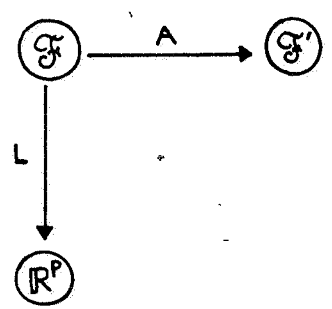
SPLINES AND KRIGING: THEIR FORMAL EQUIVALENCE

G. Matheron

Ecole des Mines, Fontainebleau

The aim of this paper is to demonstrate that two methods of interpolation (kriging and spline functions) are equivalent. They are equivalent in the sense that any fitted curve obtained using the spline functions can be identified with a fit obtained using kriging and vice versa. Although the first of these two problems (finding the kriging system which is equivalent to a given spline function) is relatively easy to solve, the second is more difficult. To make it easier to follow the proof, the problem will first be presented using the index notation traditionally used in geostatistics; and then the proof will be repeated in the more abstract algebraic terminology used in the theory of spline functions.

DEFINITION OF THE SPLINE PROBLEM



Suppose that we have two Hilbert spaces (F and F') and a continuous linear transformation A which maps F onto F'. (In practice, F and F' are two spaces of functions.)

Suppose we also have a continuous linear transformation L which maps F into $\mathbb{R}^p$. In practice, L is a family of N linear functionals $L_\alpha (\alpha = 1, 2, \ldots, p)$ which are continuous in F.

The usual problem of fitting a spline function can be considered in the following terms:

· find a function $f \in F$ which minimizes $\|Af\|^2$ under the conditions: $L_\alpha(f) = f_\alpha$ (the f_α are given numbers).

Let N denote the kernel of A and N_L, that of L. If we suppose that

$$N \cap N_L = 0 \qquad (1)$$

then the problem has a unique solution. We now show that this is identical to the interpolation function obtained from a suitably chosen kriging system.

To simplify the proof, we shall develop a second norm $\| \|_1$, which is equivalent in F to the original one. Let $N^\perp$ be the subspace of F which is orthogonal to the kernel N of A. The restriction from A to $N^\perp$ then is a continuous one-to-one mapping from $N^\perp$ onto F'. From the general theory of Hilbert spaces, we know that the inverse mapping (from F' onto $N^\perp$) also is continuous, so we are concerned with a homeomorphism.

Consequently, if we define a new norm $\| \|_1$ for F as follows:

$$\|f\|_1^2 = \|\Pi_N f\|^2 + \|A\Pi_N^\perp f\|^2$$

where f is any function in F, and where Π_N and $\Pi_N{}^\perp$ are the mappings which project F into N and $N^\perp$ under the original norm. This new norm does not change the topology of F. In particular, the mappings A and L are yet continuous, and the subspaces N and $N^\perp$ (and also their projections) remain unchanged.

With this new metric, F' can be identified with $N^\perp$, and similarly, the mapping A and the projections mapping $\Pi_N{}^\perp$.

Because we shall only use the new metric $\|f\|_1$ from now on, we can drop the subscript 1. So, the new metric now will be denoted by $\|f\|$.

On the other hand, each of the continuous linear functionals L_α can be identified with the (uniquely determined) function $L_\alpha \in F$ such that:

$$L_\alpha(f) = \langle L_\alpha\ f \rangle \quad \forall f \in F \qquad (2)$$

Let S denote the subspace of F of dimension $p < \infty$, generated by these functions $L_\alpha \in F$.

More generally, we associate a continuous linear functional (that is, an element L_x of F) with each point x belonging to a space E. We now suppose that the L_x span F: in practice, F is a space of functions on E and the L_x, $x \in E$ are defined by:

$$f(x) = \langle L_x, f \rangle \quad f \in F \qquad (3)$$

In this situation, the L_α are the forms L_{x_α} associated with the experimental points x_α. Note that the metric on F must be strong enough so that the convergence of $f_n \to f$ in F implies the pointwise convergence $f_n(x) \to f(x)$ for all $x \in E$. So it must be stronger than that of a space L^2 on E.

After these changes, the problem can be expressed in the following terms:

· find $f \in F$ which minimizes $\|\Pi\ f\|$ under the condition $\Pi_S\ f = f_S$ at the data points $N^\perp$ (which comes back to $\langle L_\alpha , f \rangle = f_\alpha$ at the data points).

