Generalized cross-covariances and their estimation

by

H.R. Künsch,
A. Papritz and
F. Bassi

Research Report No. 77
August 1996

Seminar für Statistik

Eidgenössische Technische Hochschule (ETH)

CH-8092 Zürich

Switzerland

Generalized cross-covariances and their estimation

H.R. Künsch,
A. Papritz and
F. Bassi

Seminar für Statistik
ETH Zentrum
CH-8092 Zürich, Switzerland

August 1996

Abstract

Generalized cross-covariances describe the linear relationships between spatial variables observed at different locations. They are invariant under translation of the locations for any intrinsic processes, they determine the co-kriging predictors without additional assumptions and they are unique up to linear functions. If the model is stationary, i.e. if the variograms are bounded, they correspond to the stationary cross-covariances. Under some symmetry condition they are equal to minus the usual cross-variogram. We present a method to estimate these generalized cross-covariances from data observed at arbitrary sampling locations. In particular we do not require that all variables are observed at the same points. For fitting a linear coregionalization model we combine this new method with a standard algorithm which ensures positive definite coregionalization matrices. We study the behavior of the method both by computing variances exactly and by simulating from various models.

KEY Words: Multivariate intrinsic processes; Co-kriging; Cross-correlations; Variance of increments; Undersampling; Coregionalization.

Running head: Generalized cross-covariances and their estimation

1 Introduction

Co-kriging is a powerful method which exploits linear relations between several spatial variables for predicting values at unsampled locations or block means. Wackernagel (1995) deals in detail with its various theoretical and practical aspects. But unlike univariate kriging there are still unresolved issues in the procedure. One has to decide which function to use for expressing linear relations between variables: the cross-variogram, the pseudo cross-variogram (Myers, 1991) or the cross-covariances. Each choice imposes certain restrictions on the models one can consider and on the sampling scheme. To estimate these functions from data we have to assume that they are invariant under translation, i.e. that they depend only on the relative positions of any two locations. For any intrinsic model the cross-variogram has this property, but the cross-covariances are translation-invariant only if the processes are stationary. The pseudo cross-variogram is intermediate: it is shift-invariant for some non-stationary models, but not for general intrinsic processes (Papritz *et al.*, 1993). However the cross-variogram has other disadvantages: it cannot be estimated unless both variables are observed at many sampling locations, and it is sufficient for co-kriging only if an additional symmetry condition holds (Condition (iii) of Myers, 1982), which is equivalent to the symmetry of the cross-covariances in the stationary case. The pseudo cross-variogram or the cross-covariances do not have either of these limitations.

In this paper we investigate a different function for expressing linear relations which is called the generalized cross-covariance. It is free from any of the above restrictions. It is shift-invariant for any intrinsic model, it can be estimated from arbitrary sampling schemes and it determines the co-kriging predictor without additional assumptions. First we define the new function and discuss its properties. Then we present a new estimation method for general sampling schemes. The first part is of more theoretical interest since the definition of the new function is not very explicit except in symmetric situations where it is equal to minus the standard cross-variogram. Although in most soil science applications the symmetry assumption is quite plausible, checking it may be helpful. Our results make this possible without assuming stationarity. Moreover the new estimation method is of high practical importance since undersampling of one variable is common in many applications. It is the same idea which was presented in Papritz *et al.* (1993), namely least squares fitting of the cloud of products of differences to estimate the expected value of these products. But here we show how one can obtain a computationally feasible solution to this minimization problem. The method can be applied to any parametric model for the generalized cross-covariance. In particular, by choosing a piecewise constant function, we obtain the experimental generalized cross-covariance. But this method does not necessarily give an estimate which is conditionally positive definite. In order to guarantee positive definiteness in the coregionalization model, we combine our estimation method with the algorithm of Goulard and Voltz (1992). When one variable is undersampled, the experimental auto- and cross-variograms are estimated for different lag distances, and thus this algorithm needs to be modified.

One may argue about how much one gains by dealing with intrinsic models instead of stationary ones. But there are many examples where the variograms appear to be unbounded over the region where observations are available. We think that in such cases one should not use a stationary model. In particular estimation of both the cross-covariances and the pseudo cross-variogram requires the mean to be estimated first. This may lead to large biases when the variograms are unbounded. One might further contest the need for modelling asymmetric cross-correlation patterns between variables observed in the two- or three-dimensional space because there is little evidence for this in the literature. But this lack is partly due to the common use of the cross-variogram for describing cross-correlation, and this function, being even, cannot describe asymmetries. For all these reasons, we believe that our method represents a significant advantage.

2 Generalized cross-covariances

We consider a bivariate random process $\mathbf{Z}(\mathbf{x}) = (Z_1(\mathbf{x}), Z_2(\mathbf{x}))'$ indexed by points $\mathbf{x} \in \mathbb{R}^2$. ($'$ denotes the transpose). The case of more than two variables is a straightforward generalization. A linear combination

$$\sum_{i=1}^{n_1} \lambda_i Z_1(\mathbf{x}_{1,i}) + \sum_{k=1}^{n_2} \mu_k Z_2(\mathbf{x}_{2,k})$$

will be called an increment (of order zero) if

$$\sum_{i=1}^{n_1} \lambda_i = \sum_{k=1}^{n_2} \mu_k = 0.$$

Thus increments are unchanged if we add arbitrary constants to each variable. Note that the usual unbiasedness conditions in co-kriging imply that the errors of the ordinary co-kriging system are increments.

Definition 2.1 $(\mathbf{Z}(\mathbf{x}))$ *is intrinsic (of order zero) if*

(i) All increments have expectation zero.

(ii) The variance of any increment does not change if we add an arbitrary displacement $\mathbf{h} \in \mathbb{R}^2$ *to all* $\mathbf{x}_{1,i}$ *and* $\mathbf{x}_{2,k}$.

Because we consider here only increments and intrinsic models of order zero, we omit the order from our terminology.

We turn now to the computation of the variance of increments. Clearly

$$\begin{aligned} \mathrm{Var}[\sum_{i=1}^{n_1} \lambda_i Z_1(\mathbf{x}_{1,i}) + \sum_{k=1}^{n_2} \mu_k Z_2(\mathbf{x}_{2,k})] &= \mathrm{Var}[\sum_{i=1}^{n_1} \lambda_i Z_1(\mathbf{x}_{1,i})] \\ +\mathrm{Var}[\sum_{k=1}^{n_2} \mu_k Z_2(\mathbf{x}_{2,k})] + 2\mathrm{Cov}[\sum_{i=1}^{n_1} \lambda_i Z_1(\mathbf{x}_{1,i}), &\sum_{k=1}^{n_2} \mu_k Z_2(\mathbf{x}_{2,k})]. \end{aligned} \tag{2.1}$$

