Thermodynamic formalism

Lecture by Amie Wilkinson
Notes by Clark Butler

November 18, 2014

Ruelle's Perron-Frobenius Theorem: Recall the statement: Let Σ_A^+ be a topologically mixing subshift of finite type and $\varphi : \Sigma_A^+ \to \mathbb{R}$ Hölder. Then there is $\lambda > 0$, $\nu \in M(\Sigma_A^+)$ a shift-invariant measure, and $h \in \Sigma_A^+$ positive such that

$$L_\varphi^*(\nu) = \lambda\nu$$
$$L_\varphi(h) = \lambda h$$
$$\nu(h) = 1$$

and for every $g \in \Sigma_A^+$,

$$\left\| \frac{1}{\lambda^m} L_\varphi^m(g) - \nu(g)h \right\| \to 0, \text{ as } m \to \infty$$

where $\|\cdot\|$ denotes the supremum norm on continuous functions.

Some new notation: $C_m(\omega) = \{\omega' \in \Sigma_A^+ : \ \omega_i = \omega_i', i = 0, \ldots, m\}$. This is $C[0; \omega_0, \ldots, \omega_m]$ in our old notation. Recall that we were working with the function space

$$\Lambda = \{u \in C(\Sigma_A^+) |\ u \geq 0,\ \nu(u) = 1, \text{ and if } \omega_i = \omega_i' \text{ for } i \in [0, m] \text{ then } u(\omega) \leq B_m u(\omega')\}$$

where $B_m = e^{\sum_{k=m}^\infty 2b\alpha^k}$, where $\mathrm{var}_k(\varphi) \leq b\alpha^k$. Let M be the smallest positive integer such that A^M has all positive entries. Recall $Q = \frac{1}{\lambda} L_\varphi$. We showed last time that $Q(\Lambda) \subset \Lambda$. We also showed that if $u \in \Lambda$ then $\inf Q^M(u) > \frac{1}{K}$, where $K = \lambda^M e^{M\|\varphi\|} B_0$. By the Schauder-Tychonoff theorem, there was an $h \in \Lambda$ such that $Q(h) = h > \frac{1}{K}$ since also $Q^M(h) = h$. Last time we also proved the inequality for $u \in \Lambda$, $\omega' \in C_m(\omega)$,

$$[L_\varphi u](\omega) \leq e^{2b\alpha^{m+1}} B_{m+1} [L_\varphi u](\omega')$$

It only remains to prove the convergence statement $\left\| \frac{1}{\lambda^m} L_\varphi^m(g) - \nu(g)h \right\| \to 0$ for $g \in C(\Sigma_A^+)$. We will prove this statement for $g = u \in \Lambda$. The general case can be found in Bowen's book; it proceeds by an approximation argument (all functions in $C(\Sigma_A^+)$ can be suitably approximated by functions in Λ).

The main step is to show that there is some $\eta \in (0, 1)$ such that for every $u \in \Lambda$,

$$Q^M(u) = \eta h + (1 - \eta) u',$$

for some $u' \in \Lambda$. This lemma can be iterated:

$$Q^{2M}(u) = \eta h + (1 - \eta) Q^M(u') = \eta h + (1 - \eta)\eta h + (1 - \eta)^2 u''$$

and in general

$$Q^{kM}(u) = C_k h + (1 - \eta)^k u_k$$

where $u_k \in \Lambda$ and $C_k = (1 - (1 - \eta)^k)$. This proves that

$$\|Q^{kM}(u) - h\| \leq (1 - \eta)^k (\|h\| + \|u_k\|) \leq (1 - \eta)^k (\|h\| + \|u_k\|) \leq (1 - \eta)^k (\|h\| + K)$$

Recall from last time we also proved an upper bound $\|u\| \leq K$ for $u \in \Lambda$. Thus $\|Q^{kM}(u) - h\| \to 0$ exponentially fast, which proves the claim since $\nu(u) = 1$. (Technically, we have only shown this along a subsequence, but it's easy to upgrade this to convergence along the full sequence using the estimates coming from $u \in \Lambda$).

