

Measure Theory

Matilde Marcolli

Caltech, Spring 2026

Ma4/104: Introduction to Fractal Geometry and Chaos

Measure Spaces

rigorous theory of “volumes” (and of probabilities)

- **measure space** (X, Σ, μ)
 - **σ -algebra** Σ on a set X
 - Σ collection of subsets of X such that
 - \emptyset and X belong to Σ
 - *complements*: $A^c = X \setminus A$ is in Σ whenever A is in Σ
 - *countable unions*: $\bigcup_{k=1}^{\infty} A_k$ is in Σ if $A_k \in \Sigma$ for all $k \in \mathbb{N}$
 - **measure** $\mu : \Sigma \rightarrow \mathbb{R}_+ \cup \{\infty\}$
 - $\mu(\emptyset) = 0$
 - *monotonicity*: $\mu(A_1) \leq \mu(A_2)$ if $A_1 \subseteq A_2$
 - *disjoint additivity*: $\mu(\bigsqcup_{k=1}^{\infty} A_k) = \sum_{k=1}^{\infty} \mu(A_k)$ additive on disjoint unions
- **probability distribution**: a measure space where $\mu(X) = 1$
 - $A \in \Sigma$ events
 - $\mu(A)$ their probabilities

Examples

- ① $\Sigma = 2^X$ all subsets of X and $\mu = \delta_x$ (Dirac delta measure supported at a chosen point $x \in X$) with

$$\delta_x(A) = \begin{cases} 1 & x \in A \\ 0 & x \notin A \end{cases}$$

- ② $\Sigma = 2^X$ and counting (cardinality) measure

$$\mu(A) = \begin{cases} \#A & A \text{ finite} \\ \infty & A \text{ infinite} \end{cases}$$

- ③ If X finite, counting measure can be normalized to a probability (uniform probability)

$$\mu(A) = \frac{\#A}{\#X}$$

Support of a measure

- measure space (X, Σ, μ)
- assume Borel measure: (X, \mathcal{T}) topological space and open sets $U \in \mathcal{T}$ of the topology are all measurable sets $U \in \Sigma$
- support of μ

$$\text{supp}(\mu) = \{x \in X \mid \mu(U) > 0 \forall U \in \mathcal{T} \text{ with } x \in U\}$$

- an open set $U \in \mathcal{T}$ has $\mu(U) > 0$ iff $U \cap \text{supp}(\mu) \neq \emptyset$ and $\mu(U) = 0$ otherwise
- Dirac delta measure δ_x has $\text{supp}(\delta_x) = \{x\}$

Continuity of measures

- limits of sets

$$\liminf_{n \rightarrow \infty} A_n = \bigcup_{k=1}^{\infty} \bigcap_{j=k}^{\infty} A_j \quad \limsup_{n \rightarrow \infty} A_n = \bigcap_{k=1}^{\infty} \bigcup_{j=k}^{\infty} A_j$$

same for increasing or decreasing sequence

- (X, Σ, μ) measure space
 - increasing sequence $A_1 \subset A_2 \subset \dots \subset A_n \subset \dots$

$$\mu\left(\lim_{n \rightarrow \infty} A_n\right) = \lim_{n \rightarrow \infty} \mu(A_n)$$

- decreasing sequence $A_1 \supset A_2 \supset \dots \supset A_n \supset \dots$

$$\mu\left(\lim_{n \rightarrow \infty} A_n\right) = \lim_{n \rightarrow \infty} \mu(A_n)$$

- arbitrary sequences

$$\mu\left(\liminf_{n \rightarrow \infty} A_n\right) \leq \liminf_{n \rightarrow \infty} \mu(A_n)$$

- increasing sequence: limit is union, use

$\bigcup_{n=1}^{\infty} A_n = A_1 \sqcup_{k=2}^{\infty} (A_k \setminus A_{k-1})$ and additivity of measure

$$\mu(\bigcup_k A_k) = \mu(A_1) + \sum_{k=2}^{\infty} \mu(A_k \setminus A_{k-1})$$

$$= \lim_k \mu(A_1 \sqcup_{k=2}^{\infty} (A_k \setminus A_{k-1})) = \lim_k \mu(A_k)$$

- decreasing sequence: limit is intersection, use $B_k = A_1 \setminus A_k$
increasing sequence

$$\mu(\bigcap_k A_k) = \mu(A_1) - \mu(\bigcup_j B_j) = \mu(A_1) - \lim_j \mu(B_j) = \lim_j \mu(A_j)$$

