Analyticity and Derivatives of Entropy Rate for HMP's

Guangyue Han
University of Hong Kong

Brian Marcus
University of British Columbia

Hidden Markov chains

- A **hidden Markov chain** $Z = \{Z_i\}$ is a stochastic process of the form $Z_i = \Phi(Y_i)$, where $Y = \{Y_i\}$ denotes a stationary finite-state Markov chain (with the probability transition matrix $\Delta = (p(j|i))$) and Φ is a deterministic function on the Markov states.

- Alternatively a **hidden Markov chain** Z can be defined as the output process obtained when passing a stationary finite-state Markov chain X through a noisy channel.

An Example from Digital Communication

At time n a binary symmetric channel with crossover probability ε (denoted by BSC(ε)) can be characterized by the following equation

$$Z_n = X_n \oplus E_n,$$

where $\oplus$ denotes binary addition, X_n denotes the binary input, E_n denotes the i.i.d. binary noise with $p_E(0) = 1 - \varepsilon$ and $p_E(1) = \varepsilon$, and Z_n denotes the corrupted output.

We further assume that $X = \{X_i\}$ be the binary input Markov chain with the probability transition matrix

$$\Pi = \begin{bmatrix} \pi_{00} & \pi_{01} \\ \pi_{10} & \pi_{11} \end{bmatrix}.$$

Since $\{Y_i\} = \{(X_i, E_i)\}$ is jointly Markov with

$$\Delta = \left[\begin{array}{c|cccc} y & (0,0) & (0,1) & (1,0) & (1,1) \\ \hline (0,0) & \pi_{00}(1-\varepsilon) & \pi_{00}\varepsilon & \pi_{01}(1-\varepsilon) & \pi_{01}\varepsilon \\ (0,1) & \pi_{00}(1-\varepsilon) & \pi_{00}\varepsilon & \pi_{01}(1-\varepsilon) & \pi_{01}\varepsilon \\ (1,0) & \pi_{10}(1-\varepsilon) & \pi_{10}\varepsilon & \pi_{11}(1-\varepsilon) & \pi_{11}\varepsilon \\ (1,1) & \pi_{10}(1-\varepsilon) & \pi_{10}\varepsilon & \pi_{11}(1-\varepsilon) & \pi_{11}\varepsilon \end{array}\right].$$

$\{Z_i\}$ is a hidden Markov chain with $Z = \Phi(Y)$, where Φ maps states $(0,0)$ and $(1,1)$ to 0 and maps states $(0,1)$ and $(1,0)$ to 1.

Entropy Rate

For a stationary stochastic process $Y = \{Y_i\}$, the **entropy rate** of Y is defined as

$$H(Y) = \lim_{n \to \infty} H_n(Y),$$

where

$$H_n(Y) = H(Y_0 | Y_{-1}, Y_{-2}, \cdots, Y_{-n}) = \sum_{y_{-n}^0} -p(y_{-n}^0) \log p(y_0 | y_{-n}^{-1}).$$

Let Y be a stationary first order Markov chain with

$$\Delta(i,j) = p(y_1 = j | y_0 = i).$$

It is well known that

$$H(Y) = -\sum_{i,j} p(y_0 = i) \Delta(i,j) \log \Delta(i,j).$$

In this talk

- Analyticity of entropy rate of a hidden Markov chain as a function of the underlying Markov chain parameters under mild positivity assumptions.
- Analyticity of hidden Markov chain in the sense of functional analysis.
- A "stabilizing" property and then obtain Taylor series expansion for "Black Hole" case.
- An example for determining domain of analyticity.
- Necessary and sufficient conditions for analyticity of the entropy rate for a hidden Markov chain with unambiguous symbol.

A random dynamical system

Let B be the number of states for Markov chain Y.

For each symbol a of the Z process, we form Δ_a by zeroing out the columns of Δ corresponding to states that do not map to a.

Let W be the simplex, comprising the vectors

$$\{w = (w_1, w_2, \cdots, w_B) \in \mathbb{R}^B : w_i \geq 0, \sum_i w_i = 1\},$$

and similarly we form W_a by zeroing out all the coordinates corresponding to the states that do not map to a.

Let $W^{\mathbb{C}}$, $W_a^{\mathbb{C}}$ denote the complex version of W, W_a, respectively.

