RLChina Reinforcement Learning Summer School

RLChina 2022

Reinforcement Learning and Dynamic Macroeconomic Models

Prof. Shenghao Zhu

University of International Business and Economics

August 24, 2022

Topics

- Single agent's problem
- Stationary distribution
- Heterogeneous agents model
- Future perspectives

Income fluctuation problems

- The agent's problem

$$\max_{\{c_t, a_{t+1}\}} E_0 \sum_{t=0}^{\infty} \beta^t u(c_t)$$

$$s.t.\ c_t + a_{t+1} = (1 + r_t)a_t + w_t l_t$$

$$c_t \geq 0, a_{t+1} \geq -b$$

where

R=1+r is the gross return of the asset.

w is the wage.

l_t is the labor efficiency endowment, which follows a stochastic process.

The Bellman equation

- Let $V(a_t, l_t)$ be the value function

$$V(a_t, l_t) = max_{c_t, a_{t+1}}\{u(c_t) + \beta E_t V(a_{+1}, l_{t+1})\}$$

$$s.t.\ \ c_t + a_{t+1} = Ra_t + wl_t,$$

$$c_t \geq 0, a_{t+_1} \geq -b$$

where l_t follows a Markov chain.

Supervised learning and unsupervised learning

- Give nn to $V(a_t, l_t)$
- Supervised learning. Using the Bellman equation to generate data
- Unsupervised learning. Using the Bellman equation as a loss function
- Ergodic set and the domain of (a_t, l_t)

The Euler equation

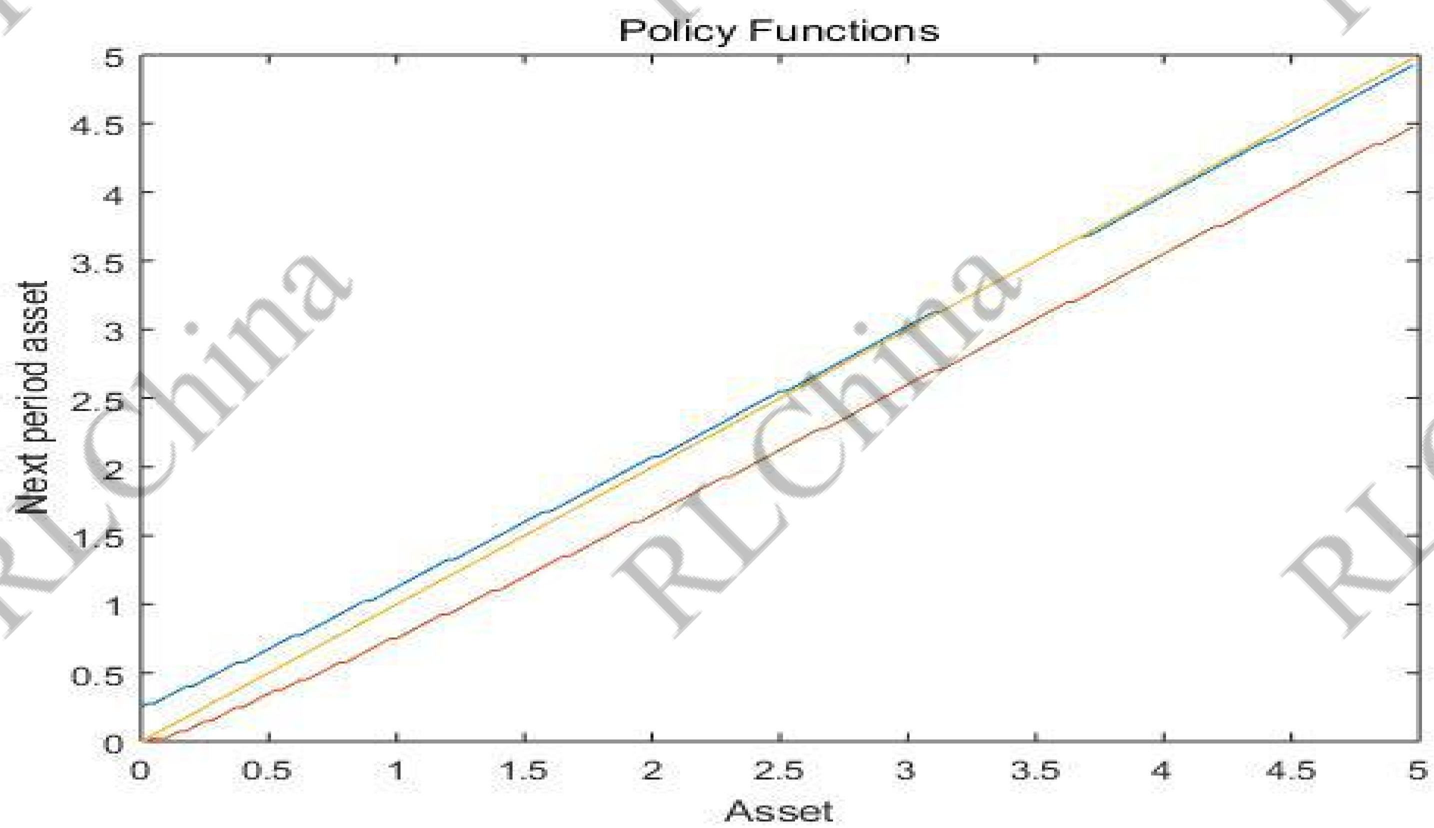
- We can also start from the policy function
- $u'(c_t) \geq \beta R E_t u'(c_{t+1})$ with equality if $a_{t+_1} > -b$
- Policy $c_t = c(a_t, l_t)$
- Transform
- $A \geq 0, H \geq 0 \; and \; AH = 0$
- $\varphi(A, H) = A + H - \sqrt{A^2 + H^2} = 0$
- Supervised learning and unsupervised learning

