used.

Cokriging also permits us to extract structural components (kriging analysis).

Lastly, it can be used to condition a multivariate Gaussian simulation.

8 Indicators

In this section, we will look at the tools of linear geostatistics when applied to the case of indicators. We will consider successively:

- the case of one indicator;
- the case of an indicator and another variable;
- the case of several indicators;

and will present, in this last case, some basic models and the cokriging of indicators.

8.1 Indicator of a set

8.1.1 Indicator variable of a set

The presence or the absence of a set A (a geological facie for instance) at a point x can be coded by the 0/1 « indicator » variable :

$$1_A(x) = \begin{cases} 1 & si \quad x \in A \\ 0 & si \quad x \notin A \end{cases}$$

Since a point belongs necessarily, either to A, or to its complement A^c , their indicators are respectively 1 and 0 or else 0 and 1. We have then:

$$1_A(x) + 1_{A^c}(x) = 1$$

and $\quad 1_A(x) 1_{A^c}(x) = 0$

Note the following properties :

$1_A(x) 1_A(y)$ is 1 if both x and y belong to A, 0 otherwise.

$$1_A(x) = [1_A(x)]^2$$

Mean of an indicator at different points :

$\frac{1}{N} \sum_i 1_A(x_i)$ = proportion of points x_i that belong to A.

$\frac{1}{V} \int_V 1_A(x)\, dx$ = proportion of points of V that belong to A.

8.1.2 Random model

The set A is now considered as a random set (a set whose shape and location are random). Its indicator variable is represented by a 0/1 Random Function. This indicator satisfies:

$E[1_A(x)] = P(x \in A)$, say p_A independent of x, if this probability is stationary (i.e. invariant under translation)

$$E[1_{A^c}(x)] = 1 - p_A$$

$$\text{var}[1_A(x)] = p_A(1 - p_A) = \text{var}[1_{A^c}(x)] \leq 0.25$$

Non-centered covariance of (the indicator of) A :

$$E[1_A(x)1_A(x+h)] = P(x \in A,\ x+h \in A)\ \text{, say}\ K_A(h)\ \text{ if stationary}$$

Centered covariance:

$$\text{cov}[1_A(x), 1_A(x+h)] = C_A(h)\ \text{ if stationary}$$

Variogram :

$$\frac{1}{2}E([1_A(x+h)-1_A(x)]^2) = \frac{1}{2}\left[P(x \in A,\ x+h \notin A) + P(x \notin A,\ x+h \in A)\right]$$

that is : $P(x \in A,\ x+h \notin A) = \gamma_A(h) = \gamma_{A^c}(h) \leq 0.5$

in the stationary case. (As its variance is bounded, an indicator which satisfies the intrinsic hypothesis, having then stationary increments, is necessarily stationary.)

Note that the variograms of a set and of its complement are the same. On the other hand, a set and its complementary do not have necessarily the same connectivity (think to a set made of disjoint – hence disconnected – lenses, surrounded by the complementary set: this being connected, as it is possible to go from one point to another one without getting off the set). It follows that a variogram is not capable of characterizing the connectivity of a set.

We have :

$$P(x+h \in A \mid x \in A) = \frac{K_A(h)}{p_A}$$

$$P(x+h \notin A \mid x \in A) = \frac{\gamma_A(h)}{p_A}$$

The variogram of a random set cannot be arbitrary. In particular we have necessarily:

$$|1_A(z)-1_A(x)| \leq |1_A(z)-1_A(y)| + |1_A(y)-1_A(x)|$$

hence: $0.5\,E[1_A(z)-1_A(x)]^2 \leq 0.5\,E[1_A(z)-1_A(y)]^2 + 0.5\,E[1_A(y)-1_A(x)]^2$

and so the triangular inequality:

$$\gamma_A(h+h') \leq \gamma_A(h) + \gamma_A(h')$$

It follows that:

$$\gamma_A(2h) \leq 2\,\gamma_A(h)$$

hence a behaviour at the origin:

$$\gamma_A(h) \propto |h|^\alpha\ \ \alpha \leq 1$$

α is related to the irregularity of the frontiers of A. It is equal to 1, its maximum (variogram being linear at short distances), for a set with regular contours. A parabolic behaviour (in h^2) at short distance is then not admissible.

8.1.3 Kriging of an indicator

Let us code the data $x_\alpha \in A, ..., x_\beta \notin A, ...$ by the indicators $1_A(x_\alpha)$, equal to 0 or 1.

The conditional probability $P(x_0 \in A | x_\alpha \in A, ..., x_\beta \notin A, ...)$ is in general not accessible. Then one can use a kriging of the indicator, for instance a simple kriging :

$$[1_A(x_0)]^K = \sum_{\alpha=1}^{n} \lambda_\alpha 1_A(x_\alpha) + \left[1 - \sum_{\alpha=1}^{n} \lambda_\alpha \right] p_A$$

which has the meaning of a probability, but can be out of [0, 1], which requires a post-processing.

8.2 Case of an indicator and of another variable

8.2.1 General case

Consider a set A represented by its indicator $1_A(x)$, and a variable Z(x). Then it is possible to compute and compare the means $E[Z(x) | x \in A]$ and $E[Z(x) | x \in A^c]$ of Z(x) on A and on its complementary A^c respectively.

Let us look at the interpretation of the cross-structures in this case:

The non-centered cross-covariance:

$$E[1_A(x)Z(x+h)] = E[Z(x+h) | x \in A]P(x \in A)$$

gives, up to the factor $P(x \in A)$, the mean of Z at distance h of a point x of A. The ratio between the cross-variogram and the simple variogram of A provides the mean deviation of Z between a pair of points distant of h across the frontier:

$$\frac{\gamma_{A,Z}(h)}{\gamma_A(h)} = E[Z(x+h) - Z(x) | x+h \in A, x \notin A]$$

(assuming this quantity is symmetrical in h).

