Estimation of Single-Agent Dynamic Discrete Choice Model 1

Nested-Fixed-Point Algorithm

Yuta Toyama

Last updated: December 6, 2018

Introduction –Why are dynamics important?

- Consumers:
  - Buy today or wait till tomorrow (durables)
  - Stockpiling from sales (storables)
  - Retirement decision
- Firms:
  - Entry/Exit
  - Investment
  - R&D, introduction of new products

Why are dynamics difficult?

- Computational burden to solving dynamic problems blows up quickly with the state space.
- Often trade-off between making the model realistic and making it dynamic.
- Other issues:
  - identication assumptions are more demanding
  - concerns about serially correlated unobservables and other forms of unobserved heterogeneity
  - solving for equilibria, multiplicity (in game settings).

Scope of the lecture

- Focus on (1) discrete choice, (2) infinite horizon, and (3) single-agent setting
- Continuous decision in dynamic setting
  - level of investment, saving, etc.
  - most common in labor and macro
  - continuous decisions important in IO as well.
- Infinite horizon vs finite horizon
  - Here: stationary environment in infinite horizon setting.
  - Other: non-stationary in finite horizon.
    - example: retirement decisions, human capital accumulation
- Single-agent vs multiple-agent
  - Multiple-agent (strategic interactions) important in IO.
  - Multiple-agent introduces another difficulty in analysis: multiplicity of equilibria and high dimensionality of state space

Crash Course on Continuous Choice Model in Dynamic Setting

- Consumption & saving

$$\max_{c_t, s_t} \mathbb{E}\left[\sum_{\tau=t}^{\infty} \beta^{\tau-t} u(c_t)|\Omega_t\right] \ \ s.t. \ c_t + s_t \leq y_t + r_t s_{t-1}$$

  - s_t :saving, r_t: interest rate, y_t: income
  - Ω_t: information available at time t

- The Bellman equation

$$V(\Omega_t) = \max_{c_t} u(c_t) + \beta\mathbb{E}\left[V(\Omega_{t+1})|\Omega_t\right]$$

- Euler equation can be used for estimation of parameters in GMM framework (Hansen 1982)

$$u'(c_t) = \beta\mathbb{E}[u'(c_{t+1})|\Omega_t]$$

Plan of the Lecture

- Today: Full solution approach
  - Rust (1987, Econometrica) in IO
  - (not covered) Keane and Wolpin (1994, JPE) in labor
- Next lecture: Two step estimators
  - Hotz and Miller (1993, REStud), Agguiregabiria and Mira (2002, EMA), Bajari, Benkard, and Levine (2007, EMA)
- Applications if time allows...

Single-Agent Dynamic Models

- Rust (1987) is one of the first papers in this literature.
- Model is quite simple, but empirical framework introduced in this paper for dynamic discrete-choice (DDC) models is still widely applied.

- Agent is Harold Zurcher, manager of bus depot in Madison, Wisconsin.
  - Each week, HZ must decide whether to replace the bus engine, or keep it running for another week.

Rust (1987)

- Engine replacement problem is an example of an optimal stopping problem, which features the usual tradeoff:
  - there are large fixed costs associated with "stopping" (replacing the engine),
  - but new engine has lower associated future maintenance costs.
- Optimal solution characterized by a threshold type of rule above which it is optimal to replace.

Model

- x_t: the bus engine's mileage
  - Limit 90 possible values for computational reasons.
  - State variable
- Action $i_t \in \{0, 1\}$ where
  - $i_t = 1$ if replace the engine
  - $i_t = 0$ if keep the engine and perform normal maintenance
- Note: For simplicity, focus on the case of only one bus (treated as independent entities in the paper)

- Per-period profit function

$$u(i_t, x_t; \theta_1) = \begin{cases} -c(x_t, \theta_c) + \epsilon_t(0) & \textit{if } \ i_t = 0 \\ -(RC - c(0, \theta_c)) + \epsilon_t(1) & \textit{if } \ i_t = 1 \end{cases}$$

