FONTAINEBLEAU/CGMM N-872

NEW TYPES OF DISJUNCTIVE KRIGING

PART I

M. ARMSTRONG Février 1984

NEW TYPES OF DISJUNCTIVE KRIGING : PART 1

INTRODUCTION : THE NEED FOR NEW MODELS.

A series of new and more general isofactorial models leading, among other things, to discrete disjunctive kriging, have been the subject of several recent articles by Matheron (1983, 1984). In the introduction to these papers, Matheron mentions two situations where the existing techniques in non-linear geostatistics cannot be applied.

The first is when the distribution under study has an atom (a peak) at the origin. A typical example of this is uranium grades which can have up to 70 or 80% of zero values. Clearly there is no way of establishing a 1-1 transformation between the real values and the Gaussian anamorphosed values under these circumstances. While it would be possible to establish an arbitrary rule to decide which zeroes are less zero than the others, this solution is rather unsatisfactory. So this is one area where a new approach is required.

Another domain where the existing non-linear methods fall down is when the variables under study are inherently discrete. A classic example of this is the stone count in diamonds.

A third type of situation requiring new models (and one not mentioned by Matheron) is that of data grouped into classes. This occurs in coal preparation and in general in ore dressing, where the data are arranged in size and density classes.

The common point between all these problems is the discrete nature of either the variable itself or of part or all of the data. There is therefore a very real need for some sort of "discrete" disjunctive kriging. The recent work by Matheron presents several models of this type as well as new models for change of support. However as these articles come after 10 years of work on the subject, they tend to be fairly heavy going. The present writer found it helpful to go back to Matheron's early unpublished notes to trace the evolution of his ideas. This paper presents an annotated translation of hitherto unpublished parts of his work, which may be helpful to other readers. The starting point is the 1973 note N-360, where disjunctive kriging was first introduced. This note is divided into the following sections :

0 - Introduction

1 - Equations of Disjunctive Kriging

2 - The representation of the bivariate distributions

3 - Isofactorial models

4 - Hermite Polynomial Models (corresponding to a normal distribution)

5 - Laguerre Polynomial Models (corresponding to a gamma distribution)

6 - Models with Polynomial Factors (in general)

7 - A general representation for these types of distributions.

Although the first four sections were published in English at the first NATO Geostatistics workshop (Matheron, 1976), the rest has languished in the Centre's archives. Of the remaining three sections, chapters 5 and 6 are the starting point for "discrete disjunctive kriging".

It would be possible to present the translations of just these two sections and leave the reader to draw the parallels with the usual disjunctive kriging based on Hermite polynomials, but this has one serious disadvantage : the section on the Hermitian case presented in the NATO papers is a very condensed version of the French original. Moreover, since chapters 4 and 5 (in French) follow exactly the same line of reasoning, it seems preferable to start with the section on Hermite models and then go on to the other types of disjunctive kriging.

DISJUNCTIVE KRIGING WITH HERMITE POLYNOMIAL MODELS

To start with, Matheron reviews some of the properties of Hermite polynomials $H_n(x)$. (Proofs can be found in any text with a chapter on orthogonal polynomials). They are defined by

$$e^{-x^2/2} H_n(x) = \frac{d^n}{dx^n} e^{-x^2/2} \qquad n = 0, 1, 2...$$

These form a set of orthogonal polynomials with respect to the normal distribution; that is,

$$\int H_n(x) H_m(x) \frac{e^{-x^2/2}}{\sqrt{2\pi}} dx = \delta_{nm} n!$$

The normed polynomials $\eta_n(x) = H_n(x)/ n!$ form an orthogonal basis for the Hilbert space $L^2(R, e^{-x^2/2}/\sqrt{2\pi})$. Matheron then goes on to develop a factorial representation for the bivariate normal p.d.f. To be more precise, he shows that the bivariate normal p.d.f. g(x,y) can be written as

$$g(x,y) = g_1(x)\ g_2(y) \left[\sum_{o}^{\infty} U_n H_n(x) H_n(y) \right] \tag{1}$$

where $g_1(x)$ and $g_2(y)$ are the marginal p.d.f. This result is clearly the key to disjunctive kriging.

A simple way of doing this is by finding the Fourier transform of g(x,y). This is

$$e^{-\frac{u^2+v^2}{2} - \rho uv} = \sum (-1)^n \rho^n \frac{u^n v^n}{n!} e^{-\frac{u^2+v^2}{2}}$$

