Mathematical Geology, Vol. 26, No. 1, 1994

On the Condition Number of Covariance Matrices in Kriging, Estimation, and Simulation of Random Fields[1]

Rachid Ababou[2], Amvrossios C. Bagtzoglou,[3] and Eric F. Wood[4]

The numerical stability of linear systems arising in kriging, estimation, and simulation of random fields, is studied analytically and numerically. In the state-space formulation of kriging, as developed here, the stability of the kriging system depends on the condition number of the prior, stationary covariance matrix. The same is true for conditional random field generation by the superposition method, which is based on kriging, and the multivariate Gaussian method, which requires factoring a covariance matrix. A large condition number corresponds to an ill-conditioned, numerically unstable system. In the case of stationary covariance matrices and uniform grids, as occurs in kriging of uniformly sampled data, the degree of ill-conditioning generally increases indefinitely with sampling density and, to a limit, with domain size. The precise behavior is, however, highly sensitive to the underlying covariance model. Detailed analytical and numerical results are given for five one-dimensional covariance models: (1) hole-exponential, (2) exponential, (3) linear-exponential, (4) hole-Gaussian, and (5) Gaussian. This list reflects an approximate ranking of the models, from "best" to "worst" conditioned. The methods developed in this work can be used to analyze other covariance models. Examples of such representative analyses, conducted in this work, include the spherical and periodic hole-effect (hole-sinusoidal) covariance models. The effect of small-scale variability (nugget) is addressed and extensions to irregular sampling schemes and higher dimensional spaces are discussed.

KEY WORDS: kriging, condition number, random fields, conditional simulation, covariance matrices, state-space estimation.

[1]Received 18 May 1992; accepted 21 May 1993.

[2]Centre d'Etude Nucléaire de Saclay, Commissariat a l'Energie Atomique, 91191 Gif Sur Yvette Cedex, France.

[3]Center for Nuclear Waste Regulatory Southwest Research Institute, Analyses, San Antonio, Texas 78238-5166.

[4]Department of Civil Engineering and Operations Research, Princeton University, Princeton, New Jersey 08540.

INTRODUCTION

The numerical implementation of geostatistical estimation and random field simulation requires solving dense linear systems involving a covariance matrix. The computational tractability of these algebraic systems can be characterized by the condition number of the covariance matrix, which depends on the underlying covariance structure and on the spatial configuration of measurements (location of data points). An analysis of the spectral condition number of covariance matrices is developed in this paper. We begin by presenting the context in which such problems arise.

Geostatistical Applications

The covariance matrix arises in the process of solving kriging systems for the posterior mean, or Best Linear Unbiased Estimate, of a given set of geostatistical data. Essentially, kriging is a particular type of spatial interpolation based on some assumed spatial structure. In ordinary kriging, the spatial structure is characterized by a stationary mean and a stationary two-point covariance function, usually with the implicit assumption that the observed variable has the statistical distribution of a Gaussian random field. In the case of positively skewed variables such as conductivity, porosity, etc., a one-to-one transform such as $\ln(x)$ can be used to make the Gaussian assumption more realistic.

Kriging is also used for conditional simulation of random fields by a superposition method (Delhomme, 1979; Journel and Huijbregts, 1978). The superposition method produces random fields conditioned on a set of measurements, by linearly combining unconditional random fields with kriged fields. As will be seen, kriging calculations involve the repeated solution of covariance matrix systems. This can be achieved by computing (only once) a triangular factorization of the matrix, and by using the triangular factors in a recursive backward-forward solution scheme for any (of the many) right-hand side vector(s).

Finally, one of the simplest methods for generating Gaussian random fields, although not necessarily the most efficient one, is the "multivariate Gaussian method" which also requires factoring a covariance matrix. Specifically, the method consists of factoring the covariance matrix C into a product of two triangular matrices M and M^T. The random field is then generated by multiplying M by a vector of replicates of independent, normalized Gaussian variables. The covariance matrix can be either the prior covariances among all grid points (unconditional generation), or the posterior covariances among all simulation points conditioned on measurement data, where the "simulation points" do not include the data points themselves (conditional generation). See Dagan (1982) for an example of application of this technique to stochastic subsurface hydrology.

Inasmuch as perfect knowledge of needed field data cannot be attained, linear estimation and conditional simulation play an important role in processing, interpolating, and statistically interpreting data. These methods are particularly important for spatially distributed models of hydrological processes, such as space-time rainfall fields, the large-scale contaminant migration in heterogeneous geological formations (nuclear waste disposal). In all the methods reviewed above, the feasibility of factorization (or inversion) depends on the covariance matrix being invertible. More precisely, the numerical stability and accuracy of matrix factorization (or inversion) depend on the condition number of the covariance matrix. The condition number ranges from one, for an identity matrix, to infinity for a non-invertible matrix. A large condition number means that the systems to be solved are very sensitive to small perturbations, in which case the estimation or simulation procedures become impractical.

Role of Covariance Matrix in Kriging

Kriging is essentially a statistical estimation or interpolation of spatially distributed data. The equivalence between kriging and state-space linear estimation theory was established by Chirlin and Wood (1982). Building on this, we will show that the linear estimation problem designated as "ordinary kriging" can be reduced to the solution of a set of *uncoupled* linear systems, where the matrix to be inverted or factored is the covariance matrix (C) whose elements are the prior covariances among data points. The usual geostatistical approach to this problem is based on a somewhat different formalism, although the results are identical.

In geostatistics, the estimated field is usually obtained by solving a *kriging system* which contains the covariance matrix as sub-matrix [Journel and Huijbregts (1978, Chap. V); De Marsily (1986, Chap. 11); Isaaks and Srivastava (1989, Chap. 12); O'Dowd (1991)]. The latter author pointed out that the condition number of the kriging matrix K is always larger than, or at least equal to, the condition number of the prior covariance matrix C. With this in mind, it will be worthwhile to reformulate the kriging problem using the state-space formalism suggested above. Thus, we will show that the kriging system can be solved directly by factoring the covariance matrix C. The condition number of this matrix intrinsically characterizes the computational difficulty of the problem.

Ill-Conditioning of Covariance Matrices

The spectral condition number of the covariance matrix can be used to characterize the numerical feasibility of inversion or factorization algorithms required in the class of simulation and kriging problems mentioned above. For a given matrix A, the spectral condition number $\kappa(A)$ is defined as the ratio of

largest to smallest eigenvalues, each taken in absolute value (see Golub and Van Loan, 1989; Press et al., 1986). That is:

$$\kappa(A) = \frac{\text{Max } |\lambda_k(A)|}{\text{Min } |\lambda_k(A)|} \tag{1}$$

To be more specific, let us assume from now on that the prior covariance function of the spatial field being simulated or estimated is stationary, that is $C(x, y) = C(\xi)$ where $\xi = y\text{-}x$ is the two-point separation or lag vector, and has unit variance, that is $C(0) = \sigma^2 = 1$. The associated covariance matrix is symmetric positive-definite. Now, it turns out that $\kappa(C)$ can be quite sensitive to the shape of the covariance function $C(\xi)$. The Gaussian covariance function appears to produce particularly ill-conditioned covariance and kriging matrices. Some pertinent observations gathered from the literature are briefly discussed below.

Ekstrom (1973), Lewis (1987), Posa (1989), and O'Dowd (1991), all indicate (in different ways) the ill-conditioned nature of Gaussian covariance systems in the context of linear estimation. Posa (1989) computed the condition number of the kriging matrix numerically for Gaussian, exponential, and spherical covariances, and concluded that the condition number of Gaussian covariance was the worst, and this for several configurations of the data points. Lewis (1987) and O'Dowd (1991) noticed that the ill-conditioned nature of the Gaussian covariance matrix (and of the associated kriging matrix) is related to the infinitely differentiable property of Gaussian covariance random fields. One curious consequence of this property is that, given information on the random field and *all* its derivatives at only one point, one can extrapolate the field with perfect accuracy at any other point (Yaglom, 1962). The extremely smooth nature of this random field is related to the very fast Gaussian decay of its spectral density at large wavenumbers (high spatial frequencies).

Based on these observations, it may be expected that under certain conditions to be determined (e.g., oversampling), the discrete eigenvalue spectrum of the Gaussian covariance matrix includes near-zero eigenvalues. These near-zero eigenvalues would lead to a very large condition number according to Eq. (1). Qualitative arguments along these lines were developed in a paper by Ekstrom (1973), for the case of numerical deconvolution of signals with highly continuous and smooth kernels.

Scope of Work

The purpose of the present work is to investigate the behavior of the condition number of the covariance matrix with respect to the following factors: (1) properties of the input covariance function, (2) sampling distance (mesh size

Δx), and (3) number of data points (grid size N). At issue is the question of choosing an appropriate covariance model. Field data usually do not provide sufficient information to resolve the exact shape of the covariance function, in which case the selected type of covariance function is chosen mostly for its convenience or some other subjective preference. Here, our aim is to find which covariance models produce the best-conditioned systems. Incidentally, this work will serve to determine some of the reasons for the excessive ill-conditioning experienced with models like the Gaussian covariance. These considerations are of practical importance for both geostatistical estimation and random field simulation.

After presenting a state-space formulation of the conditioning and kriging problem, this paper develops and analyzes new results on the numerical conditioning of the associated algebraic systems (influence of grid spacing, domain size, and covariance structure). This analysis requires the solution of eigenvalue problems. It is based, first, on results obtained by standard solution methods (characteristic polynomials and numerical decompositions), and, second, on complementary results obtained by a new solution method (spectral or continuum approximation).

It should be noted that several random field simulation procedures do not depend on the condition number of the covariance matrix. For instance, the multidimensional turning bands method of Mantoglou and Wilson (1982) and Tompson et al. (1989) generates stationary, unconditional random fields without any matrix inversion. Indeed, the turning bands method is based on one-dimensional Fourier transforms and multidimensional projections, neither of which requires matrix inversion. On the other hand, as explained earlier, inversion or factorization of a covariance matrix is required for: (1) random field generation by the multi-variate Gaussian method, (2) conditional random field generation by the superposition methods, and (3) geostatistical estimation or kriging. The inversion of covariance-type matrices is also required in signal processing packages.

In the next section, we indicate specifically how covariance matrix systems arise in ordinary kriging. This is developed by formulating kriging from the point of view of state-space linear estimation theory. The third section introduces different types of stationary covariance functions, and their associated spectral density functions. The fourth and fifth sections develop quantitative analyses of the spectral condition number of covariance matrices. Most of the analysis in these sections is devoted to the special case of one-dimensional space with uniform sampling scheme, although some preliminary extensions of results to non-uniform and multidimensional sampling schemes are indicated. Finally, a summary of our main results and conclusions is given.

