Chaos, Complexity, and Inference (36-462)

Lecture 5: Symbolic Dynamics; Making Discrete Stochastic Processes from Continuous Deterministic Dynamics

Cosma Shalizi

27 January 2009

Symbolic dynamics Reducing continuous time series to sequences of discrete symbols

Stochastic processes How to get random sequences from deterministic dynamics

Topological entropy rate How many different things can your process do?

Formal languages Ways of making your process talk

Further reading: Badii and Politi (1997) is the reference on symbolic dynamics which comes closest to what we are doing (because I learned it from there). It's mathematically advanced, however. Chapter 9 of Devaney (1992) has a more pure-mathematical introduction; if you like that, you might consider Lind and Marcus (1995); Kitchens (1998) assumes you know a lot of abstract algebra.

Symbolic Dynamics

Start with our favorite dynamical system, with a continuous state S_t and a map Φ

$$S_{t+1} = \Phi(S_t)$$

Partition $\mathcal{B}$: divide the state space up into non-overlapping **cells**, $B_0, B_1, \ldots B_{k-1}$
$b(S_t) =$ label (**symbol**) for the cell S_t is in
$= X_t$ (say)
symbol sequence X

$$\begin{aligned} X_1^\infty &= b(S_1), b(S_2), b(S_3), \ldots \\ &= b(S_1), b(\Phi(S_1)), b(\Phi^{(2)}(S_1)), \ldots \end{aligned}$$

i.e., given initial condition S_1 and partition $\mathcal{B}$, symbol sequence X_1^∞ is fixed

The Shift Map

Seen symbols, what about the dynamics?
Shift map Σ

$$\Sigma(X_1^\infty) = X_2^\infty$$

Σ shifts the symbol sequence one place over
$\Sigma^{(k)}$ shifts the symbol sequence k places

$$\Sigma^{(k)}(X_1^\infty) = X_{1+k}^\infty$$

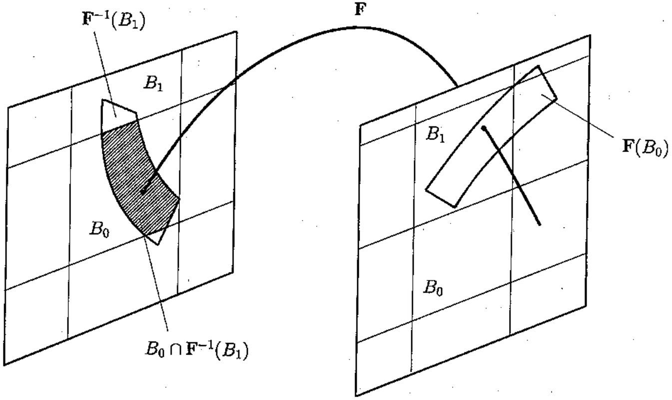
$$\begin{array}{ccc} S_t & \xrightarrow{\Phi} & S_{t+1} \\ b\downarrow & & b\downarrow \\ X_t^\infty & \xrightarrow{\Sigma} & X_{t+1}^\infty \end{array}$$

Why do this?

1. Model of finite-resolution measurements
2. “Continuous math is hard; let’s go discretize”
   - Discrete-math mathematical tools
   - Probability tools
3. Sometimes involves no real loss

Refinement of partitions

Subdivide cells according to which symbol they *will* give us



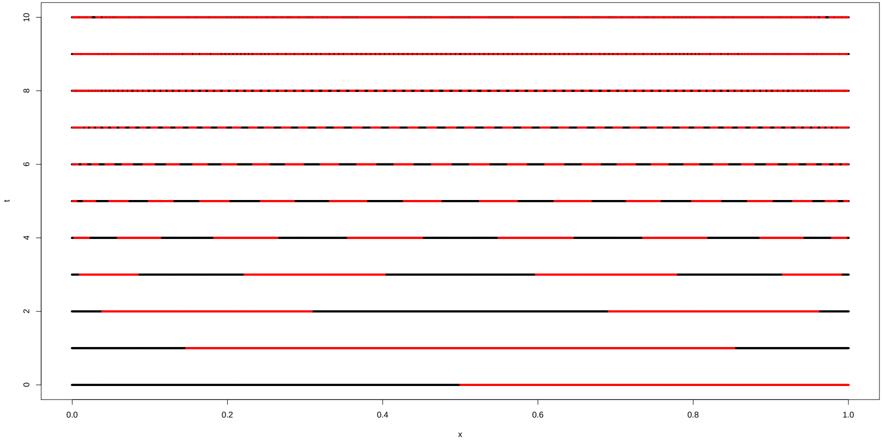
From Badii and Politi (1997, p. 71)