THE EQUATIONS FOR THE SPLINE INTERPOLATION PROBLEM

The kernel N_L of the mapping $L = (L_\alpha)$ is just the subspace $S^\perp$ of F which is orthogonal to the space S spanned by the $L_\alpha \in F$. The condition $N \cap N_L = 0$ can then be written as:

$$N \cap S^\perp = 0 \qquad (4)$$

Thus, for all $f \in N$, the relation $\Pi_S\ f = 0$ implies that $f = 0$. In other words, the restriction from Π_S to N is one-to-one mapping from N into S. Because the dimension of S is finite, it is therefore greater than or equal to that of the kernel N:

$$\boxed{\text{Dim } N \leq \text{Dim } S}$$

Let (f^ℓ) be a basis for the kernel N ($\ell = 1,2,\ldots,k \leq \dim S$) and let L_α be the given basis of S. For all x, it is clear that

$$f^\ell(x) = \langle L_x , f^\ell \rangle$$

and in particular

$$f_\alpha^\ell \equiv f^\ell(x_\alpha) = \langle L_\alpha , f^\ell \rangle$$

The condition (4) can be written in analytical terms as:

$$C_\ell\ f_\alpha^\ell = 0 \quad (\alpha = 1,2,\ldots,k) \Longrightarrow C_\ell = 0 \qquad (4')$$

(in effect, $N \cap S^{\perp}$ is made up of functions of the form $C_\ell f^\ell$, which satisfy this condition). We recognize the usual condition seen in the theory of kriging (the linear independence of the basic functions f^ℓ on the set of experimental points).

Let us now try to solve the original problem (i.e. to find the function f which minimizes $\|\Pi_{N^\perp} f\|$ subject to the condition that $< L_\alpha , f > = f_\alpha$ at the data points). This function f is such that $\Pi_N{}^\perp$ is orthogonal to all $g \in F$ such that $< g, L_\alpha > = 0$. Let $B^{\alpha\beta}$ denote the inverse matrix of $< L_\alpha L_\beta >$. For all $g \in F$, the element

$$g - < g\, L_\beta > B^{\alpha\beta} L_\alpha$$

is orthogonal to S. Because $\Pi_N{}^\perp f$ is orthogonal to this element, we see that

$$< \Pi_{N^\perp} f , g > = B^{\alpha\beta} < \Pi_{N^\perp} f , L_\alpha > < g\, L_\beta > \quad g \in F$$

and hence:

$$\Pi_{N^\perp} f = B^{\alpha\beta} < \Pi_{N^\perp} f , L_\alpha > L_\beta$$

Thus $\Pi_{N^\perp} f$ must be of the form $b^\alpha L_\alpha$, with the coefficients b^α satisfying

$$b^\alpha < L_\alpha\, f^\ell > = 0 \qquad (5)$$

(because $\Pi_{N^\perp} f$ is orthogonal to N). The function f which is the sum of $\Pi_{N^\perp} f$ and an element of N, is of the form:

$$f = b^\alpha L_\alpha + C_\ell f^\ell$$

In addition to equation (5), the coefficients b^α and C_ℓ must satisfy

$$< f , L_\alpha > = f_\alpha$$

From this, we obtain the following system of equations:

$$f = b^\alpha L_\alpha + C_\ell f^\ell$$

$$b^\alpha f^\ell_\alpha = 0 \qquad (6)$$

$$< f , L_\alpha > = f_\alpha$$

where $f^\ell_\alpha = < L_\alpha\, f^\ell >$.

The value of the function at the points $x \in E$ is $f(x) = < f L_x >$. So we have:

$$f(x) = b^\alpha < L_\alpha L_x > + C_\ell f^\ell(x)$$

$$b^\alpha f^\ell_\alpha = 0 \qquad (7)$$

$$f(x_\alpha) = f_\alpha$$

When we look at these equations more closely, we see that they are just the universal kriging equations. In fact, if we put

$$\sigma(x,y) = < L_x L_y > \quad (x,y) \in E \qquad (7')$$

$\sigma(x,y)$ is a covariance, and there is a random function $Z(x)$ in the space E which satisfies

$$< Z_x Z_y > = \sigma(x,y)$$

Let $Z^*(x)$ denote the universal kriging of $Z(x)$ in terms of the variables $Z_\alpha = Z(x_\alpha)$ in the presence of a drift $m(x) = a_\ell f^\ell(x)$. The kriging equations which give the estimated (i.e. interpolated) value $z^*(x)$ are

$$z^*(x) = b^\alpha \sigma_{\alpha x} + C_\ell f^\ell(x)$$

$$b^\alpha f^\ell_\alpha = 0 \qquad (8)$$

$$z^*(x_\alpha) = z_\alpha$$

This system is equivalent to (but not identical with) the usual one. It characterizes the kriging from the point of view of interpolation (Matheron, 1970).

Thus for $z_\alpha = f_\alpha$, this system is clearly identical to (7). In the same manner that (7') suffices to determine z^*, (7) characterizes f, and so we have:

$$z^* = f$$

We have thus shown that any spline function is equivalent to a function obtained by kriging. We now go on to show the converse.

THE CONVERSE

Let Z_x, $x \in E$ be a random function. Let H be the space generated by Z_x. Let $\sigma(x,y) = < Z_x , Z_y >$. We now consider the problem of kriging Z_x in the presence of the drift terms $a_\ell f^\ell(x)$, given the values of $Z_\alpha = x_\alpha$.

To express this problem in the form of a spline interpolation, we have to determine elements $Y^\ell \in H$ such that:

$$f^\ell(x) = < Y^\ell Z_x > \quad (x \in E) \tag{9}$$

A problem arises at this point, because these elements Y^ℓ need not exist in our initial space H (for example, if Z(x) is stationary, the functions of the form $x \to <Y,Z_x>$ are bounded because $|<Y,Z_x>| \leq ||Y|| x ||Z_x|| = \text{Constant}$).