Reward maximization

- Pick large T
- Give nn to $c_t = c(a_t, l_t)$
- Object

$$E_0 \sum_{t=0}^{T} \beta^t u(c_t)$$

Policy function



A Markovian operator

- A Markov process has a transition probability $P(z, B)$

$$P(z, B) = \Pr(z_{t+1} \in B | z_{t=z})$$

where B is Borel set in S.

- $P(z, B)$ defines an operator T^* on $\boldsymbol{\Lambda}(S, \mathrm{F})$ which is the set of probability measures on (S, F).

$$(T^*\lambda)(B) = \int_S P(z, A) d\lambda$$

for any Borel set B $\subseteq S$.

Markov process induced from the agent's problem

- The wealth accumulation process

$$a_{t+1} = Ra_t - c(a_t, l_t) + wl_t$$

- The transition probability

$$P\big((a,l), A \times l'\big) = \begin{cases} \pi(l'|l) & if\ a'(a,l) \in A \\ 0 & otherwise \end{cases}$$

for all $(a,l) \in S$.

Stationary distribution

- A distribution μ is stationary or invariant if

$$\mu(B) = \int_S P(z, B) d\mu$$

for any Borel set $\mathrm{B} \subseteq S$.

- μ is a fixed point of the operator T^*

$$\mu = T^* \mu$$

- μ is an infinite dimensional item.
- Dimension reduction
- Simulation

Continuous-time model

- The agent's problem

$$\max E_0 \int_0^\infty e^{-\rho t} u\big(c(t)\big) dt$$

$$s.t.\, dk(t) = [r(t)k(t) + w(t)l(t) - \delta k(t) - c(t) - T(t)]dt + \sigma k(t) dB(t)$$

$l(t)$ follows a two-state Poisson process $l(t) \in \{l_1, l_2\}$

$T(t)$ is the tax schedule

$B(t)$ is a standard Brownian motion

Hamilton-Jacobi-Bellman equations

- Let

$$v(k(t)) = max E_t \int_t^{\infty} e^{-\rho(\tau - t)} u(c(\tau)) d\tau$$

- HJB

$$\rho v_j(k) = max_{c,k} u(c_j) + v'_j(k) s_j(k) + \frac{1}{2} v''_j(k) \sigma^2 k^2 + \lambda_j(v_{-j}(k) - v_j(k))$$

for $j = 1,2$

where $s_j(k) = rk + wl_j - \delta k - c_j - T$

- F.O.C.

$$v'_j(k) = u'(c_j(k))$$

Two approaches in optimal control

- 1. Variational approach (Calculus variation, Pontryagin's maximum principle)
- Deep learning
- 2. Recursive structure (Bellman equation)
- Reinforcement learning

Wealth accumulation process

- Under the optimal nn of $c(k,l)$, we have
$$dk(t) = [r(t)k(t) + w(t)l(t) - \delta k(t) - c(t) - T(t)]dt$$
$$+ \sigma k(t)dB(t)$$
- Two ways to generating the stationary distribution

1. Simulation
2. PDE (histogram)

Kolmogorov Forward Equation

- The distributions $f_j(k,t), j = 1,2$ are governed by KFE

$$\frac{\partial}{\partial t} f_j(k,t) = \frac{1}{2}\frac{\partial^2}{\partial k^2}\left[\sigma^2 k^2 f_j(k,t)\right] - \frac{\partial}{\partial k}\left[s_j(k) f_j(k,t)\right] - \lambda_j f_j(k,t) + \lambda_{-j} f_{-j}(k,t)$$

for $j = 1,2$

- The stationary distribution $f_j(k), j = 1,2$

$$0 = \frac{1}{2}\frac{\partial^2}{\partial k^2}\left[\sigma^2 k^2 f_j(k)\right] - \frac{\partial}{\partial k}\left[s_j(k) f_j(k)\right] - \lambda_j f_j(k) + \lambda_{-j} f_{-j}(k)$$

for $j = 1,2$

Partial equilibrium

- Interest rate r(t) and wage rate w(t) are exogenously given
- General equilibrium