It is well known that

$$\mathrm{Var}[\sum_{i=1}^{n_1} \lambda_i Z_1(\mathbf{x}_{1,i})] = -\sum_{i=1}^{n_1}\sum_{j=1}^{n_1} \lambda_i \lambda_j \gamma_{11}(\mathbf{x}_{1,i} - \mathbf{x}_{1,j}) \tag{2.2}$$

and

$$\mathrm{Var}[\sum_{k=1}^{n_2} \mu_k Z_2(\mathbf{x}_{2,k})] = -\sum_{k=1}^{n_2}\sum_{\ell=1}^{n_2} \mu_k \mu_\ell \gamma_{22}(\mathbf{x}_{2,k} - \mathbf{x}_{2,\ell}) \tag{2.3}$$

where $\gamma_{\alpha\alpha}(\mathbf{h})$ denotes the semi-variogram of $(Z_\alpha(\mathbf{x}))$. Because $-\gamma_{\alpha\alpha}(\mathbf{h})$ acts like the usual auto-covariance in (2.2) and (2.3), it is also called a generalized auto-covariance, see Matheron (1973). Similarly a function $\nu_{12}(\mathbf{h})$ will be called a *generalized cross-covariance* if

$$\mathrm{Cov}[\sum_{i=1}^{n_1} \lambda_i Z_1(\mathbf{x}_{1,i}), \sum_{k=1}^{n_2} \mu_k Z_2(\mathbf{x}_{2,k})] = \sum_{i=1}^{n_1}\sum_{k=1}^{n_2} \lambda_i \mu_k \nu_{12}(\mathbf{x}_{1,i} - \mathbf{x}_{2,k}) \tag{2.4}$$

holds provided $\sum \lambda_i = \sum \mu_k = 0$, see Christakos (1992), Chapter 4.4. It is however far from obvious whether such a generalized cross-covariance exists for arbitrary intrinsic processes, whether it is unique and how we can find it if it exists. Some simplification occurs since for an arbitrary $\mathbf{x} \in \mathbb{R}^2$

$$\sum_{i=1}^{n_1} \lambda_i Z_1(\mathbf{x}_{1,i}) = \sum_{i=1}^{n_1} \lambda_i (Z_1(\mathbf{x}_{1,i}) - Z_1(\mathbf{x}))$$

and

$$\sum_{k=1}^{n_2} \mu_k Z_2(\mathbf{x}_{2,k}) = \sum_{k=1}^{n_2} \mu_k (Z_2(\mathbf{x}_{2,k}) - Z_2(\mathbf{x})).$$

Thus

$$\begin{aligned}&\mathrm{Cov}[\sum_{i=1}^{n_1} \lambda_i Z_1(\mathbf{x}_{1,i}), \sum_{k=1}^{n_2} \mu_k Z_2(\mathbf{x}_{2,k})] \\ &= \sum_{i=1}^{n_1}\sum_{k=1}^{n_2} \lambda_i \mu_k E[(Z_1(\mathbf{x}_{1,i}) - Z_1(\mathbf{x}))(Z_2(\mathbf{x}_{2,k}) - Z_2(\mathbf{x}))]. \end{aligned} \tag{2.5}$$

From this it follows that we have to check (2.4) only for special increments: ν_{12} is a generalized cross-covariance if

$$\begin{aligned}&E[(Z_1(\mathbf{x}+\mathbf{h}_1) - Z_1(\mathbf{x}))(Z_2(\mathbf{x}+\mathbf{h}_2) - Z_2(\mathbf{x}))] \\ &= \nu_{12}(\mathbf{h}_1 - \mathbf{h}_2) - \nu_{12}(\mathbf{h}_1) - \nu_{12}(-\mathbf{h}_2) + \nu_{12}(\mathbf{0}). \end{aligned} \tag{2.6}$$

The property (2.4) follows easily from (2.6) and (2.5) by using $\sum \lambda_i = \sum \mu_k = 0$. It would be nice if we could simplify (2.6) further and replace it by something which involves only a single displacement, but this seems not possible.

We now discuss existence and uniqueness of generalized cross-covariances. First note that if the process is stationary, then the usual cross-covariance is also a generalized cross-covariance. Next, for $\mathbf{h}_1 = \mathbf{h}_2 = \mathbf{h}$ the left hand side of (2.6) is twice the cross-variogram $\gamma_{12}(\mathbf{h})$. Hence the generalized cross-covariance determines the cross-variogram

$$\gamma_{12}(\mathbf{h}) = \nu_{12}(\mathbf{0}) - \frac{1}{2}(\nu_{12}(\mathbf{h}) + \nu_{12}(-\mathbf{h})). \tag{2.7}$$

If the symmetry condition (iii) of Myers (1982) holds, then minus the cross-variogram is a generalized cross-covariance, see Myers (1982), equation (16). To see the connection with equation (2.7), note that the symmetry condition is equivalent to $\nu_{12}(\mathbf{h}) = \nu_{12}(-\mathbf{h})$ and constants do not matter for the generalized cross-covariance, as discussed below. But if the symmetry condition does not hold, the cross-variogram does not determine the generalized cross-covariance. This is intuitively obvious from (2.7) and we will give such an example shortly. Next assume that

$$\frac{1}{2}\mathrm{Var}[Z_1(\mathbf{x}+\mathbf{h}) - Z_2(\mathbf{x})]$$

is independent of $\mathbf{x}$ for all $\mathbf{h}$. Then this expression is called the pseudo cross-variogram $\gamma_{12}^p(\mathbf{h})$ (Myers, 1991), and $-\gamma_{12}^p(\mathbf{h})$ is also a generalized cross-covariance, see Formula (5') of Myers (1991) or Formula (6) of Papritz *et al.* (1993). As discussed in Papritz *et al.* (1993) the pseudo cross-variogram is shift-invariant for some, but not for all intrinsic non-stationary processes. It is easy to give examples for which neither the cross-variogram nor the pseudo cross-variogram nor the cross-covariances is adequate. Let $(Y^{(1)}(\mathbf{x}))$ and $(Y^{(2)}(\mathbf{x}))$ be two independent univariate processes with linear semi-variogram $g(\mathbf{h}) = ||\mathbf{h}||$ and define

$$\begin{aligned} Z_1(\mathbf{x}) &= a_{11}Y^{(1)}(\mathbf{x}) + a_{12}Y^{(2)}(\mathbf{x}), \\ Z_2(\mathbf{x}) &= a_{21}Y^{(1)}(\mathbf{x}-\mathbf{u}_1) + a_{22}Y^{(2)}(\mathbf{x}-\mathbf{u}_2). \end{aligned}$$

Then it is easily checked that

$$\nu_{12}(\mathbf{h}) = -a_{11}a_{21}||\mathbf{u}_1 + \mathbf{h}|| - a_{12}a_{22}||\mathbf{u}_2 + \mathbf{h}||.$$

Thus $\nu_{12} \neq \gamma_{12}$. Also $\mathrm{Cov}[Z_1(\mathbf{x}+\mathbf{h}), Z_2(\mathbf{x})]$ and $\mathrm{Var}[Z_1(\mathbf{x}+\mathbf{h}) - Z_2(\mathbf{x})]$ both depend on $\mathbf{x}$. It is even true that $\mathrm{Var}[\lambda_1 Z_1(\mathbf{x}+\mathbf{h}) - \lambda_2 Z_2(\mathbf{x})]$ depends on $\mathbf{x}$ for any λ_1 and λ_2, so scaling of the variables does not help. Hence the following result is a genuine extension to all cases considered before.

