Mathematical Geology, Vol. 29, No. 5, 1997

Six Factors Which Affect the Condition Number of Matrices Associated with Kriging[1]

George J. Davis[2] and Max D. Morris[3]

Determining kriging weights to estimate some variable of interest at a given point in the field involves solving a system of linear equations. The matrix of this linear system is subject to numerical instability, and this instability is measured by the matrix condition number. Six parameters in the kriging process have been identified which directly affect this condition number. Analysis of a series of 648 experiments gives some insight on these parameters, and how the condition number relates to kriging variance.

KEY WORDS: stability, conditioning, robustness.

INTRODUCTION

Many papers have been published that give advice to scientists involved with kriging. The advice ranges from sampling adequately (Webster and Oliver, 1992), to selecting a data network (Morris, 1991; Yfantis, Flatman, and Behar, 1987), to new applications of kriging (Oliver and others, 1992).

The condition number of the kriging matrix also has been a popular topic. Diamond and Armstrong (1984) discuss condition numbers in a geostatistical context, and show examples of ill-conditioning using close data points. Posa (1989) considers the effect of selecting one variogram model over another on the condition number. Included are some experiments with various data densities and ranges. O'Dowd (1991) considers the condition number of kriging systems with various nugget and sill values for a fixed data configuration. Finally, Ababou, Bagtzoglou, and Wood (1994) discuss the condition number of covariance matrices for five different covariance models.

This paper is an attempt to study systematically the condition number as a function of all of the parameters in the kriging exercise. Specifically, we discuss

[1] Received 29 December 1995; revised 4 October 1996.

[2] Mathematics and Computer Science, Georgia State University, Atlanta, Georgia 30303. e-mail: gdavis@math.cs.gsu.edu

[3] Mathematical Sciences Section, CSM Division, Oak Ridge National Laboratory, Oak Ridge, Tennessee 37830. e-mail: morrismd@ornl.gov

0882-8121/97/0700-0669$12.50/1

the effect of the (i) data network, (ii) variogram model, (iii) nugget effect, (iv) range, (v) selecting kriging weights using the covariance matrix or semivariogram matrix, and (vi) sill value. A series of 648 experiments were performed to examine the change in the condition number as each of these six parameters are changed.

In the next section, we present an overview of the kriging problem, and examine the six factors. We then present a general discussion of mathematical conditioning and numerical stability. Relevant quantities are defined and described in the context of geostatistics. In the next section, we take a detailed look at the results of the experiments. This leads to the final section of conclusions.

OVERVIEW OF THE KRIGING PROBLEM

Kriging is an important technique for estimating spatially distributed data. Kriging was used first almost 40 years ago (Cressie, 1990) and is applied to mining (Krige, 1951; Journel and Huijbregts, 1978), meteorology (Cressie, 1990), estimating the risk of childhood cancer (Oliver and others, 1992), and many other fields. Given a set of sample points within some closed region, and measurements of some variable of interest at these points, kriging is used to predict the variable of interest at any other point in the region. Kriging can take many forms, and in this paper we consider the theory and application of ordinary kriging. The method proceeds through well-defined stages, each of which requires care and some amount of judgment. The quality of the final predicted values depends on these judgments.

First, a network of sample points must be determined. Several authors have considered the effects of different data networks (Morris, 1991; Yfantis, Flatman, and Behar, 1987) and the importance of sampling adequately (Webster and Oliver, 1992). It is critical that enough points are selected to estimate the sample semivariogram meaningfully:

$$\hat{\gamma}(h) = \frac{1}{2m(h)} \sum_i (z(x_i) - z(x_i + h))^2$$

Here, h is the distance between two sample points, $m(h)$ is the number of pairs of points separated by distance, h, and z is the variable of interest, regarded as a realization from a stationary stochastic process. The sample semivariogram is used to model the semivariogram

$$\gamma(h) = \tfrac{1}{2}E(z(x) - z(x + h))^2$$

where E is the expected value.

Having a plot of observed values of $\hat{\gamma}(h)$, these points are fit to one of the

permissible models for the variogram. There are many possible selections, and among the more popular are spherical, exponential, and Gaussian models.

Typically these functions are scaled to have limiting values of 1 as h gets large, and in this example 1 is termed the "sill" value. Sill values may not always be 1. Scaling of the data (and consequently of the variogram functions) can result in arbitrary values for sill. This, in turn, can lead to potentially dangerous numerical instability in the resulting kriging matrix. This instability is manifested by a kriging matrix with a large condition number, which is caused in this example by the interaction of large and small numbers.

