Introduction to Structural Econometrics

Lesson 2: Dynamic Programming and Recursive Models

Guillaume Wilemme

Aix-Marseille School of Economics - PhD course

Sept 2017

Table of contents

Plan

1. General questions
2. Some theory about dynamic programming
   - The Bellman equation
   - Continuous choice: an example
   - Discrete choice: an example
3. Application: Rust (1987)

Schedule

- October 3rd, room 23, 10-12am
  What is Structural Econometrics?
- October 10th, room 23, 10-12am
  Dynamic Programming and Recursive Models
- October 24th, room 23, 10-12am
  Monte-Carlo Methods
- November 7th, room 23, 10-12am
  Simulation-based Estimation
- November 14th, room 15, 10-12am
  State-Space Representation and Latent Variables
- November 21st, room 16, 10-12am
  Numerical tools

Evaluation

- Write only a R file
- It must contain
  1. one or several functions `model(..)` that generate a dataset given a list of parameters `theta`
  2. a graphical part, to illustrate the issue
  3. a bad way to estimate the parameters if possible (like a regression with an endogenous variable)
  4. a better (good?) way to estimate the parameters
- all the functions must be commented
- about the subject: find something related to your research interests, send me an email if you need help
- Deadline before or after Christmas vacations

Plan

Dynamic programming

1. We consider recursive theoretical models that are applied in structural econometrics.
2. For the theory, a good textbook is Ljunqvist and Sargent's Recursive Macroeconomic Theory.
3. By recursive models, I mean models that are solved by the agents through backward induction.
   - but also by the econometrician!

A general framework

- Time is discrete and the agent has VNM preferences. She lives from $t = 1$ to $t = T$ (possibly $T = +\infty$).
- Expected utility at time t writes

$$U(s_t) = \sum_{\tau=t}^{T} \beta^{\tau-t} \mathbb{E}^*_{s_\tau} \left[u^*(s_\tau) \middle| s_t \right]. \quad (1)$$

- s_t is the vector of *state variables* (or predetermined or backward-looking). Note that s_0 is given.
- The problem becomes interesting when there are choices involved. The star superscript denotes the operators under optimal decisions (e.g. indirect utility)

Optimal decision

- The agent can influence the state by making state-contingent decisions $d_\tau(s_\tau)$ at time τ. Denote $\mathbf{d}_t = \{d_\tau\}_{\tau=t,..,T}$. The individual program is the following:

$$U(s_t) = \max_{\mathbf{d}_t \in \mathbf{C}_t} \sum_{\tau=t}^{T} \beta^{\tau-t} \mathbb{E}_{s_\tau} \left[u(s_\tau, d_\tau(s_\tau)) \Big| s_t, \mathbf{d}_t \right]. \quad (2)$$

- $\mathbf{d}_t$ is a set of *control variables* (or jump or forward-looking).
  - affect utilities
  - affects the transition probabilities between states
- $\mathbb{E}_{s_\tau}$ is an expectation operator over probabilities $p(s_\tau | s_{t-1}, \mathbf{d}_t)$.

The Bellman equation

- A convenient writing of the program is the Bellman equation formulation:

$$U(s_t) = \max_{d_t \in \mathcal{C}_t} \left\{ u(s_t, d_t) + \beta \mathbb{E}_{s_{t+1}} \left[U(s_{t+1}) | s_t, d_t \right] \right\} \qquad \text{for } t < T, \quad (3)$$

- with a termination condition if $T < \infty$, $U(s_T) = \max_{d_T \in \mathcal{C}_T} u(s_T, d_T)$
- The way to characterize the program depends on the set $\mathcal{C}_t$, continuous or discrete.
- Contemplate the recursive structure

Example with savings

- a worker earn a stcoachastic income y_t and choose to consume and save:

$$U_t(a_t, y_t) = \max_{\{c_t | a_{t+1}=(1+r)(a_t+y_t-c_t)\}} \{u(c_t) + \beta \mathbb{E}_{y_{t+1}} U_{t+1}(a_{t+1}, y_{t+1})\} \tag{4}$$

- with the termination condition, $U(s_T) = \max_{d_T \in \mathcal{C}_T} u(s_T, d_T)$

Solution

- Without limited liability constraints, the problem is solved by substituting c_t and writing two conditions:
  - the first-order condition in a_{t+1}
  - the envelop theorem in a_{t+1}:

$$-\frac{u'(c_t)}{1+r} + \beta \mathbb{E}_{y_{t+1}} \frac{\partial U_{t+1}}{\partial a_{t+1}}(a_{t+1}, y_{t+1}) = 0, \tag{5}$$

$$\frac{\partial U_t}{\partial a_t}(a_t, y_t) = u'(c_t). \tag{6}$$

- We search $c_t^*(a_t, y_t)$

Solution (II)