COVARIANCE MATRIX AND STATE-SPACE FORMULATION OF KRIGING

In the geostatistical formulation of kriging [Journel and Huijbregts (1978, Chap. V); De Marsily (1986, Chap. 11); Isaaks and Srivastava (1989, Chap. 12)], a spatial field $y(\mathbf{x})$ is estimated at any desired point $\mathbf{x_0}$ from a linear combination of the values of $y(\mathbf{x})$ at measurement points (noted as $\mathbf{y_D}$). The kriging system is obtained by minimizing the error variance subject to the constraint that the estimate be unbiased, i.e., the mean of the estimator is equal to the mean of the random field $y(\mathbf{x})$. This leads to the $(N + 1) \times (N + 1)$ linear system:

$$\begin{bmatrix} [\mathbf{C}]_{N,N} & [\mathbf{1}]_{N,1} \\ [\mathbf{1}]^T_{1,N} & 0 \end{bmatrix} \cdot \begin{bmatrix} [\lambda]_{N,1} \\ -\mu \end{bmatrix} = \begin{bmatrix} [\mathbf{c_0}]_{N,1} \\ \mathbf{1} \end{bmatrix} \tag{2}$$

where $[\mathbf{C}]$ is the $N \times N$ matrix of covariances among the N data points, $[\mathbf{c_0}]$ is the N-vector of covariances between point $\mathbf{x_0}$ and the N data points, $[\lambda]$ is the N-vector of unknown kriging weights, μ is the unknown Lagrange multiplier used because of the unbiasedness constraint, and "$[\mathbf{1}]$" stands for a vector or matrix whose elements are all equal to one.

Note that the kriging matrix $\mathbf{K}$ appearing on the left-hand side of (2) includes $\mathbf{C}$ as sub-matrix. The last row and the last column of $\mathbf{K}$ express the unbiasedness constraint. Both $\mathbf{C}$ and $\mathbf{K}$ are nonsingular matrices. However, while $\mathbf{C}$ is symmetric positive-definite with positive real eigenvalues, $\mathbf{K}$ is indefinite with exactly one negative eigenvalue (Posa, 1989). Most importantly, as pointed out earlier, the condition number of the covariance matrix $\mathbf{C}$ is always smaller than, or at worst equal to, that of the kriging matrix $\mathbf{K}$ (O'Dowd, 1991).

The solution of the ordinary kriging system (2) can be reduced to the easier numerical problem of inverting matrix $\mathbf{C}$. This simplification arises quite naturally when the kriging equations are reformulated from the point of view of state-space linear estimation theory. Chirlin and Wood (1982) established the equivalence between kriging and state-space linear estimation in a general setting. Furthermore, we show below that the problem of linear estimation of N_E estimation points, given N_D data points, can be treated in a computationally efficient way by using block-matrix algebra. This technique was used for instance in the CSIMUL code for conditional simulation and linear estimation of multidimensional fields (Ababou and Wood, unpublished).

Let N be the total number of output "data," including both actual data and estimated data ($N = N_D + N_E$). The state-space formulation allows us to write the posterior estimate of $y(\mathbf{x})$ in the form of an N-vector, $\mathbf{y}^*$, possessing the unbiasedness property *a priori*. For simplicity, we give the relevant equations in the absence of measurement noise. The state-space vector of estimated values, including also the known measurements at "data points," is given by:

$$\mathbf{y}^* = \mathbf{m} + \mathbf{G} \cdot (\mathbf{y_D} - \mathbf{H} \cdot \mathbf{m}) \tag{3}$$

where $\mathbf{m}$ represents the mean (N-vector), $\mathbf{y_D}$ the data values (N_D-vector), $\mathbf{H}$ the measurement matrix (an $N_D \times N$ rectangular matrix such that $\mathbf{y_D} = \mathbf{H} \cdot \mathbf{y}$), and $\mathbf{G}$ is the so-called gain matrix (an $N \times N_D$ rectangular matrix). Note that the "no bias" condition is automatically satisfied, since $\langle \mathbf{y}^* \rangle = \mathbf{m}$ by construction. The gain matrix itself satisfies the matrix equation:

$$\mathbf{G} \cdot (\mathbf{H} \cdot \mathbf{P} \cdot \mathbf{H}^T) = \mathbf{P} \cdot \mathbf{H}^T \tag{4}$$

where $\mathbf{P}$ represents the full $N \times N$ matrix of covariances. Note that $\mathbf{P}$ includes as a sub-matrix the covariance matrix $\mathbf{C}$, which is the $N_D \times N_D$ matrix of covariances among data points. The computational difficulty resides in the solution of Eq. (4).

The system represented by Eq. (4) can be decomposed by ordering the states in vector $\mathbf{y}$ such that all estimation points (E) come first, and all data points (D) come last. Using the formalism of block-matrix algebra, we let:

$$\mathbf{y} = \begin{bmatrix} \mathbf{y_E} \\ \mathbf{y_D} \end{bmatrix}; \mathbf{P} = \begin{bmatrix} \mathbf{P_{EE}} & \mathbf{P_{ED}} \\ \mathbf{P_{DE}} & \mathbf{P_{DD}} \end{bmatrix}; \mathbf{G} = \begin{bmatrix} \mathbf{G_{ED}} \\ \mathbf{G_{DD}} \end{bmatrix} \tag{5}$$

where subscripts E and D indicate sub-vectors or sub-matrices relative to estimation points and data points, respectively. In particular, $\mathbf{P_{DD}}$ represents the data-data covariance matrix, that is: $\mathbf{P_{DD}} = \mathbf{C}$. After some manipulations, it can be shown that Eq. (4) leads to the sub-system:

$$\mathbf{P_{DD}} \cdot \mathbf{G_{DE}} = \mathbf{P_{DE}} \tag{6}$$

where both $\mathbf{G_{DE}}$ and $\mathbf{P_{DE}}$ are $N_D \times N_E$ rectangular matrices. Note that $\mathbf{G}_{DE} = \mathbf{G}_{ED}^T$; $\mathbf{G}_{DE}$ is the rectangular gain matrix, which represents the relative reduction of variance, at estimation points (E), due to information originating from measurements at data points (D).

Now, denote by $\mathbf{g}$ any particular column-vector of the rectangular gain sub-matrix $\mathbf{G_{DE}}$, and by $\mathbf{p}$ the corresponding column-vector of the rectangular covariance sub-matrix $\mathbf{P_{DE}}$. That is, $\mathbf{g}$ and $\mathbf{p}$ represent the jth column-vector of each matrix ($1 \leq j \leq N_E$). Computing the gain sub-matrix amounts to solving N_E linear systems of the form:

$$\mathbf{C} \cdot \mathbf{g} = \mathbf{p} \tag{7}$$

where $\mathbf{C}$ represents the square covariance sub-matrix $\mathbf{P_{DD}}$. Furthermore, the posterior covariance matrix among estimated values, $\mathbf{P}_{EE}^*$, can be computed directly by:

$$\mathbf{P}_{EE}^* = \mathbf{P_{EE}} - \mathbf{G_{ED}} \cdot \mathbf{P_{DE}} \tag{8}$$

where $\mathbf{G_{ED}} = \mathbf{G}_{DE}^T$. The diagonal of the posterior covariance matrix, $\mathbf{P}_{EE}^*$, yields the variances of estimation errors.

In brief, Eq. (3), (7), and (8), give the complete solution of the kriging system, provided solution of N_E linear systems all involving the same covariance matrix, $\mathbf{C}$, of size $N_D \times N_D$. The accuracy of inversion of $\mathbf{C}$ by direct elimination, direct factorization, or iterative methods, depends on its spectral condition number, which depends itself on the discrete eigenvalue spectrum of the matrix, as indicated by Eq. (1). As explained earlier, the accuracy of inversion of conditioning and kriging systems based on the state-space formulation (without Lagrange multipliers) is always better or, at worse, equal to that based on the classical formulation (with Lagrange multipliers). The state-space formulation is adopted here for this reason, among others.

COVARIANCES AND SPECTRAL DENSITIES

Covariance Functions and Spectral Density Functions

Figure 1 depicts different types of one-dimensional covariance functions which will be investigated later. Multidimensional counterparts to these covariance functions can also be defined, although care must be taken to satisfy certain admissibility conditions (Vanmarcke, 1983; Christakos, 1984). For example, the ellipsoidal-anisotropic Gaussian covariance function is defined for any number of spatial dimension by:

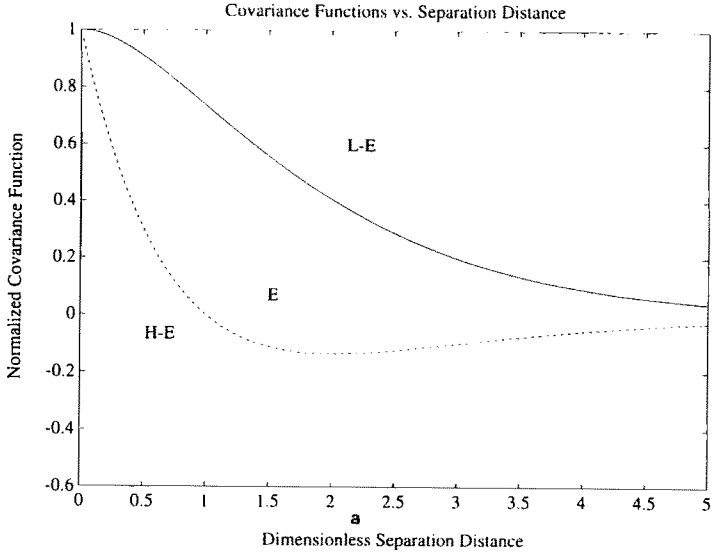
$$\mathbf{C}(\xi) = \sigma^2 \prod_{m=1}^{m=D} \exp\left\{-\frac{1}{2}\left(\frac{\xi_m}{\ell_m}\right)^2\right\} \tag{9}$$

where ξ is the separation vector, D is the dimension of space ($D = 1, 2$, or 3), and ℓ_m is a characteristic fluctuation scale along direction x_m. An alternative set of length scales, the integral correlation scales, can be obtained by integrating the covariance function to infinity and dividing by the valance. Thus, the one-sided integral correlation scale λ_m is related to the above-defined fluctuation scales by the relation:

$$\lambda_m = \sqrt{\frac{\pi}{2}}\,\ell_m \simeq 1.25\,\ell_m \tag{10}$$

However, the integral scales can vanish for other types of covariance functions (hole-functions below). Therefore, we prefer to use the fluctuation scales (ℓ_m), rather than the integral correlation scales (λ_m), as characteristic length scales of heterogeneity.

The spectral density function can be obtained from the Wiener-Khintchine theorem (Yaglom, 1962), by computing the Fourier transform of the covariance function. In the case of the Gaussian covariance, the spectral density is also Gaussian, that is



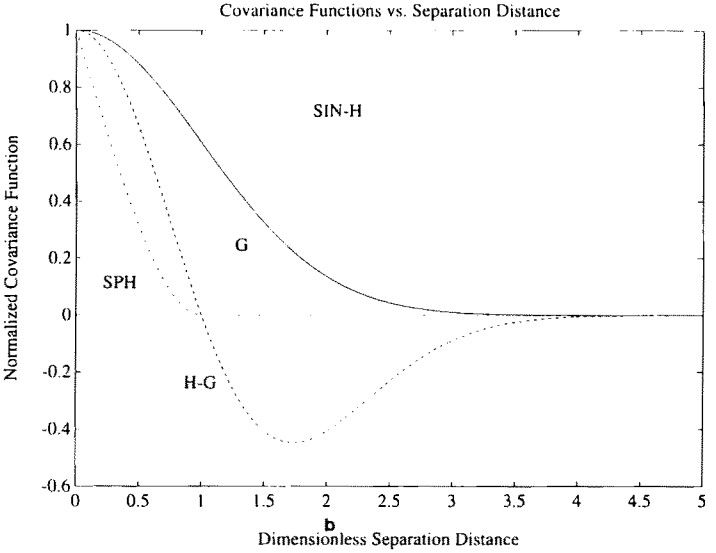


Fig. 1. One-dimensional covariance functions considered in this study. (a) Solid: linear-exponential (Markovian); dot: exponential; dash: hole-exponential. (b) Solid: Gaussian; dash: hole-Gaussian; dot: hole-sinusoidal; dash-dot: spherical.

$$\mathbf{S(k)} = \sigma^2 \prod_{m=1}^{m=D} \left(\frac{\ell_m}{\sqrt{2\pi}} \right) \exp\left\{ -\frac{1}{2} (k_m \ell_m)^2 \right\} \tag{11}$$

where **k** represents the wave-vector. Note in particular that both $C(\xi)$ and $S(\mathbf{k})$ must be even functions, that is $C(-\xi) = C(\xi)$ and $S(-\mathbf{k}) = S(\mathbf{k})$, and this regardless of the particular form of the covariance function at hand. See for instance Vanmarcke (1983) for more details. The one-dimensional Gaussian covariance is obtained by setting $D = 1$ in the above equations.