The three variograms also have a parameter a, termed the range of influence. Usually selected as a percentage of the largest possible h value in the sample, a indicates when the variogram has effectively reached its sill. In the exponential and Gaussian model, a is by convention the distance at which the variogram value is approximately 95% of the sill value. Precisely determining this range of influence parameter from data can be difficult, but small changes in a will not dramatically influence the goodness-of-fit.

For the range of values of h, these variogram curves have similar shape, and it may not be obvious which curve best fits the observed values. Posa (1989) has said, "In many geostatistical case studies, no physical reason exists to prefer one variogram model to another which fits the experimental variogram equally well. Besides, in many situations, small changes in the parameters of the selected variogram model produce an equally acceptable fit to a given experimental variogram." However, it has been noted previously (Diamond and Armstrong, 1984; Abalou, Bagtzoglou, and Wood, 1994), that the numerical consequences of these selections can be severe.

By observing the plots of values, the underlying variogram function may not appear to pass through the origin, but tends toward the positive y-axis. This can be incorporated into the model through the "nugget effect." If the point of intersection with the y-axis appears to be at c_0, then the standard variogram functions can be formulated as:

$$\text{spherical:} \quad \gamma(h) = \begin{cases} c_0 + c_1 \left(\dfrac{3h}{2a} - \dfrac{1}{2} \left(\dfrac{h}{a} \right)^3 \right) & \text{if } h \le a \\ c_0 + c_1 & \text{if } h > a \end{cases},$$

$$\text{for } h \neq 0 \qquad \gamma(0) \equiv 0$$

$$\text{exponential:} \quad \gamma(h) = c_0 + c_1(1 - e^{-3h/a})$$

$$\text{Gaussian:} \quad \gamma(h) = c_0 + c_1(1 - e^{-3h^2/a^2})$$

In these formulae, c_1 is selected so that $(c_0 + c_1)$ is the sill value. For reasons that we discuss later, we rescale the data so that the sill value is 1.

Given any point x, the variable of interest $z(x)$ is predicted as a linear combination of the values of z in the network:

$$\hat{z}(x) = \sum_i w_i z(x_i)$$

The w_i are termed the kriging weights, and are the solution of the linear system

$$\begin{bmatrix} \Gamma & \mathbf{1} \\ \mathbf{1}^T & 0 \end{bmatrix} \begin{bmatrix} \mathbf{w} \\ \lambda \end{bmatrix} = \begin{bmatrix} \mathbf{r} \\ 1 \end{bmatrix}$$

Here, Γ is the semivariogram matrix defined by $\Gamma_{ij} = \gamma(|x_i - x_j|)$, $\mathbf{1}$ is the column vector of all ones, $\mathbf{r}$ is defined by $r_i = \gamma(|x - x_i|)$, and $\mathbf{w}$ is the (unknown) vector of weights. The additional unknown λ, a Lagrange parameter, is present to ensure that the sum of the weights is one, and the estimation is unbiased. Alternatively, we could solve

$$\begin{bmatrix} C & \mathbf{1} \\ \mathbf{1}^T & 0 \end{bmatrix} \begin{bmatrix} \mathbf{w} \\ -\lambda \end{bmatrix} = \begin{bmatrix} \mathbf{s} \\ 1 \end{bmatrix}$$

for the same weights $\mathbf{w}$, with C representing the covariance matrix $C_{ij} = \text{cov}(x_i, x_j)$, $s_i = \text{cov}(x, x_i)$.

The mathematical values of the w_i are identical in each of the two formulations, but the numerical behavior of the systems can differ. The subject of this paper is to study the numerical stability of these two linear systems.

CONDITIONING AND STABILITY

Errors in the numerical solution of a problem can arise from many sources, including measurement errors and errors from floating point arithmetic. Two major concepts related to detecting and assessing error are conditioning of a problem and stability of an algorithm. These concepts are often used interchangeably and erroneously. Conditioning describes the sensitivity of a problem to changes in its data, and stability describes the sensitivity of an algorithm to changes in its parameters. In the literature, the words may be interchanged, but one concept refers to the properties of a problem and the other to those of an algorithm. This is an important distinction.

Consider an extreme example. Suppose we wish to solve the 2×2 linear system

$$x + y = 2$$
$$x + y = 1$$

Clearly no solution exists, and this conclusion is independent of any algorithm we may wish to employ. This is as ill-conditioned as a problem can get; no solution at all. Any small change in the coefficients of x and y causes a radical change in the nature of the solution.

