Estimating Dynamic Programming Models

Katsumi Shimotsu[1] Ken Yamada[2]

[1]Department of Economics
Hitotsubashi University

[2]School of Economics
Singapore Management University

Dynamic Programming (Structural) Models

- For assessing various policy proposals (e.g. pension reform, export subsidy), understanding the dynamic response of individuals and firms is important.
- Regression models are subject to the Lucas critic.
- It is desirable to model dynamic optimizing behavior on individuals/firms explicitly, along with how the state of economy evolves.

Dynamic Programming Models: Examples

- Rust (1987): bus engine replacement
- Rust and Rothwell (1995): nuclear plant operation
- Rust and Phelan (1997): retirement decision and pension/health plan
- Keane and Wolpin (1997): schooling, work and occupational choice
- Imai and Keane (2004): labor hours choice with human capital accumulation

Possible to quantitatively assess the impact of public policies that have never been implemented (counter-factual simulation)

Machine Replacement Model (Rust, 87)

Consider a mechanic who maintains a computer.
His objective is to maximize the discounted sum of utilities

$$\max_{a_1, a_2, \ldots} E\left[\sum_{t=1}^{\infty} \beta^t U(a_t, x_t, \epsilon_t; \theta)\right].$$

- $a_t \in \{0, 1\}$: machine replacement decision
- x_t: observable state variable (machine age)
- ϵ_t: state variable observable to the mechanic, but not to econometrician: additional information on machine condition
- θ: parameter of interest

Machine Replacement Model (Rust, 87)

- The utility function:

$$U(a_t, x_t, \epsilon_t; \theta) = -mc(x_t; \theta) - a_t rc(x_t; \theta) + \epsilon_t.$$

- $mc(x_t; \theta)$: machine maintenance cost
- $rc(x_t; \theta)$: machine replacement cost
- ϵ_t: unobserved (to econometrician) state variable
- Transition of x_t: $x_t = a_{t-1} + (1 - a_{t-1})(x_{t-1} + 1)$.
- The state variable evolves according to $(x_t, \epsilon_t) \sim p(x_t, \epsilon_t | x_{t-1}, \epsilon_{t-1}, a_{t-1}; \theta)$

Machine Replacement Model (Rust, 87)

$$\max_{a_1,a_2,\ldots} E\left[\sum_{t=1}^{\infty}\beta^t U(a_t,x_t,\epsilon_t;\theta)\right], \quad (x_t,\epsilon_t)\sim p(x_t,\epsilon_t|x_{t-1},\epsilon_{t-1},a_{t-1};\theta)$$

- We want to estimate θ from the data $\{a_t, x_t\}$,$t = 1,\ldots,n$.
- From the mechanic's point of view, he can solve this problem to obtain an optimal decision rule: $a = \delta(x,\epsilon;\theta)$.
- Because ϵ is unobservable to econometrician, the empirical implication of the model is the conditional choice probabilities (cf. probit model):

$$P(a|x;\theta) = \int I\{a = \delta(x,\epsilon;\theta)\}\, g(d\epsilon)$$

How to compute $P(a|x;\theta)$

- Let $s_t = (x_t, \epsilon_t)$. Define the value function as

$$W_\theta(s_t) = \max_{a_t, a_{t+1}, \ldots} E\left[\sum_{s=t}^{\infty} \beta^s U(a_s, s_s; \theta) \middle| s_t\right].$$

- We may compute the value function by finding the fixed point of the Bellman equation

$$W_\theta(s_t) = \max_{a \in A}\left\{U(a, s_t, \theta) + \beta \int W_\theta(s_{t+1}) p(ds_{t+1}|s_t, a)\right\}.$$

- $U(a, s_t, \theta)$: today's utility
- $W_\theta(s_{t+1})$: maximum future utility when tomorrow's state is s_{t+1}

How to compute $P(a|x;\theta)$

- Bellman equation

$$W_\theta(s_t) = \max_{a \in A} \left\{ U(a, s_t, \theta) + \beta \int W_\theta(s_{t+1}) p(ds_{t+1}|s_t, a) \right\}.$$

