Lecture 7: More on Dynamic Programming

1 Value function iteration

Value function iteration proceeds by constructing a sequence of value functions and associated policy functions. The sequence is created by iterating on the Bellman equation, starting from $V_0 = 0$,

$$V_{j+1}(s) = \max_{c}\{r(s,c) + \beta V_j(s')\}, \tag{1}$$

subject to $s' = g(s,c)$, s given, and continuing until V_j has converged.

For our RBC example, it can be implemented using the following algorithm:

Value function iterations

1. Discretize the state space/set the grid, $k = \{k_1, k_2, \ldots, k_i, \ldots, k_{i_k}\}$ around the steady state, (or min/max values).

2. Choose $v_0(k)$, where $v_0(\cdot)$ being a vector of length i_k.

3. For each value of k at the grid, compute the value function

$$v_{n+1}(k) = T(v_n(k)) = \max_{0<k'<Ak^\alpha+(1-\delta)k} \left[u(Ak^\alpha + (1-\delta)k - k') + \beta v_n(k')\right]$$

4. Compute the distance d between v_{n+1} and v_n, e.g.

$$d(v_{n+1}, v_n) = \sum_{i=1}^{i_k} |v_{n+1}(k_i) - v_n(k_i)| \qquad \text{or} \tag{2}$$

$$d(v_{n+1}, v_n) = \max_{i=1}^{i_k} |v_{n+1}(k_i) - v_n(k_i)| \tag{3}$$

5. Repeat until $d(v_{n+1}, v_n) \leq \epsilon$.

Once the value function iteration piece of the program is completed, the value function can be used to find the policy function, $k' = h(k)$.

- Approximating the value function and the policy rules by a finite state space requires a large number of points on this space (i_k has to be big).
- This can be very time consuming in terms of numerical calculations.
- One can reduce the number of points on the grid, while keeping a satisfactory accuracy by using interpolations on this grid.
- However, value function iteration method faces "curse of dimensionality"

⇒ if each of N state variable is discretized into n_s grid points, the value function has to be evaluated into N^{n_s} points — this demands an increasing computer memory and slows down the computation.

- Value function iteration converges at rate β.

2 Policy function iteration

The policy function iteration, so called Howards improvement algorithm, is a faster method.[1]

It consists of the following steps:

1. Guess a feasible policy, $c = h_j(s)$, $j = 0$, with s on a grid.
2. *Policy evaluation step*: Compute the value of using this rule forever as

$$V_{h_j}(s) = \sum_{t=0}^{\infty} \beta^t r(s_t, h_j(s_t)), \qquad \forall s \ ,$$

where $s_{t+1} = g(s_t, h_j(s_t))$.

$$\big(\text{or} \qquad V_{h_j}(s) = r(s, h_j(s)) + \beta V_{h_j}(g(s, h_j(s))), \qquad \forall s.\big)$$

3. *Policy improvement step*: Generate a new policy $c = h_{j+1}(s)$ that solves the two–

[1]See Judd (1998) and Ljunqvist and Sargent (2002) for more details.

period problem:

$$h_{j+1}(s) = \arg\max_{c}\{r(s,c) + \beta V_{h_j}(g(s,c))\}, \qquad \forall s\ .$$

4. Iterate over j to convergence on step 2 (i.e. $d(h_{j+1}(s), h_j(s)) \leq \epsilon$), or step 3 (i.e. $d(V_{h_{j+1}}(s), V_{h_j}(s)) \leq \epsilon$).

- The convergence rate is much faster than the value function iteration method.
- *Policy evaluation step* can be in some cases very time consuming, especially when the state space is large.

3 Stochastic case

The non-stochastic problem is a natural starting point but it is necessary to consider stochastic case. After all, even our simple RBC/stochastic growth model has dynamic stochastic element in it.