1. The stationary distribution
2. Endogenous r(t) and w(t)

Heterogeneous agents model

- Aiyagari model (Without aggregate uncertainty, Mean field game)
- Krusell-Smith model (With aggregate uncertainty, Mean field game with common shock)
- There is a continuum of agents, with name $i \in [0,1]$, in the economy.

Aiyagari model

- The capital market

$$\int_S a d\mu = K$$

- The labor market

$$\int_S l d\mu = L$$

- The production function F(K,L)
- Interest rate $r = \frac{\partial}{\partial K} F(K,L)$
- Wage rate w $= \frac{\partial}{\partial L} F(K,L)$

Krusell-Smith model

- The production function

$$F(K, L) = A_t {K_t}^{\alpha} {L_t}^{1-\alpha}$$

where A_t follow a stochastic process.

- The capital process $K_t = \int_S \ a d\mu_t$
- Then r_t and w_t are stochastic
- The state space contains the cross-section distribution μ_t, which is infinite dimensional

The Bellman equation

- $V(\mu_t, a_t, l_t)$ satisfies
$$V(\mu_t, a_t, l_t) = max_{c_t, a_{t+1}}\{u(c_t) + \beta E_{T^*} V(\mu_{t+1}, a_{t+1}, l_{t+1})\}$$
$$s.t.\ c_t + a_{t+1} = (1 + r_t)a_t + w_t l_t$$
$$a_{t+1} \geq -b$$
$$\mu_{t+1} = T^* \mu_t$$
- Given T^*, we find an optimal nn c^*
- Given c^*, we deduce T^*
- T^* is called rational expectation

HACT

- Heterogeneous agent continuous-time model
- Transition path. Given $f_j(k,0)$

$$
\begin{aligned}
\frac{\partial}{\partial t} f_j(k,t) \\
&= \frac{1}{2}\frac{\partial^2}{\partial k^2}\left[\sigma^2 k^2 f_j(k,t)\right] - \frac{\partial}{\partial k}\left[s_j(k) f_j(k,t)\right] - \lambda_j f_j(k,t) \\
&+ \lambda_{-j} f_{-j}(k,t)
\end{aligned}
$$

for $j = 1,2$

Simulation methods

- Name

$$\int_S a d\mu_t = K_t = \int_{[0,1]} {a_t}^i di$$

- Taleb, "Statistical consequences of fat tails"
- Low efficiency of convergence

KFE-A powerful tool

- Keep track of $f_j(k,t)$
- $f_j(k,t)$ is infinite dimensional
- Furthermore, $f_j(k,t)$ is STOCHASTIC
- Random measures $\{\mu_t\}_{t=0}^{\infty}$ is a stochastic process in a Hilbert space.

Perspectives

- Optimal control in infinite dimension
- Stochastic analysis
- Fixed-point theorem

Optimal control in infinite dimension

- The agent's problem

$$\max E_0 \int_0^\infty e^{-\rho t} u\big(c(t)\big) dt$$

s.t.

$$\frac{\partial}{\partial t} f_j(k,t) = \frac{1}{2}\frac{\partial^2}{\partial k^2}\left[\sigma^2 k^2 f_j(k,t)\right] - \frac{\partial}{\partial k}\left[s_j(k) f_j(k,t)\right] - \lambda_j f_j(k,t) + \lambda_{-j} f_{-j}(k,t)$$

for $j = 1,2$

Stochastic analysis

- With aggregate shock, we have KFE with stochastic coefficients
$$\frac{\partial}{\partial t} f(k,\omega,t) = \frac{1}{2}\frac{\partial^2}{\partial k^2}[\sigma^2 k^2 f(k,\omega,t)] - \frac{\partial}{\partial k}[s(k,\omega,t) f(k,\omega,t)]$$
- An SPDE

Fixed-point theorem

- Given T^*, we find an optimal nn c^*
- Given c^*, we deduce T^*
- T^* is called rational expectation
- An Euler operator $\mathbf{\Gamma} c(a, l)$ is monotone
- $\mathbf{\Gamma} c(a, l) = y$ solves
  $u'(y) = \max\{\varphi(a_t, l_t), \beta R E_t u'[c(R a_t + w l_t - y, l_{t+1})]\}$