- When ϵ_t is continuously distributed, finding a fixed point of this Bellman equation becomes difficult

- The space of s_t can be very large $\Rightarrow$ need to evaluate $W_\theta(s)$ at many points

- Numerical integration in computing $\int W_\theta(s_{t+1}) p(ds_{t+1}|s_t, a)$

How to compute $P(a|x;\theta)$

- *Conditional Independence Assumption*:
  the transition of x_t and ϵ_t can be written as

$$p(x_{t+1}, \epsilon_{t+1}|s_t, a_t) = g(\epsilon_{t+1}|x_{t+1})f(x_{t+1}|x_t, a_t).$$

- Any statistical dependence between ϵ_t and ϵ_{t+1} is transmitted entirely through x_{t+1}.
- The probability density of x_{t+1} depends only on x_t and not ϵ_t. ϵ_t is a noise “superimposed" on x_t.

How to compute $P(a|x;\theta)$

- Define the integrated value function as

$$V_\theta(x_t) = \int W_\theta(x_t, \epsilon_t) dg(\epsilon_t | x_t)$$

- $V_\theta(x)$ satisfies another Bellman equation

$$\begin{aligned} & V_\theta(x) \\ = \quad & \max_{a \in A} \left\{ \int U(a, x, \epsilon, \theta) dg(\epsilon | x) + \beta \int V_\theta(x') f(dx' | x, a) \right\}. \end{aligned}$$

- $V_\theta(x)$ is a fixed point of a separate mapping on the reduced space of x rather than the space of $s = (x, \epsilon)$.

- Under some assumptions on $U(a, x, \epsilon)$, one does not need to use numerical integration to compute the right hand side.

How to compute $P(a|x;\theta)$

- Example: $U(a,x,\epsilon;\theta) = u(a,x;\theta) + \epsilon(a)$, and $\epsilon(a)$ follows extreme value dist'n (logit error), independent across a.

- Then one does not need to use numerical integration to compute the right hand side of the Bellman equation.

- Further, $P(a|x;\theta)$ admits a multinomial logit formula

$$P(a|x;\theta) = \frac{\exp[u(a,x;\theta) + EV_\theta(a,x)]}{\sum_{j=1}^{J} \exp[u(j,x;\theta) + EV_\theta(j,x)]},$$

  where $EV_\theta(a,x) = \int V_\theta(x')f(dx'|x,a)$ = the maximum future utility when the pair of the current action and observable state is (a,x)

Nested Fixed Point (NFXP) algorithm (Rust, 87)

Given a (cross-sectional) data set $\{a_i, x_i\}_{i=1}^n$, the NFXP solves

- Inner loop: for each candidate value of θ, solve the integrated Bellman equation to compute $EV_\theta(a, x)$ and then compute the equilibrium conditional choice probabilities $P(a|x; \theta)$.
- Outer loop: $\max_\theta N^{-1} \sum_{i=1}^N \ln P(a_i|x_i; \theta)$
- Then, $\hat{\theta}$ is the MLE.
- Computational cost depends on the size of the state space (the support points of x and a).
- Computationally intensive when x may take many different values.

Alternate fixed point mapping

- Hotz and Miller (1993, ReStud), Aguirregabiria and Mira (2002, ECMA; 2007, ECMA)
- Computationally attractive alternative to the NFXP
- These procedures use the fixed point in the space of $P(a|x)$: the conditional choice probability of action a given x.
- Unlike the value function, we can obtain a good “guess" of $P(a|x)$ from the data, typically from a frequency estimator.

Relation between value function and $P(a|x)$

- Given the value function, computing $P(a|x)$ is not very difficult (logit formula).

- But computing the value function

$$W(s_t) = \max_{a_t, a_{t+1}, \ldots} E\left[\sum_{s=t}^{\infty} \beta^s U(a_s, s_s; \theta) \middle| s_t \right]$$

directly from $P(a|x)$ is difficult, even with simulations.

