Measures and Dimension
Introduction to Fractal Geometry and Chaos

Matilde Marcolli

MAT1845HS Winter 2020, University of Toronto
M 5-6 and T 10-12 BA6180

Some References

- Yakov Pesin and Vaughn Climenhaga, *Lectures on fractal geometry and dynamical systems*, American Mathematical Society, 2009
- Kenneth Falconer, *Fractal geometry*, Wiley, 2003

Topological Dimension and Hausdorff Dimension

- topological dimension
  - open coverings $\mathcal{U} = \{U_\alpha\}$ of X
  - $N(\mathcal{U}) := \sup_{x\in X} N(\mathcal{U},x)$ number $N(\mathcal{U},x)$ of open sets U_α of $\mathcal{U}$ such that $x \in U_\alpha$ (how many times the point x is covered by $\mathcal{U}$)
  - first minimize over the covering to avoid redundancy $\inf_{\mathcal{U}'\subset\mathcal{U}} N(\mathcal{U}')$ taking refinements of covering
  - then maximize over $\mathcal{U}$ to get (non-redundant) estimate of maximal complexity of the covering
  - $N(X) := \sup_{\mathcal{U}} \inf_{\mathcal{U}'\subset\mathcal{U}} N(\mathcal{U})$
  - topological dimension $\dim_{top}(X) = N(X) - 1$

• know that Hausdorff dimension behaves well with respect to Lipschitz functions

- $f : X \to Y$ with $d_Y(f(x), f(y)) \leq C\, d_X(x, y)$ then for any $Z \subseteq X$
$$\dim_H(f(Z)) \leq \dim_H(Z)$$
- for open covering $\mathcal{U} = \{U_\alpha\}$ of Z in X have $\mathrm{diam}(f(U_\alpha)) \leq C\, \mathrm{diam}(U_\alpha)$
$$\sum_\alpha (\mathrm{diam}(f(U_\alpha)))^s \leq C^s \sum_\alpha (\mathrm{diam}(U_\alpha))^s$$
$$\mu_{H,s}(f(Z)) \leq C^s \mu_{H,s}(Z)$$
so if r.h.s. finite then l.h.s. also finite, so dimension estimate
- for bi-Lipschitz functions $\dim_H(f(Z)) = \dim_H(Z)$
- no relation under more general continuous functions: for line mapped to plane via Peano curve Hausdorff dimension increases by one, under a projection decreases

- topological dimension and metrics
  - suppose topology on X is induced by a metric
  - many different metrics in general induce the same topology
  - Hausdorff dimension depends on the metric because diameters of open sets of coverings defined using metric
  - so Hausdorff dimension can vary with choice of metric (with fixed topology) and topological dimension remains the same
  - relation of Hausdorff and topological dimension: for $Z \subseteq X$

$$\dim_{\text{top}}(Z) = \inf_{d \text{ metrics}} \dim_{H,d}(Z)$$

with infimum over metric inducing the same topology

Relation between Topological and Hausdorff Dimensions

- result by Edward Marczewski (alias Szpilrajn)
  - Edward Szpilrajn, *La dimension et la mesure*, Fundamenta Mathematicae, 28 (1937) 81–89

- $\dim_{top}(Z) = \min \dim_H(\tilde{Z})$ minimum of Hausdorff dimensions of sets $\tilde{Z}$ homeomorphic to Z (inf is realized in a particular metric)

- idea of proof
  - first show that if a set Z (in an ambient metric space) has $(n+1)$-dimensional Hausdorff measure zero then it has topological dimension $\leq n$
  - then show that any such Z of topological dimension $\leq n$ and any given $s > n$, Z is homeomorphic to a set $\tilde{Z}$ inside the unit cube $\mathcal{I}^{2n+1}$ for which $\mu_{H,s}(\tilde{Z}) = 0$
  - then show that every Z with topological dimension $\leq n$ is homeomorphic to a subset $\tilde{Z}$ of the unit cube $\mathcal{I}^{2n+1}$ for which $\mu_{H,s}(\tilde{Z}) = 0$ for all $s > n$
  - this $\tilde{Z}$ then must have $\dim_H(\tilde{Z}) = n = \dim_{top}(Z)$

Topological properties from Hausdorff dimension

• Example: if $Z \subseteq \mathbb{R}$ has $\dim_H(Z) < 1$ then Z is totally disconnected

- because $\dim_H(Z) < 1$ the Hausdorff measure $\mu_{H,1}(Z) = 0$
- for $x, y \in Z$ show there are open sets U, V in $\mathbb{R}$ with $U \cap V = \emptyset$ and $U \cup V = Z$ with $x \in U$ and $y \in V$
- consider the function $f : Z \to \mathbb{R}_+$ given by $f(z) = d(x, z)$ (distance in $\mathbb{R}$)
- it is a Lipschitz function so $\dim_H(f(Z)) \leq \dim_H(Z)$, also less than one
- so also have $\mu_{H,1}(f(Z)) = 0$
- this implies $\mathbb{R} \smallsetminus f(Z)$ dense in $\mathbb{R}$ otherwise $f(Z)$ would contain some open set of $\mathbb{R}$ which would then have non-zero measure $\mu_{H,1}$ (Lebesgue measure on the line)

