A Quasi-likelihood Method for Fractal-Dimension Estimation

You-Gan Wang: CSIRO Mathematical and Information Sciences, P.O. Box 120, Cleveland, Qld 4163, Australia

Yan-Xia Lin: School of Mathematics and Applied Statistics, University of Wollongong, Northfields Ave, Wollongong, NSW 2522, Australia

M.D.E. Haywood: CSIRO Division of Marine Research, P.O. Box 120, Cleveland, Qld 4163, Australia.

Abstract This paper proposes a simple method of constructing quasi-likelihood functions for dependent data based on conditional-mean-variance relationships, and applies the method to estimating the fractal dimension from box-counting data. Simulation studies are carried out to compare with traditional methods. We also analysed the real data from the fishing ground in the Gulf of carpentaria, Australia.

1. Introduction

Fractal-dimension has been found useful in describing the "complexity" of random spatial patterns (Taylor and Taylor, 1989; Ogata and Katsura, 1991). There are various definitions of a "fractal dimension" proposed in the literature, such as box dimension, Hausdorff dimension, packing dimension. They may or may not have the same value for a given fractal depending on the properties of the fractal (see Cutler, 1993).

Estimation of the dimension of a fractal has become an important and interesting statistical problem(Roberts & Cronin, 1996). The box-counting data are often collected for estimating the fractal dimension because of its ease of application.

The quasi-likelihood (QL) method has been found useful in parameter estimation, especially when the distributions cannot be fully specified. However, the QL function for the dependent data probably has not been received much attention as it should (Wang, 1996). This paper proposes a way of constructing QL functions for dependent observations based on conditional moments, which naturally generalizes the original definition of the QL function for independent observations. This approach is then applied with this method to estimate the dimension of a fractal from box-counting data. The principle may be applied to other types of dimensions as well. We compare the QL with the least-squares method by analysing the simulated data from random fractals. We also analyse a data set collected from fishers in the Gulf of Carpentaria, Australia, to establish a relative capacity index for a selected area.

2. Quasi-likelihood

If Y is a random variable with $E(Y) = \mu$ and $\mathrm{Var}(Y) = V(\mu)$ (a known function), the quasi-likelihood (QL) function for Y, $Q(\mu; Y)$, is defined as (Wedderburn, 1974; McCullagh & Nelder, 1989, p. 325)

$$Q(\mu; Y) = \int_Y^\mu \frac{Y-t}{V(t)} dt. \tag{1}$$

If $Y = (Y_i)_{n\times 1}$ is a vector of n independent observations, the QL function is defined as $Q(\mu; Y) = \sum_{i=1}^n Q(\mu_i; Y_i)$, where $(\mu_i)_{n\times 1}$ is the corresponding mean vector of Y.

The QL approach only requires specification of the mean-variance relationship rather than a full likelihood function. It has been found extremely useful in modeling overdispersion problems, and there has been extensive development in this area, focusing on the case of independent observations.

It is of practical importance to consider the case of dependent data. McCullagh & Nelder (1989, p.332-336) constructed a QL function for dependent data. Unfortunately, their QL function, in general, is not uniquely determined, and it depends on the path of a line integral.

In general, if $Y^{(i)}$ is the vector of the first i observations, the log-likelihood function of $Y = (Y_i)_{n\times 1}$ can be expressed as

$$\sum_{i=1}^{n} \log\{f_i(Y_i|Y^{(i-1)})\}. \qquad (2)$$

Suppose $\mu(i)$ is the conditional expectation of $\mathrm{E}(Y_i|Y^{(i-1)})$, and the conditional variance $\mathrm{Var}(Y_i|Y^{(i-1)}) = v_i(\mu(i))$ can be expressed as a function of $\mu(i)$. We can then define the following conditional QL (strictly, the quasi-conditional-likelihood) function for Y_i,

$$Q_i(\mu(i); Y^{(i)}) = \int_{Y_i}^{\mu(i)} \frac{Y_i - t}{v_i(t)} dt. \qquad (3)$$

