Lecture 3, 4: Dynamic Programming

Yuya Takahashi

April 9, 11, 2019

1 Introduction

- Dynamic programming has a wide range of applications in economics.

- Since most of DP problems do not have an analytical solution, this is exactly where we should apply the numerical methods we learn in class.

2 Dynamic Programming

- Problems are simple when the state space is finite.

- Interesting not only because they are simple but because we often solve harder problems by discretization.

- To fix the idea, we start with one of optimal growth models.

2.1 Example

- Deterministic infinite-horizon Ramsey model.

- A planner with K_0 chooses a sequence of future capital stocks $\{K_t\}_{t=1}^{\infty}$ to maximize the utility of a representative household

$$U_0 = \sum_{t=0}^{\infty} \beta^t u(C_t)$$

subject to the economy's resource constraint

$$F(N, K_t) \geq C_t + I_t$$

where

$$I_t = K_{t+1} - (1-\delta)K_t$$

and non-negativity constraints on C_t and K_{t+1}.

- For convenience, define $f(K_t) = F(N, K) + (1-\delta)K$.

- Assume both u and F are strictly concave and twice continuously differentiable.

- Rewrite the problem in the Bellman equation formulation

$$V(K) = \max_{0 \le K' \le f(K)} u(f(K) - K') + \beta V(K')$$

- Once we know $V(K)$, we can solve for optimal K' for any given K: $K' = h(K)$.
  Called the policy function.

- Discretize the space of K. Let $\mathcal{K} = \{K_1, K_2, ..., K_n\}$ with $K_i < K_{i+1}$. That is, we transform the problem from solving the functional equation in the space of continuous functions to finding a vector of n elements.

- Below, we also use h_i to denote the policy when the current capital stock is K_i.

- Fixed point algorithm (value function iteration):

  Step 1 Set the initial values $\mathbf{V}^0$ and choose $\epsilon > 0$.

  Step 2 For $i = 1, ..., n$, compute

$$V_i^{l+1} = \max_j u(f(K_i) - K_j) + \beta V_j^l$$

  Step 3 If $||V^l - V^{l+1}|| < \epsilon$ go to Step 4, otherwise go back to Step 2.

  Step 4 For $i = 1, ..., n$, compute the final solution

$$h_i^* = \arg\max_j u(f(K_i) - K_j) + \beta V_j^{l+1}$$

- Linearly converges with rate β. In the value function iteration, we do not use the policy function.

- Can we use policy functions effectively?

- The value resulting from following policy h is

$$V_i = u(f(K_i) - K_{h_i}) + \beta V_{h_i}$$

for $i = 1, ..., n$.

- Define Q with zeros everywhere except for its row i and column $j = h_i$ elements, which equal one.

- The above equation is written in matrix notation

$$\mathbf{V} = \mathbf{u} + \beta Q \mathbf{V},$$

with solution

$$\mathbf{V} = [I - \beta Q]^{-1}\mathbf{u}.$$

- Policy function iteration:

Step 1 Set the initial values $\mathbf{V}^0$ and choose $\epsilon > 0$.

Step 2 For $i = 1, ..., n$, find

$$h_i^{l+1} = \arg\max_j u(f(K_i) - K_j) + \beta V_j^l.$$

Step 3 Using this policy, find

$$u_i^{l+1} = u(f(K_i) - K_{h_i^{l+1}})$$

for $i = 1, ..., n$, and update Q^{l+1} based on h^{l+1}.

Step 4 Compute

$$\mathbf{V}^{l+1} = [I - \beta Q^{l+1}]^{-1}\mathbf{u}^{l+1}$$

Step 5 $||V^l - V^{l+1}|| < \epsilon$ stop, otherwise go back to Step 2.

- Let's assume $u(C_t) = \ln C_t$ and $F(N, K) = AK^\alpha$ with $0 < \alpha < 1$.

- It is known that

$$V(K) = A + B \ln K$$

where

$$A = \frac{1}{1-\beta}\left[\ln(1-\alpha\beta) + \frac{\alpha\beta}{1-\alpha\beta}\ln\alpha\beta\right]$$
$$B = \frac{\alpha}{1-\alpha\beta}.$$

- This is a rare case (Brock-Mirman model). We can compare the computed value function with this analytical solution.

