

Using the Matérn Covariance Function for Maximum Likelihood Estimation of Fractal Dimension

Stima di Massima Verosimiglianza della Dimensione Frattale Attraverso la Funzione di Covarianza di Matérn

Simone Di Zio– Luigi Ippoliti – Giuliano Visini
Department of Quantitative Methods and Economic Theory
University G. d’Annunzio, Chieti-Pescara

Riassunto: In questo lavoro proponiamo uno stimatore di massima verosimiglianza per la dimensione frattale di un processo stocastico gaussiano. Lo stimatore si basa sulla funzione di covarianza appartenente alla classe di Matérn e le sue proprietà vengono esaminate mediante uno studio di simulazione.

Keywords: Fractal Dimension, Covariance Function, Fractal Index, Matérn Class.

1. Introduction

Starting from Mandelbrot’s work (1982), fractal based analysis have found extensive applications in almost all scientific disciplines. The pattern of fractal like structures can be described by means of the fractal dimension D . Unlike the concept of the Euclidean (topological) dimension, fractal dimension need not be an integer and it is a quantitative *scale-free* measure of irregularity or ‘roughness’ of patterns. If $\{Z(t): t \in \mathbb{R}^d\}$ is a stationary Gaussian stochastic process, whose realization is a d -dimensional geometric structure in $\mathbb{R}^{d+1}$ (i.e. a curve or profile if $d=1$ and a surface if $d=2$), the fractal dimension is always $d < D \leq d+1$. If $Z(t)$ is very smooth we have a fractal dimension close to d , while D approaches $d+1$ if the structure is extremely rough. A variety of methods have been proposed to determine the fractal dimension. Some of the most well known methods are: *box-counting method*, *walking-dividers method*, *spectral method* and *variogram method*. Each of them requires the estimation of a slope coefficient in a log-linear regression based on m points near the origin (Taylor and Taylor, 1991; Kent and Wood, 1997). In any case, the problem is that an estimate of the fractal dimension is difficult to produce directly, so an alternative approach is to calculate an estimate of a quantity called *fractal index*, δ , which is determined by the behaviour near the origin of the covariance function, and is generally more accessible than D itself. In particular, for a stationary Gaussian process $\{Z(t): t \in \mathbb{R}^d\}$ with covariance function $C(h) = \text{cov}\{Z(t), Z(t+h)\}$, the fractal dimension is given by

$$D = d + 1 - \delta/2, \quad (1)$$

when $C(h)$ satisfies the following approximation

$$C(h) \approx C(0) - a|h|^\delta \quad \text{as } |h| \rightarrow 0 \quad (2)$$

with $0 < \delta \leq 2$ (Kent and Wood, 1997). Notice that equation (2) implies that $\gamma(h) \approx a|h|^\delta$, where $\gamma(h)$ denotes the variogram function and, therefore, we might estimate δ by linear regression of $\log(\hat{\gamma}(h))$ on $\log(h)$, where $\hat{\gamma}(h)$ represents the empirical variogram. Once an estimate of δ is available, the estimate of the fractal dimension D is obtained by equation (1). In this work, as an alternative method from the ones cited above, we propose a Maximum Likelihood Estimator (MLE) of the fractal index of a profile (i.e. $d=1$). This estimator is based on the Matérn autocovariance function (Stein, 1999) that, as known, is characterized by a smoothing parameter, ν , that can be interpreted as a fractal index. The paper is outlined as follows. In Section 2 we illustrate the Maximum Likelihood Estimator based on the Matérn covariance function, while in Section 3 we perform a simulation study to assess the performance of the estimator. Finally, we conclude the paper with a discussion in Section 4.