- arbitrary sequence: use $B_k = \bigcap_{j=k}^{\infty} A_j$ increasing sequence

$$\mu(\liminf_{n \rightarrow \infty} A_n) = \mu(\bigcup_{k=1}^{\infty} B_k) = \lim_k \mu(B_k) \leq \liminf_k \mu(A_k)$$

Almost everywhere properties

- (X, Σ, μ) measure space, a property of X holds *almost everywhere* if it fails to hold on a set of μ -measure zero
- Examples:
 - X with δ_x Dirac measure: two functions $f_1, f_2 : X \rightarrow \mathbb{R}$ agree almost everywhere iff $f_1(x) = f_2(x)$ at the point x where δ_x is supported (sets of measure zero = those not containing x)
 - X with counting measure: two functions $f_1, f_2 : X \rightarrow \mathbb{R}$ agree almost everywhere iff $f_1(x) = f_2(x)$ at all points $x \in X$ (no non-empty sets of measure zero)
- Note: in the case of probability spaces probability zero does not mean impossibility
- How to construct more interesting measures than the Dirac and counting examples?

Outer measures and measures

- $\mu^* : 2^X \rightarrow \mathbb{R}_+ \cup \{\infty\}$ defined on all subsets with
 - $\mu^*(\emptyset) = 0$
 - $\mu(A_1) \leq \mu(A_2)$ if $A_1 \subseteq A_2$
 - *countable subadditivity*:

$$\mu^*\left(\bigcup_{k=1}^{\infty} A_k\right) \leq \sum_{k=1}^{\infty} \mu(A_k)$$

even when disjoint union require only \leq not additivity

- use an outer measure μ^* to construct a σ -algebra Σ and a measure space (X, Σ, μ)
 - σ -algebra determined by μ^*
$$\Sigma := \{E \subseteq X : \mu^*(A) = \mu^*(A \cap E) + \mu^*(A \cap (X \setminus E)), \forall A \subseteq X\}$$
 - $\mu = \mu^*|_{\Sigma}$ is a measure (the definition of Σ forces additivity on disjoint unions)

- σ -algebra property of

$$\Sigma := \{E \subseteq X : \mu^*(A) = \mu^*(A \cap E) + \mu^*(A \cap (X \setminus E)), \forall A \subseteq X\}$$

- \emptyset and X are in Σ
- condition defining Σ is symmetric in E and $X \setminus E$ so if a set is in Σ its complement also is
- given a sequence of sets E_k in Σ

$$\begin{aligned} \mu^*(A) &= \mu^*(A \cap E_1) + \mu^*(A \setminus E_1) = \\ \mu^*(A \cap E_1) &+ \mu^*((A \setminus E_1) \cap E_2) + \mu^*((A \setminus E_1) \setminus E_2) = \\ &\sum_{j=1}^k \mu^*((A \setminus \cup_{i=1}^{j-1} E_i) \cap E_j) + \mu^*(A \setminus \cup_{j=1}^k E_j) \\ &\geq \sum_{j=1}^k \mu^*((A \setminus \cup_{i=1}^{j-1} E_i) \cap E_j) + \mu^*(A \setminus \cup_{j=1}^{\infty} E_j) \end{aligned}$$

- estimate for all k so in the limit

$$\mu^*(A) \geq \sum_{j=1}^{\infty} \mu^*((A \setminus \cup_{i=1}^{j-1} E_i) \cap E_j) + \mu^*(A \setminus \cup_{j=1}^{\infty} E_j)$$

- then use $A \cap \cup_{j=1}^{\infty} E_j = \cup_{j=1}^{\infty} ((A \setminus \cup_{i=1}^{j-1} E_i) \cap E_j)$

$$\mu^*(A) \leq \mu^*(A \cap \cup_{j=1}^{\infty} E_j) + \mu^*(A \setminus \cup_{j=1}^{\infty} E_j)$$

and by previous also

$$\mu^*(A \cap \cup_{j=1}^{\infty} E_j) + \mu^*(A \setminus \cup_{j=1}^{\infty} E_j) \leq \mu^*(A)$$

- so $\cup_{j=1}^{\infty} E_j \in \Sigma$ closed under countable unions

- restriction $\mu = \mu^*|_{\Sigma}$ is a measure
 - disjoint union of sets E_j in Σ , use estimate

$$\mu^*(A) \geq \sum_{j=1}^{\infty} \mu^*((A \setminus \sqcup_{i=1}^{j-1} E_i) \cap E_j) + \mu^*(A \setminus \sqcup_{j=1}^{\infty} E_j)$$