Now consider the problem:

$$x + (1 + \epsilon)y = 2$$
$$x + y = 1$$

where ϵ is some small nonzero number. This problem is ill-conditioned, as a small change in the coefficients of x and y (for example, changing ϵ) will cause a large change in the solution vector $(1 - (1/\epsilon), (1/\epsilon))^T$. Again, this is independent of the actual algorithm we may wish to use to compute the answer. An unstable algorithm may make the computed solution even worse.

As a contrast, consider the well-conditioned problem:

$$x + y = 4$$
$$7y = 21$$

Small changes in the coefficients of x and y do not cause large relative changes in the solution. But think about the effect of using a poor algorithm to solve this system. Using 7-digit decimal arithmetic (roughly equivalent to single precision FORTRAN), compute the multiplicative inverse of 7, and multiply it on both sides of the second equation. The solution would be computed as $y = (0.1428571) * 21 = 2.9999991$, not the 3 we would expect. Then x would be computed as $x = 4 - y = 1.0000009$. The error in the solution is the result of the selection of an inferior algorithm. The problem is well-conditioned, and the x and y can be computed exactly in 7-digit arithmetic by first dividing 7 into 21!

Note that there is no precise dividing line between well-conditioned and ill-conditioned, or stable and unstable. These are judgments which depend on the precision of the data, the precision of the arithmetic used to solve the problem, and the desired accuracy of the answer. The inexact answer as given might be acceptable in one application, and unsatisfactory in another.

Fortunately for the general linear systems problem $A\mathbf{x} = \mathbf{b}$, much is known about conditioning and stability. A quantity termed the condition number of A defines the mathematical sensitivity of the problem. The condition number is defined as

$$\kappa(A) = \begin{Bmatrix} \|A\| \, \|A^{-1}\| & \text{if } A \text{ nonsingular} \\ \infty & \text{if } A \text{ singular} \end{Bmatrix}, \quad \text{where } \|\bullet\| \text{ is any matrix norm}$$

The main theorem on conditioning is as follows. Let $A\mathbf{x} = \mathbf{b}$ and $(A + \epsilon F)\tilde{\mathbf{x}} = (\mathbf{b} + \epsilon\mathbf{f})$. Here, F and $\mathbf{f}$ have roughly the same magnitudes as A and $\mathbf{b}$, respectively. Thus, the second system can be thought of as a perturbation of the first one, and, for any given vector norm $\|\cdot\|$, we are interested in the relative error $\|\mathbf{x} - \hat{\mathbf{x}}\|/\|\mathbf{x}\|$.

Theorem:

$$\frac{\|\mathbf{x} - \tilde{\mathbf{x}}\|}{\|\mathbf{x}\|} \leq \kappa(A)\epsilon\left(\frac{\|F\|}{\|A\|} + \frac{\|\mathbf{f}\|}{\|\mathbf{b}\|}\right)$$

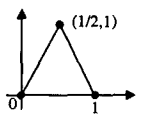
Proof (Golub and VanLoan, 1989). Generally, this theorem says that the relative change in the solution of a linear system is bounded by the condition number times the relative error in the matrix and its right-hand side. Again, this is completely independent of the algorithm used to solve the linear system. This theorem describes an upper bound on the change in the answer when errors are made in the problem.

Unstable algorithms also can contribute to the error in the solution. It is well known, for example, that Gaussian elimination without pivoting is an unstable algorithm. Examples of well-conditioned systems exist for which this algorithm produces unacceptable answers, whereas other algorithms with pivoting do not. In all of the work that follows, we use Gaussian elimination with partial pivoting. Although it is not as stable as Gaussian elimination with complete pivoting, it may be the algorithm of selection in applied numerical analysis.

If there is an overall message about stability and conditioning it would be that ill-conditioned systems are difficult to solve accurately no matter what algorithm is employed, and unstable algorithms should be avoided at all times.

Let us now bring this general discussion to typical kriging problems in geostatistics. Frequently, field measurements are imprecise, and are known to only one or two digits of accuracy. Entries in the kriging matrix system are functions of those measurements. The potential exists for condition numbers on the order of 1000 to 10000 to magnify error to the point that the computed solution could be completely meaningless. Numerical analysts may be trained to think of condition numbers of 10^6 or 10^7 as high (when using single precision FORTRAN), yet kriging condition numbers less than that could be troublesome.