  - $c(x_t, \theta_c)$: regular maintenance costs (including expected breakdown costs)
  - RC: costs of replacing an engine
  - $\{\epsilon_t(0), \epsilon_t(1)\}$: payoff shocks
  - x_t is observable to both agent and econometrician
  - ϵ_t is only observable to the agent
- ϵ_t is necessary for a coherent model, for sometimes we observe the agent making different decisions for the same value of x
- State variables: (x_t, ϵ_t)
  - Assume the first-order Markovian structure of state transition

$$P(x_{t+1}, \epsilon_{t+1} | x_t, \epsilon_t, i_t)$$

  - Note: x_t satisfies Markov property if $Pr(x_t | x_0, x_1, \cdots, x_{t-1}) = Pr(x_t | x_{t-1})$

Bellman Equation

- Bellman equation

$$V(x_t, \epsilon_t) = \max_i \underbrace{u(i, x_t, \theta_1)}_{\tilde{u}(i, x_t, \theta_1) + \epsilon_t(i)} + \beta EV(x_t, \epsilon_t, i)$$

  where

$$EV(x_t, \epsilon_t, i) = \int V(x', \epsilon') dP(x', \epsilon' | x_t, \epsilon_t, i)$$

  and P is the transition probability from (x, ϵ) to (x', ϵ'). We assume Markovian transition process.
- Choice probability:

$$\begin{aligned} P(i_t = 1 | x_t) &= E_{\epsilon_t} [\tilde{u}(1, x_t, \theta) + \beta EV(x_t, \epsilon_t, 1) + \epsilon_t(1) \\ &\quad > \tilde{u}(0, x_t, \theta) + \beta EV(x_t, \epsilon_t, 0) + \epsilon_t(0)] \end{aligned}$$

  - Almost similar to discrete choice except for $EV(x_t, \epsilon_t, 1)$, a solution to Bellman equation.

Parameters

- θ_1: parameters of cost function
- θ_2: parameters of distribution of ϵ (will be normalized)
- θ_3 : parameters of $x-$state transition function
- RC: replacement cost
- β: discount factor, will be imputed.
  - Hard to identify from the data in dynamic discrete choice models (Magnac and Thesmar 2002 EMA)

Econometric Problem

- The state variables:
  - x_t: mileage. Both agent (Harold Zurcher) and the econometrician observe this. Call this the "observed state variable".
  - ϵ_t: payoff shocks. Econometrician does not observe this. Call this the "unobserved state variable"
- Data: $\{i_t, x_t\}$ for $t = 1, \cdots, T$ for a particular bus.
- Assume that buses are homogenous and independent to pool observations.
  - A panel of bus-and-month level observations.

Difficulty: Computation

- Remember that

$$
\begin{aligned}
P(i_t = 1|x_t) \quad &= \quad E_{\epsilon_t}\left[\tilde{u}(1, x_t, \theta) + \beta EV(x_t, \epsilon_t, 1) + \epsilon_t(1)\right. \\
&\qquad \left. > \tilde{u}(0, x_t, \theta) + \beta EV(x_t, \epsilon_t, 0) + \epsilon_t(0)\right]
\end{aligned}
$$

where

$$
V(x_t, \epsilon_t) = \max_i \underbrace{u(i, x_t, \theta_1)}_{\tilde{u}(i, x_t, \theta_1) + \epsilon_t(i)} + \beta EV(x_t, \epsilon_t, i)
$$

$$
EV(x_t, \epsilon_t, i) = \int V(x', \epsilon') dP(x', \epsilon' | x_t, \epsilon_t, i)
$$

- In principle,
  - Solve the dynamic programming problem $\rightarrow$ get $EV(x_t, \epsilon_t, i)$
  - Use this to obtain $P(i_t = 1|x_t)$, write likelihood function.
- This is computationally difficult b.c.
  - Unobserved state variable ϵ_t may appear non-linearly.