3 Dynamic Markov Decision Process

3.1 Basic Framework

- Single agent decision problems can be viewed as a "game against nature"

- We will discuss dynamic models within a framework of Markov decision processes (MDP)

- Formally, a discrete-time MDP consists of the following objects

  - A time index $t \in \{0, 1, 2, ..., T\}, T \leq \infty$;
  - A state space S;
  - A decision space D;

- A family of constraint sets $\{D_t(s_t) \subseteq D\}$;

- A family of transition probabilities $\{p_{t+1}(\cdot|s_t, d_t)\}$;

- A discount factor $\beta \in (0, 1)$ and family of single period reward functions $\{(u_t(s_t, d_t)\}$ such that the utility functional U has the additively separable decomposition

$$U(\mathbf{s}, \mathbf{d}) = \sum_{t=0}^{T} \beta^t u_t(s_t, d_t),$$

where $\mathbf{s} = (s_0, ..., s_T)$ and $\mathbf{d} = (d_0, ..., d_T)$.

- The agent's optimization problem is to choose an optimal decision rule $\delta^* = (\delta_0, ..., \delta_T)$ to solve the following problem

$$\max_{\delta=(\delta_0,...,\delta_T)} E\left\{U(\mathbf{s}, \mathbf{d})\right\},$$

where expectation is with respect to the partially controlled stochastic process $\{s_t, d_t\}$ induced by the decision rule δ.

- A decision rule which has the property that it depends on the past history of the process only via the current state s_t is called *Markovian*.

- In many applications, we assume that the transition probabilities and utility functions do not directly depend on t, i.e., are the same for all t.

- These MDP's are said to be stationary.

- Notice that in finite horizon case $(T < \infty)$, stationarity does not help much as each period continuation value, $\sum_{j=t}^{T} \beta^j u(s_j, d_j)$, depends on t.

- In the infinite-horizon problems, stationary Markovian structure implies that the future looks the same whether the agent is in state s_t at time t or in state s_{t+k} at time $t + k$ provided that $s_t = s_{t+k}$.

- The only variable which affects the agent's view about the future is the current value of the state, s. Therefore, the optimal decision rule and corresponding value function should be time invariant.

- It is convenient to write the dynamic problem in the form of a Bellman equation

$$V(s_t) = \max_{d_t \in D(s_t)} \left\{ u(s_t, d_t) + \beta \int V(s_{t+1}) p(s_{t+1} | s_t, d_t) \right\}$$

where optimal decision rule satisfies

$$\delta(s_t) = \arg \max_{d_t \in D(s_t)} \left\{ u(s_t, d_t) + \beta \int V(s_{t+1}) p(s_{t+1} | s_t, d_t) \right\}$$

- The idea behind Bellman formulation is simple: let $V(s_t)$ be the value of the maximized objective function

$$\max_{\delta} E \left[\sum_{t=0}^{\infty} \beta^t u(s_t, d_t) \right]$$
$$\text{s.t. } s_{t+1} \overset{i.i.d.}{\sim} P(\cdot | s_t, d_t)$$

when the initial state is s_t.

- Then a choice d_t in the current period yields current period reward $u(s_t, d_t)$ and implies that the system next period will be in the state s_{t+1}, where s_{t+1} will be chosen by "nature" according to the distribution assumed above.

- The maximum expected payoff that can be obtained from this position, with s_{t+1} unknown, is $\beta E[V(s_{t+1})|s_t, d_t]$. In other words, we split the problem into current reward and continuation value.

- Under some regularity conditions (i.e. $u(s, d)$ is jointly continuous and bounded in (s, d) and $D(s)$ is a continuous correspondence) it is possible to show that

$$\Gamma(W)(s) = \max_{d \in D(s)} \left\{ u(s, d) + \beta \int W(s')p(s'|s, d) \right\}$$

where Γ is an operator $\Gamma : C(S) \to C(S)$ and $C(S)$ is a Banach space, is a contraction mapping which, by Contraction Mapping Theorem, has a unique fixed point.

- Solution methods:

  - Finite horizon MDP's are typically solved by backward recursion, i.e. starting from the terminal period;

  - There is a wide variety of techniques for solving infinite-horizon problems. We mostly focus on policy/value functions iteration.

- Clearly, a solution to the MDP depends on the set of parameters. Therefore, a typical estimation procedure would search over the parameter vector such that the model predictions (optimal policy) matches observed behavior as close as possible.

3.2 Rust (1987): An empirical model of Harold Zurcher

- Harold Zurcher (HZ) is a superintendent of maintenance at the Madison (Wisconsin) Metropolitan Bus Company.

- HZ decides when to replace old bus engines with new ones.

- The engine replacement problem features a standard trade-off between minimizing replacement costs versus minimizing maintenance costs and unexpected engine failures.

- This is obviously a techniques paper as the actual application is not of significant interest, but

  - the model can be applied to other discrete capital investment decisions;

  - many of the techniques introduced can be used for other dynamic problems.

- General problem fits into "optimal stopping" framework: there is a critical cutoff mileage level x^* above which replacement takes place

- Data: Rust observes 162 busses in the fleet of Madison Metro over time. On each bus there are data on

  - monthly mileage;
  - date, mileage and list of components repaired or replaced each time a bus visits the company shop;

- On average, bus engines were replaced after 5 years with over 200,000 elapsed miles with considerable variation in the time and mileage at which replacement occurs.