Theorem 2.1 *If $(\mathbf{Z}(\mathbf{x}))$ is an intrinsic process with $\gamma_{11}(\mathbf{h})$ and $\gamma_{22}(\mathbf{h})$ both continuous at $\mathbf{h} = \mathbf{0}$, then there exists a generalized cross-covariance ν_{12}. It is unique up to linear functions.*

Since the proof uses more advanced mathematical machinery, we put it into the Appendix. It shows that a valid generalized cross-covariance can be represented in the form (A.6) where C_{12} and Q_{12} are measures satisfying (A.3) and (A.4), C_{12} is even and Q_{12} is odd. The first integral is equal to the cross-variogram, and Q_{12} is zero if the symmetry condition (iii) of Myers (1982) holds. In the stationary case C_{12} and Q_{12} are finite measures and we can integrate each term on the right hand side of (A.6) separately. Then we obtain the usual cross-covariance plus a linear function. It is easy to see that when a linear function $c_0 + c_1 h_1 + c_2 h_2$ with arbitrary coefficients c_0, c_1, c_2 is added to a generalized cross-covariance, the result is still a generalized cross-covariance. This corresponds to the non-uniqueness of generalized covariances for univariate intrinsic processes of higher order (cf. Matheron, 1973).

By combining (2.1)-(2.3) and (2.4) we find

$$\begin{aligned} \mathrm{Var}[\sum_{i=1}^{n_1} \lambda_i Z_1(\mathbf{x}_{i,1}) + \sum_{k=1}^{n_2} \mu_k Z_2(\mathbf{x}_{2,k})] &= -\sum_{i=1}^{n_1}\sum_{j=1}^{n_1} \lambda_i\lambda_j\gamma_{11}(\mathbf{x}_{1,i} - \mathbf{x}_{1,j}) \\ -\sum_{k=1}^{n_2}\sum_{\ell=1}^{n_2} \mu_k\mu_\ell\gamma_{22}(\mathbf{x}_{2,k} - \mathbf{x}_{2,\ell}) &+ 2\sum_{i=1}^{n_1}\sum_{k=1}^{n_2} \lambda_i\mu_k\nu_{12}(\mathbf{x}_{1,i} - \mathbf{x}_{2,k}). \end{aligned} \tag{2.8}$$

This formula has two main uses. First we can derive the equations for co-kriging when γ_{11}, γ_{22} and ν_{12} are given. This is by now well known, so we omit the details. Second we can use it for estimating ν_{12} as well as γ_{11} and γ_{22}, although other methods are available for estimating the latter. This will be discussed in the next section. It means that although the generalized cross-covariances are defined only through a rather abstract existence theorem, they can be estimated from data. So the concept is also useful for a practitioner.

3 Estimation of generalized cross-covariances

3.1 General fitting via least squares

We assume that we have a parametric model $\nu_{12}(\mathbf{h}) = \nu_{12}(\mathbf{h}; \boldsymbol{\theta})$ and want to estimate the parameter $\boldsymbol{\theta}$ (which usually contains several components) from data $(Z_1(\mathbf{x}_{1,i}), i = 1, \ldots n_1)$ and $(Z_2(\mathbf{x}_{2,i}), i = 1, \ldots n_2)$. Define

$$P_{ijk\ell} = (Z_1(\mathbf{x}_{1,i}) - Z_1(\mathbf{x}_{1,j}))(Z_2(\mathbf{x}_{2,k}) - Z_2(\mathbf{x}_{2,\ell})) \tag{3.1}$$

and

$$\begin{aligned}\pi_{ijk\ell}(\boldsymbol{\theta}) = \nu_{12}(\mathbf{x}_{1,i} - \mathbf{x}_{2,k}; \boldsymbol{\theta}) - \nu_{12}(\mathbf{x}_{1,i} - \mathbf{x}_{2,\ell}; \boldsymbol{\theta}) \\ -\nu_{12}(\mathbf{x}_{1,j} - \mathbf{x}_{2,k}; \boldsymbol{\theta}) + \nu_{12}(\mathbf{x}_{1,j} - \mathbf{x}_{2,\ell}; \boldsymbol{\theta}).\end{aligned} \tag{3.2}$$

Then by equation (2.4)

$$E[P_{ijk\ell}] = \pi_{ijk\ell}(\theta). \tag{3.3}$$

We now fit the $\pi_{ijk\ell}$'s to the cloud of the $P_{ijk\ell}$'s by least squares, i.e. we propose to estimate $\boldsymbol{\theta}$ by minimizing

$$Q_{12}(\boldsymbol{\theta}) = \sum_{i=1}^{n_1}\sum_{j=1}^{n_1}\sum_{k=1}^{n_2}\sum_{\ell=1}^{n_2}(P_{ijk\ell} - \pi_{ijk\ell}(\boldsymbol{\theta}))^2, \tag{3.4}$$

cf. Formula (24) in Papritz *et al.* (1993). There we assumed the symmetry condition which implies that $\nu_{12}(\mathbf{h}) = -\gamma_{12}(\mathbf{h})$, but this is not necessary. The drawback of (3.4) is that in this form we need approximately $\frac{1}{4}n_1^2 n_2^2$ multiplications and $\frac{1}{2}n_1^2 n_2^2$ additions to compute $Q_{12}(\boldsymbol{\theta})$ for one value of $\boldsymbol{\theta}$ even if we ignore the operations needed to compute $\pi_{ijk\ell}(\boldsymbol{\theta})$. Except for n_1 and n_2 very small, this is not feasible. However by simple algebraic manipulations we can convert the expression defining $Q_{12}(\boldsymbol{\theta})$ into a form which needs far less operations to compute. For this we write

$$\rho_{ik}(\boldsymbol{\theta}) = \nu_{12}(\mathbf{x}_{1,i} - \mathbf{x}_{2,k}; \boldsymbol{\theta})$$

so that

$$\pi_{ijk\ell}(\boldsymbol{\theta}) = \rho_{ik}(\boldsymbol{\theta}) - \rho_{i\ell}(\boldsymbol{\theta}) - \rho_{jk}(\boldsymbol{\theta}) + \rho_{j\ell}(\boldsymbol{\theta}).$$

Then because

$$\pi_{ijk\ell}(\boldsymbol{\theta}) = -\pi_{ij\ell k}(\boldsymbol{\theta}) = -\pi_{jik\ell}(\boldsymbol{\theta}) = \pi_{ji\ell k}(\boldsymbol{\theta})$$

we obtain

$$
\sum_{i=1}^{n_1}\sum_{j=1}^{n_1}\sum_{k=1}^{n_2}\sum_{\ell=1}^{n_2}\pi_{ijk\ell}(\boldsymbol{\theta})^2 = 4\sum_{i=1}^{n_1}\sum_{j=1}^{n_1}\sum_{k=1}^{n_2}\sum_{\ell=1}^{n_2}\pi_{ijk\ell}(\boldsymbol{\theta})\rho_{ik}(\boldsymbol{\theta})
$$

$$
= 4n_1n_2\sum_{i=1}^{n_1}\sum_{k=1}^{n_2}\rho_{ik}(\boldsymbol{\theta})^2 - 4n_1\sum_{i=1}^{n_1}\sum_{k=1}^{n_2}\sum_{\ell=1}^{n_2}\rho_{ik}(\boldsymbol{\theta})\rho_{i\ell}(\boldsymbol{\theta})
$$

$$
-4n_2\sum_{i=1}^{n_1}\sum_{j=1}^{n_1}\sum_{k=1}^{n_2}\rho_{ik}(\boldsymbol{\theta})\rho_{jk}(\boldsymbol{\theta}) + 4\sum_{i=1}^{n_1}\sum_{j=1}^{n_1}\sum_{k=1}^{n_2}\sum_{\ell=1}^{n_2}\rho_{ik}(\boldsymbol{\theta})\rho_{j\ell}(\boldsymbol{\theta})
$$

$$
= 4n_1n_2\sum_{i=1}^{n_1}\sum_{k=1}^{n_2}\rho_{ik}(\boldsymbol{\theta})^2 - 4n_1\sum_{i=1}^{n_1}(\sum_{k=1}^{n_2}\rho_{ik}(\boldsymbol{\theta}))^2
$$

$$
-4n_2\sum_{k=1}^{n_2}(\sum_{i=1}^{n_1}\rho_{ik}(\boldsymbol{\theta}))^2 + 4(\sum_{i=1}^{n_1}\sum_{k=1}^{n_2}\rho_{ik}(\boldsymbol{\theta}))^2.
$$

