Large deviations and non-uniform specification properties

Vaughn Climenhaga

University of Houston

April 14, 2013

Joint work with Daniel J. Thompson (Ohio State)
and Kenichiro Yamamoto (Tokyo Denki University)

The talk in one slide

Setting: X a shift space on a finite alphabet (generalises naturally)

Theorem (Known results)

Suppose X has specification. Then

1. *bounded distortion $\Rightarrow$ unique equilibrium state + Gibbs*
2. *Gibbs $\Rightarrow$ large deviations principle*

Goal: Same results with non-uniform versions of above properties

Key idea:

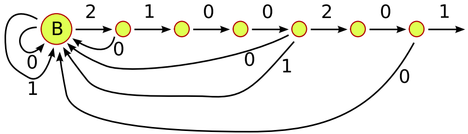
- $\mathcal{L}$ the language of X (space of finite orbit segments)
- Only require properties for $\mathcal{G} \subset \mathcal{L}$
- Get results if $\mathcal{G}$ is "big enough"

Shift spaces, languages, and sets of words

Shift space: closed, shift-invariant set $X \subset \mathcal{A}^{\mathbb{N}}$ ($\mathcal{A}$ finite: alphabet)

- Finite word $w \in \mathcal{A}^* = \bigcup_{n \geq 0} \mathcal{A}^n \rightsquigarrow$ cylinder $[w]$
- Language of X is $\mathcal{L} = \{w \in \mathcal{A}^* \mid [w] \neq \emptyset\}$.

Example: $\beta > 1 \rightsquigarrow X = \Sigma_\beta$ is coding space for $x \mapsto \beta x \pmod 1$



Sequence determined by $1 = \sum_{n=1}^{\infty} a_n \beta^{-n}$

$\mathcal{L} = \{$labels of paths starting at **B**$\}$

Consider subsets $\mathcal{D} \subset \mathcal{L}$ (points + times) / (orbit segments)

- $\mathcal{G} = \{$labels for paths starting and ending at **B**$\}$
- $\mathcal{C}^s = \{$labels for paths that never return to **B**$\}$

Specification

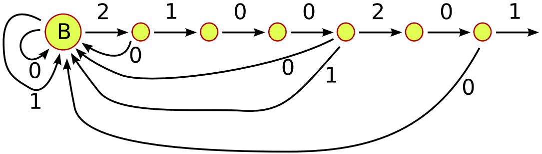
Various transitivity/mixing properties for (X, σ):

full shift $\Rightarrow$ (irreducible) Markov $\Rightarrow$ (weak) specification $\Rightarrow$ transitive

Definition: $\mathcal{D} \subset \mathcal{L}$ has specification if $\exists \tau$ (gluing time) s.t. words from $\mathcal{D}$ can be glued together with connecting words of length $\leq \tau$

- $\forall w^1, \dots, w^k \in \mathcal{D}$ there exist $v^1, \dots, v^k \in \mathcal{L}$ such that $w^i v^i w^{i+1} \cdots v^{j-1} w^j \in \mathcal{D}$ for all $1 \leq i < j \leq k$

Example: For the β-shifts, $\mathcal{G}$ has specification, but $\mathcal{L}$ does not



Large deviations and thermodynamics

$\mathcal{M}(X) = \{\text{Borel prob. measures on } X\}$ $\mathcal{E}_n(x)(\varphi) = S_n\varphi(x)$

- Empirical measures: $\mathcal{E}_n(x) = \frac{1}{n}\sum_{k=0}^{n-1} \delta_{\sigma^k x}$

Large deviations: Study decay of $m\{x \mid \mathcal{E}_n(x) \in U\}$

- $m \in \mathcal{M}(X)$ is reference measure, and $U \subset \mathcal{M}(X)$

Pressure of φ on $\mathcal{D} \subset \mathcal{L}$ is $P(\mathcal{D}, \varphi) = \lim \frac{1}{n} \log \left(\sum_{\mathcal{D}_n} e^{\varphi_n(w)}\right)$

- $\mathcal{D}_n = \{w \in \mathcal{D} \mid |w| = n\}$ $\varphi_n(w) = \sup_{x\in[w]} S_n\varphi(x)$

Variational principle: $P(\varphi) = \sup\{h(\mu) + \int \varphi \, d\mu \mid \mu \in \mathcal{M}_\sigma(X)\}$

- $\mathcal{M}_\sigma(X) = \{\mu \in \mathcal{M}(X) \mid \mu \text{ is } \sigma\text{-invariant}\}$
- Supremum achieved by equilibrium states