The condition number can be sensitive to scaling. In particular, consider the effect of the sill value in the following example. Select a three-point data network over the unit square as indicated next. Suppose the experimental var-



iogram is determined to have sill value 1, nugget value 0.30, and an exponential shape. That is, $\gamma(h) = 0.30 + 0.70(1 - e^{-3h/a})$. If $a = 1$, the condition number of the resulting kriging matrix is about 3.1 clearly well conditioned. If the sill value is rescaled to 1000, the variogram becomes $\gamma(h) = 300 + 700(1 - e^{-3h/a})$, and the condition number jumps to 1.26×10^6. The high condition number was caused by the artificial scaling, which then poses a significant numerical threat. To avoid this difficulty, we always take the sill value to be 1.

If the scaling has been done properly, what can be done about high condition numbers? Nothing, once the matrix system has been formulated. However, several selections besides the sill must be made before the system is finalized, and these selections can affect drastically the condition number of the resulting matrix. It is these selections which we now investigate.

FACTORS WHICH AFFECT CONDITIONING

We have identified six factors which affect the condition number of the kriging system. They are: (i) data network/data spacing, (ii) selection of the variogram model, (iii) nugget effect, (iv) selection of a range of influence parameter a, (v) selection of solving the matrix system using Γ or C, and as previously mentioned, and the (vi) sill value.

We have performed 648 numerical experiments in an attempt to discover how these factors affect the condition number of the kriging matrix. We started with a unit square as our sampling region, and considered three different data networks at four different densities. We selected data points as vertices of triangles, squares, and hexagons, with approximately 16, 25, 36, and 100 data points in the region. Because of border effects, we could not select exactly the same number of points for each configuration. Three different variogram models were used: spherical, exponential, and Gaussian, with three different nugget effects: 0, 0.15, and 0.3. Three different ranges of influence a were used, expressed as 0.5, 0.75, or 1.0 times the maximum distance between two points on the network. We also considered a wider range of nugget effects and ranges of influence for square networks. The trends observed were merely extensions of the trends reported in what follows. No new insight was gained from the extended situations. The condition numbers of the kriging matrices based on Γ and C, referred to as K_Γ and K_C, respectively, were recorded. As sill values were 1.

The data from the experiments then were examined by holding five of the factors constant, and considering the effect of differing the sixth. In this way some insight can be gained on the effect of each individual factor on the condition number of the kriging matrix.

One of the few trends that was "absolute" across all the experiments was that increasing the number of data points n in the unit square will result in a

higher condition number in K_Γ and K_C. To see why, consider a three-point data network and, assuming no nugget,

$$K_\Gamma = \begin{bmatrix} 0 & \gamma(|x_1 - x_2|) & \gamma(|x_1 - x_3|) & 1 \\ \gamma(|x_2 - x_1|) & 0 & \gamma(|x_2 - x_3|) & 1 \\ \gamma(|x_3 - x_1|) & \gamma(|x_3 - x_2|) & 0 & 1 \\ 1 & 1 & 1 & 0 \end{bmatrix}$$

Now, if we define $|x_1 - x_2| = \equiv \epsilon$, and $|x_1 - x_3| \equiv \nu$, consider the effect of letting x_1 approach x_2, or $\epsilon \to 0$. Then,

$$K_\Gamma = \begin{bmatrix} 0 & \gamma(\epsilon) & \gamma(\nu) & 1 \\ \gamma(\epsilon) & 0 & \gamma(|x_2 - x_3|) & 1 \\ \gamma(\nu) & \gamma(|x_3 - x_2|) & 0 & 1 \\ 1 & 1 & 1 & 0 \end{bmatrix}$$

$$\to \begin{bmatrix} 0 & 0 & \gamma(\nu) & 1 \\ 0 & 0 & \gamma(\nu) & 1 \\ \gamma(\nu) & \gamma(\nu) & 0 & 1 \\ 1 & 1 & 1 & 0 \end{bmatrix}$$

which is exactly singular. Therefore, although it is important to sample adequately from a statistical viewpoint, too fine a data network can have a negative numerical impact.

In a regular data network, it is of far less consequence exactly how the points are selected. In all of our experiments the condition numbers from square, triangular, and hexagonal networks were close, approximately within a factor of two. These minor variations probably were caused more by the border effects of different values of n than by any fundamental mathematical structure. Here is a typical example selected from a square grid with an exponential variogram and nugget value 0.30, and range of influence a equal to one half the maximum point separation. Figure 1 shows both the effect of increasing n and the relatively little difference in the three data networks. The solid line represents square networks with $n = 16$, 25, 36, and 100; the dashed line triangular networks with $n = 17$, 27, 39, and 105; and the dash-dot line hexagonal networks with $n = 18$, 28, 36, and 112. Results for K_Γ and for other kriging parameter values are similar.

