Math 254A Lecture 28 Notes

Daniel Raban

June 2, 2021

1 Variational Principles for the Entropy Rate

1.1 Recap

Last time, we showed that

$$\begin{aligned} s(\mu) &:= \inf_{W,U \ni \mu} \lim_{B \uparrow \mathbb{Z}^d} \frac{1}{|B|} \log |\{x \in A^B : P_x^W \in U\}| \\ &= \begin{cases} h(\mu) := \lim_B \frac{1}{|B|} H(\mu_B) & \text{if } \mu \in P^T \\ -\infty & \text{otherwise.} \end{cases} \end{aligned}$$

Here, we extend h by $h(\mu) = -\infty$ if $\mu \notin P^T$. Then $h : M(A^{\mathbb{Z}^d}) \to [-\infty, \log|A|]$ is concave and upper semicontinuous, and the set $\{h > -\infty\} = \{h \geq 0\} = P^T$. The upper bound $\log|A|$ is achieved when $\mu = \mathrm{Unif}_A^{\times \mathbb{Z}^d}$.

Now, we will see two variational principles.

1.2 The first variational principle

Theorem 1.1. *Let $\psi : A^{\mathbb{Z}^d} \to \mathbb{R}^r$ depend only on coordinates in a finite $W \subseteq \mathbb{Z}^d$. For $x \in \mathbb{R}^r$, let*

$$s(\psi, y) = \inf_{V \ni x} \lim_{B \uparrow \mathbb{Z}^d} \frac{1}{|B|} \log |\{x \in A^B : \tfrac{1}{|B|} \Psi_B(x) \in V\}|,$$

where the inf is over open, convex neighborhoods of x in $\mathbb{R}^r$. Then

$$\begin{aligned} s(\psi, y) &= \sup\{h(\mu) : \mu \in P^T, \langle \psi, \mu \rangle = y\} \\ &= \sup\{h(\mu) : \mu \in P, \langle \psi, \mu \rangle = y\}. \end{aligned}$$

with the convention that $\sup \varnothing = -\infty$.

Proof.

$$\begin{aligned}\frac{1}{|B|}\Psi_B(x) &= \frac{1}{|B|}\sum_{v+W\subseteq B}\psi(T^v x)\\ &= \frac{1}{|\{v : v+W\subseteq B\}|}\sum_{v+W\subseteq B}\psi(T^v x) + o(|B|)\\ &= \langle\psi, P_x^W\rangle + o(|B|),\end{aligned}$$

This gives ($\geq$): For any $V \subseteq \mathbb{R}^r$ and finite $W \subseteq \mathbb{Z}^d$, we have

$$\frac{1}{|B|}\log|\{x \in A^B : \langle\psi, P_x^W\rangle \in V\}|.$$

The condition $\langle\psi, P_x^W\rangle \in V$ defines any convex neighborhood of any μ such that $\langle\psi,\mu\rangle = y$. So taking $\lim_B$ of the above, we get that it is $\geq h(\mu)$ for any such μ.

Now consider ($\leq$). Let

$$h = \sup\{h(\mu) : \mu \in P, \langle\psi,\mu\rangle = y\}.$$

The set $\{\mu \in P : \langle\psi,\mu\rangle = y\}$ is compact, so there exists a window W and open convex sets $U_1, \dots, U_r$ in $P(A^W)$ such that $\{\mu \in P : \langle\psi,\mu\rangle = y\} \subseteq \bigcup_i\{\mu \in P : \mu_W \in U_i\}$ and

$$\frac{1}{|B|}\log|\{x : P_x^W \in U_i\}| \leq (h+\varepsilon) + o(1)$$

for all i. Finally, by compactness again, if $V \subseteq \mathbb{R}^r$ is a small enough neighborhood of y, then

$$\bigcup_i\{\mu \in P : \mu_W \in U_i\} \supseteq \{\mu : \langle\psi,\mu\rangle \in V\}$$

So

$$\frac{1}{|B|}\log|\{x : \langle\psi, P_x^W\rangle \in V\}| \leq \max_i \frac{1}{|B|}\log|\{x : P_x^W \in U_i\}| + \frac{\log s}{|B|} \leq h + \varepsilon$$

as $B \uparrow \mathbb{Z}^d$. Since $\varepsilon > 0$ is arbitrary, we get $s(\psi, y) = h$, as desired. □