Using the conditional argument as in (2), the overall QL for Y can thus be defined as

$$Q(\mu; Y) = \sum_{i=1}^{n} Q_i(\mu_i; Y^{(i)}).$$

By taking partial derivatives with respect to β, the QL function results in the following estimating equations

$$\sum_{i=1}^{k} \frac{Y_i - \mu(i)}{v_i(\mu(i))} \frac{\partial \mu(i)}{\partial \beta_j} = 0, \text{ for } 1 \le j \le p. \qquad (4)$$

The estimating equations can also be written in a matrix form of

$$D'V^{-1}(Y - \mu) = 0, \qquad (5)$$

in which V is a diagonal matrix with i-th element v_i, and μ is the vector of $\{\mu(i)\}_{n\times 1}$.

Let $\mathcal{F}_i$ be the standard filtration generated by $Y_j, 1 \le j \le i$, and

$$Z_i = \frac{Y_i - \mu(i)}{v_i(\mu(i))}.$$

Clearly, $(Z_i, \mathcal{F}_i)$ is a martingale difference. Standardizing this martingale difference results in the estimating equations given by (4). Therefore, from the results of Godambe & Heyde (1987, p.236) and Heyde (1987), the QL estimating functions are optimal with respect to both the fixed sample criteria and the asymptotic criterion, and the central limit theorem holds for the estimates under appropriate regularity conditions. Multivariate Gauss approximation can then be used to evaluate the variances of the estimates (see also Liang and Zegger, 1986; Godambe & Heyde, 1987; Lin and Heyde, 1992 and 1997).

Let us now consider some examples to see how the proposed QL approach works.

Example 1: First order autoregressive process

Consider a stochastic process of

$$Y_i = h(Y^{(i-1)}, \theta) + \epsilon_i,$$

where $h(.,.)$ is a smooth function, $Y_0 = \epsilon_0$ and ϵ_i are *i.i.d.* with mean 0 and variance σ^2 (the density function is unknown). The parameter θ is of interest. Clearly, $\mu(1) = 0$ and $\mu(i) = \mathrm{E}(Y_i|Y^{(i-1)}) = h(Y^{(i-1)}, \theta)$ for $i \ge 2$. The QL estimating equation corresponding to (4) is thus

$$\sum_{i\ge 1}\{Y_i - h(Y^{(i-1)}, \theta)\}\partial h(Y^{(i-1)}, \theta)/\partial\theta = 0.$$

In particular, if $h(Y^{(i-1)}, \theta) = \theta Y_{i-1}$, the above equation becomes $\sum_i Y_{i-1}(Y_i - \theta Y_{i-1})$, which is the same as obtained by Heyde (1987) and McCullagh & Nelder (1989, p.340-341). By imposing assumptions on higher moments, Heyde (1987) also obtained a combined estimating function for θ.

Example 2: Bienayme-Galton-Watson branching process
Let $Y_0 = 1$, and $Y_{i+1} = Y_{i,1} + Y_{i,2} + ... + Y_{i,Y_i}$, where $Y_{i,j}, 1 \leq j \leq Y_i$ are i.i.d., each with the same offspring distribution, and are independent of Y_i. The offspring distribution has a mean θ, which is of interest, and variance σ^2.

The conditional mean $\mathrm{E}(Y_{i+1}|Y_i) = \theta Y_i$ and the condition variance $\mathrm{Var}(Y_{i+1}|Y_i) = Y_i\sigma^2$. Therefore, the estimating equation from the generalised quasi-likelihood becomes

$$\sum_{i=1}^{n}(Y_i - \theta Y_{i-1}) = 0.$$

Godambe & Heyde (1987) obtained the same result using some optimal estimation criteria. This solution is also the maximum likelihood estimate when the offspring has a power-series distribution (Godambe & Heyde, 1987).

3. Estimation of Box Dimension

Box dimension, also known as capacity dimension, is the most widely used index for measuring the complexity or irregularity of a fractal. It gives an idea of the relative size of the object which is too irregular to be measured by classical geometry. The dimension has also been described as the amount of Euclidean space that the fractal set fills or a measure of its roughness (Hall and Wood, 1993). Fractal dimension is analogous to the length of a line, or the area of a square and allows comparisons to be made with other fractals and with classical shapes.

