HAUSDORFF DIMENSION

ADAM MORGAN

1. Dimension on $\mathbb{R}^n$

The main aim of these notes is to define the Hausdorff dimension on $\mathbb{R}^n$. This allows us to assign, to *every* subset U of $\mathbb{R}^n$, a real number, the "Hausdorff dimension" of U. It is worth pausing to consider what properties we might want this notion of dimension to have. There are a number of subsets of $\mathbb{R}^n$ for which everyone would agree on what their dimension should be. The linear subspaces of $\mathbb{R}^n$ are the most obvious ones, and the dimension of these should be their dimension as a vector space. For example, lines should have dimension 1 and planes dimension 2. Secondly we have curves and surfaces, whose dimension should be 1 and 2 respectively since they "look like" lines and planes. This is formalized by the notion of an $m-$dimensional manifold and curves and surfaces will be 1 and 2 dimensional sub-manifolds of $\mathbb{R}^n$ respectively. If these curves and planes are cut out by polynomial equations then we can give them the Zariski topology and there is a notion of dimension here which again gives the same answer. It would certainly be nice to define a notion of dimension that agrees with these existing (and intuitive) notions where both are defined. What is perhaps more important, however, is that our notion of dimension has the properties that we might want. These could include:

- If $V \subseteq U$ then $\dim V \leq \dim U$ "monotonicity"
- $\dim(V \cup U) = \max\{\dim(V), \dim(U)\}$ "stability"
- $\dim\left(\bigcup_{i=1}^{\infty} U_i\right) = \sup\{\dim(U_i)\}$ "countable stability"
- If U is countable then $\dim(U) = 0$
- If U is open then $\dim(U) = \dim(\mathbb{R}^n)$
- Our dimension should transform nicely under e.g. translation, rotation (homeomorphism? diffeomorphism?)

We might also ask that our dimension is easy to compute as one of the things we classically use dimension for is as a quick test to see if two spaces can be related in some way.

2. Minkowski/Box-Counting Dimension

Perhaps the most intuitive notion of dimension for arbitrary bounded subsets of $\mathbb{R}^n$ is that of box-counting dimension. Suppose we take a bounded set $U \subseteq \mathbb{R}^n$ and divide $\mathbb{R}^n$ into boxes of edge length $\frac{1}{m}$ and let $N_m(U)$ be the number of boxes that intersect U. If U is a line segment then we expect that $N_m(U)$ will grow like m as m becomes large, whilst if U is a disk, say, we expect $N_m(U)$ will grow like m^2 for m large. In general, for subsets U of $\mathbb{R}^n$ whose dimension we think should be d, we expect that $N_m(U)$ should grow like m^d for large m. This suggests that the quantity

$$\lim_{m\to\infty} \frac{\log(N_m(U))}{\log(m)}$$

might be a reasonable definition of the dimension of U.

More precisely, given a bounded subset $\phi \neq U \subseteq \mathbb{R}^n$, fix $\delta > 0$ and divide $\mathbb{R}^n$ into boxes of edge length δ (with a vertex of one box at the origin and sides directed along the co-ordinate axes, say) and define $N_\delta(U)$ to be the number of boxes that intersect U. We then define the **upper** and **lower box-counting dimensions** of U to be the quantities

$$\overline{\dim_B}(U) := \limsup_{\delta\to 0} \frac{\log\left(N_\delta(U)\right)}{-\log(\delta)}$$

and

$$\underline{\dim_B}(U) := \liminf_{\delta\to 0} \frac{\log\left(N_\delta(U)\right)}{-\log(\delta)}$$

respectively. If these values agree, then we define the **Minkowski** or **box-counting dimension** of U, $\dim_B(U)$, to be their common value.

Remark: The box-counting dimension, if it exists, need not be an integer.

Properties of, and problems with, box-counting dimension

Lemma 2.1: *Let $U, V \subseteq \mathbb{R}^n$ and suppose that $\dim_B(U)$ and $\dim_B(V)$ both exist. Then*

(i) *If $V \subseteq U$ then $\dim_B(V) \leq \dim_B(U)$*
(ii) $\dim_B(U \cup V) = \max\{\dim_B(U), \dim_B(V)\}$

Proof. (i) follows from the fact that $N_\delta(V) \leq N_\delta(U)$ for all $\delta > 0$. (ii) follows from the inequality

$$\max\{N_\delta(U), N_\delta(V)\} \leq N_\delta(U \cup V) \leq 2\max\{N_\delta(U), N_\delta(V)\}$$