Now $(-1)^n u^n e^{-u^2/2}$ is the Fourier transform of $\frac{d^n}{dx^n} e^{-x^2/2}$, that is of $H_n(x)\ g_1(x)$. So it follows that

$$g(x,y) = \Phi(x,y)\ g_1(x)\ g_2(y) \tag{2}$$

where

$$\Phi(x,y) = \sum \frac{\rho^n}{n!} H_n(x) H_n(y)$$

So for the normal distribution, the "factors" are the (normed) Hermite polynomials, and the corresponding Eigen values are ρ^n, n = 0,1... (The eigen values appear naturally when the theory is developed in terms of projections). The chapter on the normal distribution also contains three small sections. The first shows how to find the conditional expectation for a stationary random function which can be transformed (by an anamorphosis) to a multivariate normal distribution ; that is when the joint distribution of the random variables $Z(x_1),...Z(X_m)$ have a multivariate normal distribution for any m sample locations $x_1,...x_m$. Since this is rather restrictive, the second section treats the case where only the bivariate distributions are normal. This is just the usual disjunctive kriging. In the third section, Matheron shows how to calculate the estimation variance in the preceding two cases. As it is not always possible to transform a random function to bivariate or multivariate normality, he then goes on to consider the gamma distribution and its orthogonal polynomials.

DISJUNCTIVE KRIGING WITH LAGUERRE POLYNOMIAL MODELS

The first paragraph of this chapter is devoted to finding a model for the corelation between two random variables Z(x) and Z(x+h), each with a gamma distribution.

Consider a random measure μ on $\mathbb{R}^n$, which is stationary and orthogonal, and has a gamma distribution, that is, for all Borel balls $B \subset \mathbb{R}^n$ with volume $V(B) = V < \infty$

$$E\left[e^{-\lambda\mu(B)}\right] = e^{-\psi(\lambda)}$$

where $\psi(\lambda) = \theta V \log (a+\lambda)$.

Let Z(x) denote the regularization of μ corresponding to a specified compact ball B ; i.e.

$$Z(x) = \int \mu(d\xi)\ 1_B(x+\xi)$$

where $1_B(.)$ is the indicator function for the ball B. The transitive covariogram of B is

$$K(h) = \int 1_B(x)\ 1_B(x+h)\ dx$$

(For more information on transitive covariograms see Matheron (1970)). Matheron then shows that

$$E\left[e^{-\lambda Z(x)-\nu Z(x+h)}\right]$$

$$= \exp\left\{-\theta K(h)\ \psi(\lambda+\nu) - \theta(K(o)-K(h))(\psi(\lambda)+\psi(\nu))\right\} \qquad (3)$$

This is done by splitting up Z(x) and Z(x+h) into disjoint components.

$$Z(x) = \mu(B_x \cap B_{x+h}) + \mu(B_x \setminus B_x \cap B_{x+h})$$

$$Z(x+h) = \mu(B_x \cap B_{x+h}) + \mu(B_{x+h} \setminus B_x \cap B_{x+h})$$

So the random variables X = Z(x) and Y = Z(x+h) can be split into three components, X_1, Y_1 and W :

$$Z(x) = W + X_1$$

$$Z(x+h) = W + Y_1$$

where X_1, Y_1 and W are three independent gamma variables with parameter values of $\theta(K(o) - K(h))$ for X and Y, and of $\theta K(h)$ for W. Equation (3) follows from this.

$$E\left[e^{-\lambda Z(x)-\nu Z(x+h)}\right]$$

$$= E\left[e^{-\lambda(W+X_1) - \nu(W+Y_1)}\right]$$

$$= E\left[e^{-(\lambda+\nu)W - \lambda X_1 - \nu Y_1}\right]$$

$$= \exp\left\{-\theta\ K(h)\ \psi\ (\lambda+\nu) - \theta\left[K(o) - K(h)\right](\psi(\lambda) + \psi(\nu))\right\}$$

So we see that Z(x) and Z(x+h) are correlated gamma variables with parameter $\alpha = \theta K(o) = \theta\ \nu(B)$. The correlation coefficient ρ between Z(x) and Z(x+h) is $\rho = K(h)/K(o)$. So the parameter value for W is $\rho\alpha$, while that for X_1 and Y_1 is $\alpha(1-\rho)$. For simplicity the scale parameter has been set to 1.

Laguerre Polynomials :

The system of orthogonal polynomials relative to the gamma distribution is defined by :

$$\frac{d^n}{dx^n}\ x^{n+\alpha-1}\ e^{-x} = (-1)^n\ n!\ x^{\alpha-1}\ e^{-x}\ L_n(x)$$

for a specified parameter value $\alpha > 0$. It can be shown that these polynomials are orthogonal. Integration by parts gives

$$\frac{1}{\Gamma(\alpha)} \int_0^\infty L_n(x)\, L_m(x)\, x^{\alpha-1}\, e^{-x}\, dx = \delta_{nm}\, C_n$$

where $C_n = \frac{\Gamma(n+\alpha)}{n!\Gamma(\alpha)}$. Consequently the normal polynomials $\ell_n = L_n/\sqrt{C_n}$ form an orthogonal basis for the Hilbert space $L^2(R_{+1}, \frac{1}{\Gamma(\alpha)} x^{\alpha-1} e^{-x})$.