Of course, it is possible to compare the spatial variability (variogram) of Z(x) computed, on the whole domain, on A (the points x and x+h within A), on its complementary, or across A and A^c.

If one is interested in the behaviour of Z(x) within A, it is also possible to consider the new variable (restraining Z(x) to A) $Z(x)1_A(x)$, equal to Z(x) within A and to 0 outside. This will be the following case.

8.2.2 Indicator and variable being simultaneously zero

Here we consider a set A represented by its indicator $1_A(x)$, and a variable Z(x) equal to 0 outside A. We then have the relation:

$$Z(x) = Z(x)1_A(x)$$

The ratio between the non-centered cross-covariance and the non-centered covariance of A :

$$\frac{E[1_A(x)Z(x+h)]}{E[1_A(x)1_A(x+h)]} = E[Z(x+h) | x \in A, x+h \in A]$$

is equal to the mean of Z within the intersection between the set A and its translation by h. The ratio between the cross-variogram and the simple variogram of A can then be written (again assuming symmetry):

$$\frac{\gamma_{A,Z}(h)}{\gamma_A(h)} = E[Z(x+h) \mid x+h \in A, x \notin A]$$

Its evolution in function of h measures the effect of the borders of A on Z. It is constant when there is no border effect. This is, in particular, the case when Z(x) can be considered as resulting from the restriction to A of a variable independent of A (its variographic structure is then accessible from the sole data of Z(x) in A).

8.2.3 Service variables

By service variables, we mean a pair of additive variables (then suited to linear estimation), the ratio of which is a variable of interest which is not additive, typically in mining:

T(x), represents the ore,

Q(x), the metal,

with ratio $Z(x) = \frac{Q(x)}{T(x)}$ equal to the grade in metal.

This is the classical case of a stratiform or vein-type orebody studied in 2D with:

T(x) thickness of the layer at 2D point x,

Q(x) metal accumulation on the thickness of the layer,

with ratio equal to the grade of the thickness.

Consider a block defined by the support v of its 2D size. By definition, its mean grade Z(v), ratio between metal and ore, is:

$$Z(v) = \frac{\int_v Q(x)dx}{\int_v T(x)dx} = \frac{\frac{1}{v}\int_v Q(x)dx}{\frac{1}{v}\int_v T(x)dx}$$

that is, the ratio between regularized Q and T :

$$Z(v) = \frac{Q(v)}{T(v)}$$

or else: $Z(v) = \frac{1}{v}\int_v Z(x)\left(\frac{T(x)}{\frac{1}{v}\int_v T(y)dy}\right)dx$

So the mean grade is not the direct mean of grades, but the mean of grades weighted by thicknesses.

Similarly, one can deduce an estimation of the grade (at a point or on a block) from a linear estimation of the service variables thickness and accumulation (for instance by kriging, or better, cokriging).

Due to the relation Q(x)=Z(x)T(x), the 3 variables are not independent and can present various relations, so it is instructive to examine their scatterplots. In particular:

Z(x) can be negatively correlated to T(x), notably if T(x) and Q(x) are independent;

Z(x) can, on average, increase with T(x) (positive correlation), then Q(x) also;

Z(x) can be uncorrelated to or independent of T(x), for instance if Q(x) varies proportionally to T(x).

The structural relations, also, can be various. In particular the observation of a cross-structure between T and Q being proportional to the simple structure of T corresponds to a model with residual with T being self-krigeable:

$$Q(x) = aT(x) + b + R(x)$$

This can be notably the case when, on average, the metal varies linearly with the ore :

$$E[Q(x)|T(x)] = aT(x) + b \text{ (proportional T(x) if b = 0)}$$

that is:

$$E[Z(x)|T(x)] = a + b/T(x) \text{ (constant if b = 0)}$$

Yet another example of service variables is given by the recoverable reserves at cut-off grade z, that is (denoting by Z(x) a 3D grade):

- the ore, given by the indicator $1_{Z(x)\geq z}$ (equal to 0 or 1), this indicator regularized (for instance, the proportion of ore on the height of benches), or $1_{Z(v)\geq z}$ for a selection support v;
- the metal, given by $Z(x)1_{Z(x)\geq z}$ (equal to 0 or Z(x)), this metal regularized (for instance on the height of benches), or $Z(v)1_{Z(v)\geq z}$,

the ratio between these two service variables (or their estimations) representing the corresponding recoverable grade. We are here in the case of an indicator and of another variable (the metal) being simultaneously zero.

Considering several possible cut-offs, as well as a change of support (going from samples at quasi-punctual support x to a selection block v) is part of non-linear geostatistics.

8.3 Indicators of several sets

Clearly distinguish:

- independent random sets, then:

$$P(x \in A,\ x \in B) = P(x \in A)\ P(x \in B)$$

- disjoint sets $A \cap B = \varnothing$, then:

$$1_A(x) 1_B(x) = 0 \text{ and so: } P(x \in A,\ x \in B) = 0$$

- nested sets $A_2 \subset A_1$:

$$1_{A_2}(x) \le 1_{A_1}(x)$$

8.3.1 Cross-variography

non-centered cross-covariance:

$$E[1_A(x) 1_B(x+h)] = P(x \in A,\ x+h \in B) \text{ that is } K_{AB}(h) \text{ if stationary}$$

cross-covariance :

$$C_{AB}(h) = K_{AB}(h) - p_A p_B$$

cross-variogram:

$$\gamma_{AB}(h) = \frac{1}{2} E[1_A(x+h) - 1_A(x)][1_B(x+h) - 1_B(x)]$$

8.3.2 Independent sets

Then we have $C_{AB}(h) = 0$ and $\gamma_{AB}(h) = 0$.