The one-dimensional covariance functions of Fig. 1 were plotted as functions of the dimensionless lag ξ/ℓ. Selected covariances are also given analytically in Table 1, along with the associated spectral densities. Both were normalized for a unit-variance field and a unit fluctuation scale. To obtain dimensional expressions, let $\xi \leftarrow \xi/\ell$, $k \leftarrow k\ell$, $C \leftarrow \sigma^2 C$, and $S \leftarrow \sigma^2 \ell S$ in Table 1. Important differences exist among the seven models. The hole-gaussian and hole-exponential covariances exhibit negative correlation at intermediate lags $\xi \approx \ell$, while the other functions exhibit positive correlations for all lags. The hole-sinusoidal exhibits negative correlation at lags $\xi > \pi$ with the maximum negative correlation at $\xi = 3\pi/2$. Some functions decay quickly to zero at large lags $\xi >> \ell$ (Gaussian and hole-Gaussian). Other functions decay quickly at small lags $\xi << \ell$ (exponential, hole-exponential, and spherical). The linear-exponential covariance decays slowly both at small lags and large lags. These

Table 1. One-Dimensional Covariance Functions $C(\xi)$ and Spectral Density Functions $S(k)$ (Dimensionless Formulation)

Covariance model	Covariance function	Spectral density function
1. Hole-exponential	$(1 - \lvert\xi\rvert) \exp\{-\lvert\xi\rvert\}$	$\frac{2}{\pi} \frac{k^2}{(1 + k^2)^2}$
2. Exponential	$\exp\{-\lvert\xi\rvert\}$	$\frac{1}{\pi} \frac{1}{1 + k^2}$
3. Linear-exponential	$(1 + \lvert\xi\rvert) \exp\{-\lvert\xi\rvert\}$	$\frac{2}{\pi} \frac{1}{(1 + k^2)^2}$
4. Hole-Gaussian	$(1 - \lvert\xi^2\rvert) \exp\left\{-\frac{1}{2}\lvert\xi^2\rvert\right\}$	$\frac{1}{\sqrt{2\pi}} k^2 \exp\left\{-\frac{1}{2} k^2\right\}$
5. Gaussian	$\exp\left\{-\frac{1}{2}\lvert\xi^2\rvert\right\}$	$\frac{1}{\sqrt{2\pi}} \exp\left\{-\frac{1}{2} k^2\right\}$
6. Spherical	$1 - (3/2)\xi + (1/2)\xi^3$ for $0 < \xi < 1$	$\frac{3}{\pi k^2}\left\{\frac{1}{2} + \frac{(1 - \cos k)}{k^2} - \frac{\sin k}{k}\right\}$
	0 otherwise	

differences suggest that similar differences may exist in the eigenvalue spectra and condition numbers of the corresponding covariance matrices. The exponential covariance corresponds to a first-order Markovian time-process, and the linear-exponential corresponds to a second-order Markovian process in one-dimensional space (see for instance Vanmarcke, 1983). The exponential and Gaussian models, in particular, have been frequently used in geostatistics and hydrology. There exists, however, a number of other covariance models which have been extensively used in geostatistics. Examples are the triangular model and the cubic model which is parabolic at the origin and thus more robust than the Gaussian. Furthermore, the covariance structure is often corrupted, in practice, by the nugget effect or small-scale noise. The effect of adding a nugget to the above covariance models will be examined briefly in terms of conditioning in Appendix B.

Covariance Matrices

Covariance matrices can be viewed as discretized versions of covariance functions. Given an arbitrary configuration of points $\{x_1, x_2, \cdots, x_N\}$, the associated covariance matrix $\mathbf{C} = [c_{ij}]$ is of the form

$$c_{ij} = c_{j-i} = c_{i-j} = \mathbf{C}(|x_j - x_i|) \tag{12}$$

Note that this matrix is always symmetric and positive-definite, owing to the general properties satisfied by any admissible covariance function. Furthermore, because of the assumed stationarity, we have $C(x, y) = C(y - x) = C(\xi)$. The covariance matrix is therefore a Toeplitz matrix, having constant elements along each of its diagonals (see Golub and Van Loan 1989; Press et al., 1986).

For simplicity, we will focus mainly on the case of a regular grid in one-dimensional space (or time), unless stated otherwise. The "grid" describes the configuration of data points in the case of kriging and conditional random field simulation by the superposition method. This is the case we use for illustration. Alternatively, however, the "grid" would represent all simulation points in the case of random field simulation by the multi-variate Gaussian method. These methods were described earlier.

Two grid-related quantities of special interest are the size of the mesh or sampling distance (Δx), and the number of grid points or sampling points (N). A third quantity of interest, related to the previous ones, is the size of the domain, $L = (N - 1)\Delta x$. Without loss of generality, we consider the case of a unit variance random field such that $C(0) = 1$, and we normalize all length scales by the fluctuation scale, ℓ. That is, Δx and L will be expressed from now on in terms of fluctuation scale units.

With these simplifications, the Gaussian covariance matrix takes the form:

$$c_{ij} = c_{j-i} = \exp\{-\tfrac{1}{2}(j - i)^2 \Delta x^2\} \tag{13}$$

The covariance matrices arising from other covariances models listed in Table 1 can be expressed in a similar fashion. Note the remarkably simple form taken by the corresponding spectral density functions (listed in Table 1). As far as we know, the *discrete* eigenvalue spectrum and the spectral condition number of these covariance matrices cannot generally be expressed in such simple form. Their behavior is investigated in the remaining sections.

EXACT AND NUMERICAL EVALUATION OF EIGENVALUE SPECTRA

The Eigenvalue Problem

Let $\mathbf{C}$ be an $N \times N$ covariance matrix, as defined by Eq. (12). For example, the Gaussian covariance matrix was given in Eq. (13). By definition, the eigenvalues (λ) of matrix $\mathbf{C}$ are such that there exist eigenvectors ($\mathbf{y} \neq 0$) satisfying the system:

$$(\mathbf{C} - \lambda \mathbf{I}) \cdot \mathbf{y} = 0 \tag{14}$$

The necessary and sufficient condition for a non-zero solution to (14) is that the determinant of $\mathbf{C} - \lambda \mathbf{I}$ be null. This gives an Nth-order polynomial equation for the λs:

$$\mathbf{P}_N(\lambda) = \text{Det}\,(\mathbf{C} - \lambda \mathbf{I}) = 0 \tag{15}$$

where P_N is the characteristic polynomial of matrix $\mathbf{C}$. Since $\mathbf{C}$ is symmetric positive-definite, it is known that there are N real positive roots, i.e., N real positive eigenvalues. However, the roots of $P_N(\lambda)$ are difficult to evaluate in closed-form except for small matrices.

To gain some insight on the "spectral behavior" of matrix $\mathbf{C}$, we have used a combination of analytical and numerical techniques. For small matrices ($N = 2, 3, 4$), we have developed closed form results for the eigenvalues and the associated condition number. For larger $N \times N$ matrices, say with N up to several hundreds, we used the Singular Value Decomposition (SVD) method as implemented in the MATLAB package (MATLAB, 1990; Press et al., 1986). The numerical results obtained with the SVD method were checked with the closed form expressions, at least for $N < 5$. All the covariance models listed in Table 1 were investigated in this fashion. These analyses, presented below, will be completed in the next section based on a new closed form expression for large matrices (continuum approximation).

Exact Closed Form Results for Small N

Exact expressions for the condition number of covariance matrices of size $N < 5$ are obtained by solving the characteristic polynomial Eq. (15). The closed form results, and their derivation, are briefly summarized in Appendix A.

The exact condition numbers given by Eq. (A1)-(A4) of Appendix A are plotted as functions of mesh size (Δx) in Figs. 2 and 3. The effect of increasing N from $N = 2$ to $N = 4$ is shown on Fig. 2 for the case of the Gaussian covariance function, while Fig. 3 depicts the influence of the shape of the covariance function for the case $N = 4$.

Figure 3 shows that the condition number decreases monotonically for monotonic covariance functions, and non-monotonically for hole-covariance functions. The condition number of the hole-exponential model is consistently lowest, and that of the Gaussian model is consistently highest, in the range of small sampling distance ($\Delta x \leq 0.60$). In the range of $0.6 < \Delta x \leq 0.75$ the spherical model surpasses the hole-exponential model. For all models, the condition number must go to infinity as $\Delta x \to 0$ while N is held fixed, since the covariance matrix becomes a singular matrix with all elements equal to one.

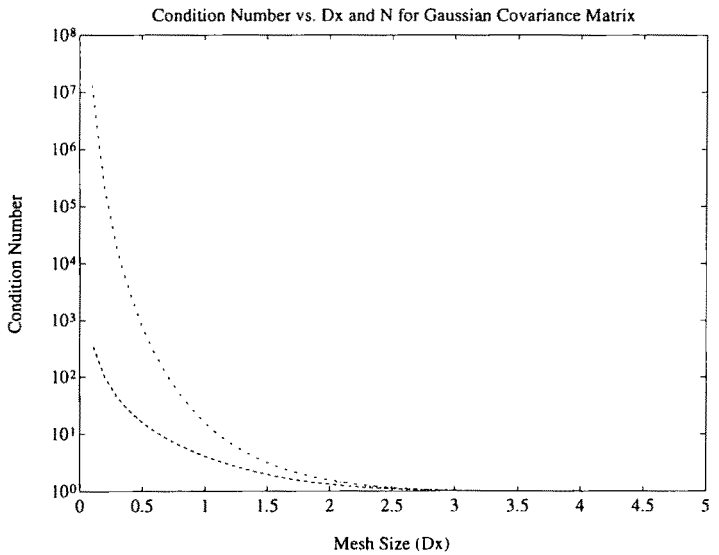


Fig. 2. Condition number κ vs. mesh size Δx for different values of N in the case of the Gaussian covariance (exact closed form results). Dash: $N = 2$; dot: $N = 3$; dash-dot: $N = 4$.