- applied to $A = \sqcup_j E_j$ gives

$$\mu^*(\sqcup_{j=1}^{\infty} E_j) \geq \sum_{j=1}^{\infty} \mu^*(E_j)$$

- together with subadditivity of the outer measure gives additivity on disjoint unions in Σ

Carathéodory construction

- **General idea:**

- 1 start with a class of sets \mathcal{S} for which have a good definition of what an extension ℓ (length, area, volume) should be (example: intervals on the real line and their length)
- 2 define an outer measure on all subsets of the ambient X by setting

$$\mu^*(A) := \inf_{A \subset \cup_k S_k} \sum_k \ell(S_k)$$

infimum over all sequences of sets in the sample class whose union covers A

- 3 then pass to σ -algebra and measure (X, Σ, μ) defined by this outer measure
- 4 this measure μ will still have the property that it agrees with the original ℓ on sets $S \in \mathcal{S}$ belonging to the chosen sample class, $\mu(S) = \ell(S)$

Existence of non-measurable sets

- distinction between the σ -algebra Σ defined by an outer measure and the full set 2^X of all subsets of X lies in the possible existence of non-measurable sets
- for the case of $\mathcal{S} =$ intervals in the real line and $\ell =$ length, with μ the Lebesgue measure on \mathbb{R} (translation invariant measure) existence of non-measurable sets not constructive: use of the axiom of choice
- existence of non-measurable sets on the sphere S^2 with respect to the rotation invariant measure related to Banach-Tarski paradox of duplication of the sphere



Non-measurable sets: Banach-Tarski paradox

- F_2 free group on two generators a, b
- all words of arbitrary length in the letters a, b, a^{-1}, b^{-1} with no cancellations (no a preceded or followed by a^{-1} and no b preceded or followed by b^{-1})
- additional “empty word \emptyset (identity element of the group)
- product = concatenation of words (with elimination of cancellations)

$$F_2 = \{\emptyset\} \cup \mathcal{W}(a) \cup \mathcal{W}(b) \cup \mathcal{W}(a^{-1}) \cup \mathcal{W}(b^{-1})$$

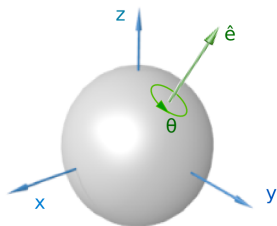
$\mathcal{W}(x)$ = words starting with letter x

can also write as

$$F_2 = \mathcal{W}(a) \cup a\mathcal{W}(a^{-1}) \quad \text{or} \quad F_2 = \mathcal{W}(b) \cup b\mathcal{W}(b^{-1})$$

piece $a\mathcal{W}(a^{-1})$ contains all other possibilities because of cancellation of aa^{-1} , same for $b\mathcal{W}(b^{-1})$

- realize F_2 as a subgroup of rigid motions of the sphere
 - choose two different axes of rotation e_1, e_2 (straight lines through the center of the sphere)
 - choose two angles θ_1, θ_2 that are not rationally related (θ_i/π lin independent over \mathbb{Q})
 - group generated by $a = R_{e_1, \theta_1}$ and $b = R_{e_2, \theta_2}$



- choose a point P on the sphere not on the axes e_1, e_2
- take orbits of P under transformations in each set $\mathcal{W}(a), \mathcal{W}(b), \mathcal{W}(a^{-1}), \mathcal{W}(b^{-1})$
- these orbits do not cover the whole sphere S^2
- keep choosing new points outside of the orbits already obtained and take their orbits under the same sets of transformations until the whole S^2 is covered
- Note: this requires infinitely many choices (axiom of choice is implicit here)
- Then choose a point x_α from each of the orbits constructed (Note: again infinitely many choices)
- obtain 4 disjoint sets covering the sphere S^2

$$A = \cup_{\alpha} \mathcal{W}(a)x_{\alpha}, \quad B = \cup_{\alpha} \mathcal{W}(b)x_{\alpha},$$

$$C = \cup_{\alpha} \mathcal{W}(a^{-1})x_{\alpha}, \quad D = \cup_{\alpha} \mathcal{W}(b^{-1})x_{\alpha}$$