2. Maximum Likelihood Estimator

Given a zero mean stationary Gaussian process $\{Z(t): t \in \mathfrak{R}\}$ with covariance function $C_\theta(h)$, the likelihood function is just the joint density of the n observations $\mathbf{Z}=(Z(t_1), \dots, Z(t_n))$ viewed as a function of the unknown parameter vector $\boldsymbol{\theta}$ characterising the covariance structure. Thus, the maximum likelihood estimate (MLE) of the parameter vector $\boldsymbol{\theta}$ is any vector of values for the parameters that maximizes the following log-likelihood function

$$l(\boldsymbol{\theta} | \mathbf{Z}) = -\frac{n}{2} \log(2\pi) - \frac{1}{2} \log \det(\mathbf{C}_\theta) - \frac{1}{2} (\mathbf{Z}^T \mathbf{C}_\theta^{-1} \mathbf{Z}) \quad (4)$$

A practical model for $C_\theta(h)$ is provided by the Matérn class of covariance functions,

$$C_\theta(h) = \frac{\pi^{1/2} \phi}{2^{\nu-1} \Gamma(\nu+0.5) \alpha^{2\nu}} (\alpha|h|)^\nu \mathbf{K}_\nu(\alpha|h|) \quad (5)$$

where ν , ϕ , α are positive parameters and $\mathbf{K}_\nu$ is a modified Bessel function of the second kind of order ν (Stein, 1999). One of the most important reasons for adopting this function is that ν plays the role of a smoothing parameter. In fact, the larger ν is, the smoother Z is and, consequently, it provides information about the roughness of the process (Stein, 1999). Furthermore, when $0 < \nu < 1$ it can be shown that in a neighbourhood of 0

$$C_\theta(h) \approx \lambda - \eta|h|^{2\nu} \quad (6)$$

where

$$\lambda = \frac{\Gamma(\nu) \sqrt{\pi} \phi}{\Gamma(\nu+0.5) \alpha^{2\nu}} \text{ and } \eta = \frac{\Gamma(\nu) \Gamma(1-\nu) \phi}{2\nu \Gamma(2\nu)}$$

Accordingly, if $\hat{\nu}$ is the ML estimate of ν , we have that comparing (6) with (2) we can propose $2\hat{\nu}$ as an estimator of the fractal index to be plugged into (1) to obtain the estimate of the fractal dimension as $\hat{D} = d + 1 - \hat{\nu}$. Notice also that standard asymptotic theory for MLE's suggests that calculating the Fisher information matrix $\mathbf{I}_\theta$ and its inverse is a fruitful way of learning about the behaviour of MLEs. For a zero mean Gaussian process the Fisher information matrix takes on a fairly simple form. Specifically, if $\mathbf{Z} \sim \text{MVN}(0, \mathbf{C}_\theta)$ then the (i,j)th element of $\mathbf{I}(\theta)$ is $\mathbf{I}_{i,j}(\theta) = \frac{1}{2} \text{tr} \{ \mathbf{C}^{-1} \mathbf{C}_i \mathbf{C}^{-1} \mathbf{C}_j \}$, where $\mathbf{C}_i = \frac{\partial \mathbf{C}_\theta}{\partial \theta_i}$ (Mardia and Marshall, 1984). To carry out this calculation requires differentiating the modified Bessel function $\mathbf{K}_\nu$ with respect to ν , and this can be done numerically.

3. Experimental Results

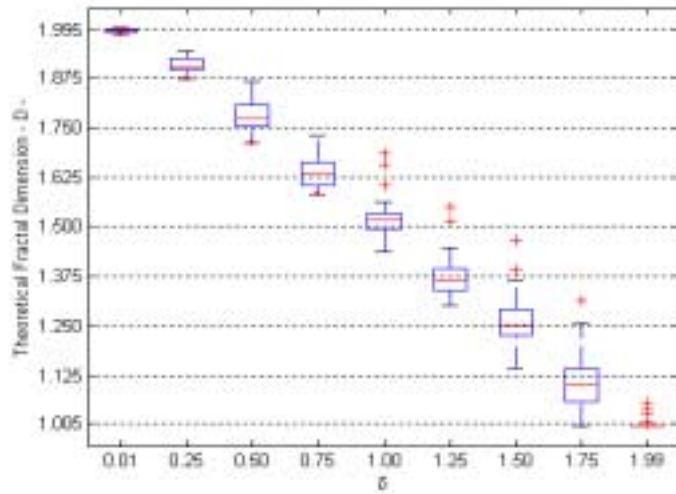
To assess the precision and accuracy of the ML estimates of fractal dimension we have performed a simulation study. In particular, using a covariance function of the type $C_\theta(h) = \exp(-\beta|h|^\delta)$, which was also used by Kent and Wood (1997) as well as Davies and Hall (1999) in their simulation studies, we have simulated 100 realizations of the zero mean Gaussian process Z for various values of $\delta \in (0, 2)$. Each realization consisted of $n=500$ equally spaced observations $Z(0), Z(0.002), \dots, Z(0.999)$ and was simulated using the *Cholesky decomposition* procedure. Table 1 and Figure 1 show, for nine different values of δ , the variations, mean, bias, standard error and mean squared error (MSE) of fractal dimension D estimated by the ML with the Matérn covariance function.