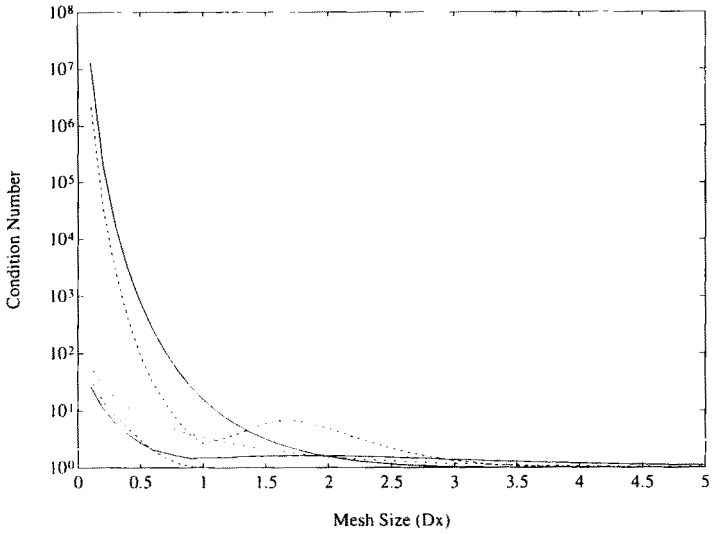


Fig. 3. Condition number κ vs. mesh size Δx for $N = 4$ and various covariance functions (exact closed form results). Solid: Gaussian; dash: hole-Gaussian; dot: linear-exponential; dash-dot: exponential; solid: hole-exponential; dash: spherical.

For the Gaussian and hole-Gaussian models, the condition number grows very quickly as Δx decreases.

The singular behavior of Gaussian-type covariances is due to the fact that the minimum eigenvalue remains very close to zero at small Δx. This is illustrated in Fig. 4 for the Gaussian model. In contrast, better-behaved models such as the hole-exponential have a minimum eigenvalue that increases rapidly as Δx increases away from zero: see Fig. 5, compared to Fig. 4. These important differences are directly related to the behavior of covariance functions at the origin. The Gaussian and hole-Gaussian covariances have zero slope at the origin, while the exponential and hole-exponential both have significant non-zero slopes at the origin (Fig. 1 and Table 1). The "flatness" of the covariance curves at the origin explains the particularly bad conditioning of the Gaussian models in the case of small or moderate sampling distance relative to fluctuation scale ($\Delta x \leq 0.75$ in dimensionless terms).

The behavior of a relatively strong periodic hole-effect is studied with the help of the hole-sinusoidal covariance model given by:

$$C(\xi) = \frac{\sin \xi}{\xi} \tag{16}$$

For this covariance model the relative amplitude of the hole effect is $\omega = 21.2\%$

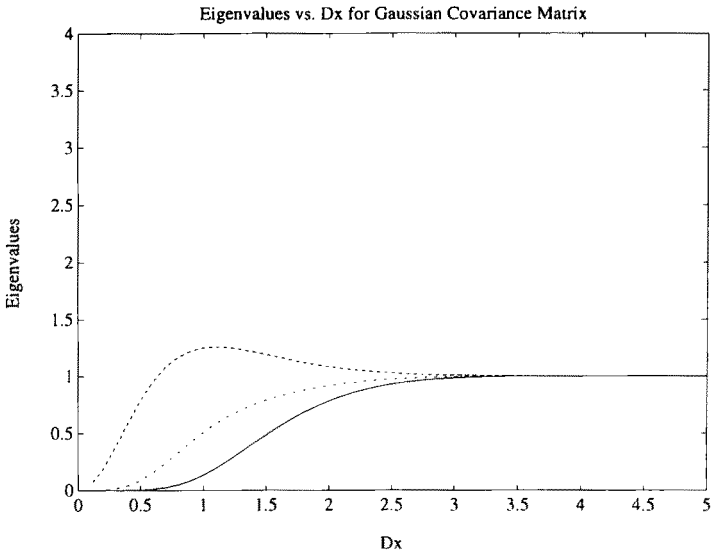


Fig. 4. Eigenvalues λ vs. mesh size Δx for $N = 4$ in the case of the Gaussian covariance (exact closed form results). There are four eigenvalues, and the condition number is the ratio of largest to smallest.

and is obtained at $\xi \approx 3\pi/2$. This model is positive definite, has a sill, and exhibits parabolic behavior at the origin. The non-monotonic behavior of the condition number is clearly shown in Fig. 6a. Similarly, the dampening periodic behavior of the eigenvalues is shown in Fig. 6b. The parabolic behavior of the model near the origin yields a high condition number in the order of 10^3, almost two orders of magnitude worse than the better behaved covariance models.

Numerical Results for Large N

We have further investigated the relation between condition number (κ) and mesh size (Δx) as the size of the domain (L) remains fixed. Figure 7 shows a comparison of curves $\kappa(\Delta x)$ for the covariance models of Fig. 1 and Table 1. Here, the spectral condition numbers were computed using the SVD algorithm as implemented in MATLAB. The dimensionless domain size is $L = 10$ and N varies with $1 + L/\Delta x$, so that $N = 101$ for $\Delta x = 0.1$. Overall, the results of Fig. 7 for fixed domain size suggest the following ranking of covariance models, from best-conditioned to worst-conditioned: (1) hole-exponential, (2) exponential, (3) linear-exponential, (4) hole-Gaussian, (5) Gaussian. This ranking gives more importance to the region of ill-conditioning, corresponding here to $\Delta x \leq 0.75$ and $N \geq 14$.

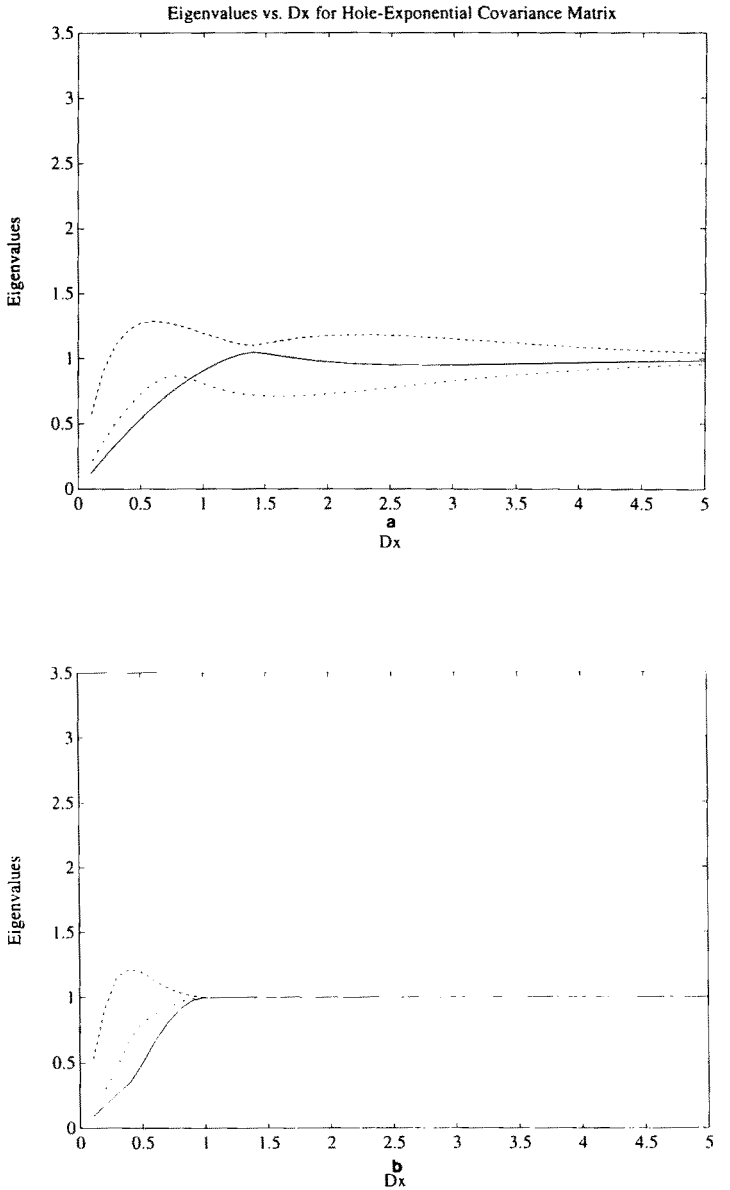


Fig. 5. Eigenvalues λ vs. mesh size Δx for $N = 4$ in the case of (a) the hole-exponential; and (b) the spherical covariance (exact closed form results). There are four eigenvalues, and the condition number is the ratio of largest to smallest.

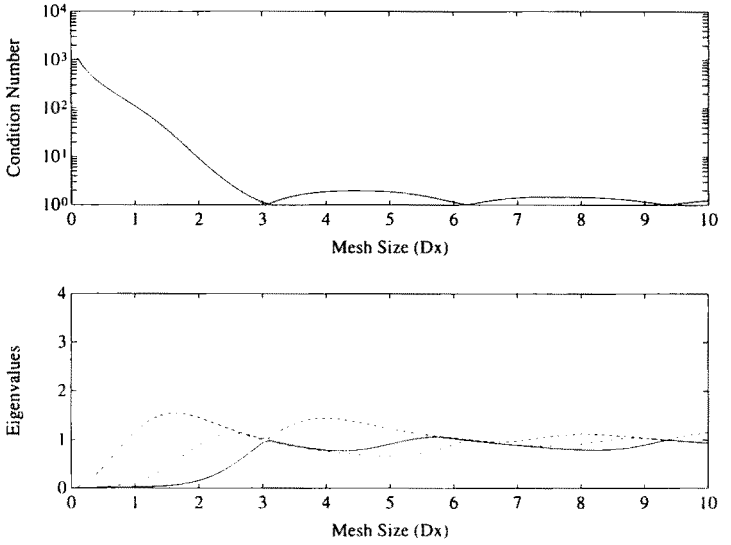


Fig. 6. Behavior of the hole-sinusoidal covariance model. (a) Condition number; and (b) Eigenvalues λ vs. mesh size Δx for $N = 4$ (exact closed form results).

Furthermore, numerical computations of $\kappa(\Delta x)$ were also carried out for certain irregular grids obtained by perturbing regular grids at only one location, i.e., by adding or deleting just one sampling point. For values of $\Delta x > 0.75$, the results are totally unaffected for all types of covariance functions. This type of "perturbation," however, lowers the condition number by 7 to 8 orders of magnitude at values of $\Delta x \leq 0.5$ for the Gaussian and hole-Gaussian covariance functions. To illustrate this phenomenon, we present in Fig. 8—analogous to Fig. 7—the condition number of a perturbed grid. Here, Δx represents the mesh size of the regular grid. The perturbed grid was obtained by adding one sampling point at a distance $\Delta x/10$ to the right of the central node. These results indicate the sensitivity of the Gaussian models to small differences in sampling schemes. However, other results were not so conclusive, and we have not pursued further investigation of such grid perturbation schemes.

Returning to the case of regular grids, we now investigate the relation between condition number (κ) and number of points (N) for a fixed mesh size (Δx), using again the SVD numerical approach. The resulting curves $\kappa(N)$ are depicted in Figs. 9–13. Each figure corresponds to a particular covariance model, with different curves $\kappa(N)$ corresponding to different values of mesh size.

