DDC

Junrui Lin

NYU MSQE

March 21, 2023

Nested Fixed Point Theorem (Rust, J., 1987)

Setup of *Regenerative Optimal Stopping Problem*

- Zurcher chooses discrete actions i_t each period, $i_t = 1(replace), i_t = 0(keep)$
- We observe regenerative process $\{i_t, x_t\}$, the likelihood of the process can be written as $l(i_1, ..., i_t, x_1, ..., x_t|\theta)$
- Define differentiable cost function $c(x_t, \theta_1)$, the flow utility will be

$$u(x_t, i_t, \theta_1) = \begin{cases} -c(x_t, \theta_1) & i_t = 0 \\ -s - c(0, \theta_1) & i_t = 1 \end{cases}$$

- The question is also called Markov Decision Problem (MDP), because the state (milage of the engine x_t) has Markov property

$$p(x_{t+1}|x_t, i_t, \theta_2) = \begin{cases} \theta_2 \exp\{\theta_2(x_{t+1} - x_t)\} & i_t = 0, x_{t+1} \geq x_t \\ \theta_2 \exp\{\theta_2(x_{t+1})\} & i_t = 0, x_{t+1} \geq 0 \\ 0 & \text{otherwise} \end{cases}$$

Nested Fixed Point Theorem (Rust, J., 1987)

Now we solve the problem

- First we can write out the Bellman equation

$$V_\theta(x_t) = \max_{i_t} \mathbb{E}\left[u(x_t, i_t, \theta_1) + \beta \int_0^\infty V_\theta(x_{t+1}) dG(x_{t+1}|x_t, i_t, \theta_2)\right]$$

- In Rust (1986), he proved there is an optimal stationary, Markovian replacement policy $\Pi = \{f, f, ...\}$

$$i_t^* = f(x_t, \theta) = \begin{cases} 1 & x > \gamma(\theta_1, \theta_2) \\ 0 & x \leq \gamma(\theta_1, \theta_2) \end{cases}$$

  where

$$s(1-\beta) = \int_0^{\gamma(\theta_1,\theta_2)} [1 - \beta \exp\{-\theta_2(1-\beta)x'\}] \frac{\partial c(x', \theta_1)}{\partial x'} dx'$$

- Then we can form the likelihood to use MLE and solve for θ

Nested Fixed Point Theorem (Rust, J., 1987)

It looks good, but there are several drawbacks

- Assuming the monthly mileage $(x_{t+1} - x_t)$ has i.i.d. exponential does't reflect real data, but using other flexible distribution will ruin previous explicit solution
- The cutoff replacement rule doesn't fit the data, in the data, there are replacement in many values of x_t, potential explanation: x_t is just one dimension of the quality of the engine, there are unobserved characteristics ε_t that affect the replacement decision
- So the decision rule should be something like

$$i_t = f(x_t, \theta) + \varepsilon_t$$

  where ε_t is observed by the agent but not the econometrician.
- However this solution makes the estimation very challenging: the structural model assume optimization, but the solution means agent deviates the optimal behavior randomly.

Nested Fixed Point Theorem (Rust, J., 1987)

- Given Markovian state variables (x_t, ε_t) embodied by the transition probability $p(\theta_2, \theta_3)$—also relate to the optimal control f

$$p(x_{t+N}, \varepsilon_{t+N}, ..., x_{t+1}, \varepsilon_{t+N} | x_t, \varepsilon_t) = \prod_{i=1}^{N-1} p(x_{i+1}, \varepsilon_{i+1} | x_i, \varepsilon_i, f_i(x_i, \varepsilon_i), \theta_2, \theta_3)$$

