RESEARCH REPORTS

Department of Statistics

University of Leeds

Leeds LS2 9JT

Some Statistical Inference Problems in kriging II : Theory

by

K.V. MARDIA

No.4. July 1980

Some Statistical Inference Problems in kriging II : Theory

K.V. Mardia

University of Leeds

SUMMARY

In simple kriging, various models are fitted at present by some ad-hoc methods. We explore the maximum likelihood estimators for fitting the parameters when the stochastic process is Gaussian. First, we deal with isotropic models, and estimate the nugget effect, sill and range. A set of unbiased estimators of the nugget effect and sill are also given when the range is infinite or known. Tests for important hypotheses such as no-nugget effect, etc. are developed. Secondly, we deal with the fitting of anisotropic models under geometric anisotropy, and in particular, the maximum likelihood estimator of the anisotropy parameter is given. Various tests of the hypothesis of isotropy are developed. Particular cases such as the spherical scheme, doubly geometric scheme are described. Their numerical applications are given elsewhere.

1. Introduction

Let $\underset{\sim}{z}' = (z_1, \ldots, z_n)$ be a single realization. The location of z_i is at $\underset{\sim}{x} = \underset{\sim}{x}_i$, $x \epsilon R_p$. In practice, these points may be on a line or a plane. The vector $\underset{\sim}{z}$ is a 'stacked' vector. Suppose that $\underset{\sim}{z} \sim N(\mu\underset{\sim}{1}, \underset{\sim}{\Sigma})$.

We assume that

$$\underset{\sim}{\Sigma} = (\sigma_{ij}), \ \sigma_{ii} = \sigma^2,$$

$$\sigma_{ij} = \sigma(d_{ij}),$$

where $d_{ij} = (\underset{\sim}{x}_i - \underset{\sim}{x}_j)' \underset{\sim}{\Lambda} (\underset{\sim}{x}_i - \underset{\sim}{x}_j)$, $\underset{\sim}{\Lambda} > 0$.

For this paper, we assume that the form of $\sigma(.)$ is known but it contains unknown parameters. (We replace the Mahalanobis distance by the L_1 distance for the doubly geometric process).

For $\underset{\sim}{\Lambda} = \underset{\sim}{I}$, we have the case of isotropy.

Note that if

$$\underset{\sim}{\Sigma} = (\sigma^2 - \psi) \underset{\sim}{\Sigma}^* + \psi \underset{\sim}{I},$$

where $\underset{\sim}{\Sigma}^*$ is a correlation matrix, then ψ is called the nugget term and σ^2 is the sill. Further, if $\sigma(.) = 0$ for $h > \alpha$ but $\sigma(.) \neq 0$ for $h < \alpha$ then α is called the range. If the range is known to be infinite then there are only two parameters σ^2 and ψ to be estimated. Note that the common notation in geostatistics for σ^2, ψ and α is $C + C_o$, C_o and a respectively.

A simple departure from isotropy in two dimensions can be indicated by

$$\underset{\sim}{\Lambda} = \begin{bmatrix} 1 & 0 \\ 0 & \lambda \end{bmatrix}, \ \lambda^2 = 1 - \phi, \ \phi < 1,$$

where λ is called the anisotropy ratio (commonly denoted by k), $\lambda > 0$ and ϕ will be called the anisotropy parameter.

In general, the preferred direction is known, and therefore this model is adequate for most practical situations, e.g. it is expected that there will be more variability in the vertical direction than in the horizontal direction. We discuss some aspects of the overall model. In particular, we may write

$$\underset{\sim}{\Lambda} = \text{diag}\ (\lambda_1, \lambda_2),$$

and consider σ^2, λ_1 and λ_2 as parameters. Its extension to $p = 3$ is also indicated.

Note that the spherical scheme is

$$\sigma(\underset{\sim}{h}) = (\sigma^2 - \psi)\ \{1 - \tfrac{3}{2}\ \frac{|h|}{a} + \frac{|h|^3}{2a^3}\}, \qquad |h| < a.$$

We also discuss the following particular schemes for two dimensions.

(i) Gaussian scheme:

$$\sigma(h,k) = \sigma^2\ \delta_1^{\ h^2}\ \delta_2^{\ k^2},$$

(ii) Doubly geometric scheme:

$$\sigma(h,k) = \sigma^2 \lambda^{|h|}\ \nu^{|k|}.$$

This can be generated by $z_{ij} = \lambda z_{i-1,j} + \nu z_{i,j-1} - \lambda\nu z_{i-1,j-1} + \varepsilon_{ij}$, where ε_{ij} are independently distributed.

Note that

$$\hat{\mu} = (\underset{\sim}{1}'\underset{\sim}{\Sigma}^{-1}\underset{\sim}{z})/(\underset{\sim}{1}'\underset{\sim}{\Sigma}^{-1}\underset{\sim}{1}).$$

For our discussion we shall assume $\mu = 0$, since if $\underset{\sim}{\Sigma}$ is known then $\hat{\mu}$ is known.