- Idea is to

  - construct a (parametric) model which predicts the time and mileage at which engine replacement occurs;

  - using the model predictions (conditional on parameter values) find parameters that "fit" the data.

- Some assumptions of structural model:

  - mileage evolves exogenously;

  - HZ treats each bus independently in his decision making;

$\Rightarrow$ HZ decides only whether to replace engine or not and does not make joint replacement & utilization decisions.

3.2.1 Basic model setup

- Notation:

- θ denotes a set of unknown parameters to estimate.
- x_t (observed state variable) $=$ engine accumulated mileage at t.
- i_t is investment decision, i.e.

$$\begin{cases} i_t = 1, & \text{if replace engine in month } t; \\ i_t = 0, & \text{otherwise.} \end{cases}$$

- Decision set, $D(x_t)$: a finite set of allowable values of the control variable i_t when state variable is x_t.
- Suppose x_t evolves exogenously according to a first order Markov process, i.e., transition probability is given by $p(x_{t+1}|x_t, i_t, \theta)$ (same for every bus).

- Immediate reward function, $u(x_t, i_t, \theta)$: a single-period utility of decision i when state variable is x_t. Example:

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

where $c(x_t, \theta_1),\ s.t.\ \frac{\partial C(\cdot)}{\partial x} > 0$, is the expected costs of operating a bus with mileage x_t (including maintenance and social costs of breakdowns); R is the cost of replacement (i.e. a new engine) which occurs immediately. Note $\theta = (R, \theta_1)$.

- Assume that HZ makes decisions to maximize the expected present discounted sum of future utilities over an infinite horizon (β is a discount factor), i.e.

$$V(x_t, \theta) = \sup_{\Pi} E\left\{ \sum_{j=t}^{\infty} \beta^{j-t} [u(x_j, h_j, \theta)] \,\middle|\, x_t, \Pi, \theta \right\},$$

where $\Pi = \{h_t, h_{t+1}, h_{t+2}, ...\}$, $h_t \in D(x_t)$ for all t, and the expectation is taken over future x_t's, which evolve according to Markov process.

- Note that

  - the problem is formulated for one bus;

  - x_t completely summarizes state at time t: (1) the expected value of future utilities only depends on x_t and not on its past values, and (2) x_t is all that is relevant for the physical state of the bus.

- The problem above can be written in the form of Bellman equation,

$$V(x_t, \theta) = \max_{i_t} \left\{ u(x_t, i_t, \theta) + \beta E[V(x_{t+1}, \theta) | x_t, i_t, \theta] \right\},$$

  with the optimal policy function defined as

$$h(x_t, \theta) = \arg\max_{i_t} \{ u(x_t, i_t, \theta) + \beta E[V(x_{t+1}, \theta) | x_t, i_t, \theta] \}.$$

- Numerical solution to Bellman equation (Value function iteration):

  - start with any arbitrary function $V^0(\cdot)$;

- apply contraction mapping (to get $V^1(\cdot)$)

$$V^1(\cdot) = T(V^0(\cdot)) = \max_{d_t}\{u(s_t, d_t) + \beta E[V^0(\cdot)]\};$$

- repeat previous step with a new argument $V^1(\cdot)$ to get $V^2(\cdot)$;

- continue until the difference between $V^l(\cdot)$ and $V^{l+1}(\cdot)$ is below some tolerance level;

- resulting $V(\cdot)$ and corresponding policy function give the solution.

\- Example with a single discrete state variable x_t:

- suppose x_t is discrete, e.g. $x_t \in \mathbf{X} = \{1, ..., 10\}$ ($V(7)$ is value function when $x_t = 7$);

- start with arbitrary 10 numbers, e.g. $V(1) = V(2) = ... = V(10) = 0$;

- compute $V^1(\cdot) = T(V^0(\cdot))$ element by element, i.e.

$$V^1(1,\theta) = \max_{i_t}\left\{u(1,i_t,\theta) + \beta \sum_{x_{t+1}\in\mathbf{X}} V^0(x_{t+1},\theta)\,\Pr(x_{t+1}|1,i_t,\theta)\right\}$$

$$V^1(2,\theta) = \max_{i_t}\left\{u(2,i_t,\theta) + \beta \sum_{x_{t+1}\in\mathbf{X}} V^0(x_{t+1},\theta)\,\Pr(x_{t+1}|2,i_t,\theta)\right\}$$

$$..$$

$$..$$

$$..$$

$$V^1(10,\theta) = \max_{i_t}\left\{u(10,i_t,\theta) + \beta \sum_{x_{t+1}\in\mathbf{X}} V^0(x_{t+1},\theta)\,\Pr(x_{t+1}|10,i_t,\theta)\right\}$$

- which gives a new set of 10 numbers;
- repeat (now using $V^1(x_t,\theta)$ on the right hand side) to get $V^2(x_t,\theta)$;
- repeat until convergence, i.e. when $V^l(\cdot)$ and $V^{l+1}(\cdot)$ are close enough.

