A CHILD'S GUIDE TO DYNAMIC PROGRAMMING

1. Introduction

This is a simple guide to deterministic dynamic programming. In what follows, I borrow freely from King (1987) and Sargent (1987).

2. The Finite Horizon Problem

Consider the following time separable, recursive, problem:

$$\max \sum_{t=0}^{T} F_t(\mathbf{x}_t, \mathbf{u}_t) + W_0(\mathbf{x}_{T+1}) \tag{P}$$

subject to

$$\mathbf{x}_{t+1} = Q_t(\mathbf{x}_t, \mathbf{u}_t), \quad t = 0, \ldots, T; \tag{1}$$

$$\mathbf{x}_0 = \bar{x}_0; \tag{2}$$

$$\mathbf{x}_{T+1} \geq 0; \tag{3}$$

$$\mathbf{u}_t \in \Upsilon, \quad t = 0, \ldots, T. \tag{4}$$

where,

$\mathbf{x}_t$ is a n-vector of *state* variables (x_{it}).

$\mathbf{u}_t$ is a m-vector of *control* variables.

$F(\cdot)$ is a twice continuously differentiable objective function.

$Q(\cdot)$ is a vector of twice continuously differentiable *transition* function.

The different constraints are:

Equation (1) defines the transition equations for each state variables.

Equation (2) shows initial conditions for each state variable.

Equation (3) are terminal conditions for each state variable.

Equation (4) defines the feasible set for control variables.

This problem can be solved using a standard constrained optimization. The Lagrangian is

$$L = \sum_{t=0}^{T} F(\mathbf{x}_t, \mathbf{u}_t) + W_0(\mathbf{x}_{T+1}) + \sum_{t=0}^{T} \Lambda_t \left[Q(\mathbf{x}_t, \mathbf{u}_t) - \mathbf{x}_{t+1} \right],$$

where Λ_t is a n-vector of Lagrange multiplier. The first-order conditions are:

$$\frac{\partial L}{\partial \mathbf{u}_t} = \frac{\partial F_t}{\partial \mathbf{u}_t} + \frac{\partial Q_t}{\partial \mathbf{u}_t}\Lambda_t = F_{2t}(\mathbf{x}_t, \mathbf{u}_t) + Q_{2t}(\mathbf{x}_t, \mathbf{u}_t)\Lambda_t = 0, \quad t = 0, \ldots, T;$$

$$\frac{\partial L}{\partial \mathbf{x}_t} = \frac{\partial F_t}{\partial \mathbf{x}_t} + \frac{\partial Q_t}{\partial \mathbf{x}_t}\Lambda_t - \Lambda_{t-1} = F_{1t}(\mathbf{x}_t, \mathbf{u}_t) + Q_{1t}(\mathbf{x}_t, \mathbf{u}_t)\Lambda_t - \Lambda_{t-1} = 0, \quad 1, \ldots, T;$$

$$\frac{\partial L}{\partial \mathbf{x}_{T+1}} = \frac{\partial W_0}{\partial \mathbf{x}_{T+1}} - \Lambda_T = W_0'(\mathbf{x}_{T+1}) - \Lambda_T = 0.$$

Where $\partial F_t/\partial \mathbf{u}_t$ is a m-vector with $\partial F_t/\partial u_{jt}$ in the jth row (u_{jt} is the jth element of $\mathbf{u}_t$) and $\partial F_t/\partial \mathbf{x}_t$ is a n-vector with $\partial F_t/\partial x_{it}$ in the ith row. Similarly, $\partial Q_t/\partial \mathbf{u}_t$ is a $m \times n$ matrix with element $\partial Q_{it}/\partial u_{jt}$ in its ith column and jth row. Also, $\partial Q_t/\partial \mathbf{x}_t$ is a $n \times n$ matrix with element $\partial Q_{it}/\partial x_{jt}$ in its ith column and jth row. Finally, $\partial W_0/\partial \mathbf{x}_{T+1}$ is a n-vector with element $\partial W_0/\partial x_{jt}$ in its jth row.

Abstracting from second-order conditions, the maximum can be found as the solution to the set of first-order conditions. These can be solved recursively. For, example, at period T, we have

$$F_{2T}(\mathbf{x}_T, \mathbf{u}_T) + Q_{2T}(\mathbf{x}_T, \mathbf{u}_T)\Lambda_T = 0;$$

$$W_0'(\mathbf{x}_{T+1}) - \Lambda_T = 0;$$

$$\mathbf{x}_{T+1} = Q_T(\mathbf{x}_T, \mathbf{u}_T).$$

Given that $\mathbf{x}_T$ is fixed in period T, these can be solved to yield feedback rules:

$$\mathbf{x}_{T+1} = f_T(\mathbf{x}_T);$$

$$\mathbf{u}_T = h_T(\mathbf{x}_T);$$

$$\Lambda_T = \ell_T(\mathbf{x}_T).$$

Then, in period $T-1$, we have

$$F_{2T-1}(\mathbf{x}_{T-1}, \mathbf{u}_{T-1}) + Q_{2T-1}(\mathbf{x}_{T-1}, \mathbf{u}_{T-1})\Lambda_{T-1} = 0;$$