2. The Isotropic Case

2.1. Estimation

We first assume that the range is infinite or known, and there are only two parameters, σ^2 and ψ, which need estimating. Let us write $\varepsilon^2 = \psi/\sigma^2$ as the nugget coefficient. We thus have

$$\underset{\sim}{\Sigma} = \sigma^2\{(1-\varepsilon^2)\ \underset{\sim}{\Sigma}^* + \varepsilon^2 \underset{\sim}{I}\}, \tag{2.1}$$

where $\underset{\sim}{\Sigma}^*$ is known from the form of $\sigma(.)$. For example, we assume that α is known for the spherical scheme. Note that (2.1) is related to the usual Factor Model.

2.1.1. Maximum likelihood estimators

If ψ=0, we obviously have $\hat{\sigma}_0{}^2$ as the m.l.e. of σ^2 as

$$\hat{\sigma}_0{}^2 = \frac{\underset{\sim}{z}'\underset{\sim}{\Sigma}^{*-1}\underset{\sim}{z}}{n} \ . \tag{2.2}$$

The log likelihood for $\psi \neq 0$ is

$$\log L = \text{const.} - \frac{n}{2}\ \log \sigma^2 - \tfrac{1}{2} \log |\underset{\sim}{\Sigma}^*| - \tfrac{1}{2} \log |\underset{\sim}{I} - \varepsilon^2 \underset{\sim}{D}|$$

$$- \frac{1}{2\sigma^2}\ \underset{\sim}{z}'(\underset{\sim}{I}-\varepsilon^2\underset{\sim}{D})^{-1}\underset{\sim}{\Sigma}^{*-1}\underset{\sim}{z}, \tag{2.3}$$

where $\underset{\sim}{D} = \underset{\sim}{I} - \underset{\sim}{\Sigma}^{*-1}$.

Consider the spectral decomposition of $\underset{\sim}{D}$ by

$$\underset{\sim}{D} = \underset{\sim}{A}'\underset{\sim}{\Gamma}\underset{\sim}{A},$$

where $\underset{\sim}{\Gamma} = \text{diag}\ (\gamma_i)$ and $\underset{\sim}{A}$ is an orthogonal matrix. We then have

$$\log L = \text{const.} - \frac{n}{2} \log \sigma^2 - \tfrac{1}{2} \sum_{i=1}^{n} \log (1-\varepsilon^2\gamma_i)$$

$$- \frac{1}{2\sigma^2} \sum_{i=1}^{n} \frac{u_i{}^2(1-\gamma_i)}{(1-\varepsilon^2\gamma_i)}\ , \tag{2.4}$$

where $\underset{\sim}{u} = \underset{\sim\sim}{Az}$.

Hence the m.l. equations for σ^2 and ε^2 are

$$\sigma^2 = \frac{1}{n} \sum_{i=1}^{n} \frac{u_i^{\,2}(1-\gamma_i)}{(1-\varepsilon^2\gamma_i)} , \tag{2.5}$$

$$\sigma^2 \sum_{i=1}^{n} \frac{\gamma_i}{(1-\varepsilon^2\gamma_i)} = \sum_{i=1}^{n} \frac{u_i^{\,2}\gamma_i(1-\gamma_i)}{(1-\varepsilon^2\gamma_i)^2} . \tag{2.6}$$

We can solve (2.5) and (2.6) by iteration. However, on eliminating σ^2, we have

$$\frac{1}{n}\left\{ \sum_{i=1}^{n} \frac{u_i^{\,2}(1-\gamma_i)}{(1-\varepsilon^2\gamma_i)} \right\}\left\{ \sum_{i=1}^{n} \frac{\gamma_i}{(1-\varepsilon^2\gamma_i)} \right\} = \sum_{i=1}^{n} \frac{u_i^{\,2}\gamma_i(1-\gamma_i)}{(1-\varepsilon^2\gamma_i)^2} . \tag{2.7}$$

Asymptotically, $(\hat{\sigma}^2, \hat{\varepsilon}^2)$ is bivariate normal with mean vector $(\sigma^2, \varepsilon^2)$ and the covariance matrix

$$\frac{2\sigma^4}{\tau(\varepsilon)} \begin{bmatrix} \mathrm{tr}\,\{\underset{\sim}{D}^2(\underset{\sim}{I}-\varepsilon^2\underset{\sim}{D})^{-2}\} & \frac{1}{\sigma^2}\ \mathrm{tr}\,\{\underset{\sim}{D}(\underset{\sim}{I}-\varepsilon^2\underset{\sim}{D})^{-1}\} \\ & \frac{n}{\sigma^4} \end{bmatrix} , \tag{2.8}$$

where $\tau(\varepsilon) = n\ \mathrm{tr}\,\{\underset{\sim}{D}^2(\underset{\sim}{I}-\varepsilon^2\underset{\sim}{D})^{-2}\} - \{\mathrm{tr}\ \underset{\sim}{D}(\underset{\sim}{I}-\varepsilon^2\underset{\sim}{D})^{-1}\}^2$.

For $\varepsilon = 0$, we have

$$\left.\begin{aligned} \mathrm{var}(\hat{\sigma}^2) &= 2\sigma^4\ \mathrm{tr}\ \underset{\sim}{D}^2/\tau(0), \\ \mathrm{var}(\hat{\varepsilon}^2) &= 2n/\tau(0), \\ \mathrm{cov}(\hat{\sigma}^2, \hat{\varepsilon}^2) &= 2\sigma^2\ \mathrm{tr}\ \underset{\sim}{D}/\tau(0), \end{aligned}\right\} \tag{2.9}$$

where $\tau(0) = n\ \mathrm{tr}\ \underset{\sim}{D}^2 - \{\mathrm{tr}\ \underset{\sim}{D}\}^2$. (2.10)

In general, $\mathrm{tr}\ \underset{\sim}{D} = O(n)$. Hence $\tau(\varepsilon)$ is of order n^2.