8.3.3 Disjoint sets

We then have (in the symmetric case) :

$$\gamma_{AB}(h) = \frac{1}{2} E[1_A(x+h) - 1_A(x)][1_B(x+h) - 1_B(x)]$$

$$= -P(x \in A,\ x+h \in B)$$

$$\frac{-\gamma_{AB}(h)}{p_A} = P(x+h \in B \mid x \in A)$$

$$\frac{-\gamma_{AB}(h)}{\gamma_A(h)} = P(x+h \in B \mid x \in A, x+h \notin A)$$

This last ratio is constant when no border effects A-> B (i.e. from A to B).

Moreover, if the disjoint sets (say A, B and C) make a partition, their indicators are linked by the closeness relation :

$$1_A(x)+1_B(x)+1_C(x)=1$$

hence the relations previously mentioned on this subject, for instance :

$$C_{AB}(h)+C_{AC}(h)=-C_A(h) \quad \text{negative at h = 0}$$

and $\gamma_{AB}(h)+\gamma_{AC}(h)=-\gamma_A(h)\le 0$

8.3.4 Nested sets

Consider for instance:

$$A_2 \subset A_1 \subset A_0$$

The ratio between non-centered cross-covariance of two sets and non-centered covariance of the larger set is:

$$\frac{E[1_{A_1}(x)1_{A_2}(x+h)]}{E[1_{A_1}(x)1_{A_1}(x+h)]}=\frac{P(x\in A_1, x+h\in A_2)}{P(x\in A_1, x+h\in A_1)}=P(x+h\in A_2 \mid x\in A_1, x+h\in A_1)$$

it is constant when no border effect from A_2 to A_1.

We have, for the cross-variogram (symmetrical case) :

$$\gamma_{A_1A_2}(h)=\frac{1}{2}E[1_{A_1}(x+h)-1_{A_1}(x)][1_{A_2}(x+h)-1_{A_2}(x)]$$

$$=P(x\notin A_1,\ x+h\in A_2)$$

$$\frac{\gamma_{A_1A_2}(h)}{1-p_{A_1}}=P(x+h\in A_2 \mid x\notin A_1)$$

$$\frac{\gamma_{A_1A_2}(h)}{\gamma_{A_1}(h)}=P(x+h\in A_2 \mid x\notin A_1, x+h\in A_1)$$

This last ratio is constant when there is no border effect from A_2 to A_1.

8.3.5 Remark : relation between disjoint sets and nested sets

The case of nested sets can be transported to this of disjoint sets, posing for instance :

$$C = A_2$$
$$B = A_1 - A_2$$
$$A = A_0 - A_1 - A_2$$

In theory, the families of the indicators of the disjoint sets and of the nested sets are equivalent. In particular, the absence of border effects from A_2 to A_1 (nested sets) corresponds to the absence of border effects from A to B (disjoint sets).

Inversely, it is interesting to cumulate disjoint sets into nested sets only when these sets can be ordered (for instance geological facies with decreasing granulometry).

8.3.6 Cokriging of indicators

The indicators of disjoint or nested sets are not independent. In general cokriging allows one to improve the consistency between the estimations of the different indicators (however, similarly to kriging, the cokriging of an indicator can get out of [0,1]).

8.4 Some basic models

Three types of basic models are illustrated at the end of the document (figures kindly provided by Hélène Beucher, showing images and variographic figures in the two axes, output from Isatis software). Each case corresponds to 3 disjoint facies, colored in grey (G), yellow (J) and red (R), in proportion nearly equal in the three models, but different from one facies to another.

8.4.1 Model without border effect

More precisely, this is a model without border effect when going out of grey. Going out of grey, one finds yellow or red according to a probability independent of the distance. Hence for instance a cross-variogram between red and grey proportional to the variogram of grey.

This model has been constructed in the following way :

(1) set the whole space in red ;

(2) generate the yellow facies;

(3) superimpose independently the grey facies;

(it would be possible to continue to superimpose sequentially other independent sets).

Because of the hierarchy of the construction, this model is suited to a description by cumulating facies into nested sets, here: [red] $\subset$ [red + yellow] $\subset$[red + yellow + grey].

Consider then a model without border effect when going up into the nested sets $A_0 \supset A_1 \supset A_2 \ldots$

We then have for instance :

$$\frac{\gamma_{A_1 A_2}(h)}{\gamma_{A_1}(h)} = P(x+h \in A_2 \mid x \notin A_1, x+h \in A_1) = \frac{p_{A_2}}{p_{A_1}}$$

One can show that the indicators can be factorized from the residuals of indicators :

$$H_1(x) = \frac{1_{A_1}(x)}{p_{A_1}} - 1$$

$$\ldots, \quad H_{i+1}(x) = \frac{1_{A_{i+1}}(x)}{p_{A_{i+1}}} - \frac{1_{A_i}(x)}{p_{A_i}}, \ldots$$

In the isotopic case, it is sufficient to krige these residuals to deduce the cokriging of all indicators.

8.4.2 Mosaic model with independent valuations

Such a model can be built from a stationary partition of space into tiles (in the figure: a Voronoi partition, given by the zones of influence of Poisson points). Each tile is given a facies A, B or C (here grey, yellow or red), according to an a priori probability p_A, p_B or p_C, independently of the other tiles. (On the figure, all initial tiles cannot be distinguished: in particular, the large areas in grey are made of several tiles, each set – by pure randomness – in grey.)