We need to know whether, for a given function f, there exists a $Y_f \in E$ such that for all $x \in E$

$$f(x) = < Y_f \; Z(x) >$$

The necessary and sufficient condition for this is that there exists a $B < \infty$ such that:

$$(\Sigma \lambda_i \; f(x_i))^2 \leq B \sum_{i,j} \lambda_i \lambda_j \sigma(x_i, x_j) \tag{10}$$

for all finite linear combinations.

We are going to modify our original random functions Z(x) and also the space H, so that this condition is satisfied. We know that the results of the kriging will be the same if we replace Z(x) by $\tilde{Z}(x) = Z(x) + B_\ell \; f^\ell(x)$ where B_ℓ are any arbitrary random variables. For example, we can take the B_ℓ to be linearly independent and orthogonal to the Z(x). The covariance of the new random functions $\tilde{Z}(x)$ is then

$$\tilde{\sigma}(x,y) = \sigma(x,y) + K_{\ell s} \; f^\ell_x \; f^s_y$$

where $K_{\ell s} = < B_\ell, B_s >$ is a strictly positive definite matrix. It is clear that

$$(\sum_i \lambda^i \; f^\ell(x_i))^2 \leq \sum_s (\sum_i \lambda^i \; f^s(x_i))^2 = \sum_{i,j} \lambda^i \lambda^j \delta_{\ell s} \; f^\ell_{x_i} \; f^s_{x_j}$$

$$\leq \frac{1}{a} \sum_{i,j} \lambda^i \lambda^j K_{\ell s} \; f^\ell_{x_i} \; f^s_{x_j}$$

where a is the smallest eigenvalue of $K_{\ell s}$. Consequently, the condition (10) is satisfied for $\tilde{\sigma}(x,y)$.

So, for the rest of the proof, we can assume that (10) is satisfied, and, for simplicity, we shall write Z(x), $\sigma(x,y)$ instead of $\tilde{Z}(x)$ and $\tilde{\sigma}(x,y)$.

Let $Y^\ell \in H$ be the elements satisfying (9). We now are going to construct a Hilbert space F (of functions) which is isomorphic to H and which contains the functions f^ℓ.

To do this, we associate the function f_Y defined by

$$f_Y(x) = < Y, Z(x) >$$

with each $Y \in H$, and we define

$$\| f_Y \| = \| Y \| \tag{11}$$

It is clear that the space built in this manner has the desired attributes (that is, condition (10) is satisfied). Once the results obtained earlier are applied in this space, it becomes evident that the spline interpolation in F is equivalent to the interpolation given by kriging in H. The $L_\alpha \in F$ are the functions defined by

$$L_\alpha(x) = < Z_\alpha, Z_x > = \sigma_{\alpha x}$$

or, more generally

$$L_x(y) = < Z_x, Z_y > = \sigma_{xy}$$

NOTE - If N = 0 (that is, if the mapping A: $F \to F'$ in the original problem is bijective and continuous (that is bicontinuous)), we evidently come back to the situation of simple kriging. Thus, the element f minimizing $\|f\|^2$ under the condition $<f, L_\alpha> = f_\alpha$ is evidently the element of S satisfying this condition; that is

$$f = B^{\alpha\beta} \; f_\alpha \; L_\beta$$

In particular, for all $x \in E$:

$$f(x) = B^{\alpha\beta} \; f_\alpha < L_\beta \; L_x >$$

which is obviously the equation for a simple kriging considered as an interpolation.

ALGEBRAIC POINT OF VIEW

Because the spaces H and F are isomorphic, it does not matter whether we work with the functions f or with the random variables in H. The isomorphism

$$\Phi : H \to F$$

is defined as follows:

$$\Phi_Y(x) = \langle Y, Z_x \rangle \quad (Y \in E,\ x \in E) \qquad (12)$$

$$\|\Phi_Y\| = \|Y\|$$

For example, let us work in H. N is the space generated by the Y^ℓ; S is the space generated by the Z_α. We put

$$V_0 = S \cap N^\perp \qquad (13)$$

$$N^* = V_0^\perp \cap S = \Pi_S N$$

So V_0 is the space of admissible linear combinations $\lambda^\alpha Z_\alpha$ (that is, those satisfying $\lambda^\alpha f_\alpha^\ell = 0$). As N^* is the projection in S of the space N generated by the Y^ℓ, it is the space generated by the optimal estimators A_ℓ of the drift. Writing S as the direct sum

$$S = V_0 \oplus N^* \qquad (14)$$

simply shows that any $\lambda^\alpha Z_\alpha \in S$ is the sum of two orthogonal terms: a drift term $\lambda^\alpha A_\ell f_\alpha^\ell$, which belongs to N^*, and a residual $\lambda^\alpha (Z_\alpha - A_\ell f_\alpha^\ell)$ which belongs to V_0.

Similarly for S, we introduce the space S' defined as the direct sum of two orthogonal subspaces N and V_0:

$$S' = V_0 \oplus N \qquad (14')$$

Because V_0 is the same in both expressions, the duality between kriging and spline interpolation is obvious once we interchange N and N*, and S and S'.