Note that the asymptotic var $(\hat{\psi})$ can be written down from

$$\mathrm{var}(\hat{\psi}) = \sigma^4\ \mathrm{var}(\hat{\varepsilon}^2) + \varepsilon^4\ \mathrm{var}(\hat{\sigma}^2) + 2\sigma^2\varepsilon^2\ \mathrm{cov}(\hat{\sigma}^2, \hat{\varepsilon}^2)\ .$$

On expanding the terms in the likelihood to order ε^4, we obtain

$$\hat{\hat{\sigma}}^2 = \frac{\underset{\sim}{z}'\underset{\sim}{U}_o\underset{\sim}{z}}{n} + \hat{\varepsilon}^2\ \frac{\underset{\sim}{z}'\underset{\sim}{U}_1\underset{\sim}{z}}{n} + \hat{\varepsilon}^4\ \frac{\underset{\sim}{z}'\underset{\sim}{U}_2\underset{\sim}{z}}{n}\ . \tag{2.11}$$

Further, to the same order of approximation,

$$\hat{\hat{\varepsilon}}^2 = \frac{n\ \underset{\sim}{z}'\underset{\sim}{U}_1\underset{\sim}{z} - (\mathrm{tr}\ \underset{\sim}{D})\ \underset{\sim}{z}'\underset{\sim}{U}_o\underset{\sim}{z}}{\left[\{\mathrm{tr}\ \underset{\sim}{D}^2-(\mathrm{tr}\ \underset{\sim}{D})^2\}\underset{\sim}{z}'\underset{\sim}{U}_o\underset{\sim}{z} + (n+1)(\mathrm{tr}\ \underset{\sim}{D})\ \underset{\sim}{z}'\underset{\sim}{U}_1\underset{\sim}{z}-2n\ \underset{\sim}{z}'\underset{\sim}{U}_2\underset{\sim}{z}\right]}\ , \tag{2.12}$$

where $\underset{\sim}{U}_r = \underset{\sim}{D}^r\ \underset{\sim}{\Sigma}^{*-1}$

$$= \underset{\sim}{D}^r - \underset{\sim}{D}^{r+1}, \quad r = 0,1,\ldots$$

These expansions will be valid if tr $\underset{\sim}{D} < 1$. We can use these as initial estimators in iteration.

<u>Example</u>

For the linear Markov process

$$\underset{\sim}{\Sigma}^* = \gamma^{|i-j|},\ |\gamma|<1,$$

we have

$$\underset{\sim}{D} = \frac{-\gamma}{1-\gamma^2}\begin{pmatrix} \gamma & -1 & 0 \ldots\ldots & 0 \\ & 2\gamma & -1 \ldots\ldots & 0 \\ & & 2\gamma & -1 \\ & & & \gamma \end{pmatrix}.$$

Hence, $\hat{\sigma}^2$ and $\hat{\varepsilon}^2$ can be obtained from (2.5) and (2.7) and approximately from (2.11) and (2.12). For (2.9) $(\hat{\sigma}^2, \hat{\varepsilon}^2)$ is asymptotic normal with mean $(\sigma^2,0)$ and covariance matrix given by

$$\mathrm{var}(\hat{\sigma}^2) = \frac{2\sigma^4(1+2\gamma^2)}{n}\ ,\ \mathrm{var}(\hat{\varepsilon}^2) = \frac{(1-\gamma^2)^2}{n\gamma^2}\ ,$$

and $\mathrm{cov}(\hat{\sigma}^2, \hat{\varepsilon}^2) = \dfrac{-2\sigma^2(1-\gamma^2)}{n}$.

Note that

$$\mathrm{var}(\hat{\psi}) \approx \frac{\sigma^4(1-\gamma^2)^2}{n\gamma^2} + \frac{2\sigma^4\varepsilon^2}{n}\{\varepsilon^2(1+2\gamma^2)-2(1-\gamma^2)\} \ . \tag{2.13}$$

2.1.2. Unbiased Estimators

It can be shown that

$$\psi^* = (\underset{\sim}{z}'\underset{\sim}{D}\underset{\sim}{z})/(\mathrm{tr}\ \underset{\sim}{D})$$

and

$$\sigma^{*2} = \underset{\sim}{z}'\underset{\sim}{z}/n$$

are unbiased estimators of ψ and σ^2 respectively. For the doubly geometric process, it can be shown that

$$\mathrm{var}(\psi^*) = \frac{(1-\gamma^2)\sigma^4}{(n-1)\gamma^2} + \frac{\psi^2(3n-4)}{(n-1)^2}$$

$$+ \frac{(\sigma^2-\psi)^2\ (1-\gamma^2)\{n\gamma^{2n}\ (1+\gamma^2-\gamma^{2n-4}) + \gamma^{2n+2}\}}{\gamma^4(n-1)^2}$$

$$\approx \frac{(1-\gamma^2)\sigma^4}{n\gamma^2} + \frac{\psi^2}{n} \ . \tag{2.14}$$

Hence for $\psi = 0$, its asymptotic relative efficiency from (2.13) is

$$1 - \gamma^2,$$

which shows the efficiency decreases with large values of γ.