The next step is to show that the bivariate p.d.f. of two gamma variables X and Y with the correlation structure defined earlier can be written as

$$\Phi(x,y)\ \frac{x^{\alpha-1}}{\Gamma(\alpha)}\ \frac{y^{\alpha-1}}{\Gamma(\alpha)}\ e^{-x-y}$$

where $\Phi(x,y)$ is, in this case,

$$\Phi(x,y) = \Sigma\, U_n\, \frac{L_n(x)\, L_n(y)}{C_n} \tag{4}$$

This is done by obtaining the Laplace transform of the Laguerre polynomial $L_n(x)$, which is

$$\frac{1}{\Gamma(\alpha)} \int_0^\infty e^{-\lambda x}\, L_n(x)\, x^{\alpha-1}\, e^{-x}\, dx$$

$$= \frac{(-1)^n}{n!}\ \frac{1}{\Gamma(\alpha)} \int_0^\infty \frac{d^n}{dx^n} (x^{\alpha+n-1}\, e^{x})\, e^{-\lambda x}\, dx$$

$$= \frac{(-1)^n}{n!}\ \frac{\lambda^n}{\Gamma(\alpha)} \int_0^\infty x^{\alpha+n-1}\, e^{-(1+\lambda)x}\, dx$$

$$= \frac{(-1)^n}{\Gamma(\alpha)}\ \frac{\lambda^n}{n!}\ \Gamma(\alpha+n)\ .\ \frac{1}{(1+\lambda)^{\alpha+n}}$$

$$= (-1)^n\, C_n\, \frac{\lambda^n}{(1+\lambda)^{\alpha+n}} \tag{5}$$

The bivariate Laplace transform $E(e^{-\lambda X-\nu Y})$ can be calculated directly from the decomposition into X_1, Y_1 and W.

$$E[e^{-\lambda X-\nu Y}] = E[e^{-\lambda X_1 - \nu Y_1 - (\lambda+\nu)W}]$$

$$= \frac{1}{(1+\lambda)^{(1-\rho)\alpha}} \cdot \frac{1}{(1+\nu)^{(1-\rho)\alpha}} \cdot \frac{1}{(1+\lambda+\nu)^{\rho\alpha}}$$

$$= \frac{1}{(1+\lambda)^{\alpha}} \frac{1}{(1+\nu)^{\alpha}} \left[\frac{(1+\lambda)(1+\nu)}{(1+\lambda+\nu)}\right]^{\rho\alpha}$$

$$= \frac{1}{(1+\lambda)^{\alpha}(1+\nu)^{\alpha}} \left[1 - \frac{\lambda\nu}{(1+\lambda)(1+\nu)}\right]^{-\rho\alpha}$$

The last term can be expanded as a negative binomial ; viz

$$\left[1 - \frac{\lambda\nu}{(1+\lambda)(1+\nu)}\right]^{-\rho\alpha} = \sum \frac{\Gamma(\rho\alpha+n)}{n!\ \Gamma(\rho\alpha)} \cdot \left(\frac{\lambda\nu}{(1+\lambda)(1+\nu)}\right)^{n}$$

Hence

$$E[e^{-\lambda X} e^{-\nu Y}] = \sum \frac{\Gamma(\rho\alpha+n)}{n!\ \Gamma(\rho\alpha)} \frac{\lambda^n \nu^n}{(1+\lambda)^{n+\alpha}(1+\nu)^{n+\alpha}}$$

$$= \sum C_n U_n \frac{\lambda^n}{(1+\lambda)^{n+\alpha}} \frac{\nu^n}{(1+\nu)^{n+\alpha}} \qquad (6)$$

where $$U_n = \frac{\Gamma(\rho\alpha+n)}{\Gamma(\rho\alpha)} \frac{\Gamma(\alpha)}{\Gamma(\alpha+n)}$$

Back-transforming equation (6) leads to the factorial representation (4) and shows that the eigen values are U_n. In contrast to the normal distribution where $U_n = \rho^n$, here we have

$$U_n = E[T^n]$$

where T is a beta random variable $\beta(\rho\alpha, (1-\rho)\alpha)$ with mean ρ. So although a Gaussian anamorphosis transforms the marginal distribution to a normal, the bivariate distribution is not Gaussian (normal) - but this should have been obvious from the outset.

For the sake of curiosity, it is worthwhile seeing if there is a bivariate distribution with

$$\Phi(x,y) = \sum \rho^n \, \ell_n(x) \, \ell_n(y)$$

which could possibly be transformed to bi-normality.

[NOTE : The new bivariate distribution being developed here represents a diffusion-type process. It is the basis for Matheron's recent work (1983) on change of support models] .