Denoting:

$\rho(h)$ = P(x and x+h belong to the same tile)

we have:

$$K_A(h) = P(x \in A,\ x+h \in A) = p_A \rho(h) + p_A{}^2 [1-\rho(h)]$$

$$K_{AB}(h) = P(x \in A,\ x+h \in B) = p_A p_B [1-\rho(h)]$$

$$\gamma_A(h) = p_A(1-p_A)[1-\rho(h)]$$

$$\gamma_{AB}(h) = -p_A p_B [1-\rho(h)]$$

All simple and cross-variograms are identical. There is no border effect between facies.

The mosaic model with independent valuations corresponds precisely to the case where cokriging does not bring any improvement to kriging, at least in an isotopic case. As a matter of fact, the indicators are intrinsically correlated, hence cokriging coincides with kriging :

$$[1_A(x_0)]^{CK} = [1_A(x_0)]^K$$

In other words, to estimate the indicator $1_A(x_0)$, only the $1_A(x_\alpha)$ matter, not the $1_B(x_\alpha)$, which is intuitive from the construction of the model.

8.4.3 Model of diffusive type

This corresponds to a gaussian RF thresholded into grey, yellow and red : one crosses yellow to go from grey to red. Hence the existence of border effects: when going out from grey, for instance, the probability to have red increases with the distance. The facies are naturally ordered, and such a model is very much used in the oil industry to describe facies with increasing or decreasing granulometry (for instance decreasing from sand – in red – to shale – in grey – passing through silt – yellow). Because of this order, it is interesting to cumulate the facies (nested facies in the sense of increasing or decreasing granulometry).

The simple and cross-structures of the indicators are not simple models : they are a function of the simple structure of the gaussian RF and of the thresholds (see course on simulations). In practice, the fitting of such structures of facies indicators requires one to search for the structure of the hidden gaussian variable. This model is very widely used in simulation (rather than in estimation by cokriging), the facies being then directly obtained by thresholding the simulated gaussian variables.

8.4.4 Towards non-linear geostatistics

Applying thresholds to a RF leads to nested sets (for instance, the sets of points where the RF exceeds the values 0, 1, 2, 3…), the values taken by a discrete or discretized RF making a partition of space into disjoint sets (for instance, the sets of points where the RF is respectively equal to 0, 1, 2, 3…). The cokriging of indicators, then called disjunctive kriging, allows one to estimate any function of the RF.

This simplifies into kriging of indicators in the case of the mosaic model with independent valuations (characterized by an absence of destructuration of the nested sets when the cut-off varies).

Disjunctive kriging is essentially used within the frame of isofactorial models, where it is obtained by kriging separately the factors :

- residuals of indicators in the hierarchical models without border effects when going up (or down) in the levels;
- other families of factors in the models of diffusive type (Hermite polynomials in the anamorphosed gaussian case, Laguerre polynomials for gamma distributions, etc.).

See course on non-linear geostatistics.

9 References

Chilès, J.-P., and Delfiner, P. 1999. *Geostatistics: Modeling spatial uncertainty*. Wiley, New York, 695 p.

Desbarats, A. J. and Dimitrakopoulos, R. 2000. Geostatistical simulation of regionalized pore-size distributions using min/max autocorrelation factors. *Math. Geol*. Vol. 32 No 8, 919-942.

Künsch, H. R., Papritz, A., and Bassi, F. 1997. Generalized cross-covariances and their estimation. *Math. Geol*. Vol. 29 No 6, 779-799.

Switzer, P. and Green, A. A. 1984. Min/max autocorrelation factors for multivariate spatial imaging. Technical report no 6, Department of Statistics, Stanford University, 14 p.

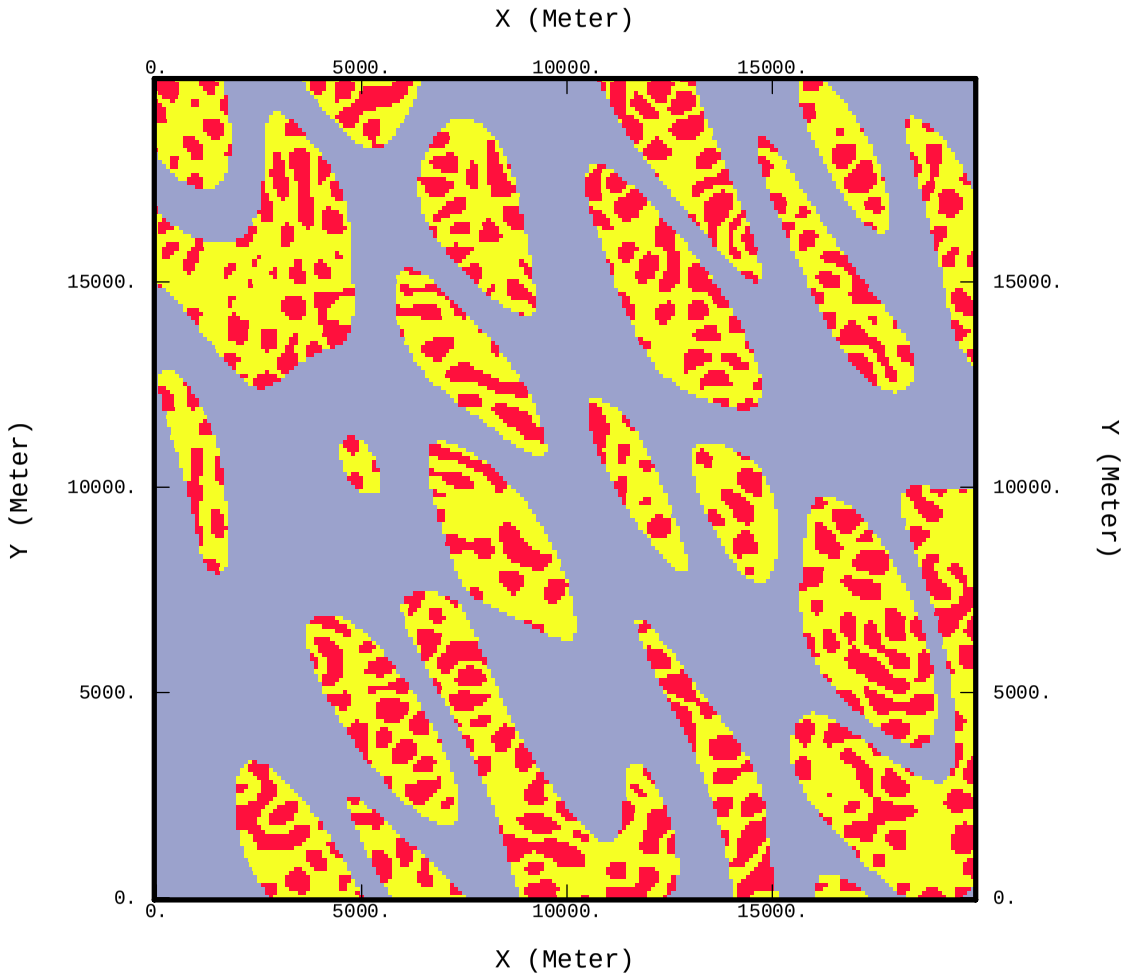
Wackernagel, H. 1993. Cours of géostatistique multivariable. Centre of Géostatistique, Ecole des Mines of Paris, Fontainebleau, 80 p.