The Operators Λ and Λ^*

Let R be the restriction of the projection mapping Π_S of the space S to the space $S' = V_0 \oplus N$. Its adjoint R* is clearly

the restriction of the projection mapping $\Pi_{S'}$ of S' to the space $S = V_0 \oplus N^*$. We now show that the operators have inverses. Because the proofs are the same for R and R*, we need only demonstrate one of the two, R. Any $Y \in S'$ is the sum of an element belonging to V_0 and to one from N:

$$Y = \Pi_{V_0} Y + \Pi_N Y$$

The first term is invariant under the mapping R. Because $\Pi_{V_0} Y$ already belongs to S

$$R\,\Pi_{V_0} Y = \Pi_S \Pi_{V_0} Y = \Pi_{V_0} Y$$

The second term:

$$R\,\Pi_N Y = \Pi_S \Pi_N Y$$

belongs to $\Pi_S N = N^*$; it therefore is orthogonal to the first. Consequently $R\,Y = 0$ implies that $\Pi_{V_0} Y = 0$ and $\Pi_S \Pi_N Y$. The first relation ($\Pi_{V_0} Y = 0$) indicates that $Y \in N$, because Y is in $S' = V_0 \oplus N$. The second then can be rewritten as $\Pi_S Y = 0$. But under our hypothesis that $N \cap S = 0$, the fact that $Y \in N$ and that $\Pi_S Y = 0$ implies that $Y = 0$. Hence R is injective. Moreover, as we have seen, the image space is the direct sum of V_0 and N:

$$R\,S' = \Pi_S S' = V_0 \oplus \Pi_S N = V_0 \oplus N^*$$

Then $R\,S' = S$ and so R is surjective.

Thus, R is bijective. Let Λ_0^* be its inverse: similarly, R* is bijective and has an inverse, the dual of Λ_0^* which we shall denote by Λ_0.

We are now going to see that Λ_0 is the operator associated with kriging and that Λ_0^* is the operator associated with spline interpolation.

Characterization of Λ and Λ^*

Λ_0 maps S' onto S. So, for all $Z' \in S'$, $\Lambda_0 Z' = Z$ is the only element of S having Z' as its projection in S'. Similarly, for all $Z \in S$, $\Lambda_0^* Z = Z'$ is the only element of S' such that $\Pi_S Z' = Z$.

We extend Λ_0 and Λ_0^* over the whole of H by putting

$$\Lambda = \Lambda_0 \Pi_{S'} \quad , \quad \Lambda^* = \Lambda_0^* \Pi_S$$

Thus (see Fig. 1):

for all $Z \in H$, ΛZ is the only element of S having the same projection in S' as Z, and
$\Lambda^* Z$ is the only element of S' having the same projection in S as Z.

(Note that $\Lambda = \Pi_S \Lambda_0 \Pi_{S'}'$, has $\Pi_{S'}' \Lambda_0^* \Pi_S = \Lambda_0^* \Pi_S$ as its adjoint, which is just the operator that we have termed Λ^*).

Under these conditions, for all $Z \in H$, ΛZ is the kriged estimate of Z (using the sample points Z_α and in the presence of a drift belonging to $\Phi(N)$, where Φ is the isomorphism $H \to F$).

In fact, let Z^* be the kriged estimate of Z; that is the unique element of S which minimizes $\| Z - Z^* \|$ under the constraints:

$$Z^* \in S$$
$$< Z - Z^*, Y > = 0 \quad Y \in N$$

This element Z^* is characterized by the additional condition: $Z - Z^*$ must belong to the space orthogonal to $S \cap N^\perp = V_0$. After taking the condition $< Z - Z^*, Y > = 0$ for all $Y \in N$ into account, we see that Z^* is characterized by the following two conditions:

$$Z^* \in S$$
$$< Z - Z^*, Y > = 0 \quad \forall Y \in N \oplus V_0 = S'$$

Now it is clear that ΛZ satisfies these conditions. For a start, we know that $\Lambda Z \in S$ by construction. Now if $Y \in S'$, we determine that

$$< \Pi_{S'}(Z - \Lambda Z), Y > = < Z - \Lambda Z, Y >$$

But, from the definition of Λ, $\Pi_{S'} \Lambda Z = \Pi_{S'} Z$. And so we have that

$$< Z - \Lambda Z, Y > = 0$$

We now consider the dual problem. We shall show that Λ^* is the operator associated with the spline interpolation. To be able to work in H, we consider the isomorphism $F \to H$. To find the

