FAST MULTIDIMENSIONAL INTERPOLATIONS

Franklin G. Horowitz[1,2], Peter Hornby[1,2], Don Bone[3], Maurice Craig[4]

[1]CSIRO Division of Exploration & Mining, Nedlands, WA 6009 Australia
[2]Australian Geodynamics Cooperative Research Centre, Nedlands, WA 6009 Australia
[3]CSIRO Division of Information Technology, Canberra, ACT 2601 Australia
[4]CSIRO Division of Exploration & Mining, Floreat Park, WA 6014 Australia

ABSTRACT

We have developed a high-performance, flexible scheme for interpolating multi-dimensional data. The technique can reproduce exactly the results obtained, for example, from Ordinary Kriging and related techniques in 3D, or from Thin-Plate Splines (Briggs' minimum-curvature algorithm) in 2D. Moreover, compared to traditional implementations of these algorithms, our method enjoys large computational-cost savings.

The new approach produces an interpolation that obeys a Partial Differential Equation (PDE). The PDE may arise from physically based arguments, but its form can vary widely. It might be specified only implicitly (as through a Model Variogram), or be nonlinear (although a performance penalty could then apply).

While a formal equivalence between Kriging and Splines has been known for some time (Matheron (1980)), the present derivation, from radial basis functions, further illuminates this connection. Thus, for example, we can make explicit the PDEs that underlie some of the Model Variograms most often used in Geostatistics. Besides its practical utility, the work thereby acquires a theoretical interest.

BACKGROUND

The aim of this paper is to explore the close connections between several seemingly different topics: Kriging; radial basis function (RBF) interpolations; Spline interpolations; and the solutions of Partial Differential Equations (PDEs). We demonstrate that these techniques can produce families of identical interpolations.

We start with some notation. Throughout this paper, we denote vectors in a lower case boldface font, and matrices in an upper case boldface. Let $\boldsymbol{x}_i \in \Re^n$ be the position of a control point at which some scalar value u_i is known. We denote by $u(\boldsymbol{x})$ an interpolation which passes through the set of N control points $u(x_i) = u_i;\ \ 1 \le i \le N$.

KRIGING

Simple Kriging

In the traditional formulation of simple Kriging (and its derivatives; Isaaks and Srivastava (1989)), one estimates $u(\boldsymbol{x}_0)$ at some interpolation point $\boldsymbol{x}_0$ from a weighted linear combination of the u_i

$$u(\boldsymbol{x}_0) \approx (\boldsymbol{w}_0, \boldsymbol{u}) \ , \tag{1}$$

where $\boldsymbol{u}$ is the vector with components u_i and $(\boldsymbol{a}, \boldsymbol{b})$ denotes the inner (dot) product between two vectors. The weight vector $\boldsymbol{w}_0$ is found by solving the linear system

$$\boldsymbol{C}\boldsymbol{w}_0 = \boldsymbol{b}_0 \, , \tag{2}$$

where the $C_{ij} = \phi(\boldsymbol{x}_i - \boldsymbol{x}_j)$ are the covariances between control points in the assumed model, and $\boldsymbol{b}_0 = \phi(\boldsymbol{x}_0 - \boldsymbol{x}_j)$ gives the covariances between the control points and $\boldsymbol{x}_0$.

We choose simple Kriging for its mathematical simplicity, and because most (if not all) of the other Kriging techniques can be derived from simple Kriging via additional constraints and/or via integrals of the resulting interpolation (e.g. ordinary Kriging, universal Kriging or Kriging with a trend, block Kriging).

Dual Kriging

In the dual Kriging formulation (Matheron (1982), or Royer and Vieira (1984), for an English example), one forms the estimate $u(\boldsymbol{x}_0)$ from a single set of weights $\boldsymbol{f}$,

$$u(\boldsymbol{x}) \approx (\boldsymbol{b}_0, \boldsymbol{f}) \ . \tag{3}$$

The unknown $\boldsymbol{f}$ is obtained by solving

$$\boldsymbol{C}^{\dagger} \boldsymbol{f} = \boldsymbol{u} \, . \tag{4}$$

Details are given in Appendix 1, where dual formulation analogues of techniques like Kriging with a trend are also outlined.

Reprinted with corrections from Chapter 9 (pp. 53-56) of the 26th Proceedings of the Application of Computers and Operations Research in the Mineral Industry (APCOM 26), R.V. Ramani (ed.), Soc. Mining, Metall., and Explor. (SME), Littleton, Colorado, U.S.A., 538 p., 1996.