Wackernagel, H. 1995. *Multivariate geostatistics : an introduction with applications*. Springer, Berlin. 256 p.

10 Index

(referring to sections)

Model without border effect (when going out of grey)



Proportions (%) :

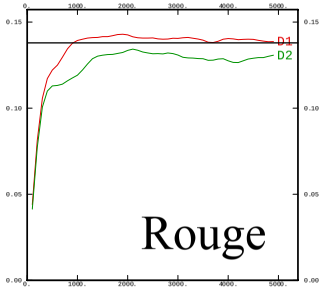
grey gris 54.5

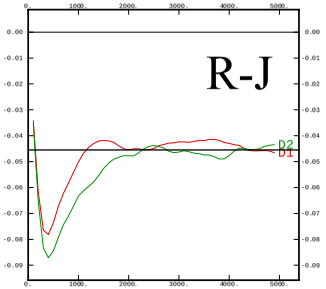
yellow jaune 30.3

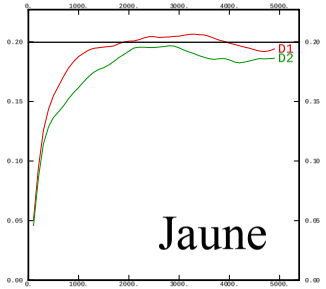
red rouge 15.2

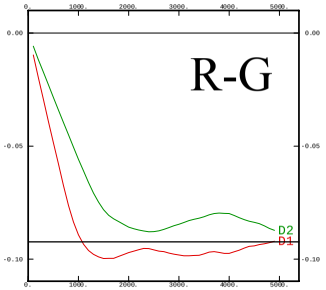
Model without border effect (when going out of grey)

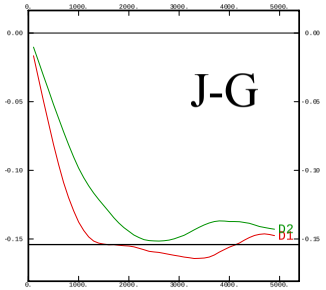
- simple and cross variograms

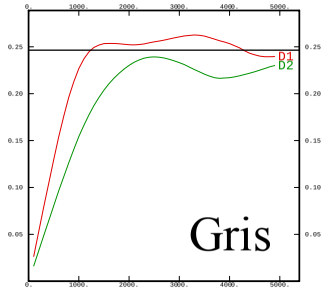






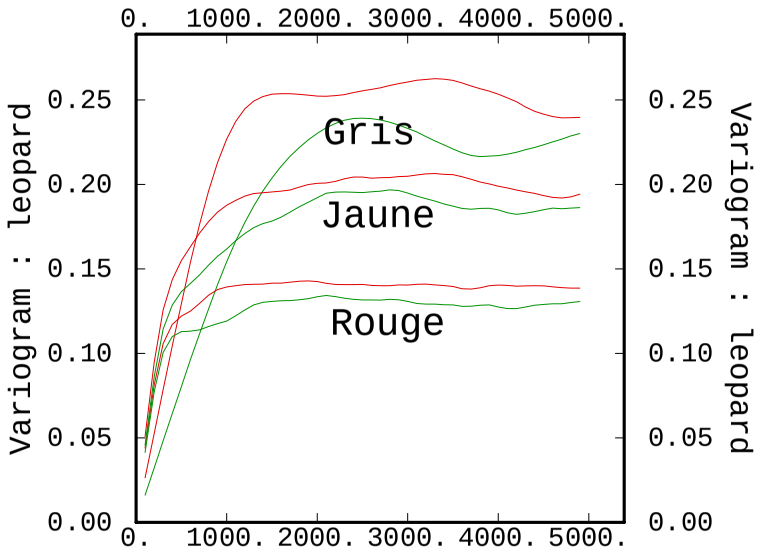




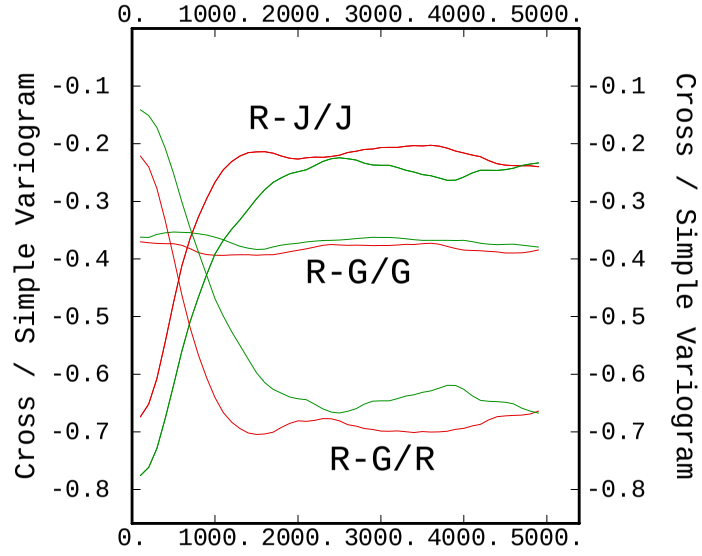


Model without border effect (when going out of grey)