- If x_t is continuous variable one can

  - artificially discretize the problem, e.g. $x_t' = 1$ if $x_t \subset (0.5, 1.5)$, convert $p(x_{t+1}|x_t, i_t, \theta)$ to $\Pr(x_{t+1}'|x_t', i_t, \theta)$ and $u(x_t, i_t, \theta)$ to $u(x_t', i_t, \theta)$ and proceed as in the example above;

  - keep the problem continuous, but only solve for the value function at a finite number of points:

$$V^l(5, \theta) = \max_{i_t} \left\{ u(5, i_t, \theta) + \beta \int V^{l-1}(x_{t+1}, \theta) p(x_{t+1}|5, i_t, \theta) \right\}$$

    where we need to simulate an integral on the RHS;

  - one way is to use simulation (alternatively one can use quadrature approximation), i.e.

$$V^l(5, \theta) = \max_{i_t} \left\{ u(5, i_t, \theta) + \beta \frac{1}{NS} \sum_{ns} V^{l-1}(x_{t+1}^{ns}, \theta) \right\}$$

  - since x_{t+1}^{ns} are different from the points at which we have computed $V^{l-1}(\cdot)$, some sort of interpolation is required.

- As one might expect, given $\frac{\partial c(\cdot)}{\partial x_t}$ the optimal policy function has the form

$$h(x_t, \theta) = \begin{cases} 1 \text{ if } x_t \geq \gamma(\theta), \\ 0 \text{ if } x_t < \gamma(\theta) \end{cases}$$

  where the constant γ represents a threshold value of mileage (optimal stopping barrier) such that whenever current mileage on the bus exceeds γ it is optimal to replace the old bus engine.

- Rust (1987) shows that under specific choice of functional form for $p(x_{t+1}|x_t, i_t, \theta)$ (namely, that monthly mileage $(x_{t+1} - x_t)$ has an i.i.d. exponential distribution) it is possible to derive likelihood function, which is simply the probability density of the controlled stochastic process $\{i_t, x_t\}$.

- There are two major problems with this approach:

  1. Data might be inconsistent with the distributional assumption (which is actually the case for the Rust's data). There is no explicit solution to the optimal control problem if one wants to be more flexible with the distributional assumptions.

2. Earlier assumption that the physical state of a bus is completely characterized by a single variable, accumulated mileage x_t results in a degenerate hazard function for bus engine replacement: the probability of replacing a bus engine is 0 in the interval $(0, \gamma)$ and 1 thereafter. Note that this suggests that the engines on all buses must be replaced at the same level of mileage, which is completely refuted by the data: mileage at replacement varies from a minimum of 82,400 to a maximum 387,300.

- One way to solve the problem is to just add an error term ε in order to reconcile the difference between $h(x, \theta)$ and the observed choice i_t, i.e.,

$$i_t = h(x_t, \theta) + \varepsilon_t.$$

- Note that such an approach would only be internally consistent if the decision maker randomly departs from the optimal solution.

- A more plausible interpretation of observed deviations from the optimal stopping rule is that there are other unobserved by us but observed by the agent state variables.

- If we allow for unobserved state variables, i.e., if we assume that the agent's replacement decisions are based on other information ε_t which we have not observed, the dynamic model typically has no analytical solution.

- Rust (1987) suggests a way to deal with the problem by introducing an error term to a structural model with an interpretation of ε_t as an unobservable (by econometricians) state variable.

- Called "structural unobservables".

- With this, we have non-degenerate density. Without this, the likelihood would always be zero for all parameters.

3.2.2 Alternative model setup

- Update of the previous model

  - Decision set, $D(x_t)$: a finite set of allowable values of the control variable i_t when state variable is x_t.

  - Unobserved state variables, $\varepsilon_t = \{\varepsilon_t(i)|i \in D(x_t)\}$: a vector of state variables observed by agent but not by the econometrician. $\varepsilon_t(i)$ is a component of utility of an alternative i in period t which is known only by the agent.

  - Observed state variables, $x_t = \{x_t(1), ..., x_t(K)\}$: a $K-$dimensional vector of state variables observed by both the agent and the econometrician.

  - Immediate reward function, $u(x_t, i; \theta) + \varepsilon_t(i)$: a single-period utility of decision i when state variables are (x_t, ε_t). θ is a vector of unknown parameters to be estimated. Notice that ε_t enters utility function in an additively separable way.