Efficiency

From the standpoint of the operations count required to compute an interpolation, a direct implementation of equations (3) and (4) offers a slight advantage over equations (1) and (2). Recall that $\boldsymbol{C}$ is $N \times N$.

The dual formulation requires solving only the single linear system (4), which costs $O(N^3)$ operations for practical algorithms (e.g. Press, et al. (1992) for a readable discussion). In order to evaluate the interpolation at M different grid points, one needs a further $O(MN)$ operations (M evaluations of (3), each of which requires an $O(N)$ dot product between $\boldsymbol{f}$ and $\phi(\boldsymbol{x}_0 - \boldsymbol{x}_i)$). Thus a direct implementation of dual Kriging requires a grand total of

$$O(N^3) + O(MN) \tag{5}$$

operations.

The simple formulation requires solving the linear system (2) at each of the M grid points. A direct implementation requires $O(MN^3)$ operations. Evaluation of (1) over the grid requires a further $O(MN)$ operations, yielding a total cost of

$$O(MN^3) + O(MN) \tag{6}$$

operations. A more sophisticated algorithm might be to solve (2) by inverting $\boldsymbol{C}$ once, costing $O(N^3)$. Then, over each of M grid points, it would evaluate $\boldsymbol{C}^{-1}\boldsymbol{b}_0$. Each matrix multiplication requires $O(N^2)$ operations, resulting in $O(MN^2)$. The total cost of this algorithm being

$$O(N^3) + O(MN^2) \ , \tag{7}$$

and it might be more efficient than the direct implementation (depending on how M scales with N). Remember, sampling is typically pointwise (spot sampling) or linewise (drill strings, traverses, flight lines), and hence M could be as high as $O(N^2)$ or even $O(N^3)$, although $O(N)$ is probably more common.

Clearly, the dual formulation is more efficient than either of the simple formulation algorithms. Equally clearly, when $N \sim 10,000$ (common for real world applications), *none* of these algorithms (as stated) is practical on commonly available computers. This observation has led to the development of many localised algorithms, whose primary computational reason for existence is to force N down to something manageable (like 40 or 50). Of course these algorithms require large and complex bookkeeping components, for such operations as nearest neighbourhood searches. We will not consider the ramifications of this modification here.

RBF INTERPOLATION

The theoretical underpinnings of radial basis function interpolation are described in Powell (1992), and the references found therein. Briefly, although the method is not formulated in the language of spatial statistics, the actual interpolation is computed nearly identically to dual Kriging,

$$u(\boldsymbol{x}_0) \ = \ \sum_i f_i \phi(\boldsymbol{x}_0 - \boldsymbol{x}_i) + P_h(\boldsymbol{x}) \ . \tag{8}$$

Here ϕ is as before, and $P_h(\boldsymbol{x})$ is some polynomial of the components of $\boldsymbol{x}$. The f_i and the coefficients of the polynomial $P_h(\boldsymbol{x})$ satisfy an augmented system

$$\hat{\boldsymbol{C}}^{\dagger}\hat{\boldsymbol{f}} \ = \ \hat{\boldsymbol{u}} \, . \tag{9}$$

The only difference between equations (8)-(9) and equations (3)-(4) is the addition of the polynomial term $P_h(\boldsymbol{x})$, resulting in the augmentation of the simple Kriging system of equations (Appendix 1). When $P_h(\boldsymbol{x}) = 0$ we recover (3), and when $P_h(\boldsymbol{x}) \neq 0$, we have the dual Kriging analogue of universal Kriging (also known as Kriging with a trend).

In RBF parlance, Kriging's model covariance functions, $\phi(\boldsymbol{x}_0 - \boldsymbol{x}_i)$, are the radial basis functions upon which the technique is founded. Just as in Kriging, practitioners of RBF interpolation choose the form of ϕ from a small set of functions known to be admissible (principally by inducing the matrix $\hat{\boldsymbol{C}}^{\dagger}$ to be invertible). The efficiency of RBF interpolation is identical to that of dual Kriging (e.g. described by expression (5)).

