Markov Processes

Chase Abram*

September 1, 2022

1 Intuition and Motivation

Markov processes[1] are "forgetful", in that they only need the most recent history to know the distribution of the next draw. We often find these useful in macro, for when we want to model that something depends only on where it was yesterday, not the day before. Maybe realized TFP has a random component, but otherwise only depends deterministically on TFP today. Maybe we view agents decisions in equilibrium as only depending on their state today, not states from earlier.

To be transparent, we are also interested in these because they are analytically nice, as we'll see below. We could argue that they provide a good balance between being able to nest somewhat complicated dynamics, and tractable analytics. If you're skeptical, I'd encourage you to go find examples where we need more than just yesterday's knowledge to learn about the randomness today. The cases certainly exist, but it's not so easy to argue that a Markov approximation is necessarily bad.

2 Basics

Mathematically, a **Markov process** satisfies

$$P[X_t \in A \mid X_{t-1}, X_{t-2}, \ldots] = P[X_t \in A \mid X_{t-1}]$$

We'll focus on finite, discrete Markov processes. Therefore we can model the transition probabilities with a matrix. Consider the following:

$$\begin{bmatrix} 0.2 & 0.8 & 0 \\ 0 & 0.9 & 0.1 \\ 1.0 & 0 & 0 \end{bmatrix}$$

The (i, j) entry tells the probability of moving to state j, given the current state is i.[2] So there is a 0.8 probability of moving to state 2, given that you are currently at state 1. If you are in state 3, you will always move to state 1.

*Please send corrections to abram@uchicago.edu.

[1] I here mean first-order. You can make any order process a first-order by a state re-definition anyway, similarly how all VARs are VAR(1)s if you stack right.

[2] Some definitions use the transpose of this. As long as you keep straight what's going on, it's not an issue.

Armed with this approach alone, we can start defining other things. Let a Markov matrix be Q. Perhaps most importantly, the stationary distribution of a Markov process is a vector π satisfying $\pi' Q = \pi'$. This is a (left) eigenvector of Q with eigenvalue 1, and the interpretation is that if you have some distribution over the states, then have the whole distribution transition, the new distribution is the same.

Connected to this idea, Markov processes make it really easy to consider multiperiod transition. Q^2 has entries which the probability of moving from i to j in exactly 2 periods. So if I wanted to know the probability of moving from i to j in exactly 57 periods, this problem is a priori tricky, but with a Markov process we just consider the (i, j) entry of Q^{57}.

Powers of Q are also useful for finding the stationary distribution. Note that π' is a fixed point of Q. It turns out we can find (the unique) π by considering

$$\lim_{n\to\infty} \pi_0' Q^n$$

where π_0 is any distribution, provided the process is ergodic, meaning the only ergodic set is the entire sample space.

The simple idea behind the proof for this statement is that we can consider the eigenvalues of Q. We will find that one eigenvalue is always 1, and the other are ≤ 1. If only one eigenvalue is 1, then distributions must converge towards the corresponding eigenvector as Q is applied, and this will happen when the operator is ergodic.

3 Ergodicity

Ergodicity is intuitively the idea that you can start from anywhere, and if you transition enough, you can get anywhere. For example, the following is obviously ergodic

$$\begin{bmatrix} 0.2 & 0.7 & 0.1 \\ 0.3 & 0.3 & 0.4 \\ 0.4 & 0.5 & 0.1 \end{bmatrix}$$

because it is immediately possible to get anywhere from anywhere. But this is also ergodic

$$\begin{bmatrix} 0.2 & 0.8 & 0 \\ 0 & 0.9 & 0.1 \\ 0.7 & 0 & 0.3 \end{bmatrix}$$

I can get from 1 to 3 in 2 steps, 2 to 1 in 2 steps, and 3 to 2 in 2 steps. Or note that the square of the above matrix has all positive entries.

An **ergodic set** is such that all points in the set can be reached from anywhere else in the set, and no smaller subset satisfies this propery. Formally, a set A is ergodic if both the following hold.

(i) The probability of exiting A is zero, given that you start in A.

(ii) There does not exist a proper subset $B \subsetneq A$ such that the above property holds.

Note that we call a process ergodic if it only has one ergodic set, and that set is the whole space. Consider the following

$$\begin{bmatrix} 0.2 & 0.8 & 0 \\ 0.8 & 0.2 & 0 \\ 0 & 0 & 1.0 \end{bmatrix}$$

There are 2 ergodic sets: $\{1, 2\}$ and $\{3\}$. Note that the stationary distribution is no longer unique, since both $\pi = [0, 0, 1]$ and $\pi = [0.5, 0.5, 0]$ work[3]. The takeaway: if you start in an ergodic set, you never get to leave (with positive probability)!

