ACTA STEREOL 1990; 9/1: 9-14
PROC ECS 5 FREIBURG I. BR. 1989
REVIEW ARTICLE

INTRODUCTION TO MARKOV SPATIAL MODELS

Adrian Baddeley
Centre for Mathematics & Computer Science
Kruislaan 413, Amsterdam 1098 SJ, Netherlands

ABSTRACT

We draw the attention of stereologists to some recent developments in the theory of random spatial patterns (Arak and Surgailis, 1989; Baddeley and Møller, 1989; Clifford, 1989; Ripley, 1989) which may be useful in the study of spatial interactions in a multiphase structure. This paper is intended as a non-mathematical introduction to the field. We describe the basic concept of a Markov chain, then briefly survey the new developments.

KEYWORDS:

Covariance; Markov processes; nearest-neighbour interaction; random mosaics; spatial statistics.

1 INTRODUCTION

When studying a material containing several phases, or a biological tissue consisting of different compartments, it is natural to ask whether the arrangement of phases is 'completely random', or whether some phases tend to be associated.

As a 'volumetric' measure of phase association one can use the covariance function (Serra, 1982). The cross-covariance $C_V^{\alpha,\beta}(r)$ of phases α, β at displacement r is defined as the probability that a randomly-chosen point x in three dimensions will have phase label α while simultaneously the point $x + r$ will have phase label β. If the structure is isotropic, this can be estimated stereologically, from the same cross-covariance on a random section:

$$\hat{C}_V^{\alpha,\beta}(r) = C_A^{\alpha,\beta}(r) \tag{1}$$

see Greco et al (1979) for an application.

Kroustrup et al (1988) develop an analysis of *surface affinities* which is appropriate when the phases of interest are in contact with each other. The total area of interface surface between two specified phases α, β is estimated using standard stereological methods:

$$S_V(\beta : \alpha) = \frac{4}{\pi} B_A(\beta : \alpha) \tag{2}$$

These can then be studied for evidence of non-random association.

Another approach, potentially more powerful than the preceding ones, might be to take the whole record of phase labels along a linear transect and analyse this record as a 1-dimensional random process. Transect methods are especially popular in botany, forestry and ecology: see, e.g. Matern(1987), Pielou (1964, 1965), Ripley (1977). We discuss this approach further in §4.

2 MARKOV CONCEPT

This section explains the basic concept of a Markov chain. We illustrate it with a non-stereological example: the problem of analysing a passage of English text, such as

```
once upon a time there was a woodcutter who lived ...
```

As a first step we could count the frequency of each letter, finding that (say) 12% of the letters are e, 6% are t and so on. A useful way to visualise these frequencies is to *simulate* a stream of random letters drawn from the given probability distribution. This can be done with the help of a computer random number generator:

```
r,e.gtaa tt eo set s F ue ioNb ostheealmitoldtyee e
hsyt h r o. hsh,w,o N.nu. "ni euMuieoelhemed coo a,
n,  fo oen o ooa isivrhno f   so yb ys sbeg nefols
```

As a description of the original text, this simulation is equivalent to the table of letter frequencies. The random text reveals the limitations of such an analysis. Letter frequencies in text could be compared to volume fractions V_V of phases in a multiphase material in stereology; they contain information about proportions but no information about 'pattern'. See the recent comments of Weibel (1989).

For more detailed information, we need to study the dependence between successive characters. The number of times a letter n is followed by a letter c, divided by the total number of appearances of the letter n, is an estimate of the probability of a "transition" from n to c. Once again, this frequency information can be represented by simulating a stream of random text, this time using the transition probabilities to generate the random sequence:

```
thoncl Hedba Che wany ss be iler man y cll o he.
ft graingad Shee de t br con haved povedo ourwo
" mindy.." he buns f  all t  of nou.  Min an in
```

This is an example of a **Markov chain** (see e.g. Feller, 1965, chap. XV). A Markov chain is more formally defined as a sequence of random "states"

$$X_1, X_2, \ldots$$

with the so-called 'short memory property' that the conditional distribution of X_n given all the previous states $X_1, X_2, \ldots X_{n-1}$ depends only on the immediately preceding state X_{n-1}:

$$\mathbf{P}(X_n | X_1, X_2, \ldots X_{n-1}) = \mathbf{P}(X_n | X_{n-1}) \tag{3}$$

If the previous state was $X_{n-1} = \alpha$ and the new state is $X_n = \beta$ we have observed a *transition* from α to β. "Short memory" means that the next transition does not depend on the previous history of transitions, only on the current state.

