Finite State Markov Decision Processes

T. Kam

Outline

1. **Motivation**
2. **Markov Decision Process**
3. **Markov Chain preliminaries**
   - Conditional probabilities
   - Unconditional probability distributions
   - Long-run Properties of Markov Chain
4. **Bellman functional equation**
5. **Example: Stochastic optimal growth model**

Motivation

1. Equilibrium business cycle models aggregate risk
2. Partial equilibrium investment, asset pricing models with aggregate risk
3. Heterogeneous agent models with individual decision-maker risk
4. Dynamic Industrial Organization models (entry/exit), firm-specific risk

MDP

Model as dynamic stochastic decision problem (Markov Decision Process):

- Simple extension of previous deterministic dynamic programming problem.
- Exogenous finite-state Markov chain "shock" ε_t.
- ε_t perturbs *transition function* for the (endogenous) state vector.

Markov Chains

Definition (Markov property)

A stochastic process $\{\varepsilon_t\}$ has the Markov property if for all $k \geq 1$ and all t,

$$\Pr\left(\varepsilon_{t+1} \middle| \varepsilon_t, \varepsilon_{t-1}, ..., \varepsilon_{t-k}\right) = \Pr\left(\varepsilon_{t+1} \middle| \varepsilon_t\right).$$

$$\Pr\left(\text{Future } \middle| \text{Present \& Past}\right) = \Pr\left(\text{Future } \middle| \text{Present}\right)$$

Definition

A *time invariant (homogeneous) Markov chain* is defined by

1. Finite state space for the Markov chain $S = \{s_1, ..., s_n\}$.
2. An $n \times n$ transition/stochastic matrix P with typical element:

   $$P_{ij} = \Pr\left(\varepsilon_{t+1} = s_j \middle| \varepsilon_t = s_i\right)$$

   for $i, j = 1, ..., n$.
3. A $1 \times n$ vector λ_0 – containing the initial distribution across states – whose i-th element is the probability of being in state i at time 0:

   $$\lambda_{0i} = \Pr\left(\varepsilon_0 = s_i\right).$$

For this definition to be valid, matrix P and vector λ must always satisfy the following:

1. $P_{ij} \geq 0$,
2. The rows of P are probability distributions and so must sum to 1:

$$\begin{aligned}\sum_{j=1}^{n} P_{ij} =& 1 \\ =& \Pr\left\{ (\varepsilon_1 = s_1 | \, \varepsilon_0 = s_i) \cup (\varepsilon_1 = s_2 | \, \varepsilon_0 = s_i) \right. \\ & \left. \cdots \cup (\varepsilon_1 = s_n | \, \varepsilon_0 = s_i) \right\},\end{aligned}$$

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

Conditional probabilities. Probability of moving from state i to state j in one period is:

$$P_{ij} = \Pr\left(\varepsilon_{t+1} = s_j \middle| \varepsilon_t = s_i\right).$$

We can show that

$$\Pr\left(\varepsilon_{t+k} = s_j \middle| \varepsilon_t = s_i\right) = P_{ij}^{(k)}.$$

where $P_{ij}^{(k)}$ is the i, j element of the matrix $P^k = \underbrace{PPP \cdots P}_{k \text{ times}}$.

Unconditional probabilities. Unconditional probability distributions of ε_t for each period t.

- Start at $t = 0$ with given initial unconditional distribution λ_0

$$\lambda_1 = \lambda_0 P$$

- so we have

$$\begin{aligned}\lambda_2 &= \lambda_1 P = \lambda_0 P^2 \\ &\vdots \\ \lambda_t &= \lambda_0 P^t\end{aligned}$$

  where the ith element of λ_t is $\Pr(\varepsilon_t = s_i)$.

- Or we have

$$\lambda_{t+1} = \lambda_t P.$$

Steady state distribution [Appendix B]

1. Existence of stationary distribution
2. If stochastic matrix P strictly positive elements, stationary distribution is unique.
3. If stochastic matrix P^τ, for some $\tau > 1$, strictly positive, stationary distribution is unique.

Theorem (MC-LLN)

Let $h : S \to \mathbb{R}$. If $\{\varepsilon_t\}$ is a Markov chain (P, λ_0) on the finite set $S = \{s_1, ..., s_n\}$ such that it is asymptotically stable with stationary distribution λ^*, then as $T \to \infty$,

$$\frac{1}{T}\sum_{t=0}^{T} h(\varepsilon_t) \to \sum_{j=1}^{n} h(s_j)\lambda^*(s_j)$$

with probability one.

MDP: Notation

- $x_t \in X$ — current endogenous state vector (e.g. capital in the growth model);
- $u_t \in A$ — current action vector (e.g. consumption/next-period capital;
- P — *stochastic matrix* containing the conditional probability of moving from one state to another;
- S — space of finite exogenous states.
- $\Delta(S)$ — space of distributions over S. Probability simplex.
- $\lambda_0 \in \Delta(S)$ — initial distribution of the finite states.
- $\varepsilon_t \sim \text{Markov}(P, \lambda_0)$

Model of controlled Markov (endogenous) state:

$$x_{t+1} = F(x_t, u_t, \varepsilon_{t+1}).$$

Timing assumption:

- Start period t,
- Random variable ε_t is observed as $\varepsilon_t^o = s_i$ where $i = 1, ..., n$.
- Given F known, current endogenous state x_t known.
- Decision u_t taken given known states (x_t, ε_t^o).
- Induces realization of x_{t+1}.
- Enter period $t + 1$ ...