- so $\exists r \in \mathbb{R} \smallsetminus f(Z)$ with $0 < r < f(y)$
- take the open sets

$$U = f^{-1}([0, r)) = \{z \in Z \,|\, d(x, z) < r\}$$

$$V = f^{-1}((r, \infty)) = \{z \in Z \,|\, d(x, z) > r\}$$

disjoint $U \cap V = \emptyset$ and cover $Z = U \cup V$ (because $r \in \mathbb{R} \smallsetminus f(Z)$ so there is no point $z \in Z$ with $d(x, z) = r$)

- $x \in U$ and $y \in V$ (because r chosen with $r < f(y) = d(x, y)$) so x, y in two different components of Z (totally disconnected)

Box Counting Dimension (upper bounds on the Hausdorff dimension)

• instead of ϵ-coverings with $\mathrm{diam}(U_\alpha) \leq \epsilon$ varying (as in construction of Hausdorff measure) use a fixed size of covering only

- Z compact subset of ambient X (so can always extract finite open subcoverings)
- $\mathcal{D}(Z,\epsilon)$ = set of all open coverings of Z by open sets U_α with $\mathrm{diam}(U_\alpha) = \epsilon$
- then have for $N(\mathcal{U})$ = number of open sets in a covering $\mathcal{U}$

$$r(Z,\epsilon,s) := \inf_{\mathcal{D}(Z,\epsilon)} \sum_\alpha (\mathrm{diam}(U_\alpha))^s = \epsilon^s \inf_{\mathcal{U}\in\mathcal{D}(Z,\epsilon)} N(\mathcal{U})$$

$$= \epsilon^s N(Z,\epsilon), \quad \text{where } N(Z,\epsilon) := \inf_{\mathcal{U}\in\mathcal{D}(Z,\epsilon)} N(\mathcal{U})$$

- unlike construction of Hausdorff measure now this function has no monotonicity in ϵ because $\epsilon' < \epsilon$ coverings are not also ϵ coverings here
- so the limit when $\epsilon \to 0$ need not exist: need to take $\liminf$ and $\limsup$
- define

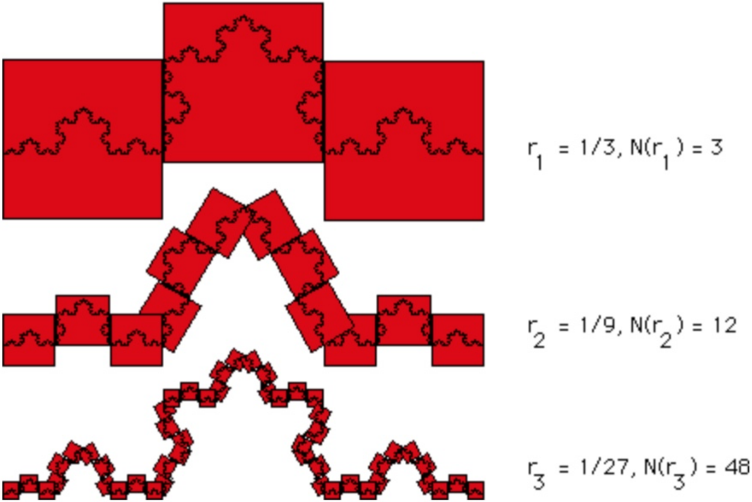
$$\underline{r}(Z, s) := \liminf_{\epsilon \to 0} r(Z, \epsilon, s) \quad \text{and} \quad \overline{r}(Z, s) := \limsup_{\epsilon \to 0} r(Z, \epsilon, s)$$

- these satisfy $\underline{r}(\emptyset, s) = \overline{r}(\emptyset, s) = 0$ and for $Z_1 \subseteq Z_2$

$$\underline{r}(Z_1, s) \leq \underline{r}(Z_2, s) \quad \text{and} \quad \overline{r}(Z_1, s) \leq \overline{r}(Z_2, s)$$

- but no subadditivity because cannot decompose an ϵ-cover as union of other coverings of smaller diameters

box counting for the Koch snowflake



- define box counting dimension as

$$\underline{\dim}_B(Z) := \inf\{s \in \mathbb{R}_+ \,|\, \underline{r}(Z,s) = 0\} = \sup\{s \in \mathbb{R}_+ \,|\, \underline{r}(Z,s) = \infty\}$$

$$\overline{\dim}_B(Z) := \inf\{s \in \mathbb{R}_+ \,|\, \overline{r}(Z,s) = 0\} = \sup\{s \in \mathbb{R}_+ \,|\, \overline{r}(Z,s) = \infty\}$$

- properties: $\underline{\dim}_B(\emptyset) = \overline{\dim}_B(\emptyset) = 0$ and for $Z_1 \subseteq Z_2$

$$\underline{\dim}_B(Z_1) \leq \underline{\dim}_B(Z_2) \quad \text{and} \quad \overline{\dim}_B(Z_1) \leq \overline{\dim}_B(Z_2)$$