- Transition probability, $p(x_{t+1}, \varepsilon_{t+1}|x_t, \varepsilon_t, i_t, \theta)$: Markov transition density for state variables (x_t, ε_t) when alternative i_t is selected.

- Assume that HZ makes decisions to maximize the expected present discounted sum of future utilities,

$$V(x_t, \varepsilon_t) = \sup_{\Pi} E\left\{\sum_{j=t}^{\infty} \beta^{j-t}[u(x_j, h_j, \theta) + \varepsilon_j(h_j)]|x_t, \varepsilon_t, \theta\right\},$$

where $\Pi = \{h_t, h_{t+1}, h_{t+2}, ...\}$, $h_t \in D(x_t)$ for all t, and the expectation is taken with respect to the controlled stochastic process $\{x_t, \varepsilon_t\}$.

- This infinite-horizon MDP can be written as a Bellman equation

$$V(x_t, \varepsilon_t) = \max_{i \in D(x_t)} \{u(x_t, i, \theta) + \varepsilon_t(i) + \beta EV(x_t, \varepsilon_t, i)\}$$

where

$$\begin{aligned} EV(x_t, \varepsilon_t, i) &\equiv E[V(x_{t+1}, \varepsilon_{t+1})|x_t, \varepsilon_t, i] \\ &= \int_y \int_\eta V(y, \eta) p(y, \eta|x_t, \varepsilon_t, i, \theta) \end{aligned}$$

is a function of (x_t, ε_t, i).

- Under certain regularity conditions (see Rust, 1988) the solution to this problem is given by a stationary decision rule, $i_t = h(x_t, \varepsilon_t, \theta)$, where optimal control h is defined by

$$h(x_t, \varepsilon_t, \theta) = \arg\max_{i \in D(x_t)} \{u(x_t, i, \theta) + \varepsilon_t(i) + \beta EV(x_t, \varepsilon_t, i)\}$$

3.2.3 Estimation

- In order to estimate parameters of the model Rust (1987) employs a nested fixed-point estimation algorithm

- The algorithm consists of two loops:

  - Outer loop: search over parameter values

- Inner loop: for a given set of parameters, solve DP problem and match predicted choices using likelihood function

- If we do not impose additional restrictions on $p(x_{t+1}, \varepsilon_{t+1}|x_t, \varepsilon_t, i_t, \theta)$, i.e. we allow serial correlation in the variables and allow x_t to be correlated with ε_t, we need to solve the problem using "brute force" approach, which is prohibitive.

- Additional (more restrictive) assumptions on $p(x_{t+1}, \varepsilon_{t+1}|x_t, \varepsilon_t, i_t, \theta)$ may significantly simplify computations.

- Conditional Independence (CI) assumption:

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

- Implications of the CI assumption:

- x_t is a sufficient statistic for ε_{t+1} $\Rightarrow$ any statistical dependence between ε_t and ε_{t+1} is transmitted entirely through the vector x_{t+1}.

- The probability density of x_{t+1} depends only on x_t and not ε_t.

\- CI is a strong assumption, but it provides considerable computational benefits:

1. $EV(\cdot)$ is not a function of ε_t, so that we avoid integration over the unknown function $EV(\cdot)$. Intuitively, under CI x_t summarizes all relevant information about the future.

2. $EV(\cdot)$ is a fixed point of a separate contraction mapping on the reduced state space.

\- Under CI, the Bellman equation can be written as

$$V(x_t, \varepsilon_t, \theta) = \max_{i \in D(x_t)} \left\{ u(x_t, i, \theta) + \varepsilon_t(i) + \beta \int_y \int_\eta V(y, \eta) p(y, \eta | x_t, i, \theta) \right\}$$

- Define "alternative specific" value functions

$$\begin{cases} \bar{V}^0(x_t, \theta) = u(x_t, 0, \theta) + \beta E[V(x_{t+1}, \varepsilon_{t+1}, \theta) | x_t, i_t = 0, \theta], \\ \bar{V}^1(x_t, \theta) = u(x_t, 1, \theta) + \beta E[V(x_{t+1}, \varepsilon_{t+1}, \theta) | x_t, i_t = 1, \theta] \end{cases} \tag{1}$$

- $\bar{V}^0(x_t)$ and $\bar{V}^1(x_t)$ stand for value of not replacing and replacing the bus engine, net of the error terms ε_{0t} and ε_{1t} respectively.

