A THEORETICAL APPROACH TO FRACTAL DIMENSION OF RANDOM PROCESSES

Mehrdad Salmasi, M. Modarres-Hashemi

Department of Electrical and Computer Engineering, Isfahan University of Technology
Isfahan, Iran

Abstract: Fractal geometry is a new field in mathematics. Recently, it has found its role in signal processing. Fractal dimension is one of the features of fractal sets, which is widely used in different aspects of signal processing. There are just practical ways to estimate the fractal dimension of an arbitrary random process. In these methods, the fractal dimension of sample functions are evaluated and based on these calculations the mean of the fractal dimension is determined for the random process. In this paper a theoretical method is proposed for evaluating the fractal dimension of a group of random processes. It is shown that without using sample functions, the mean of fractal dimension can be estimated precisely.

Key-Words: Fractal geometry, Fractal dimension, Random processes, Theoretical method, Histogram, Mean, Variance.

1 Introduction

Fractal geometry is a field for investigating rough and irregular sets. So irregular signals such as clutter, noise, chaos,… can be studied using this new field. There is not any exact definition for a fractal and they are described by their special characteristics. The most important feature of a fractal set is its fractal dimension. This feature is used in different applications and so different methods have been introduced to calculate it. There are efficient algorithms for calculating the fractal dimension of deterministic signals but random processes lack such algorithms. In this case, the usual method is based on generation of sample functions and evaluation of their fractal dimension. So the mean of these dimensions is considered as the fractal dimension of random process. Drawbacks of this method have encouraged us to innovate a new theoretical method.

In section 2 the basic definitions of the fractal geometry is presented. The theoretical method is studied in section 3 and some simulation results are put in section 4. Finally in section 5 based on the simulation results some conclusions are proposed.

2 Fractal geometry

Fractal has come from the word ‘fractus’, which means broken. Before 19th century there was no attitude to study rough and irregular sets. But step-by-step the needs for exploring these sets came out. At 1980 professor Benoit Mandelbrot introduced the bases of fractal geometry.

There is no exact definition for fractal sets. A set F is a fractal if it has some of the following characteristics [1]:

1-F has a fine structure.
2-F is too irregular to be studied by classic geometry.
3- Usually F has a self-similar structure.
4-Usually F can be described with an easy statement.
5-Fractal dimension of F is greater than its topological dimension.

Some authors have proposed the fifth statement as a definition for fractal but there are some cases that contradict this definition. So it is seen that there is not a common definition for a fractal set. Fractal dimension is a measure of how rough a set is [2]. When the set is more irregular, its fractal dimension is higher. There are different definitions for fractal dimension. The best theoretical definition is Hausdorf dimension. Since it is too difficult to compute this dimension, new definitions of fractal dimension have been introduced. Similarity dimension, divider dimension, box-counting dimension are three other definitions. Among them the box-counting dimension is the best practical one and It can be easily evaluated. Fractal dimension is defined for every set F as following:

$$\dim F_B = \lim_{\delta \to 0} \frac{\log(N_\delta(F))}{-\log(\delta)} \qquad (1)$$

where $N_\delta(F)$ is the number of δ-mesh cubes that intersect with F.

When $F \subset R^2$, There is an easy algorithm for calculating the box dimension [3].

Step1: normalize the signal into unit square.

Step2: choose a sequence of δ_k ’s such that they tend to zero.

Step3: for each δ_k grid the unit square into squares with length δ_k.

Step4: count the number of grid squares $N_{\delta_k}(F)$, which intersect with normalized signal.

Step5: Plot log($N_{\delta_k}(F)$) vs. –log(δ) and find the slope of regression line.

3 Fractal Dimension of Random Processes

Random processes play an important role in signal processing and so knowing their fractal properties is essential for a fractal analysis [4,5]. The first question is that if random processes possess fractal characteristics or not. It can be investigated by plotting log($N_{\delta_k}(F)$) vs. –log(δ) because in real fractals this plot is a straight line. If the random processes are fractals, the second question is that what their fractal dimension is.

As the algorithm presented in previous section can be applied for all signals, it is not necessary to answer the first question. This algorithm can be utilized even for non-fractal signals. The fractal dimension evaluated in this way, gives a measure of roughness in signals. So it is applicable for non-fractal signals.