In particular, Figs. 9 and 10 correspond to the "worst" and "best" covariance models identified above, i.e., the Gaussian and hole-exponential models

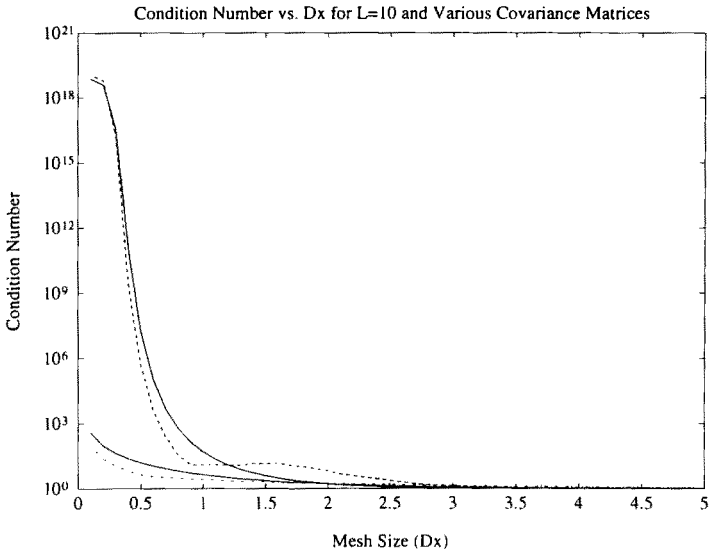


Fig. 7. Condition number κ vs. mesh size Δx in the case of fixed domain size ($L = 10$) for various covariance functions. Solid: Gaussian; dash: hole-Gaussian; dot: linear-exponential (Markovian); solid: exponential; dash-dot: hole-exponential.

respectively. These figures indicate how the matrix condition degrades as more sampling points are added at regular intervals in order to extend the size of the sampled domain. For all models, the condition number appears to reach a finite asymptotic value κ_∞ as $N \to \infty$ while Δx is held fixed, i.e., in the limit of infinite sampling domain. This asymptotic condition number is itself a rapidly increasing function of sampling density (Δx^{-1}), particularly for the Gaussian models. Concerning the hole-covariance models (Figs. 10 and 13), the rate of increase of κ with N is highest for $\Delta x = 1$, a phenomenon not encountered with monotonic models. Since lag $\xi = 1$ corresponds to zero-crossing of the hole-covariance, taking $\Delta x = 1$ yields a covariance matrix whose first off-diagonal elements are all null.

Based on the region of large (asymptotic) condition number, for $N \geq 50$ or so, the hole-exponential model is ranked best, followed by the exponential, the linear-exponential, the hole-Gaussian, and the Gaussian models, in that order. This ranking, which confirms that proposed earlier based on Fig. 7, has practical consequences. The condition number characterizes numerical stability. Using the hole-exponential instead of the Gaussian covariance function, one can achieve a relative improvement in numerical stability ranging from 2 to 14 orders of magnitude (compare Figs. 9 and 10). However, the striking differences ob-

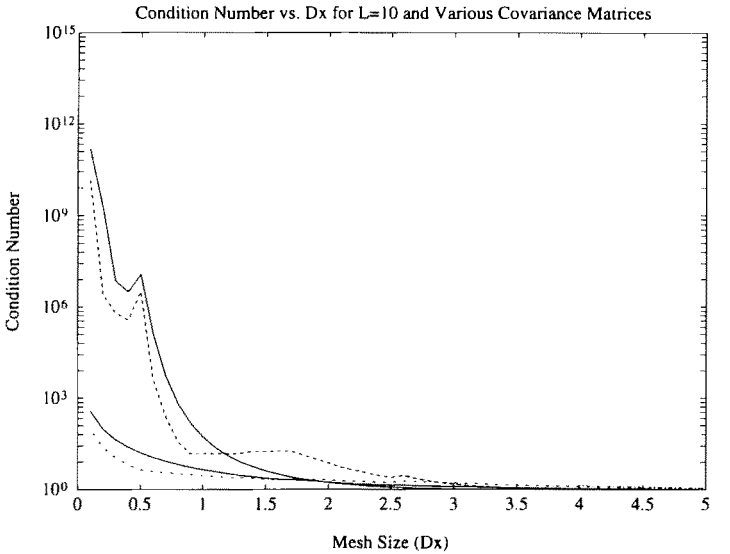


Fig. 8. Condition number κ vs. mesh size Δx for a fixed domain size and a perturbed grid (non-uniform sampling). Solid: Gaussian; dash: hole-Gaussian; dot: linear-exponential; solid: exponential; dash-dot: hole-exponential.

served between different covariance models are not fully understood at this stage. The behavior of $\kappa(N, \Delta x)$ must be related in some fashion to the spectral properties of the underlying covariance model. This question is investigated in the next section.

CONTINUUM FORMULATION OF SPECTRAL EIGENVALUE PROBLEM

In this section, we develop a new analysis in order to understand the influence of the covariance model on the condition number of the covariance matrix. We obtain a closed form functional relation between the condition number and the spectral density function, for given values of grid size and mesh size. This relation, although approximate, turns out to be qualitatively correct for a broad range of parameters, particularly in the case of monotonic covariance functions. Application of the new relation to some of the covariance models of Table 1 and Fig. 1 will enable us to better understand the results of the previous sections, and to examine possible extensions to multidimensional space.

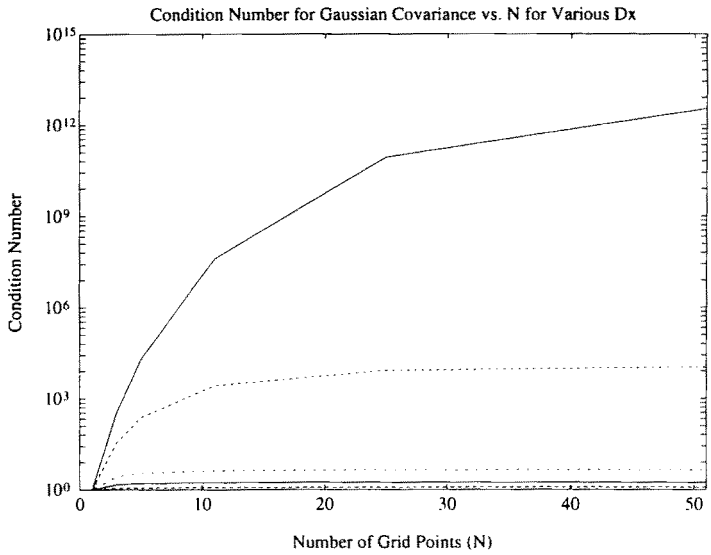


Fig. 9. Condition number κ vs. number of sampling points N, for different values of sampling distance Δx, for the Gaussian covariance. Solid: $\Delta x = 0.4$; dash: $\Delta x = 0.7$; dot: $\Delta x = 1.0$; dash-dot: $\Delta x = 1.5$; solid: $\Delta x = 2.0$; dash: $\Delta x = 3.5$.

Continuum Approximation of Eigenvalues and Condition Number

We start with the premise that the spectral density function, being the Fourier transform of the covariance function, constitutes in some sense a continuous-space version of the discrete eigenvalue spectrum of the covariance matrix. This intuitive idea can be made more precise by considering the limit form of the matrix eigenvalue problem (14) as $\Delta x \to 0$ and $L \to \infty$.

First note that, owing to the stationarity assumption, the matrix eigenvalue problem can be written in the form of a discrete convolution:

$$\sum_{\mathbf{j}=\mathbf{1}}^{\mathbf{j}=\mathbf{N}} c_{i-j}\mathbf{y}_j = \lambda \mathbf{y}_i (\mathbf{i} = 1, \cdots, \mathbf{N})$$

$$\mathbf{c} * \mathbf{y} = \lambda \mathbf{y} \tag{17}$$

In the second equation, the star operator "*" stands for discrete convolution, and $\mathbf{c}$ designates an infinite-length vector $[c_k]$, with $-\infty \le k \le +\infty$, such that the c_ks are equal to covariance values at lag k for $k \in \{-(N-1), \cdots, 0, \cdots, +(N-1)\}$, and are identically zero elsewhere.

Now, let $\Delta x \to 0$ and $L \to \infty$, implying also that $N \to \infty$, since $N = 1 + L/\Delta x$. Taking these limits in Eq. (17), writing $x_i = -L/2 + (i-1)\Delta x$

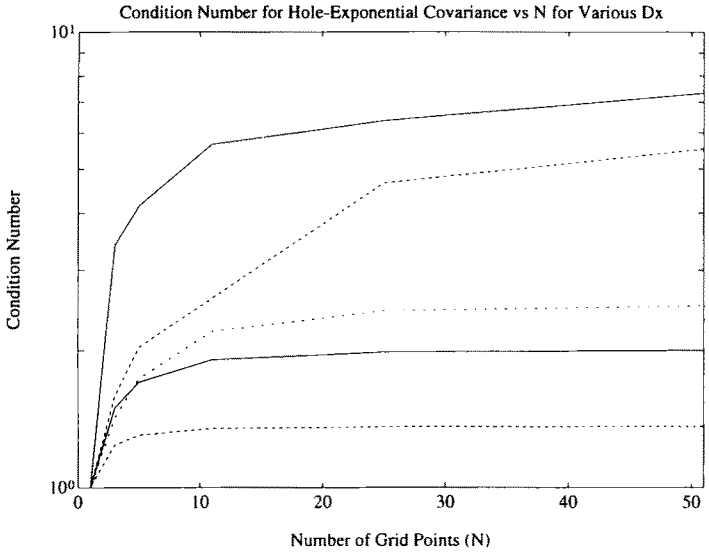


Fig. 10. Condition number κ vs. number of sampling points N, for different values of sampling distance Δx, for the hole-exponential covariance. Solid: $\Delta x = 0.4$; dash: $\Delta x = 0.7$; dot: $\Delta x = 1.0$; dash-dot: $\Delta x = 1.5$; solid: $\Delta x = 2.0$; dash: $\Delta x = 3.5$.

and $x_j = -L/2 + (j - 1)\Delta x$, and multiplying both sides by Δx, yields a convolution integral of the form:

$$\int_{-\infty}^{+\infty} \mathbf{C}(\mathbf{x} - \mathbf{x}') \cdot \mathbf{y}(\mathbf{x}')\, d\mathbf{x}' = \Lambda \cdot \mathbf{y}(\mathbf{x})$$

$$\Lambda = \lim_{\substack{\Delta x \to 0 \\ L \to \infty}} (\lambda \Delta \mathbf{x}) \tag{18}$$

where Λ, as defined above, represents a continuous eigenvalue spectrum for the continuum version of the matrix eigenvalue problem. The limit of $(\lambda \Delta x)$ does not necessarily vanish since λ depends on Δx. The function $C(\xi)$ represents as before the covariance function [e.g., Eq. (9)]. Applying the usual Fourier convolution theorem to the convolution integral yields the identity:

$$\overline{\mathbf{C}}(\mathbf{k}) \cdot \overline{\mathbf{y}}(\mathbf{k}) = \Lambda \cdot \overline{\mathbf{y}}(\mathbf{k}) \tag{19}$$

where k stands for wavenumber (or frequency), and the overbar sign designates Fourier transforms from x-space to k-space. By the Wiener–Khintchine theorem (Yaglom, 1962), the Fourier transform of the covariance function $C(\xi)$ is the spectral density $S(k)$ of the random field. This yields the key relation:

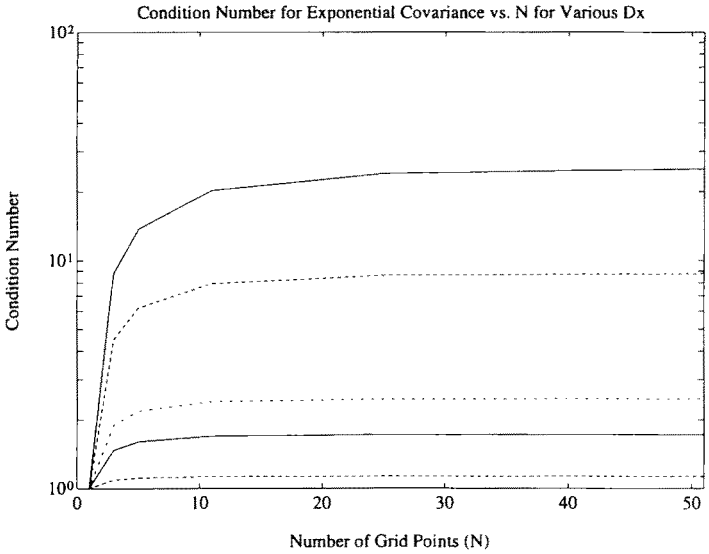


Fig. 11. Condition number κ vs. number of sampling points N, for different values of sampling distance Δx, for the exponential covariance. Solid: $\Delta x = 0.4$; dash: $\Delta x = 0.7$; dot: $\Delta x = 1.0$; dash-dot: $\Delta x = 1.5$; solid: $\Delta x = 2.0$; dash: $\Delta x = 3.5$.

$$\Lambda(\mathbf{k}) = \mathbf{S}(\mathbf{k}) \tag{20}$$

In order to obtain a closed form expression for the *discrete* eigenvalue spectrum, and to capture the effects of discretization and finite domain size, we now discretize and cut off the spectral density function as indicated in Fig. 14. To account for finite domain size, we discretize the wavenumber space into finite steps Δk, and we apply the low wavenumber cut-off:

$$\mathbf{k}_{\min} = \Delta \mathbf{k} = \frac{\pi}{\mathbf{L}} \tag{21}$$

This choice is based on the fact that the largest wave capable of representing fluctuations on the scale L has a wavelength equal to $2L$. In addition, to account for the discrete character of space, we apply a high wavenumber cut-off:

$$\mathbf{k}_{\max} = \frac{\pi}{\Delta \mathbf{x}} \tag{22}$$

The latter choice is based on the fact that the smallest wavelength capable of sensing fluctuations on the scale Δx is equal to $2\Delta x$ (sampling theorem). Thus, $k_{\max}$ represents the Nyquist frequency. If W is wavelength, we have $k = 2\pi / W$.

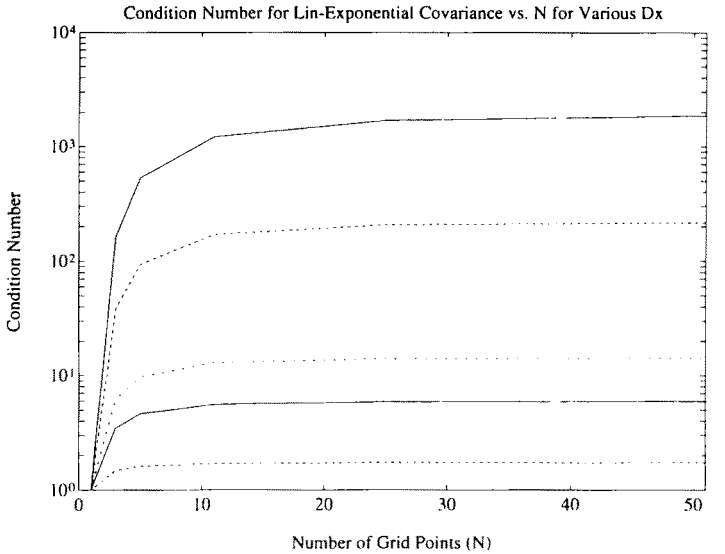


Fig. 12. Condition number κ vs. number of sampling points N, for different values of sampling distance Δx, for the linear-exponential covariance. Solid: $\Delta x = 0.4$; dash: $\Delta x = 0.7$; dot: $\Delta x = 1.0$; dash-dot: $\Delta x = 1.5$; solid: $\Delta x = 2.0$; dash: $\Delta x = 3.5$.

Based on Eqs. (20)–(22), the discrete eigenvalues spectrum λ_n is therefore approximated by:

$$\lambda_{\mathbf{n}} \approx \frac{1}{\Delta \mathbf{x}} \mathbf{S}(\mathbf{k_n}) \tag{23}$$

In this equation, the k_ns should take values $n\pi/[(N-1)\Delta x]$, with $1 \leq n \leq N-1$ in order to have $k_{\min} \leq k_n \leq k_{\max}$. However, this leads to the loss of one eigenvalue. To re-establish the correct number of eigenvalues, we use the approximation $L = (N-1)\Delta x \approx N\Delta x$. This yields:

$$\mathbf{k_n} = \frac{\mathbf{n}}{\mathbf{N}} \frac{\pi}{\Delta \mathbf{x}}; \mathbf{n} = 1, \cdots, \mathbf{N} \tag{24}$$

Equations (23)–(24), taken together, define the approximate eigenvalue spectrum, which should become exact in the limit of small mesh size and infinite domain ($\Delta x \to 0$ and $L \to \infty$). The condition number corresponding to this approximate eigenvalue spectrum is given by:

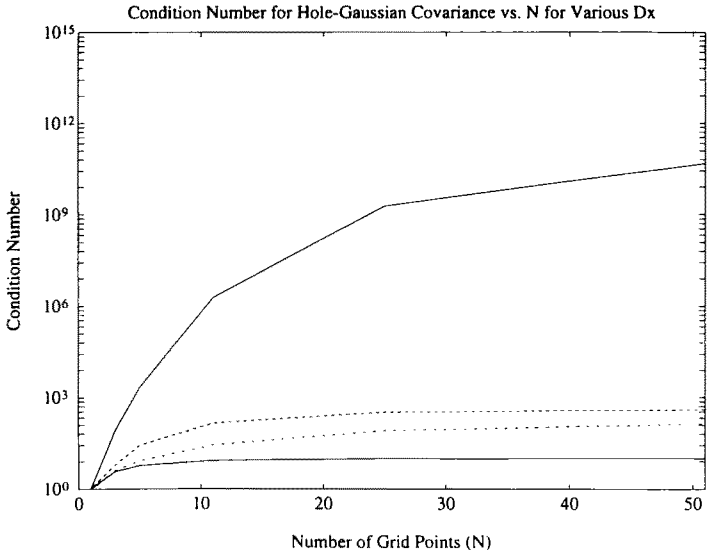


Fig. 13. Condition number κ vs. number of sampling points N, for different values of sampling distance Δx, for the hole-Gaussian covariance. Solid: $\Delta x = 0.4$; dash: $\Delta x = 0.7$; dot: $\Delta x = 1.0$; dash-dot: $\Delta x = 1.5$; solid: $\Delta x = 2.0$; dash: $\Delta x = 3.5$.

$$\kappa = \frac{\underset{n=1\cdots N}{\text{Max}} \; \mathbf{S}(\mathbf{k}_n)}{\underset{n=1\cdots N}{\text{Min}} \; \mathbf{S}(\mathbf{k}_n)} \tag{25}$$

Note that, due to the positivity of the spectral density function, the absolute values are not needed in Eq. (25).

We expect Eq. (25) to be more accurate for relatively large domains, $L > O(1)$, and small sampling distances, $\Delta x < O(1)$. In dimensional terms, the first condition means that the sampled domain must be larger than the fluctuation length scale; the second condition means that the sampling density, expressed as number of samples per fluctuation length scale, must be larger than one.

Application of Continuum Approximation to Monotonic Covariance Models

The Gaussian, Exponential, and linear-exponential models all have monotonic covariance and spectral density functions, as can be seen from Fig. 1 and Table 1. Since the spectral density is a monotonically decreasing function of wavenumber, Eq. (25) leads to a particularly simple expression:

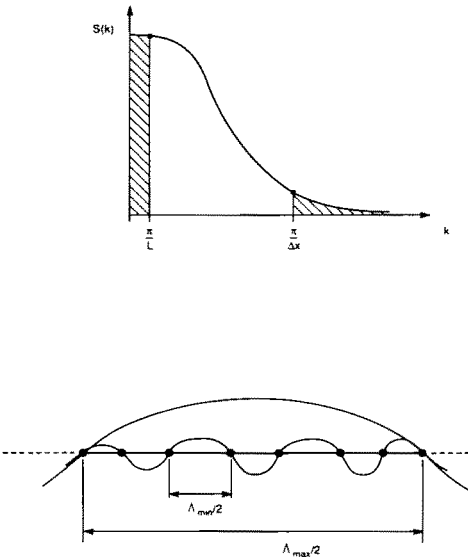


Fig. 14. Low and high wavenumber truncations of spectral density function to account for finite domain and mesh size in the continuum approach. Top: spectral cut-offs in Fourier space. Bottom: largest and smallest half-wavelengths in physical space.

$$\kappa \approx \frac{\mathbf{S}(\mathbf{k}_{\min})}{\mathbf{S}(\mathbf{k}_{\max})} \approx \frac{\mathbf{S}\left(\frac{\pi}{\mathbf{L}}\right)}{\mathbf{S}\left(\frac{\pi}{\Delta \mathbf{x}}\right)} \approx \frac{\mathbf{S}\left(\frac{1}{\mathbf{N}}\frac{\pi}{\Delta \mathbf{x}}\right)}{\mathbf{S}\left(\frac{\pi}{\Delta \mathbf{x}}\right)} \tag{26}$$

Based on Eq. (26), it is easily seen that the correct limits $\kappa \to 1$ are obtained as $\Delta \mathbf{x} \to \infty$ and as $N \to 1$, and this in spite of the continuum approximation which normally requires small Δx, large L, and therefore large N. We give in Table 2 the specific expressions obtained for each monotonic covariance model as a function of Δx and L, assuming $L \approx N\Delta x$. There are two limit cases of practical interest:

- Limit of infinite sampling domain (L) for fixed mesh size (Δx):

$$\kappa_\infty(\Delta \mathbf{x}) = \lim_{L \to \infty} \kappa(\Delta \mathbf{x}, \mathbf{L}) \tag{27}$$

- Limit of infinite sampling density (Δx^{-1}) for fixed domain size (L):

$$\kappa_\infty(\mathbf{L}) = \lim_{\Delta x \to 0} \kappa(\Delta \mathbf{x}, \mathbf{L}) \tag{28}$$

Table 2. Condition Number $\kappa(L, \Delta x)$ for Monotonic Covariance Models (Continuum Approximation)

Name of covariance model	Condition number $\kappa(L, \Delta x)$
Linear-exponential (Markovian)	$\left[\dfrac{1 + \left(\dfrac{\pi}{\Delta x}\right)^2}{1 + \left(\dfrac{\pi}{L}\right)^2}\right]^2$
Exponential	$\dfrac{1 + \left(\dfrac{\pi}{\Delta x}\right)^2}{1 + \left(\dfrac{\pi}{L}\right)^2}$
Gaussian	$\exp\left\{\dfrac{1}{2}\left[\left(\dfrac{\pi}{\Delta x}\right)^2 - \left(\dfrac{\pi}{L}\right)^2\right]\right\}$

For all three models, the limit of infinite domain yields a finite, asymptotic condition number. To obtain this limit value $\kappa_\infty(\Delta x)$, let $L \to \infty$ in Table 2. Remarkably, we see that the condition number eventually reaches a maximum as the sampled domain size is increased by adding more and more sampling points at constant intervals. The limit value is itself an increasing function of sampling density (Δx^{-1}). However, all three models have infinite condition number, $\kappa_\infty(L) = \infty$, in the limit of infinite sampling density. That is, the condition number goes to infinity as more and more sampling points are added within a domain of fixed size L.