- Assume additive separability:

$$U(a, s) = u(a, x) + \epsilon.$$

Alternate fixed point mapping

- The Bellman equation

$$W(s_t) = \max_{a \in A} \left\{ u(x_t, a) + \epsilon(a) + \beta \int W(s_{t+1}) p(ds_{t+1} | s_t, a) \right\}.$$

- Define the integrated value function as

$$V(x_t) = \int W(x_t, \epsilon_t) dp(\epsilon_t | x_t)$$

- With additive separability, $V(x)$ satisfies another Bellman equation

$$\begin{aligned} & V(x) \\ = & \int \max_{a \in A} \left\{ u(x, a) + \epsilon(a) + \beta \int V(x') f(dx' | x, a) \right\} dp(\epsilon_t | x). \end{aligned}$$

Alternate fixed point mapping

- Mapping from $P(a|x)$ to the value function

$$V(x) = \sum_{a \in A} P(a|x) \left\{ u(x,a) + E[\epsilon(a)|x,a] + \beta \int_X V(x')f(dx'|x,a) \right\}$$

where $E[\epsilon(a)|x,a]$ is the expected value of ϵ conditional on action a is chosen.

- $E[\epsilon(a)|x,a]$ admits a simple closed-form representation.
- Combining this with the mapping from the value function to P gives a mapping $\Psi(\theta, P)$: $P(a|x)_{\{a,x\}} \rightarrow P(a|x)_{\{a,x\}}$.
- True P is the fixed point of this mapping.

Nested pseudo-likelihood (NPL) estimator (Aguirregabiria and Mira, 02, 07)

- Define $P = P(a|x)_{\{a,x\}}$: vector of conditional choice probabilities for all a and x.

- Start from a guess of P, $\tilde{P}_0$. Typically, a frequency estimator.

- Estimate θ by

$$\tilde{\theta}_1 = \arg\max \sum_{i=1}^{n} \ln \Psi(\theta, \tilde{P}_0)(a_i|x_i).$$

- Update P by $\tilde{P}_1 = \Psi(\tilde{\theta}_1, \tilde{P}_0)$.

- Update θ and P further by $\tilde{\theta}_2 = \arg\max \sum_{i=1}^{n} \ln \Psi(\theta, \tilde{P}_1)(a_i|x_i)$, $\tilde{P}_2 = \Psi(\tilde{\theta}_2, \tilde{P}_1)$, ...

Nested pseudo-likelihood (NPL) estimator (Aguirregabiria and Mira, 02, 07)

- Once P is fixed, evaluating $\Psi(\theta, P)$ for different θ does not require computing a fixed point.
- In a single agent model (for example, bus engine model), $\tilde{\theta}_1$ is first-order equivalent to the MLE.
- $\tilde{\theta}_j$ approaches to the MLE in a higher-order sense as j increases. (Kasahara and Shimotsu, 2008)
- In multiple agent (game-theoretic) model, $\tilde{\theta}_1$ is not the MLE. Iterating the updating may give a better estimator.

Game-theoretic models

- We may view the bus engine model as a model with a fixed point constraint
- $$\hat{\theta}_{ML} = \arg\max_{\theta \in \Theta} \frac{1}{n} \sum_{i=1}^{n} \ln P\left(a_i | x_i; \theta\right)$$
  $$s.t. \qquad P = \Psi(\theta, P)$$
- Game-theoretic models are also characterized by a fixed point.

Game-theoretic models

- Many markets are characterized as a competition among a few firms with differentiated products.
- In such markets, strategic interaction between firms become important.
- We want to understand how firms compete with each other in such markets.
- Implications for competition policy, entry regulation, etc.
- “Structural economic models" explicitly model firms’ strategic interaction (and dynamic choice)

Simple static game of entry

- Two potential entrants to a market (for example, 5th generation smartphone, Docomo and KDDI).
- If both firms enter, both firms earn positive profit. If only one firm enters, it earns larger profit.
- Payoff matrix

		firm 2	
		in	out
firm 1	in	$(1, 1)$	$(3, 0)$
	out	$(0, 3)$	$(0, 0)$

- Nash equilibrium: given the other player's action, I have no incentive to change my action $\Rightarrow$ (in,in)

Games with private information

- But firm 1 may not know everything about firm 2.
- Payoff matrix with a random component

		firm 2	
		in	out
firm 1	in	$(\varepsilon_1 - \theta, \varepsilon_2 - \theta)$	$(\varepsilon_1, 0)$
	out	$(0, \varepsilon_2)$	$(0, 0)$

  ε_1 and ε_2 are drawn independently from uniform distribution$[0, 1]$. Both firms know θ.

