Dynamic Macroeconomic Theory

Thomas J. Sargent

Harvard University Press
Cambridge, Massachusetts
and London, England
1987

1 Dynamic Programming

This chapter introduces the basic ideas and methods of dynamic programming and displays the restrictions on a dynamic system and the objective function that must be met for dynamic programming to be applicable. Where these restrictions are satisfied, dynamic programming provides a powerful method for studying dynamic optimization. The required restrictions permit the analyst to break what is in general a single exceedingly large dimensional optimization problem into a collection of much smaller optimization problems that can be solved sequentially. That step usually affords computational simplicity and often provides analytical insights.

The restrictions on objective functions and the dynamic system required for dynamic programming are satisfied in many formulations of private agents' investment problems. There is a class of multiagent problems (differential games), however, in which the structure of interactions among different agents' decision problems prevents one or more agents' problems from conforming to the restrictions required for dynamic programming. For these agents, therefore, optimal decisions must be computed not sequentially but simultaneously. The inapplicability of sequential methods to such decision problems is called time inconsistency and was first studied in macroeconomic contexts by Kydland and Prescott (1977) and Calvo (1978).

Although the main ideas of dynamic programming are simple, the details can involve sophisticated mathematical arguments. In this chapter things have been kept at a heuristic level of presentation, with the hope of communicating the main ideas quickly and enabling the reader to use these techniques to solve problems. More thorough presentations of the subject are listed at the end of the chapter; in particular see Bertsekas (1976); Bertsekas and Shreve (1978); Lucas, Prescott, and Stokey (forthcoming); Bellman (1957); and Chow (1981).

1.1 A General Intertemporal Problem

Consider the following general intertemporal optimization problem under certainty. Let x_t be an $(n \times 1)$ vector of *state* variables at time t, $t = 0, 1, \ldots, T+1$. Let u_t be a $(k \times 1)$ vector of *control* variables at time t, $t = 0, \ldots, T$. (The terms "state" and "control" are ambiguous in the context of the problem of this section. A precise description of them will be postponed, pending consideration of the special problem of Section 1.2, in which they are well motivated.) The problem is to choose $u_0, u_1, \ldots, u_T, x_1, \ldots, x_{T+1}$ to maximize an objective function

$$R(x_0, u_0, x_1, u_1, \ldots, x_T, u_T, x_{T+1}), \tag{1.1}$$

subject to x_0 given and subject to a system of constraints connecting the controls and the states, which we write in the implicit form

$$G(x_0, u_0, x_1, u_1, \ldots, x_T, u_T, x_{T+1}) \geq 0. \tag{1.2}$$

In (1.2) we imagine that G is a collection of $(T+1)n$ functions. We imagine that R and G are sufficiently smooth and that R is sufficiently concave to permit the method of Kuhn and Tucker to be applied. We then have a standard classical constrained-optimization problem, which can be solved by forming the following Lagrangian and maximizing with respect to $u_0, u_1, \ldots, u_T, x_1, x_2, \ldots, x_{T+1}$:

$$\begin{aligned} J = {} & R(x_0, u_0, x_1, u_1, \ldots, x_T, u_T, x_{T+1}) \\ & + \mu' G(x_0, u_0, x_1, u_1, \ldots, x_T, u_T, x_{T+1}), \end{aligned} \tag{1.3}$$

where μ is a $[(T+1)n \times 1]$ vector of Lagrange multipliers. The solution of this problem can be represented as a set of functions $u_0 = H_0(x_0)$, $u_1 = H_1(x_0), \ldots, u_T = H_T(x_0)$, with the optimal controls expressed as a function of the initial given state x_0, and a set of functions $x_1 = w_1(x_0)$, $x_2 = w_2(x_0), \ldots, x_{T+1} = w_{T+1}(x_0)$, with subsequent states expressed as a function of the initial state x_0. It is a standard feature of this problem that the optimal controls $u_0, u_1, \ldots, u_T$ as functions of x_0 must be determined simultaneously. This feature can be verified by obtaining the first-order necessary conditions for maximizing (1.3) and by studying the structure of the Jacobian matrix for the system of first-order necessary conditions. The system of first-order conditions in general fails to be recursive or block recursive, so that the optimal values for u_t and x_t are simultaneously determined.

1.2 A Recursive Problem

For dynamic problems in which the horizon T is large, it would be convenient if the problem could somehow be specialized to avoid the need to compute all of the controls simultaneously. This consideration has led to the following specialization of (1.1) and (1.2), which permits a recursive approach to the computation of the optimal controls.

We assume that $r_t(x_t, u_t)$ is a concave function and that the set $\{x_{t+1}, x_t, u_t : x_{t+1} \leq g_t(x_t, u_t), u_t \in R^k\}$ is convex and compact. We thus replace (1.1) and (1.2) with the problem of maximizing by choice of $(u_0, x_1, u_1, \ldots, x_{T+1})$ the function

$$r_0(x_0, u_0) + r_1(x_1, u_1) + \ldots + r_T(x_T, u_T) + W_0(x_{T+1}), \tag{1.4}$$

subject to x_0 given and the "transition" equations

$$\begin{aligned} x_1 &= g_0(x_0, u_0) \\ x_2 &= g_1(x_1, u_1) \\ &\cdot \\ &\cdot \\ &\cdot \\ x_{T+1} &= g_T(x_T, u_T). \end{aligned} \tag{1.5}$$

The function $r_t(x_t, u_t)$ is called the one-period return function at t, whereas the function $g_t(x_t, u_t)$ is called the transition function at t. The structure of the transition equations (1.5) motivates the labeling of x_t as state and u_t as control variables. The state vector x_t constitutes a complete description of the current position of the system. As far as the current and future returns $r_s(x_s, u_s)$ for $s \geq t$ are concerned, past values of u_v and x_v for $v < t$ add no information beyond that contained in x_t. This result is a consequence of the particular time separable structure of (1.4) and (1.5). The control vector u_t contains variables under the partial control of the problem solver that impinge on x_{t+1}, given x_t. In general for a given problem, the appropriate definition of the state is not unique, there being alternative ways of completely describing the current position of the system. Many of the admissible definitions of the state will include redundancies.

In (1.4) and (1.5) the functions $r_t(x_t, u_t)$, $W_0(x_{T+1})$, and $g_t(x_t, u_t)$ are assumed to be sufficiently smooth to permit the use of Lagrange's method.

Forming the Lagrangian, we have

$$L = r_0(x_0, u_0) + r_1(x_1, u_1) + \ldots + r_T(x_T, u_T) + W_0(x_{T+1}) + \lambda_0'[g_0(x_0, u_0) - x_1] + \lambda_1'[g_1(x_1, u_1) - x_2] + \ldots + \lambda_T'[g_T(x_T, u_T) - x_{T+1}], \tag{1.6}$$

where λ_t is an $(n \times 1)$ vector of Lagrange multipliers for $t = 0, \ldots, T$ and the prime denotes transposition.