Let ε be the current value of a vector of random variables, so called "shocks", and let $\varepsilon \in \Psi$.[2]

The Bellman equation can be written:

$$V(s,\varepsilon) = \max_{s' \in \Gamma(s,\varepsilon)} r(s,s',\varepsilon) + \beta E_{\varepsilon'|\varepsilon} V(s',\varepsilon') \tag{4}$$

for all (s,ε), where we used both transition equation and police function

$$s' = g(s,\varepsilon) \tag{5}$$

$$c = h(s,\varepsilon) \tag{6}$$

Note that in writing expectations above ($E_{\varepsilon'|\varepsilon}$) we assumed that

- the stochastic process is exogenous as the distribution of ε' depends on ε but not on the current state and control,

[2] We will assume that Ψ is a finite set.

- ε follows a first-order Markov process, that is the distribution of ε' depends on only the realized value of ε,[3]

- the distribution of ε' conditional on ε is stationary, i.e. $\varepsilon'|\varepsilon$ is time invariant.

The conditional probability of $\varepsilon'|\varepsilon$ is given by a transition matrix, Π, with

$$\pi_{ij} \equiv Prob(\varepsilon' = \varepsilon_j | \varepsilon = \varepsilon_i)$$

For Π to be a probability transition matrix it must be the case that

1. $\pi_{ij} \in [0, 1]$, and

2. $\sum_j \pi_{ij} = 1$ for each i, that is rows sums up to 1.

Then

Theorem 1. *If $r(s, s', \varepsilon)$ is real-valued, continuous, concave and bounded, $0 < \beta < 1$ and the constraint set compact and convex, then:*

1. *there exists the unique value function $V(s, \varepsilon)$ that solves* (4)

2. *there exists a stationary policy function, $\phi(s, \varepsilon)$.*

Proof: As in case of deterministic case, use Blackwell's Theorem.

The first-order condition for (4) is

$$r_{s'}(s, s', \varepsilon) + \beta E_{\varepsilon'|\varepsilon} V_{s'}(s', \varepsilon') = 0. \tag{7}$$

Use (4) to determine $V_{s'}(s', \varepsilon')$

$$V_s(s, \varepsilon) = r_s(s, s', \varepsilon)$$

and an Euler equation:

$$r_{s'}(s, s', \varepsilon) + \beta E_{\varepsilon'|\varepsilon} r_{s'}(s', s'', \varepsilon') = 0. \tag{8}$$

[3] If values of shocks from previous periods were relevant for the distribution of ε', then they could simply be added to the state vector.

Euler equation has the usual interpretation: the expected sum of the effects of a marginal variation in the control in the current period (s) must be zero.

(Note that this is a version of Benveniste and Scheinknan (1979) equation.
For deterministic case with with h being policy function, g transition equation, and V defined as

$$V(s) = r(s, h(s)) + \beta V(g(x, h(x)))$$

the Benveniste and Scheinknan formula is given by

$$V'(s) = r_s(s, h(s)) + \beta g_s(s, h(s))V'(g(x, h(x)))$$

which, if transition equation is defined such that $g_s = 0$, reduces to

$$V'(s) = r_s(s, h(s)).)$$

If a policy is optimal, there is not other value of control c that, *in expectation*, make the agent better off.
After the realization of ε', there may have been better decisions for the agent, that is

$$r_{s'}(s, s', \varepsilon) + r_{s'}(s', s'', \varepsilon'') \neq 0, \qquad \forall \varepsilon' \tag{9}$$

but these ex post errors were not predicable given the information available to the agent.

3.1 RBC model

Assume that shocks are in the production function and, as before, assume that labor is inelastically supplied at one unit per household (we consider more general endogenous labor supply formulation later).

