# ECN 200D—Week 1 Lecture Notes
# Markov Chains

William M Volckmann II

April 4, 2017

## 1 Stochastic Growth Model

Consider a general growth model,

$$V(k) = \max_{k' \in \Gamma(k)} \left\{ F(k, k') + \beta E\big[V(k')\big] \right\}.$$

This model is deterministic. We will want to consider a similar model in a stochastic environment. Let $z_t$ denote exogenous shocks. We will suppose that $z_t$ is countable in outcomes and, for now, finite. Then we can write the growth model as

$$V(k, z) = \max_{k' \in \Gamma(k,z)} \left\{ F(k, k', z) + \beta E\big[V(k', z')|z\big] \right\}.$$

Since $z$ is random, we'll need to impose some structure on the nature of its randomness. In our case, we will consider Markov randomness.

## 2 Markov Chains

**Definition 1.** A stochastic process $\{x_t\}$ is said to have the **Markov property** if for all $k \geq 1$ and all $t$,

$$\mathbb{P}(x_{t+1}|x_t, x_{t-1}, \ldots, x_{t-k}) = \mathbb{P}(x_{t+1}|x_t).$$

So probabilistically, the outcome of tomorrow's state only depends on today's state, not on the entire history of states. We will assume the Markov property and characterize the process by a **Markov chain**.

A **time-invariant** Markov chain consists of three things.

**(a)** An $n$-dimensional state space (i.e. possible realizations of $x_t$) consisting of canonical vectors $e_i$, $i = 1, \ldots, n$. For instance, if the third state is realized, then $x_t = e_3$, which has a 1 in the 3rd position:

$$x_t = e_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}.$$

**(b)** An $n \times n$ **transition matrix** $P$, which records the probabilities of moving from one value of the state to another in one period. For instance, the $ij$th element $P_{ij}$ is the probability of moving from state $i$ today to state $j$ tomorrow. In other words,

$$P_{ij} = \mathbb{P}(x_{t+1} = e_j | x_t = e_i).$$

**(c)** An $n \times 1$ initial probability distribution $\pi_0$ whose $i$th element is the probability of being in state $i$ at time zero. So the $i$th row of $\pi_0$ is

$\pi_{0i} = \mathbb{P}(x_0 = e_i)$.

Suppose we are in state $x_t = e_i$. Then the probability that we move to *some* state $j$ tomorrow must sum to 1,

$$\sum_{j=1}^{n} P_{ij} = 1.$$

Otherwise there's some probability that we'd just kind of stop. A matrix $P$ that satisfies this condition is called a **stochastic matrix**. Similarly, the probability that there is *some* initial state must be one:

$$\sum_{i=1}^{n} \pi_{0i} = 1.$$

Otherwise there's some probability that we'd never even start.

**Example 1.** Let's consider the probability of moving from state 1 to state 2 in *two* periods. Suppose that there are $n = 2$ possible states. We need to go from state 1 to some intermediate state $h$, and then from $h$ to 2. The relevant probabilities are

$$\begin{aligned}\mathbb{P}(x_{t+2} = e_2 | x_t = e_1) = &\mathbb{P}(x_{t+1} = e_1 | x_t = e_1)\mathbb{P}(x_{t+2} = e_2 | x_{t+1} = e_1) \\ &+ \mathbb{P}(x_{t+1} = e_2 | x_t = e_1)\mathbb{P}(x_{t+2} = e_2 | x_{t+1} = e_2) \\ =&P_{11}P_{12} + P_{12}P_{22}.\end{aligned}$$

The stochastic matrix is

$$P = \begin{bmatrix} P_{11} & P_{12} \\ P_{21} & P_{22}. \end{bmatrix} \quad \Longrightarrow \quad P^2 = \begin{bmatrix} P_{11}P_{11} + P_{12}P_{21} & P_{11}P_{12} + P_{12}P_{22} \\ P_{21}P_{11} + P_{22}P_{21} & P_{21}P_{12} + P_{22}P_{22} \end{bmatrix}.$$

Notice that $P^2_{12}$ is exactly what we want. More generally, the probability of moving from $i$ to $j$ in $k$ periods is determined by

$$\mathbb{P}(x_{t+k} = e_j | x_t = e_i) = P_{ij}^k. \qquad \blacksquare$$

**Example 2.** Let the state variable $z_t$ denote total factor productivity in period $t$. We want to forecast it. GDP growth $y_t$ can be either in a boom or a bust, where $e_1$ indicates that we're in a boom, $e_2$ a recession. The boom state has $y_1 = 1.2$ and the recession state has $y_t = -0.4$.

