

MARKOV DYNAMICS

QUANTITATIVE ECONOMICS 2023

Piotr Żoch

November 20, 2023

MOTIVATION

- The assumption that wage offers in the job search model are i.i.d. was very restrictive.
- We want to develop tools that allow us to tackle more general problems.
- For example, we want to have some persistence in the wage process.
- We focus on an important class of processes: distribution of the future state depends on the current state (and not on the past).
- Markov chains.

PLAN

- Introduction to Markov chains.
- Finite state approximation.

MARKOV CHAINS

MARKOV CHAINS

- Let $X = \{x_1, \dots, x_n\}$ be a finite set of states.
- We will consider random processes taking values in X .
- We call X the **state space** of the process.

MARKOV CHAINS

- Let $(X_t) := (X_t)_{t \geq 0}$ be a sequence of random variables taking values in X .
- An operator $P \in \mathcal{L}(\mathbb{R}^X)$ is called a **Markov operator** on X if:
 1. P is positive;
 2. $P\mathbb{1} = \mathbb{1}$.
- Let $\mathcal{M}(\mathbb{R}^X)$ be the set of all Markov operators on X .
- $P \in \mathcal{M}(\mathbb{R}^X)$ if and only if:
 1. $P(x, x') \geq 0$ for all $x, x' \in X$;
 2. $\sum_{x' \in X} P(x, x') = 1$ for all $x \in X$.
- (X_t) is a **Markov chain** on X if there exists a $P \in \mathcal{M}(\mathbb{R}^X)$ such that

$$\mathbb{P}(X_{t+1} = x' \mid X_0, X_1, \dots, X_t) = P(X_t, x') \quad \text{for all } x' \in X.$$

MARKOV CHAINS

- This condition says:
 1. Transition from X_t to X_{t+1} **does not** require any information about the past.
 2. P contains **all** information needed to go from X_t to X_{t+1} .
- We often call (X_t) **P-Markov** if it satisfies this condition.
- We often call P the **transition matrix** of the Markov chain.
- X_0 or its distribution ψ_0 is the **initial condition** of (X_t) .
- Sometimes we write

$$\mathbb{P} \left(X_{t+1} = x^j \mid X_t = x_i \right) = P(x_i, x_j) = P_{ij}.$$

MARKOV CHAIN SIMULATION

Algorithm Simulate MC

```
1: procedure MCSIM
2:    $t \leftarrow 1$ 
3:   Draw  $X_1$  from  $\psi_1$ 
4:   while  $t < \infty$  do
5:     Draw  $X_{t+1}$  from  $P(X_t, \cdot)$ 
6:      $t \leftarrow t + 1$ 
7:   end while
8: end procedure
```

Note: we start from 1, not 0, to be consistent with the notation we use in the code.

EXAMPLE: INVENTORY DYNAMICS

- **S-s** dynamics.
- A firm starts period t with inventory X_t .
- If $X_t \leq s$, the firm orders S units of the good.
- There is i.i.d. demand D_t with distribution $\phi \in \mathcal{D}(\mathbb{Z}_+)$.
- Dynamics of X_t are described by

$$X_{t+1} = \max \{X_t - D_t, 0\} + S \cdot \mathbb{1} \{X_t \leq s\}.$$

- Observe that

$$\mathbb{P}(X_{t+1} = x' \mid X_0, X_1, \dots, X_t) = P(X_t, x') \quad \text{for all } x' \in \{0, 1, \dots, s + S\}.$$

- $X = \{0, 1, \dots, s + S\}$ is the **state space**. **Goal**: find P .

EXAMPLE: INVENTORY DYNAMICS

- Let $h(x, d) = \max \{x - d, 0\} + S \cdot \mathbb{1} \{x \leq s\}$.
- We have $X_{t+1} = h(X_t, D_t)$.
- This means that

$$P(x, x') = \mathbb{P}(h(x, D_t) = x') = \sum_d \mathbb{1} \{h(x, d) = x'\} \phi(d)$$

for all $(x, x') \in X \times X$.

- (X_t) is a Markov chain on $\{0, 1, \dots, s + S\}$ with transition matrix P .
- Note: we assumed $X_0 \leq S + s$.

EXAMPLE: JOB SEARCH

- Return to the McCall job search model. w^* is the reservation wage.
- State $X \in \{0, 1\}$ - unemployed (0) or employed (1) at the beginning of period t .
- We have $\mathbb{P}(X_{t+1} = 1 \mid X_t = 1) = 1$ and $\mathbb{P}(X_{t+1} = 0 \mid X_t = 1) = 0$ - if employed, stay employed.
- $\mathbb{P}(X_{t+1} = 1 \mid X_t = 0)$ equals the probability of accepting a job offer, $\mathbb{P}(W_t \geq w^*)$.
- $\mathbb{P}(X_{t+1} = 0 \mid X_t = 0)$ equals the probability of declining a job offer, $\mathbb{P}(W_t < w^*)$.
- How to find P ?