3.1 Practical method [6]

The easy way of calculating the fractal dimension of a random process is as follows:

1-Generate a large number of sample functions.

2-calculate the fractal dimension of each sample function.

3-find the average of them.

This method has been widely used but has some disadvantages. For example a large number of sample functions are needed for calculating the fractal dimension and we can't find a precise Pdf for the fractal dimension. The distribution of fractal dimension is a key parameter in some applications and so this is an important drawback of the practical method.

3.2 Theoretical method [7]

In this part a theoretical method is proposed. The process is considered as a discrete-time process, which is specified in N equally spaced time points [8]. In each time i, it is a random variable a_i. The random variables are independent and have identical Pdf in the range [0,1]. (e.g., uniform Pdf ,a truncated normal to [0,1],…).

The fractal dimension algorithm is exerted on this kind of processes.

As the Pdf of random variables restricts to [0,1], normalization is not needed.

Suppose a fixed δ. Referring to this δ, a grid is considered for the unit square, so the unit square is divided into $B=\left\lfloor \frac{1}{\delta} \right\rfloor$ rows and $B=\left\lfloor \frac{1}{\delta} \right\rfloor$ columns. N_δ is a random variable and its statistics is needed for evaluating the fractal dimension. A random variable S_i is defined as follows:

$S_i \equiv$ The number of squares that intersect with the signal in i'th column, for $i \in \{1,2,...,B\}$.

The S_i 's are iid random variables [7], so for each i we define: $\mu \equiv E(S_i)$ and $\sigma^2 \equiv VAR(S_i)$.

It is clear that N_δ is the sum of S_i 's, i.e.

$$N_\delta = S_1 + S_2 + ... + S_B \tag{2}$$

As the S_i 's are iid and (2), the central limit theorem indicates that the distribution of $\frac{S_1+S_2+...+S_B-B\mu}{\sigma\sqrt{B}}$ converges to standard normal Pdf as $B \to \infty$. As a consequence, for B large enough N_δ has a gaussian Pdf with:

$$E(N_\delta) = B\mu \tag{3}$$

$$VAR(N_\delta) = \sigma^2 B \tag{4}$$

So knowing μ and σ, the Pdf of N_δ is completely determined. As S_i 's are iid, it is enough to find $E(S_1)$ and $VAR(S_1)$.

The random variable M_i is defined as:

$$M_i \equiv \begin{cases} 0: \text{if in the first column, the signal doesn't intersect with the square in the i'th row.} \\ 1: \text{if in the first column, the signal intersect with the square in the i'th row.} \end{cases}$$

It is obvious that:

$$S_1 = M_1 + M_2 + ... + M_B \tag{5}$$

The number of points in each column is:

$$N_C \equiv \left\lfloor \frac{N}{\delta} \right\rfloor \tag{6}$$

Suppose $p_i, i=1,2,...,B$ is the probability that a_1 is situated in the i'th row. Then [7]:

$$P(M_i = 1) = 1-(1-p_i)^{N_C} \tag{7}$$

and the definition shows that:

$$E(M_i) = P(M_i = 1) \tag{8}$$

Equation (5) yields:

$$E(S_1) = \sum_{i=1}^{B} E(M_i) \tag{9}$$

Incorporating (7), (8) in (9) results in

$$\mu \equiv E(S_i) = \sum_{i=1}^{B} [1-(1-p_i)^{N_C}] \tag{10}$$

Using (3) and (10)

$$E(N_\delta) = B\sum_{i=1}^{B} [1-(1-p_i)^{N_C}] \tag{11}$$

Equation (11) is a theoretical approach for calculating the mean of N_δ.

The next step is the evaluation of the $\sigma^2 \equiv VAR(S_i)$.

