NUMERICAL ASPECTS OF THE UNIVERSAL KRIGING METHOD FOR HYDROLOGICAL APPLICATIONS

CARLOS LÓPEZ
Centro de Cálculo
Facultad de Ingeniería, CC 30, Montevideo, Uruguay
JAVIER SAMPER
E.T.S. de Ingenieros de Caminos, Canales y Puertos.
Campus de Elviña, 15192 La Coruña, Spain

Abstract. Many hydrological variables usually show the presence of spatial drifts. Most often they are accounted for by either universal or residual kriging and usually assuming low order polynomials. The Universal Kriging matrix ($\mathbf{M}$), which includes in it the values of the polynomials at data locations (matrix $\mathbf{F}$), in some cases may have a too large condition number and can even be nearly singular due to the fact that some columns are close to be linearly dependent. These problems are usually caused by a combination of pathological data locations and an inadequate choice of the coordinate system. As suggested by others, we have found that an appropriate scaling can alleviate the problem by significantly reducing the condition number of $\mathbf{M}$ (cond($\mathbf{M}$)). This scaling, however, does not affect the linear independence of this matrix. We show that a QR factorization of matrix $\mathbf{F}$ leads to a significant improvement on cond($\mathbf{M}$). An alternative to drift polynomials consists on using a set of functions derived from the eigenvectors of the variogram matrix (Γ). Although their potential usefulness as interpolating functions remains to be ascertained, they are optimal from the point of view of optimizing cond($\mathbf{M}$). In fact we are able to provide a rigorous proof for an upper bound of cond($\mathbf{M}$). Applications of the theoretical developments to hydraulic head data from an alluvial aquifer are also presented.

1. Introduction

The sensitivity of the kriging method in general, and the universal kriging in particular, either to input data or to variogram models have been analyzed by Diamond and Armstrong (1984), Posa (1989) and O'Dowd (1991). Diamond and Armstrong (1984) reported results on the relationship between the relative change in the kriging weights and the condition number of the ordinary kriging matrix. We remind the reader that the

condition number of the matrix of a system of equations measures the sensitivity of the solution to modifications on the input data (Atkinson, 1989). If the condition number is too large, the significance of the results (in this case, the universal kriging weights) can be completely lost. Posa (1989) studied the condition number of the ordinary kriging method for various well-known covariance models. He noticed that unstable situations occur when Gaussian models without a nugget are used. O'Dowd (1991) showed that for ordinary kriging the condition number of the matrix $\mathbf{M}$ has a lower bound, which depends on the covariance model and data location. He also showed some evidence that a suitable way to reduce cond($\mathbf{M}$) consists on scaling the sill (when defined) or conversely substituting the column and row of 1's to a different constant value. His findings were based upon theoretical results for a pure nugget variogram. Here we analyze different alternatives to lower the condition number of the universal kriging (UK) matrix. We first summarize briefly the UK theory and discuss some of the problems and inconsistencies of the standard UK matrix. Alternative approaches to overcome such problems include: (1) the factorization of matrix $\mathbf{F}$, and (2) the use of the eigenvectors of the variogram matrix as drift functions. The performance of these approaches is evaluated in terms of both the condition number and the estimation capabilities. To that purpose we use hydraulic head data from the Andújar alluvial aquifer in Southern Spain.

2. Universal kriging theory

Let Z(x) be a non-stationary random function. Its mathematical expectation $E\left[Z(x)\right]=m(x)$ exists and can be modeled as a linear combination of basis functions $f_l(x), l=1,...,\mathrm{P}$. The difference $\varepsilon(x)= Z(x)-m(x)$ is assumed to be a zero mean second order stationary random function having a variogram γ_ε (Samper and Carrera, 1990):

$$Z(x)=m(x)+\varepsilon(x); E(\varepsilon(x))=0 \text{ and } m(x)=\sum_{1}^{N} a_l f_l(x) \qquad (1)$$

where a_l are the drift coefficients. Eq. (1) is assumed to hold locally (otherwise, residual kriging might be applied). It is also assumed that the variogram is known. The UK estimator is a linear combination of the data values,

$$Z^*(x)=\sum_{1}^{N}\lambda_i Z(x_i)=\sum_{1}^{N}\lambda_i Z_i$$

In order to be unbiased, this estimator must satisfy:

$$E(Z^*) = E(\sum_{1}^{N}\lambda_i Z_i) = E(Z) = E\left[m(x)+\varepsilon(x)\right] = E\left[m(x)\right] = m(x)$$