- Only firm 1 observes ε_1, and only firm 2 observes ε_2 (private information).

How to determine the equilibrium

		firm 2	
		in	out
firm 1	in	$(\varepsilon_1 - \theta, \varepsilon_2 - \theta)$	$(\varepsilon_1, 0)$
	out	$(0, \varepsilon_2)$	$(0, 0)$

- Suppose firm 1 thinks that firm 2 enters with probability P_2 = firm 1's subjective probability ("belief")
- Firm 1 enters the market if
  $P_2(\varepsilon_1 - \theta) + (1 - P_2)\varepsilon_1 > 0 \Rightarrow \varepsilon_1 > \theta P_2$
- Because $\varepsilon_1 \sim U[0, 1]$, firm 1 enters the market with probability $1 - \theta P_2$ when his belief is P_2.

How to determine the equilibrium

- Firm 1 enters the market with probability $1 - \theta P_2$ when his belief is P_2.
- Firm 2 enters the market with probability $1 - \theta P_1$ when his belief is P_1.
- Bayesian perfect equilibrium: both players' belief must be consistent with their action.

$$P_1 = 1 - \theta P_2, \quad P_2 = 1 - \theta P_1 \Rightarrow P_1 = P_2 = 1/(1+\theta)$$

- Best response mapping: $[0,1]^2 \rightarrow [0,1]^2$
  $\Psi(\theta, P) = (1 - \theta P_2, 1 - \theta P_1)'$.
  Then BPE is a fixed point of $\Psi : P^* = \Psi(\theta, P^*)$

Estimation of θ in this model

- Suppose we have iid data of the entry decision (a_{1i}, a_{2i}) for $i = 1, \ldots, n$, and we want to estimate the value of θ.

- If we assume the data are in the equilibrium,
  we can estimate θ by

$$\hat{\theta} = \frac{1}{2}\left[\frac{1-\hat{P}_1}{\hat{P}_1} + \frac{1-\hat{P}_2}{\hat{P}_2}\right].$$

Dynamic discrete game

- N firms = potential entrants
- Entry/exit choice: $a_{it} \in A = \{0, 1\}$.
- Firm i's profit in period t:

$$\tilde{\Pi}_i(a_t, S_t, a_{i,t-1}, \epsilon_t; \theta)$$

  - All the firms' current decision: $a_t = (a_{1t}, ..., a_{Nt})'$
  - Market demand condition: S_t (observable)
  - Past entry decision: $a_{i,t-1}$
  - Private shocks: $\epsilon_t = (\epsilon_{1t}, ..., \epsilon_{Nt})'$.
- θ: parameter of interest

Dynamic discrete game

- Profit function of firm i:

$$\tilde{\Pi}_i(a_t, S_t, a_{t-1}, \epsilon_{it}; \theta) = \theta_{RS} \ln S_t - \theta_{FC,i} - \theta_{EC}(1 - a_{i,t-1})$$

$$- \theta_{RN} \ln(1 + \sum_{j \neq i} a_{jt}) \; + \; \epsilon_{it1}$$

- θ_{RS}: Revenue parameter
- $\theta_{FC,i}$: Operating cost
- θ_{EC}: Entry cost
- θ_{RN}: Degree of strategic substitution

Dynamic discrete game

- Dynamic optimization by firm i

$$\max_{a_{i1},a_{i2},\ldots} \quad E\left[\sum_{t=0}^{\infty}\beta^t \Pi_i(a_t, S_t, a_{i,t-1}, \epsilon_t;\theta) \middle| S_t, a_{t-1};\theta\right]$$

- Assume the state variable follows a Markov process
- Markov decision problem given his belief
- Stationary solution

Dynamic discrete game

- Empirical implication: firm i's conditional choice probabilities $= P_i(a|x)$.
- $P_i = \{P_i(a|x)\}_{(a,x)}$: firm i's conditional choice probabilities for all possible x
- The conditional choice probabilities of all the firms: $P = (P_1, \ldots, P_N)$
- For a given θ, an equilibrium is characterized by a fixed point of the best response mapping

$$P = \Psi(\theta, P)$$

- We assume $\Psi(\theta, P)$ has a unique fixed point for the moment.