Classical (uniform) results

Bowen (1974): If (X,σ) has specification and φ is Hölder, then:

- φ has a unique equilibrium state $\mu \in \mathcal{M}_\sigma(X)$
- μ is Gibbs: $K \leq \frac{\mu[w]}{e^{-nP(\varphi)+S_n\varphi(x)}} \leq K'$ for all $x \in [w]$, $w \in \mathcal{L}_n$

Young (1990): If (X,σ) has specification and m is Gibbs for φ, then we have a large deviations principle with reference measure m:

$$U \subset \mathcal{M}(X) \text{ open } \Rightarrow \underline{\lim}_{n\to\infty} \frac{1}{n} \log m\{x \mid \mathcal{E}_n(x) \in U\} \geq \sup_{\mu\in U} q(\mu)$$

$$F \subset \mathcal{M}(X) \text{ closed } \Rightarrow \overline{\lim}_{n\to\infty} \frac{1}{n} \log m\{x \mid \mathcal{E}_n(x) \in F\} \leq \sup_{\mu\in F} q(\mu)$$

Rate function $q(\mu) = \begin{cases} h(\mu) + \int \varphi \, d\mu - P(\varphi) & \mu \in \mathcal{M}_\sigma(X) \\ -\infty & \mu \notin \mathcal{M}_\sigma(X) \end{cases}$

Motivating idea

Similar theorems in non-uniform setting given following condition:

- "$\mathcal{G} \subset \mathcal{L}$ has good properties, and every word in $\mathcal{L}$ can be transformed into a word in $\mathcal{G}$ without too much fuss"

For uniqueness, this means every $\mathcal{G}^M$ has specification, and

- Transform $w \in \mathcal{L}$ to $v \in \mathcal{G}$ by removing "bad bits" from ends (Decompose as $w = u^p v u^s$)
- u^p, u^s come from a list $\mathcal{C} \subset \mathcal{L}$ of "obstructions", and list is "thermodynamically small" $(P(\mathcal{C}, \varphi) < P(\varphi))$

For large deviations, this means $\mathcal{G}$ has spec, m Gibbs on φ, and

- $\mathcal{L} \rightsquigarrow \mathcal{G}$ by making edits (insertions, deletions, changes)
- Number of edits $\leq g(|w|)$, where $\frac{g(n)}{n} \to 0$

Decompositions and uniqueness

Decomposition of $\mathcal{L}$: sets $\mathcal{C}^p, \mathcal{G}, \mathcal{C}^s \subset \mathcal{L}$ such that $\mathcal{L} = \mathcal{C}^p\mathcal{G}\mathcal{C}^s$.

$$\mathcal{G}^M = \{uvw \in \mathcal{L} \mid u \in \mathcal{C}^p, v \in \mathcal{G}, w \in \mathcal{C}^s, |u|, |w| \leq M\}$$

Theorem (C.–Thompson, 2012)

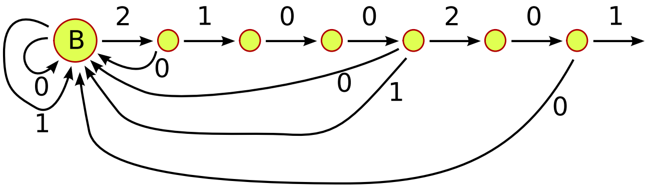
Suppose $\mathcal{L}$ has a decomposition such that

1. φ *has bounded distortion on* $\mathcal{G}$
2. $\mathcal{G}^M$ *has specification for every* M
3. $P(\mathcal{C}^p \cup \mathcal{C}^s, \varphi) < P(\varphi)$

Then φ *has a unique equilibrium state* μ*. It is Gibbs on each* $\mathcal{G}^M$*.*

Example: β-shift

$$\mathcal{C}^p = \emptyset$$
$$\mathcal{G} = \{\text{words (paths) starting and ending at } B\}$$
$$\mathcal{C}^s = \{\text{words (paths) starting at } B \text{ and never returning}\}$$



- $\mathcal{L} = \mathcal{C}^p \mathcal{G} \mathcal{C}^s$
- $\mathcal{G}^M$ corresponds to paths ending in first M vertices, so $\mathcal{G}^M$ has specification for each M
- $h(\mathcal{C}) = 0$, where $\mathcal{C} = \mathcal{C}^p \cup \mathcal{C}^s$
- In fact, $P(\mathcal{C}, \varphi) < P(\varphi)$ for every Hölder φ

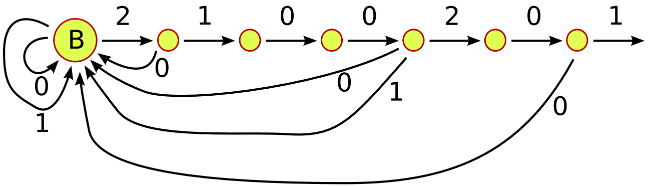
Statistical specification properties

Large deviations results have been obtained for β-shift and other systems by using statistical specification properties.