From (7), the Laplace transform of $\Phi(x,y)$ is

$$\Phi(\lambda,\nu) = \sum \frac{\rho^n}{n!} \frac{\Gamma(n+\alpha)}{\Gamma(\alpha)} \frac{\lambda^n \nu^n}{(1+\lambda)^{n+\alpha} (1+\nu)^{n+\alpha}}$$

Remembering the form for the negative binomial expansion, we see that

$$\Phi(\lambda,\nu) = \frac{1}{[(1+\lambda)(1+\nu)]^{\alpha}} \left[1 - \frac{\rho \lambda \nu}{(1+\lambda)(1+\nu)}\right]^{-\alpha}$$

$$= [1 + \lambda + \nu + (1-\rho)\lambda \nu]^{-\alpha}$$

We now have to check that this Laplace transform does correspond to some distribution (an infinitely divisible one, at that). To do this we consider the following expansion :

$$\Phi(\lambda,\nu) = \frac{(1-\rho)^{\alpha}}{(1+\lambda\,(1-\rho))^{\alpha}\,(1+\mu\,(1-\rho))^{\alpha}} \left[1 - \frac{\rho}{(1+\lambda\,(1-\rho))(1+\nu\,(1-\rho))}\right]^{-\alpha}$$

$$= \frac{(1-\rho)^{\alpha}}{\Gamma(\alpha)} \sum \frac{\Gamma(\alpha+n)}{n!} \frac{\rho^n}{(1+\lambda\,(1-\rho))^{\alpha+n}\,(1+\nu\,(1-\rho))^{\alpha+n}}$$

This is the Laplace transform of a mixture of two gamma distributions with parameter $\alpha+n$ where the parameter n of the gamma distributions varies according to a negative binomial distribution with p.d.f.

$$\Pr\,(N=n) = \frac{\Gamma(n+\alpha)}{n!\,\Gamma(\alpha)} \rho^n (1-\rho)^{\alpha}$$

For the sake of curiosity, it is worthwhile seeing if there is a bivariate distribution with

$$\Phi(x,y) = \sum \rho^n \ell_n(x) \ell_n(y)$$

which could possibly be transformed to bi-normality.

[NOTE : The new bivariate distribution being developed here represents a diffusion-type process. It is the basis for Matheron's recent work (1983) on change of support models] .

From (7), the Laplace transform of $\Phi(x,y)$ is

$$\Phi(\lambda,\nu) = \sum \frac{\rho^n}{n!} \frac{\Gamma(n+\alpha)}{\Gamma(\alpha)} \frac{\lambda^n \nu^n}{(1+\lambda)^{n+\alpha} (1+\nu)^{n+\alpha}}$$

Remembering the form for the negative binomial expansion, we see that

$$\Phi(\lambda,\nu) = \frac{1}{[(1+\lambda)(1+\nu)]^{\alpha}} \left[1 - \frac{\rho \lambda \nu}{(1+\lambda)(1+\nu)}\right]^{-\alpha}$$

$$= [1 + \lambda + \nu + (1-\rho)\lambda \nu]^{-\alpha}$$

We now have to check that this Laplace transform does correspond to some distribution (an infinitely divisible one, at that). To do this we consider the following expansion :

$$\Phi(\lambda,\nu) = \frac{(1-\rho)^{\alpha}}{(1+\lambda (1-\rho))^{\alpha} (1+\mu (1-\rho))^{\alpha}} \left[1 - \frac{\rho}{(1+\lambda (1-\rho))(1+\nu (1-\rho))}\right]^{-\alpha}$$

$$= \frac{(1-\rho)^{\alpha}}{\Gamma(\alpha)} \sum \frac{\Gamma(\alpha+n)}{n!} \frac{\rho^n}{(1+\lambda (1-\rho))^{\alpha+n} (1+\nu (1-\rho))^{\alpha+n}}$$

This is the Laplace transform of a mixture of two gamma distributions with parameter $\alpha+n$ where the parameter n of the gamma distributions varies according to a negative binomial distribution with p.d.f.

$$Pr (N=n) = \frac{\Gamma(n+\alpha)}{n!\, \Gamma(\alpha)} \rho^n (1-\rho)^{\alpha}$$

Then the conditional p.d.f. of the gamma variates is of the form

$$x_1^{\alpha+n-1} \quad e^{-x_1} \quad dx_1$$

After changing variables to

$$X = X_1(1-\rho) \qquad Y = Y_1(1-\rho)$$

this p.d.f. becomes

$$\left(\frac{x}{1-\rho}\right)^{\alpha+n-1} \quad e^{-x/(1-\rho)} \frac{dx}{1-\rho}$$

Consequently the joint (unconditional) p.d.f. of X and Y is

$$f(x,y)dx\ dy = \sum_{n=o}^{\infty} \frac{\Gamma(n+\alpha)}{n!\,\Gamma(\alpha)} \rho^n (1-\rho)^\alpha \left(\frac{x\ y}{(1-\rho)^2}\right)^{\alpha+n-1} e^{-x+y/(1-\rho)} \frac{dx}{(1-\rho)} \frac{dy}{(1-\rho)}$$

$$= \sum_{o}^{\infty} \frac{\Gamma(n+\alpha)}{n!\,\Gamma(\alpha)} \rho^n \frac{(xy)^{\alpha+n-1}}{(1-\rho)^{2n+\alpha}} e^{-(x+y)/(1-\rho)} dx\ dy \qquad (7)$$

It is not difficult to see that the Laplace transform given earlier corresponds to this p.d.f.