The precise form of $\kappa(\Delta x, L)$ in Table 2 clearly shows that the Gaussian model diverges much faster than the other models as sampling density increases while L is held fixed. This is in agreement with previous numerical results. In fact, comparing the analytical results of Table 2 to the numerical results of Fig. 7 for $\Delta x = 0.1$, $L = 10$, $N = 101$, excellent agreement is observed for all monotonic models ($\kappa \approx 10^3$ for the exponential model, $\kappa \approx 10^6$ for the linear-exponential model, and $\kappa \approx 10^{19}$ for the Gaussian model). In the case of the Gaussian covariance, analytical results are depicted in Fig. 15, to be compared to Fig. 9. Both figures show κ vs. N for different values of Δx. The analytical values of κ from Fig. 15 agree fairly well with the numerical results of Fig. 9, at least for sufficiently large domains ($L \approx N\Delta x >> 1$).

Case of Non-Monotonic (Hole-Covariance) Models

The behavior of non-monotonic covariance models is more complex. Indeed, the hole-covariances of Fig. 1 are associated with hole-spectral densities $S(k)$ having a maximum at the normalized wavenumber $k = 1$, as can be inferred

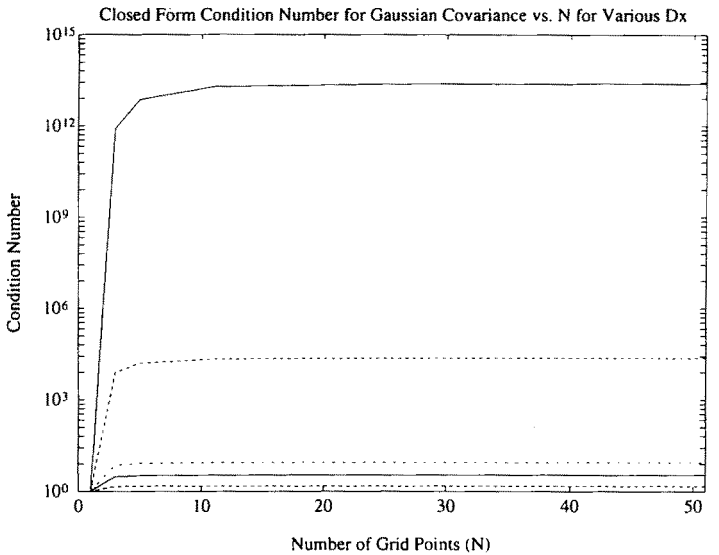


Fig. 15. Condition number κ vs. number of points N, for different values of sampling distance Δx, for the Gaussian covariance (continuum approximation results). Solid: $\Delta x = 0.4$; dash: $\Delta x = 0.7$; dash-dot: $\Delta x = 1.5$; solid: $\Delta x = 2.0$; dash: $\Delta x = 3.5$.

from Table 1. Complete analytical results can be obtained by utilizing Eq. (25) while taking into account the existence of the local maximum of $S(k)$. The behavior of κ depends on whether $(\Delta x, L)$ are greater or smaller than π. Analysis of the hole-Gaussian model indicates that the solution is more sensitive to the continuum approximation than in the case of monotonic models. For instance, the incorrect limit $\kappa \to N$ is obtained as $\Delta x \to \infty$ (the correct limit is $\kappa \to 1$). For fixed sampling distance $\Delta x \leq \pi$, the limit of infinite domain size is $\kappa_\infty(\Delta x) = \infty$, which seems correct. However, the question of the range of validity of the continuum approximation for hole-covariance models remains open.

Case of Multidimensional Sampling Schemes

The results obtained with the continuum approximation can be extended, with some caution, to the case of two- and three-dimensional grids with uniform rectangular mesh. Care must be taken to correctly translate the domain and mesh geometry from physical space to Fourier space. The most natural way to accomplish this may be to assume a standard node ordering in both physical space and Fourier space (XYZ sweeps), and to apply Fourier space cut-offs to each wavenumber component separately.

The case of the multidimensional anisotropic Gaussian covariance is particularly straightforward, owing to the fact that both the covariance and spectral density are spatially separable functions [see Eqs. (9)–(11)]. Thus, consider the two-dimensional case with anisotropic fluctuation scales (ℓ_x, ℓ_y). Normalizing all length scales with ℓ_x in the x-direction, and ℓ_y in the y-direction, yields isotropic covariance and spectral density functions. With this scaling, assume that the sampled domain has dimensionless size $L_x \times L_y$, dimensionless mesh size $\Delta x \times \Delta y$, and number of grid points $N_x \times N_y$. The domain boundaries and the sampling mesh are both rectangular, but not necessarily with the same aspect ratio. Replicating the one-dimensional continuum approximation developed earlier, we obtain the following two-dimensional result (Gaussian model):

$$\kappa = \exp\left\{\frac{1}{2}\left[\left(\frac{\pi}{\Delta \mathbf{x}}\right)^2 + \left(\frac{\pi}{\Delta \mathbf{y}}\right)^2 - \left(\frac{\pi}{\mathbf{L_x}}\right)^2 - \left(\frac{\pi}{\mathbf{L_Y}}\right)^2\right]\right\} \tag{29}$$

Various limits of infinite sampling density and infinite domain can be easily inferred from Eq. (29). Thus, increasing the dimensionless sampling density along just one direction leads to a singular problem with infinite condition number. On the other hand, increasing the dimensionless size of the domain by adding more points at regular intervals along either one direction, or both directions simultaneously, leads to a finite asymptotic condition number. These results parallel those obtained earlier for the one-dimensional Gaussian model.

SUMMARY AND CONCLUSIONS

In the preceding sections, we showed that the solution of ordinary kriging equations for purposes of geostatistical estimation can be reduced to the repeated inversion of a stationary covariance matrix. We also pointed out that kriging was required in certain types of conditional random field simulators, and that inversion of the covariance matrix was also required in multivariate Gaussian generators. Based on these observations, we set out to study the discrete eigenvalue spectrum and the spectral condition number of any such covariance matrix using three different methods: (1) exact analytical solution of the characteristic polynomial equation (limited to small matrices with only few grid points), (2) numerical singular value decomposition of the matrix (quasi-exact and tractable for large matrices), and (3) new analytical solution based on a continuum approximation (tractable for all grid sizes but approximate). Our conclusions can be summarized as follows.

The condition number diverges as sampling density increases while domain size is fixed. It also increases with domain size as the sampling interval remains fixed, but, unlike the previous case, a finite asymptotic value is reached. The asymptotic condition number is a rapidly increasing function of the sampling

density, especially for covariance models which are "flat" at the origin (e.g., Gaussian). For non-monotonic covariance models, the rate of increase of the condition number with the number of sampling points is highest at a sampling distance of one characteristic fluctuation length. This sampling distance value corresponds to zero-crossing of the hole-covariance functions. Here, we implicitly express all length scales, such as sampling distance and domain size, in fluctuation scales units. Thus, the "sampling density" is the number of samples per fluctuation scale.

We ranked the covariance models listed in Table 1 as follows, from "best" to "worse" conditioned: (1) hole-exponential and spherical, (2) exponential, (3) linear-exponential, (4) hole-Gaussian, and (5) Gaussian. This ranking is based mainly upon condition number comparisons in the region of ill-conditioning, that is, at small-to-moderate sampling distances.

An alternative closed form solution was developed for the spectral condition number, based on a continuum approximation. There is good agreement with the numerical results for the case of monotonic covariance functions. The analytically predicted condition number is correct for sufficiently large domains. In the limit of infinite domain, the existence of a finite asymptotic condition number is confirmed. The divergent limit of infinite sampling density seems also in agreement with numerical results, and there is excellent quantitative agreement with numerical results at large sampling densities.

In multidimensional space, based on the continuum approximation results, we expect that the behavior of the condition number will be similar, e.g., when taking limits of infinite sampling density and/or infinite domain size along one or more directions. In the case of a two-dimensional ellipsoidal-anisotropic Gaussian covariance, increasing the sampling density along just one direction, relative to the fluctuation scale in that direction, leads to divergence.

The effect of small-scale variability (nugget) was studied theoretically and numerically in Appendix B. As it could be intuitively expected, adding a small-scale variance to the diagonal of the covariance matrix improves its condition number. This was verified numerically for the exponential covariance model which improved its condition number by up to a factor of 20 near the region of ill-conditioning.

The case of two and three spatial dimensions deserves further investigation, as well as the possible effects of irregular sampling. For regular sampling grids, the continuum-based approximation can offer a simple and easy way to quickly evaluate the condition number of the covariance matrix for alternative covariance models and sampling densities. Note that the condition number always degrades as sampling density increases. This suggests that the magnitude of the condition number is not only a measure of numerical instability, but also a heuristic measure of oversampling.

ACKNOWLEDGMENTS

This study was funded by the U.S. Nuclear Regulatory Commission under the Research Projects "Stochastic Analysis of Large-Scale Flow and Transport in Unsaturated Fractured Rock" (FIN B6664) and "Performance Assessment" (FIN L1152); however, the views expressed in this paper are those of the authors. Comments and suggestions of three anonymous referees are greatly appreciated.

APPENDIX A. CONDITION NUMBER FOR $N < 5$: EXACT RESULTS IN CLOSED FORM

Exact expressions for the condition number of covariance matrices of size $N < 5$ are obtained by solving the characteristic polynomial Eq. (15). The solutions, in terms of eigenvalues and condition number, are briefly developed here.

For convenience, we use an auxiliary variable $\mu = 1 - \lambda$, and we define the associated polynomial $Q_N(\mu) = P_N(1 - \lambda)$. Recall that c_i stands for the covariance of lag $\xi = i\Delta x$, and that $c_0 = 1$ and $-1 \leq c_i \leq +1$ for a unit-variance field. The condition number (κ) measures the spread of eigenvalues in terms of absolute values, and can be obtained from Eq. (1) once the eigenvalue spectrum is obtained from Eq. (15).