□

This shows that the box-counting dimension has a number of the properties we want. Moreover, it can be shown that $\dim_B(U)$ exists and is equal to m when U is a compact, smooth, $m-$dimensional sub-manifold of $\mathbb{R}^n$, so this notion is consistent with the usual notion of dimension. However, the box-counting dimension has a serious drawback in that it fails to be countably stable. Indeed, let $U := \mathbb{Q} \cap [0,1] \subseteq \mathbb{R}$. Then

$$N_\delta(U) = \left\lceil \frac{1}{\delta} \right\rceil$$

and so $\dim_B(U)$ exists and is equal to one. However, it is immediate that a finite set has zero box-counting dimension. This is clearly not what we want; not only does countable stability fail, but a set which clearly "ought" to have dimension zero turns out to have dimension one. This limits the utility of box-counting dimension (although it is easy to calculate) and forces us to look for new notions of dimension on $\mathbb{R}^n$.

3. Hausdorff Measure and Hausdorff Dimension on $\mathbb{R}^n$

(Have in mind that this generalizes to arbitrary metric spaces)

Consider a line segment L in $\mathbb{R}^2$. Then with the Lebesgue measure, L has measure zero. However, if one identifies the line containing L with $\mathbb{R}$ and uses the Lebesgue measure here, L has finite and positive measure. Similarly, if we take a disc in $\mathbb{R}^3$ then this has zero measure with the Lebesgue measure on $\mathbb{R}^3$, but non-zero and positive measure when we identify the plane containing the disc as $\mathbb{R}^2$ and use the Lebesgue measure here. This suggests that the "correct" measure to use on a set that "should have dimension d" is the Lebesgue measure on $\mathbb{R}^d$, provided we can make sense of this. The Hausdorff measures on $\mathbb{R}^n$ allow us to rigorously give meaning to this idea for all $d \in \mathbb{R}_{\geq 0}$ and all subsets of $\mathbb{R}^n$. The Hausdorff dimension of a set $U \subseteq \mathbb{R}^n$ is then the value of d such that the measure corresponding to d is the "correct" one to measure U with. We now make this discussion precise.

Recall that we defined the Lebesgue (outer) measure on $\mathbb{R}$ by:

> Given $U \subseteq \mathbb{R}$, let $\mathcal{C}_U$ be the collection of (infinite) sequences $\{(a_i, b_i)\}_{i=1}^{\infty}$ such that $U \subseteq \bigcup_{i=1}^{\infty}(a_i, b_i)$. Then define the **Lebesgue outer measure** of U by
> $$\lambda^{\star}(U) := \inf\left\{\sum_{i=1}^{\infty}(b_i - a_i) \mid \{(a_i, b_i)\}_{i=1}^{\infty} \in \mathcal{C}_U\right\}$$

The definition of the Hausdorff (outer) measure is similar. First, given $\phi \neq U \subseteq \mathbb{R}^n$, define the **diameter** of U to be

$$|U| := \sup\{|x - y| \ : \ x, y \in U\}$$

Now let $s \geq 0$ and $\delta > 0$. We define, for $U \subseteq \mathbb{R}^n$,

$$\mathcal{H}_{\delta}^{s}(U) := \inf\left\{\sum_{i=0}^{\infty}|U_i|^s \ : \ \{U_i\}_{i=1}^{\infty} \text{ is a } \delta\text{−cover of } U\right\}$$

where $\{U_i\}_{i=1}^{\infty}$ is said to be a $\delta-$cover of U if

(i) $U \subseteq \bigcup_{i=1}^{\infty} U_i$

(ii) $|U_i| \leq \delta$ for all i

Note that the $H_{\delta}^{s}(U)$ are non-decreasing as $\delta \searrow 0$ as fewer covers become admissible. We define the $s-$**dimensional Hausdorff (outer) measure** of U to be

$$\mathcal{H}^s(U) := \lim_{\delta\to 0}\mathcal{H}_{\delta}^{s}(U) = \sup_{\delta>0}\mathcal{H}_{\delta}^{s}(U)$$

which may be (and often is as we shall see) 0 or ∞. It can be checked that this does indeed define an outer measure on $\mathbb{R}^n$ and, as with the Lebesgue measure, we can define a $\sigma-$algebra, containing $\mathcal{B}(\mathbb{R}^n)$, on which $\mathcal{H}^s$ honestly is a measure. If $s = n$ is a positive integer then (as is highly plausible) it can be checked that $\mathcal{H}^n$ on $\mathbb{R}^n$ is just a constant multiple of λ_n, the $n-$dimensional Lebesgue measure. However, it is the possibility of s not being an integer that leads to the definition of Hausdorff dimension.