This means that the kriging coefficients λ_i must satisfy the following equations

$$\sum_{1}^{N} \lambda_i \, m(x_i) = m(x) \tag{2}$$

Assuming that m(x) can be appropriately described by a function subspace named f, Eq. (2) can be rewritten as (Samper and Carrera, 1990):

$$\sum_{i=1}^{N} \lambda_i \left(\sum_{l=1}^{P} a_l \, f_l(x_i) \right) = \sum_{l=1}^{P} a_l \left(\sum_{i=1}^{N} \lambda_i \, f_l(x_i) \right) = \sum_{l=1}^{P} a_l \, f_l(x)$$

which holds if

$$\sum_{i=1}^{N} \lambda_i \, f_l(x_i) = f_l(x) \qquad l = 1..P$$

Kriging equations are obtained after minimizing the estimation error

$$\sum_{j=1}^{N} \lambda_j \gamma_\varepsilon(x_i - x_j) + \sum_{l=1}^{P} \mu_l \, f_l(x_i) = \gamma_\varepsilon(x_i - x) \qquad i = 1,\ldots,N(x)$$

where μ_l is the l-th Lagrange multiplier and N(x) is the size of the kriging neighborhood which is not necessarily equal to N because typically only the closest points are considered. Our calculations assume that N(x)=N, therefore our estimator is global. The system to be solved can be written in matrix form as

$$\mathbf{M} \begin{bmatrix} \underline{\lambda} \\ \underline{\mu} \end{bmatrix} = \begin{bmatrix} \Gamma & \mathbf{F} \\ \mathbf{F}^T & 0 \end{bmatrix} \begin{bmatrix} \underline{\lambda} \\ \underline{\mu} \end{bmatrix} = \begin{bmatrix} \underline{\gamma} \\ \underline{f} \end{bmatrix} \tag{3}$$

where $\underline{\lambda}$ is the vector of kriging coefficients, $\underline{\mu}$ is the vector of Lagrange multipliers, $\mathbf{F}$ is a rectangular matrix with entries $f_l(x_i)$, $\underline{\gamma}$ is a vector of variogram values $\gamma_j = \gamma_\varepsilon(x - x_j)$ and $\underline{f}$ is the vector of values of $f_l(x)$ at data locations.

3. Some inconsistencies of the standard approach

This standard approach has some problems which are discussed below.

3.1. LACK OF HOMOGENEITY ON THE ENTRIES OF KRIGING MATRIX

The units of matrix Γ are the same as those of the variogram. In the standard Universal Kriging method, $\mathbf{F}$ is formed by evaluating monomials at data points. Therefore, the entries of matrix $\mathbf{F}$ are derived from powers of distance, i.e., their units are different than those of Γ. Moreover, the entries of matrix $\mathbf{F}$ have very different absolute values. For this reason, numerical instabilities may arise.

3.2. STRONG DEPENDENCY ON THE ORIGIN OF COORDINATES

For a two-dimensional variable, the entries of matrix $\mathbf{F}$ are $f_{il} = f_l(x_i) = (x_i)^m (y_i)^{l-m}$. Their values depend strongly on the origin of coordinates. In fact, the position of the origin affects strongly the condition number of the system which in turn affects the precision of the estimator Z^* (Diamond and Armstrong, 1984). In some exploratory calculations using synthetic data, we performed a standard change of variables by mapping every (x,y) location to the range (-S,S),(-S,S) being S the sill, and used such standardized coordinates for creating $\mathbf{F}$. Such simple mapping leads to a decrease of the condition number of the full matrix from 10^8 to 10^4.

3.3. LINEAR INDEPENDENCE OF THE COLUMNS OF $\mathbf{F}$

The columns of matrix $\mathbf{F}$ are built from the values of known functions (typically monomials) at data locations. They should be linearly independent in order to uniquely determine the drift coefficients a_l. Such property is usually not checked because most programs use monomials, which are supposed to be independent by themselves. However, when considering arbitrary (x,y) data coordinates, such property does not necessarily hold. It should be noticed that there are *at most* N(x) independent column vectors (monomials) in the N(x) dimension space, so P is strictly bounded.