PARTIAL DIFFERENTIAL EQUATIONS

Links to RBFs and Kriging

We now motivate the connection between dual Kriging, radial basis functions, and linear PDEs by relating the PDE formalism to the RBF formalism. Suppose that, for a given ϕ, there exists an operator L such that

$$L\phi(\boldsymbol{x} - \boldsymbol{x}_i) \ = \ \delta(\boldsymbol{x} - \boldsymbol{x}_i) \ , \tag{10}$$

where δ is Dirac's delta, and

$$LP_h(\boldsymbol{x}) \ = \ 0 \ . \tag{11}$$

For example, if L is a differential operator, ϕ is the Green's function of L, and $P_h(\boldsymbol{x})$ is a homogeneous solution to the PDE. Admittedly, (10)-(11) are strong assumptions to make, however, our aim here is to outline an idea, and we do not wish to cloud the simplicity of that idea with mathematical trickery dealing with difficult cases.

Applying L to (8) gives

$$Lu = \sum_i f_i \delta(\boldsymbol{x} - \boldsymbol{x}_i) \ . \qquad (12)$$

This corresponds to an operator equation, where u is the response of a system (described by L) to a set of point forces of magnitude f_i at points $\boldsymbol{x}_i$. The coefficients of $P_h(\boldsymbol{x})$ determine, or are determined by, the boundary conditions, depending upon your point of view. Equation (12), together with adequate BCs, renders the problem positive definite. Apart from the question of BCs, this problem is equivalent to the Kriging (or RBF) problem. In fact, this is potentially an interesting route to a proof of the existence of unique and stable solutions to the latter problems (at least for certain ϕ and $P_h(\boldsymbol{x})$). Thus it would seem plausible that the Green's functions of invertible PDE problems are admissible as RBFs. We have *not* shown that an L exists for arbitrary ϕ, and indeed such a statement need not be true. However, the freedom in choosing boundary conditions for the Green's function problem, together with the existence of formal inverses defined in terms of Fourier methods, can take one a lot further in this direction than one would initially think possible.

Now suppose in the following that L *is* a differential operator. Then we might consider discretising the system (12) via a Finite Difference approximation of L (more generally, one could include quadrature for an integro-differential L). In this discretisation, wherever $u(\boldsymbol{x}_i)$ is known, f_i is unknown and vice versa. So the discretisation of equation (12) needs re-arrangement such that all unknowns appear on the left hand side, and all knowns on the right hand side.

Next, one imposes boundary conditions (BCs) which implicitly determine $P_h(\boldsymbol{x})$. An example, relevant to ore grade estimation, might be to require that $u(\boldsymbol{x}) = 0$ on a boundary sufficiently distant from the region of interest.

Numerics

We have just shown that certain simple Kriging, dual Kriging and RBF interpolations are equivalent to solving a PDE subject to boundary conditions. We can now employ one of the fast PDE solvers in order to compute an interpolation.

We solve the discretised version of (12) using the full multigrid finite difference method. As discussed in (for example) Press, et al. (1992), this technique solves the problem on a grid of M points in

$$O(M) \qquad (13)$$

operations.

This is independent of the number of control points, and in fact, the method actually converges faster with increasing density of control points. Compare the scaling behaviour (13) with those of the more direct formulations of Kriging or RBF interpolations: expressions (5), (6), and (7).

As a practical matter, one commonly knows only ϕ or only L directly. Determining the missing member of the pair is straightforward upon consideration of the Fourier Transform of equation (10) (although the resulting convolution operator may be unstable).

Our current multigrid implementation has not yet reached the full potential speed of $O(M)$, due to practical difficulties involving the representation of control points on grids of different scales. The scaling behaviour of our current implementation is unquestionably better than $O(M^2)$ and probably better than $O(M \log M)$. We regard this as a difficulty with our current algorithmic strategy rather than a fundamental problem with a multigrid approach to interpolation.

Example

As a concrete example of this technique, if $L \equiv \nabla^4$, $u = u(\boldsymbol{x})$ is interpreted as displacement, and $f = f(\boldsymbol{x})$ as force $(\boldsymbol{x} \in \Re^2)$, then equation (12) would be the PDE for a two-dimensional thin plate spline (the minimum curvature interpolation of Briggs (1974), common in geophysical applications). Higher dimensional analogues of thin plate splines would simply let $(x \in \Re^n)$, and define ∇^4 appropriately. Interpolations which are stiffer than thin plate splines can be found from considering higher powers of ∇ (e.g. the spherical model variogram).