Similarly, since

$$
P_{ijk\ell} = -P_{ij\ell k} = -P_{jik\ell} = P_{ji\ell k},
$$

we have

$$
\sum_{i=1}^{n_1}\sum_{j=1}^{n_1}\sum_{k=1}^{n_2}\sum_{\ell=1}^{n_2}P_{ijk\ell}\pi_{ijk\ell}(\boldsymbol{\theta}) = 4\sum_{i=1}^{n_1}\sum_{j=1}^{n_1}\sum_{k=1}^{n_2}\sum_{\ell=1}^{n_2}P_{ijk\ell}\rho_{ik}(\boldsymbol{\theta})
$$

$$
= 4n_1n_2\sum_{i=1}^{n_1}\sum_{k=1}^{n_2}\rho_{ik}(\boldsymbol{\theta})(Z_1(\mathbf{x}_{1,i}) - \overline{Z}_1)(Z_2(\mathbf{x}_{2,k}) - \overline{Z}_2).
$$

Taking these results together we see that

$$
\frac{1}{4}Q_{12}(\boldsymbol{\theta}) = \frac{1}{4}\sum_{i=1}^{n_1}\sum_{j=1}^{n_1}(Z_1(\mathbf{x}_{1,i}) - Z_1(\mathbf{x}_{1,j}))^2\sum_{k=1}^{n_2}\sum_{\ell=1}^{n_2}(Z_2(\mathbf{x}_{2,k}) - Z_2(\mathbf{x}_{2,\ell}))^2
$$

$$
-2n_1n_2\sum_{i=1}^{n_1}\sum_{k=1}^{n_2}\rho_{ik}(\boldsymbol{\theta})(Z_1(\mathbf{x}_{1,i}) - \overline{Z}_1)(Z_2(\mathbf{x}_{2,k}) - \overline{Z}_2) + n_1n_2\sum_{i=1}^{n_1}\sum_{k=1}^{n_2}\rho_{ik}(\boldsymbol{\theta})^2
$$

$$
-n_1\sum_{i=1}^{n_1}(\sum_{k=1}^{n_2}\rho_{ik}(\boldsymbol{\theta}))^2 - n_2\sum_{k=1}^{n_2}(\sum_{i=1}^{n_1}\rho_{ik}(\boldsymbol{\theta}))^2 + (\sum_{i=1}^{n_1}\sum_{k=1}^{n_2}\rho_{ik}(\boldsymbol{\theta}))^2. \tag{3.5}
$$

All sums on the right now involve only 2 indices. This implies that the number of operations required to compute $Q_{12}(\boldsymbol{\theta})$ is only the square root of what we had before. So the computation is feasible for most data sets occurring in practice.

The simplest case occurs when $\nu_{12}(\mathbf{h};\boldsymbol{\theta})$ depends linearly on the parameters, i.e.

$$
\nu_{12}(\mathbf{h};\boldsymbol{\theta}) = \sum_{r=1}^{p}\theta_r f_r(\mathbf{h}) \tag{3.6}
$$

where $f_1(\mathbf{h}),\ldots,f_p(\mathbf{h})$ are given functions. Then $Q_{12}(\boldsymbol{\theta})$ is a quadratic function, and minimizing $Q_{12}(\boldsymbol{\theta})$ is equivalent to solving a system of linear equations. The coefficients of this linear system can be easily deduced from (3.5). We leave the details to the reader.

3.2 Experimental generalized cross-covariances

In order to estimate generalized cross-covariances without imposing some particular structure, we use the model

$$\nu_{12}(\mathbf{h};\boldsymbol{\theta}) = \sum_{r=1}^{p} \theta_r 1_{A_r}(\mathbf{h}) \tag{3.7}$$

where $A_1, A_2, \ldots, A_r$ are subsets of $\mathbb{R}^2$ and $1_A(\mathbf{h})$ is the indicator function of a set A (i.e. equal to one if $\mathbf{h} \in A$ and equal to zero otherwise). In the isotropic case the sets A_r will be concentric rings around zero. In the anisotropic case each ring will be split for instance into eight sectors.

We estimate the parameters θ_r by the least squares method described in the previous section. The normal equations are particularly simple and involve only counting pairs of observation points satisfying certain conditions. But some thought is necessary for answering the question which parameters can be estimated. First ν_{12} is determined only up to a linear function. So we can always assume that $\nu_{12}(\mathbf{0};\boldsymbol{\theta}) = 0$, and $\nu_{12}((1,0)';\boldsymbol{\theta}) = \nu_{12}((0,1)';\boldsymbol{\theta}) = 0$ in the anisotropic case. Second some parameters cannot be estimated due to the geometry of the observation points: if $\mathbf{x}_{1,i} - \mathbf{x}_{2,k} \in A_r$ for all pairs (i,k) or $\mathbf{x}_{1,i} - \mathbf{x}_{2,k} \notin A_r$ for all pairs (i,k) then θ_r cannot be estimated. In particular when two variables are never available at the same location, we cannot estimate a pure nugget component.

We consider the estimated parameter $\hat{\theta}_r$ as the experimental generalized cross-covariance for lag $\mathbf{h}$ equal to the average of pairwise differences $\mathbf{x}_{1,i} - \mathbf{x}_{2,k}$ belonging to A_r. For a more sophisticated procedure which gives an experimental generalized cross-covariance for all lags, one could replace the indicator functions in (3.7) by the basis functions in the space of splines with suitably chosen knots. But we do not pursue this here.

3.3 Estimation in the coregionalization model

In the linear coregionalization model (Journel and Huijbregts, 1978; Wackernagel 1988; Goulard and Voltz 1992), the symmetry condition is satisfied and the variograms have the form

$$\gamma_{jk}(\mathbf{h};\boldsymbol{\theta}) = \sum_{s=1}^{S} b_{jk}^s g^s(\mathbf{h}, \eta^s).$$