$$F_{1T}(\mathbf{x}_T, \mathbf{u}_T) + Q_{1T}(\mathbf{x}_T, \mathbf{u}_T)\Lambda_T - \Lambda_{T-1} = 0;$$

$$\mathbf{x}_T = Q_T(\mathbf{x}_{T-1}, \mathbf{u}_{T-1}).$$

Given the previous feedback rules, these can be solved for

$$\mathbf{x}_T = f_{T-1}(\mathbf{x}_{T-1});$$

$$\mathbf{u}_{T-1} = h_{T-1}(\mathbf{x}_{T-1});$$

$$\Lambda_{T-1} = \ell_{T-1}(\mathbf{x}_{T-1}).$$

Continuing the above recursion, we obtain feedback rules of the form

$$\mathbf{x}_{t+1} = f_t(\mathbf{x}_t), \quad t = 0, \ldots, T$$

$$\mathbf{u}_t = h_t(\mathbf{x}_t), \quad t = 0, \ldots, T;$$

$$\Lambda_t = \ell_t(\mathbf{x}_t), \quad t = 0, \ldots, T.$$

These rules or optimal policies are indexed by time. This is because, in general, these optimal rules vary through time.

3. An Economic Example

Consider the following economic example.

$$\max \ln(c_0) + \beta \ln(c_1) + \beta^2 \ln(c_2)$$

subject to

$$a_{t+1} = (1+r)a_t - c_t, \quad t = 0, 1, 2;$$

$$a_0 = \bar{a}_0;$$

$$a_3 = 0.$$

Clearly, if $a_3 = 0$, then we must have that $c_2 = (1+r)a_2$ and our objective function becomes:

$$\max \ln(c_0) + \beta \ln(c_1) + \beta^2 \ln((1+r)a_2)$$

The Lagrangian is

$$L = \ln(c_0) + \beta \ln(c_1) + \beta^2 \ln((1+r)a_2) + \lambda_0 \left[(1+r)a_0 - c_0 - a_1\right] + \lambda_1 \left[(1+r)a_1 - c_1 - a_2\right].$$

The first order conditions are

$$\frac{\partial L}{\partial c_t} = \beta^t \frac{1}{c_t} - \lambda_t = 0, \quad t = 0, 1;$$

$$\frac{\partial L}{\partial a_1} = (1 + r)\lambda_1 - \lambda_0 = 0;$$

$$\frac{\partial L}{\partial a_2} = \beta^2 \frac{1}{a_2} - \lambda_1 = 0.$$

In period 1, we have

$$\beta \frac{1}{c_1} - \lambda_1 = 0;$$

$$\beta^2 \frac{1}{a_2} - \lambda_1 = 0;$$

$$a_2 = (1 + r)a_1 - c_1.$$

These are solved to yield

$$a_2 = f_1(a_1) = (1 + r) \left(\frac{\beta}{1 + \beta} \right) a_1;$$

$$c_1 = h_1(a_1) = (1 + r) \left(\frac{1}{1 + \beta} \right) a_1;$$

$$\lambda_1 = \ell_1(a_1) = \beta \left(\frac{1 + \beta}{1 + r} \right) \frac{1}{a_1}.$$

Then, in period 0, we have

$$\frac{1}{c_0} - \lambda_0 = 0;$$

$$(1 + r)\lambda_1 - \lambda_0 = 0;$$

$$a_1 = (1 + r)a_0 - c_0;$$

which can be solved for

$$a_1 = f_0(a_0) = (1 + r) \left(\frac{\beta(1 + \beta)}{1 + \beta(1 + \beta)} \right) a_0;$$

$$c_0 = h_0(a_0) = (1 + r) \left(\frac{1}{1 + \beta(1 + \beta)} \right) a_0;$$

$$\lambda_0 = \ell_0(a_0) = \left(\frac{1 + \beta(1 + \beta)}{1 + r} \right) \frac{1}{a_0}.$$

Clearly, we have solved the whole system. Consumption and assets evolve according to the following scheme. Given a_0, consumption and assets purchased in period 0 are

$$c_0 = (1+r)\left(\frac{1}{1+\beta(1+\beta)}\right)a_0;$$

$$a_1 = (1+r)\left(\frac{\beta(1+\beta)}{1+\beta(1+\beta)}\right)a_0.$$

Then, given a_1, consumption assets purchased at period 1 are

$$c_1 = h_1(a_1) = (1+r)\left(\frac{1}{1+\beta}\right)a_1;$$

$$a_2 = f_1(a_1) = (1+r)\left(\frac{\beta}{1+\beta}\right)a_1.$$

4. Dynamic Programming

Let's recall the time separable, recursive, problem:

$$\max \sum_{t=0}^{T} F_t(\mathbf{x}_t, \mathbf{u}_t) + W_0(\mathbf{x}_{T+1}) \qquad (P)$$

