Spatial decorrelation methods: beyond MAF and PCA

Ute Mueller

Abstract In the geostatistical treatment of multivariate data sets the joint modelling of their spatial continuity is usually required. While it is possible to automate the inference of a suitable variogram model a transformation of the set of attributes into spatially uncorrelated factors that can be simulated independently, might be desirable. Standard methods used in geostatistics for this purpose are principal component analysis (PCA) and the method of minimum/maximum autocorrelation factors (MAF). Both methods have restrictions in their applicability and a more flexible approach may be more suitable such as that offered by approximate joint diagonalisation (AJD) methods common in Blind Source Separation. The application of two AJD methods to a family of experimental semivariogram matrices is explored here and the performance is assessed on a number of simulated data sets with different spatial characteristics. A comparison with MAF and PCA shows that the use of AJD algorithm results in better decorrelation than that achieved by MAF or PCA.

1 Introduction

Geostatistical data sets are usually multivariate and joint simulation or estimation requires the joint modelling of the spatial continuity. While the inference of a suitable variogram model can be automated to some extent, it may nevertheless be preferable to transform the set of attributes into spatially uncorrelated factors that can be simulated independently. The earliest decorrelation method [1] is principal component analysis (PCA), where the original data are rotated to orthogonal factors. These factors are pairwise orthogonal at the sample locations, but for separation distances other than 0 correlation between distinct factors will continue to exist, unless the coregionalisation is intrinsic. An early decorrelation method which takes account of spatial autocorrelation is the method of minimum/maximum autocorrelation factors

Ute Mueller
Edith Cowan University, 270 Joondalup Drive, Joondalup WA 6027 e-mail: u.mueller@ecu.edu.au

Ninth International Geostatistics Congress, Oslo, Norway, June 11. – 15., 2012

(MAF). MAF is based on the joint diagonalisation of a pair of not necessarily commuting symmetric matrices [2] and the determination of a non-singular matrix that jointly diagonalises the given matrices can be seen to correspond to solving a generalised eigenvalue problem [3]. In contrast to the transformation matrix obtained in the case of PCA, the diagonalising matrix obtained for MAF is not orthogonal.

The application of MAF in a geostatistical setting was first reported in [8] who assumed that the co-regionalisation could be described by a two structure linear model of co- regionalisation. It can be shown that MAF decorrelates the theoretical model exactly, but the actual data are only approximately decorrelated. MAF is thus a special case of a non-orthogonal approximate diagonaliser of a set of symmetric matrices, here the set of semivariogram or covariance matrices for a specified set of lags. A more general approach for approximate joint diagonalisation (AJD) has been developed in the context of blind source separation. For these AJD algorithms no assumptions are made beyond symmetry of the individual matrices and so they can be applied to a family of semivariogram or covariance matrices. In their application there are no restrictions on the number of matrices to be diagonalised and no assumption is made about the underlying covariance structure of the multivariate random function. The use of AJD methods to facilitate the simulation of multivariate data has been elucidated in [4, 5, 6, 7]. From these case studies it appears that the resultant realisations are close to those achieved via a full co-simulation.

In this paper we give an overview over the different AJD methods and discuss their performance on a number of simulated data sets with different spatial characteristics and compare the performance with MAF or PCA. The remaining sections of the paper are organised as follows: First a short overview over the decorrelation methods to be used and the assessment criteria for their performance are given. In Section 3 the covariance models are described on which the methods are to be tested along with the sample data sets. In the fourth section the results are presented and in Section 5 a brief conclusion is given.

2 Approximate decorrelation

The general problem to be addressed is the following: Given a family of semivariogram matrices M_i, $i = 1, \ldots, n$ find a matrix A such that for all i the matrix $M_i = A\Lambda_i A^T$ where for each i the matrix Λ_i is diagonal. If such a matrix A exists, then the family $\{M_i : i = 1, \ldots, n\}$ is said to be *jointly diagonalisable* and when the matrix A is orthogonal, its existence is equivalent to the pairwise commutativity of the matrices in the family.