As

$$VAR(S_i) = E(S_1^2) - (E(S_1))^2 \qquad (12)$$

It is enough to evaluate $E(S_1^{\,2})$.

$$E(S_1^{\,2}) = \sum_{i=1}^{B}\sum_{j=1}^{B} E(M_i M_j) \qquad (13)$$

And we have [7]

$$E(M_i M_j) = \begin{cases} 1-(1-p_i)^{N_C} - (1-p_j)^{N_C} & \text{if } i \neq j \\ \quad - (1-p_i-p_j)^{N_C} & \\ 1-(1-p_i)^{N_C} & \text{if } i = j \end{cases} \qquad (14)$$

Incorporating (10), (13), (14) in (12) results in the desired formula for $\sigma^2 \equiv VAR(S_i)$. So having a fixed δ and calculating p_i's for the specified Pdf, we can find $E(N_\delta)$ and $VAR(N_\delta)$.

Now we change our discussion to regression line [9]. Regression line is the best linear estimation of data points in a least square error case.

Given points $(x_1, y_1), (x_2, y_2), \ldots, (x_s, y_s)$ in the plane, the regression line is determined by $y = bx + c$, where

$$b = \frac{S_{xy}}{S_x^{\,2}} \qquad (15)$$

In (15)

$$S_{xy} = \sum_{i=1}^{s} (x_i - \bar{x})(y_i - \bar{y}) \qquad (16)$$

$$S_x^{\,2} = \sum_{i=1}^{s} (x_i - \bar{x})^2 \qquad (17)$$

So given some points, using (15), (16), (17) specify the slope of the regression line.

In the fractal dimension algorithm we have to choose a sequence of δ_k's such that they tend to zero. Suppose $\delta_1, \delta_2, \ldots, \delta_s$ as the sequence. For each δ_k we calculate $E(N_{\delta_k})$ and $VAR(N_{\delta_k})$. We know that N_{δ_k} has a gaussian Pdf approximately.

In this case the points are $(\log\delta_1, \log N_{\delta_1}), (\log\delta_1, \log N_{\delta_2}), \ldots, (\log\delta_1, \log N_{\delta_s})$ and the slope of the regression line is desired. If the slope of the regression line is denoted by b, then the fractal dimension of the random process (D) will be:

$$D = -b \qquad (18)$$

Replacing x_i with $\log\delta_i$ and y_i with $\log(N_{\delta_i})$ and some simplifications, the following result is derived [4]:

$$D = \sum_{i=1}^{s} c_i . \log(N_{\delta_i}) \qquad (19)$$

Where

$$c_i = \frac{\log(\delta_i) - \overline{\log(\delta_i)}}{S^2_{\log(\delta_i)}} \qquad (20)$$

As δ_i's are deterministic and known, c_i's are deterministic and known.

We have shown that N_{δ_i}'s have gaussian Pdf with evaluated parameters. This ensures us that the statistics of fractal dimension are derivable. But there is a paradox here. As N_{δ_i}'s have gaussian Pdf, they can get negative and this results in ambiguity of $\log(N_{\delta_i})$. The key to the answer is in the word “approximately gaussian”. The central limit theorem was used in approximate case. In practice N_{δ_i}'s are between B and B^2 and never get negative values.

The most important statistics of the fractal dimension is its mean.

$$E(D) = \sum_{i=1}^{s} c_i . E(\log(N_{\delta_i})) \qquad (21)$$

So knowing the Pdf of N_{δ_i}, we have:

$$E(\log(N_{\delta_i})) = \int_{x=l_1}^{l_2} (\log x) . f_{N_{\delta_i}}(x) dx \qquad (22)$$

Where

$$f_{N_{\delta_i}}(x) \approx \frac{1}{\sqrt{2\pi}\sigma_{N_{\delta_i}}} \exp\left(\frac{-(x-\mu_{N_{\delta_i}})^2}{2\sigma^2_{N\delta_i}}\right) \qquad (23)$$

In (22) the parameters l_1 and l_2 can be chosen as, $l_1 = B,\ l_2 = B^2$.

It is worth noting that equations (21), (22), (23) are basic equations for calculating the mean of the fractal dimension.

The most important characteristic of a random variable is its Pdf and referring to (19) and (23) we can find the Pdf of the fractal dimension. But it is not an easy subject and is put away.

Therefore, the main result of this section was a theoretical way to evaluate the mean of the fractal dimension for a random process.

4 performance analysis of the theoretical method

In this section, the results of the theoretical method are compared with simulation results. The random process consists of $N = 10000$ independent random variables with uniform distribution in [0,1]. Table 1 shows the mean and variance of N_δ (for some B's) extracted from simulation and theoretical method.