- **duplication of the sphere** (Banach-Tarski paradox)
 - break S^2 into the pieces A, B, C, D
 - transform $C = \cup_{\alpha} \mathcal{W}(a^{-1})x_{\alpha}$ by the rigid motion $a = R_{e_1, \theta_1}$ into $C' = \cup_{\alpha} a\mathcal{W}(a^{-1})x_{\alpha}$ and transform $D = \cup_{\alpha} \mathcal{W}(b^{-1})x_{\alpha}$ by rigid motion $b = R_{e_2, \theta_2}$ into $D' = \cup_{\alpha} b\mathcal{W}(b^{-1})x_{\alpha}$
 - by $F_2 = \mathcal{W}(a) \cup a\mathcal{W}(a^{-1}) = \mathcal{W}(b) \cup b\mathcal{W}(b^{-1})$ and $\cup_{\alpha} F_2 x_{\alpha} = S^2$ obtain two copies of S^2 by reassembling these pieces as

$$S^2 = A \cup C' = \cup_{\alpha} \mathcal{W}(a)x_{\alpha} \cup \cup_{\alpha} a\mathcal{W}(a^{-1})x_{\alpha}$$

$$S^2 = B \cup D' = \cup_{\alpha} \mathcal{W}(b)x_{\alpha} \cup \cup_{\alpha} b\mathcal{W}(b^{-1})x_{\alpha}$$

The decomposition into sets A, B, C, D necessarily involves non-measurable sets!

Hausdorff measures $\mu_{H,s}$

- ambient space $X = \mathbb{R}^N$ some fixed N with Euclidean distance (or more generally a metric space (X, d))
- Hausdorff outer measure for a given $s \in \mathbb{R}_+^*$

$$\mu_{H,s}^*(A) := \liminf_{\epsilon \rightarrow 0} \inf_{\mathcal{U}} \sum_{\alpha} (\text{diam } U_{\alpha})^s$$

where $\mathcal{U} = \{U_{\alpha}\}$ an open covering of A with diameters at most ϵ

- Hausdorff measure $\mu_{H,s}$ obtained from the Hausdorff outer measure via the Carathéodory construction

Behavior of the Hausdorff measures

• for fixed $\epsilon > 0$ define $\mu_{H,\epsilon,s}(E) := \inf_{\mathcal{U}} \sum_{\alpha} (\text{diam } U_{\alpha})^s$ for $E \subset \mathbb{R}^N$, infimum over all countable coverings $\mathcal{U} = \{U_{\alpha}\}$ with $0 < \text{diam } U_{\alpha} \leq \epsilon$

- For a given $E \subset \mathbb{R}^N$ in the fixed ambient space, the function $s \mapsto \mu_{H,s}(E)$ is non-increasing as s grows from 0 to ∞
- If $s < t$ then $\mu_{H,\epsilon,s}(E) \geq \epsilon^{s-t} \mu_{H,\epsilon,t}(E)$
- this implies that if for E at some t have $\mu_{H,t}(E) > 0$ then for all $s < t$ have $\mu_{H,s}(E) = \infty$
- also if for some s have $\mu_{H,s}(E) < \infty$ then for all $t > s$ have $\mu_{H,t}(E) = 0$
- on a fixed set E values of Hausdorff measure as a function of s starts at ∞ , somewhere jumps to 0 then remains 0; at the unique s where it jumps can have an intermediate value $0 \leq \mu_{H,s}(E) \leq \infty$

Hausdorff dimension rigorous definition

- the Hausdorff dimension $\dim_H(E)$ is the value of s where the Hausdorff measure $\mu_{H,s}(E)$ jumps from ∞ to 0

$$\mu_{H,s}(E) = \begin{cases} \infty & s < \dim_H(E) \\ 0 & s > \dim_H(E) \end{cases}$$

- s -sets and Hausdorff measure**

- s -set $E \subset \mathbb{R}^N$: if the value of the Hausdorff measure at $s = \dim_H(E)$ is neither zero nor infinity

$$0 < \mu_{H,\dim_H(E)}(E) < \infty$$

especially “well behaved” sets with respect to the Hausdorff measure: the measure $\mu_{H,s}$ defines a good s -dimensional volume on E

properties of Hausdorff measures and dimension

- affine transformations and the Hausdorff measure

- $E \subset \mathbb{R}^N$, given any $v \in \mathbb{R}^N$, $\lambda \in \mathbb{R}_+^*$ and $s \in \mathbb{R}_+$

$$\mu_{H,s}(E + v) = \mu_{H,s}(E) \quad \text{and} \quad \mu_{H,s}(\lambda E) = \lambda^s \mu_{H,s}(E)$$