P can be large

For example,

- 5 potential entrants
- all the firms' entry status in the previous period: $2^5 = 32$ support points
- market condition takes 10 different values
- x_t takes $10 \times 32 = 320$ different values
- Length of $P = 320 \times 5 = 1600$

$\Rightarrow$ finding a fixed point of Ψ can be computationally costly

Economic models with a fixed point constraint

- When $P = P(a|x)$ is the choice probability of a discrete action a conditional on x, the log-likelihood function is

$$Q_n(\theta) = \sum_{i=1}^{n} \ln P(a_i|x_i) \quad s.t. \quad P = \Psi(\theta, P)$$

Constraint optimization approach (Su and Judd, 2012)

- Write down the Lagrangian

$$\mathcal{L}(\theta, P, \lambda) = \frac{1}{n}\sum_{i=1}^{n} \ln P\left(a_i|x_i\right) + \lambda(P - \Psi(\theta, P))$$

- Solve the first-order condition

$$\begin{aligned} \nabla_\theta \mathcal{L}(\theta^*, P^*, \lambda^*) &= 0, \\ \nabla_P \mathcal{L}(\theta^*, P^*, \lambda^*) &= 0, \\ P^* - \Psi(\theta^*, P^*) &= 0 \end{aligned}$$

- According to the authors, one can solve this problem using the “NEOS Server, a free internet service which gives the user access to several state-of-the-art solvers."

NPL estimator (Aguirregabiria and Mira, 2007)

- $\tilde{P}_0$: initial guess of P

- **Step 1:** Given $\tilde{P}_{k-1}$,

$$\frac{1}{n}\sum_{i=1}^{n} \ln[\Psi(\theta, \tilde{P}_{k-1})](a_i|x_i).$$

  can be viewed as a pseudo log-likelihood function. So, estimate θ by maximizing this objective function $\Rightarrow \tilde{\theta}_k$

- **Step 2:** Given $\tilde{\theta}_k$, update the estimate of P by

$$\tilde{P}_k = \Psi(\tilde{\theta}_k, \tilde{P}_{k-1}).$$

- Iterate Steps 1-2: $\{\tilde{\theta}_k, \tilde{P}_k\}_{k=1}^{\infty}$

NPL Fixed Points: $\{\check{\theta}, \check{P}\}$

- If the sequence $\{\tilde{\theta}_k, \tilde{P}_k\}_{k=1}^{\infty}$ converges, its limit satisfies

$$
\begin{aligned}
\check{\theta} &= \arg\max_{\theta} \frac{1}{n} \sum_{i=1}^{n} \ln[\Psi(\theta, \check{P})](a_i|x_i), \\
\check{P} &= \Psi(\check{\theta}, \check{P}).
\end{aligned}
$$

- The NPL estimator is defined as $(\check{\theta}, \check{P})$ that achieves the highest pseudo-likelihood value.

- Easy to implement: standard optimization and policy iteration

Convergence of the NPL algorithm?

- Convergence of $\{\tilde{\theta}_k, \tilde{P}_k\}_{k=1}^{\infty}$?
- Empirical researchers report non-convergence of the NPL algorithm.
- For example, it can exhibit a 2-period cycle.