subject to

$$\mathbf{x}_{t+1} = Q_t(\mathbf{x}_t, \mathbf{u}_t), \quad t = 0, \ldots, T; \qquad (1)$$

$$\mathbf{x}_0 = \bar{x}_{i0}; \qquad (2)$$

$$\mathbf{x}_{T+1} \geq 0; \qquad (3)$$

$$\mathbf{u}_t \in \Upsilon, \quad t = 0, \ldots, T. \qquad (4)$$

Dynamic programming is based on Bellman's principle of optimality. It argues that the above problem can be solved by recursively solving Bellman's equations to find *time consistent* policy functions. That is, both the objective function and constraints assume that controls dated t, $\mathbf{u}_t$, influence state $\mathbf{x}_{t+s+1}$ and returns $F_{t+s}(\cdot)$ for $s \geq t$, but not earlier. Accordingly, it is possible to split the optimization problem in sequences of optimizations as suggested by Bellman's equations. These equations are

$$W_{T-t+1}(\mathbf{x}_t) = \max F_t(\mathbf{x}_t, \mathbf{u}_t) + W_{T-t}(\mathbf{x}_{t+1}), \quad t = 0, \ldots, T.$$

They yield optimal policy functions (or feedback rules)

$$\mathbf{x}_{t+1} = f_t(\mathbf{x}_t), \quad t = 0, \ldots, T;$$

$$\mathbf{u}_t = h_t(\mathbf{x}_t), \quad t = 0, \ldots, T.$$

These policy functions are self-enforcing or time consistent in that, as time advances, there is no incentive to depart from the original plan.

We wish to show that this recursion will yield the same solution as the one found in Section 1. In that section, we have solved the problem recursively backward. In a sense, we are doing the same here. Starting from the last period, period T, we define the value function for the one-period problem by

$$W_1(\mathbf{x}_T) = \max F_T(\mathbf{x}_T, \mathbf{u}_T) + W_0(\mathbf{x}_{T+1}) \quad \text{subject to} \quad \mathbf{x}_{T+1} = Q_T(\mathbf{x}_T, \mathbf{u}_T),$$

where $\mathbf{x}_T$ is given. The Lagrangian for this optimization is

$$L = F_T(\mathbf{x}_T, \mathbf{u}_T) + W_0(\mathbf{x}_{T+1}) + \Lambda_T \left[Q_T(\mathbf{x}_T, \mathbf{u}_T) - \mathbf{x}_{T+1} \right].$$

The first-order conditions are

$$\frac{\partial L}{\partial \mathbf{u}_T} = F_{2T}(\mathbf{x}_T, \mathbf{u}_T) + Q_{2T}(\mathbf{x}_T, \mathbf{u}_T)\Lambda_T = 0;$$

$$\frac{\partial L}{\partial \mathbf{x}_{T+1}} = W_0'(\mathbf{x}_{T+1}) - \Lambda_T = 0;$$

These are the same first-order conditions found in Section 2 above. Then, as before, these first-order conditions and the transition function $\mathbf{x}_{T+1} = Q_T(\mathbf{x}_T, \mathbf{u}_T)$ can be used to solve for feedback rules

$$\mathbf{x}_{T+1} = f_T(\mathbf{x}_T);$$

$$\mathbf{u}_T = h_T(\mathbf{x}_T).$$

$$\Lambda_T = \ell_T(\mathbf{x}_T).$$

We can now think of the two-period problem

$$W_2(\mathbf{x}_{T-1}) = \max F_{T-1}(\mathbf{x}_{T-1}, \mathbf{u}_{T-1}) + W_1(\mathbf{x}_T) \quad \text{subject to} \quad \mathbf{x}_T = Q_{T-1}(\mathbf{x}_{T-1}, \mathbf{u}_{T-1}),$$

with $\mathbf{x}_{T-1}$ given. The Lagrangian for this optimization is

$$L = F_{T-1}(\mathbf{x}_{T-1}, \mathbf{u}_{T-1}) + W_1(\mathbf{x}_T) + \Lambda_{T-1} \left[Q_{T-1}(\mathbf{x}_{T-1}, \mathbf{u}_{T-1}) - \mathbf{x}_T \right].$$

The first-order conditions are

$$\frac{\partial L}{\partial \mathbf{u}_{T-1}} = F_{2T-1}(\mathbf{x}_{T-1}, \mathbf{u}_{T-1}) + Q_{2T-1}(\mathbf{x}_{T-1}, \mathbf{u}_{T-1})\Lambda_{T-1} = 0;$$

$$\frac{\partial L}{\partial \mathbf{x}_T} = W_1'(\mathbf{x}_T) - \Lambda_{T-1} = 0.$$

The last condition requires solving for the derivative of the previous value function. To find this derivative, we exploit a specific property of value functions. This property is an envelope condition (sometime called the Benveniste and Scheinkman condition). It is obtained as follows. First, write the value function as

$$\begin{aligned} W_1(\mathbf{x}_T) &= \max F_T(\mathbf{x}_T, \mathbf{u}_T) + W_0\left(\mathbf{x}_{T+1}\right)) \\ &= \max F_T(\mathbf{x}_T, \mathbf{u}_T) + W_0\left(Q_T(\mathbf{x}_T, \mathbf{u}_T)\right) \\ &= F_T(\mathbf{x}_T, h_T(\mathbf{x}_T)) + W_0\left(Q_T(\mathbf{x}_T, h_T(\mathbf{x}_T))\right). \end{aligned}$$