- these properties follow directly from the definition of the Hausdorff measure
- it follows (Haar measure for the group of translations) that $\mu_{H,n}$ on \mathbb{R}^n agrees (up to a constant scaling factor $C_n > 0$) with the Lebesgue measure of dimension n

- properties of the Hausdorff dimension

- *monotonicity*: for $A \leq B$ have $\dim_H(A) \leq \dim_H(B)$
- *countably stability*: $\dim_H(A) = \sup_k \dim_H(A_k)$ for $A = \cup_k A_k$ inside ambient \mathbb{R}^N
- *vanishing*: $\dim_H(A) = 0$ for A countable set
- *smooth manifolds*: $\dim_H(M) = m$ if $M \subset \mathbb{R}^N$ is a smooth manifold of dimension $m \leq N$

Lipschitz maps and Hausdorff measures/dimension

- Lipschitz map $f : (X, d_X) \rightarrow (Y, d_Y)$

$$d_Y(f(x), f(y)) \leq C d_X(x, y), \quad \forall x, y \in X$$

for some $C > 0$; contraction if $C < 1$

- homeomorphism $f : X \rightarrow Y$ bi-Lipschitz if

$$C^{-1} d_X(x, y) \leq d_Y(f(x), f(y)) \leq C d_X(x, y) \quad \forall x, y \in X$$

- Hausdorff measure and dimension

- Lipschitz maps satisfy, \forall measurable A and $s \in \mathbb{R}_+$

$$\mu_{H,s}(f(A)) \leq C^s \mu_{H,s}(A)$$

$$\dim_H(f(A)) \leq \dim_H(A)$$

- bi-Lipschitz maps satisfy, \forall measurable A and $s \in \mathbb{R}_+$

$$\dim_H(f(A)) = \dim_H(A)$$

- if $d_Y(f(x), f(y)) = C d_X(x, y)$ then

$$\mu_{H,s}(f(A)) = C^s \mu_{H,s}(A)$$

Hausdorff measure and Hausdorff dimension of Cantor sets

- $s = \dim_H C = \frac{\log 2}{\log 3}$ and $\mu_{H,s}(C) = 1$
 - first show $\mu_{H, \frac{\log 2}{\log 3}}(C) \leq 1$ hence $\dim_H C \leq \frac{\log 2}{\log 3}$
 - for this enough to cover C with 2^j intervals of length 3^{-j} as in the iterative construction of C , then $\mu_{H,s}(C) \leq 2^j 3^{-sj} = 1$ for $s = \frac{\log 2}{\log 3}$
 - then show opposite estimate $\mu_{H, \frac{\log 2}{\log 3}}(C) \geq 1$ so that $\dim_H C \geq \frac{\log 2}{\log 3}$
 - for this show that for any collection $\mathcal{J} = \{J_\alpha\}$ of intervals covering C have $\sum_\alpha \ell(J_\alpha)^s \geq 1$ for $s = \frac{\log 2}{\log 3}$

- by compactness can use finite coverings, can also take smallest intervals that cover a pair of ternary intervals of the construction of C (J contains one such interval I_1 , followed by an interval K in the complement of C followed by another such interval I_2)
- by construction of C these have $\ell(I_1), \ell(I_2) \leq \ell(K)$
- concavity of $x \mapsto x^s$ and $s = \frac{\log 2}{\log 3}$ give

$$\begin{aligned} \ell(J)^s &= (\ell(I_1) + \ell(K) + \ell(I_2))^s \geq \left(\frac{3}{2}(\ell(I_1) + \ell(I_2))\right)^s \\ &= 2\left(\frac{1}{2}\ell(I_1)^s + \frac{1}{2}\ell(I_2)^s\right) \geq \ell(I_1)^s + \ell(I_2)^s \end{aligned}$$

- so replacing J with two subintervals I_j of the construction of C does not increase the sum of s -powers of lengths
- keep doing this operation until reduce to a covering of C by intervals of equal length 3^{-j} that include all the 2^j intervals in the construction of C so that $\sum_j \ell(J)^s \geq 1$ for $s = \frac{\log 2}{\log 3}$ for this covering (hence for the arbitrary initial one as well)

Hausdorff measures and self-similarity

- **affine self-similar fractals** $K = f_1(K) \cup \dots \cup f_N(K)$ fixed point with f_i contractions given by **affine maps**