Suppose $t > s$ and $\delta < 1$. Then clearly $\mathcal{H}_{\delta}^{t}(U) \leq \mathcal{H}_{\delta}^{s}(U)$ whence $\mathcal{H}^t(U) \leq \mathcal{H}^s(U)$. In fact, if $\{U_i\}_{i=1}^{\infty}$ is a $\delta-$cover of U then

$$\sum_{i=1}^{\infty}|U_i|^t = \sum_{i=1}^{\infty}|U_i|^{t-s+s} \leq \delta^{t-s}\sum_{i=1}^{\infty}|U_i|^s$$

and so, in particular,

$$\mathcal{H}_{\delta}^{t}(U) \leq \delta^{t-s}\mathcal{H}_{\delta}^{s}(U)$$

If $\mathcal{H}^s(U)$ is finite then we see that $\mathcal{H}^t(U) = 0$ for all $t > s$ and so there is a "critical value" of s such that $\mathcal{H}^s(U)$ is infinite for $t < s$ and zero for $t > s$. From this, we define the **Hausdorff dimension** of U to be

$$\dim_H(U) := \sup\{s \ : \ \mathcal{H}^s(U) = \infty\}$$

At $s = \dim_H(U)$, we may have $\mathcal{H}^s(U)$ equal to 0 or ∞, or we may have $0 < \mathcal{H}^s(U) < \infty$, in which case we say that U is an $s-$**set**.

Properties of Hausdorff dimension

Proposition 3.1: *The Hausdorff dimension on $\mathbb{R}^n$ is both monotonic and countably stable (hence stable).*

Proof. For monotonicity, if $V \subseteq U \subseteq \mathbb{R}^n$ then by virtue of being an outer measure we have

$$\mathcal{H}^s(V) \leq \mathcal{H}^s(U)$$

for all $s \geq 0$, from which the result follows. Now suppose that $\{U_m\}_{m=1}^{\infty}$ is a sequence of subsets of $\mathbb{R}^n$ and fix $s \geq 0$. By monotonicity, we have $\dim_H(U_i) \leq \dim_H\left(\bigcup_{m=1}^{\infty} U_m\right)$ for all i, whence

$$\sup_{m\geq 1}\left(\dim_H(U_m)\right) \leq \dim_H\left(\bigcup_{m=1}^{\infty} U_m\right)$$

On the other hand, suppose $t > \sup_{m\geq 1}\left(\dim_H(U_m)\right)$. Then since $\mathcal{H}^t$ is countably sub-additive, we have

$$\mathcal{H}^t\left(\bigcup_{m=1}^{\infty} U_m\right) \leq \sum_{m=1}^{\infty} \mathcal{H}^t(U_m) = 0$$

whence

$$\dim_H\left(\bigcup_{m=1}^{\infty} U_m\right) \leq \sup_{m\geq 1}\left(\dim_H(U_m)\right)$$

□

Corollary 3.2: *The Hausdorff dimension of any countable set is zero. Moreover, the Hausdorff dimension of $\mathbb{R}^n$ is n and any (non-empty) open subset of $\mathbb{R}^n$ has Hausdorff dimension n as well.*

Proof. It is immediate that the Hausdorff dimension of a single point is zero, whence the same is true for countable sets by countable stability. Now let U_m be the open ball of radius m inside $\mathbb{R}^n$. Then since U_m has finite and non-zero Lebesgue measure, it follows that $0 < \mathcal{H}^n(U_m) < \infty$ whence $\dim_H(U_m) = n$. Countable stability gives the same result for $\mathbb{R}^n$. Finally, given an arbitrary open set $\phi \neq U \subseteq \mathbb{R}^n$, U has Hausdorff dimension at least n since it contains an open ball, and at most n since it is contained in $\mathbb{R}^n$. □

Remark: It can also be shown that any smooth $m-$dimensional sub-manifold of $\mathbb{R}^n$ has Hausdorff dimension m.

The Hausdorff dimension also satisfies the following invariance property.

Lemma 3.3: *Let $U \subseteq \mathbb{R}^n$ and let $f : U \to \mathbb{R}^m$ be a bi-Lipschitz map, i.e. for all $x, y \in U$,*

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

for some $0 < c_1 \leq c_2 < \infty$. Then $\dim_H(U) = \dim_H(f(U))$.

Proof. This follows easily from the definition of Hausdorff dimension. For the details see [1]. □

Finally, we make our lives easier by showing that we can compute Hausdorff dimension by restricting to covers consisting of open balls.

