

Several Methods for Calculating and Estimating the Hausdorff Dimension

Ziang Cao *

Information Science and Technology College of Hunan Agricultural University, Changsha, 410128, China

* Corresponding author Email: caoziang202301@163.com

Abstract: Among various fractal dimensions, the Hausdorff dimension is the most widely used and fundamental one. However, in many cases, calculating or estimating its value can be challenging. This article systematically introduces several methods for calculating the Hausdorff dimension, starting with its definition and properties. In method one, the dimension is estimated using upper and lower bounds. For the upper bound, it usually requires finding a specific covering, while for the lower bound, the mass distribution principle is commonly employed for estimation; In method two, the clever use of potential theory is applied to estimate the Hausdorff dimension; In method three, for some high-dimensional cases, the dimension can be reduced using the projection theorem, and then the original dimension is estimated by computing the dimension of the lower-dimensional cases; In method four, the dimension estimation of the Cartesian product of two sets is considered, and several theorems are used to provide upper and lower bounds for the dimension; In method five, the focus is on the dimension calculation of self-similar sets. For such sets, under the condition of open sets, heuristic methods can be used for calculation; In method six, analogous to self-similar sets, the estimation of the dimension for self-affine sets is summarized.

Keywords: Hausdorff Dimension; Mass Distribution Principle; Potential Theory; Self- similar Set; Self-affine Set.

1. Introduction

Fractal theory is a highly popular and dynamic new theory and discipline in the contemporary world. The concept of fractals was first proposed by mathematician Benoît B. Mandelbrot. In 1967, he published the renowned paper titled "How Long Is the Coast of Britain?" in the prestigious journal "Science" in the United States. Fractal dimension, as a quantitative representation and fundamental parameter of fractals, is an important concept in fractal theory. Among numerous fractal dimensions, the most widely used and fundamental one is the Hausdorff dimension. The development of the Hausdorff dimension can be traced back to the early 20th century when German mathematician Felix Hausdorff first defined the concept of the Hausdorff dimension. It implies that the spatial dimension can vary continuously and can be either an integer or a fraction.

Fractal geometry emerged in the latter half of the 20th century, investigating irregular structures exhibiting self-similarity or self-affinity at various scales. The Hausdorff dimension became a vital tool in fractal geometry, laying the foundation for its development. With the advancement of fractal theory, the Hausdorff dimension has been extensively applied in numerous fields such as natural sciences, biology, geography, astronomy, finance, and others. These applications demonstrate the significance of fractal dimension in describing complex systems and natural phenomena.

Let U be any non-empty subset of the n -dimensional Euclidean space R^n , define

$$|U| = \sup \{|x - y| : x, y \in U\} \quad (1)$$

as the diameter of a set. If F can be covered by a countable or finite collection of sets $\{U_i\}$, each having a diameter not exceeding δ , for every i , then $\{U_i\}$ is called an δ -covering

of F .

Definition 1.1: Let F be any subset of R^n , and let S be a non-negative number, for any $\delta > 0$, define

$$\mathcal{H}_\delta^S(F) = \inf \left\{ \sum_{i=1}^{\infty} |U_i|^S : \{U_i\} \text{ is a } \delta\text{-cover of } F \right\} \quad (2)$$

Noting that as a function of δ , $\mathcal{H}_\delta^S(F)$ is non-decreasing, when $\delta \rightarrow 0$ is given, we define the S -dimensional Hausdorff measure of the set F as follows:

$$\mathcal{H}^S(F) = \lim_{\delta \rightarrow 0} \mathcal{H}_\delta^S(F) \quad (3)$$

In traditional Euclidean space, we are accustomed to considering points as zero-dimensional, line segments as one-dimensional, planes as two-dimensional, and three-dimensional space as three-dimensional. However, for some geometric objects with peculiar shapes, this integer-dimensional description is not sufficient, and the following definition of the Hausdorff dimension allows for fractional values.

