

Chris Edmond (cpedmond@unimelb.edu.au)

Dynamic programming and the growth model

Dynamic programming and closely related recursive methods provide an important methodology for solving a wide variety of economic models. In this note, I use the optimal growth model to illustrate several important dynamic programming concepts.

Until now, we have been working with something like the following:

Sequence Problem

$$\max_{\{c_t, k_{t+1}\}_{t=0}^{\infty}} \sum_{t=0}^{\infty} \beta^t U(c_t)$$

subject to, for each $t = 0, 1, \dots$

$$c_t + k_{t+1} \leq f(k_t)$$

$$c_t, k_{t+1} \geq 0$$

given k_0

(Here I write $f(k_t)$ for the total supply of goods available at the beginning of period t , i.e., this includes un-depreciated capital).

A key feature of this problem is that it involves choosing an **infinite sequence** of consumption or capital accumulation, one for each date t . It turns out that we can replace the problem of solving for an infinite sequence with the problem of solving for a single unknown function. This is only progress because the new problem has an attractive **recursive** formulation. Under a standard set of regularity conditions, we can formulate an equivalent problem:

Recursive Problem

$$V(k) = \max_{k'} \{U(c) + \beta V(k')\}$$

subject to

$$c + k' \leq f(k)$$

$$c, k' \geq 0$$

In this problem, the utility and production functions U and f as well as the discount factor β are known, but the function V — a **value function** — is unknown. To understand the economics of this formulation, suppose that the value function V was known. Then this would just be a two-period decision problem. With strictly increasing utility, we could substitute the resource constraint with equality into the objective function and solve

$$\max_{0 \leq k' \leq f(k)} \{U[f(k) - k'] + \beta V(k')\}$$

This determines tomorrow's capital stock as a function of today's capital stock. Specifically,

$$k' = g(k) \in \arg \max_{0 \leq x \leq f(k)} \{U[f(k) - x] + \beta V(x)\}$$

The function g is known as a **policy function**. Suppose further that V is concave. Then the objective function is the sum of concave functions and hence is also concave. So the optimal policy is characterized by the first order condition

$$U'(c) = \beta V'(x)$$

That is, at an optimum, we will be trading off the marginal cost $U'(c)$ of foregone consumption against a marginal benefit $\beta V'(x)$. If V were known, we could then solve for the policy function by finding, for each k , the appropriate $k' = g(k)$ such that

$$U'[f(k) - g(k)] = \beta V'(g(k))$$

With the policy function g solved for, we can then construct the entire sequence of capital stocks $k_{t+1} = g(k_t)$ starting with the given initial condition k_0 . We can also back out consumption, namely $c_t = f(k_t) - g(k_t)$. Today's capital stock is all that we need to know in order to choose consumption optimally. The jargon for this is that today's capital stock is the only **state variable**. Finally, notice that the policy function $k_{t+1} = g(k_t)$ induces a **one-dimensional** dynamic system. From our previous analysis of the optimal growth model, it's clear that this one-dimensional dynamic system corresponds to the stable arm of the saddle-path. The policy function implies that initial consumption is $c_0 = f(k_0) - g(k_0)$.

Before turning to methods for solving for the value function, we should first develop some

intuition for the recursive formulation. To do so, turn back to the sequence problem and call $V(k_0)$ the maximized level of utility associated with optimizing behavior given initial capital k_0 . That is

$$V(k_0) \equiv \max_{\{k_{t+1}\}_{t=0}^{\infty}} \left\{ \sum_{t=0}^{\infty} \beta^t U[f(k_t) - k_{t+1}] \right\}$$

The value function depends on the primitive utility and production functions U and f , the discount factor β and takes as an argument the initial level of the state variable, k_0 . Now let the optimal sequence of capital stocks be $\{k_{t+1}^*\}_{t=0}^{\infty}$. Then

$$V(k_0) = \sum_{t=0}^{\infty} \beta^t U[f(k_t^*) - k_{t+1}^*]$$

But there is nothing special about date zero. We could just have easily defined the value function to be

$$V(k_t) = \sum_{s=t}^{\infty} \beta^{t-s} U[f(k_s^*) - k_{s+1}^*]$$

To see the special role the value function plays, we can break the sum into two pieces

$$\begin{aligned} V(k_0) &= U[f(k_0) - k_1^*] + \sum_{t=1}^{\infty} \beta^t U[f(k_t^*) - k_{t+1}^*] \\ &= U[f(k_0) - k_1^*] + \beta \sum_{t=1}^{\infty} \beta^{t-1} U[f(k_t^*) - k_{t+1}^*] \\ &= U[f(k_0) - k_1^*] + \beta V(k_1^*) \end{aligned}$$