- Agent chooses sequence of decision rule $\Pi = \{f_1, f_2, ...\}$ to maximize utility

$$V_\theta(x_0, \varepsilon_0) = \max_{\Pi} \mathbb{E}\left\{ \sum_{j=1}^{\infty} \beta^j [u(x_j, f_j, \theta_1) + \varepsilon_j(f_j)] | x_0, \varepsilon_0, \theta_2, \theta_3 \right\}$$

$$V_\theta(x_t, \varepsilon_t) = \max_{i_t} [u(x_t, i_t, \theta_1) + \varepsilon_t(i_t) + \beta \mathbb{E} V_\theta(x_{t+1}, i_t, \varepsilon_{t+1})]$$

where

$$\mathbb{E} V_\theta(x_{t+1}, i_t, \varepsilon_{t+1}) = \int_{x'} \int_{\varepsilon'} V_\theta(x', \varepsilon') dG(x', \varepsilon' | x_t, \varepsilon_t, i_t, \theta_2, \theta_3)$$

- The optimal control thus is

$$f_t(x_t, \varepsilon_t, \theta) = \arg\max_{i_t} [u(x_t, \theta_1) + \varepsilon_t(i_t) + \beta \mathbb{E} V_\theta(x_{t+1}, i_t, \varepsilon_{t+1})]$$

Nested Fixed Point Theorem (Rust, J., 1987)

$$V_\theta(x_t, \varepsilon_t) = \max_{i_t}[u(x_t, i_t, \theta_1) + \varepsilon_t(i_t) + \beta \mathbb{E} V_\theta(x_{t+1}, i_t, \varepsilon_{t+1})]$$

$$\mathbb{E} V_\theta(x_{t+1}, i_t, \varepsilon_{t+1}) = \int_{x'} \int_{\varepsilon'} V_\theta(x', \varepsilon') dp(x', \varepsilon' | x_t, \varepsilon_t, i_t, \theta_2, \theta_3)$$

$$f_t(x_t, \varepsilon_t, \theta) = \arg\max_{i_t}[u(x_t, \theta_1) + \varepsilon_t(i_t) + \beta \mathbb{E} V_\theta(x_{t+1}, i_t, \varepsilon_{t+1})]$$

The optimal control is very hard to solve

- The unobservable ε_t's support usually is unbounded, when do grid search, it creates difficulty to find fixed point of $V_\theta(x_t, \varepsilon_t)$
- Since $\mathbb{E} V_\theta(x_{t+1}, i_t, \varepsilon_{t+1})$ is an unknown function, in every iteration, we need to numerically integrate V_θ over finite grid approximation of p to get $\mathbb{E} V_\theta$. And then integrate the Bellman equation to get $P(i_t | x_t, \theta)$

Under some condition, Rust propose a MLE algorithm to solve the dynamic discrete choice (DDC) problem.

Nested Fixed Point Theorem (Rust, J., 1987)

conditional independence assumption:

$$p(x_{t+1}, \varepsilon_{t+1} | x_t, \varepsilon_t, i_t, \theta_2, \theta_3) = q(\varepsilon_{t+1} | x_{t+1}, \theta_2) p(x_{t+1} | x_t, i_t, \theta_3)$$

under this assumption

- Let $G([u(x, \theta_1) + \beta \mathbb{E}V_\theta(x)] | x, \theta_2)$ denote the social surplus function corresponding to the density $q(\varepsilon | x, \theta_2)$

$$G = \int_\varepsilon \max_i [u(x, i, \theta_1) + \beta \mathbb{E}V_\theta(x, i)] dq(\varepsilon | x, \theta_2)$$

Then

$$P(i | x, \theta) = \frac{\partial G([u(x, \theta_1) + \beta \mathbb{E}V_\theta(x)] | x, \theta_2)}{\partial u(x, i, \theta_1)}$$

We can use the following contraction mapping to solve $\mathbb{E}V_\theta$

$$\mathbb{E}V_\theta(x, i) = \int_{x'} G([u(x', \theta_1) + \beta \mathbb{E}V_\theta(x')] | x', \theta_2) dp(x' | x, i, \theta_3)$$

