

Economics D11-1
Christiano
Fall, 2000

Class Notes on Sequence and Recursive Representations of the Planning Problem

In one representation of the central planning problem (“the sequence representation”), it is characterized as a problem of finding infinite sequences which solve a particular constrained optimization problem. In another representation (“the recursive representation”), we speak of a solution as object in a function space which solves a particular functional equation. These notes discuss, at an intuitive level, the basic equivalence between the two representations. The deterministic case is covered first, after which the stochastic case is considered. For a rigorous treatment of this material, see Nancy Stokey and Robert E. Lucas, Jr., with Ed Prescott, *Recursive Methods in Economic Dynamics*, Harvard University Press, 1989.

1. Deterministic Case

1.a. Sequence Representation

The sequence representation of the planning problem:

$$\max_{\{c_t, n_t, k_{t+1}; t=0,1,2,\dots\}} \sum_{t=0}^{\infty} \beta^t u(c_t, n_t) \quad (1)$$

subject to:

$$\begin{aligned} c^g + c_t + k_{t+1} &= f(k_t, n_t), t = 0, 1, 2, 3, \dots \\ k_0 &> 0, \text{ given} \\ 0 &\leq n_t \leq 1 \\ k_t &\geq 0, t > 0 \\ c_t &\geq 0. \end{aligned} \quad (2)$$

Here, c^g denotes a (constant) level of government consumption. The maximization here is over three infinite sequences of numbers:

$$\{c_0, c_1, c_2, \dots\}, \{n_0, n_1, n_2, \dots\}, \{k_1, k_2, \dots\} \quad (3)$$

that satisfy (2). Suppose $v(k_0)$ denotes the maximized value of the objective, (1). The function, $v(\cdot)$ traces out this value for all possible values of the initial capital stock, k_0 . (Imagine the *value function*, v , was found by solving (1)-(2) repeatedly for different values of k_0 .) Writing (1) out explicitly,

$$v(k_0) = \max_{\{c_t, n_t, k_{t+1}; t=0,1,2,\dots\}} \{u(c_0, n_0) + \beta u(c_1, n_1) + \beta^2 u(c_2, n_2) + \dots\}, \quad (4)$$

subject to (2). This is a *huge* optimization problem, as it requires searching for three points in infinite dimensional space (i.e., the three infinite sequences in (3)) to maximize a function, (1), of infinitely many variables. Typically, optimizing a function of just three or four variables is difficult, and so trying to optimize a function of *infinitely* many variables is particularly sobering! Clearly, some way to simplify the problem is needed.

1.b Recursive Representation

One standard way to proceed with problems like (4) is to break it up into two pieces. Before discussing how this is done for (4), I will describe an example.

An Example

Consider the maximization problem

$$\max_{x,y} \{ -[(x - .5)^2 + (y - .5)^2] \} \quad (5)$$

subject to

$$x \geq 0, y \geq 0, x + y \leq 2. \quad (6)$$

Note that the constraint set forms a triangle in two-dimensional Euclidean space. The problem has a unique solution at $x = .5, y = .5$. If you didn't know the solution, you could approach the problem of finding it in (at least) two different ways. One is to search directly over points in the constraint set, (6), for the x, y combination which maximizes (5). This is the approach that first comes to mind when the planning problem is viewed as a sequence problem, as in (1)-(2). The second approach is recursive, and goes as follows.

Choose an x inside the constraint set, (6), i.e., $0 \leq x \leq 2$. Then optimize (5) with respect to y subject to $0 \leq y \leq 2 - x$. Now pick another value of x and reoptimize over y . Stop when you've found an x that results in the highest possible value of the objective function. The fact that this algorithm works reflects that the optimization problem in (5)-(6) can equivalently be represented as follows:

$$\max_{0 \leq x \leq 2} \left(\max_{0 \leq y \leq 2-x} \{ -[(x - .5)^2 + (y - .5)^2] \} \right)$$

or, alternatively,

$$\max_{0 \leq x \leq 2} \left(-(x - .5)^2 + \{ \max_{0 \leq y \leq 2-x} [-(y - .5)^2] \} \right) \quad (7)$$

A more explicit, but also notationally more cumbersome, way to write (7) is as follows. Define, for any $0 \leq x \leq 2$, the function $h(x)$,

$$h(x) = \max_{0 \leq y \leq 2-x} [-(y - .5)^2]. \quad (8)$$

Then, (7) is equivalent with:

$$\max_{0 \leq x \leq 2} -(x - .5)^2 + h(x). \quad (9)$$

You should convince yourself that (5)-(6) and (7) are the same problem, and have the same solution. Keep this equivalence in mind as we return to the analysis of (4).

