Monday, September 21, 2015
2:00 PM

Reading:
- Karlin and Taylor, Sections 2.1-2.3

Homework 1 due Friday, October 2 at 5 PM.
  ○ note, no office hours on Tuesday, September 29 or Thursday, October 1.

Equivalence of the Stochastic Update Rule and Probability Transition Matrix Formulation for FSDT Markov Chains

Some probability rules we will use in the proofs:

1) Computing probabilities involving conditioned random variables:
$$P(g(X,Y) = c|X = x) \qquad = \qquad P(g(x,Y) = c|X = x)$$
(plugging in known information about the random variable into the "unknown" expression)
Caveats:
- This actually can go wrong if X is continuous and g is not (Borel paradox)
- Don't drop the condition after you use it!

And of course this rule extends to the case when X and/or Y may be a collection of random variables.

2) If X and Y are independent random variables (or collections of random variables) then:
$$P(h(Y) = k|X = x) \qquad = \qquad P(h(Y) = k)$$

Also, $h(Y)$ is independent of $f(X)$ for any deterministic functions f, h.

3) In any probability rule, including the above, one can always self-consistently add a common condition to every term in the rule. So for example,
$$P(g(X,Y) = c|X = x, B) \qquad = \qquad P(g(x,Y) = c|X = x, B)$$
where B is any event, such as $B = \{Z = z\}$.

4) We will also find it useful to introduce the notion of an indicator function which relates an event A to a random variable 1_A.

Indicator function is a function on sample space:

$$1_A(\omega) = \begin{cases} 1 & \text{if } \omega \in A \end{cases}$$

$$1_A(\omega) = \begin{cases} 1 & \text{if} \quad \omega \in A \\ 0 & \text{if} \quad \omega \notin A \end{cases}$$

Let's use these ideas to show that a stochastic process defined in terms of a stochastic update rule has the Markov property and specifies the probability transition matrix(ces).

Given a collection of functions f_n (not random!) defining dynamics according to:
$X_{n+1} = f_n(X_n, Z_n)$ with iid Z_n

Let's check the Markov property:

$$P\left(X_{n+1} = j \mid X_0 = i_0, X_1 = i_1, \ldots, X_n = i_n\right)$$

$$= P\left(f_n(X_n, Z_n) = j \mid X_0 = i_0, X_1 = i_1, \ldots, X_n = i_n\right)$$

$$\text{plug in known info}$$

$$= P\left(f_n(i_n, Z_n) \mid X_0 = i_0, X_1 = i_1, \ldots, X_n = i_n\right)$$

Now we would like to assert that Z_n is independent of $X_0, X_1, \ldots X_n$. How do we show this?

Induction argument on $k = 0, 1, 2, \ldots, n$.

X_0 is independent of all the $\{Z_n\}_{n=0}^{\infty}$ by standard assumptions for initial condition of a Markov chain.

Now suppose we have that $X_0, X_1, \ldots, X_k$ are independent of Z_n, with $k \leq n - 1$. Then:
$X_{k+1} = f_k(X_k, Z_k)$ which is a function of random variables that are independent of Z_n by the induction hypothesis, and so X_{k+1} is also independent of Z_n. This concludes the proof by induction.

Use the independence rule:

$$P(X_{n+1} = j \mid X_0 = i_0, X_1 = i_1, \ldots, X_n = i_n)$$

$$= P(f_n(i_n, Z_n) = j \mid X_0 = i_0, X_1 = i_1, \ldots, X_n = i_n)$$

$$= P(f_n(i_n, Z_n) = j) \quad \text{✗}$$

independence of X_n, Z_n
in reverse

$$= P(f_n(i_n, Z_n) = j \mid X_n = i_n)$$

$$= P(f_n(X_n, Z_n) = j \mid X_n = i_n)$$

$$= P(X_{n+1} = j \mid X_n = i_n)$$

This establishes the Markov property.

How do we obtain the probability transition matrix from the stochastic update rule?

$P_{ij}^{(n)} = P(X_{n+1} = j \mid X_n = i) = P(f_n(i, Z_n) = j)$, which can be computed by knowing the probability distribution of Z_n.

Now let's show the reverse, that a FSDT Markov chain prescribed in terms of a probability transition matrix will be describable by a stochastic update rule.

The idea of this proof is essentially the idea behind simulating random variables on a finite state space.

A general random variable Y on a finite state space is described by a set of possible values $S_Y = \{y_1, \ldots, y_M\}$ with corresponding probabilities $p_1, p_2, \ldots, p_M$ such that $\sum_{j=1}^{M} p_j = 1$.