Nested Fixed Point Theorem (Rust, J., 1987)

Example: $q(\varepsilon|x,\theta_2)$ is given by a multivariate EV distribution

$$q(\varepsilon|x,\theta_2) = \prod_{i \in C} e^{-\varepsilon(i)+\theta_2} e^{-e^{-\varepsilon(i)+\theta_2}}$$

Now let's solve the model

$$\begin{aligned} G &= \int_{\varepsilon} \max_i [u(x,i,\theta_1) + \beta \mathbb{E}V_\theta(x,i)] dq(\varepsilon|x,\theta_2) \\ &= \ln\left\{ \sum_{i \in C} \exp[u(x,i,\theta_1) + \beta \mathbb{E}V_\theta(x,i)] \right\} \\ P(i|x,\theta) &= \frac{\exp[u(x,i,\theta_1) + \beta \mathbb{E}V_\theta(x,i)]}{\sum_{i \in C} \exp[u(x,i,\theta_1) + \beta \mathbb{E}V_\theta(x,i)]} \\ \mathbb{E}V_\theta(x,i) &= \int_{x'} \ln\left\{ \sum_{i \in C} \exp[u(x',i,\theta_1) + \beta \mathbb{E}V_\theta(x',i)] \right\} dp(x'|x,i,\theta_3) \end{aligned}$$

Nested Fixed Point Theorem (Rust, J., 1987)

under this assumption (cont.)

- The likelihood function of the sample can be written as

$$l(x_1, ..., x_T, i_1, ..., i_T | x_0, i_0, \theta) = \prod_{t=1}^{T} P(i_t | x_t, \theta) p(x_t | x_{t-1}, i_{t-1}, \theta_3)$$

  Then use MLE to solve for θ

Example (cont.)

- The likelihood function of the sample can be written as

$$l = \prod_{t=1}^{T} \frac{\exp[u(x_t, i_t, \theta_1) + \beta \mathbb{E} V_\theta(x_t, i_t)]}{\sum_{i \in C} \exp[u(x_t, i_t, \theta_1) + \beta \mathbb{E} V_\theta(x_t, i_t)]} p(x_t | x_{t-1}, i_{t-1}, \theta_3)$$

  Then use MLE to solve for θ

CCP (Hotz, V. J., & Miller, R. A., 1993)

Despite Nested Fixed Point Theorem is very general in estimating structural DDC model, but it's computationally heavy (we need to do DP for V with every parameter vector). Hotz and Miller proposed alternative method (Conditional Choice Probability) to do the estimation without solving the Bellman equation.

Setup

- Let $H_t = (b_0, ..., b_{t-1})$, and b_t is generated through $F_j(H_{t+1}|H_t)$ when choose d_j
- For each period t, the utility associate with the choice j is u_{tj}, define $u_j^*(H_t) := \mathbb{E}(u_{tj}|H_t)$
- Similar to before, there is stochastic utility components, $u_{tj} = u_j^*(H_t) + \varepsilon_{tj}$, the distribution of $\varepsilon_t = (\varepsilon_{t1}, ..., \varepsilon_{tJ}) \sim G(\varepsilon_t|H_t)$
- The agent choose $\{d_1, ..., d_T\}$ to maximize the objective function

$$\mathbb{E}_0(\sum_{t=0}^{T}\sum_{j=0}^{J} d_{jt}[u_j^*(H_t) + \varepsilon_{tj}])$$

CCP (Hotz, V. J., & Miller, R. A., 1993)

Setup (cont.)