2.2. Hypothesis of no nugget effect

We can test the hypothesis

$$H_o : \varepsilon = 0$$

against

$$H_1 : \varepsilon \neq 0 \ .$$

We can easily write the likelihood ratio test. For small ε^2, we have

$$-2 \log_e \Lambda = n \log_e(\hat{\sigma}_1^2/\hat{\sigma}_0^2) - \hat{\varepsilon}^2\{\mathrm{tr}\ \underset{\sim}{D} + \hat{\varepsilon}^2\ \mathrm{tr}\ \underset{\sim}{D}^2\},$$

where $\hat{\sigma}_0^2$ is given by (2.2) and $\hat{\sigma}_1^2$ and $\hat{\varepsilon}^2$ are given by (2.11) and (2.12) respectively.

An alternative test is to use

$$\hat{\psi} \sim N(0,\ 2n/\tau(o)),$$

where $\tau(o)$ is given by (2.10). For the doubly geometric process, $\tau(o) = 2n^2\gamma^2/(1-\gamma^2)^2$. This method can easily be extended for the doubly geometric process with γ, σ^2 and ε^2 as the parameters to be estimated, i.e. $\hat{\gamma}$ is the m.l.e. of γ in $\tau(o)$. (see § 3.3 on isotropy).

2.3. Finite Range

If the range α is finite and unknown, $\underset{\sim}{\Sigma}^*$ will be a function of α. However, we assume in (2.1) that σ^2 and ψ do not involve α. We can simplify the likelihood (2.3) slightly by using

$$\log |\underset{\sim}{I} - \varepsilon^2\ \underset{\sim}{D}| = -\sum \frac{\varepsilon^{2r}\ \mathrm{tr}\ \underset{\sim}{D}^r}{r},$$

$$(\underset{\sim}{I} - \varepsilon^2 \underset{\sim}{D})^{-1} = \sum \varepsilon^{2r}\ \underset{\sim}{D}^{-r}.$$

By neglecting ε^6 and higher powers, we can write down the maximum likelihood equation for α, and proceed as in (2.11) and (2.12). However, even for $\varepsilon = 0$, we can only make progress numerically. It is useful to note that

$$\frac{\partial \underset{\sim}{\Sigma}^{-1}}{\partial \alpha} = -\underset{\sim}{\Sigma}^{-1} \left|\frac{\partial \underset{\sim}{\Sigma}}{\partial \alpha}\right| \underset{\sim}{\Sigma}^{-1}.$$

In general $\partial\underset{\sim}{\Sigma}/\partial\alpha$ can be computed.

Example.

Consider only two points z_1 and z_2, unit distance apart, from the spherical scheme

$$\sigma(h) = 1 - \tfrac{3}{2}\ \frac{h}{\alpha} + \frac{h^3}{2\alpha^3}\ ,\ 0<h\leq\alpha,$$

$$= 0 \qquad\qquad\qquad ,\ h>\alpha\ .$$

We have

$$-2 \log L = z_1{}^2 + z_2{}^2 \quad \text{for } \alpha<1$$

and

$$-2 \log L = 2 \log(1-\rho^2) - \frac{1}{2(1-\rho^2)} \{z_1{}^2-2\rho z_1 z_2+z_2{}^2\} \text{ for } \alpha\geq 1,$$

where $\rho = \sigma(1)$. The m.l.e. of ρ for $\alpha\geq 1$ is a solution of a cubic. (For known value of ρ, α can be obtained from $\rho = \sigma(1)$ which is again a cubic). Hence, even for this simple case, we cannot make much analytical progress.

3. Geometric Anisotropy

Under geometric anisotropy, we have

$$\underset{\sim}{\Sigma} = (\sigma^2-\psi)\ \underset{\sim}{\Sigma}^* + \psi\ \underset{\sim}{I}, \tag{3.1}$$

where the correlogram is given by

$$\sigma(\underset{\sim}{h}) = \sigma(\underset{\sim}{h}'\underset{\sim}{\Lambda}\underset{\sim}{h}),\ \sigma(0) = I, \tag{3.2}$$

and $\underset{\sim}{\Lambda}$ is a p×p symmetric matrix.

Thus

$$\sigma^*_{ij} = \sigma\{(\underset{\sim}{x}_i-\underset{\sim}{x}_j)'\ \underset{\sim}{\Lambda}(\underset{\sim}{x}_i-\underset{\sim}{x}_j)\}\ .$$

The parameters to be estimated are σ^2, ψ and $\underset{\sim}{\Lambda}$, i.e. $2 + \frac{p(p+1)}{2}$ in all. We will concentrate here on the case for p = 2 but the method can be generalised. We initially take $\sigma^2 = 1$, $\psi = 0$, and

$$\underset{\sim}{\Lambda} = \begin{bmatrix} 1 & 0 \\ 0 & 1-\phi \end{bmatrix}. \tag{3.3}$$

where ϕ will be called the anisotropy parameter; the larger is ϕ, the greater is the anisotropy.