In particular, it is easy to see that (4) is equivalently written as:

$$v(k_0) = \max_{\{c_0, n_0, k_1\}} \left(u(c_0, n_0) + \beta \left[\max_{\{c_t, n_t, k_{t+1}; t=1, 2, \dots\}} \{ u(c_1, n_1) + \beta u(c_2, n_2) + \dots \} \right] \right), \quad (10)$$

subject to (2). In words, the equivalence between (10) and (4) can be stated as follows: "Maximizing (1) with respect to all control variables simultaneously is equivalent to the following procedure. Consider a particular set of feasible values of c_0, n_0, k_1 . Then maximize the term after the first plus sign in (10). Then consider another set of values for c_0, n_0, k_1 . Proceed in this

way until a set of values for c_0, n_0, k_1 which attains the maximum, i.e., $v(k_0)$, is found." Problem (10) is to (4) what (7) is to (5)-(6).

The important thing to note is that the problem in square brackets in (10) has *exactly* the same form as (4). The only difference is that the initial capital stock for the former problem is k_1 , whereas the initial capital stock for (4) is k_0 . Therefore, the object in square brackets in (10) is $v(k_1)$, where the $v(\cdot)$ is as defined in (4). Doing the relevant substitution, we can rewrite (10):

$$v(k_0) = \max_{\{c_0, n_0, k_1\}} (u(c_0, n_0) + \beta v(k_1)), \quad (11)$$

subject to $c_0 \geq 0, 0 \leq n_0 \leq T, c_0 + c^g + k_1 = f(k_0, n_0)$. Evidently, *if* we had the function $v(\cdot)$, then solving the planning problem would be straightforward. Just solve (11) for c_0, n_0 and k_1 . Then, to get c_t, n_t, k_{t+1} , solve

$$\max_{\{c_t, n_t, k_{t+1}\}} (u(c_t, n_t) + \beta v(k_{t+1})), t = 1, 2, 3, \dots,$$

subject to (2). Under a general set of circumstances, the sequences, $\{c_t, t = 0, 1, \dots\}, \{n_t, t = 0, 1, 2, \dots\}, \{k_{t+1}, t = 0, 1, 2, \dots\}$, obtained in this way coincide with the sequences, (3), which solve (1)-(2).

A more straightforward and mathematically useful way to characterize this solution strategy is as follows. First, we need a function, $v(\cdot)$, which solves

$$v(k) = \max_{\substack{\{c, n, k'\} \\ c \geq 0, 0 \leq n \leq 1, c + c^g + k' \leq f(k, n)}} (u(c, n) + \beta v(k')), \quad (12)$$

where no time subscript means this period's value and a prime means next period's value. Second, find the functions

$$k' = g(k), n = h(k), c = q(k) \quad (13)$$

which attain the maximum in (12). Equation (12) is the recursive representation of the planning problem. Clearly, with the functions, (13), in hand, we can compute sequences $\{c_t, t = 0, 1, \dots\}, \{n_t, t = 0, 1, 2, \dots\}, \{k_{t+1}, t = 0, 1, 2, \dots\}$. These coincide with the sequences, (3), which solve (1)-(2). For a formal proof (and a precise statement of assumptions) that the solution of the sequence problem, (1)-(2), coincides with the sequences generated by the functions obtained by solving the functional equation, (12), see Stokey and Lucas with Prescott, chapter 4. (Equation (12) is a *functional equation* because it is an equation in which the unknown is a function, $v(\cdot)$.)