- but this notion of dimension is not associated to a measure, so only have a weaker property than additivity

$$\underline{\dim}_B(\cup_\alpha Z_\alpha) \geq \sup_\alpha \underline{\dim}_B(Z_\alpha) \quad \text{and} \quad \overline{\dim}_B(\cup_\alpha Z_\alpha) \geq \sup_\alpha \overline{\dim}_B(Z_\alpha)$$

box counting is easier to compute than the Hausdorff dimension

- have $r(Z,\epsilon,s) = \epsilon^s N(Z,\epsilon)$: find where it jumps from ∞ to 0

$$r(Z,\epsilon,s) = \exp(s \log \epsilon + \log N(Z,\epsilon))$$

place where this exponent changes sign, so in the limit

$$\underline{\dim}_B(Z) = \liminf_{\epsilon} \frac{-\log N(Z,\epsilon)}{\log \epsilon}$$

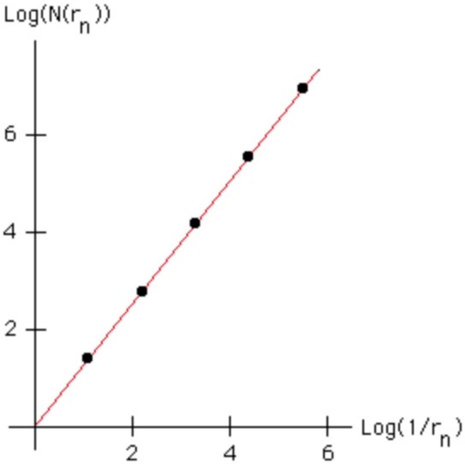
$$\overline{\dim}_B(Z) = \limsup_{\epsilon} \frac{-\log N(Z,\epsilon)}{\log \epsilon}$$

- relation to Hausdorff dimension

$$\mu_{H,s}(Z) \leq \underline{r}(Z,s) \leq \overline{r}(Z,s) \;\Rightarrow\; \dim_H(Z) \leq \underline{\dim}_B(Z) \leq \overline{\dim}_B(Z)$$

upper bounds for the Hausdorff dimension

log plot of box counting dimension for the Koch snowflake



How good (or bad) an approximation?

- Example: $Z = \mathbb{Q} \cap [0,1]$
  - $\dim_H(Z) = 0$ because for a single point $\mu_{H,s}(\text{point}) = 0$ for all $s \in \mathbb{R}_*$ and subadditivity so for countable sets also
  - what about $\underline{\dim}_B(Z)$ and $\overline{\dim}_B(Z)$?
  - general fact: if $\bar{Z}$ is the closure of Z (union with all accumulation points) then $\underline{\dim}_B(Z) = \underline{\dim}_B(\bar{Z})$ and $\overline{\dim}_B(Z) = \overline{\dim}_B(\bar{Z})$
  - reason: if a family of balls $B(x_i, \epsilon/2)$ covers Z then balls $B(x_i, \epsilon)$ cover $\bar{Z}$ (points either in Z or accumulation points of Z so arbitrarily close to Z)
  - so estimate
  $$N(\bar{Z}, 2\epsilon) \leq N(Z, \epsilon) \leq N(\bar{Z}, \epsilon)$$
  second by inclusion $Z \subseteq \bar{Z}$, so same box counting dimension
  - but then for $Z = \mathbb{Q} \cap [0,1]$ have
  $$\underline{\dim}_B(Z) = \underline{\dim}_B([0,1]) = 1 = \overline{\dim}_B([0,1]) = \overline{\dim}_B(Z)$$

An even more striking example of the gap between Hausdorff and box counting dimensions

- Claim: given any two real numbers $0 < \alpha < \beta \leq 1$ can construct a set $A \subset [0,1]$ such that

$$\dim_H(A) = 0, \quad \underline{\dim}_B(A) = \alpha, \quad \overline{\dim}_B(A) = \beta$$

- construction of the set A:
  - use a countable A so that certainly $\dim_H(A) = 0$
  - A is an increasing sequence of points in $[0,1]$ starting at 0 constructed as

$$0, a_1, 2a_1, \ldots, b_1 a_1, b_1 a_1 + a_2, b_1 a_1 + 2a_2, \ldots, b_1 a_1 + b_2 a_2, \ldots$$

  - first b_1 gaps of length a_1 then b_2 gaps of length a_2, etc.

- choice of the sequences $\{a_k\}$ and $\{b_k\}$:
  - set A constructed as above

$$\ldots, \sum_{k=1}^{b} b_k a_k + a_{n+1}, \ldots, \sum_{k=1}^{n} b_k a_k + b_{n+1} a_{n+1}, \ldots$$

  - use a rapidly decreasing sequence for a_n for example $a_n = e^{-n}$
  - the choice of b_k sequence of non-negative integers will be made to obtain α and β as box counting dimensions
  - Note: the endpoints between the differently spaced parts of the sequence $A = \{x_n\}$ are

$$T_n = \sum_{k=1}^{n} b_k a_k, \quad \text{the limit } T := \lim_n T_n = \lim_n x_n$$

is the limit point of A as a sequence

- select properties for the sequence $\{b_n\}$:
  - $\{b_n\}$ monotonically increasing sequence of positive integers with $\lim_n b_n = \infty$
  - convergent series $\sum_{n=1}^{\infty} b_n a_n < 1$ (so $T \in [0,1]$)
  - want the sequence of partial sums $S_n := \sum_{k=1}^{n} b_k$ to satisfy

$$\liminf_{n\to\infty} \frac{1}{n} \log S_n = \alpha \quad \text{and} \quad \limsup_{n\to\infty} \frac{1}{n} \log S_n = \beta$$