The structure of a Markov chain is completely described by its transition probabilities

$$p_{n,\alpha,\beta} = \mathbf{P}(X_n = \beta | X_{n-1} = \alpha) \tag{4}$$

and we usually assume the chain is *homogeneous in time*, meaning that $p_{n,\alpha,\beta} = p_{\alpha,\beta}$ does not depend on n. These probabilities can then be estimated by counting the frequencies of transitions in a long sequence of observations.

Note that the definition of 'state' is rather abstract, and we can define a 'state' to be anything appropriate to the problem. Returning to the example of English text, instead of regarding a single character as a state, we could take characters a pair at a time, so that the text

```
Once upon a time
```

is interpreted as `On, nc, ce, ...` The transition probabilities would then describe the frequency with which a pair of characters on is followed by a character c, say. Following is a simulation from this Markov chain:

```
I ke wassit re barounks brolurn then,
 all (whilike colustat I sherey 'win
witerk, cally dere; eaveas a Punk up
wer mark. I the fouggian't do thirs.
```

This can be extended to three-character states:

```
On he fruit books a contic wer rings a bettle homes
suit get offed throughest as told bear fore as of
hiking wing the of when, madly he is in the accome,
who it borate rite low whoever acces I doorwegians.
```

Alternatively we can define a state to consist of an entire word, and count the frequency with which a given word is followed by another given word. A simulation from the corresponding Markov chain follows:

```
to return a double room for breath, out
a mascara pen and announced, in the door
opened and, gasping for the waitress came
by accident, but he's well aware of you
```

And for a two-word-state Markov chain:

```
book, sees plenty of parking directly in front of
the pall bearers and plunged out of a lisp, which
made her more beautiful to him. A few days later,
```

And for a three-word-state Markov chain:

```
the lobby. After a minute he comes back,
with the girl on his arm.  "Fancy meeting
my wife here," he says to the pig
```

The main point of this section is that a complex process such as English text can be represented or described, tolerably well, by a simple random model. Increasingly faithful descriptions can be made by redefining the notion of 'state' or equivalently by adding higher levels of dependence.

3 SPATIAL MARKOV PROCESSES

A Markov chain is usually regarded as an ordered sequence progressing in "time" (see, e.g. Serra, 1982, p. 551), but time can also be viewed as one-dimensional "space". The probability of an entire sequence of states

$$\alpha\alpha\beta\alpha\beta\alpha\beta\beta\beta\ldots$$

can be computed by multiplying the conditional probabilities:

$$c.\prod_n p_{X_n,X_{n+1}} = c.p_{\alpha,\alpha}p_{\alpha,\beta}p_{\beta,\alpha}p_{\alpha,\beta}p_{\beta,\alpha}p_{\alpha,\beta}p_{\beta,\beta}p_{\beta,\beta}\cdots \quad (5)$$

a product of terms involving only successive pairs of states. We can therefore re-formulate this as a spatial process:

- "space" consists of a line of discrete "sites" numbered $1, 2, 3, \ldots$.
- at each site n, there is a "state" X_n.
- the conditional distribution of the state at site n *given the states at all other sites*, depends only on the states at the two immediate neighbour sites $n-1$ and $n+1$.

$$\mathbf{P}(X_n|X_i, \text{all } i \neq n) = \mathbf{P}(X_n|X_{n-1}, X_{n+1}) \quad (6)$$

It can be verified that this is equivalent to the original definition. The concept of 'transition' between successive states in time has been replaced by the concept of spatial 'interaction' between adjacent sites.

In this form, Markov chains can be generalised to random processes with an arbitrary collection of "sites" with arbitrary connections, i.e. an arbitrary *graph*. For example, the points of a two-dimensional rectangular or hexagonal grid could be taken as the "sites" for a spatial Markov process.

Say that two sites m, n in a finite graph are *neighbours* if they are joined by an edge in the graph, and denote this by $m \sim n$. A Markov process on a graph G is a system of random state values $\{X_n : n \in G\}$ such that

> The conditional distribution of X_n given all the other values X_i depends only on the sites that are neighbours of n:

$$\mathbf{P}(X_n|X_i, \text{all } i \neq n) = \mathbf{P}(X_n|X_m : m \sim n) \quad (7)$$

An important example is the *Ising process* developed as a model for quantum magnetisation. Sites form a regular (usually square) grid in two dimensions. Each site n represents an atom. Atoms are magnetised either "up" (+1) or "down" (-1). The magnetisation of each atom is random but affected by its immediate neighbours.