- Then value function for the problem can be written as

$$V(x_t, \varepsilon_t, \theta) = \max_{i_t} \left\{ \bar{V}^0(x_t, \theta) + \varepsilon_{0t}, \bar{V}^1(x_t, \theta) + \varepsilon_{1t} \right\} \tag{2}$$

- By moving (2) one period forward and substituting it into (1) we can write

$$\begin{cases} \bar{V}^0(x_t, \theta) = u(x_t, 0, \theta) + \beta E[\max\limits_{i_{t+1}} \left\{ \begin{array}{c} \bar{V}^0(x_{t+1}, \theta) + \varepsilon_{0t+1}, \\ \bar{V}^1(x_{t+1}, \theta) + \varepsilon_{1t+1} \end{array} \right\} | x_t, 0, \theta], \\ \bar{V}^1(x_t, \theta) = u(x_t, 1, \theta) + \beta E[\max\limits_{i_{t+1}} \left\{ \begin{array}{c} \bar{V}^0(x_{t+1}, \theta) + \varepsilon_{0t+1}, \\ \bar{V}^1(x_{t+1}, \theta) + \varepsilon_{1t+1} \end{array} \right\} | x_t, 1, \theta] \end{cases}$$

- Rust (1988) shows that the following is a joint contraction mapping

$$\begin{pmatrix} \bar{f}^0(x_t) \\ \bar{f}^1(x_t) \end{pmatrix} = \begin{pmatrix} u(x_t, 0, \theta) + \beta E[\max\limits_{i_{t+1}} \left\{ \begin{array}{c} \bar{f}^0(x_{t+1}) + \varepsilon_{0t+1}, \\ \bar{f}^1(x_{t+1}) + \varepsilon_{1t+1} \end{array} \right\} |x_t, 0, \theta], \\ \\ u(x_t, 1, \theta) + \beta E[\max\limits_{i_{t+1}} \left\{ \begin{array}{c} \bar{f}^0(x_{t+1}) + \varepsilon_{0t+1}, \\ \bar{f}^1(x_{t+1}) + \varepsilon_{1t+1} \end{array} \right\} |x_t, 1, \theta] \end{pmatrix},$$

and will converge to a unique fixed point $(\bar{V}^0(x_t, \theta), \bar{V}^1(x_t, \theta))$.

- Note that now the size of the state space is much smaller, i.e. solution can be obtained much quicker: if each of the state variables has 10 discrete points the "redefined" contraction is computed on 20 points, while the previous formulation would be computed on 1000 points.

- More generally, any state variable that do not affect continuation values do not have to be in the "effective" state space, i.e. the one we need to loop over in the contraction mapping.

- Under CI, we can write likelihood,

$$
\begin{aligned}
\mathcal{L}(\theta) &= \Pr(x_1, \ldots, x_T, i_1, \ldots, i_T | x_0, i_0, \theta) \\
&= \Pr(i_1 | x_1, x_0, i_0, \theta) p(x_1 | x_0, i_0, \theta) \\
&\times \Pr(i_2 | x_2, x_1, x_0, i_1, i_0, \theta) p(x_2 | x_1, x_0, i_1, i_0, \theta) \\
&\times \Pr(i_3 | x_3, x_2, x_1, x_0, i_2, i_1, i_0, \theta) p(x_3 | x_2, x_1, x_0, i_2, i_1, i_0, \theta) \times \cdots \\
&= \Pr(i_1 | x_1, \theta) p(x_1 | x_0, i_0, \theta) \Pr(i_2 | x_2) p(x_2 | x_1, i_1, \theta) \times \cdots \\
&= \prod_{t=1}^{T} \Pr(i_t | x_t, \theta) p(x_t | x_{t-1}, i_{t-1}, \theta)
\end{aligned}
$$

- Two components of the likelihood are

  - $p(x_t | x_{t-1}, i_{t-1}, \theta)$, which is a primitive of the model, i.e. assumption on how the state variable evolves; and

  - $\Pr(i_t | x_t, \theta)$, which is given by the solution to the dynamic programming problem, i.e.,

$$\Pr(i_t = 1|x_t, \theta) = \Pr\begin{pmatrix} u(x_t, 0, \theta) + \varepsilon_{0t} + \beta E[V(x_{t+1}, \varepsilon_{t+1})|x_t, \varepsilon_t] \leq \\ u(x_t, 1, \theta) + \varepsilon_{1t} + E[V(x_{t+1}, \varepsilon_{t+1})|x_t, \varepsilon_t] \end{pmatrix}$$
$$= \iint \mathbf{1}(\bar{V}^0(x_t, \theta) + \varepsilon_{0t} \leq \bar{V}^1(x_t, \theta) + \varepsilon_{1t}) p(\varepsilon_{0t}, \varepsilon_{1t}|x_t, \theta)$$

- Suppose that $p(x_{t+1}, \varepsilon_{t+1}|x_t, \varepsilon_t, i_t, \theta) = q(\varepsilon_{t+1}|\theta) p(x_{t+1}|x_t, i_t, \theta)$, i.e. independence of ε_t, which is not correlated with anything, then

$$\Pr(i_t = 1|x_t, \theta) = F_\varepsilon(\bar{V}^1(x_t, \theta) - \bar{V}^0(x_t, \theta), \theta)$$