We now prove the lemma: Fix $\eta \in (0,1)$, whose exact value we will determine later. Given $u \in \Lambda$, set

$$u' = \frac{Q^M(u) - \eta h}{1 - \eta}$$

We need to show that for η small enough, $u' \in \Lambda$. It's easy to check that $\nu(u') = 1$. If $\eta < \frac{1}{\|h\|K}$, then $u' \geq 0$, using the lower bound $Q^M(u) \geq \frac{1}{K}$. It suffices to show that for $v = (1-\eta)u'$, for every $m \geq 0$, and for every $\omega, \omega' \in C_m(\omega)$,

$$v(\omega) \leq B_m v(\omega')$$

We can rewrite this as

$$\eta \left(B_m h\left(\omega\right) - h\left(\omega'\right)\right) \leq B_m Q^M u\left(\omega\right) - Q^M u\left(\omega'\right) \quad (*)$$

Let $u_1 = Q^{M-1}\left(u\right)$. Using the inequality we proved to show that $Q(\Lambda) \subset \Lambda$,

$$L_\varphi u_1\left(\omega\right) \leq e^{b\alpha^{m+1}} B_{m+1} L_\varphi u_1\left(\omega'\right)$$

$$\Rightarrow Q^M u\left(\omega\right) \leq e^{b\alpha^{m+1}} B_{m+1} Q^M u\left(\omega'\right)$$

Now $h \in \Lambda$ implies that $h\left(\omega\right) \leq B_m h\left(\omega'\right)$, and so inequality $(*)$ can be proved by finding an η such that

$$\eta \leq \frac{B_m - e^{b\alpha^{m+1}} B_{m+1}}{K\left\|h\right\|\left(B_m - B_m^{-1}\right)}.$$

Set $c_m := e^{b\alpha^m}$. Then we want to choose η such that:

$$\eta \leq \frac{B_m - c_{m+1} B_{m+1}}{K\left\|h\right\|\left(B_m - B_m^{-1}\right)}.$$

Now:

$$\frac{B_m - c_{m+1} B_{m+1}}{K\left\|h\right\|\left(B_m - B_m^{-1}\right)} \leq \frac{B_m - c_m B_{m+1}}{K\left\|h\right\|\left(B_m - B_m^{-1}\right)}$$

$$= \frac{B_m\left(1 - c_m^{-1}\right)}{K\left\|h\right\|\left(B_m - B_m^{-1}\right)}$$

$$= \frac{\left(1 - c_m^{-1}\right)}{K\left\|h\right\|\left(1 - B_m^{-2}\right)}$$

But as $m \to \infty$,

$$\frac{\left(1 - c_m^{-1}\right)}{K\left(1 - B_m^{-2}\right)\left\|h\right\|} \to \frac{1}{K\left\|h\right\|},$$

Hence for m large enough, choosing $\eta \leq \frac{1}{2K\|h\|}$ is sufficient, and multiplying by a sufficiently small constant gives the inequality for all m.

An alternate proof of Ruelle's Perron-Frobenius can be found in an article of Pollicott titled *A complex Ruelle-Perron-Frobenius theorem and two counterexamples.*

Gibbs measures: We keep the same assumptions as those for Ruelle's Perron-Frobenius theorem. Last time we showed that if we set $d\mu_\varphi = hd\nu$, μ_φ is a σ-invariant measure. We now want to show this measure has some nice properties. First we have a detour on ergodicity and mixing.