For each symbol a, define the vector-valued function f_a on W by

$$f_a(w) = w\Delta_a / r_a(w),$$

where $r_a(w) = w\Delta_a\mathbf{1}$. For any fixed n and z_{-n}^0, define

$$x_i = x_i(z_{-n}^i) = p(y_i = \cdot\,|z_i, z_{i-1}, \cdots, z_{-n}). \tag{1}$$

Then, $\{x_i\}$ satisfies the random dynamical iteration

$$x_{i+1} = f_{z_{i+1}}(x_i), \tag{2}$$

starting with the stationary distribution of Y.

Blackwell showed that

$$H(Z) = -\int \sum_a r_a(w) \log r_a(w) dQ(w), \tag{3}$$

where Q, known as **Blackwell's measure**, is the limiting probability distribution, as $n \to \infty$, of $\{x_0\}$ on W. Moreover, Q satisfies

$$Q(E) = \sum_a \int_{f_a^{-1}(E)} r_a(w) dQ(w). \tag{4}$$

Analyticity Theorem

Theorem 1. *Suppose that the entries of Δ are analytically parameterized by a real variable vector $\vec{\varepsilon}$. If at $\vec{\varepsilon} = \vec{\varepsilon}_0$,*

1. *Every column of Δ is either all zero or strictly positive –and–*
2. *For all a, Δ_a has at least one strictly positive column,*

then $H(Z)$ is a real analytic function of $\vec{\varepsilon}$ at $\vec{\varepsilon}_0$.

Remark Note that Theorem 1 holds when Δ is positive.

Contraction Property of the Random Dynamical System

- For each a and any two points $w, v \in W_a^{\mathbb{C}}$, define the following metric:

$$d_a(w,v) = \max\{d_{a,H}(w,v), d_{a,E}(w,v)\},$$

where

$$d_{a,H}(w,v) = d_H(w_{I_p(a)}, v_{I_p(a)}), \qquad d_{a,E}(w,v) = d_E(w_{I_z(a)}, v_{I_z(a)}).$$

- For $a_1, a_2 \in A$, the mapping $f_{a_2} : W_{a_1} \to W_{a_2}$ is a contraction mapping under the metric defined above.

- Applying mean value theorem, one can show that on certain neighborhood of W_{a_1} in $W_{a_1}^{\mathbb{C}}$ will be a contraction mapping as well.

- There is a universal contraction coefficient for all pairs (a_1, a_2), denoted by ρ.

Proof. Extend the dynamical system from real to complex:

$$x_{i+1}^{\vec{\varepsilon}} = f_{z_{i+1}}^{\vec{\varepsilon}}(x_i^{\vec{\varepsilon}}), \tag{5}$$

starting with

$$x_{-n-1}^{\vec{\varepsilon}} = p^{\vec{\varepsilon}}(y_{-n-1} = \cdot\ ). \tag{6}$$

If the extension is "small" enough, we can prove $x_i^{\vec{\varepsilon}}$ stay within f_a-contracting complex neighborhood. As a consequence, there is a positive constant L'', independent of n_1, n_2,

$$|\log p^{\vec{\varepsilon}}(z_0|z_{-n_1}^{-1}) - \log p^{\vec{\varepsilon}}(\hat{z}_0|\hat{z}_{-n_2}^{-1})| \leq L\rho^n, \tag{7}$$

if $z_j = \hat{z}_j$ for $j = 0, 1, \cdots, n$.

Further choose the extension and σ with $1 < \sigma < 1/\rho$ such that

$$\sum_{z_{-n-1}^0} |p^{\vec{\varepsilon}}(z_{-n-1}^0)| \leq \sigma^{n+2}. \tag{8}$$