In our case we have $M_i = \Gamma_{\mathbf{Z}}(\mathbf{h}_i)$, where $\mathbf{Z}$ denotes a vector random function and $\Gamma_{\mathbf{Z}}(\mathbf{h}_i)$ denotes the experimental semivariogram matrix calculated at the i^{th} lag vector $\mathbf{h}_i$.

2.1 PCA

The solution proposed by the PCA method consists of determining the eigenvalue decomposition of the variance-covariance matrix M of the data, $M = V\Lambda V^T$ and transforming the data according to $\mathbf{F}(\mathbf{u}) = \mathbf{Z}(\mathbf{u})V\Lambda^{-1/2}$. The resultant experimental semivariogram matrices of the factors are related to experimental semivariogram matrices of the raw data by

$$\Gamma_F(\mathbf{h}_i) = \Lambda^{-1/2}V^T\Gamma_{\mathbf{Z}}(\mathbf{h}_i)V\Lambda^{-1/2}. \tag{1}$$

2.2 MAF

For MAF, the variance-covariance matrix M and a semivariogram matrix $\Gamma_{\mathbf{Z}}(\mathbf{h}_0)$ at a chosen lag $\mathbf{h}_0$ are diagonalised jointly by congruence, in fact we have $AMA^T = I$ and $A\Gamma_{\mathbf{Z}}(\mathbf{h}_0)A^T = \Lambda_1$. The matrix A is non-singular and is given by $A = V_1^T\Lambda^{-1/2}V$ where V and Λ are the orthogonal matrix and diagonal matrix derived from the eigenvalue decomposition of M and V_1 is the orthogonal matrix which diagonalises $\Lambda^{-1/2}V^T\Gamma_{\mathbf{Z}}(\mathbf{h}_0)V\Lambda^{-1/2}$. The resultant experimental semivariogram matrices of the MAF factors are related to experimental semivariogram matrices of the raw data by

$$\Gamma_F(\mathbf{h}_i) = V_1^T\Lambda^{-1/2}V^T\Gamma_{\mathbf{Z}}(\mathbf{h}_i)V\Lambda^{-1/2}V_1. \tag{2}$$

The matrix derived via the MAF method transforms the variance-covariance matrix to the identity matrix and the semivariogram matrix for lag $\mathbf{h}_0$ to a diagonal matrix, and if MAF is used on a semivariogram model comprised of no more than two nested structures (2SLMC), decorrelation is exact. In all other cases, we have an approximate spatial decorrelation, and this is even the case when the decorrelation is carried out based on experimental semivariogram matrices which come from a co-regionalisation where the assumption of a 2SLMC is justified.

2.3 Joint approximate diagonalisation

Decorrelation via a method such as PCA corresponds to orthogonal joint diagonalisation (OJD), while decorrelation via MAF is a non-orthogonal joint diagonalisation (NOJD) method. The OJD problem can be solved making use of a single matrix while for the NOJD problem two matrices are required ([9]). In general, exact diagonalisation cannot be achieved for more than a pair of matrices, unless, special conditions are satisfied: in the OJD problem the pairwise commutativity of the matrices and in the NOJD problem a case such as the one described in [10] where the underlying linear model of co-regionalisation is given by two nested structures one of which is an intrinsic LMC.

As the derivation of the transformation matrix A is based on experimental semivariogram matrices, the presence of noise must be accounted for. To do so, the transformation matrix is derived from a family of semivariogram matrices matrices. A similar approach is taken in Blind Source Separation (see for example [11]) for which many of the AJD algorithms were developed. Fixed point iteration is used to determine the matrix A that best jointly diagonalises the given family of symmetric matrices according to some cost criterion. In the case of OJD the cost function is set to be

$$C_1(A) = \sum_{i=1}^{n} \|A^T M_i A - \mathrm{diag}(A^T M_i A)\|_F^2 \tag{3}$$

where the subscript F denotes the Frobenius norm of the matrix. The criterion used for NOJD is not very different from it, one version [12] is to put

$$C_2(A, \Lambda_1, \Lambda_2, \ldots \Lambda_n) = \sum_{i=1}^{n} \|M_i - {A^{-1}}^T \Lambda_i A^{-1})\|_F^2 \tag{4}$$

and a more general form is to use, as in the Uniformly Weighted Exhaustive Diagonalisation with Gauss iterations (UWEDGE) method ([13])

$$C_3(A, V) = \sum_{i=1}^{n} \|V^T M_i V - A\Lambda_{i,V} A^T\|_F^2. \tag{5}$$

Here $\Lambda_{i,V} = \mathrm{diag}(V^T M_i V)$ and the matrices A and V are called the *mixing* matrix and *demixing* matrix respectively. There are numerous algorithms for blind source separation [11]. Here we will consider RJD [14], a OJD algorithm and UWEDGE [13], a more recent NOJD algorithm.