Key Assumption: Conditional Independence

- Conditional independence: The transition density of $\{x_t, \epsilon_t\}$ satisfies

$$p(x', \epsilon' | x, \epsilon, i) = p^1(\epsilon' | x) \times p^2(x' | x, i)$$

- Interpretations:
  - Given x_t, ϵ_t is independent over time (no serial correlation)
  - Given x and i, x' is independent of ϵ
    - ϵ today has no direct effect on tomorrow's state.
    - x and i are sufficient to predict future status x'.

Likelihood

- Likelihood for a single bus

$$
\begin{aligned}
L(\{x_t, i_t\}_{t=1}^{T}|\theta) &= \prod_{t=1}^{T} Prob(i_t, x_t|\{x_\tau, i_\tau\}_{\tau=0}^{t-1}, \theta) \\
&= \prod_{t=1}^{T} Prob(i_t, x_t|x_{t-1}, i_{t-1}; \theta) \\
&= \prod_{t=1}^{T} Prob(i_t|x_t; \theta) \times \prod_{t=1}^{T} Prob(x_t|x_{t-1}, i_{t-1}; \theta_3)
\end{aligned}
$$

  - The second equality: The first-order Markovian feature of the problem
  - The third equality: Conditional independence assumption
  - Remember $\theta = (\theta_1, \theta_3)$
    - θ_1: payoff parameter, θ_3: transition parameter

- Note: we have a panel data of multiple buses. The likelihood

$$L\left(\left\{\{x_t^b, i_t^b\}_{t=1}^T\right\}_{b=1}^B |\theta\right) = \prod_{b=1}^B L(\{x_t^b, i_t^b\}_{t=1}^T|\theta)$$

where b is an index for bus. I omit b hereafter for simplicity.

Log-Likelihood

- The log-likelihood is additively separable in the two components:

$$\log(L) = \sum_{t=1}^{T} \log Prob(i_t|x_t;\theta) + \sum_{t=1}^{T} \log Prob(x_t|x_{t-1}, i_{t-1};\theta_3)$$

- State transition parameter θ_3 can be estimated separately.
- This two step estimation makes implementation easier.

Step 1: State transition of x_t

- When $i_t = 0$ (not replacement)

$$p(x_{t+1} = x_t + 0 | x_t, i_t = 0, \theta_3) = \theta_{30}$$
$$p(x_{t+1} = x_t + 1 | x_t, i_t = 0, \theta_3) = \theta_{31}$$
$$p(x_{t+1} = x_t + 2 | x_t, i_t = 0, \theta_3) = 1 - \theta_{30} - \theta_{31}$$

- When $i_t = 1$ (replacement)

$$p(x_{t+1} = 0 | x_t, i_t = 1, \theta_3) = \theta_{30}$$
$$p(x_{t+1} = 1 | x_t, i_t = 1, \theta_3) = \theta_{31}$$
$$p(x_{t+1} = 2 | x_t, i_t = 1, \theta_3) = 1 - \theta_{30} - \theta_{31}$$

- These parameters can be easily estimated (without solving DP)

Step 2: Payoff parameter

- Remember

$$V(x_t, \epsilon_t) = \max_i \tilde{u}(i, x_t, \theta_1) + \epsilon_t(i) + \beta EV(x_t, \epsilon_t, i_t)$$

$$EV(x_t, \epsilon_t, i_t) = \int V(x', \epsilon') dP(x', \epsilon' | x_t, \epsilon_t, i_t)$$

- Conditional independence implies

$$\begin{aligned} EV(x_t, \epsilon_t, i_t) &= \mathbb{E}\left[V(x_{t+1}, \epsilon_{t+1}) | x_t, \epsilon_t, i_t\right] \\ &= \int V(x_{t+1}, \epsilon_{t+1}) dP(\epsilon_{t+1}) dP(x_{t+1} | x_t, i_t) \\ &= \int V_\sigma(x_{t+1}) dP(x_{t+1} | x_t, i_t) \end{aligned}$$

where $V_\sigma(x_{t+1}) \equiv \int V(x_{t+1}, \epsilon_{t+1}) dP(\epsilon_{t+1})$.