  - a *no long tail* condition: want the tail $[T_n, T] \cap A$ of the set A to satisfy

$$\frac{T_n - T}{a_n S_n} \leq C$$

    for all $n \in \mathbb{N}$ and for some $C > 0$

- why these properties for $\{b_n\}$ work:
  - by $\sum_{n=1}^{\infty} b_n a_n < 1$ know $A \subset [0,1]$
  - given $\epsilon > 0$ choose $n = n(\epsilon)$ such that

$$e^{-(n+1)} < \epsilon \leq e^{-n}$$

  - the first S_n points of A are all separated by a gap of at least e^{-n} so

$$N(A,\epsilon) \geq S_n$$

    (need at least one set of size ϵ for each of these points for covering)
  - number of intervals of length ϵ needed to cover the rest $[T_n, T] \cap A$ is $\leq (T - T_n)/e^{-(n+1)}$ (because $e^{-(n+1)} < \epsilon$ and this counts how many $e^{-(n+1)}$-gaps in $[T_n, T]$)
  - since chosen $a_n = e^{-n}$ have by short tail estimate

$$\frac{T - T_n}{e^{-(n+1)}} = \frac{1}{e} \frac{T - T_n}{a_n} \leq \frac{C}{e} S_n$$

- putting these two estimates together get (for $n = n(\epsilon)$ as above)

$$N(A,\epsilon) \leq (1 + \frac{C}{e})\, S_n$$

- so estimate for log ratio computing box counting dimensions

$$\frac{\log(S_n)}{n+1} \leq \frac{\log N(A,\epsilon)}{-\log\epsilon} \leq \frac{\log(S_n) + \log(1+\frac{C}{e})}{n}$$

- so obtain that $\underline{\dim}_B(A)$ and $\overline{\dim}_B(A)$ are computed, respectively by

$$\liminf_n \frac{\log(S_n)}{n} = \alpha \quad \text{and} \quad \limsup_n \frac{\log(S_n)}{n} = \beta$$

- the requested properties for $\{b_n\}$ can all be satisfied:
  - b_n first grows like $e^{\alpha n}$ until the first $n = n_1$ where

$$\left| \frac{1}{n_1} \log S_{n_1} - \alpha \right| < \frac{1}{2}$$

  - then b_n grows like $e^{\beta n}$ (for instance $b_{n+1} = M b_n$ for some $e^{\beta} < M < e$ unti it reaches $e^{\beta n}$ then continue as $e^{\beta n}$) until the first $n = n_2$ where

$$\left| \frac{1}{n_2} \log S_{n_2} - \beta \right| < \frac{1}{2}$$

  - then b_n stays constant until equal again to $e^{\alpha n}$ and then keeps growing like $e^{\alpha n}$ until next $n = n_3$ where again close to α up to another factor of $1/2$ and so on

$$\left| \frac{1}{n_{4k+1}} \log S_{n_{4k+1}} - \alpha \right| < \frac{1}{2^{k+1}}$$

  similarly for proximity to β so have liminf and limsup as prescribed

- have monotonicity of the b_n (non-strict)
- series convergence: $b_n \leq e^{\beta n}$ and $a_n = e^{-n}$ so

$$\sum_n b_n a_n \leq \frac{e^{\beta-1}}{1-e^{\beta-1}} < 1$$

- no long tail condition:

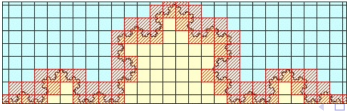
$$\frac{T - T_n}{a_n S_n} \leq \sum_{k=n+1}^{\infty} \frac{e^{-k} b_k}{e^{-n} b_n} \leq \sum_{k=n+1}^{\infty} \left(\frac{M}{e}\right)^k < \infty$$

where used $S_n \geq b_n$ and $\frac{b_{n+1}}{b_n} \leq M$

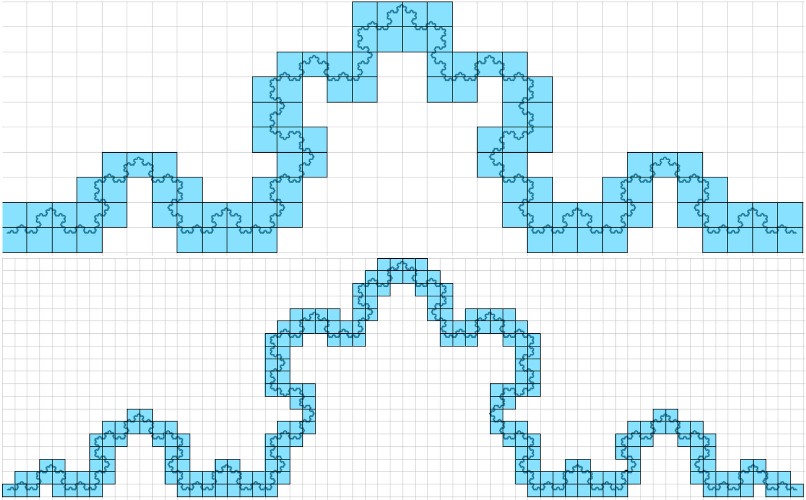
- so this construction of the sequence $\{b_n\}$ satisfies all the required properties