The first-order necessary conditions for this problem are

$$\frac{\partial L}{\partial u_t} = \frac{\partial r_t}{\partial u_t}(x_t, u_t) + \frac{\partial g_t(x_t, u_t)}{\partial u_t}\lambda_t = 0, \qquad t = 0, \ldots, T \tag{1.7a}$$

$$\frac{\partial L}{\partial x_t} = \frac{\partial r_t(x_t, u_t)}{\partial x_t} + \frac{\partial g_t}{\partial x_t}(x_t, u_t)\lambda_t - \lambda_{t-1} = 0, \qquad t = 1, \ldots, T \tag{1.7b}$$

$$\frac{\partial L}{\partial x_{T+1}} = W_0'(x_{T+1}) - \lambda_T = 0 \tag{1.7c}$$

$$x_{t+1} = g_t(x_t, u_t), \qquad t = 0, 1, \ldots T. \tag{1.7d}$$

Here $\partial r_t/\partial u_t$ is a $(k \times 1)$ vector with $\partial r_t/\partial u_{ti}$ in the ith row, where u_{ti} is the element in the ith row of u_t. Also, $\partial g_t/\partial u_t$ is a $(k \times n)$ matrix with $\partial g_{ti}/\partial u_{th}$ in the ith column and hth row, where g_{ti} is the ith row of g_t and u_{th} is the hth row of u_t. Solving (1.7b) for λ_{t-1} and shifting forward one period, we have

$$\lambda_t = \frac{\partial r_{t+1}(x_{t+1}, u_{t+1})}{\partial x_{t+1}} + \frac{\partial g_{t+1}(x_{t+1}, u_{t+1})}{\partial x_{t+1}}\lambda_{t+1}.$$

Using this and (1.7c) recursively to eliminate λ_t, $t = 0, \ldots, T$, from (1.7a), we obtain the following system:

$$\frac{\partial r_t}{\partial u_t}(x_t, u_t) + \frac{\partial g_t(x_t, u_t)}{\partial u_t}\left\{\frac{\partial r_{t+1}}{\partial x_{t+1}} + \frac{\partial g_{t+1}}{\partial x_{t+1}}\left[\frac{\partial r_{t+2}}{\partial x_{t+2}} + \frac{\partial g_{t+2}}{\partial x_{t+2}} \cdot \left(\frac{\partial r_{t+3}}{\partial x_{t+3}} + \frac{\partial g_{t+3}}{\partial x_{t+3}}\left\{\ldots + \frac{\partial g_T}{\partial x_T}[W_0'(x_{T+1})]\right\}\right)\right]\right\} = 0 \tag{1.8a}$$

$$t = 0, \ldots, T-1$$

$$x_{t+1} = g_t(x_t, u_t), \qquad t = 0, \ldots, T-1 \tag{1.8b}$$

$$\frac{\partial r_T}{\partial u_T}(x_T, u_T) + \frac{\partial g_T(x_T, u_T)}{\partial u_T}W_0'(x_{T+1}) = 0 \tag{1.8c}$$

$$x_{T+1} = g_T(x_T, u_T), \tag{1.8d}$$

where in (1.8a) it is understood that g_t and r_t both have arguments (x_t, u_t).

In the special case in which $r_t(x_t, u_t)$ is quadratic, g_t is linear, and $\partial g_t/\partial x_t \equiv 0$, Equations (1.8a)–(1.8b) can be solved to yield a system of second-order difference equations in the vector x_t, subject to the initial condition that x_0 is given, and the terminal conditions (1.8c)–(1.8d). A further specialization results if the functions r_t and g_t are assumed to be time invariant so that (1.8) yields a set of time-invariant linear difference equations. In this case, the equations can be solved using methods similar to those illustrated in Sargent (1986, chap. 9). For more general specifications, however, it is useful to have an alternative method of solving the problem or at least of characterizing the solution, because nonlinear difference equations are generally very difficult to solve directly.

To motivate this method, notice the special structure of system (1.8), which is depicted in Table 1.1. The structure is special because (x_s, u_s) for $s < t$ does not appear directly in the marginal conditions and transition laws dated t and later. This fact makes it feasible to use the following "backward" recursive solution strategy.

Given x_T, the (subsystems of the) last two equations of system (1.8), namely (1.8c) and (1.8d), form a system of $(n + k)$ equations in (x_{T+1}, u_T). We solve these equations for x_{T+1} and u_T as functions of x_T, say,

$$x_{T+1} = f_T(x_T), \qquad u_T = h_T(x_T), \tag{1.9}$$

where $f_T(x_T) \equiv g_T[x_T, h_T(x_T)]$. Next, use $u_T = h_T(x_T)$ to eliminate u_T from the preceding two (subsystems of) equations in (1.8), namely (1.8a) and

Table 1.1 The structure of system (1.8)

Define the implicit functions ϕ_t^1, ϕ_t^2, $t = 0, 1, 2, \ldots, T$, by

$$\phi_t^1(x_t, u_t, x_{t+1}, u_{t+1}, \ldots, x_T, u_T, x_{T+1}) = \frac{\partial r_t}{\partial u_t} + \frac{\partial g_t}{\partial u_t}\left\{\frac{\partial r_{t+1}}{\partial x_{t+1}} + \frac{\partial g_{t+1}}{\partial x_{t+1}}\left(\ldots + \frac{\partial g_T}{\partial x_T}\right)[W_0'(x_{T+1})]\right\} = 0$$

and $\phi_t^2(x_t, u_t, x_{t+1}) = x_{t+1} - g_t(x_t, u_t) = 0$, respectively. Then (1.8) can be represented as

$$\begin{aligned}
&\phi_0^1(x_0, u_0, x_1, u_1, x_2, \ldots, x_T, u_T, x_{T+1}) = 0\\
&\phi_0^2(x_0, u_0, x_1) = 0\\
&\quad\phi_1^1(x_1, u_1, x_2, \ldots, x_T, u_T, x_{T+1}) = 0\\
&\quad\phi_1^2(x_1, u_1, x_2) = 0\\
&\qquad\qquad\phi_T^1(x_T, u_T, x_{T+1}) = 0\\
&\qquad\qquad\phi_T^2(x_T, u_T, x_{T+1}) = 0
\end{aligned}$$

(1.8b) for $t = T - 1$,

$$\frac{\partial r_{T-1}(x_{T-1}, u_{T-1})}{\partial u_{T-1}} + \frac{\partial g_{T-1}(x_{T-1}, u_{T-1})}{\partial u_{T-1}} \cdot \left[\frac{\partial r_T}{\partial x_T}(x_T, u_T) + \frac{\partial g_T}{\partial x_T}(x_T, u_T)\, W_0'(x_{T+1})\right] = 0 \tag{1.10}$$

$$x_T = g_{T-1}(x_{T-1}, u_{T-1}),$$

and solve these equations for u_{T-1} and x_T each as functions of x_{T-1}:

$$x_T = f_{T-1}(x_{T-1}), \qquad u_{T-1} = h_{T-1}(x_{T-1}). \tag{1.11}$$

One can continue recursively in this way, solving for a collection of feedback rules of the form

$$u_t = h_t(x_t), \qquad t = T, T-1, T-2, \ldots, 0, \tag{1.12}$$

where $u_t = h_t(x_t)$, $x_{t+1} = f_t(x_t)$ solve the equations

$$\frac{\partial r_t}{\partial u_t}(x_t, u_t) + \frac{\partial g_t}{\partial u_t}\left\{\frac{\partial r_{t+1}}{\partial x_{t+1}} + \frac{\partial g_{t+1}}{\partial x_{t+1}}\left[\frac{\partial r_{t+2}}{\partial x_{t+2}} + \frac{\partial g_{t+2}}{\partial x_{t+2}} \cdot \left(\frac{\partial r_{t+3}}{\partial x_{t+3}} + \frac{\partial g_{t+3}}{\partial x_{t+3}}\left\{\ldots + \frac{\partial g_T}{\partial x_T}[W_0'(x_{T+1})]\right\}\right)\right]\right\} = 0, \tag{1.13}$$

and $x_{s+1} = g_s(x_s, u_s)$ for $s = t, t+1, \ldots, T$, given that $u_{s+1} = h_{s+1}(x_{s+1})$ for $s = t, t+1, \ldots, T-1$.