- Further assumption: i.i.d. logit error of $\{\epsilon_t(1), \epsilon_t(2)\}$.

Role of logit error

- Remember that expected value of utility under logit

$$\begin{aligned} V_\sigma(x_t) &= \mathbb{E}_{\epsilon_t}[V(x_t, \epsilon_t)] \\ &= \log\left(\exp(v_0(x_t)) + \exp(v_1(x_t))\right) \end{aligned}$$

- Define **the choice-specifc value function**

$$v_i(x_t) = \tilde{u}(i, x_t, \theta_1) + \beta \int V_\sigma(x_{t+1}) dP(x_{t+1}|x_t, i)$$

- Thus, we can write down the Bellman equation in terms of $V_\sigma(x_t)$, which is one-dimensional!

$$V_\sigma(x_t) = \log\left[\sum_{i\in\{0,1\}} exp(\tilde{u}(i, x_t, \theta_1) + \beta \int V_\sigma(x_{t+1}) dP(x_{t+1}|x_t, i))\right]$$

Choice Probability

- Under logit assumption:

$$Prob(i_t|x_t) = \frac{\exp(v_i(x_t))}{\exp(v_0(x_t)) + \exp(v_1(x_t))}$$

where

$$v_i(x_t) = \tilde{u}(i, x_t, \theta_1) + \beta \int V_\sigma(x_{t+1}) dP(x_{t+1}|x_t, i)$$

and V_σ can be obtained by Bellman equation

$$V_\sigma(x_t) = \log \left[\sum_{i \in \{0,1\}} exp(\tilde{u}(i, x_t, \theta_1) + \beta \int V_\sigma(x_{t+1}) dP(x_{t+1}|x_t, i)) \right]$$

  - Note: we already have estimated $P(x_{t+1}|x_t, i)$

Estimation Steps

1. Set a discount factor β
2. Estimate the state transition parameter θ_3.
3. Search over $\theta_1 = (RC, \theta_c)$ to maximize the likelihood function. When evaluating the likelihood for each candidate value of θ_1
   3.1 Solve the Bellman equation by contraction mapping for a candidate θ_1. Obtain $V_\sigma(x)$
   3.2 Use this to calculate the choice probabilities, evaluate log-likelihood.

Nested Fixed-Point Algorithm

- Outer loop: Search over different parameter values $\hat{\theta}_1$
- Inner loop: For a guess $\hat{\theta}_1$, compute the value function by solving the Bellman equation, then compute the log-likelihood.

- Note 1: Procedure is in spirit similar to BLP method, in which we first solve ξ_{jt} by contraction mapping, and then evaluate the objective function.

- Note 2: The logit assumption helps a lot. Computation of probabilities and the integration of optimal choices in the Bellman equation (log-sum expression)

- Note 3: Computation can still be intense. We will discuss alternatives in the next lecture.