What then to do with states not in an ergodic set? They are **transient**. Formally a state is transient if the probability of exiting and never returning is positive. For any state space, we can partition the states into a set of ergodic sets, and a set of transient states. Consider the following matrix

$$\begin{bmatrix} 0.5 & 0.2 & 0.3 \\ 0 & 0.2 & 0.8 \\ 0 & 0.7 & 0.3 \end{bmatrix}$$

The first state is transient, because if the process ever leaves it, then it will never return. Note this has nothing to do with how likely this event is to happen, since the following example satisfies the same property, though if one starts in state 1, it will take a long time to ever exit.

$$\begin{bmatrix} 1 - 10^{-10} & 10^{-10} & 0 \\ 0 & 0.2 & 0.8 \\ 0 & 0.7 & 0.3 \end{bmatrix}$$

4 Cycles

Lastly, we might care about cycles. A set of sets $\{A_i\}$ is a **cycle** if there is probability 1 that that state cycles through $A_1 \to A_2 \to \cdots \to A_n \to A_1$, given that it starts in A_1. Note that the sets may be singletons, and the union of each cycle forms an ergodic set.

5 Problems

5.1 Simple Erogodic

Consider the following Markov transition matrix

$$A = \begin{bmatrix} 0.1 & 0.5 & 0.4 \\ 0.3 & 0.7 & 0 \\ 0 & 0.2 & 0.8 \end{bmatrix}$$

(i) (Understanding check) What is the probability of going from state 2 to state 1 in one step?

(ii) What is the probability of being in state 1 after exactly two steps, given that you start in state 2?

[3] These form the extremal points of a convex hull of distributions that are all stationary.

(iii) What about exactly 100 steps?

(iv) What about exactly 1000 steps?

(v) Numerically, what do you find as the transition matrix for the first n steps, as $n \to \infty$? What can you say about how the initial distribution matters?

5.2 Random Walk

Consider the following biased random walk with reflection on the boundaries.

$$B = \begin{bmatrix} b & 1-b & 0 & \cdots & & & 0 \\ b & 0 & 1-b & 0 & \cdots & & 0 \\ 0 & b & 0 & 1-b & 0 & \cdots & 0 \\ \vdots & & \ddots & & & & \vdots \\ 0 & \cdots & & b & 0 & 1-b & 0 \\ 0 & \cdots & & & b & 0 & 1-b \\ 0 & \cdots & & & & b & 1-b \end{bmatrix}$$

(i) What is the stationary distribution when $b = \frac{1}{2}$?

(ii) What if $b = 0.6$?

(iii) Try $b = 0.51$

(iv) One common criticism in macro is that we focus too much on details that make only a tiny difference for each agent, so how could these details possibly make a big difference in the macroeconomy? If we take the above Markov process to be some reduced-form model of general equilibrium for something of interest, what can we conclude about the effects of small amounts of bias ($b = 0.5 \pm \epsilon$) on the equilibrium distribution?

(v) Evaluate the following statement: "If a change in assumptions has a small impact on household choices, then it will also have a small impact on the implied general equilibrium."

(vi) Evaluate the following statement: "If a change in assumptions has a large impact on household choices, then it will also have a large impact on the implied general equilibrium."

6 Continuous time

Increasingly, we want to model in continuous time. The above formulation works well for discrete time models, since there is a clear arrival time (once per period), and it is clear how we transition. In continuous time, we instead allow for Poisson rate of arrival, call it λ, and consider the flow between states, conditional on arrival, call this matrix P. In discrete time, the transition matrix operated on a given distribution, so we had[4]

$$\pi_{t+1} = Q' \pi_t$$

[4]Switching to left eigenvector notation because I am more comfortable.

and the stationary distribution then satisfies $\bar{\pi} = Q'\bar{\pi}$. In continuous time, the operator tells how mass should flow from each point to each other point, so we have a differential equation.

$$\dot{\pi}(t) = P'\pi(t)$$

and the stationary distribution the satisfies $0 = P'\bar{\pi}$. The solution is then $\pi(t) = \exp(tP')\pi(0)$, and if the process is ergodic, π will converge to its stationary distribution and $\pi(0)$ will cease to matter as $t \to \infty$.

We may convert between continuous and discrete time using the following, where Q is the discrete time transition matrix, I is the identity, λ is the rate of shock arrival, and P is the continuous time infinitesimal generator (the "transition matrix" for continuous time).

$$P = \lambda(Q - I)$$