The problems discussed in Sections 3.1 and 3.2 lead to a large cond($\mathbf{M}$), while that of Section 3.3 may preclude to get a meaningful solution.

4. Suggested approaches

The problems discussed in the previous section can be overcome in several ways. Here we have considered three possibilities:

1. A proper transformation of the coordinate system in order to balance the columns of $\mathbf{F}$. This can be done by transforming the coordinate system (translation and scaling) so that any every point (x,y) is mapped into a square whose x and y limits are within the range (-S,S). Actual results illustrating the effect of this mapping are presented in Section 5.
2. Choosing the monomials in such a way that the orthogonality of the columns of $\mathbf{F}$ is ensured. Since there is no unique option for such a choice, we suggest to build the columns of $\mathbf{F}$ from a set of linearly independent and orthogonal vectors which are dimensionally consistent with matrix Γ. Such a set can be derived from the singular value decomposition (eigenvectors and eigenvalues) of matrix Γ. With this set of functions, the problems discussed in Sections 3.1, 3.2 and 3.3 are overcome. However, the resulting drift does not have a simple structure as that described in Eq. (1). With this method it is possible to optimize cond($\mathbf{M}$). In fact, in Section 4.1 we derive theoretical bounds for cond($\mathbf{M}$).

3. A QR factorization of matrix $\mathbf{F}$. With this approach it is possible to solve directly the problems mentioned in Sections 3.2 and 3.3. The lack of homogeneity of the entries of matrix $\mathbf{M}$ (Section 3.1) is managed in an indirect manner.

The first procedure is fairly simple and needs no additional explanation. We will describe with some detail the other two approaches. The performance of all three approaches using hydraulic head data is evaluated in Section 5.

4.1. SINGULAR VALUE DECOMPOSITION OF MATRIX Γ

Let $\mathbf{E}$ and Λ be the matrices containing the eigenvectors and eigenvalues of matrix Γ, respectively. These matrices are related through $\Gamma = \mathbf{E}\Lambda\mathbf{E}^{-1} = \mathbf{E}\Lambda\mathbf{E}^{T}$. Notice that $\mathbf{E}$ is orthogonal because Γ is symmetric. Matrix Λ is a diagonal matrix containing the eigenvalues. Let $\mathbf{D}_{PxP}$ be a square diagonal matrix holding P eigenvalues and let $\mathbf{G}_{NxP}$ be a matrix formed using the corresponding P columns of $\mathbf{E}$. We suggest to build matrix $\mathbf{F}$ such that

$$\mathbf{F} = \mathbf{G}.\mathbf{D} \tag{4}$$

Since the eigenvalues of Γ share its units, so does matrix $\mathbf{F}$. The columns of $\mathbf{G}$ are orthogonal among themselves, $\mathbf{G}^{T}.\mathbf{G} = \mathbf{I}_{P}$ ($\mathbf{I}_{P}$ is the identity matrix of size P) and orthogonal to all the columns of $\mathbf{E}$, that is, $\mathbf{G}^{T}.\mathbf{E} = \begin{bmatrix}\mathbf{I}_{P} & \mathbf{0}\end{bmatrix}$ (here $\mathbf{0}$ denotes a matrix having all its entries equal to zero). Another relevant property is

$$\mathbf{D}\begin{bmatrix}\mathbf{I}_{P} & \mathbf{0}\end{bmatrix}\Lambda^{-1}\begin{bmatrix}\mathbf{I}_{P} & \mathbf{0}\end{bmatrix}^{T} = \mathbf{I}_{P}$$

From Eq. (3) it follows that

$$\begin{cases}\Gamma\underline{\lambda} + \mathbf{F}\underline{\mu} & = \underline{\gamma} \\ \mathbf{F}^{T}\underline{\lambda} & = \underline{f}\end{cases} \Rightarrow \begin{cases}\underline{\lambda} + \Gamma^{-1}\mathbf{F}\underline{\mu} & = \Gamma^{-1}\underline{\gamma} \\ \mathbf{F}^{T}\underline{\lambda} & = \underline{f}\end{cases}$$