1.3 Bellman's Equations

The equations of system (1.13) have interpretations as the marginal conditions from the following sequence of problems. Define the value function for a one-period problem $W_1(x_T)$ by

$$W_1(x_T) = \max_{u_T}\{r_T(x_T, u_T) + W_0(x_{T+1})\}, \tag{1.14}$$

subject to $x_{T+1} = g_T(x_T, u_T)$, with x_T given. We form the Lagrangian for this problem, and the first-order conditions can be expressed, after the Lagrange multiplier has been eliminated, as

$$\frac{\partial r_T}{\partial u_T}(x_T, u_T) + \frac{\partial g_T(x_T, u_T)}{\partial u_T} W_0'(x_{T+1}) = 0, \tag{1.15}$$

which precisely matches the marginal condition for u_T in (1.8). Equation (1.15) and the transition law $x_{T+1} = g_T(x_T, u_T)$ are to be solved jointly for $u_T = h_T(x_T)$. Now imagine substituting the solution $u_T = h_T(x_T)$ of (1.15) and (1.8d) into (1.14) to get

$$W_1(x_T) = r_T[x_T, h_T(x_T)] + W_0(g_T[x_T, h_T(x_T)]). \tag{1.16}$$

Formally, differentiating (1.16) gives

$$W_1'(x_T) = \left(\frac{\partial r_T}{\partial x_T} + \frac{\partial g_T}{\partial x_T} W_0'\right) + \frac{\partial h_T}{\partial x_T}\left[\frac{\partial r_T}{\partial u_T} + \frac{\partial g_T}{\partial u_T} W_0'(x_{T+1})\right],$$

where all functions dated T are evaluated at $[x_T, h_T(x_T)]$ and which by virtue of (1.15) becomes

$$W_1'(x_T) = \frac{\partial r_T}{\partial x_T}[x_T, h_T(x_T)] + \frac{\partial g_T}{\partial x_T}[x_T, h_T(x_T)] W_0'(g_T[x_T, h_T(x_T)]). \tag{1.17}$$

Because we have not shown that $\partial h_T/\partial x_T$ exists, this argument is informal or heuristic and should be regarded as only a way of remembering the correct answer. Correct arguments are given by Benveniste and Scheinkman (1979) and Lucas (1977).

Now define the value function for the two-period problem $W_2(x_{T-1})$ as

$$W_2(x_{T-1}) = \max_{u_{T-1}}\{r_{T-1}(x_{T-1}, u_{T-1}) + W_1(x_T)\}, \tag{1.18}$$

subject to $x_T = g_{T-1}(x_{T-1}, u_{T-1})$, with x_{T-1} given. If we proceed as with the problem defined by (1.14), the first-order condition for the problem on the right side of (1.18) can be expressed as

$$\frac{\partial r_{T-1}}{\partial u_{T-1}}(x_{T-1}, u_{T-1}) + \frac{\partial g_{T-1}(x_{T-1}, u_{T-1})}{\partial u_{T-1}} W_1'(x_T) = 0.$$

If we use formula (1.17) for $W_1'(x_T)$, this equation becomes

$$\frac{\partial r_{T-1}}{\partial u_{T-1}}(x_{T-1}, u_{T-1}) + \frac{\partial g_{T-1}(x_{T-1}, u_{T-1})}{\partial u_{T-1}} \cdot \left(\frac{\partial r_T}{\partial x_T}[x_T, h_T(x_T)] + \frac{\partial g_T}{\partial x_T}[x_T, h_T(x_T)] W_0'(g_T[x_T, h_T(x_T)])\right) = 0. \tag{1.19}$$

This equation and the transition law $x_T = g_{T-1}(x_{T-1}, u_{T-1})$ are to be solved jointly for $u_{T-1} = h_{T-1}(x_{T-1})$, $x_T = f_{T-1}(x_{T-1})$. Again proceeding as above, we can obtain

$$W_2'(x_{T-1}) = \frac{\partial r_{T-1}}{\partial x_{T-1}}[x_{T-1}, h_{T-1}(x_{T-1})] + \frac{\partial g_{T-1}}{\partial x_{T-1}} W_1'(g_{T-1}[x_{T-1}, h_{T-1}(x_{T-1})]), \tag{1.20}$$

or, using (1.17),

$$\begin{aligned} W_2'(x_{T-1}) = {} & \frac{\partial r_{T-1}}{\partial x_{T-1}}[x_{T-1}, h_{T-1}(x_{T-1})] \\ & + \frac{\partial g_{T-1}}{\partial x_{T-1}}[x_{T-1}, h_{T-1}(x_{T-1})] \\ & \cdot \left\{ \frac{\partial r_T}{\partial x_T}[x_T, h_T(x_T)] \right. \\ & \left. + \frac{\partial g_T}{\partial x_T}[x_T, h_T(x_T)]\, W_0'[f_T(x_T)] \right\}, \end{aligned}$$

where x_T is evaluated at $x_T = f_{T-1}(x_{T-1}) = g_{T-1}[x_{T-1}, h_{T-1}(x_{T-1})]$. Notice that Equation (1.19) is precisely the version of the marginal condition in (1.8) for u_{T-1}.

The pattern for the recursion is now set. We iterate on the following functional equation in the value functions

(1.21) $$W_{j+1}(x_{T-j}) = \max_{u_{T-j}}\{r_{T-j}(x_{T-j}, u_{T-j}) + W_j(x_{T-j+1})\},$$

subject to $x_{T-j+1} = g_{T-j}(x_{T-j}, u_{T-j})$, x_{T-j} given. The functional equation (1.21) is a version of Bellman's equation—named after Richard Bellman (1957). The idea is to proceed recursively and to work backward, first solving the one-period problem with $j+1=1$, deducing $W_1(x_T)$, then solving the two-period problem with $j+1=2$, deducing the two-period value function $W_2(x_{T-1})$. The process is repeated until we obtain the $(T+1)$-period value function $W_{T+1}(x_0)$. This procedure gives the optimal value of the problem as a function of the initial state x_0. Along the way we have calculated the optimal feedback rules $u_{T-j} = h_{T-j}(x_{T-j})$, $j = 0, 1, \ldots, T$. The preceding argument suggests that this backward recursion generates the same marginal conditions as the original problem (1.8). Indeed, the backward recursion technique always solves the original problem if a solution exists.

The derivative of the value functions obeys the recursion

$$\begin{aligned} W_{j+1}'(x_{T-j}) = {} & \frac{\partial r_{T-j}}{\partial x_{T-j}}[x_{T-j}, h_{T-j}(x_{T-j})] \\ & + \frac{\partial g_{T-j}}{\partial x_{T-j}} W_j'(g_{T-j}[x_{T-j}, h_{T-j}(x_{T-j})]). \end{aligned}$$

Comparing this equation with (1.7b) and (1.7c), we find that $W_j'(x_{T+1-j}) = \lambda_{T-j}$. The Lagrange multipliers λ_{T-j} in (1.6) thus give the marginal value of the state variables for the j-period problem.

The following observations supply another perspective on the recursive nature of our problem. Let us simply define the $(T+1)$-period value function $W_{T+1}(x_0)$ by

(1.22) $$\begin{aligned} W_{T+1}(x_0) = \max_{u_0, u_1, \ldots, u_T} & \{r_0(x_0, u_0) + r_1(x_1, u_1) + \ldots + r_T(x_T, u_T) \\ & + W_0(x_{T+1})\}, \end{aligned}$$

where the maximization is understood to be subject to $x_{t+1} = g_t(x_t, u_t)$, $t = 0, \ldots, T$, and x_0 given. Notice that the objective function and constraints (transition equations) have been specialized to have the key property that controls dated t influence states x_{s+1} and returns $r_s(x_s, u_s)$ for $s \geq t$ but not earlier. This key property gives the problem its recursive structure. In particular, the property makes it legitimate to cascade the maximization operator and to write (1.22) as

(1.23) $$\begin{aligned} W_{T+1}(x_0) = {} & \max_{u_0}\{r_0(x_0, u_0) + \max_{u_1}\{r_1(x_1, u_1) + \max_{u_2}\{r_2(x_2, u_2) \\ & + \ldots + \max_{u_T}\{r_T(x_T, u_T) + W_0(x_{T+1})\} \ldots \}\}\}, \end{aligned}$$

where the maximization over u_t is understood to be subject to $x_{t+1} = g_t(x_t, u_t)$ with x_t given. Equation (1.23) indicates that the original large optimization problem on the right side of (1.22) can be broken up into $(T+1)$ smaller problems. First, the problem in the innermost brackets is solved, the optimizer being $u_T = h_T(x_T)$ and the optimized value being $W_1(x_T)$. Then the problem in the second innermost brackets is solved for $u_{T-1} = h_{T-1}(x_{T-1})$ with optimized value $W_2(x_{T-1})$. This process of proceeding from the problems in the innermost brackets outward is equivalent to iterating on Bellman's functional equation (1.21).