For a spatial set the box-counting procedure is carried out by covering the set E with a collection of squares with a small side-length δ. Effectively, this means laying a grid of side length δ over the set and counting the minimum number of squares (N_δ) necessary to cover the set. In general, this number is proportional to the inverse of the grid size (Cutler, 1993), that is, $N_\delta \sim (1/\delta)^d$, as $\delta \to 0$, where d is the dimension of the object. The box-counting dimension $\beta(E)$, is defined as

$$\beta(E) = \lim_{\delta \to 0} \frac{\log(N_\delta(E))}{\log(1/\delta)}. \quad (6)$$

In order to estimate the dimension for a particular set E, N_{δ_i} may be obtained for a series of δ_i, $1 \leq i \leq k$. Here δ_i's are in decreasing order. When δ_k is small, one can rely on the following estimator

$$\hat{\beta}_1 = -\frac{\log(N_{\delta_k})}{\log(\delta_k)}. \quad (7)$$

This method uses only the last observed number and may not work well especially when the convergence in (6) is slow (Hall and Wood, 1993). An alternative method is to use the regression

$$y_i = \beta_0 + \beta_1 x_i + \epsilon_i,$$

where $y_i = \log(N_{\delta_i})$ and $x_i = \log(\delta_i)$. The the slope β_1 represents the box-counting dimension (Cutler 1993).

We now consider the case when $\delta_i = \delta^i$ for some constant δ $(0 < \delta < 1)$. We will use N_i for N_{δ_i} when there is no confusion. Clearly, N_{i+1} is generated from a sum of N_i observations, each taking a value between 1 and δ^{-2} (in a 2-dimension space). If we let $N_0 = 1$, we may write N_{i+1} as $N_{i,1}+N_{i,2}+...+N_{i,N_i}$, where $N_{i,j}, 1 \leq j \leq N_i$ are assumed to have the same offspring distribution. This is similar to the well known Bienayme-Galton-Watson branching process. The major difference is that it may be inappropriate to assume $N_{i,j}$ for $1 \leq j \leq N_i$ to be independent of each other. The offspring distribution has a mean $\theta = \delta^{-\beta}$, which is of interest, and variance σ_i^2.

If a box is not empty, and when being subdivided into $m \times m$ sub-boxes, we would expect m^β non-empty sub-boxes among these m^2 sub-boxes. This suggest that $\mathrm{E}(N_{i+1}|N_i) = \theta N_i$, in which $\theta = \delta^{-\beta}$. Denote the conditional variance $\mathrm{Var}(N_{i+1}|N_i)$ as V_i. The estimating equation from the generalised quasi-likelihood is

$$\sum_{i=1}^{k} \frac{(N_{i+1} - \theta N_i)N_i}{V_i} = 0. \qquad (8)$$

The conditional variance is

$$V_i = \sum_{j=1}^{N_i} \mathrm{Var}\{N_{i,j}\} + \sum_{j \neq l} \mathrm{cov}\{N_{i,j}, N_{i,l}\}.$$

It appears to be appropriate to assume that V_i is a quadratic function of N_i. The overdispersion parameters in the quadratic function have to be estimated as well. Iterative procedures can then be used to update the parameter θ and re-estimating the overdispersion parameters. If the number of data points is small, such procedures are not possible. A simpler function of V_i has to be used. In particular, if we assume V_i is proportional to N_i^2, we obtain an analytical estimator

$$\hat{\theta}_{QL} = \frac{\sum_{i=1}^{k} N_{i+1}/N_i}{k-1}. \qquad (9)$$

4. A Simulation Study

Let us now consider the fractal set generated from a unit square in two dimensions and a vector of probabilities $(p_1, \cdots, p_9)$ $(\sum_{j=1}^{9} p_j = 1)$. The unit square is divided into nine smaller squares and the probabilities $p = (p_1, p_2, \cdots, p_9)$ are assigned to these squares.