2.4 Performance measures

The spatial decorrelation of the factors will be assessed via the quantitative measures introduced in [15]. These are

- the absolute deviation from diagonality $\zeta(\mathbf{h})$, defined as the sum of squares of the off-diagonal elements of the factor experimental semivariogam matrix at lag $\mathbf{h}$:

$$\zeta(\mathbf{h}) = \sum_{k=1}^{n} \sum_{j=1, j\neq k}^{n} \gamma_{\mathbf{F}}(\mathbf{h}, k, j))^2 \tag{6}$$

- the relative deviation from diagonality $\tau(\mathbf{h})$ which compares the absolute sum of off-diagonal elements with the sum of the absolute values of the diagonal elements of the factor experimental semivariogram matrix $\Gamma_{\mathbf{F}}(\mathbf{h}))$ for each lag $\mathbf{h}$:

$$\tau(\mathbf{h}) = \frac{\sum_{k=1}^{n} \sum_{j=1, j\neq k}^{n} |\gamma_{\mathbf{F}}(\mathbf{h}, k, j))|}{\sum_{k=1}^{n} |\gamma_{\mathbf{F}}(\mathbf{h}, k, k))|} \tag{7}$$

- the spatial diagonalisation efficiency $\kappa(\mathbf{h})$ which compares the sum of squares of off-diagonal elements in $\Gamma_{\mathbf{F}}(\mathbf{h}))$ with those of the attribute semivariogram matrix $\Gamma_{\mathbf{Z}}(\mathbf{h}))$

$$\tau(\mathbf{h}) = 1 - \frac{\sum_{k=1}^{n}\sum_{j=1,j\neq k}^{n}(\gamma_{\mathbf{F}}(\mathbf{h},k,j)))^2}{\sum_{k=1}^{n}\sum_{j=1,j\neq k}^{n}(\gamma_{\mathbf{Z}}(\mathbf{h},k,j)))^2} \tag{8}$$

Here $\gamma_{\mathbf{F}}(.,k,j)$ denotes the experimental cross-semivariogram for the factors F_j and F_k. The averages of the three measures over the set of lags will be denoted by $\bar{\zeta}$, $\bar{\tau}$ and $\bar{\kappa}$ respectively. For a "good" decorrelation values for $\bar{\zeta}$ and $\bar{\tau}$ should be close to 0, while that for $\bar{\kappa}$ should be above 0.9.

3 Test models and sample locations

For testing purposes three linear models of co-regionalisation were constructed in each case based on 5 attributes. The first model is an example of an intrinsic co-regionalisation, the second has two nested structures with non-commuting coefficient matrices and the third 3.
Model 1 is given by

$$\Gamma_1(\mathbf{h}) = C_1(0.3\text{nugget}(\mathbf{h}) + 0.7\text{spher}_{40}(\mathbf{h})) \tag{9}$$

with

$$C_1 = \begin{bmatrix} 1.0000 & 0.4392 & 0.3332 & 0.7143 & 0.6226 \\ 0.4392 & 1.0000 & 0.6869 & 0.5380 & 0.4705 \\ 0.3332 & 0.6869 & 1.0000 & 0.3195 & 0.4524 \\ 0.7143 & 0.5380 & 0.3195 & 1.0000 & 0.4661 \\ 0.6226 & 0.4705 & 0.4524 & 0.4661 & 1.0000 \end{bmatrix}.$$