The preceding argument implies that the optimal policies $u_t = h_t(x_t)$, $t = 0, \ldots, T$ have a self-enforcing character in the following sense. Consider the "remainder" of the objective function at some time $s > 0$, namely,

(1.24) $$\max_{u_s, u_{s+1}, \ldots, u_T}\{r_s(x_s, u_s) + \ldots + r_T(x_T, u_T) + W_0(x_{T+1})\},$$

subject to $x_{t+1} = g_t(x_t, u_t)$, $t = s, \ldots, T$, with x_s given. Then the solution of the maximum problem (1.24) is simply to use the remaining functions $u_s = h_s(x_s)$, $s = t, \ldots, T$ that were computed for the original problem. Furthermore, the maximized value of (1.24) is $W_{T-s+1}(x_s)$. Thus as time advances, there is no incentive to depart from the original plan. This self-enforcing character of optimal policies is known as Bellman's principle of optimality. Optimal policies that have this property are said to be time consistent. This property is special, is a consequence of the recursive character of the problem (1.4)–(1.5) and will not characterize the solutions of more general problems.

It is a feature of the solution to problem (1.4)–(1.5) that in general a different policy function $u_t = h_t(x_t)$, mapping the state at t into the control at t, is to be used at each date $t = 0, \ldots, T$. This is a consequence of two features of the problem: the fact that the horizon T is finite and the fact that the functions $r_t(x_t, u_t)$ and $g_t(x_t, u_t)$ have been permitted to depend on time in arbitrary ways. For many practical applications it is inconvenient that the policy function varies over time. One would like to discover contexts in which the same policy function is used for each period t. In the interests of achieving this objective, we now specialize problem (1.4)–(1.5) with the aim of generating conditions under which the policy functions h_j converge as $j \to -\infty$. We assume that

$$(1.25) \qquad \begin{aligned} r_t(x_t, u_t) &= \beta^t r(x_t, u_t), \qquad 0 < \beta < 1 \\ g_t(x_t, u_t) &= g(x_t, u_t). \end{aligned}$$

With this specification, Bellman's equation (1.21) becomes

$$W_{j+1}(x_{T-j}) = \max_{u_{T-j}}\{\beta^{T-j} r(x_{T-j}, u_{T-j}) + W_j(x_{T-j+1})\}.$$

Multiplying both sides by β^{j-T} gives

$$(1.25') \qquad \beta^{j-T} W_{j+1}(x_{T-j}) = \max_{u_{T-j}}\{r(x_{T-j}, u_{T-j}) + \beta \cdot \beta^{j-1-T} W_j(x_{T-j+1})\}.$$

Now define the current value function

$$V_{j+1}(x_{T-j}) = \beta^{j-T} W_{j+1}(x_{T-j}).$$

Notice that for $j = T$, we have $V_{T+1}(x_0) = W_{T+1}(x_0)$. Also notice that the current value function can be directly defined as

$$\begin{aligned} V_{j+1}(x_{T-j}) = \max_{u_{T-j}, u_{T-j+1}, \ldots, u_T} &\{r(x_{T-j}, u_{T-j}) + \beta r(x_{T-j+1}, u_{T-j+1}) \\ &+ \ldots + \beta^j r(x_T, u_T) + \beta^{j+1} V_0(x_{T+1})\}. \end{aligned}$$

In terms of the current value function, (1.25′) asserts that Bellman's equation becomes

$$(1.26) \qquad V_{j+1}(x_{T-j}) = \max_{u_{T-j}}\{r(x_{T-j}, u_{T-j}) + \beta V_j(x_{T-j+1})\},$$

subject to $x_{T-j+1} = g(x_{T-j}, u_{T-j})$ and x_{T-j} given. More compactly, we can write (1.26) as

$$(1.27) \qquad V_{j+1}(x) = \max_{u}\{r(x, u) + \beta V_j(\tilde{x})\},$$

subject to $\tilde{x} = g(x, u)$, x given, where the tilde denotes next-period values. Under particular conditions, iterations on (1.27) starting from any bounded and continuous initial V_0 converge as $j \to \infty$. See Bertsekas (1976, chap. 6) or Lucas, Prescott, and Stokey (forthcoming). The argument used to prove this claim is outlined in Section A.8. In this case the limit function $V = \lim_{j\to\infty} V_j$ satisfies the following version of Bellman's equation:

$$(1.28) \qquad V(x) = \max_{u}\{r(x, u) + \beta V(\tilde{x})\},$$

where the maximization is subject to $\tilde{x} = g(x, u)$, with x given. The limiting value function V that solves (1.28) turns out to be the optimal value function for the infinite horizon problem

$$(1.29) \qquad V(x_0) = \max_{\{u_t\}_{t=0}^{\infty}} \sum_{t=0}^{\infty} \beta^t r(x_t, u_t),$$

where the maximization is subject to $x_{t+1} = g(x_t, u_t)$, with x_0 given. Problem (1.29) is a version of a discounted dynamic programming problem. Under various particular regularity conditions,[1] it turns out that (1) the functional equation (1.28) has a unique strictly concave solution; (2) this solution is approached in the limit as $j \to \infty$ by iterations on (1.26) starting from any bounded and continuous initial V_0; (3) there is a unique and time-invariant optimal policy of the form $u_t = h(x_t)$, where h is chosen to maximize the right side of (1.28); (4) off corners, the limiting value function V is differentiable with

$$(1.30) \qquad V'(x) = \frac{\partial r}{\partial x}[x, h(x)] + \beta \frac{\partial g}{\partial x}[x, h(x)]\, V'(g[x, h(x)]).$$

This is a version of the formula of Benveniste and Scheinkman (1979). It is a great convenience of the specialization (1.25) of the objective function and transition functions, and also a convenience of the specification of an infinite horizon, that they imply a time-invariant policy function $u_t = h(x_t)$, for it is a routine practice in economics to seek setups in which agents use time-invariant decision rules. (Ample econometric considerations recommend or require such setups.)

The preceding results provide two methods for solving the functional

1. Alternative sets of regularity conditions work. One set of sufficient conditions is (1) r is concave and bounded, (2) the constraint set generated by g is convex and compact, that is, the set of $\{x_{t+1}, x_t, u_t: x_{t+1} \le g(x_t, u_t)\}$ for admissible u_t is convex and compact. See Lucas (1977), and Bertsekas (1976) for further details of convergence results. See Benveniste and Scheinkman (1979) and Lucas (1977) for the results on differentiability of the value function. A proof of the uniform convergence of iterations on (1.27) is contained in Section A.7 of the Appendix.

equation (1.28). The first method is constructive and simply involves iterating on (1.26), starting from $V_0 = 0$, until V_j has converged. The second method involves guessing a solution V and verifying that it is a solution to (1.28). The second method relies on the uniqueness of the solution to (1.28), but because it also relies on luck in making a good guess, it is not generally available. In the examples below, the guess-and-verify method is often used. The reader should, however, be alerted to the fact that the objective functions and constraints of these problems have been especially rigged so that the method will work. Essentially there are only two classes of specifications of preferences and constraints for which the method will work, namely, variants of specifications with linear constraints and quadratic preferences or Cobb-Douglas constraints and logarithmic preferences.

In many problems, there is no unique way of defining states and controls, and several alternative definitions lead to the same solution of the problem. Sometimes the states and controls can be defined in such a way that x_t does not appear in the transition equation, so that $\partial g_t/\partial x_t \equiv 0$. In this case, the system (1.8a)–(1.8b) simplifies to

$$\frac{\partial r_t}{\partial u_t}(x_t, u_t) + \frac{\partial g_t}{\partial u_t}(u_t) \cdot \frac{\partial r_{t+1}(x_{t+1}, u_{t+1})}{\partial x_{t+1}} = 0, \qquad x_{t+1} = g_t(u_t).$$

The first equation is a version of what is called an Euler equation. Under circumstances in which the second equation can be inverted to yield u_t as a function of x_{t+1}, using the second equation to eliminate u_t from the first equation produces a second-order difference equation in x_t.