- Significant simplification occurs if we assume that in addition to independence ε_t's are distributed Type 1 extreme values, then the gain is two-fold:

  - in the DP problem (recall property of iid extreme value)

$$E \max\{\delta_1 + \varepsilon_1, \ldots, \delta_J + \varepsilon_J\} = \mu + \ln\left(\sum_{j=1}^{J} \exp(\delta_j)\right)$$

then

$$\begin{cases} \bar{V}^0(x_t,\theta) = u(x_t,0,\theta) + \beta E[\max_{i_{t+1}} \begin{Bmatrix} \bar{V}^0(x_{t+1},\theta) + \varepsilon_{0t+1}, \\ \bar{V}^1(x_{t+1},\theta) + \varepsilon_{1t+1} \end{Bmatrix} |x_t,0,\theta], \\ \bar{V}^1(x_t,\theta) = u(x_t,1,\theta) + \beta E[\max_{i_{t+1}} \begin{Bmatrix} \bar{V}^0(x_{t+1},\theta) + \varepsilon_{0t+1}, \\ \bar{V}^1(x_{t+1},\theta) + \varepsilon_{1t+1} \end{Bmatrix} |x_t,1,\theta] \end{cases}$$

$$\Longrightarrow$$

$$\begin{cases} \bar{V}^0(x_t,\theta) = u(x_t,0,\theta) + \beta E[\mu + \ln\{e^{\bar{V}^0(x_{t+1},\theta)} + e^{\bar{V}^1(x_{t+1},\theta)}\}|x_t,0,\theta], \\ \bar{V}^1(x_t,\theta) = u(x_t,1,\theta) + \beta E[\mu + \ln\{e^{\bar{V}^0(x_{t+1},\theta)} + e^{\bar{V}^1(x_{t+1},\theta)}\}|x_t,1,\theta] \end{cases}$$

where $\mu = 0.5772$ is called the Euler constant.

- in the probability calculation

$$\Pr(i_t = 1|x_t,\theta) = \frac{\exp(\bar{V}^1(x_t,\theta))}{\exp(\bar{V}^0(x_t,\theta)) + \exp(\bar{V}^1(x_t,\theta))}$$

- CI and logit assumptions simplify things a lot, but they are strong assumptions

3.2.4 Integrated Value Functions

- Recall our Bellman equation

$$V(x_t, \varepsilon_t) = \max_{i \in D(x_t)} \{u(x_t, i, \theta) + \varepsilon_t(i) + \beta E[V(x_{t+1}, \varepsilon_{t+1})|x_t, \varepsilon_t, i]\}$$

- Let $\mathcal{P}_{x_t} \equiv \mathsf{Pr}(i_t = 1|x_t, \theta)$.

- Under CI, we can use the integrated value function formulation; i.e. the value at the beginning of the time period before ε_t realizes

$$\begin{aligned} V(x_t) = \mathcal{P}_{x_t} & \left[u(x_t, 1, \theta) + \widetilde{\varepsilon}_1^{\mathcal{P}_{x_t}} + \beta E[V(x_{t+1})|0]\right] \\ & + (1 - \mathcal{P}_{x_t})[u(x_t, 0, \theta) + \widetilde{\varepsilon}_0^{\mathcal{P}_{x_t}} + \beta E[V(x_{t+1})|x_t]] \end{aligned}$$

where $\widetilde{\varepsilon}_1^{\mathcal{P}_{x_t}} = E[\varepsilon_{1t}|\text{replace}]$ and $\widetilde{\varepsilon}_0^{\mathcal{P}_{x_t}} = E[\varepsilon_{0t}|\text{not replace}]$.

- We know that

$$\widetilde{\varepsilon}_1^{\mathcal{P}x_t} = \mu - \ln\left(\mathcal{P}_{x_t}\right)$$
$$\widetilde{\varepsilon}_0^{\mathcal{P}x_t} = \mu - \ln\left(1 - \mathcal{P}_{x_t}\right).$$

- In matrix notation,

$$\mathbf{V} = \Pi + \beta G^{\mathcal{P}}\mathbf{V}$$

where the j-th element of Π is

$$\Pi_j = \mathcal{P}_{x_t}(u(x_t, 1, \theta) + \widetilde{\varepsilon}_1^{\mathcal{P}x_t}) + (1 - \mathcal{P}_{x_t})(u(x_t, 0, \theta) + \widetilde{\varepsilon}_0^{\mathcal{P}x_t})$$

with $x_t = x_t\left(j\right)$, and (i, j) element of $G^{\mathcal{P}}$ is

$$G_{ij}^{\mathcal{P}} = \mathcal{P}_x p(x'|0, 1, \theta) + (1 - \mathcal{P}_x)p(x'|x, 0, \theta)$$

with $x = x_t(i)$ and $x' = x_t(j)$.