So

$$V(k_0) = \max_{0 \leq k_1 \leq f(k_0)} \{U[f(k_0) - k_1] + \beta V(k_1)\}$$

But the dates $t = 0$ and $t = 1$ are arbitrary, so it is very common to simply denote the capital stock today by k and the capital stock in the next period by k' . Then

$$V(k) = \max_{0 \leq k' \leq f(k)} \{U[f(k) - k'] + \beta V(k')\}$$

This is an example of a **Bellman equation**. It is one equation in an unknown function V and is an example of a **functional equation**. As we have already seen, if we knew the value function, solving the maximization problem for the policy function $k' = g(k)$ becomes an essentially trivial exercise.

The prime ($'$) notation emphasizes the recursive and stationary nature of the problem: $V(k)$ is the value of the capital stock today and $V(k')$ is the value of the capital stock tomorrow, so $\beta V(k')$ is the value of the capital stock tomorrow in terms of utility today. Given beginning-of-period capital k , we can optimize by just choosing a single capital stock k' to take into the next period — since we know that tomorrow, we will also **proceed optimally** and get utility value $V(k')$ for any capital that we do in fact carry over. Hence, at an optimum, we will be trading off the marginal cost $U'(c)$ against a marginal benefit $\beta V'(k')$. [There are two uses of a prime here, one to denote a derivative and another to denote a next-period-value. Hopefully this won't cause any confusion].

Later we will discuss methods for solving this functional equation problem. But there are a number of interesting characteristics of the optimal growth model that can be deduced without an explicit solution for the value function.

A. First order and envelope conditions

As already discussed, an optimal policy $g(k)$ is characterized by the first order condition

$$U'[f(k) - g(k)] = \beta V'(g(k))$$

But this is not much help, since V and hence V' are unknown. We can make progress using the **envelope theorem**. Specifically, the total derivative of the value function V with respect to a parameter k is equal to the partial derivative of the objective function with respect to the parameter evaluated at the optimal policy $g(k)$. [See the "aside" at the end of this note for more on the envelope theorem]. Using this result,

$$V'(k) = U'[f(k) - g(k)]f'(k)$$

(this is sometimes known as the **envelope condition**). But we actually need the derivative of the value function evaluated at the point $k' = g(k)$, namely

$$V'(g(k)) = U'[f(g(k)) - g(g(k))]f'(g(k))$$

We can now plug this into the first order condition above to get

$$U'[f(k) - g(k)] = \beta U'[f(g(k)) - g(g(k))]f'(g(k))$$

This is just the standard consumption Euler equation. To see this, recall that $k_{t+1} = g(k_t)$ and $c_t = f(k_t) - g(k_t)$ so we just have

$$U'(c_t) = \beta U'(c_{t+1}) f'(k_{t+1})$$

This is exactly the consumption Euler equation from the sequence problem (remember that I have been using $f(k_t)$ to denote the total supply of goods available at the beginning of period t , including any un-depreciated capital).

B. Steady state

Interestingly enough, we can solve for the steady state of the model without knowing either the value function or the policy function. A steady state is defined just as in the sequence problem, namely $\bar{k} = k' = k$ so that $\bar{k} = g(\bar{k})$. The steady state capital stock is a fixed point of the policy function. Steady state consumption is then backed out from the resource constraint, $\bar{c} = f(\bar{k}) - \bar{k}$. To find the steady state capital stock, just solve

$$1 = \beta f'(\bar{k})$$

which we can do without knowing either the value function or the policy function.

C. Value function iteration

Now recall the Bellman equation

$$V(k) = \max_{0 \leq k' \leq f(k)} \{U[f(k) - k'] + \beta V(k')\}$$

The right hand side of this equation defines an **operator** that maps a given function V into a new function. This operator is often denoted T and we write TV for the new function created by evaluating T at V . That is,

$$TV(k) \equiv \max_{0 \leq k' \leq f(k)} \{U[f(k) - k'] + \beta V(k')\}$$

The as-yet unknown value function V is in fact a **fixed point** of the operator equation

$$V = TV$$

Intuitively, you could think of the function V as a vector and the Bellman operator T as something like a matrix that maps vectors into vectors — this analogy is pretty loose, of course, because a matrix is a linear operator and the Bellman operator T is not.