- Let $d_s^0 = (d_{s1}^0, ..., d_{sJ-1}^0)'$ denote the agent's optimal choice in period s
- Define the conditional valuation function of choose j in period t

$$v_j(H_t) = \mathbb{E}_0[\sum_{s=t+1}^{T} \sum_{j=1}^{J} d_{sj}^0 [u_j^*(H_t) + \varepsilon_{jt}] | H_t, d_{tj} = 1]$$

so $d_{tk}^0 = 1$ when $k = \arg\max_j \{u_j^*(H_t) + \varepsilon_{jt} + v_j(H_t)\}$

- The conditional probability the agent chooses k

$$p_k(H_t) = Pr\{k = \arg\max_j \{u_j^*(H_t) + \varepsilon_{jt} + v_j(H_t)\} | H_t\}$$

define $p(H_t) = (p_1(H_t), ..., p_{J-1}(H_t))$

CCP (Hotz, V. J., & Miller, R. A., 1993)

From the setup, we can compute the conditional probability of making choice 1, given H_t

$$\begin{aligned}
p_1(H_t) &= \mathbb{E}(d_{t1}^0 = 1|H_t) \\
&= \int_{\varepsilon_1=-\infty}^{\infty} \int_{\varepsilon_2=-\infty}^{\varepsilon_1+u_{t1}^*+v_{t1}-u_{t2}^*-v_{t2}} \dots \int_{\varepsilon_J=-\infty}^{\varepsilon_1+u_{t1}^*+v_{t1}-u_{tJ}^*-v_{tJ}} dG(\varepsilon_1, ..., \varepsilon_J|H_t) \\
&= \int_{-\infty}^{\infty} G_1(\varepsilon_1, [\varepsilon_1 + u_{t1}^* + v_{t1} - u_{t2}^* - v_{t2}], ..., [\varepsilon_1 + u_{t1}^* + v_{t1} - u_{tJ}^* - v_{tJ}]|H_t)d\varepsilon_1
\end{aligned}$$

Define $v = (v_1, ..., v_{J-1})'$ and function $Q_j(v, H_t)$

$$Q_j(v, H_t) = \int G_j([\varepsilon_j + u_{tj}^* + v_j - u_{t1}^* - v_1], ..., \varepsilon_j, ..., [\varepsilon_j + u_{tj}^* + v_j - u_{tJ}^*]|H_t)d\varepsilon_j$$

Let $v(H_t) = (v_{t1} - v_{tJ}, ..., v_{tJ-1} - v_{tJ})$, $p(H_t) = (p_1(H_t), ..., p_{J-1}(H_t))$, $Q(v, H_t) = (Q_1(v, H_t), ..., Q_{J-1}(v, H_t))$, we have

$$p(H_t) = Q(v(H_t), H_t)$$

CCP (Hotz, V. J., & Miller, R. A., 1993)

The estimation requires $Q(v, H_t)$ is invertible in v, so that $v(H_t)$ can be expressed as a function of $p(H_t)$——$v_j(H_t) = Q_j^{-1}(p_t(H_t), H_t)$

Proposition 1: For each H_t, the mapping $Q(v, H_t)$ is invertible in v

$$\mathbb{E}(\sum_{j=1}^{J} d_{jt}^0[u_j^*(H_t) + \varepsilon_{tj}]|H_t) = \sum_{j=1}^{J} p_j(H_t)[u_j^*(H_t) + \mathbb{E}(\varepsilon_{tj}|H_t, d_{tj}^0 = 1)]$$

And $\mathbb{E}(\varepsilon_{tj}|H_t, d_{tj}^0 = 1)$ can be written as

$$\begin{aligned} W_j(p_t, H_t) &= \int \frac{\varepsilon_j}{p_j(H_t)} G_j([\varepsilon_j + u_{tj}^* + v_j - u_{t1}^* - v_1], ..., \varepsilon_j, ..., [\varepsilon_j + u_{tj}^* + v_j - u_{tJ}^*]|H_t) d\varepsilon_j \\ &= \int \frac{\varepsilon_j}{p_j(H_t)} G_j([\varepsilon_j + u_{tj}^* - u_{t1}^* + Q_{tj}^{-1} - Q_{t1}^{-1}], ..., \varepsilon_j, ..., [\varepsilon_j + u_{tj}^* - u_{tJ}^* + Q_{tj}^{-1}]|H_t) d\varepsilon_j \end{aligned}$$