In math, work out $\Phi^{-1} B_i$ for each cell B_i

Now the new partition $\mathcal{B}^2$ is all the sets $B_i \cap \Phi^{-1} B_j$

Refinement: knowing the cell in $\mathcal{B}^2$ tells you the cell in $\mathcal{B}$, but not the other way

Example: Take the $r = 1$ logistic map with $B_0 = [0, 0.5)$, $B_1 = [0.5, 1]$. Call these L and R, so we don't mix them up with other things, and color them red and black



shows $\mathcal{B}, \Phi^{-1}\mathcal{B}, \Phi^{-2}\mathcal{B}, \ldots$

Generating Partitions

A partition is **generating** if the cells of $\mathcal{B}$, $\mathcal{B}^2$, $\mathcal{B}^3, \ldots$ keep getting smaller forever
or: the *infinite* symbol sequence X_1^∞ corresponds to a *unique* initial condition S_1
then we are back in change-of-coordinates land:

$$\begin{array}{ccc} S_t & \xrightarrow{\Phi} & S_{t+1} \\ b \downarrow & & \uparrow b^{-1} \\ X_t^\infty & \xrightarrow{\Sigma} & X_{t+1}^\infty \end{array}$$

Example: The example partition for the logistic map is generating

for all r, not just $r = 1$

Where did all the details go?

Most of these maps get pretty complicated pretty quickly
e.g. try writing out Φ^{20} for the logistic map
but Σ^{20} is as trivial as Σ
Trick: the complexity has moved out of the dynamical map to the state space — now the space of symbol sequences — and, possibly, probability distributions on the sequence space
This lets us use different mathematical tools

When are there generating partitions?

For one-dimensional maps, make a generating partition by putting boundaries at the "critical" points, i.e. maxima, minima, vertical asymptotes
For higher-dimensional maps, there are fewer general rules; don't always exist

Estimating generating partitions

"Symbolic false nearest neighbors" (Kennel and Buhl, 2003): if you have a generating partition, then close symbol sequences should only come from close points in the state-space

1. Reconstruct your state space
2. Start with an initial partition
3. Calculate distances among symbols sequences and distances among state points
4. Find "false symbolic neighbors"
5. Tweak partition boundaries to reduce the number of false neighbors
6. Iterate to convergence

Another approach ("symbolic shadowing"): similar symbol sequences should imply close *trajectories* in state space (Hirata *et al.*, 2004)

Discrete Stochastic Processes

... from fully deterministic continuous dynamics
If S_1 has a distribution, then so do $X_1 = b(S_1)$, $X_2 = b(\Phi(S_1))$, ..., $X_t = b(\Phi^{(t-1)}(S_1))$, ...
In general the X_t will be dependent on each other
$\Rightarrow$ symbol sequences are stochastic processes
Studying these processes can tell us about the dynamical system
Symbolic dynamics tells us about how stochastic processes arise

Again with the Logistic Map

$r = 1$
$B_0 = [0, \frac{1}{2})$, $B_1 = [\frac{1}{2}, 1]$
so X_t are binary variables (values L, R)
S_1 in invariant distribution
Claim: X_1^∞ is a sequence of IID, with $P(X_t = L) = 0.5$
Translation: the logistic map gives us perfect coin-tossing

The usual argument for this

1. Under the invariant distribution, $P(B_0) = 1/2$
2. Under the invariant distribution $P(\Phi^{-1}B_i|B_j) = 1/2$ for all i, j
   boundaries are not evenly placed, but distribution isn't uniform, cancels out
   so $b(S_t)$ and $b(S_{t+1})$ are independent
3. In fact $P(\Phi^{-k}B_i|\Phi^{k-1}B_{j_k}, \Phi^{k-2}B_{j_{k-1}}, \ldots B_{j_1}) = 1/2$
   so $b(R_{t_1}), b(R_{t_2}), b(R_{t_3}), \ldots$ are all independent