- if K is an s -set: $0 < \mu_{H, \dim_H(K)}(K) < \infty$ then

$$\begin{aligned}\mu_{H, \dim_H(K)}(K) &= \mu_{H, \dim_H(K)}(\cup_i f_i(K)) \\ &\leq \sum_i \mu_{H, \dim_H(K)}(f_i(K)) = \sum_i \lambda_i^{\dim_H(K)} \mu_{H, \dim_H(K)}(K)\end{aligned}$$

- self-similarity dimension $\dim_{\text{self-sim}}(K)$ is unique $s > 0$ such that $1 = \sum_{i=1}^N \lambda_i^s$
- $j(s) = \sum_{i=1}^N \lambda_i^s$ decreasing function with $j(0) = N > 1$ and $j(s) \rightarrow 0$ for $s \rightarrow \infty$
- so above inequality gives $\sum_i \lambda_i^{\dim_H(K)} \geq 1$
- obtain **upper bound** on Hausdorff dimension:

$$\dim_{\text{self-sim}}(K) \geq \dim_H(K)$$

- **not sharp** in general due to overlaps between the images $f_i(K)$

Moran condition

Theorem (P. Moran, 1945)

Suppose that $A \subset \mathbb{R}^d$ is a compact attractor of an IFS $\mathcal{F} = \{f_1, \dots, f_k\}$ of similarity transformations with $0 < \lambda_j < 1$. Assume that either $f_j(A)$ are disjoint for $j = 1, \dots, k$ or that A obtained in the following way: Suppose Ω_1 is an open bounded set and $\Omega_2^j = f_j(\Omega_1)$ be disjoint open sets for $j = 1, \dots, k$ contained in Ω_1 . Similarly let $\Omega_3^{j\ell} = f_\ell(\Omega_2^j)$ for $\ell = 1, \dots, k$ be disjoint in all j and so on. Suppose A is the intersection of

$$\overline{\Omega_1}, \quad \overline{\cup_j \Omega_2^j}, \quad \overline{\cup_{j\ell} \Omega_3^{j\ell}}, \quad \dots$$

Then $\dim_H(A)$ is the similarity dimension, namely, the unique $s > 0$ solving

$$1 = \lambda_1^s + \dots + \lambda_k^s.$$

For many affine self-similar examples this condition holds and the Hausdorff dimension can be computed directly as the self-similarity dimension

Vitali covering theorem

- given a sufficiently large collection of sets that covers E it is possible to select a subcollection of *disjoint* sets that *almost* covers E in a measure sense
 - *semidisjoint* collection $\mathcal{C} = \{C_\alpha\}$ of sets if no C_α contained inside another C_β
 - collection $\mathcal{C} = \{C_\alpha\}$ *Vitali class* for E if for all $x \in E$ and all $\delta > 0$ there is a $C_\alpha \in \mathcal{C}$ such that $x \in C_\alpha$ and $0 < \text{diam}(C_\alpha) \leq \delta$
- **Vitali covering result:** $E \subset \mathbb{R}^N$ Hausdorff $\mu_{H,s}$ -measurable and \mathcal{C} Vitali class for E of closed sets: possible to select finite or countable subcollection $\mathcal{C}' = \{C_k\} \subset \mathcal{C}$ of disjoint sets such that either $\sum_k \text{diam}(C_k)^s = \infty$ or $\mu_{H,s}(E \setminus \cup_k C_k) = 0$;
if $\mu_{H,s}(E) < \infty$ can also obtain for a chosen $\epsilon > 0$ that $\mu_{H,s}(E) \leq \sum_k \text{diam}(C_k)^s + \epsilon$

Probability measures on Bernoulli shifts

- **alphabet** \mathfrak{A} finite set with $\#\mathfrak{A} = k$, **shift space**

$$\Sigma_k^+ := \{a_1 a_2 \cdots a_n \cdots \mid a_k \in \mathfrak{A}, \forall k \in \mathbb{N}\}$$

infinite sequences of letters in the alphabet \mathfrak{A}

- **cylinder sets** $\mathcal{C}_{w_1, \dots, w_n} \in \mathcal{C}$

$$\mathcal{C}_{w_1, \dots, w_n} := \{a = a_1 a_2 \cdots a_n \cdots \in \Sigma_k^+ \mid a_k = w_k, k = 1, \dots, n\}$$