NOTE : The gamma distribution with $\alpha = 1/2$ is the square of a normal distribution. And so in this case (and only in this one) the Gaussian anamorphosis gives a bivariate normal distribution.

To conclude the chapter on the gamma distribution, Matheron shows that the Laguerre polynomials can be kriged separately to obtain the disjunctive kriging estimator.

In summary, this approach led Matheron to look for random functions with the following properties :

a) Z(x) must be stationary. For a given x, there must be a complete countable set of orthogonal functions χ_n associated with the distribution w(dz) of Z(x). The χ_n form the basis for the Hilbert space $L^2(\mathbb{R},w)$.

b) For any two points x and y, the joint p.d.f. of Z(x) and Z(y) must be of the form

$$\Phi(z,z')\ w(dz)\ w(dz')$$

where $$\Phi(z,z') = \sum U_n(x,y)\ \chi_n(z)\ \chi_n(z').$$

The following section is devoted to the search for other models with polynomial factors.

LOOKING FOR OTHER MODELS WITH POLYNOMIAL FACTORS

Several problems arise when one is looking for random functions satisfying the conditions listed above. Since the bivariate p.d.f. $f_{\alpha\beta}(Z,Z')$ plays the same role in disjunctive kriging as the covariance does in ordinary kriging, some sort of generalization of Bochner's theorem is required to characterize the families of bivariate distributions which can be associated with random functions. To be more precise, for a given family $F(dz,dz';x_1,x_2)$ of bivariate distributions with $x_1, x_2 \in \mathbb{R}^n$, what is the condition that guarantees the existence of a R.F. $Z(x)$ on $\mathbb{R}^n$ having $F(dz,dz';x_1,x_2)$ as the bivariate distribution of $Z(x_1)$, $Z(x_2)$. Clearly for all $x \in \mathbb{R}^n$ the marginal probability $w(.) = F(., R, x, x')$ must be independent of x.

By analogy with the covariance which must be positive definite, we could postulate a similar condition. That is, for any $x_1,\ldots,x_k \in \mathbb{R}^n$ and $f_i \in L^2(R',w_i)$ $(i = 1,\ldots,k)$, we could require that

$$E\left[\left(\sum f_i(Z(x_i))\right)^2 \right] \geq 0$$

i.e.

$$\sum_i \sum_j \int f_i(Z)\, f_j(Z')\, F(dz, dz';x_i, x_j) \geq 0 \tag{8}$$

But although this condition is necessary, it is hardly sufficient. To see this we consider a partition B_k of R. The all or nothing R.F.s $1_{B_k}(Z(x))$ associated with a random partition A_k of $\mathbb{R}^n$:

$$A_k = \{x : Z(x) \in B_k\}$$

and

$$\Pr(x \in A_k, x' \in A_{k'}) = F(B_k, B_{k'}; x, x') \tag{9}$$

Condition (8) means that the covariances $C_{kk'}(x,x') = F(B_k, B_{k'}; x,x')$ form the covariance matrix of a vectorial R.F. $(Y_1(x),\ldots,Y_k(x))$ but there is no reason to believe that this vectorial R.F. need necessarily be the indicator function of a random partition.

However this suggests that the following condition should be necessary and sufficient : that for every countable partition B_k there exists a random partition A_k satisfying (9).

This draws our attention to the following problem. Under what conditions is a family $C_{kk'}(x,x')$ the covariance matrix of a random partition. One particular, simpler case of this is : under what conditions is the function $C(x,x')$ the covariance of a random set?

This problem is by no means trivial. For example, covariances of the form $\exp\{-|x-x'|^2\}$ cannot be associated with any random set. (As they have a second derivative, they have to represent a differentiable R.F.).

Matheron notes that these problems are rather difficult and goes on to attack the question from a different point of view. When dealing with ordinary variograms and covariances , the difficulty of testing whether a given function was positive definite and hence could be used as a covariance, meant that new covariance models are usually developed by construction from a known regionalized variable. Similarly the easiest way to produce bivariate distributions with the required properties is by regularizing a stationary orthogonal measure μ e.g.

$$Z(x) = \int k(x+\xi)\ \mu(d\xi)$$

where k is given.