Premultiplying the first row by $\mathbf{F}^{T}$, and using the second, one obtains

$$\mathbf{F}^{T}\Gamma^{-1}\mathbf{F}\underline{\mu} = \mathbf{F}^{T}\Gamma^{-1}\underline{\gamma} - \underline{f} \tag{5}$$

This equation holds for any matrix $\mathbf{F}$. By using Eq. (4), the left-hand side reduces to

$$\mathbf{F}^{T}\Gamma^{-1}\mathbf{F} = \mathbf{D}^{T}\mathbf{G}^{T}\mathbf{E}\Lambda^{-1}\mathbf{E}^{T}\mathbf{G}\mathbf{D} = \mathbf{D}^{T}\begin{bmatrix}\mathbf{I}_{P} & \mathbf{0}\end{bmatrix}\Lambda^{-1}\begin{bmatrix}\mathbf{I}_{P} \\ 0\end{bmatrix}\mathbf{D} = \mathbf{D}^{T}\mathbf{D}^{-1}\mathbf{D} = \mathbf{D}$$

The first term on the right-hand-side of Eq. (5) can be rewritten as

$$\mathbf{F}^{T}\Gamma^{-1}\underline{\gamma} = \mathbf{D}^{\mathbf{T}}\mathbf{G}^{\mathbf{T}}\mathbf{E}\Lambda^{-1}\mathbf{E}^{\mathbf{T}}\,\underline{\gamma} = \mathbf{D}^{\mathbf{T}}\mathbf{G}^{\mathbf{T}}\mathbf{E}\Lambda^{-1}\mathbf{E}^{\mathbf{T}}\,\mathbf{G}\,\mathbf{G}^{\mathbf{T}}\underline{\gamma} = \mathbf{G}^{T}\underline{\gamma}$$

Consequently, $\underline{\mu}$ can be derived easily from the following diagonal system

$$\mathbf{D}\underline{\mu} = \mathbf{G}^{T}\underline{\gamma} - f$$

The eigenvalues and eigenvectors of matrix $\mathbf{M}$ can be expressed in terms of those of Γ in a closed form. Appendix 1 contains the details of such expressions. This appendix contains also a proof for the upper bound of the condition number of the kriging matrix:

$$cond(\mathbf{M}) \leq 2.6\, cond(\Gamma)$$

O'Dowd (1991) stated that cond($\mathbf{M}$) cannot be less than cond(Γ). Here we were able to prove that cond($\mathbf{M}$) has a well defined upper bound for the approach based on the singular value decomposition of Γ.

4.2. QR ORTHOGONALIZATION OF MATRIX F

The QR factorization of matrix $\mathbf{F}$ renders a unique set of matrices $\mathbf{Q}$ and $\mathbf{R}$ which satisfy

$$\mathbf{F} = \mathbf{QR}$$

$$\mathbf{Q}^T\mathbf{Q} = \mathbf{I}$$

where R is an upper triangular matrix. It is straightforward to rewrite the system of equations (3) in a slightly different manner so that the overall matrix has a lower condition number. The new form of the system of equations reads as

$$\begin{bmatrix} \Gamma & \mathbf{Q} \\ \mathbf{Q}^T & 0 \end{bmatrix} \begin{bmatrix} \underline{\lambda} \\ \underline{\mu}^* \end{bmatrix} = \begin{bmatrix} \underline{\gamma} \\ \underline{f}^* \end{bmatrix} \tag{6}$$

where $\underline{\mu}^*$ and $\underline{f}^*$ are defined in the following manner

$$\underline{f}^* = \left(\mathbf{R}^{-1}\right)^T \underline{f} \; ; \qquad \underline{\mu}^* = \mathbf{R}\ \underline{\mu} \tag{7}$$

This means that instead of solving Eq. (3) one can solve Eq. (6) and obtain the expressions of $\underline{\lambda}$ and $\underline{\mu}^*$. From Eq. (7) one can derive $\underline{\mu}$ from $\underline{\mu} = \mathbf{R}^{-1}\underline{\mu}^*$.