- choose a probability distribution on the finite set \mathfrak{A} , that is, a point $P = (p_a)_{a \in \mathfrak{A}}$ in the k -simplex: $p_a \geq 0$ and $\sum_{a \in \mathfrak{A}} p_a = 1$
- use cylinder sets as sample class for the Carathéodory construction by assigning probability

$$\mathbb{P}(\mathcal{C}_{w_1, \dots, w_n}) := p_{w_1} \cdots p_{w_n}$$

- binary case $\mathfrak{A} = \{0, 1\}$ and shift space Σ_2^+ gives $P = (p, 1 - p)$ and

$$\mathbb{P}(C_{w_1, \dots, w_n}) = p^{a_n(w)}(1 - p)^{b_n(w)}$$

with $a_n(w) = \#\{w_i = 0\}$ and $b_n(w) = \#\{w_i = 1\}$ in the string $w = w_1 \dots w_n$

- \mathbb{P} is additive on disjoint unions of cylinders

$$C_{w_1, \dots, w_n} = C_{w_1, \dots, w_n, 1} \sqcup \dots \sqcup C_{w_1, \dots, w_n, k}$$

decomposition of an n -cylinder into a disjoint union of $(n + 1)$ -cylinders and

$$\begin{aligned} \mathbb{P}(C_{w_1, \dots, w_n}) &= p_{w_1} \cdots p_{w_n} = p_{w_1} \cdots p_{w_n} \sum_{a \in \mathfrak{A}} p_a = \\ &= \sum_{a \in \mathfrak{A}} \mathbb{P}(C_{w_1, \dots, w_n, a}) \end{aligned}$$

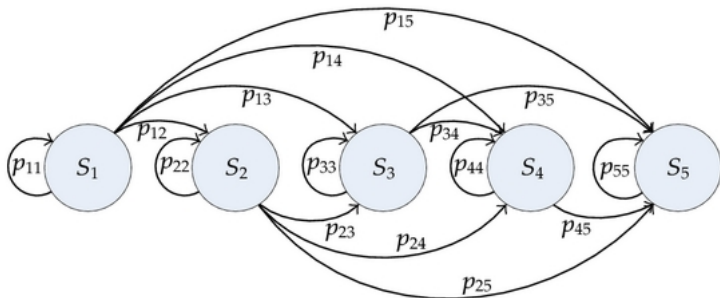
- Bernoulli probability measure on Σ_k^+ obtained by applying the Carathéodory construction (outer measure, σ -algebra, and measure) to the assignment of cylinder sets \mathcal{C} with the probabilities $\mathbb{P}(\mathcal{C}_{w_1, \dots, w_n}) := p_{w_1} \cdots p_{w_n}$
- Bernoulli measures on the Cantor set C using the coding of the Cantor set by Σ_2^+ with coding map $h : \Sigma_2^+ \rightarrow C$, define Bernoulli measure on C by $\mu_P(A) = \mathbb{P}(h^{-1}(A))$ for assigned $P = (p, 1 - p)$.
- Hausdorff measure $\mu_{H,s}$ special case of Bernoulli measure where $p_i = \lambda_i^s$ contraction ratios $\sum_i \lambda_i^s = 1$ for $s =$ self-similarity dimension

Markov measures

- probabilities with memory of past events (one step), unlike Bernoulli measures (no memory)
- shift space Σ_k^+ (one-side) infinite words on k letter
- data of a Markov chain
 - Root probabilities: of letters in the alphabet $\pi = (\pi_1, \dots, \pi_k)$ with $\pi_i \geq 0$ and $\sum_{i=1}^k \pi_i = 1$
 - Stochastic matrix of transitions: $P = (p_{ij})_{i,j=1,\dots,k}$ with $p_{ij} \geq 0$ and $\sum_{j=1}^k p_{ij} = 1$
 - Stationary vector: $\pi P = \pi$ left Perron–Frobenius eigenvector of P

$$\sum_{i=1}^k \pi_i p_{ij} = \pi_j$$

Markov chain



$$P = \begin{bmatrix} p_{11} & p_{12} & p_{13} & p_{14} & p_{15} \\ 0 & p_{22} & p_{23} & p_{24} & p_{25} \\ 0 & 0 & p_{33} & p_{34} & p_{35} \\ 0 & 0 & 0 & p_{44} & p_{45} \end{bmatrix} \text{ where } p_{ij} \geq 0 \text{ and } \sum_j p_{ij} = 1$$

Markov measure on the shift space Σ_k^+

- outer measure by assigning measure to *cylinder sets* $\mathcal{C}_{w_1, \dots, w_n}$

$$\mu_{P, \pi}(\mathcal{C}_{w_1, \dots, w_n}) := \pi_{w_1} p_{w_1 w_2} p_{w_2 w_3} \cdots p_{w_{n-1} w_n}$$