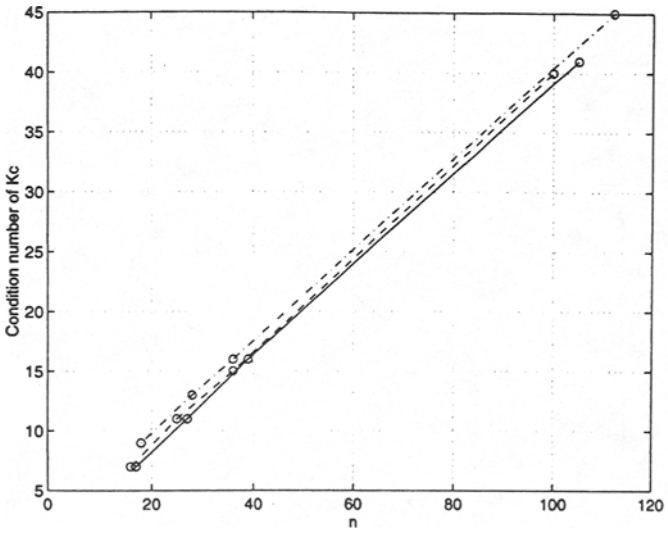


Figure 1. Effect of increasing n and changing data network.

It usually is believed that both the spherical exponential variograms have "reasonable" numerical behavior, and Gaussian variograms can experience problems. Generally, it is true in our experiments that the exponential variogram produced condition numbers less than (and close to) those of the spherical, which, in turn, were less than those of Gaussian. In fact, for matrix K_C, the exponential was marginally better than spherical every time, and the gap closed for larger nugget values. However, there are several situations in which the Gaussian variogram actually produced the best conditioned matrix K_Γ. Figure 2 shows the condition numbers of K_Γ for the spherical (solid), exponential (dashed), and Gaussian (dash–dot) variograms over square grids of various sizes with nugget value 0.30, and a set to the maximum point separation. The graph shows Gaussian with the highest condition number for small n, and lowest condition number for large n. Note also how the condition number for the spherical variogram is close to (and marginally worse than) that of the exponential.

There are times, of course, where the Gaussian variogram produces badly conditioned matrices. This happens for large values of n with no nugget effect. The matrices produced in some of these situations were numerically singular, the only time such high condition numbers were experienced.

Increasing the nugget value had a beneficial effect on the condition number in all situations. Thus, leaving all other factors constant, an increased nugget value resulted in a smaller condition number. In the limiting situation of a pure nugget effect, K_Γ would be all ones except for the main diagonal of zero, and

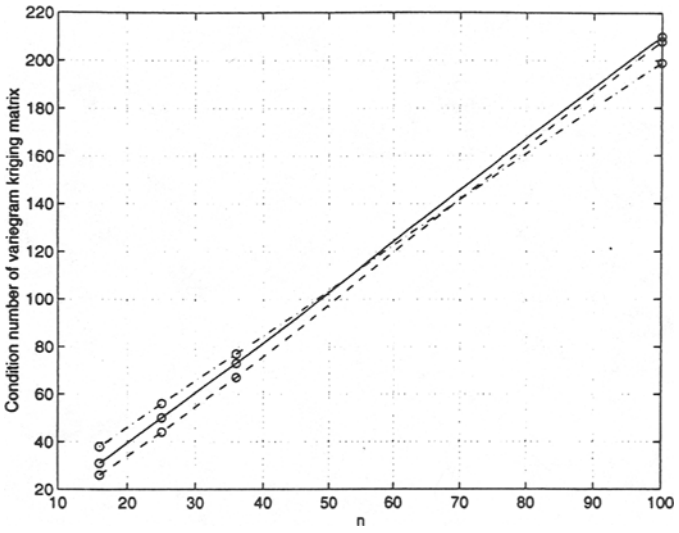


Figure 2. Effect of different variogram models.

the condition number of such a K_Γ is $(n - 1)$. The corresponding K_C has an $(n - 1) \times (n - 1)$ identity matrix as a principal submatrix, and its condition number is $(1 + \sqrt{4n - 3})/2$, as confirmed by O'Dowd (1991).