3.1. Anisotropy Parameter

Let us write,

$$d^2{}_{ij} = (x_i - x_j)^2 + (y_i - y_j)^2,$$

where (x_i, y_i) is the location of $\underset{\sim}{z}_i$.

Further,

$$\Sigma_o = (\sigma(d^2{}_{ij})). \tag{3.4}$$

We have

$$\underset{\sim}{\Sigma}^* = (\sigma\{d^2{}_{ij} - \phi(y_i - y_j)^2\}) = \underset{\sim}{\Sigma}_1 \quad \text{say.}$$

Using Taylor's expansion, we have to order ϕ^2

$$\underset{\sim}{\Sigma}_1^{-1} = (\underset{\sim}{\Sigma}_0^{-1} + \phi\underset{\sim}{W}_1 - \phi^2 \underset{\sim}{W}_2), \tag{3.5}$$

where

$$\underset{\sim}{W}_1 = \underset{\sim}{V}_1\underset{\sim}{\Sigma}_0^{-1}, \; \underset{\sim}{W}_2 = (\underset{\sim}{V}_2 - \underset{\sim}{V}_1^2)\,\underset{\sim}{\Sigma}_0^{-1}, \tag{3.6}$$

with

$$\underset{\sim}{V}_1 = \underset{\sim}{\Sigma}_0^{-1}\underset{\sim}{U}_1, \quad \underset{\sim}{V}_2 = \underset{\sim}{\Sigma}_0^{-1}\underset{\sim}{U}_2,$$

$$\underset{\sim}{U}_1 = ((y_i - y_j)^2 \quad \sigma'(d^2{}_{ij})),$$

$$\underset{\sim}{U}_2 = \tfrac{1}{2}((y_i - y_j)^4 \quad \sigma''(d^2{}_{ij})),$$

where σ', σ'' are the first and second derivatives of $\sigma(h)$ with respect to d^2_{ij}.

We also have to the order of ϕ^2

$$|\underset{\sim}{\Sigma}_1| = |\underset{\sim}{\Sigma}_0| \; [1-\phi \text{ tr } \underset{\sim}{V}_1 + \tfrac{1}{2}\phi^2\{2 \text{ tr } \underset{\sim}{V}_2 + (\text{tr } \underset{\sim}{V}_1)^2 - \text{tr } \underset{\sim}{V}_1{}^2\}] \quad (3.7)$$

Using (3.6) and (3.7) in the likelihood function for $\underset{\sim}{z}$ from $N(0, \underset{\sim}{\Sigma}_1)$, we obtain

$$\hat{\phi} = b_1/2b_2, \quad (3.8)$$

where

$$b_1 = \text{tr } \underset{\sim}{V}_1 - \underset{\sim}{z}'\underset{\sim}{W}_1\underset{\sim}{z},$$

$$b_2 = \text{tr } \underset{\sim}{V}_2 - \tfrac{1}{2} \text{ tr } \underset{\sim}{V}_1{}^2 - \underset{\sim}{z}'\underset{\sim}{W}_2\underset{\sim}{z} .$$

Further, the likelihood ratio test of isotropy ($H_0:\phi = 0$, $H_1:\phi \neq 0$) is given by

$$-2 \log \Lambda = b_1{}^2/4b_2 \quad (3.9)$$

which is asymptotically distributed as $\chi_1{}^2$.

<u>Example</u>

We consider z's on a 2×2 grid with unit distances, i.e. $\underset{\sim}{z} = (z_{11}, z_{12}, z_{21}, z_{22})$ with

$$\sigma(h,k) = \beta^{h^2+(1-\phi)k^2},$$

where β is known. The answer remains the same for the 2×2 case with

$$\beta^{|h| + (1-\phi)|k|},$$

i.e. after replacing $h^2 + (1-\phi)k^2$ by $|h| + (1-\phi)|k|$ in the above discussion. We find that

$$b_1 = \frac{2a\beta}{(1-\beta^2)^3}\Big[2\beta(1-\beta^2)^2 - \beta M_{00} + (1+\beta^2)\, M_{01}$$

$$+ 2\beta^2 M_{10} - \beta(1+\beta^2)(M_{11}+M_{1,-1})\Big],$$

$$b_2 = \frac{-a^2\beta}{(1-\beta^2)^4}\Big[4\beta(1-\beta^2)^2 - 2\beta(1+\beta^2)M_{00} + (1+6\beta^2+\beta^4)M_{01}$$

$$+ 4\beta^2(1+\beta^2)M_{01} - \beta(1+6\beta^2+\beta^4)(M_{11}+M_{1,-1})\Big],$$

where $a = -\log_e\beta$, $M_{rs} = \sum_{i=r+1}^{m}\sum_{j=s+1}^{n} Z_{i-r,j-s}Z_{ij}$. Since b_1 and b_2 contain β, it is not easy to interpret either $\hat{\phi}$ or the likelihood ratio test.

3.2. Extension

We now consider a slightly extended problem. Let us assume

$$\sigma(\underset{\sim}{h}) = \sigma(\lambda_1 h^2 + \lambda_2 k^2)$$

$$= \sigma\{\lambda(h^2+(1-\phi)\ k^2)\}\ . \qquad (3.10)$$

If λ is small, we can use the idea of expansion as above, i.e. Taylor series expansion in λ_1 and λ_2 will lead to an approximate estimator.

