EC 232 TA Session 6
Markov Processes

Feng Lin

11/08/2021

Definition of Markov Processes

A (first-order) Markov process is a stochastic process with the property

$$\Pr(X_{t+1} \in A | X_t, \ldots, X_{t-s}) = \Pr(X_{t+1} \in A | X_t),$$

for $t = 2, 3, \ldots; s = 1, 2, \ldots; A$ is, loosely speaking, some arbitrary set of values that X can take.

- Why do we care about Markov processes?
  - It is useful to model some stochastic process that depends only on the most recent history. For instance, we may want to assume that realized TFP has a random component, but otherwise only depends deterministically on TFP today.
  - It is mathematically convenient. We could argue that they provide a good balance between being able to nest somewhat complicated dynamics, and tractable analysis.

We will focus on the discrete Markov processes in this session. You will see this in operation in Lucas Tree toward the end of this quarter.

Recap: Conversion to First-Order Markov

Prof. Stokey stated in class that a Markov process of any finite order can be written as a first-order process by expanding the state space.

Example: If the distribution of X_{t+1} depends on $X_t, \ldots, X_{t-1}$, then we can define a new state as $\tilde{X}_t = (X_t, \ldots, X_{t-1})$ and use this new state instead.

Transition Function

A transition function gives you the probability that the next period's X lies in the set A, given that the current shock is x:

$$Q(A; x) = \Pr\left(X_{t+1} \in A | X_t = x\right).$$

In the discrete case, we have a transition matrix

$$Q = [q_{ij}],$$

where q_{ij} represents the probability of transitioning from j given current state is i (each row sums to 1).

From Distribution to Distribution

If we know current state, it would be straightforward to get the distribution for the next state conditional on current state using the transition function. However, many times we only know the distribution of the current state and want to learn about the distribution of the next state.

Let $\Phi_t(x)$ be the CDF and $\phi_t(x)$ be the PDF of state at time t, we have

$$\Phi_{t+1}(x) = \int Q(x; \tilde{x})\phi_t(\tilde{x})d\tilde{x}.$$

In the discrete case, let Φ_t (a row vector) be the probability of each state, and we have

$$\Phi_{t+1} = \Phi_t Q$$

- Applying this repeatedly, we have $\Phi_{t+s} = \Phi_t Q^s$.

Property: Stationary Distribution

We say that Φ is a stationary distribution if

$$\Phi(x) = \int Q(x; \tilde{x})\phi(\tilde{x})d\tilde{x};$$

and in the discrete case (where Φ denote a row vector),

$$\Phi = \Phi Q.$$

If the probability distribution over the initial state is Φ, then it is also Φ in every successive period.

Property: Ergodic Set

Let S denote the set of discrete states. A set $E \subseteq S$ is called an ergodic set if $\Pr(E|s_i) = 1$ for $s_i \in E$, and if no proper subset of E has this property. (RMED p.321)

- Less strictly speaking, it is the smallest set such that if you start from the set you will never leave the set. Note that there can be multiple ergodic sets.
- We call a process ergodic if it has only one ergodic set.

Example:

$$\begin{bmatrix} \frac{1}{2} & \frac{1}{2} & 0 & 0 \\ \frac{1}{2} & \frac{1}{2} & 0 & 0 \\ 0 & 0 & \frac{1}{3} & \frac{2}{3} \\ 0 & 0 & \frac{1}{3} & \frac{2}{3} \end{bmatrix} \qquad \begin{bmatrix} 0 & 0 & \frac{1}{3} & \frac{2}{3} \\ 0 & 0 & \frac{1}{3} & \frac{2}{3} \\ \frac{1}{2} & \frac{1}{2} & 0 & 0 \\ \frac{1}{2} & \frac{1}{2} & 0 & 0 \end{bmatrix}$$

Property: Transient State

A state is called transient if there is a positive probability of leaving and never returning. (RMED p.322)

Example:

$$\begin{bmatrix} \frac{1}{2} & \frac{1}{4} & \frac{1}{4} \\ 0 & \frac{1}{3} & \frac{2}{3} \\ 0 & \frac{1}{3} & \frac{2}{3} \end{bmatrix}$$

Property: Cycle

A cycle is a set of subsets such that the states cycle through the subsets with probability 1. (Did not find the exact reference in RMED; cyclically moving subsets, RMED p.323)

Example:

$$\begin{bmatrix} 0 & 0 & \frac{1}{3} & \frac{2}{3} \\ 0 & 0 & \frac{1}{3} & \frac{2}{3} \\ \frac{1}{2} & \frac{1}{2} & 0 & 0 \\ \frac{1}{2} & \frac{1}{2} & 0 & 0 \end{bmatrix}$$

One More Comment on Formulating DP

In general, it matters whether we make a choice before or after information arrives in each period.

- If the choice must be made before information (shock) arrives, then there should in general be an expectation inside the optimization problem (i.e. within the max or min).
- If the choice is made after the information (shock) arrives, the the agent takes into account that, for any given realization, the optimal choice will be made, so the max will be inside the expectation.