Let

$$H_n^{\vec{\varepsilon}}(Z) = -\sum_{z_{-n}^0} p^{\vec{\varepsilon}}(z_{-n}^0) \log p^{\vec{\varepsilon}}(z_0|z_{-n}^{-1}),$$

then we have

$$|H_{n+1}^{\vec{\varepsilon}}(Z) - H_n^{\vec{\varepsilon}}(Z)| = |\sum_{z_{-n-1}^0} p^{\vec{\varepsilon}}(z_{-n-1}^0) \log p^{\vec{\varepsilon}}(z_0|z_{-n-1}^{-1}) - \sum_{z_{-n}^0} p^{\vec{\varepsilon}}(z_{-n}^0) \log p^{\vec{\varepsilon}}(z_0|z_{-n}^{-1})|$$

$$= |\sum_{z_{-n-1}^0} p^{\vec{\varepsilon}}(z_{-n-1}^0)(\log p^{\vec{\varepsilon}}(z_0|z_{-n-1}^{-1}) - \log p^{\vec{\varepsilon}}(z_0|z_{-n}^{-1}))| \leq \sigma^2 L(\rho\sigma)^n.$$

Thus, for $m > n$,

$$|H_m^{\vec{\varepsilon}}(Z) - H_n^{\vec{\varepsilon}}(Z)| \leq \sigma^2 L''((\rho\delta)^n + \ldots + (\rho\delta)^{m-1}) \leq \frac{\sigma^2 L(\rho\sigma)^n}{1-\rho\sigma}.$$

$\square$

Entropy Rate Again

Let $\mathcal{X}$ denote the set of left infinite sequences with finite alphabet. For real $\vec{\varepsilon}$, consider the measure $\nu^{\vec{\varepsilon}}$ on $\mathcal{X}$ defined by:

$$\nu^{\vec{\varepsilon}}(\{x_{-\infty}^{0} : x_0 = z_0, \cdots, x_{-n} = z_{-n}\}) = p^{\vec{\varepsilon}}(z_{-n}^{0}). \tag{9}$$

Note that $H(Z)$ can be rewritten as

$$H^{\vec{\varepsilon}}(Z) = \int -\log p^{\vec{\varepsilon}}(z_0 | z_{-\infty}^{-1}) d\nu^{\vec{\varepsilon}}. \tag{10}$$

Main Idea

Let F^θ denote the Banach space consisting of all the "exponential forgetting" functions on $\mathcal{X}$ with certain norm. For every $f \in F^\theta$, there is an unique **equilibrium state** μ_f. Ruelle Showed that the mapping from f to μ_f is analytic.

In our case, for $f(\vec{\varepsilon}, z) = \log p^{\vec{\varepsilon}}(z_0|z_{-\infty}^{-1})$, one can prove $\mu_{f(\vec{\varepsilon},\cdot)} = \nu^{\vec{\varepsilon}}$ as in (9).

$$\vec{\varepsilon} \to \log p^{\vec{\varepsilon}}(z_0|z_{-\infty}^{-1}) \to \nu^{\vec{\varepsilon}}$$

In order to show $\nu^{\vec{\varepsilon}}$ is analytic with respect to $\vec{\varepsilon}$, we only need to show that $\vec{\varepsilon} \mapsto f(\vec{\varepsilon}, z)$ is analytic as a mapping from real parameter space to F^ρ.

Analyticity in a Strong Sense

Theorem 2. *Suppose that the entries of* Δ *are analytically parameterized by a real variable vector* $\vec{\varepsilon}$*. If at* $\vec{\varepsilon} = \vec{\varepsilon}_0$*,* Δ *satisfies conditions* 1 *and* 2 *in Theorem 1, then the mapping* $\vec{\varepsilon} \mapsto \nu^{\vec{\varepsilon}}$ *is analytic at* $\vec{\varepsilon}_0$ *from the real parameter space to* $(F^{\rho})^*$*.*

So, under certain assumptions, a hidden Markov chain **itself** is analytic, which, in principle, implies analyticity of other statistical quantities.

Theorem 1 is an immediate corollary of Theorem 2:

Proof. The map

$$\Omega \to F^{\rho} \times (F^{\rho})^{*} \to \mathbb{R}$$

$$\vec{\varepsilon} \mapsto (f^{\vec{\varepsilon}}, \nu^{\vec{\varepsilon}}) \mapsto \nu^{\vec{\varepsilon}}(f^{\vec{\varepsilon}})$$

is analytic at $\vec{\varepsilon}_0$, as desired. $\square$

Stabilizing Property of Derivatives in Black Hole Case

Suppose that for every $a \in A$, Δ_a is a rank one matrix, and every column of Δ_a is either strictly positive or all zeros. We call this the *Black Hole* case.