The position of a point is determined by an iterative procedure: (i) First randomly select a box of length 1/3 (each with probability p_j); (ii) Divide the selected box into nine sub-boxes of length $1/3^2$; (iii) Randomly select a sub-box among the nine (each with probability p_j); and (iv) Further divide the seleced sub-box into nine sub-boxes with length $1/3^3$, and repeat the procedures similar to (i) to (iii) until the length of the sub-box is $1/3^5$. We let one, two or three p_is be 0 and the rest p_i's have the same probability. The box-dimension of this type of fractal set is $\log(9-n)/\log(3)$ where $n = 1$, 2 or 3 is the number of p_i's with zero probability.

Figure 1 shows three fractal sets with 3000 points corresponding to the cases $n = 1$, 2 and 3. In order to estimate the dimension via the box-counting method, a grid of side length $1/3^i$ is laid over the fractal set, for each i $(i = 1, \cdots, 5)$, and the number of non-empty boxes, N_i are obtained.

Table 1: Actual and estimated fractal dimensions of the fractal sets shown in Figure 1. The standard deviations are multiplied by 1000 and given in brackets.

n	Actual	Ratio	LS	QL
1	1.893	1.537 (0.7)	1.505 (1.0)	1.602 (1.1)
2	1.771	1.524 (0.9)	1.512 (1.3)	1.559 (1.3)
3	1.631	1.493 (1.2)	1.495 (1.0)	1.500 (1.0)

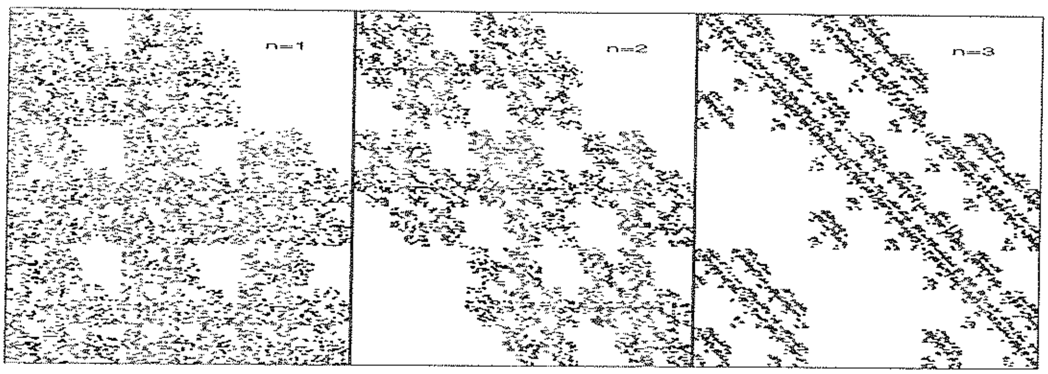

Figure 1: Three fractal sets with box-dimension $\log(9-n)/\log(3)$. Each consists of 3000 points.

Table 1 shows the mean and standard deviation of the estimates by various estimators based on 100 simulations. The ratio method is based on equation (7). Interestingly enough, it works quite well and is better than the LS method in this case. Overall, it appears that the QL method is the best.

5. Data from the Northern Prawn Fishery, Australia

The Northern Prawn Fishery has annual export earnings between \$100million and \$150 million. Some areas in the fishery produce more prawns than others. One of the reasons for this high spatial variation in catches is that survival is related to habitat type. In particular, areas of rough bottom may be associated with higher catch rates. We hypothesise that higher catch rates are related to the complexity of the reef areas rather than just the total area. Fishers in the Northern Prawn Fishery use GPS plotters for navigation and to record features such as areas of high catch, reef, rough bottom etc. We created a map of untrawlable ground within the fishery by collecting point data representing reef and rough bottom from 30 fishers and converted this data to a grid with a cell size of 200×200m. The data covered the whole of the northern prawn fishery. We hope to use box-dimension to measure the complexity for each area, which may be used to explain the variations in catches from different areas.

To demonstrate how this might work, we selected a 1° degree square (60×60 nautical miles) having an area of 12254 km^2 to the northeast of Vanderlin Island (Figure 2). Areas regarded as reef or rough bottom (untrawlable ground) covered 425 km^2 of the selected study area.