Notice that the orthogonality property of matrix $\mathbf{Q}$ has not been used yet. In fact, it is possible to re-scale its columns by a factor s in order to balance the entries of the system matrix. Moreover, the matrix can be factorized using the fact that $\Gamma = \mathbf{E}\Lambda\mathbf{E}^T$. Therefore, after introducing the scale factor s the system matrix can be rewritten as

$$\begin{bmatrix} \Gamma & s\mathbf{Q} \\ s\mathbf{Q}^T & \mathbf{0} \end{bmatrix} = \begin{bmatrix} \mathbf{E} & \mathbf{0} \\ \mathbf{0} & \mathbf{I} \end{bmatrix} \begin{bmatrix} \Lambda & s\mathbf{E}^T\mathbf{Q} \\ s\mathbf{Q}^T\mathbf{E} & \mathbf{0} \end{bmatrix} \begin{bmatrix} \mathbf{E}^T & \mathbf{0} \\ \mathbf{0} & \mathbf{I} \end{bmatrix} = \mathbf{BLB}^T \tag{8}$$

Matrix $\mathbf{L}$ has the same eigenvalues as matrix $\mathbf{M}$ because $\mathbf{B}$ is orthogonal (Atkinson, 1989). Notice that $\mathbf{L}$ has the eigenvalues λ_i of Γ in the diagonal and terms containing the scale factors in the rest of the first N rows. This fact can be exploited by using the Gershgorin theorem [see Dahlquist *et al.* (1980)] which states that matrix $\mathbf{L}$ has N eigenvalues $e_i(s)$ which verify:

$$e_i(s) = \lambda_i + O(s) \qquad i = 1, \ldots, N \tag{9}$$

where $O(s)$ is an infinitesimal of first order on s (the limit of $O(s)/s$ as *s* goes to zero is a constant). The eigenvalues corresponding to the P remaining rows (those with zeros in the diagonal) behave like

$$e_i(s) = O(s) \qquad i = N+1, \ldots, N+P \tag{10}$$

Dahlquist *et al*. (1980) also give a proof that the eigenvalues of a perturbed matrix are continuous in terms of the perturbation factor. From Equations (9) and (10) it is possible to relate the condition numbers of matrices Γ and $\mathbf{L}$. For small s, the largest eigenvalues of $\mathbf{L}$ will coincide with those of Γ, while the smallest eigenvalues will be on the order of s. Therefore, for small s the condition number of $\mathbf{L}$ will behave like O($1/s$). For large enough s, the condition number grows as $\sqrt{s}$. For intermediate values of s, we noticed that the evolution of the minimum eigenvalue in terms of s has an absolute minimum s^* for some moderate s. We tried to estimate such a value in terms of some norm of matrix Γ, and for some limited analyses we concluded that a good guess is simply $s^* = \|\Gamma\|_1 / N$, where the term in the numerator is the norm of matrix Γ (the maximum row sum of the absolute values). The optimum value s^* is bounded by the sill (when it is defined).

5. Results for the Andújar case

Previous theoretical analyses have been applied to hydraulic head data from the Andújar alluvial aquifer in Southern Spain where a uranium mill tailings has caused a uranium plume in the aquifer (Samper and Carrera, 1995). This case study is a small-scale field problem in which spatial variability plays a major role. Geostatistical analyses of hydraulic heads in this aquifer have been carried out by ENRESA (1992) for the purpose of optimizing the head monitoring network and by Juanes and Samper (1996) who used residual kriging for estimating the hydraulic head field. The study area covers about 12 km^2 and is bounded by the Guadalquivir river on the North and a low permeability mountain on the south. The actual data set used to test the three procedures is that of Juanes and Samper (1996) which consists of 82 head data measured in April 1991. In their geostatistical analysis they identified a global linear drift. Residuals fitted an exponential semivariogram with a sill of 0.23 m^2 and an effective range of about 1 Km. Using the same data, we have performed universal kriging analyses using four different methods: (1) linear polynomials, (2) scaled linear polynomials, (3) the method of eigenvalues , and (4) linear polynomials with a QR factorization. The upper-left plot in Figure 1 shows a contour map of the residuals (raw data minus spatial drift). This plot was prepared using generalized least squares with first order polynomials.

Figure 1 shows also the plots of the spatial distribution of three eigenvectors (factors) of matrix Γ. These surfaces show some general patterns. In general, the eigenvectors associated to the largest eigenvalues show the largest spatial variation with clear global drifts (see the upper right plot in Figure 1). Eigenvectors corresponding to the smaller eigenvalues tend to show erratic

spatial patterns. The interpretation of the spatial distribution of other eigenvectors is not straightforward because they do not reflect any spatial feature of the random function, but rather are determined by the semivariogram and the data locations. Their interpretation deserves additional analyses which have not yet been addressed and will be presented elsewhere.