- We obtain the Euler equation,

$$\frac{u'(c_t^*(a_t, y_t))}{1+r} = \beta \mathbb{E}_{y_{t+1}} u' \left[c_{t+1}^* \left((1+r)(a_t + y_t - c_t^*(a_t, y_t)), y_{t+1} \right) \right]$$

- We can recover c_t from c_{t+1} starting from the termination condition $c_T^*(a_T, y_T) = a_T + y_T$.
- For $c_{T-1}^*(a, y)$,

$$u'(c_{T-1}^*) = (1+r)\beta \mathbb{E}_{y_T} u' \left((1+r)(a + y - c_{T-1}^*) + y_T \right)$$

  - we have to solve this equation for each pair (a, y)
  - In addition, there is an integral to compute if y_t is continuously distributed.
  - This is just for computing c_{T-1}^*.
- The take-away message is that it can be hard to solve a dynamic continuous choice model, even in this simple case.

Example with labor participation

- A risk-neutral individualchoose to enter the labor market.
- Her utility from working is w_t (deterministic), while her utility from remaining out of the labor force is h.

$$U_t(0) = h + \max\left\{-c + \beta\left[\lambda U_{t+1}(1) + (1-\lambda)U_{t+1}(0)\right], \beta U_{t+1}(0)\right\} \tag{7}$$

$$U_t(1) = w_t + \beta \max\left\{U_{t+1}(1), U_{t+1}(0)\right\}. \tag{8}$$

- The termination conditions are $U_T(0) = h$ and $U_T(1) = w_T$.

Solution

- The agent chooses a participation decisions $d_t^*(0)$ and $d_t^*(1)$ for $t < T$.
- From the termination condition, $d_{T-1}(1) = \mathbf{1_{h<w_T}}$ and $d_{T-1}(0) = \mathbf{1_{c<\beta\lambda(w_T-h)}}$.
- For any $t < T$, the worker takes the decision by comparing the values,

$$d_t(0) = \mathbf{1_{c<\beta\lambda(U_{t+1}(1)-U_{t+1}(0))}}, \tag{9}$$
$$d_t(1) = \mathbf{1_{U_{t+1}(1)>U_{t+1}(0)}}. \tag{10}$$

- d_t is simple to obtain given U_{t+1}

Numerical solution

- To compute $d_0(0)$ for instance, the computer has to make 2^{T-1} comparisons.
- If you consider 40 periods as different years on the labor market, the number is higher than 10^{12}.
- On my computer with the Julia software (faster than R presumably), it takes me 5min!
- This is just for a binary choice and two states.
- Remember that we are just evaluating the function for a given set of parameters. In order to estimate, we need to evaluate it several times.

Plan

The subject

- Harold Zuchner is an engine repairer in a bus company,
- 10 years of observations for different bus
- Contribution
  1. methodological
  2. replacement investment in "bottom-up" framework, estimate a demand function for engine replacement

What do we observe

- Each bus is an individual.
- For each bus, we observe the monthly mileage x_t and the engine replacement decision i_t. An observation consists of a sequence of (x_t, i_t). Here is an example

$$(0, 0, .., 0, 1, 0, 0, ..) \text{ for the } \{i_t\} \quad (11)$$

$$(106, 252, .., 2553, 0, 130, 276, ..) \text{ for the } \{x_t\} \quad (12)$$

A model

- x_t is a state variable. The utility function regarding maintenance is defined by

$$u(x_t, i_t, \theta_1) = \begin{cases} -c(x_t, \theta_1) & \text{if } i_t = 0 \\ -RC - c(0, \theta_1) & \text{if } i_t = 1 \end{cases} \tag{13}$$

- x_t is typically auto-correlated, we suppose the pdf to be

$$p(x_{t+1} | x_t, i_t, \theta_2) \tag{14}$$

A model (II)

- The Bellman equation:

$$V_\theta(x_t) = \max \left\{ \begin{array}{l} u(x_t, 0, \theta_1) + \beta \int V_\theta(x_{t+1}) p(x_{t+1}|x_t, i_t = 0, \theta_2) dx_{t+1} \\ u(x_t, 1, \theta_1) + \beta \int V_\theta(x_{t+1}) p(x_{t+1}|x_t, i_t = 1, \theta_2) dx_{t+1} \end{array} \right\} \tag{15}$$

