Topological entropy and pressure (Thermodynamic formalism)

We can consider the set of all values:

$h_\mu(T)$: μ ranges over all T-invariant probability measures

Note: In some examples (eg continued fraction transformations) this is not bounded above

However, when it is bounded above we can define:

Definition: $h_{top}(T) = \sup\{h_\mu(T) \mid \mu = T\text{-invariant prob.}\}$

is the topological entropy

Remark: There are other definitions of $h_{top}(T)$ and the equivalence to those is one form of the "variational principle".

Example: Let $\begin{cases} \Sigma = \{1, \dots, k\}^{\mathbb{Z}} \\ \sigma : \Sigma \to \Sigma, \ (\sigma x)_n = x_{n+1} \end{cases}$

be the (full) shift on K-symbols

If μ_p is the Bernoulli measure associated to the probability vector $p = (p_1, \dots, p_k)$

then $h_{\mu_p}(T) = -\sum_{i=1}^{k} p_i \log p_i$

Lemma: $h_{\mu_p}(T)$ is maximized over all T-invariant probability measures by $p = (1/k, \dots, 1/k)$.

Proof. This is easy (for example using Lagrange multipliers for

$$(p_1, \dots, p_k) \longmapsto \sum_{i=1}^{k} f(p_i)$$

(subject to $(p_1, \dots, p_k) = (1/k, \dots, 1/k)$). $\square$

However, we also have that:

Lemma $h_\mu(T)$ is maximized over all T-invariant probability measures by μ_p with $p = (1/k, \dots, 1/k)$

Proof: Recall that for any T-invariant probability measure we have that if $\alpha = \{[1], \dots, [k]\}$

$$h_\mu(T) = h_\mu(T, \alpha) = \inf_n \left\{ \frac{H_\mu(\alpha^{(n)})}{n} \right\} \leq H_\mu(\alpha).$$

If we let $p = (p_1 = \mu[1], \dots, p_k = \mu[k])$ then $h_\mu(T) \leq h_{\mu_p}(T)$, ie, every entropy is dominated by that of a Bernoulli measure. Thus, $h(\mu)$ is maximized at μ_p with $p = (1/k, \dots, 1/k)$

(Remark: If $\sigma: \Sigma_A \to \Sigma_A$ is a subshift of finite type then $h_\mu(T)$ is maximized at a Markov measure.

~~Example~~ If A = $k\times k$ matrices with entries 0-1.

Let $\Sigma = \left\{ x = (x_n)_{n=-\infty}^{+\infty} \in \{1,\dots,k\}^{\mathbb{Z}} \;\middle|\; \begin{matrix} A(x_n, x_{n+1}) = 1, \\ \forall n \in \mathbb{Z} \end{matrix} \right\}$

Let $\sigma : \Sigma \to \Sigma$ by $(\sigma x)_n = x_{n+1}$.

- Assume A is aperiodic (ie, $\exists N \geq 1$ such that $A^N > 0$)

By the Perron-Frobenius Theorem:

- $Av = \lambda v$ where $\begin{cases} \lambda > 0 \\ v = (v_1, \dots, v_k) \text{ with } v_i > 0 \end{cases}$

- We define a stochastic matrix by

$$P_{ij} = \frac{A(i,j)}{\lambda} \frac{v_j}{v_i}$$

- Let $p = (p_1, \dots, p_k)$ be the left eigenvector $pP = p$.

The associated Markov measure (Parry measure): $\mu([x_0, \dots, x_{n-1}]) = p_{x_0} P(x_0, x_1) \dots P(x_{n-2}, x_{n-1})$ is the (unique) measure maximizing the entropy (as $\log \lambda = h_{top}(\sigma)$)

<u>Pressure</u>: Let $T : X \to X$ be a continuous map on a compact metric space. Let $f : X \to \mathbb{R}$ be continuous.

We define the <u>pressure</u> of f:

$$P(f) = \sup \left\{ h_\mu(T) + \int f \, d\mu \;\middle|\; \mu = T\text{-invariant} \right\}$$

Example: Let $\sigma : \Sigma \to \Sigma$ be the (full) shift on $\Sigma = \prod_{0}^{\infty} \{1, \ldots, k\}$ i.e., $(\sigma x)_n = x_{n+1}$.