Here each g^s is a semivariogram of an intrinsic univariate process. The unknown parameters $\boldsymbol{\theta}$ are on one hand the coregionalization matrices (b_{jk}^s) which have to be positive definite for all s and the parameters η^s which fix typically the range and/or the shape of each g^s. In principle, we could apply our general estimation method directly, but the nonlinearity in η might be delicate, and the estimated coregionalization matrices are not guaranteed to be positive definite. We therefore prefer to build a procedure based on existing algorithms which avoid these problems. We use our new method only to compute an experimental cross-variogram. Then we estimate iteratively the parameters η^s with the coregionalization matrices kept fixed and vice versa. In the first step we use nonlinear least squares for the difference between experimental and theoretical variograms. For this step there exists quite a large experience and it generally behaves well. In the second step we use the algorithm of Goulard and Voltz (1992) which iteratively minimizes the criterion

$$WSS = \sum_r w(\mathbf{h}_r) tr[(V(\hat{\Gamma}(\mathbf{h}_r) - \Gamma(\mathbf{h}_r;\boldsymbol{\theta})))^2]$$

under the side condition that all (b^s_{jk}) must be positive semi-definite. Here $\hat{\Gamma}$ is the matrix of experimental variograms, w are some weight functions and V is a diagonal matrix of weights. Compared to its standard version there are two particular features which needed attention. First, as explained above, if the sampling locations of the two variables are disjoint and g^1 corresponds to a nugget effect then the experimental cross-variogram is determined only up to the addition of $const.g^1(\mathbf{h})$ where $const.$ is arbitrary. Thus it is not possible to estimate b^1_{12}. We tackle this problem as follows: At the beginning of the i-th iteration step we replace $\hat{\gamma}_{12}$ by $\hat{\gamma}_{12} - const.^{(i)} g^1$ where

$$const.^{(i)} = \sum_r w_r[\hat{\gamma}_{12}(\mathbf{h}_r) - \gamma_{12}(\mathbf{h}_r;\boldsymbol{\theta}^{(i-1)})] / \sum_r w_r$$

and $\boldsymbol{\theta}^{(i-1)}$ are the parameters estimated in the $(i-1)$-th step. During the i-th iteration we then minimize WSS as usual with respect to b^1_{11}, b^1_{22} and b^1_{12} subject to the condition that the first two coefficients should be non-negative. The coefficient b^1_{12} tends to zero with proceeding iteration and therefore does not interfere in estimating b^1_{11} and b^1_{22}. Second, in the undersampled case, the experimental variograms $\hat{\gamma}_{11}$, $\hat{\gamma}_{12}$ and $\hat{\gamma}_{22}$ are all available at different lags. But for the criterion WSS above we need them for the same lags. We solve this by setting the missing values of the experimental auto- and cross-variograms equal to $\gamma_{ij}(\mathbf{h};\boldsymbol{\theta}^{(i-1)})$ at the begin of the ith iteration step. A Fortran program which puts all these steps together will be integrated into GENSTAT. It is also available upon request from the second author.

4 Numerical experiments

4.1 The experimental cross-variogram

We consider here the situation where we have observed a bivariate Gaussian random process at coincident locations on the regular grid $\{1, 2, \ldots n\}^2$. In this situation three experimental cross-variograms available: the standard one which averages the products

$$(Z_1(\mathbf{x}_i) - Z_1(\mathbf{x}_j))(Z_2(\mathbf{x}_i) - Z_2(\mathbf{x}_j))/2$$

over pairs $(\mathbf{x}_i, \mathbf{x}_j)$ in a given lag class, the new one introduced in section 3.2 and the one based on the empirical cross-covariances $\hat{C}_{12}(\mathbf{h})$, which we obtain by averaging

$$(Z_1(\mathbf{x}_i) - \overline{Z}_1)(Z_2(\mathbf{x}_j) - \overline{Z}_2)$$

over pairs $(\mathbf{x}_i, \mathbf{x}_j)$ in a given lag class. Our aim is to compare these three estimators with respect to bias and standard deviation. They are all certain quadratic forms in the observations. We thus can compute their expectation and because of the Gaussian distribution also their variance in closed form. We first look at the case where all auto- and cross-variograms are linear:

$$\gamma_{ij}(\mathbf{h}) = \beta_{ij}||\mathbf{h}||.$$

We chose $\beta_{11} = \beta_{22} = 1$ and $\beta_{12} = 0.7$ although this is not essential for our conclusions (note that the bias is proportional to β_{12} and the standard deviation is proportional to $(\beta_{11}\beta_{22} + \beta_{12}^2)^{1/2}$). The results for $n = 4, 8, 12, 16$ and lag classes $(i - 0.5, i + 0.5]$ are given in Tables 1 – 3. Clearly the estimator based on experimental cross-covariances is the worst both with respect to bias and random variability. Of course this is not surprising since our model is not stationary. But it supports our claim that one should not rely on experimental cross-covariances when variograms are unbounded. A similar behavior is expected also for the experimental pseudo cross-variogram. In contrast to this our new estimator has negligible bias and shows the same random variability as the standard estimator except for the lag classes which are small compared to the observation region. For these lags the standard estimator is considerably better. We explain this behavior as follows: our new estimator has to estimate the cross-variogram simultaneously for all lag classes, and the large variability for larger lags affects also the smaller lags. An improvement can be obtained from the following idea: Instead of computing the estimator based on all observation locations, we can compute it for smaller, overlapping subsets and then average over all the results. By doing this we obtain estimates only for small lag classes, but it is for these that we want an improvement. We implemented this idea by choosing as the subsets all two by two squares of neighboring locations. Then we have only one lag class, the first one. The standard deviation of the resulting estimator is .42 for $n = 4$, .22 for $n = 8$, .15 for $n = 12$ and .12 for $n = 16$ respectively. Comparison with Table 1 shows that this estimator is as good as the standard one.

A further estimator is given by

$$\hat{\gamma}_{12}(\mathbf{h}) = \hat{\gamma}_{12}^p(\mathbf{0}) - \hat{\gamma}_{12}^p(\mathbf{h})$$

where $\hat{\gamma}_{12}^p(\mathbf{h})$ denotes the experimental pseudo cross-variogram which is obtained by averaging

$$(Z_1(\mathbf{x}_i) - \overline{Z}_1 - Z_2(\mathbf{x}_j) + \overline{Z}_2)^2/2$$

over pairs $(\mathbf{x}_i, \mathbf{x}_j)$ in the lag class corresponding to $\mathbf{h}$. A comparison of this with the previous estimators is more diffcult since the conclusions depend on the choice of β_{ij}. For instance the bias is not proportional to β_{12}, but proportional to $\beta_{11} + \beta_{22} - 2\beta_{12}$ and thus is large if the two processes are independent or negatively dependent. In the special situation $\beta_{11} = \beta_{22} = 1$ and $\beta_{12} = 0.7$ which is favorable for this estimator the biases are approximately half as big as for the estimator based on empirical covariances and the standard deviations are similar to the ones for the standard experimental cross-variogram.

We did the same calculations also for exponential variograms

$$\gamma_{ij}(\mathbf{h}) = \beta_{ij}(1 - \exp(-||\mathbf{h}||/2)).$$

This model is stationary, and we thus expect that all three experimental cross-variograms are comparable. The results show that this is indeed the case: both the biases and the standard deviations of all three estimators are very close. We thus omit the corresponding tables.

4.2 Simulation results for some coregionalization models

We consider here all possible 9 combinations of 3 models and 3 sampling configurations. In the first model

$$\gamma_{jk}(\mathbf{h}) = \beta_{jk}(1 - \exp(-||\mathbf{h}||/\eta)) \tag{4.1}$$

with $\beta_{11} = \beta_{22} = 1, \beta_{12} = 0.7, \eta = 2$, in the second model

$$\gamma_{jk}(\mathbf{h}) = \beta^1_{jk} 1_{[||\mathbf{h}||>0]} + \beta^2_{jk}(1 - \exp(-||\mathbf{h}||/\eta)) \tag{4.2}$$

with $\beta^1_{11} = \beta^1_{22} = 0.25,\ \beta^1_{12} = 0.175,\ \beta^2_{11} = \beta^2_{22} = 0.75,\ \beta^2_{12} = 0.525$ and $\eta = 2$. Finally in the third model

$$\gamma_{jk}(\mathbf{h}) = \beta_{jk}||\mathbf{h}||^\eta \tag{4.3}$$

with $\beta_{11} = \beta_{22} = 1,\ \beta_{12} = 0.7,\ \eta = 1$. The unknown parameters are all the β's and η.