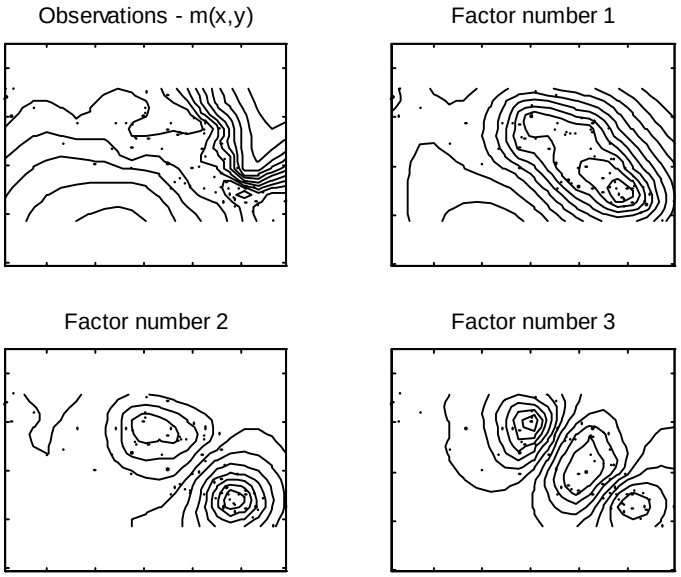

Figure 1. Contour map of Andújar raw hydraulic head data (upper left plot) and contour maps of three eigenvectors (factors) used to represent the spatial drift in the eigenvector method (remaining plots). Dots indicate data locations.

In applying the QR factorization method, we investigated the optimum value of the scale factor "s". Figure 2 shows a plot of the minimum eigenvalue of $\mathbf{M}$ as a function of s/s* (s* is our suggested estimate for the optimum). In this figure the star indicates the true optimum value. One can see that the true optimum is slightly greater than s*. The condition number of $\mathbf{M}$ as a function of s is shown in Figure 3.

Table 1 summarizes the results obtained with the four approaches. The fourth column in this table indicates the average value of the ratio between cond($\mathbf{M}$) and cond(Γ). Unscaled monomials lead to much greater condition numbers (eight orders of magnitude in this case!). The other methods give similar results. The results for scaled monomials are slightly better that those of the QR factorization, which in turn are very close to those obtained with eigenvectors.

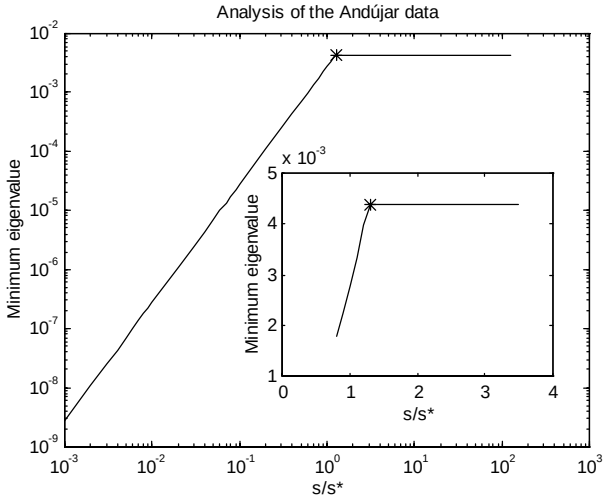

Figure 2 Evolution of the minimum eigenvalue for different choices of the s/s* ratio for the Andújar dataset. s stands for the scale factor used, while s* is the value suggested.The star in the figure indicates the optimum value.