We can often use something like the following procedure to solve for the value function:

Step 1. Guess some initial value function, say V_0 .

Step 2. Use T to construct a new value function $V_1 = TV_0$, that is,

$$V_1(k) = TV_0(k) \equiv \max_{0 \leq k' \leq f(k)} \{U[f(k) - k'] + \beta V_0(k')\}$$

Test to see if V_1 is sufficiently close to V_0 according to some metric. If so, we have a fixed point and stop. If not:

Step 3. Use T to construct a new value function $V_2 = TV_1 = TTV_0$. Test to see if we have a fixed point. If not, keep iterating on the value functions $V_{j+1} = TV_j = T^j V_0$ for $j \geq 1$. Under suitable regularity conditions, the iterates $T^j V_0$ converge to a unique V as $j \rightarrow \infty$. That limit is the fixed point, $V = TV$, and is independent of the initial guess V_0 .

To see more concretely how this works, suppose our initial guess is $V_0(k) = 0$ all k . Then

$$V_1(k) = TV_0(k) \equiv \max_{0 \leq k' \leq f(k)} \{U[f(k) - k'] + \beta 0\} = U[f(k)]$$

where the last equality follows because if $V_0(k) = 0$, the optimal policy is $k' = 0$ all k . But this means $V_1(k) = U[f(k)] \neq 0$, so we need to keep going. We construct a new iterate

$$V_2(k) = TV_1(k) \equiv \max_{0 \leq k' \leq f(k)} \{U[f(k) - k'] + \beta U[f(k')]\}$$

and proceed to test whether $V_2(k)$ is the same as $V_1(k)$. If not, we need to keep going. In practice, we may need to iterate quite a few times.

This iterative procedure works especially well when the operator T is a **contraction mapping**. [In this context, mapping is a synonym for operator]. A contraction mapping T has the property that for two functions V, W the distance between TV and TW is less than the distance between V and W . That is, a contracting operator brings functions closer together. In the context

of our iterative procedure, if T is a contraction then the iterates $T^j V_0$ are getting closer and closer together as j increases.

Sufficient conditions for the operator T to be a contraction are (i) that the operator is **monotone**, so if two functions V, W satisfy $V \leq W$, then $TV \leq TW$; and (ii) that the operator exhibits **discounting**, so if $a > 0$ is a constant and V is a function then $T(V + a) \leq TV + \beta a$ for some constant $\beta \in (0, 1)$. The Bellman operator for the growth model is in fact a contraction mapping.

In practice, value function iteration is easy to implement on a computer. We will discuss how to do this in the coming lectures.

D. Aside on the envelope theorem

Consider the choice of maximizing a concave function $U(x, \theta)$ by choice of $x \in X$ taking as given some parameter $\theta \in \Theta$. Associated with this problem is a maximum value function, given by

$$V(\theta) = U(\hat{x}(\theta), \theta) = \max_{x \in X} U(x, \theta)$$

where the optimal policy $\hat{x}(\theta)$ is

$$\hat{x}(\theta) \in \arg \max_{x \in X} U(x, \theta)$$

Since $U(x, \theta)$ is concave, this optimal policy is characterized by the first order condition

$$\frac{\partial U(\hat{x}(\theta), \theta)}{\partial x} = 0$$

Now consider the marginal effect of a change in θ on the value function. The total derivative of the value function with respect to the parameter θ is

$$V'(\theta) = \frac{\partial U(\hat{x}(\theta), \theta)}{\partial x} \hat{x}'(\theta) + \frac{\partial U(\hat{x}(\theta), \theta)}{\partial \theta}$$

But the optimal policy satisfies the first order condition, so we conclude

$$V'(\theta) = \frac{\partial U(\hat{x}(\theta), \theta)}{\partial \theta}$$

The total derivative of the value function with respect to the parameter θ is equal to the partial derivative of the utility function with respect to that parameter evaluated at the optimal policy. This is the essence of the envelope theorem.

Chris Edmond

10 August 2004