In the first configuration we have $n_1 = n_2 = 256$ with

$$x_{1,i} = x_{2,i} \ \in \{1, 2, \ldots, 16\}^2, \tag{4.4}$$

in the second configuration again $n_1 = n_2 = 256$, but

$$x_{1,i} \ \in \{1, 2, \ldots, 16\}^2,\ x_{2,i} \ \in \{1.001, 2.001, \ldots, 16.001\}^2, \tag{4.5}$$

and in the third configuration $n_1 = 400, n_2 = 100$

$$x_{1,i} \ \in \{1, 2, \ldots, 20\}^2,\ x_{2,i} \ \in \{1, 3, \ldots, 19\}^2, \tag{4.6}$$

These choices cover a wide range of different situations.

For each model and configuration we generated 200 replicates of $(Z_1(\mathbf{x}_{1,i}))$ and $(Z_2(\mathbf{x}_{2,i}))$. The standard estimator was used to compute all the experimental auto-variograms, and it also provided estimates of the experimental cross-variograms for configurations (4.4) and (4.6). For all configurations the experimental cross-variograms were computed by the new estimator using the procedure outlined in section 3.2. We used the lag classes $(0, 0.5], (0.5, 1.5], \ldots$ for all variograms. Then we fitted the linear model of coregionalization to the experimental variograms applying Goulard and Voltz's algorithm with the extensions described in section 3.3. Experimental semi-variances for lag distances ≤ 8 were used in fitting, and the all lag classes had equal weight.

The biases and standard deviations of the estimated parameters were computed from the 200 replicates. They are listed in Tables 4 – 6. Two types of comparisons are of interest here: comparing the two estimators for the same configurations and comparing the different configurations for the same estimators. For configuration (4.4) and models (4.1) and (4.2) the two estimators are practically the same. For configuration (4.6) and models (4.1) and (4.2) the new estimator is clearly better. This is true not only for the parameters of γ_{21} but also for those of γ_{22}.Surprisingly the new estimator of the range η is better only in model (4.1), but not in model (4.2). For (4.2) a smaller bias of the new estimator $\hat{\eta}$ is compensated by a larger standard deviation. With the model (4.3) the standard estimator is better for both configurations (4.4) and (4.6). This was to be expected given the results of section 4.1. The larger standard deviation of the new estimator of the experimental cross-variogram at small lags carries over to the estimator of the parameters.

Comparing the different configurations, we see that (4.6) is the best for model (4.1). The only exception here is the bias of the estimated range. For model (4.2) configuration (4.4) is the worst, but the comparison between (4.5) and (4.6) is less clear. Using (4.5) we obtain a much greater precision for the estimates of η, β_{22}^1 and β_{22}^2, but we cannot estimate β_{12}^1 at all and we loose some precision for β_{11}^2. Finally for model (4.3), configuration (4.4) is again worst, whereas (4.5) and (4.6) are about the same. The precisions for η and β_{11} are practically the same for both configurations, and (4.5) is better for β_{12}, but worse for β_{22}. As a conclusion, taking an equal number of observations for the two variables at coincident points seems to be a bad strategy. Surprisingly undersampling of one variable seems to allow reasonable estimates even for the auto-variogram of the undersampled variable. Of course for choosing between different configurations, sampling costs have also to be taken into account.

5 Conclusions

The generalized cross-covariances solve some of the difficulties associated with cross-variograms, pseudo cross-variograms and cross-covariances. The last two concepts restrict the class of models that we can consider. Our numerical experiments have clearly shown that experimental cross-covariances should be avoided when variograms are unbounded over the available lags. Unlike the cross-variogram the generalized cross-covariances can be used for modelling correlation between intrinsic random processes that depend on direction. The new method for estimating generalized cross-covariances allows us to check this, and it solves the problem of computing the experimental cross-variograms when some variables are undersampled with respect to others.

We have implemented the new estimator, and for the linear model of coregionalization where the generalized cross-covariances are equal to minus the cross-variogram, we have combined it with the algorithm of Goulard and Voltz (1992) to ensure a valid model. Tests on simulated data showed that the new estimator of the cross-variogram generally behaves well. If, however, the variograms are unbounded then it is less efficient at short lags than the standard method. But the latter requires coincident locations, and we have indicated how the former can be improved in such a situation.

In summary, we believe that our results solve some problems related so far with co-kriging and enhance this powerful tool.

A The existence of generalized cross-covariances

We shall give two proofs. The first one is short, but heavily relies on abstract analysis. According to Gelfand and Vilenkin (1964), Chapter III 5.5, there exists a complex matrix valued measure $(F_{ij})_{1\le i,j\le 2}$ on $\mathbb{R}^2$ with the properties

$$(dF_{ij}(\boldsymbol{\omega})) \text{ is positive definite,} \tag{A.1}$$

$$dF_{ij}(-\boldsymbol{\omega}) = \overline{dF_{ij}(\boldsymbol{\omega})}, \tag{A.2}$$

$$\int_{0<||\boldsymbol{\omega}||\le 1} ||\boldsymbol{\omega}||^2 dF_{ij}(\boldsymbol{\omega}) \text{ exists,} \tag{A.3}$$

$$\int_{||\boldsymbol{\omega}||>1} dF_{ij}(\boldsymbol{\omega}) \text{ exists,} \tag{A.4}$$

$$E[\sum_k \lambda_k Z_i(\mathbf{x}_k) \sum_\ell \lambda'_\ell Z_j(\mathbf{x}'_\ell)] = \int_{||\boldsymbol{\omega}||>0} \sum_k \lambda_k \exp(\sqrt{-1}\boldsymbol{\omega}\cdot\mathbf{x}_k) \sum_\ell \lambda'_\ell \exp(-\sqrt{-1}\boldsymbol{\omega}\cdot\mathbf{x}'_\ell) dF_{ij}(\boldsymbol{\omega}) \tag{A.5}$$

for any two increments $\sum_k \lambda_k Z_i(\mathbf{x}_k)$, $\sum_\ell \lambda'_\ell Z_j(\mathbf{x}'_\ell)$ where the dot denotes the inner product in $\mathbb{R}^2$.

Let us briefly comment on the differences between this and the statements in Chapter III 5.5 of Gelfand and Vilenkin (1964). Condition (A.4) is satisfied because we consider only usual random fields $(\mathbf{Z}(\mathbf{x}))$ and not generalized fields $(\mathbf{Z}(\phi))$. The relation (A.2) holds because our random fields are real and not complex. Conditions (A.1) and (A.3) are the same as in Gelfand and Vilenkin (1964). Finally (A.5) is equation (17) on p. 360 in Gelfand and Vilenkin (1964) for $\phi(\mathbf{x}) = \sum \lambda_k \delta_{\mathbf{x}_k}(\mathbf{x})$ and $\psi(\mathbf{x}) = \sum \lambda'_\ell \delta_{\mathbf{x}'_\ell}(\mathbf{x})$. The term $\sum a^{ij}_{pq} \alpha_p \overline{\beta_q}$ is missing in (A.5) since we assumed that all increments have expectation equal to zero.