Table 1: Summary statistics for ML estimates of D . The Mean, Bias, Standard Error and Mean Squared Error (MSE) are calculated for 100 realizations.

δ	Theoretical D values	Summary Statistics			
		Mean	Bias	Standard Error	MSE
0.01	1.9950	1.9931	0.0019	0.0037	0.0000
0.25	1.8750	1.9055	-0.0304	0.0195	0.0013
0.50	1.7500	1.7772	-0.0272	0.0372	0.0021
0.75	1.6250	1.6364	-0.0114	0.0367	0.0015
1.00	1.5000	1.5197	-0.0197	0.0435	0.0023
1.25	1.3750	1.3728	0.0022	0.0479	0.0023
1.50	1.2500	1.2651	-0.0151	0.0564	0.0034
1.75	1.1250	1.1096	0.0154	0.0637	0.0043
1.99	1.0050	1.0030	0.0020	0.0111	0.0001

It can be noted that, as it happens for other fractal dimension estimators proposed in literature, also our ML estimator shows a bias which depends on the theoretical values of the fractal dimension. However, very good results can be obtained when the real value of D is very close to the boundary of its parametric space (*i.e.* 1 or 2) for which the MSE achieves its minimum values.

Figure 1: Box-plots relative to the distribution of the ML estimates of D for 100 realizations.



4. Discussion

In this paper we have proposed a ML estimator based on the Matérn covariance function. Due to its behaviour at the origin (see eq. 6), the smoothing parameter ν can be interpreted as a fractal index which plays a key role in estimating the fractal dimension D . Although a deeper study on its statistical properties is needed, at a first stage it seems that it represents a competitive alternative to the aforementioned estimators. As for other techniques, the bias problem concerning the estimation of D remains a point worth noting; however, the linear relationship between the bias and D , that we have noted in an explorative analysis, seems to provide a useful information for bias correction purposes. It can also be highlighted that the MLE is more consistent than both box-counting and variogram methods since it always estimates the fractal dimension between 1 and 2, whereas the other methods can produce estimates of D greater than 2 or less than 1. Finally, another advantage of our estimator is that it does not depend on the choice of the number of data, m , which are involved in the log-linear regression.

References

- Davies S. and Hall P. (1999) Fractal analysis of surface roughness by using spatial data, *Journal of the Royal Statistical Society*, 61, 3-37.
- Kent J. T. and Wood A. T. A. (1997) Estimating the Fractal Dimension of a Local Self-similar Gaussian Process by using increments, *J. R. Statistical Soc.*, 59, 679-699.
- Mandelbrot B. B. (1982) *The Fractal Geometry of Nature*, W. H. Freeman, New York.
- Mardia K.V. and Marshall R.J. (1984) Maximum likelihood estimation of models for residuals covariance in spatial regression. *Biometrika*, 71, 135-146.
- Taylor C. C. and Taylor S. J. (1991) Estimating the Dimension of a Fractal, *Journal of the Royal Statistical Society*, 53, 353-364.
- Stein M. L. (1999) *Interpolation of Spatial Data*, Springer-Verlag, New York.