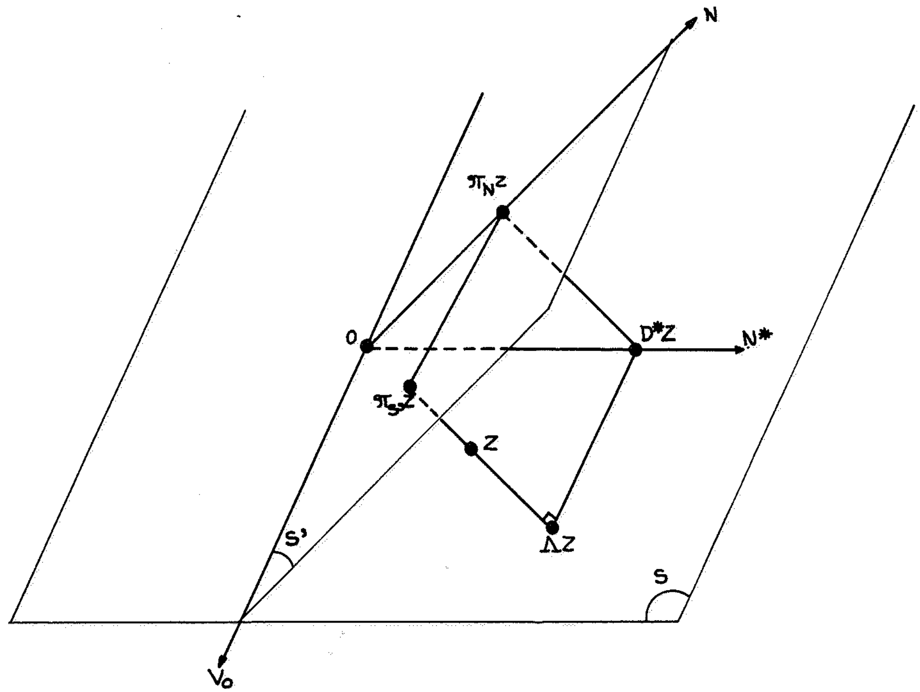


Figure 1. Characterization of Λ and Λ^*.

spline interpolator, we have to determine the $Y \in H$ which minimizes $\| \Pi_{N^\perp} Y \|$ and satisfies the conditions

$$< Y Z_\alpha > = f_\alpha$$

Let Y_f be an arbitrary element satisfying these conditions: for example the element with the minimal norm:

$$Y_f = B^{\alpha\beta} f_\alpha Z_\alpha$$

($B^{\alpha\beta}$ is the inverse of the matrix $\sigma_{\alpha\beta}$).

The unknown element Y must have the same projection in S as Y_f does. $\Pi_{N^\perp} Y$ must be orthogonal to $S^\perp$, and therefore belong to $S \cap N^\perp = V_0$. Thus $Y \in V_0 \oplus N = S'$. It is therefore the unique element in S' having the same projection in S as Y_f; that is, it must be $\Lambda^* Y_f$.

In other words, the function $f^* \in F$ which interpolates between the $f_\alpha = \langle Y_f , Z_\alpha \rangle$ and has the minimum norm, is therefore the image of this element under the isomorphism Φ, that is

$$f^* = \Phi(\Lambda^* Y_f)$$

The value of f^* in x is $f^*(x) = \langle f^* \Phi(Z_x) \rangle$. So after taking account of our isomorphism

$$f^*(x) = \langle \Lambda^* Y_f , Z_x \rangle \qquad (15)$$

But Λ^* is the dual of Λ. So we also have that

$$f^*(x) = \langle Y_f , \Lambda Z_x \rangle \qquad (15')$$

Note that ΛZ_x is an element of S, of the form $\lambda^\alpha(x) Z_\alpha$. We therefore see that $f^*(x) = \lambda^\alpha(x) \langle Y_f Z_\alpha \rangle = \lambda^\alpha(x) f_\alpha$.

ESTIMATING THE DRIFT

For all $Z \in H$, $\Pi_N Z$ represents the drift of this element and $\Lambda \Pi_N Z$ the optimal estimator of this drift. We therefore put

$$\boxed{D^* = \Lambda \Pi_N} \qquad (16)$$

From the characterization of the kriging operator Λ, $D^* Z \in N^*$ is the only element of S having the same projection in S' as the element $\Pi_N Z$. Because $\Pi_N Z$ is already in $N \subset S' = N \oplus V_0$, we see that:

$D^* Z$ is the only element of N^* having a projection in N equal to $\Pi_N Z$.

What is the relationship between the operators Λ and D^*? For a start, we note the following two relations:

$$\Pi_{V_0} = \Lambda \Pi_{V_0} = (1-D^*) \Pi_S \qquad (17)$$

The first relation is a simple consequence of the fact that V_0 is invariant under Λ. More precisely, we can write:

$$\Lambda = \Lambda \Pi_{S'} = \Lambda (\Pi_{V_0} + \Pi_N) = \Pi_{V_0} + \Lambda \Pi_N$$

because S' is the direct sum of V_0 and N. From (16), it follows that

$$\Lambda = \Pi_{V_0} + D^* \qquad (18)$$

At the same time, the operator D^* coincides with N^* within S. (In effect, if $Z \in S$, $\Pi_N Z$ is equal to the projection of the element $\Pi_{N^*} Z$ on N, hence the result that $D^* Z = \Pi_{N^*} Z$). We therefore have that

$$(I - D^*) \Pi_S = (I - \Pi_{N^*}) \Pi_S = \Pi_{V_0}$$

which is just the second relation (17). When this result is substituted into (18), we obtain:

$$\boxed{\Lambda + (I - D^*) \Pi_S + D^*} \qquad (19)$$

Here we recognize the well-known additivity theorem.