Most of the dynamic programming problems that we solve in this book are discounted dynamic programming problems.

1.4 Nonstochastic Examples

We now consider several examples of single-agent optimization problems that can be solved using dynamic programming.

Saving under Certainty

Consider the problem of a consumer in a nonrandom environment who seeks to maximize $\Sigma_{t=0}^{\infty}\beta^t u(c_t)$, $0 < \beta < 1$, subject to $A_{t+1} = R_t(A_t + y_t - c_t)$, A_0 given, where y_t, $t = 0, 1, \ldots$, is a known sequence of exponential order less than $1/\beta$ and R_t, $t = 0, 1, \ldots$, is a known and given sequence of one-period gross rates of return on nonlabor wealth. Here c_t is consumption, A_t is nonlabor wealth at the beginning of time t, and y_t is labor income at t. Labor income is assumed to be beyond the control of the agent. For concreteness let y_t equal λy_{t-1}, and say that $R_t = R > 0$ for all t, assuming that $R > \lambda > 0$. To rule out a strategy of infinite consumption supported by unbounded borrowing, we also impose the restriction that, for $t \geq 0$,

$$(1.31) \qquad c_t + \sum_{j=1}^{\infty}\left(\prod_{k=0}^{j-1} R_{t+k}^{-1}\right) c_{t+j} = y_t + \sum_{j=1}^{\infty}\left(\prod_{k=0}^{j-1} R_{t+k}^{-1}\right) y_{t+j} + A_t.$$

We define the state of the system as (A_t, y_t, R_{t-1}) and define the control at t, u_t, as $R_t^{-1} A_{t+1} = A_t + y_t - c_t$. Evidently the control u_t is gross savings. The transition equation for A_t becomes $A_{t+1} = R_t u_t$, which does not involve the state at t. The function $r_t(x_t, u_t)$ becomes $\beta^t u(A_t + y_t - R_t^{-1}A_{t+1}) = \beta^t u(A_t + y_t - u_t)$. Bellman's equation becomes

v is from
relation

$$v(A_t, y_t, R_{t-1}) = \max_{u_t}\{u(A_t + y_t - u_t) + \beta v(u_t R_t, y_{t+1}, R_t)\},$$

why small v?

where $u_t = R_t^{-1}A_{t+1}$, $y_{t+1} = \lambda y_t$, $R_t = R$. Benveniste and Scheinkman's formula (1.30) gives $\partial v(A_t, y_t, R_{t-1})/\partial A_t = u'(c_t)$. The Euler equation for u_t then becomes

$$-\beta^t u'(A_t + y_t - R_t^{-1}A_{t+1}) + \beta^{t+1}R_t u'(A_{t+1} + y_{t+1} - R_{t+1}^{-1}A_{t+2}) = 0$$

or

$$(1.32) \qquad -u'(c_t) + \beta R_t u'(c_{t+1}) = 0.$$

We seek a consumption plan that satisfies (1.32) and the "isoperimetric condition" (1.31).

As an example, suppose that $u(c_t) = \ln c_t$. Then (1.32) requires that

$$c_{t+j} = \beta^j\left(\prod_{k=0}^{j-1} R_{t+k}\right)c_t.$$

Substituting this into the left side of (1.31) gives $(1-\beta)^{-1}c_t$. Therefore (1.31) and (1.32) imply that

$$(1.33) \qquad c_t = (1-\beta)\left[y_t + \sum_{j=1}^{\infty}\left(\prod_{k=0}^{j-1} R_{t+k}\right) y_{t+j} + A_t\right],$$

so that the agent always consumes a constant fraction of his or her total human and nonhuman wealth. Equation (1.33) is valid for any sequences $\{R_t\}_{t=0}^{\infty}$, $\{y_t\}_{t=0}^{\infty}$ such that the right side converges.

To specialize (1.33) to the case in which $y_t = \lambda y_{t-1}$ and $R_t = R$, write out (1.33) as

$$c_t = (1-\beta)(A_t + y_t + R_t^{-1}y_{t+1} + R_t^{-1}R_{t+1}^{-1}y_{t+2} + \ldots).$$

Repeatedly substituting $R_{t+1}^{-1} = R^{-1}$ and $y_{t+1} = \lambda y_t$ into the above equation gives

$$c_t = (1-\beta)(A_t + y_t + R^{-1}\lambda y_t + R^{-2}\lambda^2 y_t + \ldots)$$

or

$$c_t = (1-\beta)\left[A_t + y_t\left(\frac{1}{1-\lambda R^{-1}}\right)\right],$$

where we require that $\lambda R^{-1} < 1$. That is, income is assumed to grow at a rate less than the interest rate. In the decision rule stated above consumption varies directly with current income y_t, inversely with the currently observed interest rate R, and directly with the rate of growth of income λ.

Optimal Growth

A consumer aims to maximize

$$\sum_{t=0}^{\infty} \beta^t u(c_t), \qquad 0 < \beta < 1$$

$$\text{subject to} \quad c_t + k_{t+1} = f(k_t), \qquad k_0 > 0 \text{ given}, \qquad c_t \geq 0,$$

where $u'(0) = +\infty$, $u' > 0$, $u'' < 0$, $f'(0) = +\infty$, $f'(\infty) = 0$, $f' > 0$, and $f'' < 0$. Here c_t is consumption and k_t is the stock of capital. This is a version of the problem that was studied by T. C. Koopmans (1963) and David Cass (1965).

Let the state be defined as k_t and the control as k_{t+1}. Bellman's equation is then

$$v(k_t) = \max_{k_{t+1}}\{u[f(k_t) - k_{t+1}] + \beta v(k_{t+1})\}.$$

The first-order condition is

$$-u'[f(k_t) - k_{t+1}] + \beta v'(k_{t+1}) = 0. \tag{1.34}$$

Benveniste and Scheinkman's equation (1.30) implies that $v(k_t)$ is differentiable with

$$v'(k_t) = u'[f(k_t) - k_{t+1}]f'(k_t), \tag{1.35}$$

where k_{t+1} is evaluated at the optimum $k_{t+1} = h(k_t)$.

Because $u(\cdot)$ and $f(\cdot)$ are strictly concave, it follows that $v(k)$ is strictly

concave. From this inference it follows that the optimal policy function, the solution $k_{t+1} = h(k_t)$ of (1.34), is a nondecreasing function of k_t.[2]

There is a maximum capital stock that can be sustained as a stationary equilibrium, namely that which would eventually emerge if c_t were to be zero for all t. If c_t were zero for all t, k_t would evolve according to the difference equation $k_{t+1} = f(k_t)$. Because $f'(0) = +\infty$, $f'' < 0$, and because $f'(\infty) = 0$, the equation $\bar{k} = f(\bar{k})$ has a unique positive solution. Evidently $k_{t+1} = f(k_t)$ converges to $\bar{k}$ as $t \to \infty$. [To verify this point, plot $f(k_t)$ against a 45° line.]

Let the system begin with $k_0 \in (0, \bar{k}]$. Then for $t \geq 1$, k_t must evidently remain in the bounded interval $[0, \bar{k}]$. Because the optimal policy function $h(k_t) = k_{t+1}$ is nondecreasing in k_t, it can be shown that $k_0, k_1, k_2, \ldots,$ is a monotone, bounded sequence. On the one hand, suppose that $k_1 > k_0$. Then because $h(\cdot)$ is nondecreasing, we have $k_2 = h(k_1) \geq h(k_0) = k_1$, $k_3 = h(k_2) \geq h(k_1) = k_2$, and so on. On the other hand, suppose that $k_1 < k_0$. Then $k_2 = h(k_1) \leq h(k_0) = k_1$, $k_3 = h(k_2) \leq h(k_1) = k_2$, and so on. It follows that k_t is a monotone, bounded sequence. Inasmuch as monotone, bounded sequences converge, it follows that k_t converges to a limit point $k_\infty(k_0)$ as $t \to \infty$.