TAKEAWAYS

- If exogenous states are Markov, optimal policies in dynamic problems usually induce Markov chains for the state variables.
- Instead of simulating the whole sequence of states, we can use properties of Markov chains to calculate expected values and other statistics of interest.
- In the first example: use the P matrix to get the expected time between orders or the probability of stockout for any inventory level.
- Much faster and more accurate than simulation.
- This is especially important when we have to solve the problem many times (e.g. in estimation).
- We now describe these useful properties of Markov chains.

IRREDUCIBILITY

- Let P^k be the k -th power of P .
- This is a **k-step transition matrix** of the Markov chain.
- $P^k(x, x')$ is the probability of going from x to x' in k steps.
- P is **irreducible** if and only if $\sum_{k=0} P^k \gg 0$.
- This is equivalent to: for any $x, x' \in X$ there exists k such that $P^k(x, x') > 0$ for all $x, x' \in X$.
- P -Markov chain eventually visits any state from any other state with positive probability.

STATIONARITY

- Let ψ_t be the distribution of X_t . k -th power of P .
- For a P -Markov chain (X_t) we have

$$\psi_{t+1}(x') = \sum_{x \in X} P(x, x') \psi_t(x) \quad \text{for all } x' \in X \text{ and } t \geq 0.$$

- If we treat ψ_t as a **row** vector, we can write this as

$$\psi_{t+1} = \psi_t P.$$

- This is a deterministic linear difference equation in $\mathcal{D}(X)$.

STATIONARITY

- If (X_t) is a P -Markov chain with $X_0 \sim \psi_0$ then

$$X_t \sim \psi_0 P^t \quad \text{for all } t \geq 0.$$

- We can calculate $\mathbb{E}(f(X_t))$ as

$$\mathbb{E}(f(X_t)) = \psi_0 P^t f.$$

where f is a column vector, $f \in \mathbb{R}^X$.

- Note: \mathbb{E} is the unconditional expectation, not the conditional expectation given X_0 .

STATIONARITY

- $\psi^* \in \mathcal{D}(X)$ is a **stationary distribution** for P if

$$\psi^*(x') = \sum_{x \in X} P(x, x') \psi^*(x) \quad \text{for all } x' \in X.$$

- We can also write this as

$$\psi^* = \psi^* P.$$

- If ψ^* is stationary and $X_t \sim \psi^*$ then $X_{t+k} \sim \psi^*$ for all $k \geq 0$.
- Every P -Markov chain has at **least one** stationary distribution.

STATIONARITY

- Every **irreducible** $P \in \mathcal{M}(\mathbb{R}^X)$ has a **unique** stationary distribution with $\psi^*(x) > 0$ for all $x \in X$.
- If P is irreducible with stationary distribution ψ^* , then, for any P -Markov chain (X_t) and any $x \in X$, we have

$$\mathbb{P} \left\{ \lim_{k \rightarrow \infty} \frac{1}{k} \sum_{t=0}^{k-1} \mathbb{1}\{X_t = x\} = \psi^*(x) \right\} = 1$$

- This property is called **ergodicity**.
- It says that the fraction of time spent in state x converges to the probability of being in state x in the stationary distribution.

CONDITIONAL EXPECTATIONS

- We are often interested in calculating $\mathbb{E}(f(X_{t+1}) \mid X_t = x)$.
- For a P -Markov chain we have

$$\mathbb{E}(f(X_{t+1}) \mid X_t = x) = \sum_{x' \in X} P(x, x') f(x') = P f.$$

- It follows immediately that

$$\mathbb{E}(f(X_{t+k}) \mid X_t = x) = \sum_{x' \in X} P^k(x, x') f(x') = P^k f.$$

- Note: we often see notation \mathbb{E}_t or \mathbb{E}_{X_t} . This is the conditional expectation given X_t . The notation \mathbb{E}_t implicitly assumes we already observed X_t .

CONDITIONAL EXPECTATIONS

- The **law of iterated expectations** is

$$\mathbb{E}(\mathbb{E}(Y | X)) = \mathbb{E}(Y).$$

for random variables X and Y .