The usual mathematician's argument

Integrals with the invariant distribution and pre-images of the partition get messy; avoid

1. smoothly change coordinates to go from the logistic map, with state S_t, to the tent map, with state R_t
2. this changes the invariant distribution to be the uniform distribution
3. leaves the generating partition alone
4. write R_t as a binary number
5. $b(R_t) = R$ if and only if the first digit of R_t is "1"
6. $b(R_{t+1}) = R$ if the first digit of R_t was "1" and its second digit was "0", or if the first digit was "0" and the second digit was "1"; etc. for other two-digit combinations
7. $b(R_{t+1})$ is independent of $b(R_t)$
8. in fact $b(R_{t_1}), b(R_{t_2}), b(R_{t_3}), \ldots$ are all independent

But "why think when you can do the experiment?"

EXERCISE 1: Write a program to simulate the symbolic dynamics of the logistic map with $r = 1$. Tabulate the frequencies of sub-sequences of length $2n$. Test whether X_t^{t+n-1} is independent of X_{t-n}^{t-1}.

Independent symbols is an extreme case
More general, dependence across symbols
Two aspects:

- *absolute* restrictions on what sequences can appear (today)
- *relative frequency* dependence (next lecture)

Forbidden Sequences

Take $r = 0.966$; this is moderately chaotic (Lyapunov exponent ≈ 0.42)
You can verify either by calculation or simulation that “LLL” never appears
nor “LLRR”
nor infinitely many others
These are all **forbidden**
Those which do appear are **allowed**
Also say allowed and forbidden **words** (because they’re made from letters)

Topological Entropy Rate

Every allowed word of length n implies a word of length $n-1$
$\Rightarrow$ At least as many longer words as shorter words
$W_n =$ number of allowed words of length n

$$h_0 \equiv \lim_{n\rightarrow\infty} \frac{1}{n} \log_2 W_n$$

Measures the exponential growth in the number of allowed patterns as their length grows
For any process

$$h_0 \geq 0$$

2^{h_0} says (roughly) how many choices there are for ways of continuing the typical sequence

$h_0 > 0$ is necessary for sensitive dependence on initial conditions

fixed points or periodicity $\Rightarrow h_0 = 0$

With $r = 1$, $W_n = 2^n$ so $h_0 = \log_2 2 = 1$, thus two possible continuations

Careful: $h_0 = 0$ can mean only one possible sequence, or just sub-exponential growth

EXERCISE 2: Write a program to calculate the topological entropy rate for the logistic map at any r

EXERCISE 3: How would you put a standard error on your estimate of h_0?

Languages

A **word** is a finite sequence of symbols
A **(formal) language** is a set of allowed words
A **(formal) grammar** is a collection of rules which give you all and only the allowed words

blame the linguists for mixed metaphor

See Badii and Politi (1997) for more on how this applies to dynamical systems
See Charniak (1993) for statistical aspects
See Minsky (1967); Lewis and Papadimitriou (1998) for good introductions to formal languages

Regular expressions

Simplest sorts of formal grammars; used in Unix, Perl, etc.
Basic operations:

literals e.g. L, R, etc., depending on alphabet; also "none of the above", abbreviated ε

alternation "this or that"; make an arbitrary choice from two sets
e.g. "$L|R$" means "either L or R"

concatenation string together
$L(L|R)$ means "L, followed by either L or R", means "LL or LR"

"star", repetition Repeat something zero or more times
"L^*" matches "$\varepsilon, L, LL, LLL, \ldots$"
"$(L(L|R))^*$" matches
$\varepsilon, LL, LR, LLLL, LLLR, LRLL, LRLR, \ldots$

$(LR)^*$ a period-two sequence
forbidden words include: *RR*, *LL*

$(L(L|R))^*$ "odd-place symbols must be 'L', even-place can be L or R"
or: "every other symbol must be 'L' "
forbidden words include: *RR*

periodicity here is hidden

$(LRR(RR)^*)^* \equiv (L(RR)^+)^*$ blocks of Rs, of even length, separated by isolated Ls
forbidden words include: *LL*, *LRRRL*