Lemma 3.4: *Let $U \subseteq \mathbb{R}^n$ and define $B^s_\delta(U)$ identically to $\mathcal{H}^s_\delta(U)$, save with the added restriction that the covers consist only of open balls, and $B^s(U)$ similarly. Then we have*

$$\mathcal{H}^s(U) \leq B^s(U) \leq 4^s \mathcal{H}^s(U)$$

In particular, $B^s(U)$ is zero or infinity if and only if $\mathcal{H}^s(U)$ is.

Proof. Given any $\delta-$cover of U by arbitrary sets $\{U_i\}_{i=1}^{\infty}$, define W_i to be the open ball of radius $2|U_i|$ and centre any point of U_i (which we can assume to be of non-zero diameter). Then $\{W_i\}_{i=1}^{\infty}$ is a cover of U by open balls and $|W_i| = 4|U_i|$. This yields

$$\mathcal{H}^s_{4\delta}(U) \leq B^s_{4\delta}(U) \leq 4^s \mathcal{H}^s_\delta(U)$$

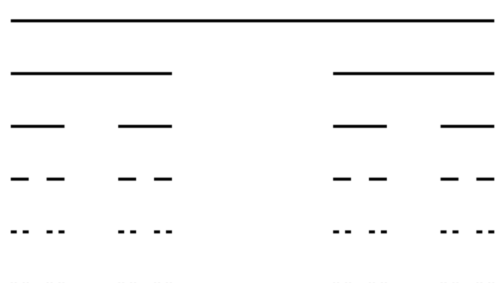
whence the result. □

Remark: In the case of $\mathbb{R}$, we can dispense with the 4^s with a little more effort.

4. The Middle-Third Cantor Set

We will illustrate the concept of Hausdorff dimension by computing the Hausdorff dimension of the middle third Cantor set, a set of independent interest.

Definition 4.1: Let $C_0 := [0,1]$. For $n \geq 0$, define C_n inductively to be C_{n-1} with the (open) middle third of each constituent interval removed, so that $C_1 = [0, \frac{1}{3}] \cup [\frac{2}{3}, 1]$ and $C_2 = [0, \frac{1}{9}] \cup [\frac{2}{9}, \frac{3}{9}] \cup [\frac{6}{9}, \frac{7}{9}] \cup [\frac{8}{9}, 1]$ etc. The picture (courtesy of [2]) is



Then define the (**middle-third**) **Cantor set** to be

$$C := \bigcap_{n=0}^{\infty} C_n$$

It is easy to see that if $x \in [0,1]$ is written in base 3 as

$$x = \sum_{n=1}^{\infty} a_n 3^{-n}$$

for $a_i \in \{0,1,2\}$ then x is in C if and only if no a_n is equal to one. In particular, C is uncountable. On the other hand, C has Lebesgue measure zero since the Lebesgue measure of C_n is

$$2^n \cdot 3^{-n} \to 0$$

as $n \to \infty$.

Proposition 4.2: *The Hausdorff dimension of the middle third Cantor set is* $s := \frac{\log(2)}{\log(3)}$. *Moreover,*

$$\frac{1}{2} \leq \mathcal{H}^s(C) \leq 1$$

Proof. Fix $n \geq 0$ and take as a cover of C the intervals comprising C_n. Denote these intervals by $U_1, ..., U_{2^n}$. Then for each $t \geq 0$ we have

$$\sum_{i=1}^{2^n} |U_i|^t = 2^n 3^{-nt} = \left(\frac{2}{3^t}\right)^n$$

For $t > \frac{\log(2)}{\log(3)}$ this gives $\mathcal{H}^s(C) = 0$ whence $\dim_H(C) \leq s$ whilst for $t = s$ we get

$$\mathcal{H}^s(C) \leq 1$$

To find a lower bound on $\dim_H(C)$ we use the remark after lemma 3.4 and restrict our covers to open intervals. Since C is compact, we can restrict to finite coverings. Fix $\delta < 1$ and let $\{U_i\}$ be a $\delta-$covering of C. For each U_i, let k_i be the largest integer such that $|U_i| < 3^{-k_i}$. Note that $k_i \geq -s \log_2(|U_i|) - 1$. Then U_i intersects at most one interval in C_{k_i}, since they are separated by at least 3^{-k_i}. Moreover, for any $j \geq k_i$, U_i intersects at most 2^{j-k_i} intervals of C_j. Since the cover is finite, we can pick $j \geq \max\{k_i\}$ whence this last statement is true for all i . However, of the 2^j intervals comprising C_j, each one has its endpoints in C and so $\bigcup_i U_i$ must intersect each of these intervals. In particular, we must have

$$2^j \leq \sum_i 2^{j-k_i}$$

and so

$$\frac{1}{2} \leq \sum_i 2^{-k_i - 1} \leq \sum_i |U_i|^s$$

Taking the infimum over all covers and the limit as $\delta \to 0$ shows that $\frac{1}{2} \leq \mathcal{H}^s(C)$ which concludes the proof. □

Remark: In fact, it can be shown that $\mathcal{H}^s(C) = 1$.