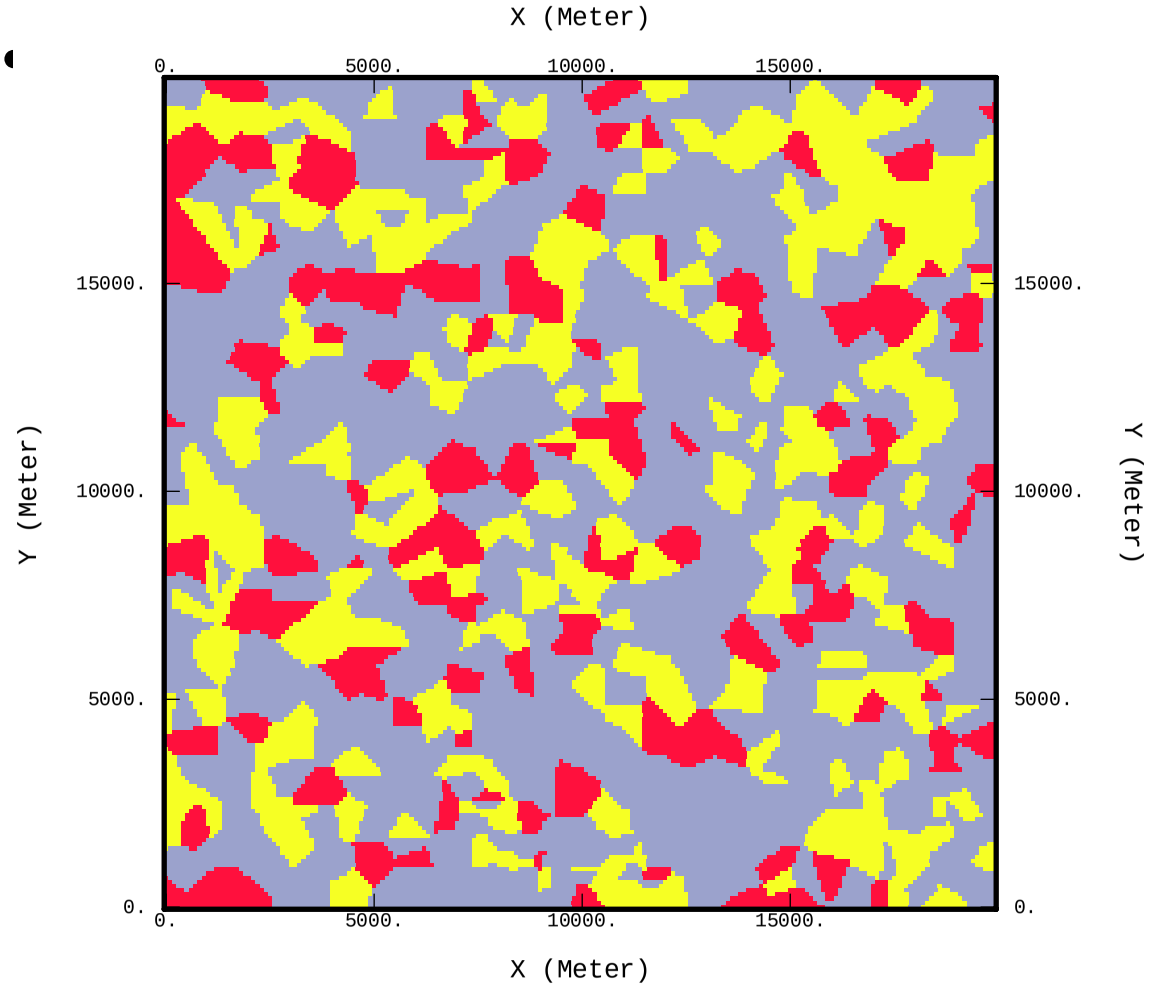
- simple variograms



ratio cross/simple



Mosaic model with independent valuations



Proportions (%) :

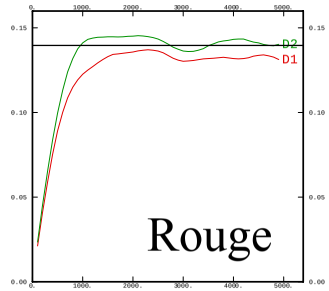
grey gris 55.9

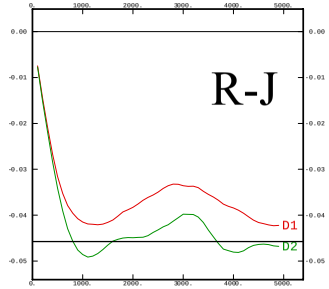
yellow jaune 27.3

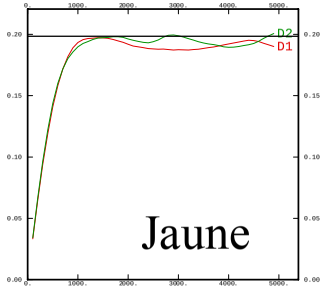
red rouge 16.8

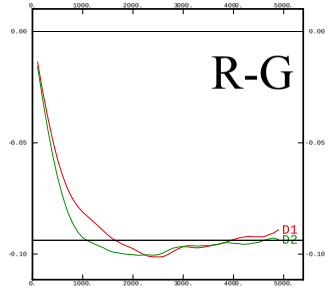
Mosaic model with independent valuations

