Week 2.

Stat 222, Spatial Statistics.
Lecture MWF 9am, Math-Sci 5203.
Professor: Rick Paik Schoenberg, frederic@ucla.edu, www.stat.ucla.edu/~frederic

DAY FOUR. Monday, 4/9/01.

1) Translations and ergodicity.

Motivation behind ergodicity: want $\bar{Y}$ to converge to $E(Y)$, and the sample covariances to converge to the true covariances $C(h)$.

Sometimes given as convergence to a stationary distribution.

Let T_h be a translation (or shift) operator by the vector h. So given a set A like $\{\omega : Z(s) > 0\}$, $T_h(A) = \{\omega : Z(s+h) > 0\}$.

And $T_h^{-1}(A)$ is the set $\{\omega : Z(s-h) > 0\}$.

Z is stationary if $P\{T_h^{-1}(A)\} = P(A)$, for any h and any measurable set A.

If Z is stationary and also,

whenever $T_h^{-1}(A) = A$, $P(A) = 0$ or 1,

then Z is ergodic.

[Question: Suppose Z is any stationary process in the plane $\mathbf{R}^2$. Think of a set A such that $T^{-1}(A) = A$.]

Why this definition?

Suppose all the Z's are guaranteed to be the same, as in the 100%-correlation process. Let A = the event that all the Z's are positive. Then $T_h^{-1}(A) = A$ for any h, but $P(A)$ might not be 0 or 1, so this type of process is not ergodic.

The basic reason for the definition is that we want to know that values of the process at far-away locations are getting close to independent. That way, as the domain of our observations expands, our sample values will get closer and closer to the true values. For a stationary *Gaussian* process, if $C(h) \to 0$ as $|h| \to \infty$, then the process is ergodic.

Note that sets like $\{Z > 0$ infinitely often $\}$ or $\{Z > 0$ eventually (i.e. outside of some sphere) $\}$ satisfy $T_h^{-1}(A) = A$, so for ergodic processes these types of events must have probability 0 or 1. This shows the essential feature of ergodic processes: they cannot have their limiting behavior be sometimes one way and sometimes another way. Note that for WN, these probabilities must be 0 or 1 (Kolmogorov's 0-1 law).

Note that in 2-d, it is not so easy to define things like *tail* events, or random walks, etc. because of the lack of a natural ordering. To make this point:

[Question: Suppose G_n is the $n \times n$ grid or *lattice* in $\mathbf{R}^2$. That is, G_n = the set of points $\{(x,y) : x = 0,1,2,\ldots,n; y = 0,1,2,\ldots,n\}$. Ordinarily we define the distance d between two points as the Euclidean distance: $d[(x_1,y_1),(x_2,y_2)] = \sqrt{(x_2-x_1)^2 + (y_2-y_1)^2}$. Think of a non-Euclidean distance d on G_n such that $d[s,t] = 0$ iff. $s = t$.]

DAY FIVE. Wednesday, 4/11/01.

1) Miscellaneous announcements.

a) Reader is ready at Course Reader Materials (1141 Westwood). 60-something dollars. Be pushy – they didn't see it on the list at first.
b) No class Friday 4/13/01 – Good Friday. Also no class Friday 4/20/01.

2) Rescaled processes.

Suppose $Z(s)$ is intrinsically stationary, and let $Y(s)$ be some rescaled version of Z, i.e. $Y(s) = g[Z(s)]$, where $g(x)$ is some function with at least two continuous derivatives. Then (see Cressie, p65) the variogram of Y can be expressed in terms of the variogram of Z:
$2\gamma_Y(h) = [g'(\mu)]^2 2\gamma_z(h)$, where $\mu = E[Z(s)]$.

3) Relative variogram.

Now, suppose that Z is NOT intrinsically stationary. Instead, suppose that the space D on which Z is defined can be broken up into disjoint subspaces D_i, and suppose that Z is intrinsically stationary on each subspace. That is, on D_i, Z has mean μ_i and variogram $2\gamma_i(h)$.

Further, suppose the variograms $2\gamma_i(h)$ all have the same basic shape, but they depend in some way on the mean μ_i. In other words, suppose that there's some positive function f such that $2\gamma_i(h)/f(\mu_i)$ doesn't depend on i.

The relative variogram is then defined as $2\gamma_i(h)/f(\mu_i)$.

A common example is when $f(x) = x^2$. In this case, if we let $g(x) = log(x)$, then $[g'(\mu_i)]^2 = 1/\mu_i^2$, so the formula for the variogram for rescaled processes above tells us that if we let $Y = g[Z(s)]$, then Y is approximately intrinsically stationary; that is, the variogram of Y is the same everywhere in D. (Why approximately? The mean of Y also must be constant.) This suggests a log transformation as a way of obtaining an intrinsically stationary process from a non-intrinsically stationary process Z.

Also, if $f(x) = x^{2-2\lambda}$, for some $\lambda > 0$, then the transformation $Y = g[Z(s)]$ with $g(x) = x^\lambda$ makes Y approximately intrinsically stationary.

4) Aggregation.

Often one is interested in looking at the average value of Z over different regions B_1, B_2, etc. One may thus define, for measurable subsets B of the space D, the process $Y(B) = \int_B Z(s)/|B|$. (Cressie calls this $Z(B)$ instead of $Y(B)$, but I'd prefer to keep Z as a process defined on locations, rather than sets.)

The question then is: what is the variogram of Y? Turns out that
$Var[Y(B_1) - Y(B_2)] = -\int_{B_1}\int_{B_1}\gamma(s-u)dsdu/|B_1|^2 - \int_{B_2}\int_{B_2}\gamma(s-u)dsdu/|B_2|^2 + \int_{B_1}\int_{B_2}\gamma(s-u)dsdu/[|B_1| \times |B_2|]$.

And, the covariogram of Y is given by

$Cov[Y(B_1), Y(B_2)] = \int_{B_1} \int_{B_2} C(s-u) ds du / [|B_1| \times |B_2|]$,
where C is the covariogram of Z.

5) Correlogram.

As in time series analysis, the correlogram is very useful in spatial data analysis. The correlogram $\rho(h)$ is defined as

$\rho(h) = C(h)/C(0)$, where $C(h)$ is the covariogram.

The correlogram tells you how much a value $Z(s)$ is correlated with a value h away (i.e. $Z(s+h)$). The correlogram has the usual properties one expects of correlations: for example $0 \leq \rho(h) \leq 1$, $\rho(0) = 1$, and $\rho(-h) = \rho(h)$.