Totally differentiating the value function yields

$$W_1'(\mathbf{x}_T) = (F_{1T} + Q_{1T}W_0') + \frac{\partial h_T}{\partial \mathbf{x}_T}\left(F_{2T} + Q_{2T}W_0'\right).$$

The initial set of first-order condition implies that

$$F_{2T}(\mathbf{x}_T, \mathbf{u}_T) + Q_{2T}(\mathbf{x}_T, \mathbf{u}_T)W_0'(\mathbf{x}_{T+1}) = 0.$$

Thus, the derivative of the value function is

$$W_1'(\mathbf{x}_T) = F_{1T}(\mathbf{x}_T, \mathbf{u}_T) + Q_{1T}(\mathbf{x}_T, \mathbf{u}_T)W_0'(\mathbf{x}_{T+1}).$$

Substituting the derivative of the value function in the last condition yields the following set of equations:

$$F_{2T-1}(\mathbf{x}_{T-1}, \mathbf{u}_{T-1}) + Q_{2T-1}(\mathbf{x}_{T-1}, \mathbf{u}_{T-1})\Lambda_{T-1} = 0;$$

$$F_{1T}(\mathbf{x}_T, \mathbf{u}_T) + Q_{1T}(\mathbf{x}_T, \mathbf{u}_T)\Lambda_T - \Lambda_{T-1} = 0;$$

$$\mathbf{x}_T = Q_T(\mathbf{x}_{T-1}, \mathbf{u}_{T-1}).$$

As before, these are the equations that permit us to write the feedback rules

$$\mathbf{x}_T = f_{T-1}(\mathbf{x}_{T-1});$$

$$\mathbf{u}_{T-1} = h_{T-1}(\mathbf{x}_{T-1});$$

$$\Lambda_{T-1} = \ell_{T-1}(\mathbf{x}_{T-1}).$$

Clearly, this recursion will yield the same solution as the one found in Section 1. Finally, it is not necessary to use Lagrange multipliers. In fact, the optimization

$$W_{T-t+1}(\mathbf{x}_t) = \max F_t(\mathbf{x}_t, \mathbf{u}_t) + W_{T-t}(\mathbf{x}_{t+1}), \quad t = 0, \ldots, T.$$

subject to

$$\mathbf{x}_{T+1} = Q_T(\mathbf{x}_T, \mathbf{u}_T), \quad t = 0, \ldots, T,$$

can be solved by substituting the transition function in the value function

$$W_{T-t+1}(\mathbf{x}_t) = \max F_t(\mathbf{x}_t, \mathbf{u}_t) + W_{T-t}\left(Q_t(\mathbf{x}_t, \mathbf{u}_t)\right).$$

In that case, the first-order condition is

$$F_{2t}(\mathbf{x}_t, \mathbf{u}_t) + Q_{2t}(\mathbf{x}_t, \mathbf{u}_t) W'_{T-t}(\mathbf{x}_{t+1}) = 0.$$

The Benveniste-Scheinkman condition is

$$W'_{T-t+1}(\mathbf{x}_t) = F_{1t}(\mathbf{x}_t, \mathbf{u}_t) + Q_{1t}(\mathbf{x}_t, \mathbf{u}_t) W'_{T-t}(\mathbf{x}_{t+1}).$$

6. An Economic Example Again

Recall our economic example:

$$\max \ln(c_0) + \beta \ln(c_1) + \beta^2 \ln((1+r)a_2)$$

subject to

$$a_{t+1} = (1+r)a_t - c_t, \quad t = 0, 1;$$

$$a_0 = \bar{a}_0.$$

The one-period problem is

$$W_1(a_1) = \max \beta \ln(c_1) + W_0(a_2);$$

subject to

$$a_2 = (1+r)a_1 - c_1;$$

$$W_0(a_2) = \beta^2 \ln((1+r)a_2).$$

Thus,

$$W_1(a_1) = \max \beta \ln(c_1) + W_0\left((1+r)a_1 - c_1\right).$$

The first-order condition is

$$\beta \frac{1}{c_1} - W_0' = 0$$

where $W_0' = \beta^2(1/a_2)$. Thus, we have

$$\beta \frac{1}{c_1} - \beta^2 \frac{1}{a_2} = 0;$$

$$a_2 = (1+r)a_1 - c_1.$$

Which yields the policy functions

$$c_1 = h_1(a_1) = (1+r)\left(\frac{1}{1+\beta}\right)a_1;$$

$$a_2 = f_1(a_1) = (1+r)\left(\frac{\beta}{1+\beta}\right)a_1.$$

The two-period problem is

$$W_2(a_0) = \max \ln(c_0) + W_1(a_1);$$

subject to

$$a_1 = (1+r)a_0 - c_0.$$

or

$$W_2(a_0) = \max \ln(c_0) + W_1\left((1+r)a_0 - c_0\right).$$

The first-order condition is

$$\frac{1}{c_0} - W_1' = 0,$$

where

$$W_1' = W_0'(1+r) = (1+r)\beta^2 \frac{1}{a_2} = \beta(1+\beta)\frac{1}{a_1}.$$

So, our equations are

$$\frac{1}{c_0} - \beta(1+\beta)\frac{1}{a_1} = 0;$$

$$a_1 = (1+r)a_0 - c_0.$$

We find the policy functions

$$c_0 = h_0(a_0) = (1+r)\left(\frac{1}{1+\beta(1+\beta)}\right)a_0;$$

$$a_1 = f_0(a_0) = (1+r)\left(\frac{\beta(1+\beta)}{1+\beta(1+\beta)}\right) a_0.$$

7. The Infinite Horizon Problem

In this section we consider an extension of the finite horizon problem to the case of infinite horizon. To simplify the exposition, we will focus our attention to stationary problems. The assumption of stationarity implies that the main functions of the problem are time invariant:

$$F_t(\mathbf{x}_t, \mathbf{u}_t) = F(\mathbf{x}_t, \mathbf{u}_t);$$
$$Q_t(\mathbf{x}_t, \mathbf{u}_t) = Q(\mathbf{x}_t, \mathbf{u}_t).$$

The result will be time invariant policy rules.

In general, to obtain an interior solution to our optimization problem, we require the objective function to be bounded away from infinity. One way to achieve boundedness is to assume discounting. However, the existence of a discount factor is neither necessary nor sufficient to ensure boundedness. Thus, we define

$$F(\mathbf{x}_t, \mathbf{u}_t) = \beta^t f(\mathbf{x}_t, \mathbf{u}_t).$$

A further assumption which is sufficient for boundedness of the objective function is boundedness of the return function in each period:

$$0 \le f(\mathbf{x}_t, \mathbf{u}_t) < k < \infty,$$

where k is some finite number.

Our infinite horizon problem is:

$$\max \sum_{t=0}^{\infty} \beta^t f(\mathbf{x}_t, \mathbf{u}_t) \qquad (P1)$$

subject to

$$\mathbf{x}_{t+1} = Q(\mathbf{x}_t, \mathbf{u}_t), \quad t = 0, \ldots, T;$$
$$\mathbf{x}_0 = \bar{x}_0.$$

The Bellman equation is

$$W_j(\mathbf{x}_t) = \beta^t f(\mathbf{x}_t, \mathbf{u}_t) + W_{j-1}(\mathbf{x}_{t+1}).$$

We can rewrite this equation in current value. To do so, simply define the current value function

$$V_j(\mathbf{x}_t) = \beta^{-t} W_j(\mathbf{x}_t)$$

such that

$$V_j(\mathbf{x}_t) = f(\mathbf{x}_t, \mathbf{u}_t) + \beta V_{j-1}(\mathbf{x}_{t+1}).$$

Often, authors will simply abstract from the time subscript and write

$$V_j(\mathbf{x}) = f(\mathbf{x}, \mathbf{u}) + \beta V_{j-1}(\mathbf{x}'),$$

where $\mathbf{x}'$ denotes next-period values.

There is a theorem in Bertsekas (1976) which states that, under certain conditions, iterations of the current value Bellman equations converge as $j \to \infty$. In this case, the limit function satisfies the following version of Bellman's equation:

$$V(\mathbf{x}) = \max f(\mathbf{x}, \mathbf{u}) + \beta V(\mathbf{x}').$$

This limiting function is the optimal value function for our problem

$$V(\mathbf{x}_0) = \max \sum_{t=0}^{\infty} \beta^t f(\mathbf{x}_t, \mathbf{u}_t),$$

where the maximization is subject to $\mathbf{x}_{t+1} = Q(\mathbf{x}_t, \mathbf{u}_t)$ and $\mathbf{x}_0$ given. It also turns out that it gives rise to a unique time-invariant optimal policy of the form $\mathbf{u}_t = h(\mathbf{x}_t)$. Finally, this limiting function generates the Benveniste-Scheinkman condition

$$V'(\mathbf{x}) = f_1(\mathbf{x}, \mathbf{u}) + \beta Q_1(\mathbf{x}, \mathbf{u}) V'(\mathbf{x}').$$

The above suggests that there are three ways to solve the optimization problem $P1$. The first method uses the Benveniste-Scheinkman condition, the second uses guesses of the value function, and the third uses iterations.

Method 1: Benveniste-Scheinkman

First, set up the Bellman equation:

$$V(\mathbf{x}) = \max f(\mathbf{x}, \mathbf{u}) + \beta V(\mathbf{x}').$$

subject to $\mathbf{x}' = Q(\mathbf{x}, \mathbf{u})$ and $\mathbf{x}_0$ given.