Bellman's equation for the infinite horizon RBC/stochastic growth model equals

$$V(A, k) = \max_{k'} u(Af(k) + (1 - \delta)k - k') + \beta E_{A'|A} V(A', k'), \qquad \forall A, k \tag{10}$$

where we used the transition equation

$$k' = Af(k) + (1 - \delta)k - c.$$

We assume that A is a bounded, discrete random variable that follows a first-order Markov process, and that the transition matrix is given by a $m \times m$ matrix Π:

$$\Pi = \begin{bmatrix} \pi_{11} & \pi_{12} & \dots & \pi_{1m} \\ \pi_{21} & \pi_{22} & \dots & \pi_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ \pi_{m1} & \pi_{m2} & \dots & \pi_{mm} \end{bmatrix}$$

For the existence of a solution we need the problem to be bounded. Define $\bar{k}$ as the solution to

$$k = A^{+}f(k) + (1 - \delta)k \tag{11}$$

where A^{+} is the largest productivity shock.

Since $c \geq 0$ transition equation implies that $\bar{k}$ is the largest amount of capital that this economy could accumulate, and that largest level of consumption is also $\bar{k}$. Thus utility is bounded by $u(\bar{k})$, and, given that the process for shocks is bounded, there exists a unique value function $V(A, k)$ that solves (10), and there is a policy function given by $k' = \phi(A, k)$.

3.2 Solution Methods

Recall that

$$V(A, k) = \max_{k'} u(Af(k) + (1 - \delta)k - k') + \beta E_{A'|A} V(A', k'), \qquad \forall A, k$$

The first-order condition (for the planner) is

$$u'(Af(k) + (1 - \delta)k - k') = \beta E_{A'|A} V_{k'}(A', k'), \qquad \forall A, k \tag{12}$$

and using

$$V_k(A,k) = u'(c)[Af(k) + (1-\delta)]$$

we get the Euler equation

$$u'(c) = \beta E_{A'|A} u'(c')[A'f(k') + (1-\delta)] \tag{13}$$

where

$$c = Af(k) + (1-\delta)k - k' \tag{14}$$

Given the law of motion of A (specified below) we have a system of first order stochastic difference equations in (c, k, A).[4]

3.2.1 Specific example: guess and verify

Assume that $U(c) = \ln c$, $f(k) = Ak^\alpha$, and that the rate of depreciation of capital is 100% ($\delta = 1$), and that A follows an AR(1) process

$$\ln A' = \rho \ln A + \varepsilon, \qquad 0 < \rho < 1.$$

Then the Euler equation is

$$\frac{1}{c} = \beta E_{A'|A} \left(\frac{A'\alpha k'^{(\alpha-1)}}{c'} \right). \tag{15}$$

We guess and verify policy function:

$$k' = \phi(A,k) = \lambda A k^\alpha$$

where λ is an unknown constant. Given the resource constraint, this implies

$$c = (1-\lambda)Ak^\alpha$$

To verify this guess and determine λ, put the proposed policy function into Euloer

[4]Earlier we log-linearized the FOC and the resource constraints around the steady state, (c^*, k^*) to find an approximate solution. The dynamic programming allows us to solve the problem directly.

equation, (15):

$$\frac{1}{(1-\lambda)Ak^{\alpha}} = \beta E_{A'|A}\left(\frac{A'\alpha k'^{(\alpha-1)}}{(1-\lambda)A'k'^{\alpha}}\right).$$

Solving for k' yields the policy function:

$$\phi(A,k) = k' = \beta\alpha A k^{\alpha} \tag{16}$$

and $\lambda = \alpha\beta$. This implies that consumption is proportional to income:

$$c = (1-\beta\alpha)Ak^{\alpha}. \tag{17}$$

One can show that the value function is given by:

$$V(A,k) = G + B\ln k + D\ln A \qquad \forall A, k$$

where G, B and D are unknown constants which we can solve for.

3.3 Decentralized problem

So far, we were considering a planner's problem. By the *Second Welfare Theorem* we know that *any efficient allocation can be decentralized as a competitive allocation.* We will introduce the concept of *recursive equilibrium.*

The firm's problem.

- Assume there is a single firm which acts competitively and employs all the capital and labor in the economy, denoted by upper case letters.
- At the beginning of the period, the firm rents capital from the household at a price of r per unit and hires labor at a wage of w per hour (static problem)
- Wage and rental rates are expressed in terms of current period output, and

  the firm takes prices as given.

