Markov Decision Processes in Economics

Stoc Proc 25-26, Based on Brock-Mirman (1972, JET)

Srikanth Pai, Asst. Prof, MSE

A Classic Problem: Growth with Uncertainty

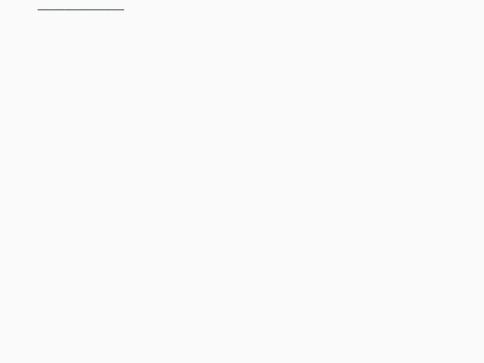
- Each period, the economy produces output using last period's capital and today's productivity.
- Productivity can be *high* or *low* and changes over time in a random but persistent way.
- A planner chooses how much of today's output to consume and how much to save as new capital for tomorrow.
- The tradeoff: enjoy consumption now vs. invest to raise uncertain future production.
- Goal: choose consumption/saving each period to maximize society's discounted well-being over time.

Formal Setup (Brock-Mirman, 1972)

- **State variables:** capital K_t , productivity z_t .
- Production: $Y_t = z_t f(K_t)$.
- Resource constraint: $Y_t = C_t + K_{t+1}$ (no depreciation for simplicity; include δ if needed).
- Shock dynamics: $z_{t+1} \sim P(\cdot \mid z_t)$ (finite-state Markov or AR(1) in logs).
- Objective:

$$\max_{\{C_t\}_{t\geq 0}} \mathbb{E}\left[\sum_{t=0}^{\infty} \beta^t u(C_t)\right], \qquad 0 < \beta < 1.$$

• Control: choose C_t (equivalently choose $K_{t+1} = z_t f(K_t) - C_t$) each period.



From Stochastic Processes to Markov Chains

Without Control: The Shock Process

- Productivity evolves randomly: z_{t+1} depends on z_t .
- This sequence $\{z_t\}$ is a stochastic process.
- To model persistence, assume z_t follows a Markov chain.

Markov Property

Definition

A process $\{z_t\}$ is Markov if

$$P(z_{t+1} \mid z_t, z_{t-1}, \ldots) = P(z_{t+1} \mid z_t).$$

- In Brock-Mirman: productivity has memory only of the last period.
- Example: 2-state productivity $\{H, L\}$ with transition matrix

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

5

Expected Rewards with a Markov Chain

- For a fixed policy, the induced process is a DTMC.
- DTMC theory: stationary distributions allow computation of long-run averages.
- With discounting, expected reward is

$$\mathbb{E}\left[\sum_{t=0}^{\infty}\beta^t u(C_t)\right].$$

Framework

Adding Control: The MDP

Adding Back Control

- State: (K_t, z_t) .
- Action: consumption choice C_t (or equivalently next capital K_{t+1}).
- Transition:

$$K_{t+1} = z_t f(K_t) - C_t, \quad z_{t+1} \sim P(\cdot \mid z_t).$$

• Reward: $u(C_t)$.

Definition: Infinite-Horizon MDP

- State space S, action space A(s).
- Transition kernel P(s' | s, a).
- Reward function r(s, a).
- Discount factor $\beta \in (0,1)$.

Definition: Infinite-Horizon MDP

- State space S, action space A(s).
- Transition kernel P(s' | s, a).
- Reward function r(s, a).
- Discount factor $\beta \in (0,1)$.

Policy

A stationary policy $\pi:\mathcal{S}\to\mathcal{A}$ maps states to actions.

Value Functions

• Value of policy π :

$$V^{\pi}(s) = \mathbb{E}_{s}^{\pi} \left[\sum_{t=0}^{\infty} \beta^{t} r(s_{t}, a_{t}) \right].$$

• Optimal value:

$$V^*(s) = \sup_{\pi} V^{\pi}(s).$$

Bellman Optimality

Bellman Optimality Equation

Theorem (Bellman)

There exists an optimal stationary policy π^* and

$$V^*(s) = \max_{a \in A(s)} \Big\{ r(s, a) + \beta \sum_{s'} P(s' \mid s, a) V^*(s') \Big\}.$$

Sketch of Proof

1. Define Bellman operator T:

$$(TV)(s) = \max_{a} \{r(s, a) + \beta \sum_{s'} P(s' \mid s, a) V(s')\}.$$

- 2. Show T is a contraction under sup norm.
- 3. Banach fixed point theorem \implies unique V^* .
- 4. Greedy policy wrt V^* is optimal.

Application: Solving

Brock-Mirman

Bellman Equation for Growth Model

$$V(K,z) = \max_C \left\{ u(C) + \beta \, \mathbb{E}[V(K',z') \mid z] \right\},$$
 subject to $K' = zf(K) - C$.

Economic Insights

- Optimal consumption-saving policy derived from Bellman equation.
- ullet Shocks are persistent \Longrightarrow precautionary saving motive.
- Links to modern RBC and DSGE models.

What the Bellman Equation Delivers in Brock-Mirman

• Closed-form policy: With log utility and Cobb–Douglas f(K), optimal consumption is a constant fraction of output:

$$C_t = (1 - \alpha \beta) z_t K_t^{\alpha}.$$

- Markov sufficiency: Optimal action depends only on current state (K_t, z_t) .
- Role of parameters:
 - Higher patience $\beta \Rightarrow$ more saving.
 - Higher productivity shock $z_t \Rightarrow$ proportional rise in both C_t and K_{t+1} .
- Big picture: Dynamic programming makes stochastic growth tractable; this recursion became the prototype for modern RBC models.