Good estimate versus easy computability

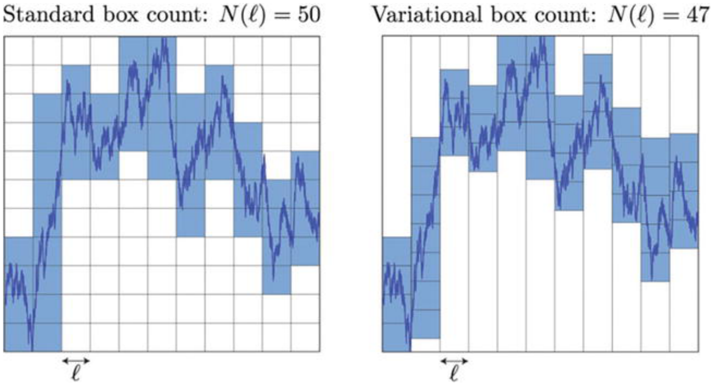
- by previous examples: box counting is not a good estimator of Hausdorff dimension, it gives a rough upper bound, which can often be much larger than the actual dimension
- it works better on more regular self-similar fractals
- it is used very frequently (most uses of fractals in applications to sciences outside of mathematics are based on box counting estimates)
- more easily computable advantage over accuracy of estimate
- even larger bound by restricting to only very simple covers by regular grids of varying sizes instead of using inf over all $\mathcal{U} \in \mathcal{D}(Z, \epsilon)$
- these larger bounds are just counting pixels that cover image at a certain resolution (much easier computation, but even coarser bound)



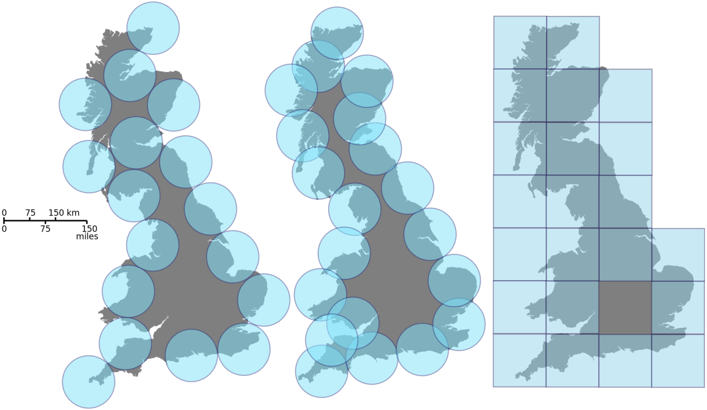
Example: grid box counting for the Koch snowflake



Example: grid and variable box counting for graph of a non-differentiable function



Comparison of box counting methods



sphere packings, sphere coverings, grid boxes: for Euclidean metric not same counting $N(\epsilon)$ but the logarithmic limit $\lim_{\epsilon\to 0} -\frac{N(\epsilon)}{\log(\epsilon)}$ cancels difference, but for other (variable) metrics comparison with grid boxed no longer equivalent; using balls generalizes for arbitrary metrics

Using measures to determine dimension

- upper bounds for the Hausdorff dimension can be obtained in various ways (eg box counting)
- lower bounds are much more difficult to obtain

• can use measures with uniform mass distribution

- measure μ on $\mathbb{R}^N$ for which there are $\alpha, k, \delta > 0$ such that every ball $B(x, r)$ of radius $r < \delta$ has measure

$$\mu(B(x,r)) \leq Kr^\alpha$$

- consider a measurable set $E \subset \mathbb{R}^N$ with $\mu(E) > 0$
- for all $\epsilon > 0$ with $0 < \epsilon < \delta$ and open covering $\mathcal{U} = \{U_j\}$ of E with $\mathrm{diam}(U_j) < \epsilon$ have

$$\sum_j (\mathrm{diam}(U_j))^\alpha \geq \sum_j \frac{\mu(U_j)}{K} \geq \frac{\mu(\cup_j U_j)}{K} \geq \frac{\mu(E)}{K} > 0$$

so have $\mu_{H,\alpha}(E) > 0$: estimate on Hausdorff dimension

$$\dim_H(E) \geq \alpha$$

Problem of constructing measures with uniform mass distribution

- on the shift space Σ_k^+ of sequences on k letters, consider the Bernoulli measure μ_P with probabilities $P = (p_1, \dots, p_k)$

$$\mu_P(\mathcal{C}(w_1, \dots, w_n) = p_{w_1} \cdots p_{w_n}$$

- realizing this space as a subset C of $\mathbb{R}$ with intervals $\mathcal{I}_{w_1,\dots,w_n}$ obtained by piecewise linear maps on an initial interval $\mathcal{I}$ with scaling factors λ_{w_i} (contraction rates) so that

$$\mathrm{diam}(\mathcal{I}_{w_1,\dots,w_n}) = \lambda_{w_1} \cdots \lambda_{w_n} \cdot \mathrm{diam}(\mathcal{I})$$