Second, find the first-order condition:

$$f_2(\mathbf{x}, \mathbf{u}) + \beta Q_2(\mathbf{x}, \mathbf{u}) V'(\mathbf{x}') = 0.$$

Three, use the Benveniste-Scheinkman condition:

$$V'(\mathbf{x}) = f_1(\mathbf{x}, \mathbf{u}) + \beta Q_1(\mathbf{x}, \mathbf{u}) V'(\mathbf{x}').$$

This should allow you to write an Euler equation. In theory, this Euler equation may be used to uncover the optimal decision rule $\mathbf{u} = h(\mathbf{x})$ that solves the problem.

Method 2: Undetermined Coefficients

This method only works on a limited set of problems. The idea is to guess the value function, and then verify this guess.

First, set up the Bellman equation:

$$V(\mathbf{x}) = \max f(\mathbf{x}, \mathbf{u}) + \beta V(\mathbf{x}').$$

subject to $\mathbf{x}' = Q(\mathbf{x}, \mathbf{u})$ and $\mathbf{x}_0$ given.

Second, guess the form of the value function: $V(\mathbf{x}) = V^g(\mathbf{x})$. Substitute the guess in the Bellman equation

$$V(\mathbf{x}) = \max f(\mathbf{x}, \mathbf{u}) + \beta V^g(\mathbf{x}').$$

subject to $\mathbf{x}' = Q(\mathbf{x}, \mathbf{u})$ and $\mathbf{x}_0$ given.

Three, perform the maximization and obtain policy rule $h^g(\mathbf{x})$. Substitute the policy rule the in value function

$$V(\mathbf{x}) = f(\mathbf{x}, h^g(\mathbf{x})) + \beta V^g(Q(\mathbf{x}, h^g(\mathbf{x})).$$

Fourth, verify that the form of $V(\mathbf{x})$ is the same as $V^g(\mathbf{x}')$:

$$V^g(\mathbf{x}) = f(\mathbf{x}, h^g(\mathbf{x})) + \beta V^g(Q(\mathbf{x}, h^g(\mathbf{x})).$$

If the guess is correct, the problem is solved. If the guess is incorrect, try the form of the value function that is suggested by the first guess as a second guess, and repeat the process.

Method 3: Iterations.

First, set up the Bellman equation:

$$V_1(\mathbf{x}) = \max f(\mathbf{x}, \mathbf{u}) + \beta V_0(\mathbf{x}').$$

subject to $\mathbf{x}' = Q(\mathbf{x}, \mathbf{u})$ and $\mathbf{x}_0$ given.

Second, set $V_0 = g(\mathbf{x})$, for a function $g(\cdot)$, as a starting point. Solve the remaining optimization problem

$$V_1(\mathbf{x}) = \max f(\mathbf{x}, \mathbf{u}) + \beta g(\mathbf{x}')$$

subject to $\mathbf{x}' = Q(\mathbf{x}, \mathbf{u})$. This should yield a policy function $\mathbf{u} = h_1(\mathbf{x})$.
Third, substitute the policy function in the return function

$$V_1(\mathbf{x}) = f(\mathbf{x}, h_1(\mathbf{x})) + \beta g(Q(\mathbf{x}, h_1(\mathbf{x}))$$

and write the new Bellman equation

$$V_2(\mathbf{x}) = \max f(\mathbf{x}, \mathbf{u}) + \beta V_1(\mathbf{x}')$$

subject to $\mathbf{x}' = Q(\mathbf{x}, \mathbf{u})$. This should yield a policy function $\mathbf{u} = h_2(\mathbf{x})$.
Fourth, substitute the policy function in the return function

$$V_2(\mathbf{x}) = f(\mathbf{x}, h_2(\mathbf{x})) + \beta V_1(Q(\mathbf{x}, h_2(\mathbf{x})).$$

Write the new Bellman equation

$$V_3(\mathbf{x}) = \max f(\mathbf{x}, \mathbf{u}) + \beta V_2(\mathbf{x}')$$

subject to $\mathbf{x}' = Q(\mathbf{x}, \mathbf{u})$. This should yield a policy function $\mathbf{u} = h_3(\mathbf{x})$.
Fifth, continue iterating until the value function $V_j(\mathbf{x})$ converges to $V(\mathbf{x})$. The control rule associated with the converged value solves the problem.

8. The Economic Example in Infinite Horizon

Consider the following problem:

$$\max \sum_{0}^{\infty} \beta^t \ln(c_t)$$

subject to

$$a_{t+1} = (1 + r)a_t - c_t;$$
$$a_0 = \bar{a}_0.$$

We can solve the above problem using all three methods. We start with method 3.
Method 1: Benveniste-Scheinkman

The Bellman equation is:

$$V(a) = \max \ln(c) + \beta V(a') \quad \text{subject to} \quad a' = (1+r)a - c.$$