The preceding convergence argument leaves open the possibility that the limit point $k_\infty(k_0)$ depends on the starting point k_0. It does not do so, however, as the following argument verifies. Let k_∞ be a limit point. At the limit point, (1.34) and (1.35) hold, and $k_{t+1} = k_t = k_\infty$. The implication is that $\beta f'(k_\infty) = 1$, an equation that determines a unique optimal stationary value k_∞. Note that the "gross rate of return" $f'(k_\infty) = \beta^{-1}$ in the stationary state and is independent of the specifics of the current-period utility function and the production function. Note also that the optimal stationary capital stock depends on $f(\cdot)$ and β but not on $u(\cdot)$.

We now specialize this example by following Brock and Mirman (1972) and considering the particular functional forms $u(c) = \ln c$ and $f(k) = Ak^\alpha$, $A > 0, 0 < \alpha < 1$. We will use the guess-and-verify method for this problem. The guess may not seem an obvious one. The inspiration for the guess can be

2. From (1.35), we have that $v'(k)$ is continuous. This follows from the continuity of $h(k)$, $f(k)$, and $f'(k)$. For two levels k_i of k, $i = 1, 2$, consider the first-order condition $u'[f(k_i) - h(k_i)] = \beta V'[h(k_i)]$. Assume that $k_1 \geq k_2$ and that $h(k_1) < h(k_2)$. By strict concavity of $v(\cdot)$ and continuity of $v'(\cdot)$, it follows that (1) for all $h(k_i)$, $v'[h(k_i)]$ is well defined, and (2) $h(k_1) < h(k_2)$ implies $v'[h(k_1)] > v'[h(k_2)]$. Therefore, $u'[f(k_1) - h(k_1)] > u'[f(k_2) - h(k_2)]$. By strict concavity of u, the preceding inequality holds if and only if $f(k_1) - h(k_1) < f(k_2) - h(k_2)$, or equivalently, $0 < h(k_2) - h(k_1) < f(k_2) - f(k_1) \leq 0$. This is a contradiction produced by the assumption that for $k_1 \geq k_2$, $h(k_1) < h(k_2)$. Therefore $h(k)$ is nondecreasing in k. (The argument in this note was constructed by Rodolfo Manuelli.)

understood by working Exercise 1.1 at the end of the chapter. For this example we make the guess

(1.36) $$v(k) = E + F \ln k,$$

where E and F are undetermined coefficients. For this guess, the first-order necessary condition (1.34) implies the following formula for the optimal policy $\tilde{k} = h(k)$, where $\tilde{k}$ is next period's value and k is this period's value of the capital stock:

(1.37) $$\tilde{k} = \frac{\beta F}{1 + \beta F} A k^{\alpha}.$$

Substituting (1.37) into the right side of (1.35) gives

(1.38) $$v'(k) = (1 + \beta F)\alpha k^{-1}.$$

Differentiating (1.36) gives

(1.39) $$v'(k) = F k^{-1}.$$

Equating (1.38) and (1.39) permits one to solve for F, $F = \alpha/(1 - \alpha\beta)$. Substituting this expression for F back into (1.36) and (1.37) gives

(1.40) $$v(k) = E + \frac{\alpha}{1 - \alpha\beta} \ln k$$
$$\tilde{k} = A\beta\alpha k^{\alpha}.$$

The fact that expressions (1.38) and (1.39) for $v'(k)$ have identical functional forms both verifies the original guess (1.36) and permits one to solve for the undetermined coefficient F. An alternative procedure for verifying the guess involves substituting (1.37) into Bellman's functional equation and equating the result to the right side of (1.36). Solving the resulting equation for E and F again gives $F = \alpha/(1 - \alpha\beta)$ and now gives

$$E = (1 - \beta)^{-1}\left[\ln A(1 - \alpha\beta) + \frac{\beta\alpha}{1 - \alpha\beta} \ln A\beta\alpha\right].$$

In Exercise 1.1, the reader is asked to construct the same solution (1.37) to the functional equation, using the method of iterating on Bellman's equation (1.26) starting from $v_0(k) = 0$. For this purpose it is useful to note that the term $F = \alpha/(1 - \alpha\beta)$ that appears in (1.40) can be interpreted as a geometric sum $\alpha[1 + \alpha\beta + (\alpha\beta)^2 + \ldots]$.

Equation (1.40) shows that the optimal policy is to have capital move according to the difference equation $k_{t+1} = A\beta\alpha k_t^{\alpha}$, or $\ln k_{t+1} = \ln A\beta\alpha + \alpha \ln k_t$. Because $\alpha < 1$, we know that k_t converges as $t \to \infty$ for any positive initial value k_0. The stationary point is given by the solution of $k_\infty = A\beta\alpha k_\infty^{\alpha}$, or $k_\infty^{\alpha - 1} = (A\beta\alpha)^{-1}$. Notice that this example obeys the general conclusion established above that k_∞ is determined from the solution of $\beta f'(k_\infty) = 1$.

1.5 The Optimal Linear Regulator Problem

We now consider a special class of dynamic programming problems in which the return functions r_t are quadratic and the transition functions g_t are linear. This specification leads to the widely used optimal linear regulator problem. We consider the special case in which the return functions r_t and transition functions g_t are both time invariant. The problem is to maximize over choice of $\{u_t\}_{t=0}^{\infty}$ the criterion

(1.41) $$\sum_{t=0}^{\infty} \{x_t' R x_t + u_t' Q u_t\},$$

subject to $x_{t+1} = Ax_t + Bu_t$, x_0 given. Here x_t is an $(n \times 1)$ vector of state variables, u_t is a $(k \times 1)$ vector of controls, R is a negative semidefinite symmetric matrix, Q is a negative definite symmetric matrix, A is an $(n \times n)$ matrix, and B is an $(n \times k)$ matrix. We guess that the value function is quadratic, $V(x) = x'Px$, where P is a negative semidefinite symmetric matrix.

Using the transition law to eliminate next period's state, Bellman's equation becomes

(1.42) $$x'Px = \max_u \{x'Rx + u'Qu + (Ax + Bu)'P(Ax + Bu)\}.$$

The first-order necessary condition for the maximum problem on the right side of (1.42) is

(1.43) $$(Q + B'PB)u = -B'PAx,$$

which implies the feedback rule for u:

(1.44) $$u = -(Q + B'PB)^{-1}B'PAx$$

or

(1.45) $$u = -Fx,$$

where $F = (Q + B'PB)^{-1}B'PA$. Substituting the optimizer (1.45) into the right side of (1.42) and rearranging gives

$$P = R + A'PA - A'PB(Q + B'PB)^{-1}B'PA. \tag{1.46}$$

Equation (1.46) is called the algebraic matrix Riccati equation.

Under particular conditions, Equation (1.46) has a unique negative semidefinite solution, which is approached in the limit as $j \to \infty$ by iterations on the matrix Riccati difference equation:[3]

$$P_{j+1} = R + A'P_jA - A'P_jB(Q + B'P_jB)^{-1}B'P_jA, \tag{1.47}$$

starting from $P_0 = 0$. Equation (1.47) is derived much like (1.46) except that one starts from the iterative version of Bellman's equation (1.26) rather than from the asymptotic version (1.28).

A modified version of problem (1.41) is the discounted optimal linear regulator problem, to maximize

$$\sum_{t=0}^{\infty} \beta^t (x_t'Rx_t + u_t'Qu_t), \qquad 0 < \beta < 1, \tag{1.48}$$

subject to $x_{t+1} = Ax_t + Bu_t$, x_0 given. For this problem Bellman's recursive equation (1.26) implies the following matrix Riccati difference equation modified for discounting:

$$P_{j+1} = R + \beta A'P_jA - \beta^2 A'P_jB(Q + \beta B'P_jB)^{-1}B'P_jA. \tag{1.49}$$

The algebraic matrix Riccati equation is modified correspondingly. The value function for the infinite horizon problem is simply $V(x_0) = x_0'Px_0$, where P is the limiting value of P_j resulting from iterations on (1.49) starting from $P_0 = 0$. The optimal policy is $u_t = -Fx_t$, where $F = \beta(Q + \beta B'PB)^{-1}B'PA$.