In class, we discussed a related method for finding the functions g and h , which does not involve finding $v(\cdot)$ first, but which can be motivated using (12). First, if $u(\cdot, \cdot)$ is strictly increasing in c and decreasing in n , then

$$q(k) = f(k, h(k)) - c^g - g(k), \quad (14)$$

so that the only functions that need to be found are g and h . Second, if assumptions are satisfied for decisions to be in the interior of their respective constraint sets and for $v(\cdot)$ to be differentiable, then the first order conditions for n and k' associated with the maximization in (12) are

$$u_c(c, n)f_n(k, n) + u_n(c, n) = 0 \quad (15)$$

$$u_c(c, n) = \beta v_k(k') \quad (16)$$

$$v_k(k) = u_c(c, n)f_k(k, n) \quad (17)$$

where c is understood to be defined by (14), and a subscript by a function means the partial derivative with respect to that argument. Use (15) to implicitly define n as a function of k and k' :

$$n = H(k, k'). \quad (18)$$

Substituting (18) and (17) into (16), we get

$$v(k, k', k'') = 0, \quad (19)$$

where

$$v(k, k', k'') \equiv u_c(c, n) - \beta u_c(c', n')f_k(k', n'). \quad (20)$$

(The v functions in (19) and (12) are different objects.) The g function we seek has the property

$$v(k, g(k), g[g(k)]) = 0 \quad (21)$$

for all k . Although this strategy has the advantage of not requiring $v(\cdot)$, it has the problem that in general, there are more than one g functions that satisfy (21), and some means has to be found to identify which is the one we wish, i.e., which is the one that attains the maximum in (12).

2. Stochastic Problem.

2.a Sequence Representation

Suppose now that rather than being constant, government consumption is stochastic. To keep things simple, suppose it can take on only two values, “low”, c_l^g , and “high”, c_h^g :

$$c_t^g \in \{c_l^g, c_h^g\} \quad (22)$$

Let s^t denote a history of government spending up to time t :

$$s^t = (c_0^g, c_1^g, c_2^g, c_3^g, \dots, c_t^g). \quad (23)$$

Only one value of s^0 is possible, since c_0^g is observed at date 0. However, for $t > 0$, several histories, s^t , are possible. For example, there are two possible values of s^1 : (c_0^g, c_l^g) and (c_0^g, c_h^g) . There are four possible values of s^2 : (c_0^g, c_l^g, c_l^g) , (c_0^g, c_l^g, c_h^g) , (c_0^g, c_h^g, c_l^g) , (c_0^g, c_h^g, c_h^g) . Let the probability, as of date 0, of any history be denoted by:

$$probability(s^t) \equiv \mu(s^t), t = 1, 2, \dots \quad (24)$$

These probabilities satisfy $\sum_{s^t} \mu(s^t) = 1$ for $t = 1, 2, 3, \dots$, where \sum_{s^t} denotes the sum over all possible date t histories. We assume that c_t^g is first order Markov. That is, the distribution of c_t^g , conditional on all past values of c_t^g , is only a function of the period $t-1$ realization of c_t^g .

To help you visualize the possible histories, you should at this point draw them in the form of a tree diagram. For a particular date, the set of possible histories includes all the possible sequence of branches one could possibly travel down to get to that date. Obviously, for low values of t , there are relatively few s^t 's. The number of s^t 's increases without bound as t increases.

The planner's problem is to select consumption, labor and investment decisions corresponding to each date, state pair in the tree diagram. Evidently, we need a notation which allows us to distinguish not just different dates, but different histories too. Thus, let consumption at date t , with history s^t , be denoted by $c(s^t)$. Similarly, denote employment in date t , with history s^t , by $n(s^t)$. Finally, let the capital stock available at the beginning of period $t+1$, given a date t history of s^t be denoted by $k(s^t)$. (Note: when we index by dates only, and not histories, then we refer to this object as k_{t+1} . This notational convention is standard.) As in the deterministic case, we assume that date t utility is discounted by β^t . In addition, assume it is expected utility that the planner cares about, so that we weigh utility for a given history,

s^t , by its probability, $\mu(s^t)$. Thus, the planner ranks alternative state-date consumption and labor sequences according to:

$$\sum_{t=0}^{\infty} \beta^t \sum_{s^t} \mu(s^t) u(c(s^t), n(s^t)). \quad (25)$$