Using (A.5) it is not difficult to obtain an explicit expression for the generalized cross-covariance. Write $dF_{12}(\boldsymbol{\omega}) = dC_{12}(\boldsymbol{\omega}) + \sqrt{-1}dQ_{12}(\boldsymbol{\omega})$ (decomposition into real and imaginary part). Then by (A.2) dC_{12} is even and dQ_{12} is odd. Thus

$$E[\sum \lambda_k Z_1(\mathbf{x}_k) \sum \lambda'_\ell Z_2(\mathbf{x}'_\ell)] =$$
$$\int_{||\boldsymbol{\omega}||>0} \sum_{k,\ell} \lambda_k \lambda'_\ell \cos(\boldsymbol{\omega}\cdot(\mathbf{x}_k - \mathbf{x}'_\ell)) dC_{12}(\boldsymbol{\omega}) - \int_{||\boldsymbol{\omega}||>0} \sum_{k,\ell} \lambda_k \lambda'_\ell \sin(\boldsymbol{\omega}\cdot(\mathbf{x}_k - \mathbf{x}'_\ell)) dQ_{12}(\boldsymbol{\omega}).$$

Note that both integrals exist because of $\sum_k \lambda_k = \sum_\ell \lambda'_\ell = 0$ and (A.3). But we cannot directly exchange the integral and the sum because the integrals would then no longer exist. However

$$\sum_{k,\ell} \lambda_k \lambda'_\ell \cos(\boldsymbol{\omega}\cdot(\mathbf{x}_k - \mathbf{x}'_\ell)) = \sum_{k,\ell} \lambda_k \lambda'_\ell [\cos(\boldsymbol{\omega}\cdot(\mathbf{x}_k - \mathbf{x}'_\ell)) - 1]$$

because $\sum \lambda_k = \sum \lambda'_\ell = 0$, and $\cos(\boldsymbol{\omega}\cdot\mathbf{h}) - 1$ is integrable by (A.3) and (A.4). Similarly

$$\sum_{k,\ell} \lambda_k \lambda'_\ell \sin(\boldsymbol{\omega}\cdot(\mathbf{x}_k - \mathbf{x}'_\ell)) = \sum_{k,\ell} \lambda_k \lambda'_\ell [\sin(\boldsymbol{\omega}\cdot(\mathbf{x}_k - \mathbf{x}'_\ell)) - \boldsymbol{\omega}\cdot(\mathbf{x}_k - \mathbf{x}'_\ell) 1_{[||\boldsymbol{\omega}||\le 1]}],$$

and $\sin(\boldsymbol{\omega}\cdot\mathbf{h})-\boldsymbol{\omega}\cdot\mathbf{h}1_{[||\boldsymbol{\omega}||<1]}$ is integrable because it is bounded and behaves like $-(\boldsymbol{\omega}\cdot\mathbf{h})^3/6$ for $||\boldsymbol{\omega}|| \to 0$. Hence a generalized cross-covariance is given by

$$\nu_{12}(\mathbf{h}) = \int_{||\boldsymbol{\omega}||>0} [\cos(\boldsymbol{\omega}\cdot\mathbf{h}) - 1] dC_{12}(\boldsymbol{\omega}) - \int_{||\boldsymbol{\omega}||>0} [\sin(\boldsymbol{\omega}\cdot\mathbf{h}) - \boldsymbol{\omega}\cdot\mathbf{h}1_{[||\boldsymbol{\omega}||\leq 1]}] dQ_{12}(\boldsymbol{\omega}), \tag{A.6}$$

and the first proof is completed.

The second proof is by direct construction. However checking the necessary details is long and delicate, so we only sketch the arguments. The idea is to use (2.6) for sufficiently many $\mathbf{h}_1$'s and $\mathbf{h}_2$'s in order to obtain ν_{12}.

We first decompose ν_{12} into an even and an odd part:

$$\begin{aligned} \nu_{12}^{+}(\mathbf{h}) &= (\nu_{12}(\mathbf{h}) + \nu_{12}(-\mathbf{h}))/2, \\ \nu_{12}^{-}(\mathbf{h}) &= (\nu_{12}(\mathbf{h}) - \nu_{12}(-\mathbf{h}))/2, \end{aligned}$$

so that $\nu_{12}(\mathbf{h}) = \nu_{12}^{+}(\mathbf{h}) + \nu_{12}^{-}(\mathbf{h})$. Then using (2.6) for $\mathbf{h}_1 = \mathbf{h}_2 = \mathbf{h}$ and putting $\nu_{12}(\mathbf{0}) = 0$ without loss of generality we find

$$\nu_{12}^{+}(\mathbf{h}) = \frac{1}{2} E[(Z_1(\mathbf{x}+\mathbf{h}) - Z_1(\mathbf{x}))(Z_2(\mathbf{x}+\mathbf{h}) - Z_2(\mathbf{x}))].$$

So ν_{12}^{+} is simply the cross-variogram. Next choosing $\mathbf{h}_1 = \mathbf{h}, \mathbf{h}_2 = \mathbf{h}/2$ in (2.6) we obtain

$$\nu_{12}^{-}(\mathbf{h}/2) = \frac{1}{2}\nu_{12}^{-}(\mathbf{h}) + \frac{1}{2}\nu_{12}^{+}(\mathbf{h}) + \frac{1}{2}E[(Z_1(\mathbf{x}+\mathbf{h}) - Z_1(\mathbf{x}))(Z_2(\mathbf{x}+\mathbf{h}/2) - Z_2(\mathbf{x}))]. \tag{A.7}$$

Without loss of generality we may take $\nu_{12}^{-}(\mathbf{e}_1) = \nu_{12}^{-}(\mathbf{e}_2) = 0$ where $\mathbf{e}_1 = (1,0)$ and $\mathbf{e}_2 = (0,1)$ (ν_{12} is determined only up to a linear function). So (A.7) gives $\nu_{12}^{-}(2^{-n_i}\mathbf{e}_i)$ for $n = 1, 2, \ldots$. But by (2.6) we obtain $\nu_{12}(\mathbf{h}_1 + \mathbf{h}_2)$ from $\nu_{12}(\mathbf{h}_1)$ and $\nu_{12}(-\mathbf{h}_2)$. So by repeated application of (2.6) we find $\nu_{12}(k_1 2^{-n_1}, k_2 2^{-n_2})$, i.e. we have constructed ν_{12} for all lags with dyadic coordinates. Note that we can arrive at the same lag $(k_1 2^{-n_1}, k_2 2^{-n_2})$ by many different sequences of intermediate steps. So for a rigorous proof we would have to show that the result does not depend on this. Finally in order to go from dyadic to general lags we have to show that ν_{12} restricted to dyadic lags is continuous. Essentially this can be done by using Schwarz' inequality and the continuity of γ_{11} and γ_{22} repeatedly. We omit the details.

References:

Christakos, G., 1992, *Random Field Models in Earth Sciences:* Academic Press, San Diego.

Gelfand, I. M., and Vilenkin, N. Ya., 1964, *Generalized Functions, Vol. 4, Applications of Harmonic Analysis:* Academic Press, New York.

Goulard, M., and Voltz, M., 1992, Linear Coregionalization Model: Tools for Estimation and Choice of Cross-Variogram Matrix: Math. Geol., v. 24, p. 269–286.

Journel, A. G., and Huijbregts, C. J., 1978, *Mining Geostatistics:* Academic Press, London.

Matheron, G., 1973, The intrinsic random functions and their applications: Adv. Appl. Probab. v. 5, 439–468.