We now illustrate this idea by the doubly geometric series,

$$\sigma(\underset{\sim}{h}) = \sigma^2\ \lambda^{|h|}\ \nu^{|k|},$$

which is a reparameterisation of (3.10).

Under H_0 : $\lambda = \nu = \gamma$, say,

H_1 ; $\lambda \neq \nu$,

where σ^2 is unknown, it is found that the likelihood ratio test for an m×n grid is

$$-2\log\Lambda = n\log(1-\hat{\lambda}^2) + m\log(1-\hat{\nu}^2) - (m+n)\log(1-\hat{\gamma}^2)$$

$$+ mn\log(\hat{\sigma}_0{}^2 - \hat{\sigma}_1{}^2) \sim \chi_1{}^2,$$

where $\hat{\lambda}$, $\hat{\nu}$, $\hat{\gamma}$, $\hat{\sigma}_0^2$, $\hat{\sigma}_1^2$ are the m.l.e's given by the following equations.

$$\hat{\lambda} = \frac{B_8 - \nu B_3 + \nu^2 B_6}{B_9 - 2\nu B_4 + \nu^2 B_7}, \tag{3.11}$$

$$\hat{\nu} = \frac{B_2 - \lambda B_3 + \lambda^2 B_4}{B_5 - 2\lambda B_6 + \lambda^2 B_7}, \tag{3.12}$$

$\hat{\gamma}$ is the solution of the cubic equation

$$2B_7\gamma^3 - 3(B_4+B_6)\gamma^2 + (2B_3+B_5+B_9)\gamma - (B_2+B_8) = 0, \tag{3.13}$$

$$\hat{\sigma}_0^2 = \frac{1}{mn}\left[B_1 - 2(B_2+B_8)\gamma + (2B_3+B_5+B_9)\gamma^2 - 2(B_4+B_6)\gamma^3 + B_7\gamma^4\right], \tag{3.14}$$

$$\hat{\sigma}_1^2 = \frac{1}{mn}\Big[B_1 - 2\nu B_2 + 2\nu\lambda B_3 - 2\nu\lambda^2 B_4 + \nu^2 B_5 - 2\nu^2\lambda B_6 + \nu^2\lambda^2 B_7$$

$$-2\lambda B_8 + \lambda^2 B_9\Big], \tag{3.15}$$

where $B_1 = M_{00}$, $B_2 = M_{01}$, $B_3 = M_{1,-1} + M_{11}$,

$B_4 = M_{01} - M_{1.}(1) - M_{m.}(1)$, $B_5 = M_{00} - M_{.1} - M_{.n}$,

$B_7 = M_{00} - M_{1.} - M_{m.} - M_{.1} - M_{.n} + Z_{11}^2 + Z_{1n}^2 + Z_{m1}^2 + Z_{mn.}^2$

B_6, B_8, B_9 can be obtained by symmetry.

Finally, $M_{i.} = \sum_{j=1}^{n} Z_{ij}^2$,

$M_{i.}(1) = \sum_{j=2}^{n} Z_{i,j}\; Z_{i,j-1}$, etc.

It is found that an asymptotic test after variance stabilisation is given by

$$(\sin^{-1}\lambda - \sin^{-1}\nu)\sqrt{mn} \sim N(0,1). \tag{3.16}$$

For large m, n we find that

$$\hat{\lambda} = \frac{(1+\hat{\nu}^2)\ r_{10} - \hat{\nu}(r_{11} + r_{1,-1})}{1-2\hat{\nu}r_{01}+\hat{\nu}^2}, \tag{3.17}$$

$$\hat{\nu} = \frac{(1+\hat{\lambda}^2)r_{01} - \hat{\lambda}(r_{11}+r_{1,-1})}{1-2\hat{\lambda}r_{10}+\hat{\lambda}^2}, \tag{3.18}$$

where $r_{ij} = M_{ij}/M_{00}$. These equations can be iterated with

$\hat{\lambda}_0 = r_{10}, \quad \hat{\nu}_0 = r_{01}$.

Similarly, $\hat{\gamma}$ satisfies for large m, n the equation

$$2\gamma^3-3(r_{10}+r_{01})\gamma^2+2(1+r_{1,-1}+r_{11})\gamma-(r_{10}+r_{01}) = 0. \tag{3.19}$$

For large m and n, and for small $\hat{\lambda}$, $\hat{\nu}$ and $\hat{\gamma}$, it is found that

$$-2 \log \Lambda \simeq \frac{mn(r_{10}-r_{01})^2}{(1-r_{1,1}-r_{1,-1})}\left[1+\frac{(r_{10}-r_{01})^2}{(1-r_{11}-r_{1,-1})}+\frac{2(r_{10}+r_{01})^2}{(1+r_{11}+r_{1,-1})}\right] \tag{3.20}$$

or, approximately, $-2 \log \Lambda \simeq mn\ (r_{10} - r_{01})^2/(1-r_{1,1}-r_{1,-1})$.