The conditional probabilities $\mathbf{P}(X_n | X_m : m \sim n)$ describe the probability distribution of the state of an atom given the states of the atoms immediately surrounding it. For the Ising process these probabilities are

$$\mathbf{P}(X_n = +1 | X_m : m \sim n) = \frac{e^{\gamma N(+)}}{e^{\gamma N(+)} + e^{\gamma N(-)}} \tag{8}$$

where γ is an *interaction parameter* and $N(+)$ = number of neighbours with value +1, $N(-)$ = number of neighbours with value −1. Depending on the value of γ, the conditional distribution of X_n can be strongly or weakly influenced by the neighbouring values: $\gamma = 0$ produces no interaction, so that all states are stochastically independent; $\gamma = \infty$ produces a deterministic 'majority vote' in which X_n must agree with the majority of its neighbours.

The probability of the entire pattern of +1's and −1's in an Ising process is

$$\mathbf{P}(X_n : n \in G) = c \exp(\gamma \sum_{n \sim m} X_n X_m) \tag{9}$$

i.e. a product of interaction terms associated with each pair of neighbours.

An interesting alternative definition of the spatial Markov process is as follows. Consider any subset A of sites in the graph. Define the *boundary* of A to be the set ∂A of sites b such that $b \sim a$ for some $a \in A$, but $b \notin A$. Then a random process $\{X_n : n \in G\}$ on the graph G is said to have the *spatial Markov property* if the conditional distribution of the states inside A given the states outside A, depends only on the states in the boundary region:

$$\mathbf{P}(X_n : n \in A | X_i : i \notin A) = \mathbf{P}(X_n : n \in A | X_m : m \in \partial A) \tag{10}$$

The celebrated **Hammersley-Clifford theorem** states that any spatial Markov process on a discrete graph can be expressed in terms of interactions between pairs of neighbours (and possibly triples of neighbours, quadruples, ...) For further information see Kindermann & Snell (1980).

4 MARKOV PROCESSES IN CONTINUOUS SPACE

Pielou (1964, 1965) studied two-dimensional patterns of vegetation using transect methods, in which the vegetation type (phase label) is recorded along a linear transect of the pattern. She proposed that the transect data be analysed as if they came from a (one-dimensional) Markov process.

This was criticised by Bartlett (1964) on the grounds that there was no known 2-dimensional random process with the property that any transect is Markov. Switzer (1965) then constructed such a process, as follows. Generate random lines in the two-dimensional plane (according to a Poisson process). The lines divide the plane into polygons. Fill the polygons with random colours α, β independently of each other. It is easily seen that line transects of this process are one-dimensional Markov processes.

Recently, Switzer's construction has been generalised to a more flexible class of random mosaics by Arak and Surgailis (1989) and Clifford and Middleton (Clifford, 1989) A *mosaic* is a function on the continuous plane that takes a finite number of values $\alpha, \beta, \ldots$ and has discontinuities only at linear boundaries. A *Markov random mosaic* is a random mosaic with the spatial Markov property that

> For any domain D in the plane, the mosaic inside D is conditionally independent of the mosaic outside D, if we are given the phase state at all points of the boundary of D and also the angles of the discontinuity lines where they hit D.

Random mosaics might be used as models of minerals or other materials in which all phase interfaces are planar. At present it is not well understood how to perform statistical inference for Markov random mosaics, or even how to simulate them efficiently.

Markov random fields (i.e. "grey-level-valued" functions on continuous n-dimensional space) have been extensively studied by the Russian school of probability theory (see e.g. Rozanov, 1982). 'Semi-Markov' random spatial processes were introduced by Matheron, Jeulin and Serra (Serra, 1982).

5 MARKOV POINT PROCESSES

Another approach to stochastic modelling of two- or three-dimensional spatial patterns, appropriate for describing the locations of cells, cell bodies etc., is to idealise the locations of the objects as points (possibly with additional information attached to each point). The outcome of a point process is thus a pattern of random points $\{X_1, \ldots, X_n\}$ in the window of observation. Techniques for analysing point patterns are presented e.g. by Ripley (1977) and Diggle (1983, 1986).