Suppose that the (infinitely divisible) distribution associated with μ is defined by

$$E\left[e^{-\lambda\mu(V)}\right] = e^{V\psi(\lambda)}$$

then

$$E\left[e^{-\lambda Z(x)-\nu Z(x+h)}\right] = \exp\left\{\int \psi(\lambda k(\xi) + \nu k(\xi+h))\,d\xi\right\}$$

In particular if $k = 1_B$, then

$$E\left[e^{-\lambda Z(x)-\nu Z(x+h)}\right]$$

$$= \exp\left\{[K(o) - K(h)]\ [\psi(\lambda) + \psi(\nu)] + K(h)\ \psi(\lambda+\nu)\right\}$$

where $K(h)$ is the transitive covariogram of k. Substituting $\Phi(\lambda) = \exp\{K(o)\ \psi(\lambda)\}$ and $\rho = K(h)/K(o)$ gives

$$\Phi(\lambda,\nu) = \Phi(\lambda)^{1-\rho}\ \Phi(\nu)^{1-\rho}\ \Phi(\lambda+\nu)^{\rho} \qquad (10)$$

Models with Polynomial Factors

The next step is to find distributions satisfying (10) and having polynomial factors χ_n. In the preceding chapters it was shown that the normal distribution and the gamma distribution have these properties. Matheron now goes on to show that the same is true of the Poisson distribution and the negative binomial. Moreover, these are the only non-trivial distributions which satisfy (10).

To start with, he shows that if a distribution w(dx) has a set of polynomials χ_n which form a basis for the Hilbert space $L^2(\mathbb{R}^n,w)$, the following properties are equivalent to ensure that the bivariate distribution of X and Y has a symmetric p.d.f. of the form

$$\Phi(x,y)\ w(dx)\ w(dy)$$

1) $\Phi(x,y)$ is of the form $\sum U_n\ \chi_n(x)\ \chi_n(y)$

2) For all $n \geq 0$ $E[X^n|Y]$ is a polynomial of degree n in y.

Proof : If 1) holds, then

$$E[X^n|Y] = \sum_p U_p\ \chi_p(y) \int \chi_p(x)\ x^n\ w(dx)$$

Since $\int \chi_p(x)\ x^n\ w(dx) = 0$ for all $p > n$,

$$E[X^n|Y = y] = \sum_{p\leq n}^{\chi} U_p < x^n, \chi_p > \chi_p(y)$$

where $< x^n, \chi_p >$ denotes $\int \chi_p(x)\ x^n\ w(dx)$.

This expression is a polynomial of degree $\leq$ n. In fact, it is precisely of degree n, because $< x^n, \chi_p >$ cannot be zero since

$$x^n = \sum_{p=o}^{n} < x^n, \chi_p > \chi_p(x)\ .$$

Conversely, suppose 2) holds. We have to show that the orthogonal polynomials $\chi_n(x)$ satisfy

$$E[\chi_n(x)|y] = U_n\ \chi_n(y)$$

The space $\mathcal{P}_{n+1}$ of dimension n+1 made up of polynomials of degree $\leq$ n is invariant under the operator $E[X|Y]$. Moreover this operator maps $\mathcal{P}_{n+1}$

onto itself, since the degree of the polynomials is conserved. Consequently, this symmetric operator has n+1 orthogonal eigen vectors in $\mathcal{P}_{n+1}$, which are just the χ_k , k = 0, 1,...,n. So 1) follows.

The next step is to look for the conditions under which distributions of the form (10) have polynomial factors. From the discussion given above, it is clear that 2) must hold.

From the relation

$$\int e^{-\lambda y} E[X^n | Y = y] \ w(dy) = (-1)^n \frac{\partial^n}{\partial \nu^n} \Phi(\lambda,\nu)\Big|_{\nu=o}$$

it is clear that the Laplace transform Φ must satisfy

$$\Phi(\lambda)^{1-\rho} \frac{\partial^n}{\partial \lambda^n} [\Phi(\lambda)]^{\rho} = \sum_{o}^{n} A_{n_k} \frac{\partial^k}{\partial \lambda^k} \Phi(\lambda)$$

for all n. For n = 1, this is of the form

$$\rho \Phi' = a_1 \Phi + b_1 \Phi'$$

This can always be satisfied by putting $a_1 = 0$ and $b_1 = \rho$. For n = 2, put $\Phi = e^{\psi(\lambda)}$ to obtain a differential equation of the form :

$$\psi'' = a \psi'^2 + b \psi' + c$$

If a = b = 0, ψ is a second degree polynomial, which leads to the normal distribution. If a = 0 but b $\neq$ 0, then $\chi(\lambda) = A e^{b\lambda} + B$; this corresponds to the Poisson distribution.

If a $\neq$ 0, the quadratic $ax^2 + bx + c$ can have 0, 1 or 2 real roots. If there is one real root α,

$$\psi'' = a(\psi' - \alpha)^2$$

The solution is then $\psi = \alpha\lambda - \frac{1}{a} \log (1+c\lambda)$, which corresponds to a gamma distribution (possibly translated and transposed).