- condition for μ_P to satisfy the uniform mass hypothesis: need to have a constant $K > 0$ and an $\alpha > 0$ independent of $w_1, \dots, w_n$ such that

$$\frac{\mu_P(\mathcal{I}_{w_1,\dots,w_n})}{\mathrm{diam}(\mathcal{I}_{w_1,\dots,w_n})^\alpha} \le K$$

- for this to work need to take $p_i = \lambda_i^\alpha$
- because of probability normalization $\sum_i p_i = 1$ need to have $\alpha = \dim_{\text{self-sim}}(C)$ because

$$\sum_i \lambda_i^\alpha = 1$$

- also any subset of $E \subseteq C$ that has $\mu_P(C) > 0$ for this measure has $\dim_{\mathrm{self-sim}}(E) = \alpha$
- in this case for Cantor sets self-similarity dimension same as Hausdorff dimension

Pointwise Dimension: more refined notion of dimension that varies with the point (behavior of balls under non-uniform measures)

- pointwise dimension

$$d_\mu(x) := \lim_{r\to 0} \frac{\log \mu(B(x,r))}{\log r}$$

is defined if the limit exists

- it's a property of a measure μ not of a set
- if the limit exists then μ-volumes of balls around point x

$$r^{d_\mu(x)+\epsilon} \leq \mu(B(x,r)) \leq r^{d_\mu(x)-\epsilon}$$

for $0 < r < \delta < 1$ where δ depends on x in general

Example: Cantor set shift space Σ_2^+ with Bernoulli measure $P = (p, q = 1 - p)$

- Cantor set C in $[0,1]$ with two contractions of equal ratio λ
- initial intervals $\mathcal{I}_1, \mathcal{I}_2$ of lengths λ (like middle third, but λ not necessarily $1/3$)
- alphabet $A = \{0, 1\}$
- balls are cylinder sets $\mathcal{C}(w_1, \ldots, w_n)$ with measure

$$\mu_P(\mathcal{C}(w_1, \ldots, w_n)) = p^{a_n(w)} \, q^{n - a_n(w)}$$

- $a_n(w)$ = number of digits equal to 1 in the string $w = w_1, \ldots, w_n$
- $n - a_n(w)$ = number of digits equal to 0 in string w
- if uniform measure $p = q = \frac{1}{2}$ then

$$d_{\mu_P}(x) = \lim_{n \to \infty} \frac{\mu_P(\mathcal{C}(w_1, \ldots, w_n))}{n \log \lambda} = \frac{\log 2}{-\log \lambda} = \dim_H(C)$$

- here $\log r = n \log \lambda$ since length of interval $\mathcal{I}_{w_1, \ldots, w_n}$ is λ^n

Pointwise dimension $d_\mu(x)$ and Hausdorff dimension of $\mathrm{supp}(\mu)$

- when limit defining $d_\mu(x)$ does not exist take liminf and limsup $\underline{d_\mu}(x)$ and $\overline{d_\mu}(x)$
- generalization of uniform mass distribution property
- μ measure on $\mathbb{R}^N$, subset $E \subset \mathbb{R}^N$ with $\mu(E) > 0$ and for some $\alpha > 0$

$$\underline{d_\mu}(x) \geq \alpha$$

for μ-almost every $x \in E$, then

$$\dim_H(E) \geq \alpha$$

- check that for all $\epsilon > 0$ have $\mu_{H,\alpha-\epsilon}(E) > 0$
  - μ-almost everywhere in E have

$$\alpha \leq \underline{d_\mu}(x) = \liminf_{r\to 0} \frac{\log \mu(B(x,r))}{\log r}$$

so $\exists \delta$ such that for $0 < r < \delta < 1$

$$\alpha - \epsilon \leq \frac{\log \mu(B(x,r))}{\log r} \quad \Rightarrow \quad \mu(B(x,r)) \leq r^{\alpha-\epsilon}$$

- given $n \in \mathbb{N}$ take

$$E_n := \{x \in E \,|\, \mu(B(x,r)) \leq r^{\alpha-\epsilon}, \ \forall 0 < r < \frac{1}{n}\}$$

- have $\mu(\cup_n E_n) = \mu(E)$ because almost everywhere property
- then for some n must have $\mu(E_n) > 0$
- show that $\mu_{H,\alpha-\epsilon}(E_n) > 0$ (then also $\mu_{H,\alpha-\epsilon}(E) > 0$)
- any open covering of E_n by balls $B(x_i, r_i)$ of radii $r_i \leq \delta_n$ gives

$$\sum_i \mathrm{diam}(B(x_i, r_i))^{\alpha-\epsilon} = 2^{\alpha-\epsilon} \sum_i r_i^{\alpha-\epsilon} \geq$$

$$2^{\alpha-\epsilon} \sum_i \mu(B(x_i, r_i)) \geq 2^{\alpha-\epsilon} \mu(\cup_i B(x_i, r_i)) \geq 2^{\alpha-\epsilon} \mu(E_n)$$