Upon substituting the optimal control $u_t = -Fx_t$ into the law of motion $x_{t+1} = Ax_t + Bu_t$, we obtain the optimal "closed-loop system" $x_{t+1} = (A - BF)x_t$. This difference equation governs the evolution of x_t under the optimal control. The system is said to be stable if $\lim_{t\to\infty} x_t = 0$ starting from any initial $x_0 \in R^n$. Assume that the eigenvalues of $(A - BF)$ are distinct, and use the eigenvalue decomposition $(A - BF) = C\Lambda C^{-1}$ where the columns of C are the eigenvectors of $(A - BF)$ and Λ is a diagonal matrix of eigenvalues of $(A - BF)$. Write the above equation as $x_{t+1} = C\Lambda C^{-1}x_t$. The solution of this difference equation for $t > 0$ is readily verified by repeated substitution to be $x_t = C\Lambda^t C^{-1}x_0$. Evidently, the system is stable for all $x_0 \in R^n$ if and only if the eigenvalue of $(A - BF)$ of maximum absolute value is strictly less than unity in absolute value. When this condition is met, $(A - BF)$ is said to be a "stable matrix."

3. If the eigenvalues of A are bounded in modulus below unity, this result obtains, but much weaker conditions also suffice. See Bertsekas (1976, chap. 4) and Sargent (1981).

A literature is devoted to characterizing the conditions on A, B, R, and Q under which the optimal closed-loop system matrix $(A - BF)$ is stable. These results are described in detail in Sargent (1981) and may be briefly described here for the undiscounted case $\beta = 1$. Heuristically, the conditions on A, B, R, and Q that are required for stability are as follows. First, A and B must be such that it is *possible* to pick a control law $u_t = -Fx_t$ that drives x_t to zero eventually, starting from any $x_0 \in R^n$ ["the pair (A, B) must be stabilizable"]. Second, the matrix R must be such that the controller *wants* to drive x_t to zero as $t \to \infty$. Notice from (1.41) that, if R is strictly negative definite, the controller will want to drive x_t to zero, because $x_t'Rx_t < 0$ for $x_t \neq 0$. If x_t does not approach zero, then the objective function is $-\infty$. When R is not strictly negative definite, however, the possibility emerges that the planner does not care whether some components of x_t fail to go to zero as $t \to \infty$. To attain stability of $(A - BF)$, it is necessary both for the planner to dislike it that some components of x_t threaten not to go to zero in the absence of countervailing control actions and for (A, B) to be such that the controller has the ability to drive those components to zero as $t \to \infty$ by an appropriate choice of F.

These conditions are discussed under the subjects of controllability, stabilizability, reconstructability, and detectability in the literature on linear optimal control. (For continuous-time linear systems, these concepts are described by Kwakernaak and Sivan 1972; for discrete-time systems, see Sargent 1981.) These conditions subsume and generalize the transversality conditions used in the discrete-time calculus of variations (see Sargent 1986). That is, the case when $(A - BF)$ is stable corresponds to the situation in which it is optimal to solve "stable roots backward and unstable roots forward." See Sargent (1986, chap. 9). Hansen and Sargent (1981) describe the relationship between Euler equation methods and dynamic programming for a class of linear optimal control systems. Also see Chow (1981).

The conditions under which $(A - BF)$ is stable are also the conditions under which x_t converges to a unique stationary distribution in the stochastic version of the linear regulator problem (see Section 1.8).

1.6 Stochastic Control Problems

We now consider a modification of problem (1.29) to permit uncertainty of particular kinds. We modify the transition equation and consider the problem, to maximize

$$E_0 \sum_{t=0}^{\infty} \beta^t r(x_t, u_t), \qquad 0 < \beta < 1, \tag{1.50}$$

subject to

$$x_{t+1} = g(x_t, u_t, \epsilon_{t+1}), \tag{1.51}$$

x_0 known and given at $t = 0$, where ϵ_t is a sequence of independently and identically distributed random variables with cumulative probability distribution function prob$\{\epsilon_t \leq e\} = F(e)$ for all t; $E_t(y)$ denotes the mathematical expectation of a random variable y, given information known at t. At time t, x_t is assumed to be known, but x_{t+j}, $j \geq 1$ is not known at t. That is, ϵ_{t+1} is realized at $(t+1)$, after u_t has been decided at t. In problem (1.50)–(1.51), uncertainty is injected by assuming that x_t follows a random difference equation.

Problem (1.50–1.51) continues to have a recursive structure, stemming jointly from the additive separability of the objective function (1.50) in pairs (x_t, u_t) and from the difference equation characterization of the transition law (1.51). In particular, controls dated t affect returns $r(x_s, u_s)$ for $s \geq t$ but not earlier. This feature implies that dynamic programming methods remain appropriate.

The problem is to maximize (1.50) subject to (1.51) by choice of a "policy" or "contingency plan" $u_t = h(x_t)$. The version of Bellman's functional equation corresponding to (1.28) becomes

$$V(x) = \max_u \{r(x, u) + \beta E[V[g(x, u, \epsilon)]|x]\}, \tag{1.52}$$

where $E\{V[g(x, u, \epsilon)]|x\} = \int V[g(x, u, \epsilon)]dF(\epsilon)$ and where $V(x)$ is the optimal value of the problem starting from x at $t = 0$. The solution $V(x)$ of (1.52) can be found by iterating on

$$V_{j+1}(x) = \max_u \{r(x, u) + \beta E[V_j[g(x, u, \epsilon)]|x]\}, \tag{1.53}$$

starting from any bounded continuous initial V_0. Under various particular regularity conditions, there obtain versions of the same four properties listed in Section 1.3. See Lucas, Prescott, and Stokey (forthcoming) or the framework presented in the Appendix.

The first-order necessary condition for the problem on the right side of (1.52) is

$$\frac{\partial r(x, u)}{\partial u} + \beta E\left[\frac{\partial g}{\partial u}(x, u, \epsilon)V'(g(x, u, \epsilon))|x\right] = 0,$$

which we obtained simply by differentiating the right side of (1.52), passing the differentiation operation under the E (an integration) operator. Off corners, the value function satisfies

$$V'(x) = \frac{\partial r}{\partial x}[x, h(x)] + \beta E\left\{\frac{\partial g}{\partial x}[x, h(x), \epsilon]V'(g[x, h(x), \epsilon])|x\right\}.$$

In the special case in which $\partial g/\partial x \equiv 0$, the formula for $V'(x)$ becomes

$$V'(x) = \frac{\partial r}{\partial x}[x, h(x)].$$

Substituting this formula into the first-order necessary condition for the problem gives the stochastic Euler equation

$$\frac{\partial r}{\partial u}(x, u) + \beta E\left[\frac{\partial g}{\partial u}(x, u, \epsilon)\frac{\partial r}{\partial x}(\tilde{x}, \tilde{u})|x\right] = 0,$$

where tildes over x and u denote next-period values.

1.7 Examples of Stochastic Control Problems

We now give several examples of stochastic dynamic programming problems.

Consumption with a Random Return

A consumer seeks to maximize

$$E_0 \sum_{t=0}^{\infty} \beta^t u(c_t), \qquad 0 < \beta < 1,$$

subject to $A_{t+1} = R_t(A_t - c_t)$, $t \geq 0$, A_0 given, where $u'(c) > 0$, $u''(c) < 0$, and where A_t is assets at the beginning of period t, c_t is consumption at t, and R_t is the gross rate of return on assets between periods t and $(t+1)$. We assume that R_t becomes known at the beginning of period $(t+1)$, after a decision about consumption at t, c_t, must be made. Assume that R_t is governed by a first-order Markov process, with transitions governed by prob$\{R_t \leq R'|R_{t-1} = R\} = F(R', R)$. When time t decisions must be made, the consumer knows A_t and $R_{t-1}, R_{t-2}, \ldots$. To rule out perpetual borrowing at the rate of return R_t, we impose the requirement that A_t must satisfy $\lim_{t\to\infty} E_0\beta^t A_t = 0$.