The effect of the range of influence parameter a, is more difficult to understand from a matrix point of view. In one sense, decreasing a tends to have the same effect as increasing the nugget value. Consider a spherical variogram model, and think about the effect of decreasing a towards zero. This would make the variogram kriging matrix approach the pure nugget situation of an all-zero diagonal and all-one off-diagonal. Such matrices are well conditioned, as noted. Allowing off-diagonal elements to take on other values (as they do with increasing a) increases the chances for linear dependency, tending to raise the condition number.

This was observed in our experiments. For the kriging matrix K_C, increasing a while keeping all other parameters constant, increased the condition number every time. In K_Γ, however, that was not universally the situation. Figure 3 shows the condition number of K_Γ plotted against a values of 1/2, 3/4, and 1 times the maximum interpoint distance, on square grids using the spherical variogram model and nugget value 0.30. Three different values of n are plotted. The bottom line (dashed) represents $n = 16$, and shows the condition number strictly increasing with increasing a. It is difficult to see in the plot itself, but the condition numbers are 30.5, 30.78, and 30.82. The middle line (solid)

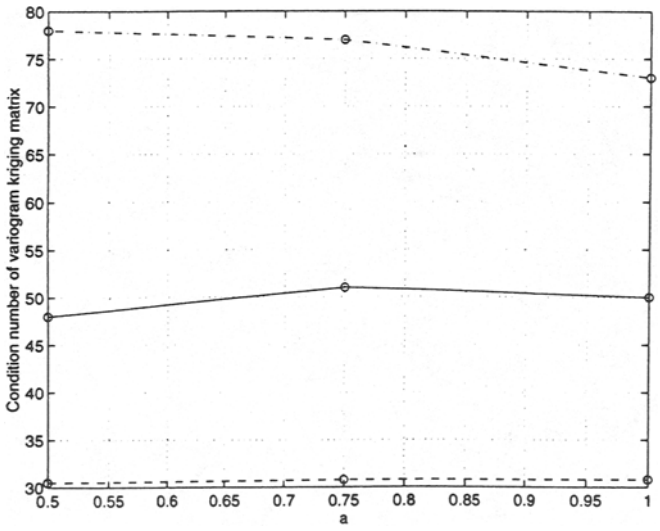


Figure 3. Effect of increasing range a.

represents $n = 25$, and shows the condition number first increasing from 48 to 51 and then decreasing to 50. The top line (dash–dot) represents $n = 36$ and shows the condition number to be decreasing from 78 to 77 to 73. Although generally it is true that increasing a will decrease the condition number, this example shows that a variety of behaviors is possible.

It was shown earlier that the kriging systems with K_Γ and K_C are mathematically equivalent. However, the condition numbers of the two matrices can differ. In our experiments, it was true that the condition number of K_C was less than that of K_Γ for all situations using the spherical or exponential model. The difference was not dramatic, and was approximately within a factor of ten each time. For some reasonably well-conditioned Gaussian variogram situations, interesting effects were observed.

Figure 4 is a plot of condition number of K_Γ (solid line) and K_C (dashed line) for various values of n on square grids, using the Gaussian variogram model, nugget value 0.15 and a the maximum interpoint distance. We see that for $n = 16$, the condition number of K_Γ is larger than that of K_C (69 vs. 61), whereas for $n = 100$ the roles have reversed (342–363). The superiority of K_Γ is the exception, although this situation shows that this is not an ironclad rule.

Although the condition number is a valuable index of the numerical difficulty of a problem, the goal of kriging is essentially statistical, that is the prediction of z, especially at sites not sampled. The uncertainty of $\hat{z}(x)$ is expressed in terms of the predictive variance or "kriging variance":

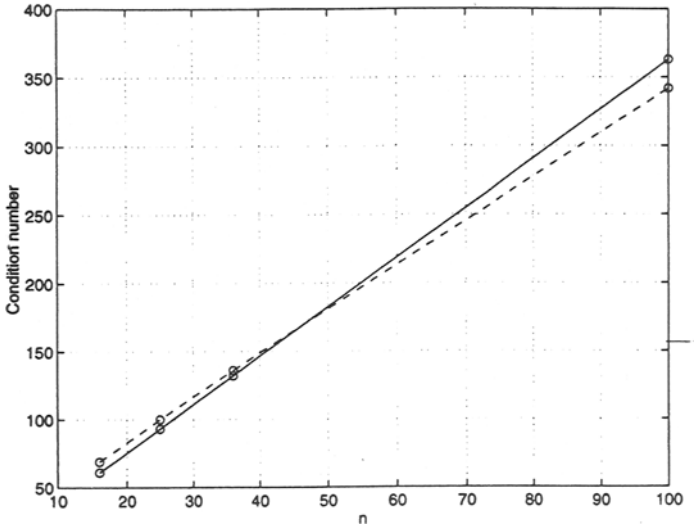


Figure 4. Condition number of two kriging matrices.