Theorem 3. *If at* $\varepsilon = \hat{\varepsilon}$*, for every* $a \in A$*,* Δ_a *is a rank one matrix, and every column of* Δ_a *is either a positive or a zero column, then for* $\vec{n} = (n_1, n_2, \cdots, n_m)$*,*

$$H(Z)^{(\vec{n})}\Big|_{\varepsilon=\hat{\varepsilon}} = H_{\lceil (|\vec{n}|+1)/2 \rceil}(Z)^{(\vec{n})}\Big|_{\varepsilon=\hat{\varepsilon}}$$

Sketch of the proof

Consider the iteration:

$$x_i = \frac{x_{i-1}\Delta_{z_i}}{x_{i-1}\Delta_{z_i}\mathbf{1}}.$$

- x_i can be viewed as a function (denoted by g) of x_{i-1} and Δ_a.
- g is a constant as a function of x_{i-1}.
- At $\varepsilon = \hat{\varepsilon}$

$$x_i = p(y_i = \cdot | z_{-\infty}^i) = \frac{x_{i-1}\Delta_{z_i}}{x_{i-1}\Delta_{z_i}\mathbf{1}} = \frac{p(y_{i-1} = \cdot)\Delta_{z_i}}{p(y_{i-1} = \cdot)\Delta_{z_i}\mathbf{1}} = p(y_i = \cdot | z_i). \tag{11}$$

- Taking n-th order derivatives, we have

$$x_i^{(n)} = \left.\frac{\partial g}{\partial x_{i-1}}\right|_{\varepsilon=\hat{\varepsilon}} (x_{i-1}, \Delta_{z_i})\, x_{i-1}^{(n)} + \text{other terms},$$

  where "other terms" involve only lower order (than n) derivatives of x_{i-1}.

- By induction, we conclude that

$$x_i^{(n)} = p^{(n)}(y_i|z_{-\infty}^i) = p^{(n)}(y_i|z_{i-n}^i).$$

  at $\varepsilon = \hat{\varepsilon}$.

- We then have that for all sequences $z_{-\infty}^0$ the n-th derivative of $p(z_0|z_{-\infty}^{-1})$ stabilizes:

$$p^{(n)}(z_0|z_{-\infty}^{-1}) = p^{(n)}(z_0|z_{-n-1}^{-1}) \;\text{ at } \varepsilon = \hat{\varepsilon}. \tag{12}$$

For n-th order derivative of $H(Z)$,

$$H^{(n)}(Z) = \lim_{k\to\infty} \sum_{z^0_{-k}} \sum_{l=1}^{n} C_{n-1}^{l-1} p^{(l)}(z^0_{-k})(\log p(z_0|z^{-1}_{-k}))^{(n-l)}$$

$$= \lim_{k\to\infty} \sum_{z^0_{-k}} \sum_{l=1}^{n} C_{n-1}^{l-1} p^{(l)}(z^0_{-k})(\log p(z_0|z^{-1}_{-n}))^{(n-l)}$$

$$= \sum_{z^0_{-n}} \sum_{l=1}^{n} C_{n-1}^{l-1} p^{(l)}(z^0_{-n})(\log p(z_0|z^{-1}_{-n}))^{(n-l)} = H_n^{(n)}(Z).$$

Interesting Property of Entropy Rate

The arguments in the proof shows that:

The stabilization of entropy rate (of ANY process) is at least twice faster than the stabilization of conditional probability.

Digital Communication Example Revisited

If all transition probabilities π_{ij}'s are positive, then at $\varepsilon = 0$,

$$\Delta = \left[\begin{array}{c|cccc} y & (0,0) & (0,1) & (1,0) & (1,1) \\ \hline (0,0) & \pi_{00} & 0 & \pi_{01} & 0 \\ (0,1) & \pi_{00} & 0 & \pi_{01} & 0 \\ (1,0) & \pi_{10} & 0 & \pi_{11} & 0 \\ (1,1) & \pi_{10} & 0 & \pi_{11} & 0 \end{array}\right].$$

$$\Delta_0 = \left[\begin{array}{c|cccc} y & (0,0) & (0,1) & (1,0) & (1,1) \\ \hline (0,0) & \pi_{00} & 0 & 0 & 0 \\ (0,1) & \pi_{00} & 0 & 0 & 0 \\ (1,0) & \pi_{10} & 0 & 0 & 0 \\ (1,1) & \pi_{10} & 0 & 0 & 0 \end{array}\right], \Delta_1 = \left[\begin{array}{c|cccc} y & (0,0) & (0,1) & (1,0) & (1,1) \\ \hline (0,0) & 0 & 0 & \pi_{01} & 0 \\ (0,1) & 0 & 0 & \pi_{01} & 0 \\ (1,0) & 0 & 0 & \pi_{11} & 0 \\ (1,1) & 0 & 0 & \pi_{11} & 0 \end{array}\right].$$