$(L^*(RR)^*)^*$ even-length blocks of Rs, separated by blocks of Ls of arbitrary length
forbidden words include: *LRRRL*

Not describable by any regular expression: every "(" must be followed eventually by a matching ")"

Expressive Machinery

Basic theorem (Kleene, 1956): every regular expression can be implemented by a machine ("automaton") with a *finite* memory; finite automata can only implement regular expressions
"implement" can mean:

- check if words match the expression
- generate words which match

these are equivalent

think about generating, it feels more like dynamics

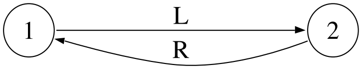
Machines as Directed Graphs

write machines as circle-and-arrow diagrams, directed graphs

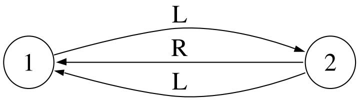
- **circles** "states"; fixes possible future sequences
  for symbolic dynamics, each state in the diagram is a *set* of states in the original, continuous state-space
- **arrows** go from circle to circle, labeled with symbols from the alphabet

paths generate words: write down the labels hit following that path
concatenation ≈ following arrows
alternation ≈ more than one out-going arrow from a circle
star ≈ loops
should distinguish allowed "start" and "stop" states

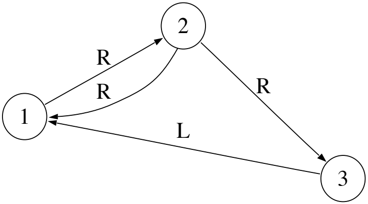
Some Machines and Expressions



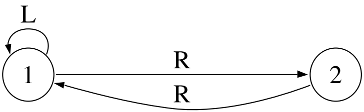
$(LR)^*$



$(L(L|R))^*$



$(L(RR)^+)^*$



$(L^*(RR)^*)^*$

Sofic Systems

Sofic: only finitely many circles (also called **finitary**)
Finite type: to determine what symbol is possible next, need only look back k symbols, for some *fixed* k
Strictly sofic: sofic, but not of finite type
$(LR)^*$ is of finite type
$(L(L|R))^*$, $(L(RR)^+)^*$ and $(L^*(RR)^*)^*$ are strictly sofic
These are the skeletons of stochastic processes
Finite type $\approx$ finite-order Markov chains
Strictly sofic $\approx$ hidden Markov, long-range correlations
next time: some statistics!

Badii, Remo and Antonio Politi (1997). *Complexity: Hierarchical Structures and Scaling in Physics*. Cambridge, England: Cambridge University Press.

Charniak, Eugene (1993). *Statistical Language Learning*. Cambridge, Massachusetts: MIT Press.

Devaney, Robert L. (1992). *A First Course in Chaotic Dynamical Systems: Theory and Experiment*. Reading, Mass.: Addison-Wesley.

Hirata, Yoshito, Kevin Judd and Devin Kilminster (2004). "Estimating a Generating Partition from Observed Time Series: Symbolic Shadowing." *Physical Review E*, **70**: 016215. doi:10.1103/PhysRevE.70.016215.

Kennel, Matthew B. and Michael Buhl (2003). "Estimating good discrete partitions from observed data: symbolic false nearest neighbors." *Physical Review Letters*, **91**: 084102. URL `http://arxiv.org/abs/nlin.CD/0304054`.

Kitchens, Bruce P. (1998). *Symbolic Dynamics: One-sided, Two-sided and Countable State Markov Shifts*. Berlin: Springer-Verlag.

Kleene, S. C. (1956). “Representation of Events in Nerve Nets and Finite Automata.” In *Automata Studies* (Claude E. Shannon and John McCarthy, eds.), vol. 34 of *Annals of Mathematics Studies*, pp. 3–41. Princeton, New Jersey: Princeton University Press.

Lewis, Harry R. and Christos H. Papadimitriou (1998). *Elements of the Theory of Computation*. Upper Saddle River, New Jersey: Prentice-Hall, 2nd edn.

Lind, Douglas and Brian Marcus (1995). *An Introduction to Symbolic Dynamics and Coding*. Cambridge, England: Cambridge University Press.

Minsky, Marvin (1967). *Computation: Finite and Infinite Machines*. Englewood Cliffs, New Jersey: Prentice-Hall.