Myers, D. E., 1982, Matrix Formulation of Co-Kriging: Math. Geol., v. 14, p. 249–257.

Myers, D. E., 1991, Pseudo-Cross Variograms, Positive-Definiteness, and Cokriging: Math. Geol., v. 23, p. 805–816.

Papritz A., Künsch H. R., and Webster R., 1993, On the Pseudo Cross-Variogram: Math. Geol., v. 25, p. 1015–1026.

Wackernagel, H., 1988, Geostatistical Techniques for Interpreting Multivariate Spatial Information. In C. F. Chung, A. G. Fabbri, and R. Sinding-Larsen (Eds.), *Quantitative Analysis of Mineral and Energy Resources:* D. Reidel Publishing Company, Dordrecht, p. 393–409.

Wackernagel, H., 1995, *Multivariate Geostatistics: an Introduction with Applications:* Springer Verlag, Berlin.

Table 1: Standard deviation for the standard experimental cross-variogram. This estimator is unbiased for all lag classes.

	lag class										
	1	2	3	4	5	6	7	8	9	10	11
n=4	.43	1.12	1.95	2.81	-	-	-	-	-	-	-
n=8	.23	.62	1.14	1.91	2.85	3.73	4.64	5.62	6.82	8.54	-
n=12	.16	.44	.81	1.38	2.09	2.83	3.67	4.52	5.49	6.47	7.39
n=16	.13	.34	.64	1.09	1.65	2.25	2.94	3.68	4.56	5.50	6.38
	lag class										
	12	13	14	15	16	17	18	19	20	21	
n=12	8.40	9.61	10.85	12.27	13.43	-	-	-	-		
n=16	7.28	8.21	9.17	10.04	11.18	12.32	13.86	14.89	16.18	17.33	

Table 2: Bias (upper figure) and standard deviation (lower figure) for the new experimental cross-variogram.

	lag class										
	1	2	3	4	5	6	7	8	9	10	11
n=4	.001	.000	-.002	.000	-	-	-	-	-	-	-
	.42	1.11	1.98	2.85	-	-	-	-	-	-	-
n=8	-.002	-.005	-.006	-.005	.003	.007	.010	.014	.021	.026	-
	.27	.60	1.09	1.86	2.83	3.77	4.80	5.83	7.07	8.63	-
n=12	-.002	-.004	-.005	-.006	-.004	-.003	-.002	.001	.006	.010	.012
	.25	.47	.78	1.30	2.01	2.76	3.62	4.51	5.55	6.66	7.69
n=16	-.001	-.002	-.004	-.005	-.005	-.005	-.004	-.003	-.002	.001	.002
	.25	.43	.66	1.04	1.56	2.13	2.82	3.57	4.48	5.46	6.39
	lag class										
	12	13	14	15	16	17	18	19	20	21	
n=12	.015	.021	.024	.027	.028	-	-	-	-	-	
	8.79	10.05	11.37	12.72	13.77	-	-	-	-	-	
n=16	.005	.007	.010	.011	.014	.016	.022	.023	.025	.025	
	7.37	8.40	9.49	10.52	11.70	12.99	14.45	15.65	16.90	18.06	

Table 3: Bias (upper figure) and standard deviation (lower figure) for the experimental cross-variogram based on experimental cross-covariances.

	lag class										
	1	2	3	4	5	6	7	8	9	10	11
n=4	-.14	-.06	.17	.34	-	-	-	-	-	-	-
	.56	1.18	1.75	2.37	-	-	-	-	-	-	-
n=8	-.21	-.28	-.27	-.20	.04	.28	.65	.89	1.31	1.87	-
	.42	.87	1.39	2.10	2.78	3.41	3.95	4.66	5.34	6.22	-
n=12	-.23	-.35	-.41	-.46	-.37	-.29	-.10	.11	.40	.80	1.10
	.38	.76	1.18	1.80	2.44	3.11	3.74	4.37	5.04	5.61	6.25
n=16	-.24	-.38	-.48	-.59	-.58	-.59	-.49	-.39	-.24	.03	.23
	.36	.70	1.08	1.62	2.18	2.79	3.40	4.04	4.77	5.43	6.10
	lag class										
	12	13	14	15	16	17	18	19	20	21	
n=12	1.38	1.79	2.30	2.81	3.22	-	-	-	-	-	
	6.95	7.67	8.32	9.04	9.62	-	-	-	-	-	
n=16	.56	.85	1.24	1.63	1.84	2.30	2.75	3.28	3.75	4.24	
	6.64	7.29	7.87	8.38	9.26	9.94	10.80	11.31	11.96	12.53	

Table 4: Biases (upper figure) and standard deviations (lower figure) of the parameters of the linear coregionalization model fitted to experimental variograms for configuration (4.4). The experimental cross-variograms were computed by the standard (OLD) and new (NEW) estimators.

	model (4.1)		model (4.2)		model (4.3)	
	OLD	NEW	OLD	NEW	OLD	NEW
η	.065	.064	.494	.506	-.075	-.076
	.784	.781	2.008	2.077	.262	.285
β^1_{11}	.034	.033	-.013	-.012	.086	.095
	.295	.293	.151	.151	.303	.334
β^1_{22}	.032	.032	-.006	-.005	.136	.143
	.283	.286	.145	.143	.315	.342
β^1_{12}	.029	.029	-.012	-.014	.058	.084
	.237	.218	.111	.114	.172	.220
β^2_{11}	-	-	.059	.061	-	-
	-	-	.319	.333	-	-
β^2_{22}	-	-	.058	.059	-	-
	-	-	.312	.320	-	-
β^2_{12}	-	-	.057	.062	-	-
	-	-	.291	.298	-	-

Table 5: Biases (upper figure) and standard deviations (lower figure) of the parameters of the linear coregionalization model fitted to experimental variograms for configuration (4.5). For this configuration the experimental cross-variogram can be computed only by the new estimator.

	model (4.1)	model (4.2)	model (4.3)
η	.077	.250	-.098
	.736	1.376	.244
β_{11}^1	.024	-.012	.080
	.266	.123	.299
β_{22}^1	.002	-.014	.124
	.248	.121	.333
β_{12}^1	.017	-	.030
	.216	-	.220
β_{11}^2	-	.072	-
	-	.303	-
β_{22}^2	-	.058	-
	-	.282	-
β_{12}^2	-	.050	-
	-	.241	-

Table 6: Biases (upper figure) and standard deviations (lower figure) of the parameters of the linear coregionalization model fitted to experimental variograms for configuration (4.6). The experimental cross-variograms were computed by the standard (OLD) and new (NEW) estimators.

	model (4.1)		model (4.2)		model (4.3)	
	OLD	NEW	OLD	NEW	OLD	NEW
η	.240	.176	.619	.481	-.017	-.035
	.775	.673	1.683	1.745	.248	.256
β_{11}^1	.054	.044	.024	.004	.046	.080
	.220	.213	.131	.134	.297	.302
β_{22}^1	.034	.023	.018	.013	.017	.056
	.247	.230	.219	.179	.271	.308
β_{12}^1	.037	.026	-.016	-.002	.011	.056
	.201	.183	.172	.132	.151	.242
β_{11}^2	-	-	.032	.046	-	-
	-	-	.273	.271	-	-
β_{22}^2	-	-	.056	.056	-	-
	-	-	.442	.406	-	-
β_{12}^2	-	-	.076	.054	-	-
	-	-	.373	.263	-	-