Note that both Δ_0 and Δ_1 are rank 1 matrix with at least one positive column, which is Black Hole case. Thus the Taylor series of $H(Z)$ when $\varepsilon = 0$ exists and can be exactly calculated. This result, as a special case of Theorem 3, recovers computational work by Zuk et. al. [2004].

Again for this example, domain of analyticity of entropy rate can be determined as follows: for given ρ with $0 < \rho < 1$, choose r and R to satisfy all the following constraints. Then the entropy rate is an analytic function of ε on $|\varepsilon| < r$.

$$0 \leq \frac{\sqrt{r(r+1)|-\pi_{00}\pi_{11}+\pi_{10}\pi_{01}|}}{\pi_{11}-|\pi_{10}-\pi_{11}|r-(|\pi_{00}-\pi_{10}-\pi_{01}+\pi_{11}|r+|\pi_{01}-\pi_{11}|)R} < \sqrt{\rho},$$

$$0 \leq \frac{\sqrt{r(r+1)|-\pi_{00}\pi_{11}+\pi_{10}\pi_{01}|}}{\pi_{01}-|\pi_{00}-\pi_{01}|r-(|\pi_{00}-\pi_{10}-\pi_{01}+\pi_{11}|r+|\pi_{01}-\pi_{11}|)R} < \sqrt{\rho},$$

$$0 \leq \frac{\sqrt{r(r+1)|-\pi_{11}\pi_{00}+\pi_{01}\pi_{10}|}}{\pi_{00}-|\pi_{01}-\pi_{00}|r-(|\pi_{00}-\pi_{10}+\pi_{11}-\pi_{01}|r+|\pi_{10}-\pi_{00}|)R} < \sqrt{\rho},$$

$$0 \leq \frac{\sqrt{r(r+1)|-\pi_{11}\pi_{00}+\pi_{01}\pi_{10}|}}{\pi_{10}-|\pi_{11}-\pi_{10}|r-(|\pi_{00}-\pi_{10}+\pi_{11}-\pi_{01}|r+|\pi_{10}-\pi_{00}|)R} < \sqrt{\rho},$$

$$0 \leq \frac{r\pi_{00}}{\pi_{01}-|\pi_{00}-\pi_{01}|r} < R(1-\rho), 0 \leq \frac{r\pi_{10}}{\pi_{11}-|\pi_{10}-\pi_{11}|r} < R(1-\rho),$$

$$0 \leq \frac{r\pi_{11}}{\pi_{10}-|\pi_{11}-\pi_{10}|r} < R(1-\rho), 0 \leq \frac{r\pi_{01}}{\pi_{00}-|\pi_{01}-\pi_{00}|r} < R(1-\rho),$$

$$2(|\pi_{00}-\pi_{01}-\pi_{10}+\pi_{11}|r+|\pi_{01}-\pi_{11}|)R+2|\pi_{10}-\pi_{11}|r+1 < 1/\rho,$$

$$2(|\pi_{10}-\pi_{11}-\pi_{00}+\pi_{01}|r+|\pi_{11}-\pi_{01}|)R+2|\pi_{00}-\pi_{01}|r+1 < 1/\rho.$$