Let $T : (X, \mu) \to (X, \mu)$ be a measure preserving transformation. T is ergodic if for every measurable B,

$$T^{-1}(B) = B \Rightarrow \mu(B) = 0 \text{ or } 1.$$

This means that the only T-invariant sets are trivial, i.e., have full measure or zero measure. T is mixing if for every B and C measurable,

$$\lim_{m\to\infty} \mu\left(T^{-m}(B) \cap C\right) = \mu(B)\,\mu(C).$$

It's easy to see that mixing implies ergodic: If B is T-invariant then taking $B = C$ in the above formula gives $\mu(B) = \mu(B)^2$, so that $\mu(B) = 0$ or 1. Mixing can be reinterpreted by defining

$$\mu\left(T^{-m}(B) \mid C\right) := \frac{\mu\left(T^{-m}(B) \cap C\right)}{\mu(C)}$$

to be the conditional measure of $T^{-m}(B)$ in C. This is the percentage of C that $T^{-m}(B)$ occupies. Mixing then implies that as $m \to \infty$, B and C become asymptotically independent; the percentage of $T^{-m}(B)$ in C is the same as the percentage of B in X.

Decay of Correlations: It's an easy exercise to show that T is ergodic if and only if for every $f \in L^2(X, \mu)$, $f \circ T = f$ implies that f is constant (where both of these statements should be taken μ-a.e.).

A similar alternative definition of mixing is the following: T is mixing if and only if for every $f, g \in L^2(X, \mu)$, as $m \to \infty$,

$$\int (f \circ T^m)(x)g(x)d\mu(x) \to \left(\int f(x)d\mu(x)\right)\left(\int g(x)d\mu(x)\right)$$

This is sometimes called a "decay of correlations" condition, in analogy with probability theory, thinking of $\{f \circ T^m\}_{m\in\mathbb{N}}$ as a sequence of random variables and $\int (f \circ T^m)(x)g(x)d\mu(x)$ as the correlations between these random variables and g. The mixing condition then implies that the random variables $\{f \circ T^m\}_{m\in\mathbb{N}}$ become asymptotically independent of any given function g.

Back to subshifts of finite type: We claim that σ acting on $(\Sigma_A^+, \mu_\varphi)$ is a mixing transformation (hence is also ergodic). It suffices to show that decay of correlations holds for a dense set in $L^2\left(\Sigma_A^+, \mu_\varphi\right)$; e.g. that for every $u, v \in C\left(\Sigma_A^+\right)$, we have

$$\lim_{m\to\infty} \int (u \circ \sigma^n)\, v d\mu \to \left(\int u d\mu_\varphi\right)\left(\int v d\mu_\varphi\right).$$

Let u, v be continuous. Then

$$\begin{aligned}
\mu(u \cdot (v \circ \sigma^m)) &= \nu(h \cdot u \cdot (v \circ \sigma^m)) \\
&= \frac{1}{\lambda^m} L_\varphi^{*m} \nu(h \cdot u \cdot (v \circ \sigma^m)) \\
&= \frac{1}{\lambda^m} \nu\left(L_\varphi^m (h \cdot u \cdot (v \circ \sigma^m))\right) \\
&= \frac{1}{\lambda^m} \nu\left(L_\varphi^m (h \cdot u) \cdot v\right) \\
&= \nu\left(\frac{1}{\lambda^m} L_\varphi^m (h \cdot u) \cdot v\right).
\end{aligned}$$

and therefore

$$\begin{aligned}
\mu\left(u\cdot\left(v\circ\sigma^{m}\right)\right)-\mu\left(u\right)\mu\left(v\right) &= \nu\left(\frac{1}{\lambda^{m}}L_{\varphi}^{m}\left(h\cdot u\right)\cdot v\right)-\nu\left(hu\right)\nu\left(hv\right)\\
&= \nu\left(\frac{1}{\lambda^{m}}L_{\varphi}^{m}\left(h\cdot u\right)\cdot v-\nu\left(hu\right)h\cdot v\right)\\
&= \nu\left(\left(\frac{1}{\lambda^{m}}L_{\varphi}^{m}\left(h\cdot u\right)-\nu\left(hu\right)h\right)\cdot v\right)
\end{aligned}$$

which goes to 0 uniformly as $m\to\infty$ by the Ruelle-Perron-Frobenius theorem.