What does the adjoint of D^*, that is, the operator

$$D = \Pi_N \Lambda^* \qquad (16')$$

represent?

It is easy to see, by duality, that D Z is the unique element of N with a projection on N^* equal to $\Pi_{N^*} Z$.

If we repeat the preceding argument replacing N and S by N^* and S', we see that:

$$\Lambda^* = \Pi_{V_0} + D = (I - D) \Pi_{S'} + D \qquad (19')$$

Two terms, $D Y_f$ and $\Pi_{V_0} Y_f = (I-D) \Pi_{S'} Y_f$ occur in the expression $\Lambda^* Y_f = Y_{f^*}$ for the spline interpolator. The first of these is the component of Y_{f^*} in N (i.e. the term $C_\ell f^\ell$ in the formula (5)). The second is the component of Y_{f^*} in V_0 (that is, the term $b^\alpha L_\alpha$ with $b^\alpha f^\ell_\alpha = 0$ in the second formula).

COKRIGING AND SMOOTHING SPLINES

We have seen that spline interpolation is equivalent to kriging. We now shall go on using a similar line of reasoning to show that smoothing spline functions are equivalent to a particular type of cokriging (filtering with error). However, the converse is not true this time.

The problem of fitting spline functions is as follows: using the same notation as before, we have a self-adjoint operator T which is strictly positive on S (i.e. for all $f \in H$, the relation $< f, T\,\Pi_S f > = 0$ implies that $\Pi_S f = 0$). Given S^f (i.e. the $f_\alpha = < f\, L_\alpha >$), we wish to determine the function $f^* \in F$ which minimizes:

$$\|\Pi_{N^\perp} f^*\|^2 + < f^* - f, T\,\Pi_S (f^* - f) > \qquad (20)$$

This element f therefore must satisfy the relation

$$\Pi_{N^\perp} f^* + T\,\Pi_S f^* = T\,\Pi_S f \qquad (21)$$

But (20) implies that $\Pi_{N^\perp} f^* \in S$, and hence that $\Pi_{N^\perp} f^* \in S \cap N^\perp = V_0$. Consequently, $f^* \in V_0 \oplus N = S'$. For $f^* \in S'$, we already have that $\Pi_{N^\perp} f^* = \Pi_{V_0} f^* = \Pi_{N^\perp} \Pi_S f^*$. So this relation (21) is therefore equivalent to

$$f^* \in S'$$

$$(\Pi_{V_0} + T)\,\Pi_S f^* = T\,\Pi_S f \qquad (22)$$

The latter of these two relations uniquely determines $\Pi_S f^*$ because T and even more so $\Pi_{V_0} + T$ is strictly positive on S.

In addition, any f^* in S' is determined uniquely if its projection $\Pi_S f^*$ in S is known. In effect, we have $f^* = \Lambda^* \Pi_S f^*$. Therefore, the relations (22) and also (21) have a unique solution when $\Pi_S f$ is given.

The simplest way to show that the system (22) is equivalent to cokriging is to use index notation.

If f is an element of F, T f will be of the form:

$$T\,\Pi_S f = T^{\alpha\beta} < f\, L_\alpha > L_\beta$$

where the matrix $T^{\alpha\beta}$ is strictly positive. Analytically, the system can be written as follows:

The first condition: $f^* \in S' = V_0 \oplus N$ is equivalent to

$$f^* = b^\alpha L_\alpha + c_\ell f^\ell$$

$$b^\alpha f^\ell_\alpha = 0$$

As $\Pi_{N^\perp}$ is identical to Π_{V_0} on S', we see that $\Pi_{N^\perp} f^* = b^\alpha L_\alpha$ (from the condition on $b^\alpha f^\ell_\alpha$). So the second equation (22) can be written as

$$b^\beta L_\beta + L_\beta T^{\alpha\beta} < L_\alpha , b^\gamma L_\gamma + c_\ell f^\ell >$$

$$= L_\beta T^{\alpha\beta} < L_\alpha f >$$

In other words, we have

$$b^\beta + b^\gamma T^{\alpha\beta} < L_\alpha L_\gamma > + c_\ell T^{\alpha\beta} f^\ell_\alpha = T^{\alpha\beta} f_\alpha$$

If we put $\sigma_{\alpha\gamma} = < L_\alpha L_\gamma >$, and if we let $S_{\alpha\beta}$ denote the inverse to the matrix T, the condition then becomes

$$b^\beta (S_{\alpha\beta} + \sigma_{\alpha\beta}) + c_\ell f^\ell_\alpha = f_\alpha$$

All in all, the function f is given by the system:

$$f^* = b^\alpha L_\alpha + c_\ell f^\ell$$

$$b^\alpha f^\ell_\alpha = 0$$

$$b^\beta (S_{\alpha\beta} + \sigma_{\alpha\beta}) + c_\ell f^\ell_\alpha = f_\alpha \qquad (23)$$

Now this system characterizes a particular sort of cokriging considered as an interpolation.