DISCUSSION

This general approach has the obvious capability to put smooth interpolation onto a physical basis. If an interpolation is required to satisfy (12) where L is known from physical arguments, a direct application of our technique will yield a physically realistic solution. However, beyond this, it allows some Kriging interpolations to be calculated quickly using PDE solution methods.

Kriging methods form the basis of modern conditional simulation techniques, which seek to characterize the variability of the predicted quantity in addition to its mean value. It is interesting to contemplate the relationship between these conditional simulations, and the PDE formulation. If there is an equivalent PDE formulation, what might it look like? A Gibbs spatial process perhaps?

Since multigrid is an iterative scheme, given a closed

form expression for L (and its finite difference representation) we believe that there is no fundamental reason to restrict application of this technique to linear L. For some classes of nonlinear PDEs, the only challenge to this general scheme should be algorithmic details, and a possibility of scaling performance that is slightly worse than $O(M)$.

We close with the following remark: since some Kriging problems can clearly be solved by solving PDEs, we suspect that there are many potential synergies to be found in applications (such as flow field simulation in petroleum engineering) where the output from an interpolation specifies the geometry for further physical simulation. Both aspects of the problem could now be solved from within the same software base.

APPENDIX 1

RBF Equivalent of Ordinary Kriging

The linear system for simple Kriging arises from a minimization problem. The conditions that the variance is minimized with respect to the unknown parameters w_i, are expressed as equation (2). For ordinary Kriging, to impose an unbiased solution, equation (2) must be solved with the constraint

$$\sum w_i = 1, \tag{A1}$$

which also ensures that a constant function is interpolated exactly. Together, equations (2) and (A1) are solved via the method of Lagrange multipliers, yielding a new augmented system (in partitioned matrix format)

$$\begin{bmatrix} \boldsymbol{C} & \mathbf{1} \\ \mathbf{1}^\dagger & 0 \end{bmatrix} \begin{bmatrix} \boldsymbol{w}_0 \\ \mu \end{bmatrix} = \begin{bmatrix} \boldsymbol{b}_0 \\ 1 \end{bmatrix}. \tag{A2}$$

We rewrite equation (A2) as

$$\hat{\boldsymbol{C}}\hat{\boldsymbol{w}}_0 = \hat{\boldsymbol{b}}_0. \tag{A3}$$

Now, in the Kriging formulation, the value of the interpolation at a point $\boldsymbol{x}_0$ is expressed as

$$u(\boldsymbol{x}_0) = (\boldsymbol{w}_0, \boldsymbol{u})_N = (\hat{\boldsymbol{w}}_0, \hat{\boldsymbol{u}})_{N+1}. \tag{A4}$$

Here, the inner (dot) product between N dimensional vectors α, β is denoted by $(\alpha,\beta)_N$. The augmented vector $\hat{\boldsymbol{u}}$ must be constructed such that the equality between the left two terms of equation (A4) is preserved. The only augmentation that preserves that equality is

$$\hat{\boldsymbol{u}} = \begin{bmatrix} \boldsymbol{u} \\ 0 \end{bmatrix}. \tag{A5}$$

Given this, we now derive a radial basis function equivalent to the ordinary Kriging system. Start with the ordinary Kriging expression for the value at some interpolation point

$$u(\boldsymbol{x}_0) = (\hat{\boldsymbol{w}}_0, \hat{\boldsymbol{u}})_{N+1}. \tag{A6}$$

Since we have assumed $\hat{\boldsymbol{C}}$ is invertible, and given that it is also symmetric, it follows that there exists an $\hat{\boldsymbol{f}}$ such that

$$\hat{\boldsymbol{C}}^\dagger \hat{\boldsymbol{f}} = \hat{\boldsymbol{u}}. \tag{A7}$$

These are the radial basis function equations augmented by a constant offset. This corresponds to the case of Equation (8) with $P_h(\boldsymbol{x}) = const$. This augmented RBF representation can be evaluated at $\boldsymbol{x}_0$ to give

$$u(\boldsymbol{x}_0) = (\hat{\boldsymbol{b}}_0, \hat{\boldsymbol{f}})_{N+1}. \tag{A8}$$

But by (A3) we know that this can be rewritten as

$$u(\boldsymbol{x}_0) = (\hat{\boldsymbol{C}}\hat{\boldsymbol{w}}_0, \hat{\boldsymbol{f}})_{N+1}. \tag{A9}$$

Rewriting (A9) as a matrix manipulation, regrouping, and transposing, we find

$$u(\boldsymbol{x}_0) = (\hat{\boldsymbol{w}}_0, \hat{\boldsymbol{C}}^\dagger \hat{\boldsymbol{f}})_{N+1}. \tag{A10}$$

Substituting from (A7) it is found

$$u(\boldsymbol{x}_0) = (\hat{\boldsymbol{w}}_0, \hat{\boldsymbol{u}})_{N+1}. \tag{A11}$$

But this is exactly the ordinary Kriging solution for $\boldsymbol{x}_0$. Therefore, if there is a solution to the augmented RBF defined by (A7) then it is equivalent to the ordinary Kriging solution and vice versa.