Model 2 is given by

$$\Gamma_2(\mathbf{h}) = C_{21}\text{nugget}(\mathbf{h}) + C_{22}\text{spher}_{50}(\mathbf{h}) \tag{10}$$

where

$$C_{21} = \begin{bmatrix} 0.3393 & 0.1725 & 0.1389 & 0.2587 & 0.2600 \\ 0.1725 & 0.2149 & 0.1099 & 0.1718 & 0.1074 \\ 0.1389 & 0.1099 & 0.2929 & 0.1436 & 0.1474 \\ 0.2587 & 0.1718 & 0.1436 & 0.3482 & 0.1677 \\ 0.2600 & 0.1074 & 0.1474 & 0.1677 & 0.3772 \end{bmatrix}$$

and

$$C_{22} = \begin{bmatrix} 0.6607 & 0.0496 & 0.4400 & 0.3510 & 0.3524 \\ 0.0496 & 0.7851 & 0.2283 & 0.3621 & 0.3254 \\ 0.4400 & 0.2283 & 0.7071 & 0.4738 & 0.3747 \\ 0.3510 & 0.3621 & 0.4738 & 0.6518 & 0.3299 \\ 0.3524 & 0.3254 & 0.3747 & 0.3299 & 0.6228 \end{bmatrix}.$$

Model 3 is given by

$$\Gamma_3(\mathbf{h}) = C_{31}\mathrm{spher}_5(\mathbf{h}) + C_{32}\mathrm{spher}_{15}(\mathbf{h}) + C_{33}\mathrm{spher}_{45}(\mathbf{h}) \tag{11}$$

where

$$C_{31} = \begin{bmatrix} 0.3251 & 0.1972 & 0.2146 & 0.1291 & 0.0881 \\ 0.1972 & 0.3559 & 0.1732 & 0.1666 & 0.2489 \\ 0.2146 & 0.1732 & 0.3238 & 0.1688 & 0.0996 \\ 0.1291 & 0.1666 & 0.1688 & 0.3506 & 0.1787 \\ 0.0881 & 0.2489 & 0.0996 & 0.1787 & 0.2996 \end{bmatrix}$$

$$C_{32} = \begin{bmatrix} 0.1922 & 0.1108 & 0.1102 & 0.0983 & 0.0777 \\ 0.1108 & 0.2406 & 0.1181 & 0.1349 & 0.1034 \\ 0.1102 & 0.1181 & 0.3035 & 0.1205 & 0.1864 \\ 0.0983 & 0.1349 & 0.1205 & 0.2182 & 0.0716 \\ 0.0777 & 0.1034 & 0.1864 & 0.0716 & 0.1953 \end{bmatrix}$$

and

$$C_{33} = \begin{bmatrix} 0.4827 & 0.1337 & -0.0037 & 0.0960 & 0.3039 \\ 0.1337 & 0.4035 & 0.2095 & 0.2089 & 0.3052 \\ -0.0037 & 0.2095 & 0.3727 & 0.2786 & 0.2014 \\ 0.0960 & 0.2089 & 0.2786 & 0.4312 & 0.2683 \\ 0.3039 & 0.3052 & 0.2014 & 0.2683 & 0.5051 \end{bmatrix}.$$

For each of the models 100 non-conditional simulations were generated via the turning bands method ([16]), and four random sample designs (Fig. 1) representing 1, 2, 5 and 10 percent of all locations were applied.

4 Results

Two sets of experiments were performed. The first was based on the semivariogram models introduced in Section 3 to assess the quality of the spatial decorrelation achieved in the absence of noise, and the second on the application of the diagonalisation approaches on the families of semivariogram matrices computed from the simulated sampled data.