Let Y(x) be a random function and let σ_{xy} be its covariance. Suppose that we have

$$Z_\alpha = Y_\alpha + \varepsilon_\alpha$$

where the ε_α are "errors" which are orthogonal to the random function and which satisfy

$$E(\varepsilon_\alpha) = 0 \quad , < \varepsilon_\alpha , \varepsilon_\beta > = S_{\alpha\beta}$$

We now consider the cokriging of Y(x) from the data Z_α. The cokriging estimator is $Y^*(x) = \lambda^\alpha(x)\, Z_\alpha$ where the weighting factors $\lambda^\alpha(x)$ are given by the system

$$\lambda^{\alpha}(x)\ (S_{\alpha\beta} + \sigma_{\alpha\beta}) = \sigma_{\beta x} + \mu_{\ell}(x)\ f^{\ell}_{\beta}$$

$$\lambda^{\alpha} f^{\ell}_{\alpha} = f^{\ell}_{x} \qquad (24)$$

By solving this system explicitly and comparing the solution with (23), we can show that $Y^*(x)$ is necessarily equal to $f^*(x)$ (after replacing Z_α by f_α). However, it is simpler to put $\varepsilon_x = 0$ for $x \neq x_\alpha$ and to note that for all x except the experimental points x_α, the cokriging estimate $Y^*(x)$ of $Y(x)$ is identical to $Z^*(x)$, where $Z^*(x)$ is obtained by kriging the Z_α. By using the characterization of $Z^*(x)$ considered as an interpolator, we then obtain the following system:

$$Y^*(x) = b^{\alpha}\ \sigma_{\alpha x} + C_{\ell}\ f^{\ell}(x)$$

$$b^{\alpha}\ f^{\ell}_{\alpha} = 0$$

$$b^{\alpha}(S_{\alpha\beta} + \sigma_{\alpha\beta}) + C_{\ell}\ f^{\ell}_{\alpha} = \dot{z}_{\alpha} \qquad (25)$$

It is clear that $Z_\alpha = f_\alpha$, the solution f^* of (23) satisfies

$$f^*(x) = < f^*\ L_x > = y^*(x)$$

which establishes the correspondence between cokriging and smoothing spline functions.

We note that one advantage of cokriging is that the matrix $S_{\alpha\beta}$ represents the matrix of error covariance and in no way is arbritrary. In contrast to this, the choice of the matrix $T_{\alpha\beta}$ used in fitting the spline curves is arbitrary. In most situations, $T^{\alpha\beta}$ is taken to be a diagonal matrix, and so we seek to minimize

$$\|Af\|^2 + \Sigma\ W_{\alpha}\ (f^*_{\alpha} - f_{\alpha})^2$$

From this, we get the impression that the W_α must be taken inversely proportional to the error variances:

$$W_{\alpha} = \frac{C}{\|\varepsilon_{\alpha}\|^2}$$

But C remains completely arbitrary. In contrast to this, the choice of the covariance matrix is in no way arbitrary. In the situation where the ε_α are orthogonal, it is

$$S_{\alpha\beta} = \|\varepsilon_{\alpha}\|^2\ \delta_{\alpha\beta}$$

and any arbitrariness disappears.

SOME CONCLUDING REMARKS

From a purely formal point of view, we have demonstrated the equivalence between kriging and spline interpolation. However, in practice, the links between the two are not nearly so apparent. On one hand, the space H associated with a given space F is always a a space of random variables generated by a random function Z_x and having a covariance $\sigma(x,y) = < L_x, L_y >$ with reasonable properties. Looking at the problem from the opposite direction, the metric belonging to the space of functions F associated with a given random function Z_x generally is difficult to handle. In any situation, one thing is certain: if we limit our consideration to spline functions associated with metrics defined by differential operators, these correspond to a limited class of random functions. Only in rare situations can the metric associated with a given random function be defined in terms of differential operators. So in practice, kriging provides an interpolation method which is more general and more powerful than spline interpolation. Moreover, the normally used spline functions are just particular situations of kriging interpolators.

As an example, we shall consider a space F of functions on $\mathbf{R}^n$, which has a differential operator A having the required properties of positive definiteness. For the purposes of the demonstration, we shall limit ourselves to the situation where A is an iterated Laplacian operator:

$$A\ f = \Delta^p\ f$$

The space F cannot be the space L^2 (because A would not be continuous), nor can it be a subspace of L^2 equipped with a stronger norm of the type

$$\|f\|^2 + \|Af\|^2$$

because in this situation, the mapping of F into L^2 would not be subjective.

The technique here is to go from the domain D_A of the operator A considered as an operator on L^2 (A is closed but not continuous). We note that the kernel of A into L^2 is 0 (it is obvious from the Fourier transformation that the equation $\Delta^p f = 0$ has no solutions in L^2), consequently $\|A f\|$ is a norm on D_A. For F we take the Hilbert completion of D_A provided with the norm $\|A f\|$. This makes F isomorphic to the closure of A D_A in L^2.