Then the agent's expected utility

$$U(p_t, H_t) = \sum_{k=1}^{J} p_k(H_t)[u_{tk}^* + W_k(p_t, H_t)]$$

CCP (Hotz, V. J., & Miller, R. A., 1993)

Example: Sterilizing Decision Problem

- In each time t, couples' make decision from two choices: $d_{t1} = 1$ (not sterilize), $d_{t2} = 1$ (sterilize)
- Sterilizing is irreversible, if $d_{t1} = 0$, $\forall s > t,\ d_{s1} = 0$
- Outcome variable b_t is birth, if not sterilize, there is probability α that a children is born, if sterilize, women can't bear children.

$$F(H_{t+1} = (H_t, 1)|H_t) = \begin{cases} \alpha & d_{t1} = 1 \\ 0 & d_{t1} = 0 \end{cases}$$

- Define $\tilde{H}_t = \sum_{s=0}^{t} b_t$, the utility

$$u_1^*(H_t) = u_2^*(H_t) = \beta^t(\delta_1 \tilde{H}_t + \delta_2 \tilde{H}_t^2)$$

CCP (Hotz, V. J., & Miller, R. A., 1993)

The couple's decision problem is choosing $\{d_s\}_{s=0}^{T}$ (or equivalently find the optimal τ to sterilize) to maximize

$$\mathbb{E}_0[\sum_{t=0}^{T}\beta^t(\delta_1\tilde{H}_t+\delta_2\tilde{H}_t^2+d_t\varepsilon_{t1}+(1-d_t)\varepsilon_{t2})]$$

where ε_{tj} is i.i.d. T1EV, which implies the indirect utility of sterilizing in period t is

$$v_2(H_t)=\beta^t(\gamma+\delta_1\tilde{H}_t+\delta_2\tilde{H}_t^2)/(1-\beta^{T-t})$$

where $\gamma\approx 0.577$ is the Euler's constant. And the indirect utility of not sterilizing in period t is

$$v_1(H_t)=\max_{\{d_s\}_{s=t+1}^{T}}\mathbb{E}[\sum_{s=t+1}^{T}\beta^s(\delta_1\tilde{H}_s+\delta_2\tilde{H}_s^2+d_s\varepsilon_{s1}+(1-d_s)\varepsilon_{s2})|H_t]$$

CCP (Hotz, V. J., & Miller, R. A., 1993)

Therefore the optimal decision rule is

$$d_t = \begin{cases} 0 & \text{if } \varepsilon_{t1} - \varepsilon_{t2} \geq \beta^{-t} v_1(H_t) - \beta^{-t} v_2(H_t) \\ 1 & \text{if } \varepsilon_{t1} - \varepsilon_{t2} < \beta^{-t} v_1(H_t) - \beta^{-t} v_2(H_t) \end{cases}$$

This implies the conditional probability of choosing not to sterilize in t is

$$p_1(H_t) = \frac{1}{1 + \exp[\beta^{-t} v_1(H_t) - \beta^{-t} v_2(H_t)]}$$

$$p_2(H_t) = \frac{\exp[\beta^{-t} v_1(H_t) - \beta^{-t} v_2(H_t)]}{1 + \exp[\beta^{-t} v_1(H_t) - \beta^{-t} v_2(H_t)]}$$

Follows *Proposition 1*

$$\begin{aligned} Q^{-1}(p_1(H_{t+1}), H_{t+1}) &= \beta^{t+1} \ln\left(\frac{1}{\exp[\beta^{-t-1} v_1(H_{t+1}) - \beta^{-t-1} v_2(H_{t+1})]}\right) \\ &= v_2(H_{t+1}) - v_1(H_{t+1}) \end{aligned}$$