The simulation proves the accuracy of theoretical formulas for N_δ.

Table 1: A comparison between simulation and theory for N_δ in different B's.

B	Mean theory	Mean Sim.	Var. Theory	Var. Sim.
25	625	625	5.06-5	1e-4
50	2456	2446.6	40.167	46.8
75	4685.6	4616.8	503.23	516.76
100	6339.7	6197.3	974	960.3

The histogram of N_δ for $B = 50$ has been sketched in figure 1. The distribution is similar to gaussian Pdf, which is one of the results of previous section.

In table 2 the fractal dimension of the random process has been evaluated with both methods. Different ranges have been chosen for B.

The theoretical approach has accurate values comparing with simulation results. The sample functions are not needed for theoretical method and this is the great advantage of this

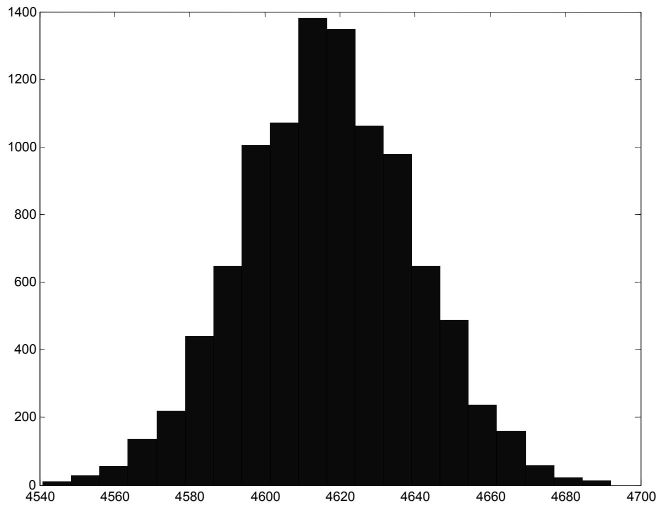

Figure 1:The histogram of N_{75} with simulation

Table 2: Comparison of the Fractal dimension, evaluated with theoretical and simulation methods.

B range	Mean of FD theory	Mean of FD Sim	Error in mean FD
20:5:100	1.7293	1.7142	0.8%
30:10:120	1.4959	1.4685	1.8%
20:10:80	1.8513	1.8379	0.7%

5 Conclusion

In spite of the approximations, the new theoretical method accurately obeys the real values of the fractal dimension processes. The privilege of this method is in its independence. No sample function is needed for calculating the mean of the fractal dimension. Remembering thousands of sample functions, which should be used to obtain a practical fractal dimension of the random process, the advantages of this method are brightened.

There are some cases, in which the Pdf of the fractal dimension is needed. In these situations, theoretical methods can give a precise formula for the Pdf, where practical methods just approximate the Pdf.

Here some advantages of this new method were introduced, but there is a long way for improving this method.

References:

[1] K. Falconer: "*Fractal Geometry*", John Wiley & Sons, 1990.

[2] M. F. Barnsley: "*Fractals Everywhere*", Academic Press, 1993.

[3] S. Haykin et al.: "Fractal characteristics of sea scattered signals and detection of sea surface targets", *IEE Proc.*, Vol.140, No.4, pp243-250, August 1993.

[4] J. Tang, Z. Zhu: "High resolution radar detection based on fractal dimension", *IEEE Proc.*, pp 369-372, 1996.

[5] F. Berizzi et al.: "Fractal theory of sea scattering", *CIE proc. on Radar*, pp 661-665,1996.

[6] W. Ying et al.: "Fractal Dimension of Random Sequences", *3rd International conf. on signal processing*, vol.1 , pp 257-260, 1996.

[7] M. Salmasi: "*Applications of fractal geometry in radar detection*", M.S. Theses, Isfahan University of Technology, In preparation.

[8] A. Papoulis: "*Probability, Random Variables and Stochastic Processes*", McGrow-Hill,1991

[9] J. K. Bhattacharyya, R. A. Johnson: '*Statistical Concepts and Methods*", John Wiley & Sons, 1977.