Remarks

- Decision maker can no longer just plan a deterministic sequence of actions.
- Makes a sequence of alternative state-contingent actions.
- To be able to make such a comprehensive list of state-contingent actions, decision maker must be able to form "correct" expectations of random future states.
- Apply von-Neumann-Morgernstern expected utility model of decision making in risky environments.

MDP: Bellman Equation

Suppose the current state is $(x_t, \varepsilon_t) = (x, s_i)$, where $i \in \{1, ..., n\}$.

So now the Bellman equation is given by

$$V(x, s_i) = \sup_{x' \in \Gamma(x, s_i)} U(x, x', s_i) + \beta \sum_{j=1}^{n} P_{ij} V(x', s_j)$$

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

Modified Bellman Operator

- Same mechanics as deterministic MDP we considered before. Just with a little modification.
- For every $x \in X$, there is a vector

$$\mathbb{R}^n \ni \mathbf{v}(x) = (V(x, s_1), ..., V(x, s_n)) \equiv (V_1(x), ..., V_n(x)).$$

- Space of all vectors real-valued, continuous and bounded functions, $\mathbf{v} : X \to \mathbb{R}^n$: $[C_b(X)]^n$.

Vector space of functions: Complete metric space

- Metric space $([C_b(X)]^n, d)$.
- Distance function, $d : [C_b(X)]^n \times [C_b(X)]^n \to \mathbb{R}_+$ given either by

$$d_\infty^n(\mathbf{v}, \mathbf{v}') = \sum_{i=1}^n d_\infty(V_i, V_i') = \sum_{i=1}^n \sup_{x \in X} |V(x, s_i) - V'(s, s_i)|,$$

or

$$d_\infty^{\max}(\mathbf{v}, \mathbf{v}') = \max_{i \in \{1, \ldots, n\}} \{d_\infty(V_i, V_i')\}.$$

- Complete metric space $([C_b(X)]^n, d)$.

Intuition:

- Fix each current $\varepsilon = s_i$, for every $x \in X$.
- RHS of the Bellman equation defines a operator T_i that maps $C_b(X)$ into itself.
- So $T_i : C_b(X) \to C_b(X)$, $i = 1, ..., n$.
- This is because for fixed $\varepsilon = s_i$,
  - $\sum_{j=1}^{n} P_{ij} V(x', s_j) \in C_b(X)$,
  - i.e. is a convex combination of all $V(x', s_j) \in C_b(X)$.

So "stacking" these T_i's together we have for each $x \in X$, $T : [C_b(X)]^n \to [C_b(X)]^n$ defined as

$$T\mathbf{v}(x) = \begin{bmatrix} T_1 V(x, s_1) \\ \vdots \\ T_n V(x, s_n) \end{bmatrix}$$

$$= \begin{bmatrix} \sup_{x' \in \Gamma(x,s_1)} U(x, x', s_1) + \beta \sum_{j=1}^{n} P_{1j} V(x', s_j) \\ \vdots \\ \sup_{x' \in \Gamma(x,s_n)} U(x, x', s_n) + \beta \sum_{j=1}^{n} P_{nj} V(x', s_j) \end{bmatrix}.$$

Punchline

- Each i-th component of T, T_i is a contraction mapping on a complete metric space $(C_b(X), d_\infty)$

- Then $T : [C_b(X)]^n \to [C_b(X)]^n$ is also a contraction mapping on the complete metric space $([C_b(X)]^n, d)$

- Recall, $d := d_\infty^n$ or $d := d_\infty^{\max}$.

Theorem

There exists a unique $\mathbf{v}$ *that satisfies the Bellman principle of optimality.*

Theorem (Properties of $\mathbf{v}$)

Assume, for each fixed $\varepsilon' \in S$:

1. $U : X \times A \to \mathbb{R}$ *is bounded and continuous;*
2. $F : X \times A \to \mathbb{R}$ *is continuous; and*
3. $\Gamma : X \to P(A)$ *is compact and continuous.*

Then the value function $\mathbf{v}$ is also bounded on $X \times S$ and for each $\varepsilon \in S$, and it is continuous on X.

Theorem (Unique state-contingent optimal plan)

1. *If U is strictly increasing, V is strictly increasing on X.*
2. *Also if U is strictly concave and F is weakly concave, for each ε, then V_i is weakly concave on X, for each $i = 1, ..., n$, and there exists a unique strategy (stationary optimal strategy) for each $\varepsilon \in S$.*

Parametric forms:

$$U(c) \begin{cases} = \frac{c^{1-\sigma}-1}{1-\sigma} & \sigma > 1 \\ \\ = \ln(c) & \sigma \to 1 \end{cases}$$

and

$$f(k, A(i)) = A(i)k^{\alpha} + (1-\delta)k; \quad \alpha \in (0,1), \delta \in (0,1].$$

This is done as two separate functions. Now let's code this up as our Python Tutorial.