Let $f(x) = f(x_0, x_1)$ where $x = (x_n)_{n=0}^{\infty}$

Lemma $P(f) = h_{\mu_p}(\sigma) + \int f d\mu_p$

where μ_p is a Markov measure.

We need a simple result:

Claim: If $\underline{p} = (p_1, \ldots, p_k)$ and $\underline{q} = (q_1, \ldots, q_k)$ are probability vectors then

$$-\sum_{i=1}^{k} q_i \log q_i + \sum_{i=1}^{k} q_i \log p_i \leq 0$$

with equality iff $\underline{p} = \underline{q}$.

Proof of claim. We can rewrite the expression:

$$-\sum_{i=1}^{k} p_i \left(\frac{q_i}{p_i}\right) \cdot \log\left(q_i/p_i\right) = \sum_{i=1}^{k} p_i f(q_i/p_i)$$

(where $f(x) = -x \log x$)

$$\leq f\Big(\underbrace{\sum_i p_i (q_i/p_i)}_{=1}\Big) \quad \text{(By } f \text{ concave)}$$

$$= 0$$

Proof of ~~Lemma~~: Consider the ~~k~~ matrix:

$$Q(i,j) = \left(e^{f(i,j)}\right)_{i,j=1}^{k}$$

By the Perron-Frobenius theorem

- $Q w = \lambda w$ where $\begin{cases} \lambda > 0 \\ w = (w_1, \dots, w_k) \text{ with } w_i > 0 \end{cases}$

- We can define a stochastic matrix by

$$\overline{P}(i,j) = \frac{Q(i,j)}{\lambda} \frac{w_j}{w_i}$$

- Let $\overline{p} = (\overline{p}_1, \dots, \overline{p}_k)$ be the left eigenvector of $\overline{P}$ i.e. $\overline{p}\,\overline{P} = \overline{p}$.

Consider an arbitrary Markov measure μ_P (for P). Fix j and then we can apply the claim with $\begin{cases} p_i \to p_i P(i,j)/p_j \\ q_i \to \overline{p}_i \overline{P}(i,j)/\overline{p}_j \end{cases}$

To deduce that

$$h_{\mu_P}(\sigma) + \int \overline{f} \, d\mu_P \leq \log \lambda \; (= P(\overline{f}))$$

with equality iff $P(i,j) = \overline{P}(i,j)$.

$$\begin{cases} \overline{f}(x_0, x_1) = f(x_0, x_1) + \log w(x_1) - \log w(x_0) \\ \Rightarrow \int \overline{f} \, d\mu_P = \int f \, d\mu_P \end{cases}$$

More generally, we define a metric on Σ (or even Σ_A) by:

$$d(x,y) = \sum_{n=-\infty}^{+\infty} \frac{e(x_n, y_n)}{2^{|n|}}$$

where $x = (x_n)_{n=-\infty}^{\infty}$, $y = (y_n)_{n=-\infty}^{\infty}$ and

$$e(i,j) = \begin{cases} 1 & \text{if } i \neq j \\ 0 & \text{if } i = j \end{cases}$$

<u>Lemma</u>: If $f : \Sigma \to \mathbb{R}$ is Hölder continuous (ie, $\exists\, \alpha > 0$ and $C > 0$ such that

$$|f(x) - f(y)| \leq C\, d(x,y)^{\alpha}\,)$$

Then there is a unique σ-invariant probability measure realizing the supremum:

$$P(f) = \sup\left\{ h_m(\sigma) + \int f\, dm \;\middle|\; m = \sigma\text{-invariant probability} \right\}$$

$$\left(= h_\mu(\sigma) + \int f\, d\mu \right).$$

<u>Def^n</u> μ is called the equilibrium state for f

<u>Remark</u>: If we let $C^\alpha(\Sigma)$ be the Banach space of Hölder continuous functions then (with norm $\|\cdot\|_\alpha$) we can replace the matrix Q by an operator

$$\begin{cases} \mathcal{L}_f : C^\alpha(\Sigma) \longrightarrow C^\alpha(\Sigma) \\ \mathcal{L}_f h(x) = \sum_{\sigma y = x} e^{f(y)} h(y) \end{cases}$$

which has maximal eigenvalue $e^{P(f)}$.

Three applications.