- Pfister, Sullivan (2005)
- Yamamoto (2009)
- Varandas (2012)

All reflect idea that the gluing procedure can be weakened in a way that does not interfere too much with Birkhoff averages.

β-shifts



Given any $v \in \mathcal{L}$, can transform v into a word $u \in \mathcal{G}$ by making a single change. (Change last non-zero symbol to 0).

Thus given any $v, w \in \mathcal{L}$, the word vw may not be in $\mathcal{L}$, but can be transformed into a word in $\mathcal{L}$ by making a single change.

General method for getting a word that concatenates statistical properties of v and w, as long as $\frac{\text{number of changes}}{\text{length of word}} \to 0$.

Edit metric

Goal: Define a metric on A^* (set of all finite words) that controls how much Birkhoff sums can vary.

An edit of a word w is any of the following:

- Substition: $w = uav \mapsto w' = ubv$ $\qquad u, v \in A^*, \ a, b \in A$
- Insertion: $w = uv \mapsto w' = ubv$ $\qquad u, v \in A^*, \ b \in A$
- Deletion: $w = uav \mapsto w' = uv$ $\qquad u, v \in A^*, \ a \in A$

$\hat{d}(v, w) =$ minimum number of edits required to go from v to w.

Key property: Let D be a metric inducing the weak* topology on $\mathcal{M}(X)$. Then for every $\eta > 0$ there is $\delta > 0$ such that if $\frac{\hat{d}(v,w)}{|v|} < \delta$, then $D(\mathcal{E}_{|v|}(x), \mathcal{E}_{|w|}(y)) < \eta$ for all $x \in [v]$ and $y \in [w]$.

Edit approachability

mistake function: a non-increasing sub-linear function $g\colon \mathbb{N} \to \mathbb{N}$.
$(\frac{g(n)}{n} \to 0)$

$\mathcal{L}$ is edit approachable by $\mathcal{G} \subset \mathcal{L}$ if there exists a mistake function g such that for every $v \in \mathcal{L}$, there is $w \in \mathcal{G}$ with $\hat{d}(v, w) < g(|v|)$.

Theorem (C.–Thompson–Yamamoto, 2013)

X a shift space on a finite alphabet, $\mathcal{L}$ its language. Suppose

1. *$\mathcal{L}$ is edit approachable by $\mathcal{G}$,*
2. *$\mathcal{G}$ has specification (with good concatenations),*
3. *$m \in \mathcal{M}(X)$ is Gibbs for φ on $\mathcal{G}$.*

Then X satisfies a LDP with reference measure m and rate q^φ

In particular, every Hölder continuous φ on a β-shift

S-gap shifts

Fix $S \subset \mathbb{N}$, define shift X in terms of its language:

- $\mathcal{L} = \{0^k 1 0^{n_1} 1 0^{n_2} 1 \cdots 0^{n_j} 1 0^\ell \mid n_i \in S\}$

Natural decomposition with $h(\mathcal{C}) = 0$:

- $\mathcal{C}^p = \{0^k 1 \mid k \in \mathbb{N}\}$
- $\mathcal{G} = \{0^{n_1} 1 \cdots 0^{n_j} 1 \mid n_i \in S\}$
- $\mathcal{C}^s = \{0^\ell \mid \ell \in \mathbb{N}\}$

For S-gap shifts, every Hölder potential has $P(\varphi) > \sup \overline{\lim} \frac{1}{n} S_n \varphi$.

Same results as β-shifts: unique eq state, Gibbs on $\mathcal{G}^M$, LDP

Open questions: What about piecewise expanding interval maps? Are there coded systems with $h(\mathcal{C}) = 0$ for which some Hölder potentials have zero entropy equilibrium states?

Conclusion

Moral of the story:

Many good consequences of specification (and other properties) can still be obtained as long as properties hold on a "large enough" set of words (orbit segments)

"Large enough" means the ability to get from $\mathcal{L}$ to $\mathcal{G}$ with some "small" tinkering, where meaning of "small" depends on context

- Unique equilibrium state: only need to remove a prefix and a suffix from the word in $\mathcal{L}$, and these come from "small" lists
- Large deviations: only need to make a small number of edits