Let $\bar{y}$ be a $2 \times 1$ vector of outcomes for GDP growth,

$$\bar{y} = \begin{bmatrix} 1.2 \\ -0.4 \end{bmatrix}.$$

Then the realization $y_t$ will be given by

$$y_t = \bar{y}' x_t.$$

For instance, if it turns out that we get a boom state, i.e. $x_t = e_1$, then

$$y_t = \begin{bmatrix} 1.2 & -0.4 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} = 1.2.$$

Now notice that

$$\begin{aligned} E[x_{t+1} | x_t = e_i] &= e_1 \mathbb{P}(x_{t+1} = e_1 | x_t = e_i) + e_2 \mathbb{P}(x_{t+1} = e_2 | x_t = e_i) \\ &= e_1 P_{i1} + e_2 P_{i2} \\ &= P' e_i \\ &= \begin{bmatrix} P_{i1} \\ P_{i2} \end{bmatrix} \\ &= P_i'. \end{aligned}$$

It follows that

$$\begin{aligned} E[y_{t+1}|x_t = e_i] &= E[\bar{y}' x_{t+1}|x_t = e_i] \\ &= \bar{y}' E[x_{t+1}|x_t = e_i] \\ &= \bar{y}' P_i'. \end{aligned}$$

More generally,

$$E[y_{t+k}|x_t = e_i] = \bar{y}' P^K e_i. \qquad \blacksquare$$

**Example 3** (Unconditional Probabilities)**.** Let $\pi_t' = \mathbb{P}(x_t)$ be the $1 \times n$ vector whose $i$th element is $\mathbb{P}(x_t = e_i)$. Consider $\pi_{1,1}' = \mathbb{P}(x_1 = e_1)$. We can write this as

$$\pi_{1,1}' = \mathbb{P}(x_0 = e_1) P_{11} + \mathbb{P}(x_0 = e_2) P_{21} = [\pi_0' P]_1.$$

More generally, we can write

$$\pi_k' = \pi_0' P^k. \qquad \blacksquare$$

Finally, notice that

$$\pi_{k+1}' = \pi_0' P^{k+1} = [\pi_0' P^k] P = \pi_k' P.$$

# 3 Stationary Distributions

We have just seen that the unconditional probability distributions evolve according to the law of motion

$$\pi_{t+1}' = \pi_t' P.$$

An unconditional distribution is called **stationary** if it satisfies $\pi_{t+1} = \pi_t$, that is, if the unconditional distribution remains unaltered with the passage of time. This implies that a stationary distribution must satisfy $\pi' = \pi' P$, from which we can write

$$(I - P')\pi = 0.$$

This looks awfully eigenvectory. As long as $P$ is a stochastic matrix, we are guaranteed at least one (unit) eigenvalue and that there is at least one satisfying eigenvector $\pi$. However, uniqueness is not guaranteed.

**Example 4.** Consider the stochastic matrix

$$P = \begin{bmatrix} p & 1-p \\ 1-p & p \end{bmatrix}.$$

Suppose $\Pi' = [\pi \;\; 1-\pi]$. Then we have a stationary distribution if

$$[\pi \quad 1-\pi] \begin{bmatrix} p & 1-p \\ 1-p & p \end{bmatrix} = [\pi \quad 1-\pi].$$

Multiplying out the left hand side, this condition holds if

$$\begin{aligned} &\pi p + (1-\pi)(1-p) = \pi, \\ &\pi(1-p) + (1-\pi)p = 1-\pi. \end{aligned}$$

Subtract the top line from the bottom line and we get

$$\pi(1-pP) - (1-\pi) = 1 - 2\pi \quad \Longrightarrow \quad 1-p = 2\pi(1-p).$$

Notice that if $p = 1$, then $\pi$ can be anything. In this case, we have

$$P = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix},$$

which acts like a "sink" because we always just go to the same state. However, if $p \neq 1$, then we can solve for the unique $\pi = 1/2$. ■

## 4 Asymptotic Stationarity

Given $\pi_0$, we might want to know whether $\pi_t$ approaches a stationary distribution over time. That is, we want to test whether

$$\lim_{t\to\infty} \pi_0 P^t = \pi_\infty,$$

where $\pi_\infty$ solves $(1 - P')\pi = 0$. If for all $\pi_0$ we have $\pi_0 P^t$ limiting to $\pi_\infty$, then we say that the Markov chain is **asymptotically stationary** with a **unique invariant distribution**. This will be the case if from every state there exists a positive probability of moving to any other state (in one or more steps). In other words, if there is no situation where it becomes impossible to eventually reach state $j$ from $i$.

**Definition 2.** A set $E$ is called an **ergodic set** if

$$\mathbb{P}(x_t \in E | x_{t-1} \in E) = 1,$$

and no proper subset of $E$ has this property.

In other words, we will always go from something in $E$ to something else in $E$. And furthermore, there is no smaller subset $E' \subset E$ where if we go into $E'$, then we also stay in $E'$. (Which means we'll end up cycling through *all* elements of $E$.)

## 5 Continuous State Spaces

An example of a stochastic process with a continuous state space is

$$\ln(z_{t+1}) = \rho \ln(z_t) + \epsilon_t.$$

If $\epsilon$ is i.i.d., then $\ln(z_t)$ follows a Markov process–the conditional expectation will depend only on the last realization of the process.

In terms of computation, it is useful to discretize the continuous states. Pick some extreme values for the process, e.g. three standard deviations from the mean, to set the bounds. Then chop up that interval in to equal-sized parts.

In the case of a Markov chain with finite $n$, we can conveniently write

$$V(k, z_i) = \max_{k' \in \Gamma(k, z_i)} F(k, k', z_i) + \beta \sum_{j=1}^{n} P_{ij} V(k', z_j).$$

So the expected value is just the nice little sum we're used to seeing.