Definition 1.2: The Hausdorff dimension of the set F is defined as follows:

$$\dim_H F = \inf \{s : \mathcal{H}^s(F) = 0\} = \sup \{s : \mathcal{H}^s(F) = \infty\} \quad (4)$$

Some geometric shapes of fractal structures, such as the Koch curve and the Cantor set, can have non-integer Hausdorff dimensions. Specifically, the Hausdorff dimension of the Cantor set is denoted by $\log 2 / \log 3$. If we attempt to cover the Cantor set with sets of dimensions 0, i.e., with points, the Hausdorff measure would be infinite. On the other hand, if we try to cover the Cantor set with sets of dimensions 1, i.e., with lines, the Hausdorff measure would be 0. Hence, it is evident that the Hausdorff dimension of the Cantor set lies between 0 and 1.

It is not difficult to see that calculating the Hausdorff dimension of a set is quite challenging. In this paper, we first introduce the definition and relevant properties of the

Hausdorff dimension. Then, based on the definition and properties, we provide several methods for computing or estimating the dimension.

2. Introduction of Relevant Properties

Based on the above definition, we now proceed to introduce some related properties. Let $F \subset R^n$, and f be a mapping from F to R^m . If there exists a positive constant $c > 0, \alpha > 0$ such that for $\forall x, y \in F$, there exists

$$|f(x) - f(y)| \leq c|x - y|^\alpha \quad (5)$$

then f is said to satisfy the α -order Hölder condition. If $\alpha = 1$ holds, then f is called a Lipschitz mapping, if there exists a constant $c_1, c_2 > 0$, such that for $\forall x, y \in F$, there exists

$$c_1|x - y| \leq |f(x) - f(y)| \leq c_2|x - y| \quad (6)$$

Then f is called a double Lipschitz mapping.

Property 2.1: If $F \subset R^n$ and $f : F \rightarrow R^m$ satisfy the α -order Hölder condition, then

$$\mathcal{H}^{s/\alpha}(f(F)) \leq c^{s/\alpha} \mathcal{H}^s(F) \quad (7)$$

Property 2.2: If f is a Lipschitz mapping, then

$$\mathcal{H}^s(f(F)) \leq c^s \mathcal{H}^s(F) \quad (8)$$

If f is a double Lipschitz mapping, then

$$c_1^s \mathcal{H}^s(F) \leq \mathcal{H}^s(f(F)) \leq c_2^s \mathcal{H}^s(F) \quad (9)$$

Property 2.3 (Homogeneity): If $F \subset R^n, \lambda > 0$, then

$$\mathcal{H}^s(\lambda F) = \lambda^s \mathcal{H}^s(F) \quad (10)$$

Below, we introduce some properties of the Hausdorff dimension.

Property 2.4 (Monotonicity): If $E, F \subset R^n$ and $E \subset F$, then

$$\dim_H E \leq \dim_H F \quad (11)$$

Property 2.5 (Stability): If $F_1, F_2, \dots$ is a sequence of sets, then

$$\dim_H \bigcup_{i=1}^{\infty} F_i = \sup_{1 \leq i \leq \infty} \{\dim_H F_i\} \quad (12)$$

Property 2.6: If $F \subset R^n$ and $f : F \rightarrow R^m$ satisfy the α -order Hölder condition, then

$$\dim_H f(F) \leq \frac{1}{\alpha} \dim_H F \quad (13)$$

Property 2.7: If f is a Lipschitz mapping, then

$$\dim_H f(F) \leq \dim_H F \quad (14)$$

If f is a double Lipschitz mapping, then

$$\dim_H f(F) = \dim_H F \quad (15)$$

For detailed proofs and other properties regarding the Hausdorff measure and dimension, please refer to the reference [1].

3. Several Methods for Calculating or Estimating Dimension

In this section, we will introduce these various methods one by one and provide examples to aid in understanding. For a detailed explanation and derivation of the principles behind each method, please refer to the references [1], [2], [3], [4], [5], [6] and [7].

3.1. The Upper and Lower Bound Estimation Method

First and foremost is the most commonly used method, which determines the dimension by estimating upper and lower bounds.

(1) Estimation of the upper bound

Estimating the upper bound is often simpler than estimating the lower bound. We usually only need to find a particular covering such that

$$\sum |U_i|^s < \infty \quad (16)$$

Then

$$\dim_H F \leq s \quad (17)$$

Moreover, such coverings often have the same radius and size, making the summation calculation convenient.