- Define $EV_\theta(x_t, i_t) = \int V_\theta(x_{t+1}) p(x_{t+1}|x_t, i_t, \theta_2) dx_{t+1}$ as the expected value. We have

$$EV_\theta(x_t, i_t) = \int \max_{i'} \left\{ u(x_{t+1}, i', \theta_1) + \beta EV_\theta(x_{t+1}, i') \right\} p(x_{t+1}|x_t, i_t, \theta_2) dx_{t+1} \tag{16}$$

- To solve numerically the problem, a strategy is to find EV solution to a fixed point problem
- Contemplate the difficulty

A deterministic model

- Denote $i_t = f(x_t, \theta)$ the optimal maintenance decision.
  - Under some assumptions, i_t is a 0-1 policy depending on a threshold mileage $\underline{x}$.
  - We cannot estimate such a deterministic model
- $i_t = f(x_t, \theta) + \epsilon_t$, is not a good option
  - loose internal consistency, what is the interpretation?
  - At best, it could be a measurement error, but it does not make sense regarding the way data were collected.
- Consider that the stochastic terms ϵ_t are observed by Harold Zuchner but not by us. Random utility framework.

A stochastic model

- V writes as

$$V_\theta(x_t,\epsilon_t) = \max \left\{ \begin{array}{l} u(x_t,0,\theta_1)+\epsilon_t(0)+\beta\int V_\theta(x_{t+1})p(x_{t+1},\epsilon_{t+1}|x_t,\epsilon_t,i_t=0,\theta_2)dx_{t+1}d\epsilon_t \\ u(x_t,1,\theta_1)+\epsilon_t(1)+\beta\int V_\theta(x_{t+1})p(x_{t+1},\epsilon_{t+1}|x_t,\epsilon_t,i_t=1,\theta_2)dx_{t+1}d\epsilon_t \end{array} \right. \quad (17)$$

- Define $EV_\theta(x_t,\epsilon_t,i_t)=\int V_\theta(x_{t+1},\epsilon_{t+1})p(x_{t+1}|x_t,\epsilon_t,i_t,\theta_2)dx_{t+1}$ as the expected value. We have

$$EV_\theta(x_t,\epsilon_t,i_t) = \int \max_{i'} \left\{ u(x_{t+1},i',\theta_1)+\epsilon_t(i)+\beta EV_\theta(x_{t+1},\epsilon_{t+1},i') \right\} p(x_{t+1},\epsilon_{t+1}|x_t,\epsilon_t,i_t, \quad (18)$$

Assumption 1

Suppose $p(x_{t+1}, \epsilon_{t+1} | x_t, \epsilon_t, i_t, \theta_2) = q(\epsilon_t) p(x_{t+1} | x_t, , i_t, \theta_2)$.

$$EV_\theta(x_t, i_t) = \int \max_{i'} \{u(x_{t+1}, i', \theta_1) + \epsilon_t(i) + \beta EV_\theta(x_{t+1}, i')\} \\ p(x_{t+1} | x_t, i_t, \theta_2) dx_{t+1} q(\epsilon_{t+1}) d\epsilon_{t+1}. \quad (19)$$

Assumption 2

ϵ_t follows a centered extreme-value distribution, then

$$EV_\theta(x_t, i_t) = \int \log \left\{ \sum_{i'} \exp\left(u(x_{t+1}, i', \theta_1) + \beta EV_\theta(x_{t+1}, i')\right) \right\} p(x_{t+1}|x_t, i_t, \theta_2) dx_{t+1}. \tag{20}$$

and

$$P(i|x_t, \theta) = \frac{\exp\left(u(x_t, i, \theta_1) + \beta EV_\theta(x_t, i)\right)}{\sum_{i'} \exp\left(u(x_t, i', \theta_1) + \beta EV_\theta(x_t, i')\right)}. \tag{21}$$

Likelihood

- The stochastic model can be estimated by maximum likelihood. The likelihood of a sequence of (x_t, i_t) is

$$p(x_1|i=1,\theta_2)P(i_1|x_1,\theta)\prod_{t>1} p(x_t|x_{t-1},i_{t-1},\theta_2)P(i_t|x_t,\theta) \quad (22)$$

- Note we can split the likelihood function in two terms:
  - one with the p, $p(x_1|i=1,\theta_2)\prod_{t>1} p(x_t|x_{t-1},i_{t-1},\theta_2)$ We will use this property to estimate the model.
  - the other with the P, $\prod_{t>0} P(i_t|x_t,\theta)$.