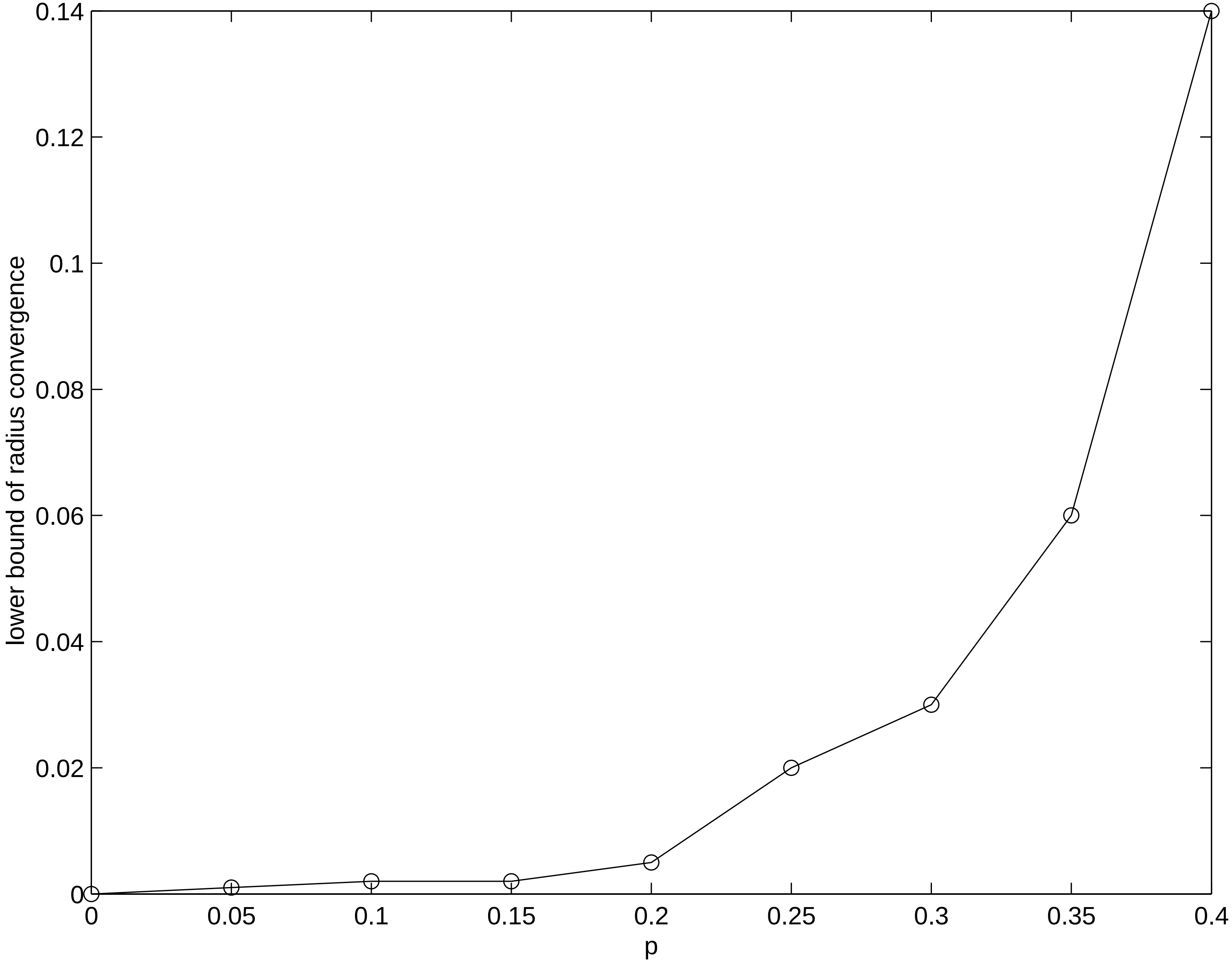

Figure 1: lower bound on radius of convergence as a function of p

Hidden Markov Chains with Unambiguous Symbol

Definition 4. A symbol a is called unambiguous if $\Phi^{-1}(a)$ contains only one element.

When an unambiguous symbol is present, the entropy rate can be expressed in a simple way. In the case of a binary hidden Markov chain, where 0 is unambiguous,

$$H(Z^\varepsilon) = p^\varepsilon(0)H^\varepsilon(z|0)+p^\varepsilon(10)H^\varepsilon(z|10)+\cdots+p^\varepsilon(1^{(n)}0)H^\varepsilon(z|1^{(n)}0)+\cdots,$$

where $1^{(n)}$ denotes the sequence of n 1's and

$$H^\varepsilon(z|1^{(n)}0) = -p^\varepsilon(0|1^{(n)}0)\log p^\varepsilon(0|1^{(n)}0)-p^\varepsilon(1|1^{(n)}0)\log p^\varepsilon(1|1^{(n)}0).$$

Example of Non-analyticity

Consider the following parameterized stochastic matrix

$$\Delta(\varepsilon) = \begin{bmatrix} \varepsilon & a-\varepsilon & b \\ g & c & d \\ h & e & f \end{bmatrix}.$$

The states of the Markov chain are the matrix indices $\{1, 2, 3\}$. Let Z^{ε} be the binary hidden Markov chain defined by: $\Phi(1) = 0$ and $\Phi(2) = \Phi(3) = 1$. We claim that $H(Z^{\varepsilon})$ is not analytic at $\varepsilon = 0$.