(2) Estimation of the lower bound

Theorem 3.1.1 (Mass Distribution Principle): Let μ be a mass distribution on F , and for some s , there exist $c > 0$ and $\delta > 0$ such that

$$\mu(U) \leq c|U|^s \quad (18)$$

holds for all sets U satisfying $|U| \leq \delta$. Then

$$\mathcal{H}^s(F) \geq \frac{\mu(F)}{c}, \text{ and} \quad s \leq \dim_H F \quad (19)$$

This is the most commonly used method for estimating the lower bound, and the mass distribution often only needs to be taken as a natural uniform distribution to satisfy the conditions, thus allowing the estimation of the lower bound. Below, we will provide an example to illustrate this method.

Example 3.1.2: Let $F_1 = F \times [0, 1] \subset R^2$ be the product of the Cantor set F with the unit interval. Then

$$\dim_H F_1 = 1 + \log 2 / \log 3 = s$$

Proof: For the upper bound, F_1 can be covered by $2^k \cdot 3^k$ squares with side length 3^{-k} , so

$$\mathcal{H}_{3^{-k}\sqrt{2}}^s(F_1) \leq 2^k 3^k (3^{-k}\sqrt{2})^s = 2^{\frac{s}{2}}$$

Let $k \rightarrow \infty$, then $\mathcal{H}^s(F_1) \leq 2^{\frac{s}{2}}$ and $\dim_H F_1 \leq s$

For the lower bound, we construct μ as the natural mass distribution on F_1

$$\mu(U) = h \cdot 2^{-k}$$

If $3^{-(k+1)} \leq |U| \leq 3^{-k}$, then U can be located at most on

one basic interval of length 3^{-k} , so

$$\mu(U) \leq |U|2^{-k} \leq |U|3^{-k \log 2 / \log 3} \leq |U|(3|U|)^{\log 2 / \log 3} \leq 3^{\log 2 / \log 3} |U|^s$$

Then, $\mathcal{H}^s(F_1) > 0$ and $s \leq \dim_H F_1$.

In the above example, for the upper bound of the dimension, we first found a covering of the target set F_1 , consisting of $2^k \cdot 3^k$ squares with side length 3^{-k} . Each set in the covering is identical, making it convenient to calculate $\sum |U_i|^s$ and determine whether it is less than infinity. For the lower bound of the dimension, we constructed a mass distribution on the target set F_1 to satisfy (3.1.1), thus estimating the lower bound of the dimension.

3.2. Potential Theory Method

For $s \geq 0$, the s -potential induced by a mass distribution μ on a point x in R^n is defined as follows:

$$\phi_s(x) = \int \frac{d\mu(y)}{|x-y|^s} \quad (20)$$

The s -energy of μ is:

$$I_s(u) = \int \phi_s(s) d\mu(x) = \iint \frac{d\mu(x) d\mu(y)}{|x-y|^s} \quad (21)$$

Theorem 3.2.1: Let F be a subset of R^n

- (1) If the mass distribution μ on F satisfies $I_s(\mu) < \infty$, then $\mathcal{H}^s(F) = \infty$ holds and $\dim_H F \geq s$.
- (2) If F is a Borel set satisfying $\mathcal{H}^s(F) > 0$, then there exists a mass distribution μ on F such that, for any $t < s$, $I_t(\mu) < \infty$ holds.

Below, we will further understand the Potential Theory Method through examples.

Example 3.2.2: Let μ be the natural mass distribution on the Cantor ternary set. Estimate the s -energy of μ for $s < \frac{\log 2}{\log 3}$, and prove $\dim_H F \geq \frac{\log 2}{\log 3}$.