The planner seeks to maximize (25) by choice of $\{c_0, n_0, k_1\}$, and $\{c(s^t), n(s^t), k(s^t)\}$, for all s^t , and all $t = 1, 2, \dots$. These two sets are an enumeration of consumption, hours worked and investment at all the dots in the figure. The planner must respect the following constraints:

$$\begin{aligned} c_0 + c_0^g + k_1 &\leq f(k_0, n_0) \\ k_0 &\geq 0, \text{ given} \\ c(s^t) + c^g(s^t) + k(s^t) &\leq f(k(s^{t-1}), n(s^t)), t = 1, 2, \dots \\ 0 &\leq n_0 \leq 1, 0 \leq c_0, 0 \leq k_1 \\ 0 &\leq n(s^t) \leq 1, 0 \leq c(s^t), 0 \leq k(s^t), t = 1, 2, \dots \end{aligned} \quad (26)$$

Denote the maximized value of (25) by $v(k_0, s^0)$. Then,

$$v(k_0, s^0) = \max_{\substack{\{c_0, n_0, k_1, \\ c(s^t), n(s^t), k(s^t), t=1, 2, \dots\}}} \sum_{t=0}^{\infty} \beta^t \sum_{s^t} \mu(s^t) u(c(s^t), n(s^t)), \quad (27)$$

subject to (26). This is the sequence representation of the planning problem when there is uncertainty. Note that, although we use only one exogenous random variable which can only take on two possible values, this is not essential. In particular, the notation in (26)-(27) is consistent with the presence of any number of exogenous random variables, as long as each has a discrete distribution. (If there were more random variables in the problem, e.g., a technology shock, its history would have to be included in s^t , and its value would be added as an argument to the production function.)

2.b Recursive Representation

We convert (26)-(27) into the recursive representation of the problem in the same way we did in the deterministic case. In particular, we break up the maximization in (27) into two parts. The first involves decisions in date 0, and the second involves the future. Unfortunately, though the result is

essentially identical to what we get in the deterministic case, it is messier to write out. Suggestions about simplifying this are solicited.

$$\begin{aligned}
v(k_0, s^0) = & \max_{\substack{\{c_0, n_0, k_1, \\ c(s^t), n(s^t), k(s^t), t=1, 2, \dots\}}} \{u(c_0, n_0) + \beta[\mu(c_i^g)u(c(c_i^g), n(c_i^g)) + \mu(c_h^g)u(c(c_h^g), n(c_h^g))] \\
& + \beta^2 \sum_{s^2} \mu(s^2)u(c(s^2), n(s^2)) + \beta^3 \sum_{s^3} \mu(s^3)u(c(s^3), n(s^3)) + \dots\} \quad (28)
\end{aligned}$$

To simplify notation, I don't list c_0^g explicitly when writing out s^t explicitly. Now,

$$\begin{aligned}
\sum_{s^2} \mu(s^2)u(c(s^2), n(s^2)) &= \mu(c_i^g, c_i^g)u(c(c_i^g, c_i^g), n(c_i^g, c_i^g)) \\
&+ \mu(c_i^g, c_h^g)u(c(c_i^g, c_h^g), n(c_i^g, c_h^g)) \\
&+ \mu(c_h^g, c_i^g)u(c(c_h^g, c_i^g), n(c_h^g, c_i^g)) \\
&+ \mu(c_h^g, c_h^g)u(c(c_h^g, c_h^g), n(c_h^g, c_h^g)) \\
&= \mu(c_i^g | c_i^g)\mu(c_i^g)u(c(c_i^g, c_i^g), n(c_i^g, c_i^g)) \\
&+ \mu(c_i^g | c_h^g)\mu(c_h^g)u(c(c_i^g, c_h^g), n(c_i^g, c_h^g)) \\
&+ \mu(c_h^g | c_i^g)\mu(c_i^g)u(c(c_h^g, c_i^g), n(c_h^g, c_i^g)) \\
&+ \mu(c_h^g | c_h^g)\mu(c_h^g)u(c(c_h^g, c_h^g), n(c_h^g, c_h^g)) \quad (29)
\end{aligned}$$

where $\mu(c_i^g | c_i^g)$ is the probability that $c_2^g = c_i^g$, conditional on $c_1^g = c_i^g$. A standard result is $\mu(c_i^g, c_i^g) = \mu(c_i^g | c_i^g)\mu(c_i^g)$. Substituting (29) into (28)