The optimization is

$$V(a) = \max \ln(c) + \beta V((1+r)a - c).$$

The first-order condition is

$$\frac{1}{c} - \beta V'(a') = 0.$$

The Benveniste-Scheinkman condition is

$$V'(a) = \beta V'(a')(1+r).$$

Accordingly,

$$V'(a) = \frac{1}{c}(1+r) \quad \text{and} \quad V'(a') = \frac{1}{c'}(1+r).$$

Then,

$$\frac{1}{c} = (1+r)\beta\frac{1}{c'}$$

or

$$\beta\frac{c_t}{c_{t+1}} = \frac{1}{1+r}.$$

This last equation is our Euler equation. We rewrite it as

$$c_{t+1} = \beta(1+r)c_t$$

or

$$c_{t+j} = (\beta(1+r))^j \, c_t.$$

To solve for the optimal policy rule, we require the transition function

$$a_{t+1} = (1+r)a_t - c_t.$$

We can rewrite it as

$$a_t = \frac{1}{1+r}(a_{t+1} + c_t).$$

If we lead this forward one period and substitute it back in the transition function, we obtain

$$a_t = \left(\frac{1}{1+r}\right)^2 a_{t+2} + \left(\frac{1}{1+r}\right)^2 c_{t+1} + \frac{1}{1+r}c_t.$$

Repeating this forward substitution suggests that

$$a_t = \left(\frac{1}{1+r}\right)^T a_{t+T} + \frac{1}{1+r} \sum_{j=0}^{T-1} \left(\frac{1}{1+r}\right)^j c_{t+j}.$$

In the limit,

$$\lim_{T\to\infty} \left(\frac{1}{1+r}\right)^T a_{t+T} = 0.$$

Thus,

$$(1+r)a_t = \sum_{j=0}^{\infty} \left(\frac{1}{1+r}\right)^j c_{t+j}.$$

We can substitute in our Euler equation to obtain

$$\begin{aligned}(1+r)a_t &= \sum_{j=0}^{\infty} \left(\frac{1}{1+r}\right)^j (\beta(1+r))^j c_t \\ &= \sum_{j=0}^{\infty} \beta^j c_t.\end{aligned}$$

Now, it is simple to show that

$$\sum_{j=0}^{\infty} \beta^j = 1 + \beta + \beta^2 + \ldots = \frac{1}{1-\beta}.$$

For example, define $S = 1+\beta+\beta^2+\ldots$. Then, $\beta S = \beta+\beta^2+\beta^3+\ldots$. Finally, $S-\beta S = 1$ and $S = 1/(1-\beta)$.

Collating these terms, we obtain our optimal policy rule

$$c_t = h(a_t) = (1-\beta)(1+r)a_t, \quad t \geq 0;$$

$$a_{t+1} = f(a_t) = \beta(1+r)a_t, \quad t \geq 0.$$

Method 2: Undetermined Coefficients

The Bellman equation is:

$$V(a) = \max \ln(c) + \beta V(a') \quad \text{subject to} \quad a' = (1+r)a - c.$$

Guess the value function: $V^g(a) = \phi + \theta \ln(a)$. The optimization is

$$V(a) = \max \ln(c) + \beta\phi + \beta\theta \ln((1+r)a - c).$$

The first-order condition is

$$\frac{1}{c} - \frac{\beta\theta}{a'} = 0.$$

This condition implies that

$$a' = \beta\theta c = (1+r)a - c.$$

Thus,

$$c = h(a) = \left(\frac{1+r}{1+\beta\theta}\right) a.$$

Substituting our policy function in the Bellman equation, we obtain

$$V(a) = \ln\left[\left(\frac{1+r}{1+\beta\theta}\right) a\right] + \beta\phi + \beta\theta\ln\left[(1+r)a - \left(\frac{1+r}{1+\beta\theta}\right) a\right].$$

Thus,

$$\phi + \theta\ln(a) = \ln\left[\left(\frac{1+r}{1+\beta\theta}\right) a\right] + \beta\phi + \beta\theta\ln\left[(1+r)a - \left(\frac{1+r}{1+\beta\theta}\right) a\right].$$

Collecting terms on the right hand side, we find that

$$\phi + \theta\ln(a) = (1+\beta\theta)\ln\left[\frac{1+r}{1+\beta\theta}\right] + \beta\phi + \beta\theta\ln(\beta\theta) + (1+\beta\theta)\ln(a).$$

Comparing both sides of the equation suggests that

$$\theta = 1 + \beta\theta \Longrightarrow \theta = \frac{1}{1-\beta};$$

$$\phi = (1+\beta\theta)\ln\left[\frac{1+r}{1+\beta\theta}\right] + \beta\phi + \beta\theta\ln(\beta\theta).$$

The last condition can be solved for ϕ.

We have thus verified our guess. The value function is

$$V(a_t) = \phi + \frac{1}{1-\beta}\ln(a_t)$$

with ϕ defined above. The optimal policy rule is

$$c_t = h(a_t) = \left(\frac{1+r}{1+\beta\theta}\right) a_t = (1-\beta)(1+r)a_t.$$

Finally, our system evolves according to

$$c_t = (1-\beta)(1+r)a_t, \quad t \geq 0;$$

$$a_{t+1} = \beta(1+r)a_t, \quad t \geq 0.$$

Method 3: Iterations.