It may be noted that for m=n=2,

$$-2 \log \Lambda = \frac{4(r_{10}-r_{01})^2}{1-2(r_{11}+r_{1,-1})}\left[1+\frac{2(r_{01}-r_{10})^2}{[1-2(r_{11}+r_{1,-1})]}\right.$$

$$\left.+\frac{4(r_{01}+r_{10})^2}{[1+2(r_{11}+r_{1,-1})]}\right]. \tag{3.21}$$

The factor of 2 in $[1-2(r_{11}+r_{1,-1})]$ in (3.21) is due to the fact that M_{00} has 4 observations while $r_{1,1}$ and $r_{1,-1}$ each have only 1 observation.

Hence, it measures the discrepancy between $\hat{\lambda}$ and $\hat{\nu}$ or rather M_{01} and M_{10}, the correlation in the x- and y- directions, i.e. the variogram in the x- and y- directions. This is similar to $\sigma^2\ \delta_1^{h^2}\ \delta_2^{k^2}$ for small λ_1 and λ_2. However, the work is not as easy.

Remark : general case.

The main aim is to estimate five parameters in (3.1) for two dimensions. We can work analytically for small ψ, λ_1, λ_2. We can use these estimators as initial estimators in a Newton-Raphson routine. We can then estimate $\underset{\sim}{\Lambda}$ if $\sigma(\underset{\sim}{h})$ has a finite range. Then

$$\sigma(r\cos\theta,\ r\sin\theta) = \sigma\left(\frac{r^2}{b(\theta)}\right),\quad r^2 \leq b(\theta),$$

$$= 1 \qquad ,\quad r^2 = b(\theta),$$

where $\{b(\theta)\}^{-1} = (\cos\theta,\ \sin\theta)'\ \ \underset{\sim}{\Lambda} \begin{pmatrix}\cos\theta \\ \sin\theta\end{pmatrix}$.

After estimating ranges, we can obtain an estimate of $\underset{\sim}{\Lambda}$ by fitting an ellipse, (see Mardia & Holmes, 1980).

4. Tests of Anisotropy Based on the Variogram

Although the previous method outlines the aim and formulates the problem, we can only do further work numerically. Also, $\underset{\sim}{\Sigma}^{-1}$ is required in calculations which can be of very large order, e.g. for a 15×15 grid, $\underset{\sim}{\Sigma}^{-1}$ is of order 225×225. Hence, for assessing isotropy, we require other methods. One of the common methods is to calculate sample variograms in four directions N-S, E-W, NE-SW, NW-SE. We assume that points are on a planar lattice for subsequent discussion. Let us assume that the Z_{ij} are normal with

$$\sigma(h,k) = \sigma^2 \quad \gamma^{|h|+|k|} \tag{4.1}$$

Suppose that $\quad \underset{\sim}{v}_1{}' = (v_{11},\ v_{12},\ \ldots,\quad v_{1p}),$

$$\underset{\sim}{v}_2{}' = (v_{21},\ v_{22},\ \ldots,\quad v_{2p})$$

be variograms in the x- and y- directions at lags 1, 2, ..., p respectively. We assume that p is known (usually taken to be 10 or so when unknown).

Independence Case

For $\gamma=0$, Zij are independent normal. It can be shown that

$$\begin{pmatrix} \underset{\sim}{v_1} \\ \underset{\sim}{v_2} \end{pmatrix} \sim N\left[\begin{pmatrix} \underset{\sim}{1} \\ \underset{\sim}{1} \end{pmatrix}\sigma^2,\ \frac{\sigma^4}{n}\begin{pmatrix} \underset{\sim}{\Sigma_{11}}, & \underset{\sim}{\Sigma_{12}} \\ \underset{\sim}{\Sigma_{12}}, & \underset{\sim}{\Sigma_{22}} \end{pmatrix}\right],$$

where

$$\underset{\sim}{\Sigma_{11}} = 4\,\underset{\sim}{I} + 8\,\underset{\sim}{1}\underset{\sim}{1}', \qquad \underset{\sim}{\Sigma_{12}} = 8\,\underset{\sim}{1}\underset{\sim}{1}'.$$

A test to measure differences in $\underset{\sim}{v_1}$ and v_2 is

$$n\,\frac{(\underset{\sim}{v_1} - \underset{\sim}{v_2})'\ (\underset{\sim}{v_1} - \underset{\sim}{v_2})}{8\sigma^4} \sim \chi^2_p. \tag{4.2}$$

Note that corr $(v_{1i}, v_{1j}) = 2/3$, $i \neq j$ although

$$\underset{\sim}{v_1} - \underset{\sim}{v_2} \sim N\left(0, \frac{8\sigma^4}{n}\right).$$

To remove the effect of scaling, i.e. σ^4, we use the variance stabilising transformation. Then

$$\frac{n}{2}\sum_{i=1}^{p}\left[\log\frac{v_{1i}}{v_{2i}}\right]^2 \sim \chi^2_p \tag{4.3}$$

This also leads to the recommendation that $\log v_{1i}/v_{2i}$ should be plotted rather than (v_{1i}, v_{2i}).

Small γ

For small γ, it can be shown that

$$\frac{n}{8\sigma^4}\left(\sum_{i=1}^{p} b_i^2 - 4\gamma \sum_{i=2}^{p} b_i b_{i-1}\right) \simeq \chi^2_p, \tag{4.4}$$

where $b_i = v_{1i} - v_{2i}$, $i=1, \ldots, p$.