- so $\mu_{H,\alpha-\epsilon}(E_n) \geq 2^{\alpha-\epsilon} \mu(E_n) > 0$
- then $\mu_{H,\alpha-\epsilon}(E) > 0$ so $\dim_H(E) \geq \alpha - \epsilon$ for all $\epsilon > 0$

$$\dim_H(E) \geq \alpha$$

Example: Cantor set with equal contraction ratios $\lambda_1 = \lambda_2 = \lambda$ and with (non-uniform) Bernoulli measure μ_P with $P = (p, q = 1 - p)$

- know Hausdorff dimension is $\dim_H(C) = \frac{\log 2}{-\log \lambda}$
- $\mu(\mathcal{I}_1) = p$ and $\mu(\mathcal{I}_2) = q$, $p + q = 1$
- pointwise dimension

$$d_{\mu_P}(x) = \lim_{n \to \infty} \frac{\log \mu(\mathcal{C}(w_1, \ldots, w_n))}{\log \ell(\mathcal{I}_{w_1, \ldots, w_n})}$$

with length of intervals $\ell(\mathcal{I}_{w_1, \ldots, w_n}) = \lambda^n$

$$d_{\mu_P}(x) = \lim_{n \to \infty} \frac{\log(p^{a_n(w)} q^{n - a_n(w)})}{n \log \lambda}$$

$$= \lim_{n \to \infty} \frac{a_n(w)}{n} \frac{\log p}{\log \lambda} + \left(1 - \lim_{n \to \infty} \frac{a_n(w)}{n}\right) \frac{\log(1-p)}{\log \lambda}$$

- crucial quantity here:

$$\alpha(x) := \lim_{n\to\infty} \frac{a_n(x)}{n}$$

  the proportion of 1's in the binary infinite sequence $x = w_1 w_2 \dots w_n \dots$
- in the case of the *uniform* Bernoulli measure $P = (\frac{1}{2}, \frac{1}{2})$ expect this frequency to be $1/2$ almost everywhere
- but lots of points where not true (e.g. points with eventually constant expansion for which $\alpha(x) = 0$ or $\alpha(x) = 1$)
- typical situation: "good" set of points where asymptotics is as expected is large in measure; "bad" set of points where different limit or no limit is large topologically (dense) but small in measure (measure zero)
- easy to see because cyclic points (with a period $w_1, \dots, w_N$ that repeats for some N) are dense but do not have the typical limit density $1/2$

• for a general Bernoulli measure μ_P on Σ_2^+ with $P = (p.q = 1 - p)$ have

$$\alpha(x) = \lim_{n\to\infty} \frac{a_n(x)}{n} = p$$

μ-almost everywhere, where p probability of digit 1

- for $n \in \mathbb{N}$ and $\epsilon > 0$ take set

$$Z_{\epsilon,n} := \{x \in C \mid \left| \frac{a_n(x)}{n} - p \right| \geq \epsilon\}$$

- a point $x \in C$ is in the "bad set" B iff $\exists \epsilon > 0$ such that $\forall N$, $\exists n \geq N$ such that $x \in Z_{\epsilon,n}$

$$B = \bigcup_{\epsilon>0} \bigcap_{N\in\mathbb{N}} \bigcup_{n\geq N} Z_{\epsilon,n}$$

- so to show $\mu_P(B) = 0$ estimate measures of the sets $Z_{\epsilon,n}$
- property defining $Z_{\epsilon,n}$ only depends on first n digits of x, so if a point $x \in Z_{\epsilon,n}$ the whole interval $\mathcal{I}_{w_1,\ldots,w_n}$ with the same first n digits is also contained in $Z_{\epsilon,n}$
- define $\mathcal{W}_{\epsilon,n} :=$ set of strings $(w_1, \ldots, w_n)$ such that $\mathcal{I}_{w_1,\ldots,w_n} \subseteq Z_{\epsilon,n}$

$$\mu(Z_{\epsilon,n}) = \sum_{(w_1,\ldots,w_n)\in\mathcal{W}_{\epsilon,n}} \mu(\mathcal{I}_{w_1,\ldots,w_n})$$

$$= \sum_{(w_1,\ldots,w_n)\in\mathcal{W}_{\epsilon,n}} p_{w_1}\cdots p_{w_n} \leq$$

$$\frac{1}{\epsilon^2} \sum_{w=(w_1,\ldots,w_n)\in\mathcal{W}_{\epsilon,n}} \left(\frac{a_n(w)}{n} - p\right)^2 p_{w_1}\cdots p_{w_n}$$

- write $(a_n(w) - np)^2 = a_n(w)^2 - 2npa_n(w) + n^2p^2 = A_1(w) + A_2(w) + A_3(w)$

$$\mu(Z_{\epsilon,n}) \leq \frac{1}{n^2\epsilon^2} \sum_w (A_1(w) + A_2(w) + A_3(w))\, p_{w_1} \cdots p_{w_n}$$

- third term:

$$\sum_w A_3(w)\, p_{w_1} \cdots p_{w_n} = \sum_{w_1,\dots,w_n} n^2p^2\, p_{w_1} \cdots p_{w_n} = n^2p^2$$