The basic pseudorandom generators on computers produce $U(0,1)$.
- see for example Pierre L'Ecuyer's webpage for suggestions

Other random variables are generated by mapping $U(0,1)$ in a suitable manner.
In particular for finite random variable Y, we can simulate it as follows:

$$Y = \sum_{j=1}^{M} y_j 1_{I_j}(U) = \begin{cases} y_1 \\ y_2 \\ \vdots \\ y_M \end{cases} \quad \text{if} \quad \begin{array}{l} U \in I_1 \\ U \in I_2 \\ \vdots \\ U \in I_M \end{array}$$

with $I_j = [\sum_{i=1}^{j-1} p_i, \sum_{i=1}^{j} p_i)$, which are a disjoint partition of the unit interval $[0,1)$.
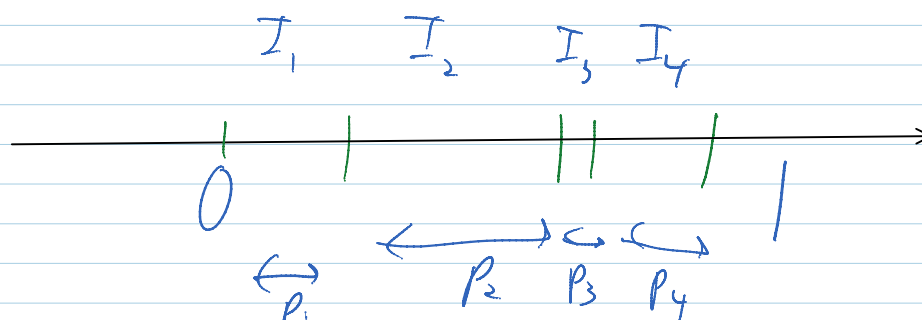
And we can directly use this idea to simulate FSDT Markov chains specified by a probability transition matrix because at each epoch n, we take the current state X_n, look at that row of the probability transition matrix $P^{(n)}$, which gives the probability distribution for the finite set of possible values of X_{n+1}. Apply the above algorithm to simulate that random variable.

And this simulation idea directly gives the mathematical proof for how to get a stochastic update rule, namely:

$$X_{n+1} = f_n(X_n, Z_n) \quad \text{with } Z_n \sim U(0,1)$$

with $f_n(i,z) = \sum_{j=1}^{M} j \, 1_{I_j}(z)$

where $I_j = [\sum_{j'=0}^{j-1} P_{ij'}^{(n)}, \sum_{j'=0}^{j} P_{ij'}^{(n)}]$.

That is, use the stochastic simulation idea with the probability distribution

$\{P_{ij}^{(n)}\}_{j=1}^{M}$ on the state space $S = \{1, 2, \ldots, M\}$.

That concludes the proof of equivalence of the stochastic update rule and probability transition matrix formulation.

In practice, a stochastic update rule can be simulated fairly directly once one knows how to simulate the Z_n. If a probability transition matrix(ces) $P^{(n)}$ are specified, then one can use the simulation scheme that was just described.

Examples of FSDT Markov chains

In these examples and in most of the rest of our consideration, we will focus on time-homogenous Markov chains which can be specified by a single stochastic update rule or single probability transition matrix

- $f_n \equiv f$
- $P^{(n)} \equiv P$

1) Two-state system ($M = 2$)

   The two states could model on/off, bound/unbound, active/inactive, sleep/wake, run/tumble

   Probability transition matrix:
   $$P = \begin{pmatrix} 1-p & p \\ q & 1-q \end{pmatrix}$$

   p is the probability to switch from state 1 to state 2 in an epoch.
   q is the probability to switch from state 2 to state 1 in an epoch.

   Stochastic update rule less natural.

2) Queueing models with maximum capacity M (Karlin and Taylor Sec. 2.2C)

   Let's consider a single server queue that handles one demand at a time.

Other demands that arrive are put into a queue to wait until the server is available. If the length of the queue (including the demand being served) exceeds M, then the demand is rejected.

Let's represent the state space S as the number of demands in the queue or in service:

$$S = \{0,1,2,\ldots,M\}$$

There are several ways we can formulate a discrete-time Markov chain model. Let's first consider taking small discrete time steps Δt, with X_n denoting the state of the system at time $n\Delta t$. For simplicity, we will assume Δt is small enough that during one epoch, only one of the following are possible:
- a new demand arrives with probability p
- service is completed with probability q
- nothing changes with probability $1 - p - q$

We're neglecting the possibility that two or more changes happen during one epoch.

Let's first write down a probability transition matrix for this model:

$$P = \begin{pmatrix}
 & 0 & 1 & 2 & \cdots & & M \\
0 & 1-p & p & 0 & \cdots & & 0 \\
1 & q & 1-p-q & p & & & \\
2 & 0 & q & 1-p-q & p & & 0 \\
\vdots & & & q & 1-p-q & & \\
 & & & & q & 1-p-q & p \\
M & 0 & 0 & & & q & 1-q
\end{pmatrix}$$

But we can also write a stochastic update rule:

$$X_{n+1} = \min\left(\left(X_n + Z_n\right)_{+}, M\right)$$

But we can also write a stochastic update rule:

$$X_{n+1} = \min\left(\left(X_n + Z_n\right)_+, M\right)$$

where Z_n iid, $Z_n = \begin{cases} 1 & \text{prob } p \\ 0 & 1-p-q \\ -1 & q \end{cases}$

$$(x)_+ = \max(x, 0)$$