Proof: Let $\phi_s(x) = \int \frac{d\mu(y)}{|x-y|^s}$, $I_s(\mu) = \int \phi_s(x) d\mu(x)$. We

let $r_k = 3^{-k}$, $F_k = \{y : r_{k+1} \leq |x-y| \leq r_k\}$, and, at the same time, it is easy to see that $\mu(F_k) \leq 2^{-k}$, hence

$$\phi_s(x) = \int \frac{d\mu(y)}{|x-y|^s} = \sum \int_{F_k} \frac{d\mu(y)}{|x-y|^s} \leq \sum r_{k+1}^{-s} \cdot 2^{-k} = \sum_{k=0}^{\infty} 3^s \cdot \left(\frac{3^s}{2}\right)^k = \frac{3^s}{1 - \frac{3^s}{2}} \leq \infty$$

At the same time,

$$I_s(\mu) = \int \phi_s(x) d\mu(x) \leq \frac{3^s}{1 - \frac{3^s}{2}} \int d\mu(x) \leq \frac{3^s}{1 - \frac{3^s}{2}} \leq \infty$$

So $\dim_H F \geq s \geq \frac{\log 2}{\log 3}$.

3.3. Projection Theorem for Dimension Estimation

Theorem 3.3.1: Let $F \subset R^n$ be a Borel set:

(1) If $\dim_H F \leq k$ holds, then for almost all $\Pi \in G_{n,k}$, there exists $\dim_H(\text{proj}_{\Pi} F) = \dim_H F$.

(2) If $\dim_H F > k$ holds, then for almost all $\Pi \in G_{n,k}$, $\text{proj}_{\Pi} F$ has a positive k -dimensional measure, and its dimension is k .

Where $G_{n,k}$ is a k -dimensional subspace or " k -dimensional plane" passing through the origin of R^n , and $\text{proj}_{\Pi} F$ represents the orthogonal projection onto the k -dimensional plane Π .

This theorem has surprising implications for estimating the dimension of sets. For example, let F be a subset of R^3 , and take the plane Π as a 2-dimensional plane. If the dimension of the orthogonal projection of set F onto plane Π is less than 2, then the original set's dimension must be the same as the Hausdorff dimension of the projected set. If the dimension of the orthogonal projection on plane Π is equal to 2, then the original set's dimension can be estimated by the same dimension as the projection.

This theory also has practical applications. When estimating the dimension of an object in a three-dimensional space, direct estimation can be challenging. Usually, we can randomly take a photograph of the object from a certain direction to form a picture. Then, by calculating the dimension of the picture, we can estimate the dimension of the original object. This significantly reduces the difficulty of estimation.

3.4. Fractal Multiplication for Dimension Estimation

For complex sets, they often can be represented as the Cartesian product of two sets. For example, in Example 3.1.2, the set $F_1 = F \times [0,1] \subset R^2$ is the Cartesian product of the Cantor ternary set F and the unit interval. It is not difficult to observe the dimension of the set F_1 , denoted as $\dim_H(F_1) = \dim_H F + \dim_H[0,1] = \dim_H F + 1$. Therefore, studying the relationship between the dimensions of two sets F and E and the dimension of their Cartesian product set $E \times F$ can facilitate the calculation and estimation of the dimensions of some complex sets. This section will introduce some relevant conclusions and how to use them to estimate the dimension of sets.

Theorem 3.4.1: Let $E \subset R^n$ and $F \subset R^m$ be Borel sets, then

$$\dim_H(E \times F) \geq \dim_H(E) + \dim_H(F) \quad (22)$$

It is not difficult to see from the theorem that the dimension of the Cartesian product set $E \times F$ is bounded below by the sum of the dimensions of sets F and E . This provides strong assistance in estimating the dimension of the product set. As for the upper bound, there are similar conclusions. However, before that, we need to briefly introduce the Box-counting dimension.

Let F be an arbitrary non-empty bounded subset of R^n , and $N_\delta(F)$ be the smallest number of sets with maximum diameter δ covering F . Then, the upper and lower Box-counting dimensions of F are respectively defined as

$$\overline{\dim}_B F = \lim_{\delta \rightarrow 0} \frac{\log N_\delta(F)}{-\log \delta} \quad (23)$$

and

$$\underline{\dim}_B F = \lim_{\delta \rightarrow 0} \frac{\log N_\delta(F)}{-\log \delta} \quad (24)$$

If these two values are equal, we denote the common value as

$$\dim_B F = \lim_{\delta \rightarrow 0} \frac{\log N_\delta(F)}{-\log \delta} \quad (25)$$

After introducing the definition of the Box-counting dimension, we now proceed to provide the upper bound estimation for the dimension of the Cartesian product set, which is related to the upper Box-counting dimension of the set.