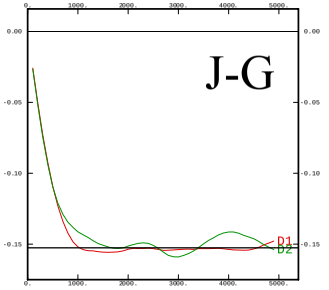
- simple and cross variograms

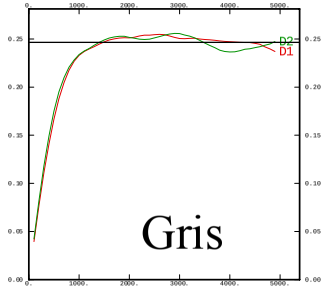






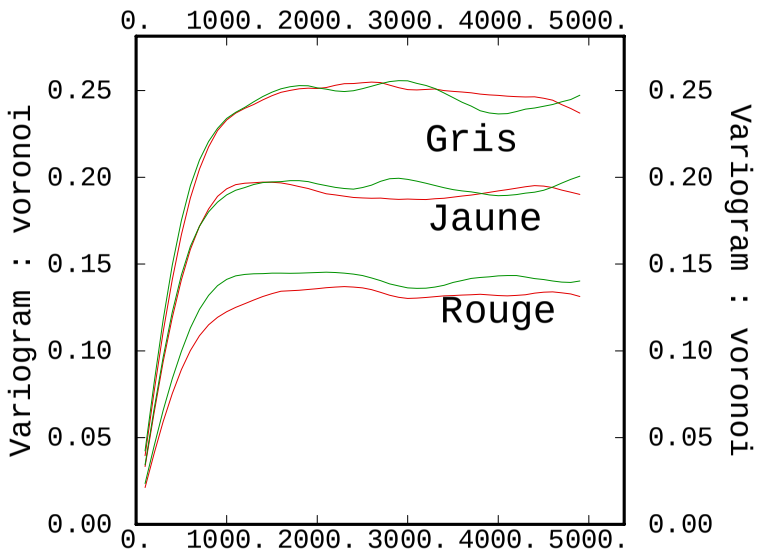




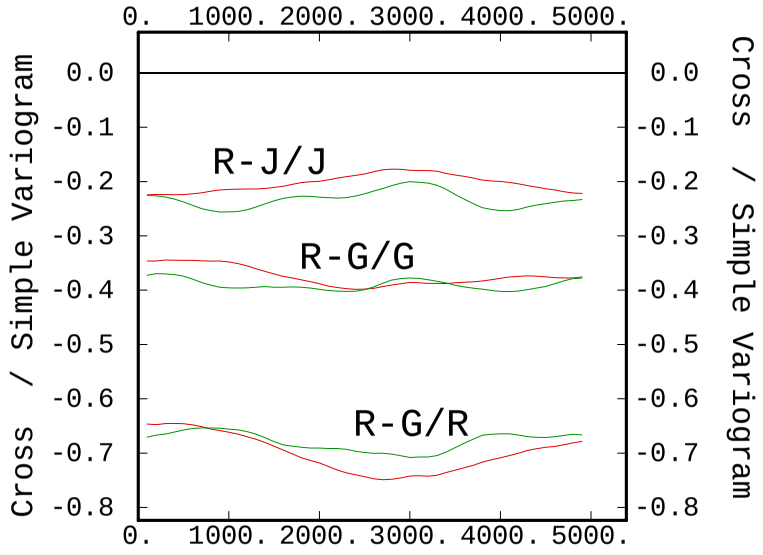


Mosaic model with independent valuations

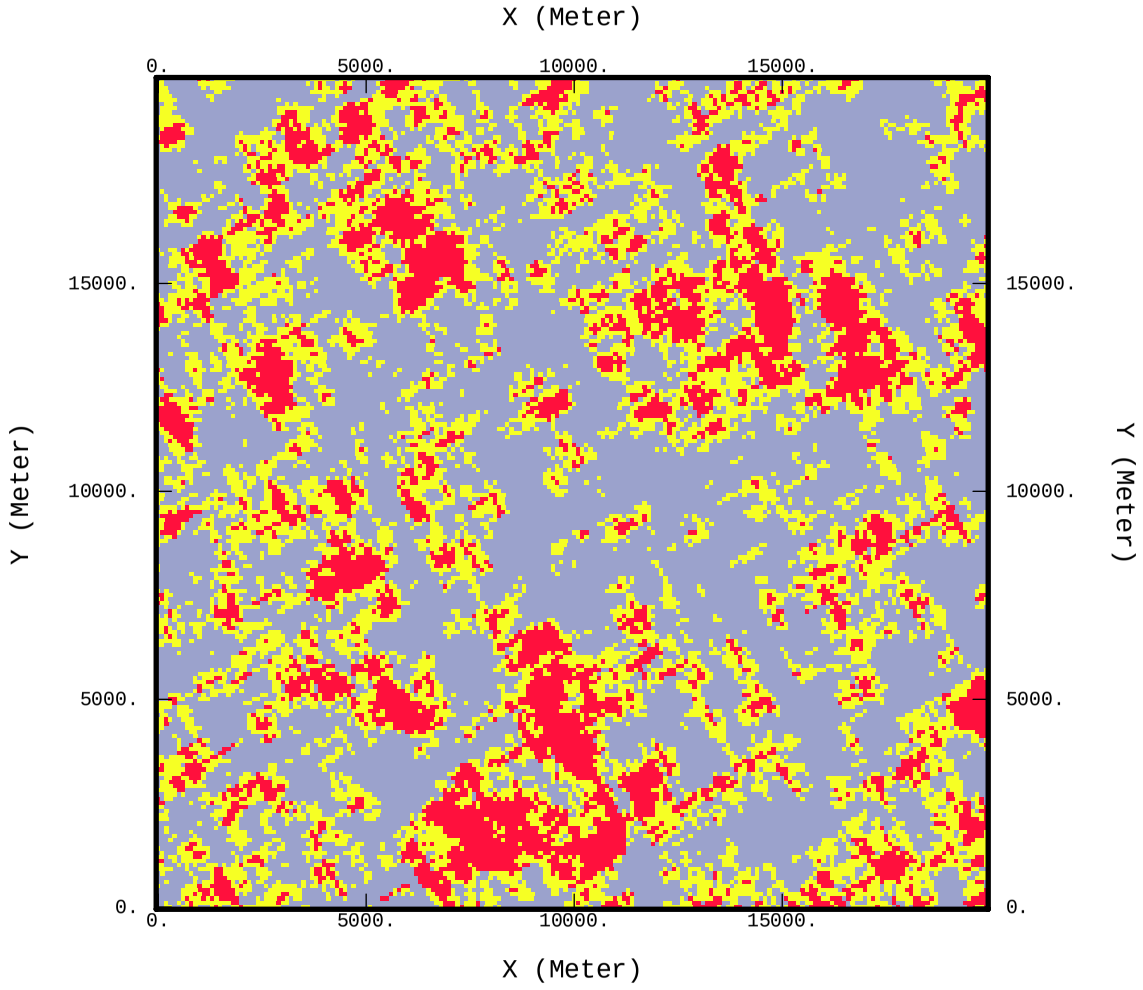
- simple variograms



ratio cross/simple



Model of diffusive type



Proportions (%):

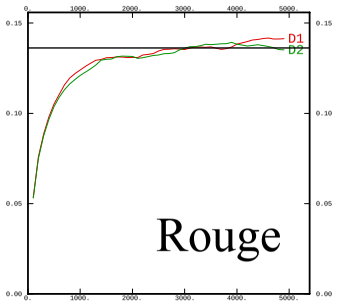