This suggests that $b_1, \ldots, b_p$ form a linear Markov chain.

$$b_i = \beta\, b_{i-1} + \varepsilon_i, \quad \beta = 2\gamma,$$

where $\varepsilon_i \sim IN(0, \sigma_\varepsilon^2)$, $\sigma_\varepsilon^2 = \sigma_b^2\ (1-\beta^2)$, $\sigma_b^2 = \frac{8\sigma^4}{n}$.

Hence, for the same order of approximation, we can use

$$\underset{\sim}{b}'\ \Sigma_{\sim b}^{-1}\ \underset{\sim}{b} \simeq \chi^2_p \tag{4.5}$$

as the criterion, i.e.

$$\frac{n}{8}\ \frac{(b_1^2 - \beta^2\ b^2_p)}{\sigma^4(1-\beta^2)} + \frac{n}{8}\sum_{i=2}^{n} \frac{(b_i - \beta b_{i-1})^2}{(1-\beta^2)\sigma^4} \simeq \chi^2_p. \tag{4.6}$$

We can estimate β either by the maximum likelihood method or the least squares method. Then one degree of freedom should be removed from χ^2. Note that (4.4) also holds for

$$\sigma^2\ \gamma^{\,h^2+k^2}.$$

For small γ (or β), it can be shown that (4.6) reduces to

$$\frac{n}{2}\ \frac{(1-\beta)^2\ u_1^2 - \beta^2\ u^2_p}{1-\beta^2} + \frac{n}{2}\sum_{i=2}^{p} \frac{(u_i - \beta u_{i-1})^2}{(1-\beta^2)} \simeq \chi^2_p, \tag{4.7}$$

where $u_i = \log v_{1i}/v_{2i}$. This statistic does not depend on the scale parameter σ^2. These are also equivalent for small γ to the statistic

$$\frac{n}{2}\sum_{i=2}^{p} \frac{(b_i - \beta b_{i-1})^2}{1-\beta^2} \simeq \chi^2_{p-1}. \tag{4.8}$$

Note that if we use only the first term $nb^2{}_1/8\sigma^4$ in (4.6) as a test statistic, it is very similar to the likelihood ratio test given by (3.20) for the doubly geometric case.

[For the power scheme h^{θ} we should use $u_i = \log [\log v_{1i})/(\log v_{2i})]$ in (4.8) in place of u_i, to remove the effect of scaling.]

Note that γ will be small when the grid size is slightly less than the range.

Remark 1.

General Case. It is expected that (4.6) with $\beta=\hat{\beta}$ should prove useful for general γ. In fact, $\Sigma_{\sim u}$ is known asymptotically for the doubly geometric scheme in terms of γ and thus (4.5) can be computed after estimating γ. The test (4.6) indicates using an autoregressive series on the u_i's. In practice,

$$u_i = \lambda_1 u_{i-1} + \lambda_2 u_{i-2} + \varepsilon_i \tag{4.9}$$

may be preferable for any practical models with two parameters such as the spherical scheme with no nugget effect. We can easily test the efficiency of (4.9) by using a generalization of (4.6).

Remark 2.

Instead of using the log-transformation, we could use $(v_{1i}-v_{2i})/M_{00} = 2(r_{2i}-r_{1i}) = c_i$ to remove the effect of scaling. The tests (4.4) to (4.6) remain identical on replacing $(v_{1i}-v_{2i})/\sigma^2$ by $2(r_{2i}-r_{1i})$, e.g.

$$\frac{n}{2} \sum_{i=2}^{n} \frac{(c_i-\hat{\beta}\, c_{i-1})^2}{1-\hat{\beta}^2} \sim \chi^2_{p-2} \tag{4.10}$$

In practice, (4.8) is found to be more robust where $\hat{\beta}$ is a least square estimator (Mardia and Gill, 1980), and then (4.8).

Remark 3.

We have concentrated on only two directions which are orthogonal. In practice, one knows in general the preferential direction of anisotropy a priori, e.g. more variation in the vertical direction of a mine than in the horizontal direction. If not, one can take the x-axis corresponding to the major axis of variation and the y-axis to the minor axis.

Remark 4

Alternatively if $\underset{\sim}{v}_3$ and $\underset{\sim}{v}_4$ are variograms in directions NE-SW and NW-SE then using the asymptotic distributions of $(\underset{\sim}{v}_1-\underset{\sim}{v}_2, \underset{\sim}{v}_3-\underset{\sim}{v}_4)$, an omnibus test can be constructed.

ACKNOWLEDGEMENTS

I wish to express my thanks to Andrew Baczkowski, David Cox, Rob Edwards, Christine Gill and John Kent for their valuable help and comments.

REFERENCES

Journel, A.G. and Huijbregts, Ch. J. (1978). Mining Geostatistics. Academic Press.

Mardia, K.V. and Gill, C.A. (1980). Some statistical inference problems in kriging I : numerical applications, J. Geol. Soc. London. (to appear).

Mardia, K.V. and Holmes, D. (1980). A statistical analysis of megalithic data under elliptic pattern. J. Roy. Statist. Soc. A (to appear).

Martin, R.J. (1979). A subclass of lattice processes applied to a problem in planar sampling. Biometrika 66, 209-217.