Extension for detrending with constrained higher order moment

Consider an interpolation of a scalar function of $\Re^2$. The augmented RBF system which includes linear functions of the coordinates in $P_h(\boldsymbol{x})$, will interpolate any linear function exactly. The equations determining the $\boldsymbol{f}_i$ and the three coefficients of $P_h(\boldsymbol{x})$ are

$$\begin{bmatrix} \boldsymbol{C} & \boldsymbol{c}_1 & \boldsymbol{c}_2 & \mathbf{1} \\ \boldsymbol{c}_1^{\dagger} & 0 & 0 & 0 \\ \boldsymbol{c}_2^{\dagger} & 0 & 0 & 0 \\ \mathbf{1}^{\dagger} & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} \boldsymbol{f} \\ \sigma_1 \\ \sigma_2 \\ \mu \end{bmatrix} = \begin{bmatrix} \boldsymbol{u} \\ 0 \\ 0 \\ 0 \end{bmatrix}, \qquad \text{(A12)}$$

where $\boldsymbol{c}_j$ is a vector of the jth coordinates of the set of control points, $\{\boldsymbol{x}_i^{(j)}\}_{i=1\ldots N}$. (Here, $\boldsymbol{x}_i^{(j)}$ denotes the jth coordinate at sample point i.) Equation (A12) adds further constraints to the first order moments of the "forces" f_i, these being $\sum_{i=1}^{N} f_i \cdot x_i^{(1)} = 0$ and $\sum_{i=1}^{N} f_i \cdot x_i^{(2)} = 0$. This is like linearly detrending the data and kriging simultaneously. The resulting interpolation is

$$u(\boldsymbol{x}_0) = (\hat{\boldsymbol{b}}_0, \hat{\boldsymbol{f}})_{N+1}, \qquad \text{(A13)}$$

where $\hat{\boldsymbol{b}}_0 = \left[\boldsymbol{b}_0 \; x_0^{(1)} \; x_0^{(2)} \; 1\right]^{\dagger}$. Clearly, higher orders may be handled similarly, so we have the dual Kriging/RBF analogue to universal Kriging/Kriging with a trend.

ACKNOWLEDGEMENT

This paper is published with the permission of the Director of the Australian Geodynamics Cooperative Research Centre.

REFERENCES

Briggs, I.C., 1974 "Machine Contouring Using Minimum Curvature," Geophysics, v. 39, no. 1, pp. 39-48.

Isaaks, E.H. and Srivastava, R.M., 1989 "An Introduction to Applied Geostatistics," Oxford University Press, New York.

Matheron, G., 1980 "Splines and Kriging - Their Formal Equivalence," in Down-to-Earth Statistics - Solutions looking for geological problems, ed. D.F. Merriam, pp. 77-95, Syracuse University Geology Contributions.

Matheron, G., 1982 "Pour une analyse krigeante des données régionalisées," Note interne, n°732, CGMM, Fontainebleau, 20 p.

Powell, M.J.D., 1992 "The Theory of Radial Basis Function Approximation in 1990", in Advances in Numerical Analysis, Vol. 2: Wavelets, Subdivision Algorithms, and Radial Basis Functions, ed. W.A. Light, pp. 105-210, Oxford University Press, New York.

Press, W.H., Teukolsky, S.A., Vetterling, W.T. and Flannery, B.P., 1992 "Numerical Recipes in C; The Art of Scientific Computing," Cambridge University Press, Cambridge.

Royer, J.J., and Vieira, P.C., 1984 "Dual Formalism of Kriging," in Geostatistics for Natural Resources Characterization, Part 2, ed. G. Verly, M.David, A.G. Journel, and A. Marechal, NATO ASI Series C: Mathematical and Physical Sciences, Vol. 122, pp 691-702, D. Reidel Pub. Co., Dordrecht.