- second term: rewriting $a_n(w) = \delta_1(w_1) + \cdots + \delta_1(w_n)$ with

$$\delta_1(w_j) = \begin{cases} 1 & w_j = 1 \\ 0 & w_j \neq 1 \end{cases}$$

$$\sum_w A_2(w)\, p_{w_1} \cdots p_{w_n} = -2np \sum_{w_1,\dots,w_n} a_n(w)\, p_{w_1} \cdots p_{w_n} =$$

$$-2np \sum_{w_1,\dots,w_n} \sum_j \delta_1(w_j)\, p_{w_1} \cdots p_{w_n} =$$

$$= -2np \sum_j \sum_{w_j} \delta_1(w_j)\, p_{w_1} \cdots p_{w_n} = -2np \sum_j \sum_{w_j} \delta_1(w_j) p_{w_j}$$

$$= -2n^2 p^2$$

- first term: $\delta_1(w_j)^2 = \delta_1(w_j)$ so write

$$a_n(w)^2 = (\sum_j \delta_1(w_j))^2 = \sum_j \delta_1(w_j) + \sum_{i \neq j} \delta_1(w_i)\delta_1(w_j)$$

then get

$$\sum_w A_1(w)\, p_{w_1} \cdots p_{w_n} =$$

$$\sum_{w_1,\ldots,w_n} (\sum_j \delta_1(w_j) + \sum_{i \neq j} \delta_1(w_i)\delta_1(w_j))\, p_{w_1} \cdots p_{w_n}$$

$$= np + \sum_{i \neq j} \sum_{w_i, w_j} \delta_1(w_i)\delta_1(w_j))\, p_{w_i} p_{w_j} = n(n-1)p^2$$

- then obtain an estimate

$$\mu_P(Z_{\epsilon,n}) \leq \frac{1}{n^2\epsilon^2}(np + n(n-1)p^2 - 2n^2p^2 + n^2p^2) = \frac{p-p^2}{n^2\epsilon^2}$$

- if $w \in Z_{2\epsilon,n}$ then either

$$a_n(w) \geq pn + 2\epsilon n \quad \text{or} \quad a_n(w) \leq pn - 2\epsilon n$$

- also $a_n(w) \leq a_m(w) \leq a_n(w) + \epsilon n$ for $n \leq m \leq (1+\epsilon)n$ so the density $a_n(w)/n$ does not change much for nearby n and can be controlled by changing ϵ slightly
- if ϵ is small enough that $p + \epsilon < 1$

$$a_m(w) \geq a_n(w) > pn + (p+\epsilon)\epsilon n + \epsilon n \geq pm + \epsilon m$$

as in first $a_n(w)$ estimate above, and if also $p - \epsilon > 0$ get also

$$a_m(w) \leq a_n(w) + \epsilon n \leq (p-\epsilon)n \leq (p-\epsilon)m$$

as in the second $a_n(w)$ estimate

- this implies inclusion $Z_{2\epsilon,n} \subset Z_{\epsilon,m}$ when $n \leq m \leq (1+\epsilon)n$
- then have

$$\bigcup_{n \geq N} Z_{2\epsilon,n} \subset \bigcup_{k \geq 0} Z_{\epsilon,n_k}$$

where n_k is given by $n_k = \lfloor (1+\epsilon)^k N \rfloor$

- then measures give

$$\mu_P(\cup_{n \geq N} Z_{2\epsilon,n}) \leq \sum_{k=0}^{\infty} \mu_P(Z_{\epsilon,n_k})$$

$$\leq \frac{p-p^2}{N\epsilon^2} \sum_{k=1}^{\infty} (1+\epsilon)^{-k} \leq \frac{p-p^2}{N\epsilon^2} \frac{1+\epsilon}{\epsilon}$$

- the N in the denominator (while ϵ is fixed) implies then

$$\mu_P(\bigcap_{N \in \mathbb{N}} \bigcup_{n \geq N} Z_{2\epsilon,n}) = 0$$

- this then implies that the measure of the "bad set" is also zero, $\mu_P(B) = 0$

Exact Dimensionality and Hausdorff Dimension

- just shown that μ_P-almost everywhere the pointwise dimension of a Cantor set with contraction ratios $\lambda_1 = \lambda_2 = \lambda$

$$d_{\mu_P}(x) = \frac{p \log p + (1-p) \log(1-p)}{\log \lambda}$$

for μ_P the Bernoulli measure with $P = (p, 1-p)$

- a measure μ is exact dimensional if there exists an α such that μ-almost everywhere $d_\mu(x) = \alpha$
- call this α the Hausdorff dimension of the measure $\dim_H(\mu)$
- we know $\dim_H(E) \geq \alpha$ whenever $\mu(E) > 0$ (in particular for $E = \mathrm{supp}(\mu)$)

- so for the Cantor set C

$$\dim_H(C) \geq \frac{p\log p + (1-p)\log(1-p)}{\log \lambda}$$

- indeed $p\log p + (1-p)\log(1-p)$ is the Shannon entropy that is maximal at $p = 1/2$ and this gives the correct Hausdorff dimension

$$\dim_H(C) = \frac{\log 2}{\log \lambda} = \max_{p\in[0,1]} \frac{p\log p + (1-p)\log(1-p)}{\log \lambda}$$

- Note the occurrence of Shannon entropy in this computation of pointwise dimensions!
- is it accidental? what is the relation between entropy and dimension?
- if the numerator is an entropy/information, what is the the denominator $\log \lambda$? (Lyapunov exponent)