Markov point processes were defined by Ripley and Kelly (1977) briefly as follows. Say that two random points X_i, X_j are *neighbours* if the distance between them is less than r ("interaction distance"). For any domain D in the plane, define the *r-boundary* of D to be the region outside D but less than r units away from D. Then a Markov point process is a random pattern of points with the property that for any domain D in the plane, the pattern of points inside D is conditionally independent of the pattern outside D, if we are given the positions of all points in the r-boundary of D.

Ripley and Kelly proved an analogue of the Hammersley-Clifford theorem (see section 3) which characterises all such processes in terms of an "interaction" between neighbouring points.

Markov point processes as defined above have the drawback that interactions between points must occur over a fixed range, and will occur between all points within that range. This has recently been overcome by Baddeley and Møller (1989), who construct random point processes in which the nature of interactions is allowed to depend 'dynamically' on the geometry of the pattern. For example, we may have interactions that only occur between points that are neighbours in the Voronoi tessellation generated by the pattern. This opens the way to constructing random point patterns with interactions that depend on spatial configuration and not just inter-point distances.

In place of "points" we may consider other geometrical objects such as circles, line segments, convex sets etc. using the standard technique of 'marking' each point with additional information giving the circle radius, line segment length and direction etc. The stochastic models obtained in this way — Markov set processes — are a natural generalisation of the Boolean model.

6 CONCLUSION

It may soon be possible to analyse spatial dependence between rock phases, biological cell types, etc. using an explicit stochastic spatial model. Advantages of explicit modelling are: the possibility of more detailed information; the ability to use estimates of model parameters as summaries of data; and the availability of statistical inference, such as tests of 'significance'.

Further theoretical work remains to be done in order to decide how to estimate the parameters of these stochastic models from stereological data, and in particular, about the relationship between a three dimensional Markov spatial process and its plane sections.

ACKNOWLEDGEMENTS

I thank Dr D Jeulin and Dr J P Kroustrup for helpful comments.

REFERENCES

Arak RJ, Surgailis D. Markov fields with polygonal realisations. *Probability Theory and Related Fields* 1989; 80: 543–579.

Baddeley A, Møller J, Nearest-neighbour Markov point processes and random sets. *International Statistical Review* 1989; 57: 89–121.

Bartlett MS. A note on spatial pattern. *Biometrics* 1964; 20: 891–892.

Clifford P. Markov random fields in statistics. *John Hammersley Festschrift*, to appear.

Diggle P. *Statistical Analysis of Spatial Point Patterns*. London: Academic Press 1983.

Diggle P. Displaced amacrine cells in the retina of a rabbit: analysis of a bivariate spatial point pattern. *J. Neuroscience Methods* 1986; 18: 115–126.

Feller W. *An introduction to Probability Theory and its Applications*, 3rd ed., Wiley 1965.

Greco A, Jeulin D, Serra J. The use of the texture analyser to study sinter structure: application to the morphology of calcium ferrites in basic sinters of rich iron ores. *J. Microscopy* 1979; **116**: 199–211.

Kindermann R, Snell JL. *Markov random fields and their applications*. American Math. Soc. (Contemporary Mathematics, vol. 1) Providence: AMS 1980

Kroustrup JP, Gundersen HJG, Vaeth M. Stereological analysis of three-dimensional structure organization of surfaces in multiphase specimens: statistical models and model-inferences. *J. Microscopy* 1988; **149**: 135–152.

Matérn B. *Spatial Variation* (3rd. ed.) Springer Verlag 1987.

Pielou EC. The spatial pattern of two-phase patchworks of vegetation. *Biometrics* 1964; **20**: 156–157.

Pielou EC. The concept of randomness in patterns of mosaics. *Biometrics* 1965; **21**: 908–920.

Ripley BD. *Spatial Statistics.* John Wiley 1977.

Ripley BD. Gibbsian interaction models. In Griffiths DA, ed. *Spatial statistics: past, present and future.* New York: Image, 1989, pp. 1–19.

Ripley BD, Kelly FP. Markov point processes. *J. London Math. Soc.* 1977 **15**: 188–192.

Rozanov Yu.A. *Markov random fields.* Springer-Verlag, NY 1982.

Serra J. *Image analysis and mathematical morphology.* Academic Press 1982.

Switzer P. A random set process in the plane with a Markovian property. *Ann. Math. Statist.* 1965; **36**: 1859–1863.

Weibel ER. Evolution of stereological methods. *J. Microscopy* 1989; **155**: 393–404.