Corollary 1.1. *For any convex, open $V \subseteq \mathbb{R}^r$,*

$$\begin{aligned}s(\psi, V) &= \lim_B \frac{1}{|B|}\log|\{x : \langle\psi, P_x^W\rangle \in V\}|\\ &= \sup_{y\in V} s(\psi, y)\\ &= \sup\{h(\mu) : \mu \in P^T, \langle\psi,\mu\rangle \in V\}\\ &= \sup\{h(\mu) : \mu \in P, \langle\psi,\mu\rangle \in V\}.\end{aligned}$$

From this, we can return to interactions giving the total potential energy $\varphi = (\varphi_F)_F$, assumed (for simplicity) to be a finite range interaction. Look at

$$|\{x \in A^B : \tfrac{1}{|B|}\Phi(x) \in I\},$$

where I is a small open interval, and $\Phi_B(x) = \sum_{F \subseteq B} \varphi_F(x_F) = \sum_{F'} |B| \langle \varphi_{F'}, P_x^W \rangle + o(|B|)$. Here, W is a big enough window to see all nonzero translates, and F' runs over one copy of each finite set $\subseteq W$ up to translation. So this set is

$$\left| \left\{ x \in A^B : \sum_{F'} \langle \varphi_{F'}, P_x^W \rangle \in I \right\} \right|.$$

$\sum_{F'} \langle \varphi_{F'}, P_x^W \rangle \in I$ is an open, convex condition in $\mathbb{R}^r$, so

$$\frac{1}{|B|} \log \left| \left\{ x \in A^B : \sum_{F'} \langle \varphi_{F'}, P_x^W \rangle \in I \right\} \right| \xrightarrow{B \uparrow \mathbb{Z}^d} \sup \left\{ h(\mu) : \mu \in P^T, \sum_{F'} \langle \varphi_{F'}, \mu \rangle \in I \approx y \right\}.$$

We can use this result to predict the most likely values of any other observable if there is a unique measure μ that maximizes $h(\mu)$ subject to the constraint $\sum_{F'} \langle \varphi_{F'}, \mu \rangle = y$.

Remark 1.1. There always exists a μ achieving the supremum if the set $\{\mu : \sum_{F'} \langle \varphi_{F'}, \mu \rangle = y\} \neq \varnothing$ by upper semicontinuity of h on the above weak* compact set.

So the key question is when we get uniqueness of that maximizer. We will discuss this next time.

1.3 A variational principle for the Fenchel-Legendre transform of h

To understand the second variational principle, we need to extend the first version from $\varphi : A^{\mathbb{Z}^d} \to \mathbb{R}^r$ to any $\psi \in C(A^{\mathbb{Z}^d})$. To apply ψ "inside a box," given $x \in A^B$, let $\widehat{x}$ be any element of $A^{\mathbb{Z}^d}$ such that $\widehat{x}_B = x$. Given B and $\psi \in C(A^{\mathbb{Z}^d})$, let

$$s_B \psi(x) = \sum_{v \in B} \psi(T^v \widehat{x}).$$

Lemma 1.1. *If $\widehat{x}, \check{x}$ are two choices of extension, then*

$$\left| \sum_{v \in B} \psi(T^v \widehat{x}) - \sum_{v \in B} \psi(T^v \check{x}) \right| = o(|B|).$$

Now a fiddly extension of the first variational principle gives

$$\frac{1}{|B|} \log |\{x \in A^B : \tfrac{1}{|B|} s_B \psi(x) \in V\}| = \sup\{h(\mu) : \mu \in P^T, \langle \psi, \mu \rangle \in V\}.$$

This version is good because we can now handle the whole Banach space $C(A^{\mathbb{Z}^d})$, which is the dual of $M(A^{\mathbb{Z}^d})$, equipped with the weak* topology. This leads to a description of the Fenchel-Lengendre transform of h:

Theorem 1.2 (2nd variational principle)**.** *On* $C(A^{\mathbb{Z}^d})$,

$$\begin{aligned} h^*(f) &:= \sup\{h(\mu) - \langle f, \mu\rangle : \mu \in M(A^{\mathbb{Z}^d})\} \\ &= \lim_{B \uparrow \mathbb{Z}^d} \frac{1}{|B|} \log \sum_{x \in A^B} e^{-s_B f(x)}. \end{aligned}$$

The $e^{-s_B f(x)}$ are the Gibbs weights that define the canonical distribution on A^B. In ergodic theory and much of mathematical physics, this limit is called the *pressure* of f (denoted $p(f)$). Caution: this is not always the physical pressure of the system.