The Bellman equation is:

$$V(a) = \max \ln(c) + \beta V(a') \quad \text{subject to} \quad a' = (1+r)a - c.$$

As a starting point, set $V_0 = \ln(a)$. The optimization is

$$V_1(a) = \max \ln(c) + \beta \ln((1+r)a - c).$$

The first-order condition is

$$\frac{1}{c} - \frac{\beta}{a'} = 0.$$

This condition implies that

$$a' = \beta c = (1+r)a - c.$$

Thus,

$$c = h_1(a) = \left(\frac{1+r}{1+\beta}\right) a.$$

Substitute this policy function in the value function to find

$$\begin{aligned} V_1(a) &= \ln(h_1(a)) + \beta \ln((1+r)a - h_1(a)) \\ &= \ln\left[\left(\frac{1+r}{1+\beta}\right) a\right] + \beta \ln\left[(1+r)a - \left(\frac{1+r}{1+\beta}\right) a\right], \\ &= \phi_1 + (1+\beta)\ln(a), \end{aligned}$$

with $\phi_1 = (1+\beta)\ln[(1+r)/(1+\beta)] + \beta\ln(\beta)$.

We can then write a new value function as

$$V_2(a) = \max \ln(c) + \beta V_1(a') = \max \ln(c) + \beta\phi_1 + \beta(1+\beta)\ln(a').$$

The optimization is

$$V_2(a) = \max \ln(c) + \beta\phi_1 + \beta(1+\beta)\ln((1+r)a - c).$$

The first-order condition is

$$\frac{1}{c} - \frac{\beta(1+\beta)}{a'} = 0.$$

This condition implies that

$$a' = \beta(1+\beta)c = (1+r)a - c.$$

Thus,

$$c = h_2(a) = \left(\frac{1+r}{1+\beta+\beta^2}\right) a.$$

Substitute this policy function in the value function to find

$$\begin{aligned} V_2(a) &= \ln(h_2(a)) + \beta\phi_1 + \beta(1+\beta)\ln((1+r)a - h_2(a)) \\ &= \ln\left[\left(\frac{1+r}{1+\beta+\beta^2}\right) a\right] + \beta\phi_1 + \beta(1+\beta)\ln\left[(1+r)a - \left(\frac{1+r}{1+\beta+\beta^2}\right) a\right] \\ &= \phi_2 + (1+\beta+\beta^2)\ln(a). \end{aligned}$$

Continuing this iterative process we find that

$$h_j(a) = \left(\frac{1+r}{\sum_{i=0}^{j}\beta^i}\right) a;$$

$$V_j(a) = \phi_j + \left(\sum_{i=0}^{j}\beta^i\right)\ln(a).$$

Then, if we let $j \to \infty$ we find that

$$h(a) = \lim_{j\to\infty} h_j(a) = (1-\beta)(1+r)a;$$

$$V(a) = \lim_{j\to\infty} V_j(a) = \phi + \frac{1}{1-\beta}\ln(a).$$

Thus, our system evolves according to

$$c_t = (1-\beta)(1+r)a_t, \quad t \geq 0;$$

$$a_{t+1} = \beta(1+r)a_t, \quad t \geq 0.$$

9. Summary

To summarize, here is the simple cookbook. Assume that you face the following problem:

$$\max \sum_{t=0}^{\infty} \beta^t f(\mathbf{x}_t, \mathbf{u}_t)$$

subject to

$$\mathbf{x}_{t+1} = Q_t(\mathbf{x}_t, \mathbf{u}_t), \quad t = 0, \ldots, T;$$
$$\mathbf{x}_0 = \bar{x}_0.$$

Here is how to proceed:

Step 1 Set up the Bellman equation:

$$V(\mathbf{x}) = \max f(\mathbf{x}, \mathbf{u}) + \beta V(\mathbf{x}').$$

subject to $\mathbf{x}' = Q(\mathbf{x}, \mathbf{u})$ and $\mathbf{x}_0$ given.

Step 2 Substitute transition function in the Bellman equation:

$$V(\mathbf{x}) = \max f(\mathbf{x}, \mathbf{u}) + \beta V\left(Q(\mathbf{x}, \mathbf{u})\right).$$

Step 3 Find the first-order condition:

$$f_2(\mathbf{x}, \mathbf{u}) + \beta Q_2(\mathbf{x}, \mathbf{u}) V'(\mathbf{x}') = 0.$$

Step 4 Find the Benveniste-Scheinkman condition:

$$V'(\mathbf{x}) = f_1(\mathbf{x}, \mathbf{u}) + \beta Q_1(\mathbf{x}, \mathbf{u}) V'(\mathbf{x}').$$

Step 5 Find the optimal policy function:

$$\mathbf{u} = h(\mathbf{x}).$$

References

Bertsekas, Dimitri, 1976, *Dynamic Programming and Stochastic Control*, New York: Academic Press.

King, Ian P., 1987, *A Simple Introduction to Dynamic Programming in Macroeconomic Models*, mimeo Queen's University.

Sargent, Thomas J., 1987, *Dynamic Macroeconomic Theory*, Cambridge: Harvard University Press.