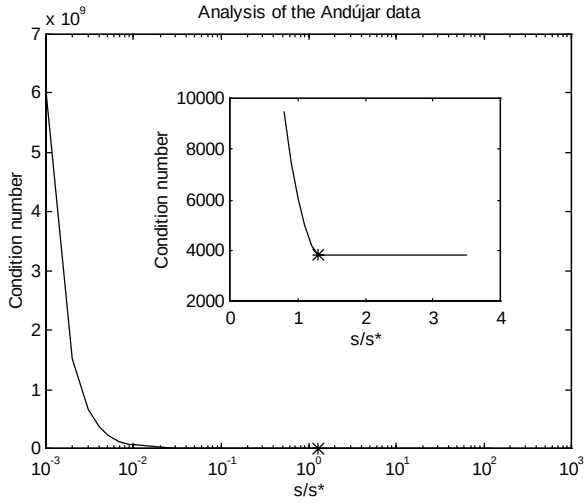

Figure 3 Evolution of the condition number for different choices of the s/s* ratio for the Andújar dataset. s stands for the scale factor used, while s* is the value suggested.The star in the figure indicates the optimum value.

For testing and comparing the estimation capabilities of the different approaches we have resorted to cross-validation. Cross-validation provides the estimation errors e_i and their corresponding standard deviations σ_i at all selected locations. Following Samper and Neuman

(1989), we use the Negative Log-likelihood (NLL) of the errors as a measure of model goodness. For Gaussian cross-validation errors, NLL contains two terms S_1 and S_2 which are defined as:

$$NLL = e^T \mathbf{V}^{-1} e + \ln|\mathbf{V}| = S_1 + S_2$$

where e is the column vector of cross-validation errors, and $|\mathbf{V}|$ is the determinant of their covariance matrix $\mathbf{V}$.

As expected, cross-validation results are identical for all methods based on polynomials. Their NLL values are all equal. The method based on eigenvectors, however, performs much worse than the others.

Table 1 Cross validation results the Andújar head data set.

Method	S_1	S_2	*NLL*	$\frac{\overline{cond(\mathbf{M})}}{cond(\Gamma)}$
Monomials	20.95	-137.8	-116.8	5.598e+008
Re-scaled monomials	20.95	-137.8	-116.8	1.016
Monomials + QR factorization	20.95	-137.8	-116.8	1.672
Eigenvectors	253.6	-138.9	114.7	1.618

6. Conclusions

We have analyzed the problem of the condition number of the universal kriging method. We have compared the performance of four methods in terms of condition number and estimation capabilities by analyzing the likelihood of cross-validation errors. Our analyses indicate that:

1. very important improvements can be attained by re-scaling the coordinates of the data to the range (-S,S),(-S,S) being S the sill when defined, or the maximum element on matrix Γ in other case.
2. a general theoretical result for the condition number of kriging matrix has been obtained by using the eigenvectors of the variogram matrix. This approach performed worse than the others, but there is room still for some improvements. The shape of the fields derived from the eigenvectors have to be studied further because they could exhibit some interesting properties yet unexplored.
3. QR orthogonalization of matrix $\mathbf{F}$ has proved to be a nearly optimum method. While the orthogonal property of the columns of $\mathbf{F}$ after the QR factorization is independent of the data set, the same may not always be the case for the re-scaling method.

7. References

Atkinson, K. 1989 An introduction to numerical analysis, John Wiley and Sons 692 pp. New Jersey, 573 p.

Dahlquist, G. and A. Bjorck 1974 Numerical Methods. Prentice-Hall, Inc. Englewood Cliffs, , n.5, p. 721-739

Diamond P. and Armstrong, M. 1984 Robustness of Variograms and Conditioning of Kriging Matrices. Math. Geol., (16): 809-822.

ENRESA, 1992 Estudio Geoestadístico de la Piezometría del Acuífero de la Fábrica de Uranio de Andújar (Internal Report), 93 pp.

Juanes, R. and J. Samper 1996 Estudio geoestadístico de los niveles piezométricos en el acuífero de la F.U.A. (Internal Report)

O'Dowd, R.J., 1991 Conditioning of coefficient matrices of ordinary kriging, Math. Geol., 23(5):721-739.

Posa, D., 1989 Conditioning of the stationary kriging matrices for some well-known covariance models. Math. Geol., 21(7):755-765

Samper, J. and S. P. Neuman, 1989 Estimation of Spatial Covariance Structures by Adjoint State Maximum Likelihood Cross Validation. 1. Theory. Water Resour. Res. 25(3):351-362

Samper, J. and J. Carrera, 1990 Geostadística: aplicaciones a la Hidrología Subterránea. Ed. CIMNE. Barcelona. Spain, 484 pp.