The representative firm chooses (K, N) to maximize profits. FOCs

$$Af_N(K, N) = w$$

and

$$Af_K(K, N) + (1 - \delta) = r.$$

Households' problem

Households have value a function

$$V(A, k, K) = \max_{k'} u(r(K)k + w(K) + Profit - k') + \beta E_{A'|A} V(A', k', K') \tag{18}$$

- k is the household's own stock of capital,
- K is the (aggregate) per capita capital stock in the economy.
- K is a state variable in households' problem since factor prices depend on the aggregate state variable through the factor demand equations: $r(K)$ and $w(K)$ depend on K.

Let $K' = H(A, K)$ denote evolution of the aggregate capital stock.

- Household takes the evolution of the aggregate state variable as given (as hh is competitive),

$\Rightarrow$ the household takes current and future factor prices as given.

The first-order condition for the household's capital decision:

$$u'(c) = \beta E_{A'|A} V_k(A', k', K') \tag{19}$$

and household uses the law of motion for K.

As $V_k = r(K)u'(c)$ we get Euler equation:

$$u'(c) = \beta E_{A'|A} r' u'(c') \tag{20}$$

Definition 1. *A **recursive equilibrium** is a factor price function: $r(K)$ and $w(K)$; individual policy functions: $h(A, k, K)$ from (18); and a law of motion for aggregate capital K: $H(A, K)$ such that*

1. *households and firms optimize,*
2. *markets clear,*
3. $H(A, k) = h(A, k, k)$.

Using first order conditions for the factor demand of the operating firm, it is easy to see that the solution to the planner's problem is a recursive equilibrium.

4 A Stochastic Growth Model with Endogenous Labor Supply

4.1 Planner's Dynamic Programming Problem

The modified planner's problem is given by

$$V(A, k) = max_{k',n} u(Af(k, n) + (1 - \delta)k - k', 1 - n) + \beta E_{A'|A} V(A', k') \qquad \forall A, k \quad (21)$$

In addition to the dynamic choice between consumption and investment, there is now also a "static" (for given k') choice of n.

Labor choice. Define $\sigma(A, k, k')$ as

$$\sigma(A, k, k') = max_n u(Af(k, n) + (1 - \delta)k - k', 1 - n), \quad (22)$$

and let $n = \hat{\phi}(A, k, k')$ be the solution. The first-order condition is given by:

$$u_c(c, 1 - n)Af_n(k, n) = u_l(c, 1 - n), \quad (23)$$

that is the marginal gain from increasing employment and consuming more equals the marginal cost of working more.

Given the current productivity shock, A, the current capital stock, k, and a level of capital next period, k', $\hat{\phi}(A,k,k')$ is a return function given the current state (A,k) and control (k') that characterizes the employment decision.

The rewritten **functional Bellman equation** becomes

$$V(A,k) = max_{k'}\sigma(A,k,k') + \beta E_{A'|A}V(A',k'), \qquad \forall_{A,k}. \tag{24}$$

- The same structure as the stochastic growth model with a fixed labor supply.
- The return function, $\sigma(A,k,k')$, is not a primitive object, but inherits its properties from the more primitive $u(c,1-n)$ and $f(k,n)$ functions.

$\Rightarrow$ There exists a unique solution, V, and a stationary policy function, $k' = h(A,k)$.

The first-order condition for the choice of the future capital stock is given by,

$$\sigma_{k'}(A,k,k') + \beta E_{A'|A}V_{k'}(A',k') = 0$$

and solving for $E_{A'|A}V_{A',k'}$ yields an Euler equation

$$-\sigma_{k'}(A,k,k') = \beta E_{A'|A}\sigma_{k'}(A',k',k'').$$

Using solution to labor problem in (22), we get

$$u_c(c,1-n) = \beta E_{A'|A}\left(u_c(c',1-n')[A'f_k(k',n')+1-\delta]\right], \tag{25}$$

where $c = Af(k,n) + (1-\delta)k - k'$.