1) Exponentially fast mixing

Let μ be the equilibrium state for $f \in C^\alpha(\Sigma)$.
$\exists C > 0$, $0 < \Theta < 1$ such that for $h_1, h_2 \in C^\alpha(\Sigma)$
we have:

$$\left|\int h_1 \circ \sigma^n \cdot h_2 \, d\mu - \int h_1 \, d\mu \int h_2 \, d\mu\right| \leq C \|h_1\|_\infty \cdot \|h_2\| \times \Theta^n$$

(Idea: we replace f by $\bar{f}$ for which μ is still an equilibrium state and $\mathcal{L}_{\bar{f}} 1 = 1$

Then $\mathcal{L}_{\bar{f}}^* \mu = \mu$ and
$$\int h_1 \circ \sigma^n \, h_2 \, d\mu = \int \mathcal{L}_{\bar{f}}^n h_2 \cdot h_1 \, d\mu = \int (\underbrace{\mu(h_2) + O(\Theta^n \|h_2\|)}_{\text{spectral gap}}) h_1 \, d\mu$$

2) Central Limit theorem

Let μ be the equilibrium state for $f \in C^\alpha(\Sigma)$.
Then for $g \in C^\alpha(\Sigma)$ (with $\int g \, d\mu = 0$), $\exists \sigma > 0$ such that

$$\lim_{n \to \infty} \mu\left(\left\{x \,\middle|\, \frac{1}{\sqrt{n}} \sum_{k=0}^{n-1} g(\sigma^k x) \in [a,b]\right\}\right) = \frac{1}{\sqrt{2\pi}\,\sigma} \int_a^b e^{-y^2/2\sigma^2} \, dy$$

(Idea: we want to use Fourier transforms to show: $\psi_n(t) = \int \exp\left(\frac{it}{\sigma\sqrt{n}} \sum_{j=0}^{n-1} f(\sigma^j x)\right) d\mu(x) \longrightarrow e^{-t^2/2}$

In fact: $\psi_n(t) = \int \mathcal{L}^n_{\left(\bar{f}+\frac{it}{\sigma\sqrt{n}}\sum_{j=0}^{n-1} f(\sigma^j x)\right)} \mathbb{1}\, d\mu(x)$.

$$\simeq \exp\left(n P\left(\bar{f}+\frac{it}{\sigma\sqrt{n}}\sum_{j=0}^{n-1} f(\sigma^j x)\right)\right)$$

$$\simeq e^{-t^2/2} \qquad \left(\begin{array}{l}\text{where } P(\bar{f})=0\\ \quad P''=\sigma^2\end{array}\right)$$

3) Hausdorff dimension

Let $T:\mathbb{C}\to\mathbb{C}$ be the polynomial map $Tz = z^2+\varepsilon$ (for $|\varepsilon|$ small).

The Julia set $J_\varepsilon = T(J_\varepsilon)$ is the limit set of the contractions $\begin{cases} T_0 : z \longmapsto +\sqrt{z-\varepsilon} \\ T_1 : z \longmapsto -\sqrt{z-\varepsilon}\end{cases}$

The dimension $d(\varepsilon) = \dim_H J_\varepsilon$ is characterized by

$$P(-d(\varepsilon)\log|T_\varepsilon' \circ \pi|) = 0$$

where $\begin{cases} \log|T_\varepsilon'| : J_\varepsilon \to \mathbb{R} \text{ is } C^\infty, \text{ and} \\ \pi : \underset{\{1,2\}}{\Sigma} \to J_\varepsilon \text{ is defined by} \\ \qquad \pi(x_n)_{n=0}^{+\infty} = \bigcap_{n=1}^{\infty} T_{x_0}T_{x_1}\dots T_{x_n} J_\varepsilon \end{cases}$

Thus $\log|T_\varepsilon'|\circ\pi : \Sigma \to \mathbb{R}$ is C^α.

This leads to:

<u>Theorem</u> (Ruelle). For ε_0 small: $\begin{cases} (-\varepsilon_0, \varepsilon_0) \longmapsto \mathbb{R} \text{ is } C^\omega \\ \varepsilon \longmapsto d(\varepsilon)\end{cases}$

(<u>Idea</u>: $\varepsilon \mapsto P(\cdot)$ is C^ω (pert^n thy) then apply Implicit f^n thm.