Theorem 3.4.2: Let $E \subset R^n$ and $F \subset R^m$ be any sets, then

$$\dim_H(E \times F) \leq \dim_H(E) + \overline{\dim}_B(F) \quad (26)$$

Theorem 3.4.3: If $\dim_B F = \dim_H F$ holds, then

$$\dim_H(E \times F) = \dim_H E + \dim_H F \quad (27)$$

The above two theorems provide an upper bound estimation for the dimension of the Cartesian product set and the conditions for equality. Let's continue discussing Example 3.1.2. It is not difficult to observe that the Cantor ternary set F satisfies $\overline{\dim}_B F = \dim_H F$. According to Theorem 3.4.3, the dimension of the Cartesian product set is equal to the sum of the dimensions of the two sets. This allows for a more convenient calculation of the dimension of set F_1 .

3.5. Dimension Calculation Method for Self-similar Sets

For some sets, they exhibit self-similarity, and we call them self-similar sets. The Cantor ternary set is an example of a self-similar set, where its parts resemble the whole. Sets like these can often be calculated for their dimension using the homogeneity of the Hausdorff measure.

Theorem 3.5.1: The family of mappings $S_1, \dots, S_k : R^n \rightarrow R$ is in a self-similar case, that is:

$$|S_i(x) - S_i(y)| = c_i |x - y| \quad (x, y \in R^n) \quad (28)$$

where $0 < c_i < 1$ (c_i is called the contraction ratio). A heuristic calculation method is as follows:

$$\mathcal{H}^s(F) = \sum_{i=1}^m \mathcal{H}^s(S_i(F)) = \sum_{i=1}^m c_i^s \mathcal{H}^s(F) \quad (29)$$

to obtain $\sum_{i=1}^m c_i^s = 1, s = \dim_H F$.

To ensure that the various parts $S_i(F)$ of F do not overlap too much, S_i needs to satisfy the following open set condition:

$$\bigcup_{i=1}^m S_i(V) \subset \{V : S_i(V) \cap S_j(V) = \emptyset, i \neq j\} \quad (30)$$

For the Cantor ternary set, $S_1 = \frac{1}{3}x, S_2 = \frac{1}{3}x + \frac{2}{3}$, it obviously satisfies the open set condition. By Theorem 3.5.1, the s-dimension is the solution of the following equation

$$1 = \left(\frac{1}{3}\right)^s + \left(\frac{1}{3}\right)^s$$

It is not difficult to deduce that $s = \log 2 / \log 3$. The dimensions of self-similar sets like the Koch curve, "Cantor dust," and others can be calculated using this method.

3.6. Dimension Estimation Method for Self-affine Sets

Theorem 3.6.1: Let $S_1, \dots, S_m$ be a contraction mapping on $D \subset R^n$ such that

$$|S_i(x) - S_i(y)| \leq c_i |x - y| \quad (x, y \in D) \quad (31)$$

for each $i, c_i < 1$. Then, there exists a unique non-empty compact subset F of S_i that is invariant under S_i , satisfying:

$$F = \bigcup_{i=1}^m S_i(F) \quad (32)$$

Moreover, if we define the transformation S on the non-empty compact set class $\mathcal{O}$, such that

$$S(E) = \bigcup_{i=1}^m S_i(E) \quad (33)$$

and if we denote S^k as the k th iteration of S , defined by $S^0(E) = E$ and $S^k(E) = S(S^{k-1}(E))$ ($k \geq 1$), where E is an arbitrary set on $\mathcal{O}$ and satisfies G for any i , $S_i(E) \subset E$, then

$$F = \bigcap_{k=1}^{\infty} S^k(E) \quad (34)$$

Theorem 3.6.2: Let $S_1, \dots, S_m$ be a contraction mapping on a closed subset D of R^n such that

$$b_i |x - y| \leq |S_i(x) - S_i(y)| \quad (x, y \in D)$$

for each $i, 0 < b_i < 1$. Assuming that F is invariant under S_i :

$$F = \bigcup_{i=1}^m S_i(F)$$

and satisfies the open set condition, then $\dim_H F \geq s$,

where $\sum_{i=1}^m b_i^s = 1$.