4.1 No noise

Semivariogram matrices at lag values from 5 to 65 were generated in 5 unit increments. For Model 1, the decorrelation via PCA is perfect, while MAF, RJD and UWEDGE fail. For Models 2 and 3 the mean decorrelation measures (Table 1) indicate that the NOJD methods do better than the OJD methods. In the case of Model 2, the decorrelation achieved via the two NOJD methods is perfect, while for the OJD

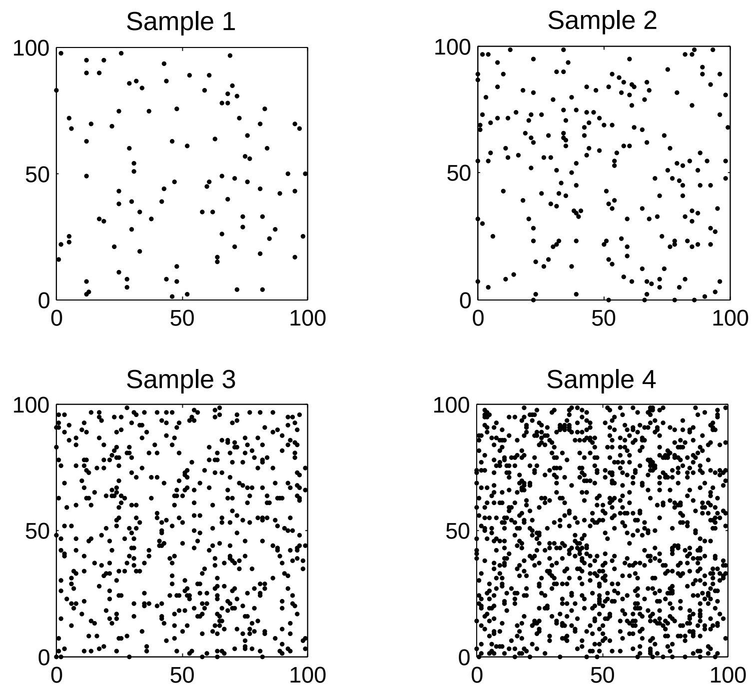

Fig. 1 Sample locations for decorrelation experiments

methods the decorrelation is only approximate, but the measures are comparable. For Model 3 the results from UWEDGE and MAF are comparable and better than those achieved by PCA or RJD. Of the latter, RJD yields better results than PCA. Using the eigenvalue criterion developed in [9] we may in addition conclude the diagonalisors determined by the OJD and NOJD methods for Models 2 and 3 are unique.

Table 1 Measures of decorrelation for theoretical models

	Model 2				Model 3			
	UWEDGE	MAF	PCA	RJD	UWEDGE	MAF	PCA	RJD
$\bar{\zeta}$	0	0	0.007	0.004	0.002	0.006	0.020	0.006
$\bar{\tau}$	0	0	0.078	0.058	0.042	0.036	0.199	0.119
$\bar{\kappa}$	1	1	0.995	0.997	0.997	0.998	0.934	0.982

4.2 Diagonalisation of experimental semivariogram matrices

To assess the ability of the algorithms to jointly diagonalise the semivariogram matrices, the semivariogram matrices were calculated at 14 lags with a lag spacing of 5 for samples 1 and 2, for samples 3 and 4 a total of 32 lags at a lag value of 2 was used. In each case the numerical measures of decorrelation were calculated.

The results for all three models are similar. Irrespective of the underlying model, the overall best decorrelation is achieved by UWEDGE, and the worst with PCA, even in the case where the underlying model exhibits intrinsic co-regionalisation. Results improve with the number of semivariogram matrices available for the decorrelation. For Model 1, the RJD method has better decorrelation results than MAF, but for models 2 and 3 MAF performs slightly better than RJD except for the smallest sample (see Fig. 2 and Table 2 for the results for the simulations run with Model 2), which underlines the need for a non-orthogonal diagonaliser.

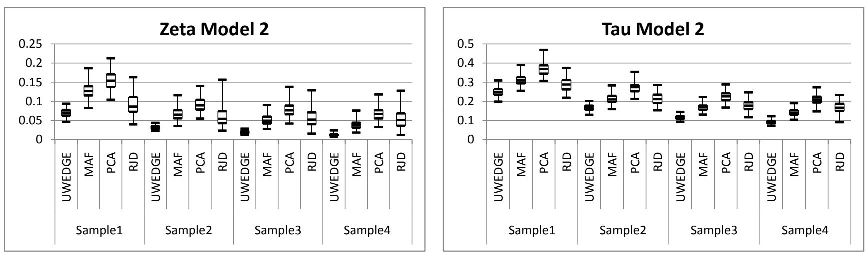

Fig. 2 Box and whisker plots for the mean absolute (left) and mean relative deviation from diagonality (right), Model 2

The transformation matrices, even though unique for each set of semivariogram matrices differ substantially between realisations.