grey gris 51.5

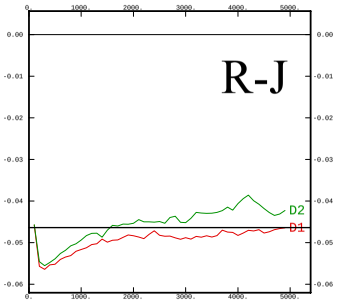
yellow jaune 30.3

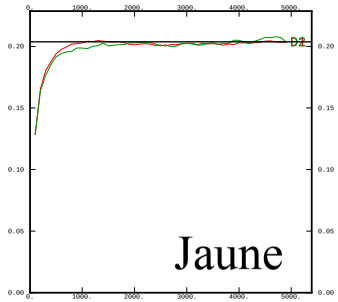
red rouge 18.2

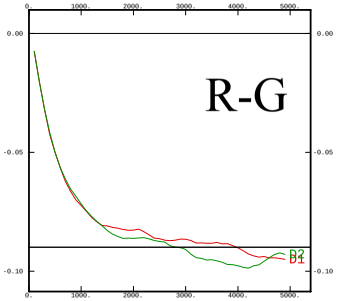
Model of diffusive type

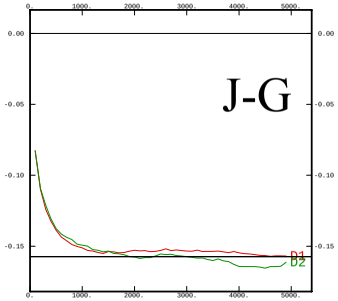
- simple and cross variograms

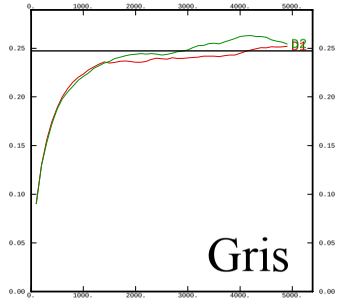






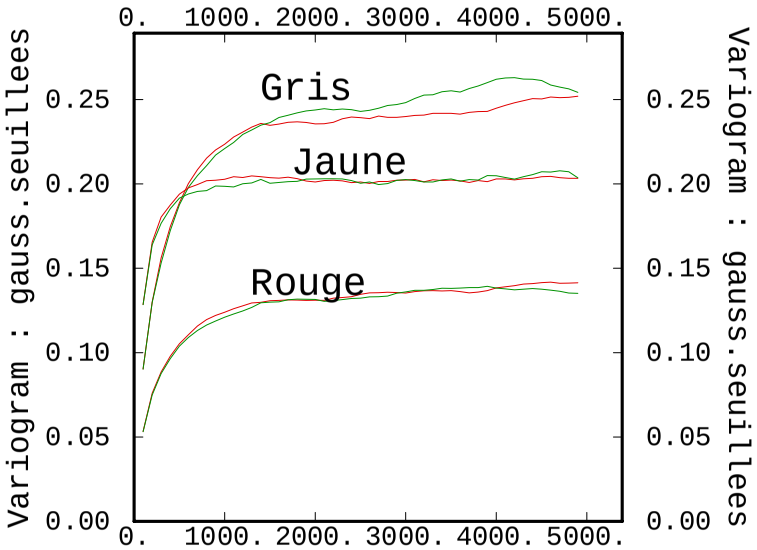






Model of diffusive type

- simple variograms



ratio cross/simple