Theorem 3.6.3: Let $S_1, \dots, S_m$ be a contraction mapping on a closed subset D of R^n , satisfying

$$|S_i(x) - S_i(y)| \leq c_i |x - y| \quad (x, y \in D) \quad (35)$$

for each $i, c_i < 1$. Then, $\dim_H F \leq s$ holds, where

$$\sum_{i=1}^m c_i^s = 1.$$

Compared to self-similar sets, self-affine sets do not strictly have equal contraction mappings, and the contraction ratios in different directions may vary. For such sets, we can use Theorem 3.6.2 and Theorem 3.6.3 to estimate their dimensions. Below, we will use examples to understand this method.

Example 3.6.4: Let $D = \left[\frac{1}{2}(1 + \sqrt{3}), (1 + \sqrt{3}) \right]$, and

let S_1 and S_2 be mappings given by $S_1(x) = 1 + 1/x$ and $S_2(x) = 2 + 1/x$ for $D \rightarrow D$, respectively. Then, $0.44 < \dim_H F < 0.66$ holds, where F is the invariant set of S_1 and S_2 .

Proof: From the mean value theorem, if $x \neq y \in D$ holds, there exists a constant $c_i \in D$ such that

$$\frac{S_i(x) - S_i(y)}{x - y} = S'_i(c_i) \quad i = 1, 2$$

which implies

$$\inf_{x \in D} |S'_i(x)| \leq \frac{|S_i(x) - S_i(y)|}{|x - y|} \leq \sup_{x \in D} |S'_i(x)|$$

that is

$$\frac{1}{2}(2 - \sqrt{3})|x - y| \leq |S_i(x) - S_i(y)| \leq 2(2 - \sqrt{3})|x - y|$$

Using Theorem 3.6.2 and Theorem 3.6.3, we estimate $0.34 < s < 1.11$. Taking two iterations of the mapping:

$$S_i \circ S_j = i + \frac{1}{j + \frac{1}{x}} = i + \frac{x}{jx + 1} \quad i, j = 1, 2$$

we can similarly estimate $0.44 < s < 0.66$. If we continue applying higher-order mappings, the dimension will approach

$$\dim_H F = 0.531.$$

4. Conclusion

Hausdorff dimension has wide-ranging applications in the study of fractal geometry, complex systems, and the peculiar structures found in the natural world. However, it is often quite challenging to calculate the Hausdorff dimension of a set directly based on its definition. In this paper, we have summarized several methods to estimate and compute the dimension of simple sets, starting from the definitions and properties of the Hausdorff measure and dimension. These methods include the upper and lower bound estimation method, potential theory method, projection theorem estimation method, fractal multiplication estimation method, self-similarity set dimension calculation method, and self-affine set dimension estimation method. These methods also have certain limitations, as calculating the Hausdorff dimension is inherently complex. As a result, there are still many methods worth exploring in this area.

References

- [1] K. J. Falconer, Fractal Geometry: Mathematical Foundations and Applications, 3rd edn. (John Wiley & Sons, Chichester, 2014).
- [2] Lu Yi. Application of bit potential theory to calculating the Hausdorff dimension [J]. Journal of Guangdong University of Mechanical Engineering, 1995 (02): 28-32.
- [3] Marstrand J M. The dimension of Cartesian product sets [M] Classics on Fractals. CRC Press, 2019: 258-266.
- [4] Yin Fengli. Principle and application of the Hausdorff measure and mass distribution for the self-similarity set [D]. Central China Normal University, 2011.
- [5] Xu Yuanfen, Jin Cheng Road. The Hausdorff dimension of the self-similarity set [J]. Journal of Jilin Normal University (Natural Science Edition), 2011, 32 (04): 111-112.
- [6] Sheng Yahui. About the Hausdorff dimension of the self-imitation set [D]. Tutor: Wen Shengyou. Hubei University, 2021.
- [7] Bárány B, Hochman M, Rapaport A. Hausdorff dimension of planar self-affine sets and measures[J]. Inventiones mathematicae, 2019, 216: 601-659.