- We get

$$\mathbf{V} = (I - \beta G^{\mathcal{P}})^{-1}\Pi.$$

- As before we can define

$$\begin{cases} \bar{V}^0(x_t,\theta) = u(x_t,0,\theta) + \beta E[V(x_{t+1},\theta)|x_t, i_t = 0, \theta], \\ \bar{V}^1(x_t,\theta) = u(x_t,1,\theta) + \beta E[V(x_{t+1},\theta)|x_t, i_t = 1, \theta] \end{cases}$$

- Then,

$$\Pr(i_t = 1|x_t,\theta) = \Pr\left(\bar{V}^1(x_t,\theta) + \varepsilon_{1t} \geq \bar{V}^0(x_t,\theta) + \varepsilon_{0t}\right) \tag{3}$$

- The policy function iteration:

1. Set $\mathcal{P}^0$ and $\epsilon > 0$.
2. Calculate
$$\mathbf{V}^l = (I - \beta G^{\mathcal{P}^{l-1}})^{-1}\Pi$$
3. Update $\mathcal{P}^l$ using (3).
4. If $||\mathbf{V}^{l+1} - \mathbf{V}^l|| < \epsilon$, stop, otherwise go back to step 2.

3.3 Application: Timmins (2002)

- Water authorities set price
- Extraction today raises cost of extraction tomorrow, so it is not optimal to set current price equal to current marginal cost
- Evidence that price is actually below static marginal cost
- Thus authority is not behaving as a planner, so what is it maximizing?

- Questions
  - how does the authority in fact choose prices
  - what are the implications of this behavior for welfare
- Assume behavioral model up to a parameter vector
- Estimate behavioral model from pricing behavior (find parameter vector that rationalizes the water authorities pricing behavior)

- Primitives of the model:

  - demand function

  - cost of extraction function

  - law of motion for aquifer height

  - Regulator preferences

- Note observations are different municipalities over time (we ignore indices)

- Demand:

$$D = \mathsf{exp}(\delta_0 - \delta_1 P + \delta_2 Inc - \delta_3 R + \delta_4 S + \varepsilon^d)$$

(ε^d not known by authorities when setting P, Rainfall is treated as stochastic process, mean is estimated)

- Extraction costs:

$$C = h^{\alpha 1} D^{\alpha_2} \mathsf{exp}(\alpha_0 + \varepsilon^c + \xi)$$

(ξ either iid or follows AR1)

- Aquifer height evolves as:

$$h_{t+1} = \gamma_0 h_t + \gamma_0 - \gamma_2 D_t + \gamma_3 R_t + \varepsilon^h$$

- Note that $\varepsilon^d, \varepsilon^c, \varepsilon^h$ are correlated: Implications for estimation?

- Consumer surplus:

$$CS = \int_P^\infty D(\cdot, p) dp$$

and tax revenue

$$TR = P * D(\cdot, P) - C(\cdot)$$

- Behavioral assumptions

  - Authority measures benefits as a weighted average of consumer surplus and revenues:

  $$\Pi = v * CS + (1 - v) * TR$$

  (if $v = .5$, authority maximizes total surplus)

  - Pricing policy chosen maximizes:

  $$E \sum_{t=t_0}^{\infty} \beta^{t-t_0} \Pi_t \qquad ((*))$$

- Authority's problem is a dynamic programming problem. Solution to (*) corresponds to the solution to the functional equation

$$V(h, x, \xi|\theta) = \sup_{p \geq 0}[E\Pi_t + \\ \beta \int V(h', x', \xi')dP(h'|\cdot)dP(x'|x)]$$

- Obtain pricing policy function:

$$P_t = P(h_t, x_t, \xi_t|\theta)$$

- Estimation:
  - Two stages
  - Stage 1: estimate demand, cost of pumping water, law of motion for aquifer height
  - Stage 2: ‘Nested Fixed Point Algorithm’:

  (i) For given parameters θ solve the sequence problem (*)

  (ii) Repeat (i) until θ makes model’s implications as close as possible to the data

- (This is computationally rather involved)

- Results
  - $v = .73$ (standard error is .005)
  - can reject null hypothesis of net surplus maximization ($v = .5$)
  - At the margin:
$$\frac{\partial CS}{\partial TR} = -.453$$
    (what does this tell us about behavior of authority)
- Counterfactual experiments:
  - Dynamic efficient pricing would yield $110 gain for households (why?)
  - Suppose low flow toilets are installed, what would be the effect?