Table 2 Average of the mean decorrelation efficiency, Model 2

	UWEDGE	MAF	PCA	RJD
Sample 1	0.933	0.885	0.867	0.928
Sample 2	0.972	0.957	0.936	0.957
Sample 3	0.990	0.980	0.961	0.968
Sample 4	0.993	0.985	0.965	0.969

5 Conclusion

The experiments discussed in the previous section indicate that there is benefit in looking beyond the use of PCA or MAF for the transformation of spatially correlated attributes into un-correlated factors. Irrespective as to the characteristics of the underlying model the NOJD methods have led to superior results compared to those employing orthogonal transformation matrices. So long as the sole purpose is a decorrelation for subsequent simulation, their use should therefore be preferred.

However, for both MAF and PCA there is a reasonably natural interpretation of the resultant factors, this is not the case for factors from UWEDGE or RJD, which may make these methods less useful for structure identification. More work is needed to assess this aspect.

References

1. Wackernagel H, 2003, Multivariate Geostatistics (3rd rev. ed.). Berlin: Springer-Verlag.
2. Switzer P, Green AA, 1984, Min/Max autocorrelation factors for multivariate spatial imaging, Palo Alto, California: Stanford University
3. Bandarian EM, Mueller UA, 2008, Reformulation of MAF as a generalised eigenvalue problem. In Ortiz J, Emery X Proceeding of the Eighth International Geostatistics Congress, Santiago, Chile, pp 1173-1178.
4. Bandarian EM, 2008, Linear Transformation Methods for Multivariate Geostatistical Simulation. PhD thesis, ECU, Perth.
5. Bandarian EM, Mueller UA, Ferreira J, Richardson S, 2010, Transformation methods for multivariate geostatistical simulation minimum/maximum autocorrelation factors and alternating columns-diagonal centres. In Dimitrakopoulos R Advances in orebody modelling and strategic mine planning I, Spectrum Series 17
6. Mueller U, Ferreira J, 2011, Multivariate geostatistical simulation via joint approximate diagonalisation: a case study , in Baafi, EY, Kininmonth, RJ and Porter, I: Proceedings of the 35th APCOM Symposium, Wollongong, 24-30 September 2011, 217-230
7. Mueller U, Ferreira J, 2012, The U-Wedge transformation method for multivariate geostatistical simulation, Math Geosci DOI 10.1007/s11004-012-9384-7
8. Desbarats JA, Dimitrakopoulos R, 2000, Geostatistical simulation of regionalized pore-size distributions using Min/Max autocorrelation factors. Math Geol 32(8):919-942
9. Afsari B, 2008, Sensitivity analysis for the problem of matrix joint diagonalisation, SIAM Journal of Matrix Analysis and Applications: 30, 1148
10. Tran TT, Murphy M, Glacken I, 2006, Semivariogram structures used in multivariate conditional simulation via minimum/maximum autocorrelation factors. Proceedings XI International Congress, IAMG, Liège
11. Theis F, Inouye, Y, 2006, On the use of joint diagonalization in blind signal processing, International Symposium on Circuits and Systems (ISCAS 2006), 21-24 May 2006
12. Yeredor A, 2002, Non orthogonal joint diagonalization in the least square sense with application in blind source separation, IEEE Trans of Sigl Proc 50(7):645-648.
13. Tichavsky P, Yeredor A, 2009, Fast approximate joint diagonalization incorporating weight Matrices. IEEE Trans on Sig Proc 57 (3): 878 - 891
14. Cardoso JK, Souloumiac A, 1996, Jacobi angles for simultaneous diagonalisation. SIAM Journal of Matrix Analysis and Applications: 17(1), 161-164.
15. Tercan AE ,1999, Importance of orthogonalization algorithm in modelling conditional distributions by orthogonal transformed indicator methods, Math Geol, 31 (2): 155-174.

16. Chilès JP, Delfiner P, 1999, Geostatistics: Modeling Spatial Uncertainty, Wiley Series in Probability and Statistics.