We then can identify the elements $f \in F$ with (equivalence classes of) functions which are not, in general, in L^2. If $A = \Delta$, we take the functions whose increments of order 2 are in L^2, and for which Δf exists in the sense of L^2. (Δf occurs as the limit in L^2, of sequences $f_n = f * \lambda_n$, where the λ_n are measures filtering the polynomials of degree ≤ 1). For $A = \Delta^p$, F is constituted of functions f such that $f * \lambda \in L^2$ for all $\lambda \in \Lambda_{2p-1}$ (filtering the polynomials of degree $\leq 2p-1$) and such that Δ^p exists as the limit of a sequence $f * \lambda_n$ $\lambda_n \in \Lambda_{2p-1}$, in L^2.

For the norm $\| A f \|$, these equivalence classes of functions form a space F isomorphic to L^2. Once $A f = \Delta^p f$ is given, it is possible to reconstitute $f * \lambda$ for all $\lambda \in \Lambda_{2p-1}$, so that the function f is in fact defined up to a polynomial $C_\ell f^\ell(x)$ of order $\leq 2p-1$. As usual in the theory of IRK-k (Matheron, 1973), this indeterminancy is not particularly important, provided that we replace the functionals L_x by the linear combinations $\Sigma \lambda^i L_{x_i} = L_\lambda$ with $\lambda \in \Lambda_{2p-1}$ (that is, $L_\lambda(P) = 0$ for polynomials of degree $\leq 2p-1$). We then can use of Fourier transformation to prove the continuity of these functionals. Hence we can associate some well-defined $L_\lambda \in F$ with the functionals. The argument given next follows the same reasoning as previously. We choose the $\lambda_\alpha \in \Lambda_{2p-1}$. The element $f \in F$ which minimizes $\|A\ f\|^2$ subject to the constraints $f(\lambda_\alpha) = f_\alpha$ at the sample points also can be determined by kriging in the isomorphic space generated by an IRF-(2p-1) $Z: \Lambda_{2p-1} \to H$ such that

$$\|Z(\lambda)\| = \|A(L_\lambda)\| \qquad (\lambda \in \Lambda_{2p-1})$$

By definition, the element $L_\lambda \in F$ must satisfy

$$< A L_\lambda, A f > = \int \lambda(dx)\ f(x) \qquad f \in F$$

The Fourier transform $\tilde{A}L_\lambda$ of AL_λ then can be determined from that of λ by

$$\tilde{A}L_\lambda = (-1)^p \frac{(\tilde{\lambda})}{(4\ \Pi^2 |u|^2)^p}$$

where $\tilde{\lambda}$ is the transform of λ. The transform $\tilde{A}L_\lambda$ exists in L^2 because of the condition $\lambda \in \Lambda_{2p-1}$. The generalized covariance $K(h)$ of the $(2p-1)$ - IRF $Z(x)$ satisfies the following relation

$$\int \lambda(dx)\ K(x-y)\ \lambda(dy) = \int (AL_\lambda)^2\ dx$$

for all $\lambda \in \Lambda_{2p-1}$.

By using the spectral measure $\chi(du)/(4\ \Pi^2 |u|^2)^{2p}$ associated with K, this can be written as

$$\int \frac{(\tilde{\lambda})^2}{(4\ \Pi^2\ |u|^2)^{2p}} \chi(du) = \int \frac{(\tilde{\lambda})^2}{(4\ \Pi^2\ |u|^2)^{2p}}\ du \quad (\lambda \in \Lambda_{2p-1})$$

Because the measure χ has no atom at the origin, we therefore have $\chi(du) = du$, that is:

$$\Delta^{2p}\ K(h) = \delta$$

For $n = 2k+1$ (i.e. spaces of odd dimension), $K(h)$ then is $|h|^{4p-n}$ (up to a multiplicative factor).

For $n = 2k$, logarithmic terms of the form $|h|^{4p-n} \log |h|$ arise. For example, using the norm $\int |\Delta\ f|^2\ dx$, we find that

$$K(h) \sim |h|^3 \qquad \text{in } \mathbb{R}^1$$
$$\sim |h|^2 \log |h| \qquad \text{in } \mathbb{R}^2$$
$$\sim |h| \qquad \text{in } \mathbb{R}^3$$

These results also can be deduced directly from the results obtained by Duchon (1976). The criteria he determined for the situation of splines minimizing $\int |\Delta\ f|^2\ dx$ in $\mathbb{R}^1$ or $\mathbb{R}^2$ can be shown easily to be equivalent to the characterization of kriging using $K(h) = |h|^3$ or $|h|^2 \log |h|$ respectively.

REFERENCES

Duchon, J., 1975, Fonctions splines et vecteurs aleatoires: Seminaire d'Analyse Numerique, no. 213.

Duchon, J., 1976, Interpolation des fonctions de deux variables par des fonctions splines du type plaque mince: RAIRO Analyse Numerique, v. 10, no. 12, p. 7-12.

Matheron, G., 1970, La theorie des variables regionalisees et ses applications: C.M.M., Fontainebleau.

Matheron, G., 1973, The intrinsic random functions and their applications: Adv. In Applied Prob., v. 5, p. 439-468.

INDEX