$$\begin{aligned}
v(k_0, c_0^g) &= \max_{c_0, n_0, k_1} \{u(c_0, n_0) + \max_{c(s^t), n(s^t), k(s^t), t > 0} [\\
&+ \beta \mu(c_i^g) (u(c(c_i^g), n(c_i^g)) + \beta \mu(c_i^g | c_i^g) u(c(c_i^g, c_i^g), n(c_i^g, c_i^g)) + \beta \mu(c_h^g | c_i^g) u(c(c_i^g, c_h^g), n(c_i^g, c_h^g))) \\
&\beta \mu(c_h^g) (u(c(c_h^g), n(c_h^g)) + \beta \mu(c_i^g | c_h^g) u(c(c_h^g, c_i^g), n(c_h^g, c_i^g)) + \beta \mu(c_h^g | c_h^g) u(c(c_h^g, c_h^g), n(c_h^g, c_h^g))) \\
&+ \beta^3 \sum_{s^3} \mu(s^3) u(c(s^3), n(s^3)) + \dots \}] \} \quad (30)
\end{aligned}$$

Next, express the probabilities in the date 3 piece of this sum in terms of probabilities conditional on the realization of c_1^g . Then, collect terms in $\beta \mu(c_i^g)$ and $\beta \mu(c_h^g)$. Do this for the date 4 piece of the sum, then the date 5, and so on. Proceeding in this way, we ultimately find

$$\begin{aligned}
v(k_0, s^0) &= \max_{c_0, n_0, k_1} \{u(c_0, n_0) \\
&+ \mu(c_i^g) \beta \max_{\substack{c(s^t), n(s^t), k(s^t), t > 0 \\ c_1^g = c_i^g}} \sum_{t=1}^{\infty} \beta^{t-1} \sum_{s^t} \mu(s^t | c_1^g = c_i^g) u(c(s^t), n(s^t)) \\
&+ \mu(c_h^g) \beta \max_{\substack{c(s^t), n(s^t), k(s^t), t > 0 \\ c_1^g = c_h^g}} \sum_{t=1}^{\infty} \beta^{t-1} \sum_{s^t} \mu(s^t | c_1^g = c_h^g) u(c(s^t), n(s^t)) \}] \} \quad (31)
\end{aligned}$$

The thing to note in (31) is that the two summations from date $t = 1$ to ∞ have exactly the same form as (27). The only difference is that the initial conditions for the two summations in (31) are k_1, c_i^g and k_1, c_h^g , respectively, while the initial condition for (27) is k_1, c_0^g (c_0^g can be either c_i^g , or c_h^g .) Thus, we can rewrite (31) as:

$$v(k_0, s^0) = \max_{c_0, n_0, k_1} \{u(c_0, n_0) + \beta E[v(k_1, s^1) | k_0, s^0]\} \quad (32)$$

where

$$E[v(k_1, s^1) | k_0, s^0] = \mu(c_i^g) v(k_1, c_i^g) + \mu(c_h^g) v(k_1, c_h^g). \quad (33)$$

Equation (32) is the stochastic analog of (11). Note from (32) that only the current and next period's value of the shock enters. (This is due in part to our first-order Markov assumption on c_i^g .) Let s and s' denote this, and next period's value of the exogenous shock. Then, (32) can be written:

$$v(k, s) = \max_{c, n, k'} \{u(c, n) + \beta E[v(k', s') | k, s]\} \quad (34)$$

This is the recursive representation of the planning problem. To solve it, first solve this for v . Then, decision rules,

$$k' = g(k, s), n = h(k, s), c = f(k, h(k, s)) - c^g - g(k, s), \quad (35)$$

which attain the maximum in (34) can be found. They can be used to trace out, given the initial conditions, (k_0, s^0) , sequences $\{c_0, n_0, k_1\}$, and $\{c(s^t), n(s^t), k(s^t)\}$, for all s^t , and all $t = 1, 2, \dots\}$. Perhaps it is not surprising, given that all we have done is to rewrite the sequence problem, that these sequences correspond to the sequences which solve the sequence representation of the planning problem. (See Stokey and Lucas, with Prescott, chapter 9, for details.) With differentiability of the value function and interior solutions, it is easily verified that the investment decision rule, $g(\cdot, \cdot)$, satisfies the obvious stochastic analog of (21). This is the basis of the model solution strategy discussed in class.