If the quadratic has no real roots :

$$\psi'' = a \left[(\psi'+\alpha)^2 + b^2 \right]$$

and $$\psi(\lambda) = \frac{1}{a} \log \cos (ab\lambda + c) - \alpha \lambda - c'$$

there is no distribution with this as its Laplace transform. Lastly, there is the case where there are two (different) real roots :

$$\psi'' = a(\psi'-\alpha)(\psi'-\beta)$$

This gives

$$\frac{\psi'-\alpha}{\psi'-\beta} = b\, e^{a(\beta-\alpha)\lambda}$$

or

$$\psi'(\lambda) = \frac{\alpha - \beta b\, e^{a(\beta-\alpha)\lambda}}{1 - b\, e^{a(\beta-\alpha)\lambda}}$$

Supposing that $a(\beta-\alpha) > 0$, which is permissible since α and β can always be reversed, the solution is

$$\psi(\lambda) = \beta\lambda - \frac{1}{a} \log\left(\frac{b - e^{-a(\beta-\alpha)\lambda}}{b - 1}\right)$$

This corresponds to a negative binomial (at least up to a linear transformation), if $|b| > 1$.

So we see that in addition to the gamma and the normal distributions, two others (the Poisson and the negative binomial) also have the properties needed for disjunctive kriging, at least for $n = 2$.

Poisson Distribution

The next step is to check that these properties hold for $n > 2$ for the Poisson distribution. Since this is a discrete distribution, the generating function $G(s)$ will be used instead of Laplace transforms. For the Poisson distribution with parameter θ

$$G(s) = e^{\theta(s-1)}$$

The bivariate distribution associated with (10) is defined by

$$G(s,t) = \sum P_{nm}\, s^n\, t^m$$
$$= G(s)^{1-\rho}\, G(t)^{1-\rho}\, G(st)^{\rho}$$

Hence

$$\text{Log } G(s,t) = (1-\rho)\left[\text{Log } G(s) + \text{Log } G(t)\right] + \rho \text{Log } G(st)$$

Substituting for G(s) etc, gives

$$\text{Log } G(s,t) = (1-\rho)\ \theta\left[s - 1 + t - 1\right] + \rho\theta\left[st - 1\right]$$

$$= \theta(t-1) + \theta(1 + \rho t - \rho)(s-1)$$

Consequently for a fixed value of n

$$\sum_n P_{nm}\ t^m = e^{-\theta}\ e^{\theta(t-1)(1-\rho)}\ \frac{\theta^n}{n!}\ (1 + \rho t - \rho)^n$$

So the generating function $G_n(t)$ of Y for a fixed value of X = n is

$$G_n(t) = (1 + \rho t - \rho)^n\ e^{\theta(1-\rho)(t-1)}$$

The conditional moments can be found by differentiating k+1 times and putting t = 0.

$$E\left[Y(Y-1)\ldots(Y-k) \mid n\right] = \theta^{k+1} + \ldots + \rho^k\ n(n-1)(n-k)\ .$$

So we see that $E\left[Y^n|X\right]$ is indeed a polynomial of degree n in x. The eigen value U_n associated with the polynomial X_n is $U_n = \rho^n$, as was the case for the normal.

To complete this section, Matheron gives the expression for the orthogonal polynomials associated with the Poisson distribution. These are

$$P_n(x) = 1 - \binom{n}{1}\frac{x}{\theta} + \binom{n}{2}\ \frac{x(x-1)}{\theta^2}\ \ldots\ldots\ldots + (-1)^n\ \frac{x\ldots\ (x-n+1)}{\theta^n}$$

and the normed polynomials are

$$w_n(x) = \sqrt{\frac{\theta^n}{n!}}\ P_n(x)$$

Negative Polynomial Distribution

As with the Poisson distribution, it is necessary to show that the conditions are satisfied for n > 2. The negative binomial has the following generating function

$$G(s) = \left(\frac{1-\alpha}{1-\alpha s}\right)^{\beta}$$

$$= (1-\alpha)^{\beta} \frac{\sum \alpha^n \, \Gamma(\beta+n) \, s^n}{n! \, \Gamma(\beta)} \qquad \begin{array}{l} 0 < \alpha < 1 \\ 0 < \beta \end{array}$$

The corresponding bivariate distribution is defined by

$$G(s,t) = E(s^X \, t^Y)$$

$$= G(s)^{1-\rho} \, G(t)^{1-\rho} \, G(st)^{\rho}$$

To obtain the conditional distribution of Y for a fixed value of X, the random variables are split into 3 independent components :

$$X = X_1 + Z$$

$$Y = Y_1 + Z$$

such that

$$G(s)^{1-\rho} = E(s^{X_1}) = E(s^{Y_1})$$

$$G(s)^{\rho} = E(s^Z)$$

X_1 and Z have negative binomial distributions and

$$Pr(X_1 = k, \, Z = p) = \frac{(1-\alpha)^{\beta} \, \alpha^{k+p} \, \Gamma(\beta\,\rho + p) \, \Gamma((1-\rho)\,\beta + k)}{p! \, \Gamma(\rho\beta) \, k! \, \Gamma((1-\rho)\,\beta)}$$

The distribution of Z for a fixed value of $X = X_1 + Z$ can be deduced from

$$Pr(Z = p \mid X = n) = \frac{Pr(Z = p, \, X_1 = n - p)}{Pr\,(X = n)}$$

$$= \binom{n}{p} \frac{\Gamma(\beta)}{\Gamma(\beta\rho)\Gamma((1-\rho)\,\beta)} \frac{1}{} \; \frac{\Gamma(\rho\,\beta + p) \, \Gamma((1-\rho)\,\beta + n - p)}{\Gamma(\beta + n)}$$

for $0 \leq p \leq n$

This can be shown to be a binomial distribution with parameters n and p where the value of p is chosen at random from a beta distribution.