Bus Data

TABLE I

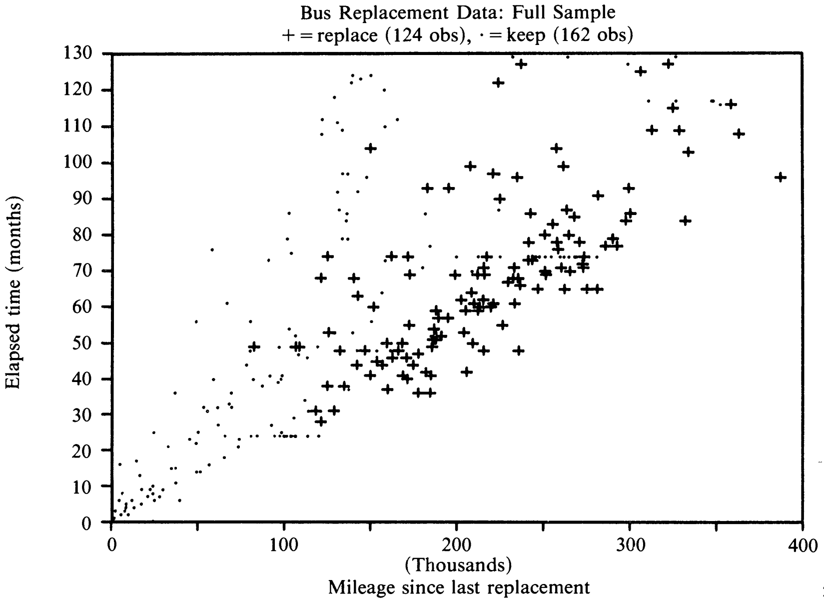
BUS TYPES INCLUDED IN SAMPLE

Bus Group	Number of Buses	Manufacturer	Engine	Model	Year	Seats	Empty Weight	Purchase Price	Estimated Value as of 10/1/84
1	15	Grumman	V6-92 series	870	1983	48	25,800	$145,097	$145,097
2	4	Chance	3208 CAT	RT-50	1981	10*	N.A.	100,775	124,772
3	48	GMC	8V71	T8H203	1979	45	25,027	92,668	125,000
4	37	GMC	8V71	5308A	1975	53	20,955	62,506	55,000
5	12	GMC	8V71	5308A	1974	53	20,955	49,975	48,000
6	10	GMC	6V71	4523A	1974	45	19,274	45,704	48,000
7	18	GMC	8V71	5308A	1972	51	20,955	43,856	45,000
8	18	GMC	6V71	4523A	1972	45	19,274	40,542	40,000

Note All buses are diesel powered and have air conditioning

* Handicap bus, outfitted with 4 long benches and accommodations for 6 wheelchairs.

Replacement Pattern



Parameter Estimates

TABLE X

STRUCTURAL ESTIMATES FOR COST FUNCTION $c(x, \theta_1) = .001\theta_{11}x$

FIXED POINT DIMENSION = 175

(Standard errors in parentheses)

Parameter		Data Sample			Heterogeneity Test	
Discount Factor	Estimates Log-Likelihood	Groups 1, 2, 3 3864 Observations	Group 4 4292 Observations	Groups 1, 2, 3, 4 8156 Observations	LR Statistic (df = 6)	Marginal Significance Level
$\beta = .9999$	RC	11.7257 (2.597)	10.896 (1.581)	9.7687 (1.226)	237.53	1.89E − 48
	θ_{11}	2.4569 (.9122)	1.1732 (0.327)	1.3428 (0.315)		
	θ_{30}	.0937 (.0047)	.1191 (.0050)	.1071 (.0034)		
	θ_{31}	.4475 (.0080)	.5762 (.0075)	.5152 (.0055)		
	θ_{32}	.4459 (.0080)	.2868 (.0069)	.3621 (.0053)		
	θ_{33}	.0127 (.0018)	.0158 (.0019)	.0143 (.0013)		
	LL	−3993.991	−4495.135	−8607.889		
$\beta = 0$	RC	8.2969 (1.0477)	7.6423 (.7204)	7.3113 (0.5073)	241.78	2.34E − 49
	θ_{11}	56.1656 (13.4205)	36.6692 (7.0675)	36.0175 (5.5145)		
	θ_{30}	.0937 (.0047)	.1191 (.0050)	.1070 (.0034)		
	θ_{31}	.4475 (.0080)	.5762 (.0075)	.5152 (.0055)		
	θ_{32}	.4459 (.0080)	.2868 (.0069)	.3622 (.0053)		
	θ_{33}	.0127 (.0018)	.0158 (.0019)	.0143 (.0143)		
	LL	−3996.353	−4496.997	−8614.238		
Myopia tests:	LR Statistic ($df = 1$)	4.724	3.724	12.698		
$\beta = 0$ vs. $\beta = .9999$	Marginal Significance Level	0.0297	0.0536	.00037		

Next Lecture

- Nested-fixed point approach (or full-solution approach) is still computationally demanding.
- Next week: Introduce two-step estimators that avoids solving the Bellman equation in estimation.