$$\begin{aligned}\operatorname{var}(\hat{z}(x) - z(x)) &= 2\sum_{i=1}^{n} w_i\gamma(x_i - x) - \sum_{i=1}^{n}\sum_{j=1}^{n} w_i w_j \gamma(x_i - x_j) \\ &= 2\mathbf{w}^T\mathbf{r} - \mathbf{w}^T\Gamma\mathbf{w} \\ &= \mathbf{w}^T(2\mathbf{r} - \Gamma\mathbf{w})\end{aligned}$$

For most used covariance functions, this quantity typically is relatively small for x near one or more sites in the sampling network, and larger for more remote x.

One natural summary measure of the quality of kriging precision available in a given situation is the kriging variance averaged over the geographic region of interest Ω:

$$\int \operatorname{var}(\hat{z}(x) - z(x))\, d\Omega$$

Solving for the kriging weights using the semivariogram matrix, we numerically estimated the given integral as the average kriging variance for 1000 randomly selected sites. The same 1000 sites were used for each calculation, and we let the other kriging parameters differ as before for a total of 324 examples. The randomly selected sites are shown in Figure 5.

Figure 6 shows the relationship between the condition number of the kriging

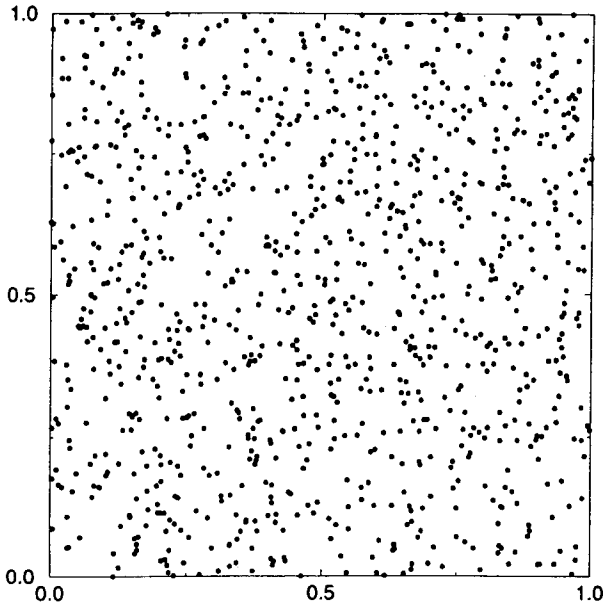


Figure 5. 1000 randomly selected sites.

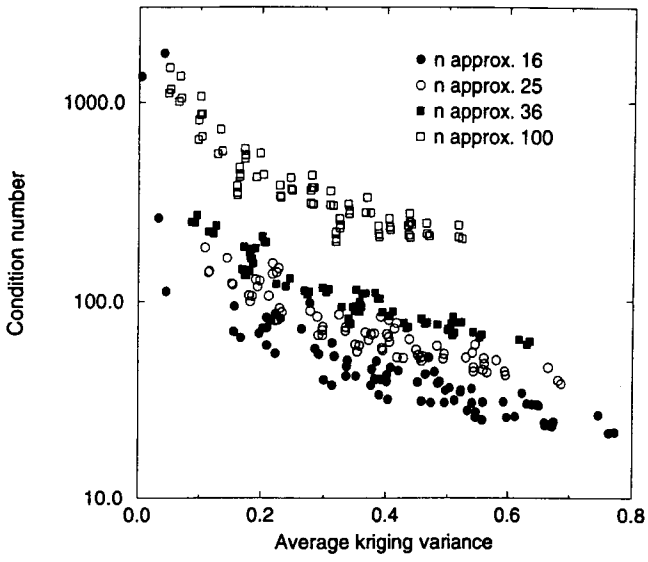


Figure 6. Average kriging variance plotted against condition number.