Property of NPL updating (Kasahara and Shimotsu, 2012)

- In a neighborhood of P^0,

$$\tilde{\theta}_j - \hat{\theta}_{NPL} = O_p(||\tilde{P}_{j-1} - \hat{P}_{NPL}||),$$
$$\tilde{P}_j - \hat{P}_{NPL} = M_{\Psi_\theta}\Psi_P(\tilde{P}_{j-1} - \hat{P}_{NPL}) + \text{smaller order terms},$$

  where

$$\Psi_P = \nabla_P \Psi(\theta, P) = \text{ Jacobian of } \Psi(\theta, P)$$
$$M_{\Psi_\theta} = I - \Psi_\theta(\Psi_\theta' \Delta_P \Psi_\theta)^{-1}\Psi_\theta' \Delta_P$$

- The convergence of $\tilde{P}_k$ depends on the dominant eigenvalue of $M_{\Psi_\theta}\Psi_P$.

Dynamic Game Example (continued)

- Profit function of firm i:

$$\tilde{\Pi}_i(a_t, S_t, a_{t-1}, \epsilon_{it}; \theta) = \theta_{RS} \ln S_t - \theta_{FC,i} - \theta_{EC}(1 - a_{i,t-1})$$
$$- \theta_{RN} \ln(1 + \sum_{j \neq i} a_{jt}) + \epsilon_{it1}$$

- θ_{RS}: Revenue parameter
- $\theta_{FC,i}$: Operating cost
- θ_{EC}: Entry cost
- θ_{RN}: Degree of strategic substitution

Dominant eigenvalue of Ψ_P and $M_{\Psi_\theta}\Psi_P$

θ_{RN}	$\rho(\Psi_P)$	$\rho(M_{\Psi_\theta}\Psi_P)$
1	0.337	0.292
2	0.693	0.595
4	1.184	1.180
6	1.479	1.478

Note: $\dim(P) = 144$ and $\dim(\theta) = 2$.

Strong strategic substitutability

$\rightarrow \{\tilde{\theta}_k, \tilde{P}_k\}$ diverges away from (θ^0, P^0).

What if Ψ is not locally contractive around P^0? (Kasahara and Shimotsu, 2012)

(1) Relaxation method:

$$[\Lambda(\theta, P)](a|x) = \{[\Psi(\theta, P)](a|x)\}^{\alpha} P(a|x)^{(1-\alpha)}, \alpha \in (0, 1)$$

— easy, works in some cases

(2) Recursive Projection Method

References I

Aguirregabiria, V. and P. Mira (2002). Swapping the nested fixed point algorithm: a class of estimators for discrete Markov decision models. Econometrica 70, 1519-1543.

Aguirregabiria, V. and Mira, P. (2007). Sequential estimation of dynamic discrete games. Econometrica 75, 1-53.

Aguirregabiria, V. and Mira. P. (2010). Dynamic discrete choice structural models: A survey. Journal of Econometrics 156, 38-67.

Hotz, J. and R. A. Miller (1993). Conditional choice probabilities and the estimation of dynamic models. Review of Economic Studies 60, 497-529.

Imai, S., Keane, M. P. (2004). Intertemporal labor supply and human capital accumulation. International Economic Review 45, 601-641.

References II

Kasahara, H. and K. Shimotsu (2008) Pseudo-likelihood Estimation and Bootstrap Inference for Structural Discrete Markov Decision Models. Journal of Econometrics 146, 92-106.

Kasahara, H. and K. Shimotsu (2012) Sequential Estimation of Structural Models with a Fixed Point Constraint. Econometrica 80, 2303-2319.

Keane, M. P. and K. I. Wolpin (1997). The Career Decisions of Young Men. Journal of Political Economy, 105, 473-522.

Rust, J. (1987). Optimal replacement of GMC bus engines: an empirical model of Harold Zurcher. Econometrica 55, 999-1033.

Rust, J., Rothwell, G. (1995). Optimal Response to a Shift in Regulatory Regime: The Case of the US Nuclear Power Industry. Journal of Applied Econometrics 10, S75-S118.

References III

Rust, J., Phelan. C. (1997). How Social Security and Medicare Affect Retirement Behavior In a World of Incomplete Markets. Econometrica 65, 781-831.

Su, Che-Lin and Judd, K. L. (2012). Constrained optimization approaches to estimation of structural models. Econometrica 80, 2213-2230.