(To see this, let p be a binomial variable with parameters n and x and let x be a r.v. with a beta distribution with parameters $\rho\beta$ and $(1-\rho)\beta$. The p.d.f. of the beta distribution is

$$f(x) = \frac{\Gamma(\beta)}{\Gamma(\rho\beta)\ \ \Gamma((1-\rho)\ \beta)}\ x^{\rho\beta-1}\ (1-x)^{(1-\rho)\beta-1}$$

$$\Pr(P = p|x) = \binom{n}{p}\ x^{p}\ (1-x)^{n-p}$$

$$\Pr(P = p) = \int_0^1 \Pr(P = p|x)\ f(x)\ dx$$

Substituting the expressions for $\Pr(P = p|x)$ and $f(x)$ into this equation and integrating gives the required relation).

The generating function of Z given X can be deduced from this :

$$F(s^N|X = n) = \sum C\ \binom{n}{p} \int s^k\ x^p\ (1-x)^{n-p}\ x^{\rho\beta-1}\ (1-x)^{\rho\beta-1}\ dx$$

where

$$C = \frac{\Gamma(\beta)}{(\rho\beta)\ \Gamma((1-\rho)\ \beta)}$$

Reversing the order of integration and summation and noting that

$$\sum \binom{n}{p}\ (sx)^p\ (1-x)^{n-p} = (1 - x + sx)^n$$

leads to

$$E(s^Z|X = n) = C \int_0^1 (1 - x + sx)^n\ x^{\rho\beta}\ \ (1-x)^{\rho\beta-1}\ dx$$

Consequently the generating function of $Y = Z + Y_1$ is

$$E\left[s^Y|X = n\right] = \left(\frac{(1-\alpha)}{(1-\alpha s)}\right)^{(1-\rho)\beta} E\left[s^Z|n\right]$$

Differentiating the conditional generating function for Z, p times and putting s = 1, gives

$$E\left[Z(Z-1)\ \ldots\ \ (Z-p+1)|X = n\right]$$

$$= n(n-1)\ \ldots\ \ (n - p + 1) \int_0 x^p\ f_\beta(x)\ dx$$

where $f_{\beta}(x)$ is the p.d.f. of a beta distribution. This shows that $E[Z^p|X = n]$ is a polynomial of degree p in n and hence so is $E[Y^p|X = n]$.

Therefore the negative binomial satisfies the requirements stated earlier. The explicit expression for the orthogonal polynomials will not be given here. However it is interesting to note that their eigen values are given by

$$U_p = \int_0^1 x^p f_{\beta}(x)\, dx = \frac{\Gamma(\rho\beta + p)\ \Gamma(\beta)}{\Gamma(\rho\beta)\ \ \Gamma(\beta+p)}$$

CONCLUSION

This shows that the basic theory for disjunctive kriging using distributions other than the normal existed in Matheron's notes as early as 1973. His subsequent work in 1975 (N-432 and N-449) shows that there are other distributions with polynomial factors but unlike the four considered here, the other distributions are not infinitely divisible and hence do not satisfy equation (10). This work, which will be presented in a subsequent note, was developed using a different approach - that of infinitesimal generators.

ACKNOWLEDGEMENT

The author would like to thank Professor Matheron for many constructive discussions during the preparation of the manuscript.

REFERENCES

Matheron, G. (1973) Le Krigeage Disjonctif. Internal Note N-360. Centre de Géostatistique, Fontainebleau. 40 p.

Matheron, G. (1983) Modèles Isofactoriels et Changement de Support. Internal Note N-821, Centre de Géostatistique, Fontainebleau. Submitted for publication.

Matheron, G. (1984) Isofactorial Models and Change of Support. NATO A.S.I. "Geostatistics for Natural Resources Characterization" , ed. G. Verly, Reidel & Co., Dordrecht.

Matheron, G. (1976) A Simple Substitute for Conditional Expectation : The Disjunctive Kriging. Proceedings, NATO A.S.I. "Advanced Geostatistics in the Mining Industry", ed. M. Guarascio et Al., Reidel & Co., Dordrecht, pp. 221-236.

Matheron, G. (1975a) Compléments sur les Modèles Isofactoriels. Internal Note N-432. Centre de Géostatistique, Fontainebleau. 21 p.

Matheron, G. (1975b) Modèles Isofactoriels Discrets et Modèles de Walsh. Internal Note N-449, Centre de Géostatistique, Fontainebleau.24 p.