5. A systematic Approach to Computing Hausdorff Dimension; Frostman's Lemma

As we saw in the previous section it can be hard to compute lower bounds on $\mathcal{H}^s(U)$, making it hard to compute the Hausdorff dimension of U. Frostman's lemma gives a general approach to doing this.

Definition 5.1: Let μ be a measure on $\mathbb{R}^n$. That is, μ is an outer measure on $\mathbb{R}^n$ for which all Borel sets are measurable. The **support** of μ is the smallest closed set X such that $\mu(\mathbb{R}^n \backslash X) = 0$. We say that μ is a **measure on** $U \subseteq \mathbb{R}^n$ if U contains the support of μ. If μ is a measure on a bounded set $U \subseteq \mathbb{R}^n$ and $0 < \mu(\mathbb{R}^n) < \infty$ we call μ a **mass distribution**.

Lemma 5.2 (Frostman): *Let U be a bounded subset of $\mathbb{R}^n$ and μ a mass distribution on U. Fix $s > 0$ and suppose that there are positive constants c and δ such that*

$$\mu(V) \leq c|V|^s$$

for all sets V with $|V| \leq \delta$. Then $\mathcal{H}^s(U) \geq \mu(U)/c > 0$, whence $\dim_H(U) \geq s$.

Proof. If $\{U_i\}_{i=1}^{\infty}$ is any $\delta-$cover of U then

$$0 < \mu(U) \leq \mu\left(\bigcup_{i=1}^{\infty} U_i\right) \leq \sum_{i=1}^{\infty} \mu(U_i) \leq c \sum_{i=1}^{\infty} |U_i|^s$$

$\square$

The following lemma serves to provide us with suitable mass distributions.

Lemma 5.3: *Let $E \subseteq \mathbb{R}^n$ be a bounded Borel set and $\mathcal{E}_0$ consist of the set E. For $k \geq 1$, let $\mathcal{E}_k$ be a collection of disjoint Borel subsets of E such that each set $U \in \mathcal{E}_k$ is contained in precisely one set of $\mathcal{E}_{k-1}$ and itself contains a finite number of sets in $\mathcal{E}_{k+1}$. Suppose further that the maximum diameter of sets in $\mathcal{E}_k$ tends to zero as k tends to infinity. Define $\mu(E) = 1$ and (inductively) if $U_1, ..., U_m$ are the sets in $\mathcal{E}_k$ contained in some set $U \in \mathcal{E}_{k-1}$, define $\mu(U_i)$ in such a way that $\sum_{i=1}^{m} \mu(U_i) = \mu(U)$. For each k, let E_k be the union of all sets in $\mathcal{E}_k$. Define $\mu\left(\mathbb{R}^n \backslash E_k\right) = 0$.*

Then μ (is well defined and) can be extended to a measure on $\mathbb{R}^n$. Moreover, the support of μ is contained in $\bigcap_{k=1}^{\infty} \bar{E}_k$.

Proof. Omitted. See [1]for details. $\square$

Corollary 5.4: Let C be the middle third Cantor set and $s := \frac{\log(2)}{\log(3)}$. Then $\mathcal{H}^s(C) \geq \frac{1}{2}$.

Proof. By lemma 5.3 we can define a mass distribution on C such that each interval comprising C_n has measure 2^{-n}. Let $\delta < 1$ and $k \geq 0$ be the largest integer such that $\delta < 3^{-k}$. If V is a set such that $|V| \leq \delta$ then V intersects at most one of the 2^k intervals comprising C_k. In particular we have

$$\mu(V) \leq 2^{-k} \leq 2|V|^s$$

On the other hand, $\mu(C) = 1$ whence Frostman's lemma gives the result. $\square$

References

[1] K. Falconer, Fractal Geometry: mathematical foundations and applications, John Wiley & Sons Ltd. 1990

[2] http://pearlsatrandomstrung.blogspot.co.uk