- seen as probability of starting in state w_1 and ending in state w_n with an oriented path in the graph of the Markov chain passing through intermediate vertices w_2, \dots, w_{n-1}
- additivity on cylinders* because of stochastic matrix property $\sum_{j=1}^k p_{ij} = 1$ for all $i = 1, \dots, k$
- use Carathéodory construction to obtain Σ -algebra and Markov measure
- Shift invariant** measure because $\pi P = \pi$

$$\mu_{P, \pi}(\sigma^{-1}(A)) = \mu_{P, \pi}(A)$$

$\sigma : \Sigma_k^+ \rightarrow \Sigma_k^+$ non-invertible one-sided shift

$$w = w_1 w_2 \cdots w_N \cdots \mapsto \sigma(w) = w_2 w_3 \cdots w_{N+1} \cdots$$

- Σ_k^+ parameterizes all possible infinite directed paths in the graph of the Markov chain
- the measure $\mu_{P,\pi}(\mathcal{C}_{w_1,\dots,w_n})$ gives the probability that a random path starts with the steps w_1, \dots, w_n
- if an entry $p_{ij} = 0$ can remove the corresponding edge in the graph: probability zero to all paths where the letter j follows the letter i (reaching state j from state i)
- Bernoulli measures: if $p_i > 0$ for $i = 1, \dots, k$ have $\text{supp}(\mu_P) = \Sigma_k^+$
- Markov measures with all $\pi_i > 0$ and all $p_{ij} > 0$ have $\text{supp}(\mu_{P,\pi}) = \Sigma_k^+$

- Markov measures with all $\pi_i > 0$ and all $p_{ij} > 0$ have $\text{supp}(\mu_{P,\pi}) = \Sigma_k^+$
- otherwise **transition matrix**

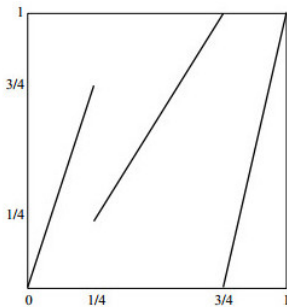
$$A_{ij} = \begin{cases} 0 & p_{ij} = 0 \\ 1 & p_{ij} > 0 \end{cases}$$

- transition matrix only measures if an edge is “accessible” to paths or not (without quantifying the probability)
- **primitive** A if $\exists n \in \mathbb{N}$ such that A^n positive $A^n_{ij} > 0$ all i, j
- primitive A implies given any two states i, j there is an oriented path in the Markov chain graph of length at least n given by an admissible word that goes from i to j
- **admissible sequences** $\Sigma_A^+ \subset \Sigma_k^+$

$$\Sigma_A^+ = \{w = w_1 w_2 \cdots w_N \cdots \in \Sigma_k^+ \mid A_{w_i w_{i+1}} = 1, \forall i \geq 1\}$$

- support of Markov measure $\text{supp}(\mu_{P,\pi}) = \Sigma_A^+$

Dynamical systems: one-dimensional Markov maps



- collection I_1, \dots, I_k of disjoint subintervals of $[0, 1]$ and map $f : I_1 \cup \dots \cup I_k \rightarrow [0, 1]$ such that $f_i = f|_{I_i}$ is homeomorphism to image $f(I_i)$
- each image $f(I_i)$ contains each interval I_j it intersects: if $I_j \cap f(I_i) \neq \emptyset$ then $I_j \subseteq f(I_i)$

Cantor sets and coding of one-dimensional Markov maps

- **Transition matrix** for one-dimensional Markov maps

$$A_{ij} = \begin{cases} 1 & f(I_i) \cap I_j \neq \emptyset \\ 0 & f(I_i) \cap I_j = \emptyset \end{cases}$$

- **Cantor set** of one-dimensional Markov map
 $f : I_1 \cup \dots \cup I_k \rightarrow [0, 1]$

$$C_A = \{x \in I_1 \cup \dots \cup I_k \mid \text{all iterates } f^{\circ n} \text{ are defined at } x\}$$

- dynamical system $f : C_A \rightarrow C_A$
- **coding map** for $f : C_A \rightarrow C_A$

$$\sigma : \Sigma_A^+ \rightarrow \Sigma_A^+ \text{ one-sided shift map}$$

- **subshifts of finite type** $\sigma : \Sigma_A^+ \rightarrow \Sigma_A^+$ models of dynamics for many different kinds of maps and fractals



Hiroshi Kawano, *Design 3-1: Color Markov Chain Pattern*, 1964