- We often have to calculate $\mathbb{E}_t \mathbb{E}_{t+1} \mathbb{E}_{t+2} \cdots$ or $\mathbb{E}(\mathbb{E}_{t+k}) \cdots$.
- Recall that

$$\mathbb{E}_t f(X_{t+k}) = \sum_{x' \in X} P^k(x, x') f(x') = P^k f$$

- It is a function of X_t , $h(X_t) := \mathbb{E}_t f(X_{t+k})$.
- We thus have

$$\mathbb{E}(\mathbb{E}_t f(X_{t+k})) = \mathbb{E}(h(X_t)) = \psi_0 P^t P^k f.$$

APPROXIMATION

APPROXIMATION

- We often start with a **continuous** state Markov models.
- For example, consider a linear Gaussian AR(1) model

$$X_{t+1} = \rho X_t + b + \nu \epsilon_t, \quad \epsilon_t \stackrel{iid}{\sim} N(0, 1), |\rho| < 1.$$

- The model has a unique stationary distribution

$$\psi^* = N\left(\mu_x, \sigma_x^2\right), \quad \mu_x := \frac{b}{1-\rho}, \quad \sigma_x^2 := \frac{\nu^2}{1-\rho^2}.$$

- **Goal:** find a discrete approximation.
- This means finding a grid X and a transition matrix P such that ψ^* is **close** to the stationary distribution of the Markov chain with transition matrix P .

APPROXIMATION

- Once we do it, we can work with the discrete approximation instead of the continuous model.
- There are several ways to do it. We will discuss two of them:
 1. Tauchen method;
 2. Rouwenhurst method.

TAUCHEN METHOD

- We start by creating the grid X .
- For now keep $b = 0$.
- Pick a number $m > 0$ and an integer N .
- Set $x_1 = -m\sigma_X$ and $x_N = m\sigma_X$.
- Set $x_i = x_1 + \frac{i-1}{N-1} (x_N - x_1)$ for $i = 2, \dots, N - 1$.
- This is our state space X .

TAUCHEN METHOD

- Now we need to find the transition matrix P . Still keep $b = 0$ for now.
- The idea is to discretize the **conditional distribution** for each point on the grid.
- Let F be the CDF of $N(0, \nu^2)$. We have

$$\mathbb{P}(t - \delta < X_{t+1} \leq t + \delta \mid X_t = x) = F(t + \rho x + \delta) - F(t + \rho x - \delta)$$

- Use this to get, for $j \in \{2, N-1\}$

$$P(x_i, x_j) = F(x_j - \rho x_i + s/2) - F(x_j - \rho x_j - s/2), \quad s := \frac{x_N - x_1}{N-1}.$$

- We also have

$$P(x_i, x_1) = F(x_1 - \rho x_i + s/2), \quad P(x_i, x_N) = 1 - F(x_N - \rho x_i - s/2).$$

TAUCHEN METHOD

- Finally, if $b \neq 0$, we simply shift each x_i by μ_x .
- Usually we pick m equal to three. Our discrete state space is three times the standard deviation of the stationary distribution in each direction.
- If $\rho \approx 1$ in AR(1) process, it is difficult to obtain a discrete approximation that matches both the unconditional variance and the first-order auto-correlation, unless the number of grid points is large.
- A wide domain needed because realizations can be far from the unconditional mean.
- Many points needed because the conditional distributions are centered around the current realizations.

ROUWENHURST METHOD

- Despite these problems, the Tauchen method is still widely used – especially with low persistence ρ .
- There is a **better** alternative: the Rouwenhurst method guarantees to match these 2 moments even with a small number of points.

ROUWENHURST METHOD

- Use a symmetric and evenly spaced state space with points on $[-A, A]$.
 - For $N = 2$ (two points) use $Z_2 = \begin{bmatrix} p & 1-p \\ 1-q & q \end{bmatrix}$
 - For $N > 2$ construct the $N \times N$ matrix Z_N recursively

$$Z_N = p \begin{bmatrix} Z_{N-1} & \mathbf{0} \\ \mathbf{0}' & 0 \end{bmatrix} + (1-p) \begin{bmatrix} \mathbf{0} & Z_{N-1} \\ 0 & \mathbf{0}' \end{bmatrix} \\ + (1-q) \begin{bmatrix} \mathbf{0}' & 0 \\ Z_{N-1} & \mathbf{0} \end{bmatrix} + q \begin{bmatrix} 0 & \mathbf{0}' \\ \mathbf{0} & Z_{N-1} \end{bmatrix}$$

- Divide all but the top and the bottom rows by 2
- To match the unconditional mean, unconditional variance, and first-order auto-correlation set

$$p = q = \frac{1 + \rho}{2}, \quad A = \sqrt{N-1} \nu$$