Let $\pi(\varepsilon) = (\pi_1(\varepsilon), \pi_2(\varepsilon), \pi_3(\varepsilon))$ be the stationary vector of $\Delta(\varepsilon)$. Since $\Delta(\varepsilon)$ is irreducible, $\pi(\varepsilon)$ is analytic in ε and positive. Now,

$$p^{\varepsilon}(0)H^{\varepsilon}(z|0) = -p^{\varepsilon}(00)\log p^{\varepsilon}(0|0) - p^{\varepsilon}(10)\log p^{\varepsilon}(1|0).$$

$$= -\pi_1(\varepsilon)\varepsilon\log\varepsilon - \pi_1(\varepsilon)(a - \varepsilon + b)\log(\pi_1(\varepsilon)(a - \varepsilon + b))$$

which is not analytic at $\varepsilon = 0$. However it can be shown that the sum of all other terms is analytic at $\varepsilon = 0$. Thus, $H(Z^{\varepsilon})$ is not analytic at $\varepsilon = 0$.

Another Example of Non-analyticity

Consider the following parameterized stochastic matrix

$$\Delta(\varepsilon) = \begin{bmatrix} e & a & b \\ f - \varepsilon & c & \varepsilon \\ g & 0 & c \end{bmatrix}.$$

The states of the Markov chain are the matrix indices $\{1, 2, 3\}$.

Let Z^ε be the binary hidden Markov chain defined by $\Phi(1) = 0$ and $\Phi(2) = \Phi(3) = 1$. We show that $H(Z^\varepsilon)$ is not analytic at $\varepsilon = 0$.

We have

$$p^\varepsilon(1|1^{(n)}0) = (ac^{n+1} + a\varepsilon(n+1)c^n + bc^{n+1})/(ac^n + a\varepsilon nc^{n-1} + bc^n)$$

$$= (ac^2 + a\varepsilon(n+1)c + bc^2)/(ac + a\varepsilon n + bc),$$

and

$$p^\varepsilon(0|1^{(n)}0) = ((f-\varepsilon)ac^n + ga\varepsilon nc^{n-1} + gbc^n)/(ac^n + a\varepsilon nc^{n-1} + bc^n)$$

$$= ((f-\varepsilon)ac + ga\varepsilon n + gbc)/(ac + a\varepsilon n + bc).$$

When $\varepsilon \to -(a+b)c/an$, the term $p^\varepsilon(1^{(n)}0)H^\varepsilon(z|1^{(n)}0) \to \infty$. Meanwhile, the sum of all the other terms is analytic. Thus, we conclude that $H(Z^\varepsilon)$ blows up when one approaches $-(a+b)c/an$ and therefore is not analytic at $\varepsilon = 0$.

Necessary and Sufficient Conditions for Analyticity

Theorem 5. *Let Δ be an irreducible stochastic $d \times d$ matrix with the following form:*

$$\Delta = \begin{bmatrix} a & r \\ c & B \end{bmatrix} \tag{13}$$

where a is a scalar and B is an $(d-1) \times (d-1)$ matrix. Let Φ be the function defined by $\Phi(1) = 0$, and $\Phi(2) = \cdots \Phi(n) = 1$. Then for any parametrization $\Delta(\varepsilon)$ such that $\Delta(\varepsilon_0) = \Delta$, letting Z^ε denote the hidden Markov chain defined by $\Delta(\varepsilon)$ and Φ, $H(Z^\varepsilon)$ is analytic at ε_0 if and only if

1. *$a > 0$, and $rB^j c > 0$ for $j = 0, 1, \cdots$.*
2. *The maximum eigenvalue of B is simple and strictly greater in absolute value than the other eigenvalues.*