Samper, J. and J. Carrera, 1995 Numerical Modeling of Remedial Actions for an Uranium Mill tailings in Spain: from model construction to prediction uncertainty. In: Groundwater Quality: Remediation and Protection. Praga, IAHS Pub. N. 225, pp. 299-310.

Appendix 1: upper bound for cond (M)

Let $\underline{x}$ be an eigenvector of Γ and λ its corresponding eigenvalue. Let $\underline{z}$ be a vector defined as:

$$\underline{z} = \left[\underline{x} \quad 0 \quad \cdots \quad \alpha \quad 0 \quad \cdots\right]^T .$$

were α is a scalar to be determined. We will find the appropriate value of α by imposing that $\underline{z}$ to be an eigenvector of $\mathbf{M}$. Notice that $\left|\underline{z}\right|^2 = \underline{z}^T \underline{z} = \underline{x}^T \underline{x} + \alpha^2 = 1 + \alpha^2$. If we denote by N the size of Γ and N+P the size of $\mathbf{M}$, one has at most P different choices for $\underline{x}$ to be proportional to a column of $\mathbf{F}$, and N-P choices for $\underline{x}$ to be orthogonal to a column of $\mathbf{F}$.

In the first case, $\mathbf{F}^T \underline{x} = \left[0 \quad \cdots \quad \lambda \quad 0 \quad \cdots\right]^T$, therefore for $\underline{z}$ to be an eigenvector of $\mathbf{M}$ it should verify

$$\mathbf{M}\underline{z} = \left[(\lambda + \alpha\lambda)\underline{x} \quad 0 \quad \cdots \quad \lambda \quad 0 \quad \cdots\right]^T = \lambda^* \underline{z} = \lambda^* \left[\underline{x} \quad 0 \quad \cdots \quad \alpha \quad 0 \quad \cdots\right]^T$$

which holds if

$$\begin{cases} \lambda(1+\alpha) = \lambda^* \\ \lambda = \lambda^* \alpha \end{cases} \Leftrightarrow \begin{cases} 1+\alpha = \lambda^* / \lambda \\ 1/\alpha = \lambda^* / \lambda \end{cases} \Leftrightarrow 1+\alpha = 1/\alpha \Leftrightarrow \alpha_{1,2} = \frac{-1 \pm \sqrt{5}}{2}$$

Therefore for each λ there are two valid eigenvalues λ_1^* and λ_2^*. This renders in turn 2P valid eigenvectors $\underline{z}_1$ and $\underline{z}_2$, which are mutually orthogonal since $\underline{z}_1^T \underline{z}_2 = \underline{x}^T \underline{x} + \alpha_1 \alpha_2 = 1 - 1 = 0$.

For the second case, $\mathbf{F}^T \underline{x} = 0$ therefore $\mathbf{M}\underline{z} = \lambda \underline{z}$ and α=0. This implies that for such cases, $\lambda = \lambda^*$ and the eigenvalues are the same. There are N-P new eigenvectors which are also orthogonal to those already found.

Let λ_{max} and λ_{min} be the largest and the smallest eigenvalues of Γ, respectively. Depending on whether the corresponding eigenvectors are or are not columns of $\mathbf{F}$, the largest eigenvalue of $\mathbf{M}$ will be bounded by $\lambda_{max}/\alpha_1 \left(\alpha_1 < 1\right)$ while the smallest will be bounded by λ_{min}/α_2. Therefore the condition number of $\mathbf{M}$ is also bounded by

$$cond(\mathbf{M}) \le \frac{\lambda_{\max}/\alpha_1}{\lambda_{\min}/\alpha_2} = \frac{\alpha_2}{\alpha_1} cond(\Gamma) = \frac{\alpha_2^2}{\alpha_1 \alpha_2} cond(\Gamma) = \frac{\alpha_2^2}{-1} cond(\Gamma) =$$

$$= \left(\frac{-1-\sqrt{5}}{2} \right)^2 cond(\Gamma) \approx 2.6 cond(\Gamma)$$

which is valid for any choice of eigenvectors included in $\mathbf{F}$.

Acknowledgments

We want to thank the Uruguay National University, the University of La Coruña and the Galician Foundation for Civil Engineering (La Coruña, Spain) for their support and funding of the travel expenses of the first author to Spain and of the second author to Uruguay.