The sample area was sequentially divided up into 4, 16, 64, 256, 1024 and 4096 boxes with the number of boxes containing fractal elements being (4, 15, 56, 197, 655, 1908). The ratio method produces the estimate as $\log(1908)/(5\log(2)) \approx 1.816$. The estimates by the regression method and the QL method are 1.790, and 1.786. In this case, these two estimates are not very different. However, we would expect differences to apply to other areas because these two methods differ in general, as the simulation results indicate.

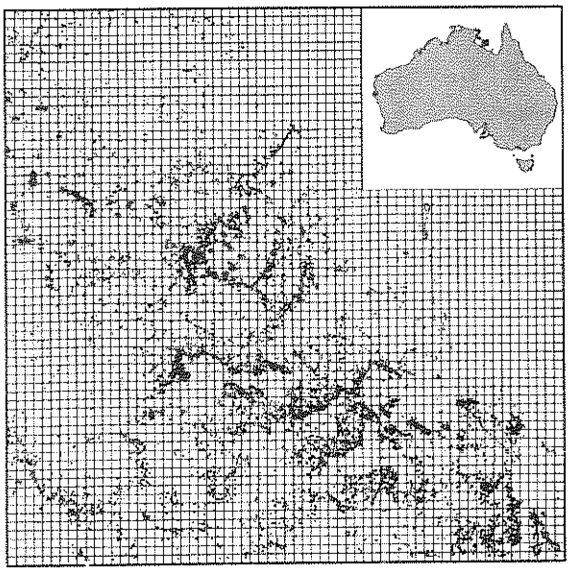

Figure 2: The study area divided into 4096 boxes with points showing the reef and rough bottom, in the Northern Prawn Fishery, Australia.

6. Discussion

We have introduced the QL method for dependent data and applied it in the context of fractal-dimension estimation. The study presented here is only a prelimary one. Further exploration of the QL approach while accounting for the correlations between observed numbers in each sub-box is of great interest. We intend to establish a fractal dimension for each stock area in the Northern Prawn Fishery, and investigate the possible relationship with the annual catches in these areas.

References

Cutler, C., A Review of the Theory and Estimation of Fractal Dimension, Dimension

Estimation and Models (pp. 1-107)., (Ed. Tong, H.), World Scientific, 1993.

Falconer, K., *Fractal Geometry - Mathematical Foundations and Applications*, John Wiley and Sons, 1990.

Godambe, V.P., and C.C. Heyde, Quasi-likelihood and optimal estimation. *Int. Statist. Rev.* 55, 231-44, 1987.

Hall, P., and C.C. Heyde, *Martingale Limit Theory and its Application.* New York: Academic Press, 1980.

Hall, P., and A. Wood, On the Performance of Box-counting Estimators of Fractal Dimension, *Biometrika* 80, 246-252, 1993.

Heyde, C.C., On combining quasi-likelihood estimating functions. *Stoch. Processes Applic.* 25, 281-287, 1987.

Liang, K.-Y., and S.L. Zeger, Longitudinal data analysis using generalized linear models. *Biometrika* 73, 13-22, 1986.

Lin, Y.-X., and C.C. Heyde, Optimal estimating functions and Wedderburn's quasi-likelihood. Comm. in Stat. : Theory and Methods, Vol 22, 303-311, 1997.

Lin, Y.-X., and C.C. Heyde, On the space of estimating functions, J. Stat. Plan. Inf. (to appear), 1997.

McCullagh, P., and J.A. Nelder, *Generalized Linear Models.* London: Chapman and Hall. Second edition, 1989.

Roberts, A.J. & A. Cronin, Unbiased estimation of multi-fractal dimensions of finite data sets. *Physica A* 233, 867-878, 1996.

Sørensen, M., On quasi-likelihood for semimartingales. *Stochastic Process. Appl.* 35, 331-346, 1990.

Taylor, C, and Taylor, S, (1991). Estimating the Dimension of a Fractal, J. R. Statist. Soc. B 53, 353-364, 1991.

Wang, Y.-G., A quasi-likelihood approach for ordered categorical data with overdispersion. *Biometrics* 52, 1252-1258, 1996.

Wedderburn, R.W.M., Quasi-likelihood functions, generalized linear models, and the Gauss-Newton method. *Biometrika* 61, 439-447, 1974.