This yields, for $N = 1, 2, 3, 4$, respectively:

$N = 1$:

$$\begin{aligned} \mathbf{Q}_1(\mu) &= 1 \\ \mu &= 0 \\ \kappa &= 1 \end{aligned} \tag{A1}$$

$N = 2$:

$$\begin{aligned} \mathbf{Q}_2(\mu) &= \mu^2 - \mathbf{c}_1^2 \\ \mu_{1,2} &= \pm \mathbf{c}_1 \\ \kappa &= \frac{1 + |\mathbf{c}_1|}{1 - |\mathbf{c}_1|} \end{aligned} \tag{A2}$$

$N = 3$:

$$\mathbf{Q}_3(\mu) = \mu^3 - (2\mathbf{c}_1^2 + \mathbf{c}_2^2)\mu + 2\mathbf{c}_1^2\mathbf{c}_2$$

$$\mu_1 = 2\mathbf{c}_1\sqrt{\frac{2 + \mathbf{r}^2}{3}}\cos\frac{\alpha}{3}$$

$$\mu_{2,3} = -2\mathbf{c}_1\sqrt{\frac{2 + \mathbf{r}^2}{3}}\cos\frac{\alpha \pm \pi}{3}$$

$$\cos\alpha = -\mathbf{r}\left(\frac{3}{2 + \mathbf{r}^2}\right)^{3/2}; \quad \mathbf{r} = \frac{\mathbf{c}_2}{\mathbf{c}_1}$$

$$\kappa = \frac{\text{Max } |1 - \mu_i|}{\text{Min } |1 - \mu_i|} \quad (\mathbf{i} = 1, 2, 3) \tag{A3}$$

$N = 4$:

$$\mathbf{Q}_4(\mu) = \{(\mathbf{c}_3 - \mu)(\mathbf{c}_1 - \mu) - (\mathbf{c}_2 - \mathbf{c}_1)^2\} \cdot \{(\mathbf{c}_3 + \mu)(\mathbf{c}_1 + \mu) - (\mathbf{c}_2 + \mathbf{c}_1)^2\}$$

$$\mu_{1,2} = +\frac{1}{2}\{\mathbf{c}_1 + \mathbf{c}_3 \pm \sqrt{(\mathbf{c}_3 - \mathbf{c}_1)^2 + 4(\mathbf{c}_1 - \mathbf{c}_2)^2}\}$$

$$\mu_{3,4} = -\frac{1}{2}\{\mathbf{c}_1 + \mathbf{c}_3 \pm \sqrt{(\mathbf{c}_3 - \mathbf{c}_1)^2 + 4(\mathbf{c}_1 + \mathbf{c}_2)^2}\}$$

$$\kappa = \frac{\text{Max } |1 - \mu_i|}{\text{Min } |1 - \mu_i|} \quad (\mathbf{i} = 1, 2, 3, 4) \tag{A4}$$

APPENDIX B. THE NUGGET EFFECTS AND ITS INFLUENCE ON CONDITION NUMBER

Nugget Covariance

For any of the covariance models considered, a pure "nugget" covariance reflecting the presence of small-scale noise can be added (see for instance Journel and Huijbregts, 1978). In terms of the random field itself, $y(x)$, this corresponds to a decomposition into two independent fields, $y(x) = y_0(x) + y_1(x)$. Here, $y_0(x)$ represents the nugget effect or small-scale variability corresponding to measurement scale, assumed less than the smallest measurement spacing being considered, and $y_1(x)$ represents larger scale variability (e.g., one of the covariance models given with scale ℓ). The resulting covariance of $y(x)$ can be expressed as:

$$\mathbf{C}(\xi) = \mathbf{C}_{00}(\xi) + \mathbf{C}_{11}(\xi) \tag{B-1}$$

where $\mathbf{C}_{11}(\xi)$ is the covariance corresponding to large-scale variability with variance σ^2. The nugget covariance, $\mathbf{C}_{00}(\xi)$, can be represented as a "coarsened white noise" with variance σ_0^2, as is usually done in geostatistical applications:

$$\mathbf{C}_{00}(\xi) = \begin{matrix} \sigma_0^2 \text{ for } |\xi| \leq \ell_0 \\ 0 \text{ for } |\xi| > \ell_0 \end{matrix} \tag{B-2}$$

or, in the limit $\ell_0 \to 0$, as a pure white noise covariance with intensity s_0^2:

$$\mathbf{C}_{00}(\xi) = s_0^2 \delta(\xi) \tag{B-3}$$

It is important to recognize that there is a definite relation between the two formulations which conveys the fact that (B-2) is a "coarsened" version of Eq. (B-3):

$$\sigma_0^2 = s_0^2/(2\ell_0) \tag{B-4}$$

Therefore, for a given intensity, the apparent nugget variance depends in fact on the "small scale" ℓ_0 (scale of measurement).

Effect of Nugget on Condition Number

For small measurement grids, Appendix A gives analytical expressions for the condition number. These expressions can be extended to the case of a nugget. For instance, in the simple case $N = 2$, we obtain:

$$\kappa(\sigma_0) = \frac{1 + |(1 - \sigma_0^2)\rho_1|}{1 - |(1 - \sigma_0^2)\rho_1|} \tag{B-5}$$

where ρ_1 represents the correlation function (without nugget) at lag one ($\xi = \Delta x$). It is not difficult to show that, for $0 \leq \sigma_0^2 \leq 1$ and $-1 \leq \rho_1 \leq +1$

$$\kappa(\sigma_0) \leq \kappa(0) \tag{B-6}$$

This equation expresses that the condition number with nugget effect is always less than that without a nugget. That is, the nugget effect improves the condition number of the system.

It could be expected intuitively that adding a small-scale variance to the diagonal of the covariance matrix improves its condition number. We now use the spectral (continuum) approximation to confirm that this is true for larger values of N as well, at least in the case of monotonic covariance models. Briefly, we start by noting that the Fourier Transform of the pure white noise (B-3) is a constant. Thus, we obtain the spectral density:

$$S(k) = S_{00}(k) + S_{11}(k) = \frac{s_0^2}{2\pi} + S_{11}(k) \tag{B-7}$$

For monotonic models, applying the approximate formula of Eq. (26) and inserting $s_0^2 = 2\ell_0\sigma_0^2$ gives:

$$\kappa(\sigma_0) \approx \frac{S\left(\frac{\pi}{L}\right)}{S\left(\frac{\pi}{\Delta x}\right)} \approx \frac{\sigma_0^2 + \frac{\pi}{\ell_0} S_{11}\left(\frac{\pi}{L}\right)}{\sigma_0^2 + \frac{\pi}{\ell_0} S_{11}\left(\frac{\pi}{\Delta x}\right)} \tag{B-8}$$

For example, using the exponential covariance model (Table 1), we obtain:

$$\kappa(\sigma_0) \approx \frac{\sigma_0^2\ell_0 + \left\{1 + \left(\frac{\pi}{L}\right)^2\right\}^{-1}}{\sigma_0^2\ell_0 + \left\{1 + \left(\frac{\pi}{\Delta x}\right)^2\right\}^{-1}} \tag{B-9}$$

Based on Eqs. (B-8) and (B-9), the following remarks can be made:

- The condition number κ goes to unity in the limit of a pure nugget, or (in other words) it tends to improve as the nugget variance increases
- The condition number κ increases to a finite limit, in the limit of infinite domain with fixed sampling density (Eq. 30). This result holds with and

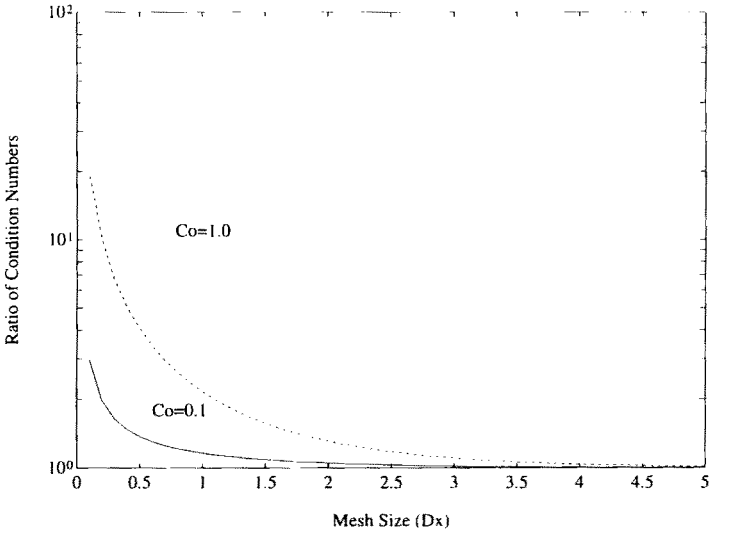


Fig. B-1. Ratio of condition number without a nugget variance over condition number with a nugget variance vs. mesh size Δx in the case of fixed domain size for an exponential covariance. Solid: $C_0 = \sigma_0^2 = 0.1$; dash: $C_0 = \sigma_0^2 = 1.0$.

without a nugget, but it can be shown from our equations that the limit value of κ is smaller when a nugget is present.

- The condition number κ increases to a finite limit, in the limit of infinite sampling density with fixed domain size (Eq. 28), provided a nugget is present. If the nugget variance vanishes, then we obtain a condition number that increases indefinitely as sampling density increases. The nugget, therefore, eliminates this singular behavior of κ.

These theoretically derived results are verified numerically for the case of an exponential covariance model with variance $\sigma^2 = 1.0$. A nugget variance $C_0 = \sigma_0^2$ of 0.1 and 1.0 is added to the main diagonal of the covariance matrix and the condition number κ is calculated. Then the ratio $[\kappa]^{C_0=0}/[\kappa]^{C_0\neq 0}$ is calculated as a function of mesh size Δx. This result is depicted in Fig. B-1. As it was shown earlier, the nugget covariance is improving the condition number of the covariance matrix, especially for small Δx. The improvement in the condition number of the exponential covariance model is calculated to be three and twenty times for $\sigma_0^2 = 0.1$ and $\sigma_0^2 = 1.0$, respectively.

REFERENCES

Chirlin, G. R., and Wood E. F., 1982, On the relationship between kriging and state estimation: Water Res. Res., v. 18, n. 2, p. 432–438.

Christakos, G., 1984, On the problem of permissible covariance and variogram models: Water Res. Res., v. 20, no. 2, p. 251–265.

Dagan, G., 1982, Stochastic modeling of groundwater flow by unconditional and conditional probabilities, 1. Conditional simulation and the direct problem: Water Res. Res., v. 18, n. 4, p. 813–833.

Delhomme, J. P., 1979, Spatial variability and uncertainty in groundwater flow parameters: A geostatistical approach: Water Res. Res., v. 15, n. 2, p. 269–280.

De Marsily, G., 1986, *Quantitative Hydrogeology*: Academic Press, New York, 440 p.

Ekstrom, M. P., 1973, A spectral characterization of the ill-conditioning in numerical deconvolution: IEEE Trans., Audio-Electroacoust., v. AU-21, n. 4, p. 344–348.

Golub, G. H., and Van Loan, C. F., 1989, *Matrix Computations (2nd ed.)*: The John Hopkins University Press, Baltimore and London, 642 p.

Isaaks, E. H., and Srivastava R. M., 1989, *An Introduction of Applied Geostatistics*: Oxford University Press, 561 p.

Journel, A. G., and Huijbregts, C. J., 1978, *Mining Geostatistics*: Academic Press, New York, 600 p.

Lewis, J. P., 1987, Generalized stochastic subdivision: ACM Trans. Graphics, v. 6, n. 3, p. 167–190.

Mantoglou, A., and Wilson, J. L., 1982, The turning bands method for the simulation of random fields using line generation by a spectral method: Water Res. Res., v. 18, n. 5, p. 1379–1394.

MATLAB User's Guide, 1990: The MathWorks, Natick, Massachusetts.

O'Dowd, R. J., 1991, Conditioning of coefficient matrices of ordinary kriging: Math. Geol., v. 23, n. 5, p. 721–739.

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

Press, W. H., Flannery, B. P., Teukolsky, S. A., and Vetterling, W. T., 1986, *Numerical Recipes: The Art of Scientific Computing*: Cambridge University Press, Cambridge, U. K., 818 p.

Tompson, A. F. B., Ababou, R., and Gelhar, L. W., 1989, Implementation of the three-dimensional turning bands random field generator: Water Res. Res., v. 25, n. 10, p. 2227-2243.

Vanmarcke, E., 1983, *Random Fields: Analysis and Synthesis*: The MIT Press, Cambridge, Massachusetts, 382 p.

Yaglom, A. M., 1962, *Stationary Random Functions* (Translated by R. A. Silverman, Ed.): Dover, New York, 235 p.