matrix and the estimated average kriging variance for the examples. Points are coded by the four nominal values of n considered. This particular variable gives a clearer separation of the points into groups than any of the other parameters considered. Especially for a given value of n, it is clear that larger condition numbers are associated with smaller average kriging variances, and vice versa. This is not particularly surprising because the types of factors which lead to larger condition numbers (e.g., little or no nugget, relatively large values for the influence parameter, etc.) generally are those which characterize a sample which is "informative" for spatial prediction purposes. However, the figure shows how "tight" this relationship is for the range of conditions studied here. For a given value of n, changes in the other five factors which lead to a meaningful increase (decrease) in the condition number will be associated with a decrease (increase) in the average kriging variance. Separation, by nominal sample sizes, of the four groups plotted shows that a given level of statistical precision is associated with larger condition numbers as the sample size increases.

CONCLUSIONS

Our experiments have shown that all regular data networks of a given density produce linear systems with comparable condition numbers. Thus, we see no significant advantage to any for the square, triangular, or hexagonal networks. An adequate sampling density is necessary for statistical purposes, but increasing density is associated clearly with increasing condition numbers. In the absence of a nugget effect, clusters of close data elements can present serious numerical problems.

The semivariogram model, nugget, and influence parameter generally are estimated, formally or otherwise, using the sample semivariogram. This process should be as objective as possible, even though the data do not always provide a firm basis for each decision, as noted earlier. Because of the statistical reliability of kriging is at least to some extent reliant on these decisions and estimates, we cannot urge users to "bias" these selections in such a way as to reduce the condition number. However, users should be aware of effects these selections have on the numerical conditioning of the kriging system.

Relatively high (low) condition numbers generally are associated with relatively low (high) average kriging variances, especially so once the value of n has been fixed. For many applications, this apparent conflict is not serious; condition numbers remain in the acceptable range for situations in which spatial predictions are precise enough for the problem at hand. However, for situations which usually seem desirable from a statistical point of view, for example, large n, small nugget, and slowly decaying correlations, numerical conditioning of kriging can become a serious problem. This suggests that statisticians interested

in working with prediction problems characterized by these apparently favorable conditions might reasonably consider whether slightly less statistically efficient techniques with better numerical properties might be preferred.

Perhaps the best advice we can give is to be mindful of the condition number when building and solving kriging systems. Software exists to calculate quickly the condition number of the matrices involved. These should be considered along with the statistical properties associated with a network and semivariogram when evaluating the practical value of kriging in application.

REFERENCES

Ababou, R., Bagtzoglou, A. C., and Wood, E. F., 1994, On the condition number of covariance matrices in kriging, estimation, and simulation of random fields: Math. Geology, v. 26, no. 1, p. 99–133.

Cressie, N., 1990, The origins of kriging: Math. Geology, v. 22, no. 3, p. 239–252.

Diamond, P., and Armstrong, M., 1984, Robustness of variograms and conditioning of kriging matrices: Math. Geology, v. 16, no. 8, p. 809–822.

Golub, G. H., and Van Loan, C. F., 1989, Matrix computations (2nd. ed): The Johns Hopkins Univ. Press, Baltimore, MD, 642 p.

Journel, A. G., and Huijbregts, Ch. J., 1978, Mining geostatistics: Academic Press, New York, 600 p.

Krige, D. G., 1951, A statistical approach to some basic mine valuation problems on the witwatersrand: Jour. Chem. Metal. Min. Soc. South Africa, v. 52, p. 119–139.

Morris, M. D., 1991, On counting the number of data pairs for semivariogram estimation: Math. Geology, v. 23, no. 7, p. 929–943.

O'Dowd, R. J., 1991, Conditioning of coefficient matrices of ordinary kriging: Math. Geology, v. 23, no. 5, p. 721–739.

Oliver, M. A., Lajaunie, Ch., Webster, R., Muir, K. R., and Mann, J. R., 1992, Estimating the risk of childhood cancer, *in* Soares, A., ed., Geostatistics Troia '92: Kluwer Acad. Publ., Dordrecht, p. 899–910.

Posa, D., 1989, Conditioning of the stationary kriging matrices for some well-known covariance models: Math. Geology, v. 21, no. 7, p. 755–765.

Webster, R., and Oliver, M. A., 1992, Sample adequately to estimate variograms of soil properties: Jour. Soil Science, v. 43, no. 1, p. 177–192.

Yfantis, E., Flatman, G. T., and Behar, J. V., 1987, Efficiency of kriging estimation for square, triangular, and hexagonal Grids: Math. Geology, v. 19, no. 3, p. 183–205.