For this problem we define the state as (A_t, R_{t-1}) and the control $\hat{u}_t$ as $(A_t - c_t)$. The transition equation for A_t is then given by $A_{t+1} = R_t(A_t - c_t) = R_t\hat{u}_t$, whereas the transition equation for R is implicitly defined by $F(R', R)$. Let $v(A_t, R_{t-1})$ be the value of the problem for a consumer with initial assets A_t when the last observed rate of return is R_{t-1}. Then Bellman's functional equation is

$$v(A_t, R_{t-1}) = \max_{\hat{u}_t}\{u(A_t - \hat{u}_t) + \beta E_t v(\hat{u}_t R_t, R_t)\}.$$

The first-order necessary condition for the problem on the right is

$$-u'(c_t) + \beta E_t v_1(\hat{u}_t R_t, R_t) R_t = 0.$$

Applying the Benveniste-Scheinkman formula to evaluate $v_1(A_t, R_{t-1})$ gives

$$v_1(A_t, R_{t-1}) = u'(c_t).$$

Using this formula in the first-order necessary condition gives the Euler equation

(1.54) $$u'(c_t) = \beta E_t u'(c_{t+1}) R_t.$$

A solution of the agent's optimization problem is a saving policy function $u_t = h(A_t, R_{t-1})$, which implies a consumption policy function $c_t = c(A_t, R_{t-1}) \equiv A_t - h(A_t, R_{t-1})$. This policy function must satisfy the Euler equation (1.54) and must imply that the boundary condition on assets $\lim_{t\to\infty} E_0 \beta^t A_t = 0$ is satisfied. Substituting the function $c(A_t, R_{t-1})$ into the Euler equation and using the transition equation gives

(1.55) $$u'[c(A_t, R_{t-1})] = \beta E_t u'[c(R_t[A_t - c(A_t, R_{t-1})], R_t) R_t].$$

This is a functional equation in the optimal policy function $c(A_t, R_{t-1})$.

To take a specific example, let $u(c)$ equal lnc and let R_t be an independently and identically distributed random process such that $1 \le ER_t < 1/\beta^2$. We guess that the optimal policy takes the form $c_t = \gamma A_t$, where γ is a constant to be determined. Substituting this guess into (1.54) gives

$$\frac{1}{\gamma A_t} = \beta E \frac{R_t}{\gamma R_t(A_t - \gamma A_t)},$$

where E is now the unconditional expectation operator. Solving for γ gives $\gamma = 1 - \beta$. The optimal policy is of the form $c_t = (1-\beta)A_t$. It can be verified that this policy satisfies the boundary condition that we have imposed on asset accumulation. The optimal policy is to consume a constant fraction of wealth, $0 < \gamma < 1$, where $\gamma = 1 - \beta$.

Under the optimal policy, assets evolve according to $A_{t+1} = R_t(1-\gamma)A_t$, which implies that

$$A_t = (1-\gamma)^t \prod_{j=0}^{t-1} R_j A_0, \qquad t \ge 1.$$

Consequently, we have that

$$c_t = \gamma(1-\gamma)^t \prod_{j=0}^{t-1} R_j A_0, \qquad t \ge 1, \qquad c_0 = \gamma A_0.$$

The optimal value of $E_0 \sum_{t=0}^{\infty} \beta^t \ln c_t$ is then given by

$$\ln \gamma A_0 + E_0 \sum_{t=1}^{\infty} \beta^t \ln\left[\gamma(1-\gamma)^t \prod_{j=0}^{t-1} R_j A_0\right].$$

Because R_j is independently and identically distributed, evaluating this expression gives

$$v(A_0, R_{-1}) = \frac{1}{1-\beta}\ln\gamma + \ln(1-\gamma)\sum_{t=0}^{\infty}\beta^t t + \sum_{t=0}^{\infty}\beta^t t E \ln R + \frac{1}{1-\beta}\ln A_0,$$

where $E \ln R$ is the expectation of $\ln R_t$ for all t. Still, $\sum_{t=0}^{\infty} t\beta^t = \beta/(1-\beta)^2$ (see Sargent 1979, Eq. 21, p. 88). Therefore the value function can be written

$$v(A_0, R_{-1}) = \frac{1}{1-\beta}\ln\gamma + \ln(1-\gamma)\frac{\beta}{(1-\beta)^2} + \frac{\beta}{(1-\beta)^2}E\ln R + \frac{1}{1-\beta}\ln A_0.$$

The value depends directly on the mean of the logarithm of the rate of return but is independent of the realization of the rate of return at the beginning of the current period. This last property is special and depends on the assumption that R_t is distributed independently and identically over time.

Dynamic Portfolio Theory

This example generalizes the preceding one to the case in which a consumer can allocate his or her assets among a set of n assets, where the ith asset bears rate of return R_{it} at time t. Here R_{it} is assumed to be a positive random gross rate of return R_{it}... variable that is bounded from above with probability 1. The consumer maximizes $E_0 \sum_{t=0}^{\infty} \beta^t u(c_t)$ by choosing contingency plans for s_{it} for $i = 1, \ldots, n$ and $t \ge 0$, subject to

$$c_t + \sum_{i=1}^{n} s_{it} = A_t, \qquad t \ge 0$$

$$A_{t+1} = \sum_{i=1}^{n} s_{it} R_{it}, \qquad t \ge 0, \qquad A_0 \text{ given}$$

$$\lim_{t\to\infty} E_0 \beta^t A_t = 0.$$

Here s_{it} is the amount of asset i purchased in period t, and c_t is consumption at t. At time t, A_t and R_{it-1}, $i = 1, \ldots, n$ are observed, but R_{it}, $i = 1, \ldots, n$ is not observed until the beginning of period $(t+1)$. We assume that R_{it} is governed by a Markov process, with transition probabilities given by $\text{prob}\{R_t \le R' | R_{t-1} = R\} = F(R', R)$, where $R_t = (R_{1t}, \ldots, R_{nt})$ and R' and R are both n-dimensional vectors. Shortly we will specialize the setup to the case in which R_t and R_{t-1} are independently distributed for all t.

We define the state for this problem as (A_t, R_{t-1}), whereas the control is now the vector $(s_{1t}, \ldots, s_{nt})$. Bellman's functional equation is

$$v(A_t, R_{t-1}) = \max_{s_{1t},\ldots,s_{nt}} \left\{ u(A_t - \sum_{i=1}^n s_{it}) + \beta E_t v\left(\sum_{i=1}^n s_{it} R_{it}, R_t\right)\right\}.$$

The first-order necessary conditions for the problem on the right are

$$u'\left(A_t - \sum_{i=1}^n s_{it}\right) = \beta E_t R_{it} v_1\left(\sum_{k=1}^n s_{kt} R_{kt}, R_t\right), \qquad i = 1, \ldots, n.$$

The Benveniste-Scheinkman formula implies that $v_1 = u'(A_t - \Sigma_{i=1}^n s_{it})$. Substituting this equation into the above first-order conditions gives

$$u'\left(A_t - \sum_{i=1}^n s_{it}\right) = \beta E_t R_{it} u'\left(\sum_{k=1}^n R_{kt} s_{kt} - \sum_{j=1}^n s_{jt+1}\right),$$
$$i = 1, \ldots, n.$$

We now want to solve for optimal policy functions $s_{it} = s_i(A_t, R_{t-1})$. Substituting the policy functions into the preceding Euler equation gives

(1.56)
$$u'\left[A_t - \sum_{i=1}^n s_i(A_t, R_{t-1})\right]$$
$$= \beta E_t R_{it} u' \left\{\sum_{k=1}^n R_{kt} s_k(A_t, R_{t-1}) - \sum_{j=1}^n s_j\left[\sum_{k=1}^n R_{kt} s_k(A_t, R_{t-1}), R_t\right]\right\},$$
$$i = 1, \ldots, n.$$

This is a set of n functional equations in the n unknown functions $s_i(A_t, R_{t-1})$, $i = 1, \ldots, n$.

We now consider the special case in which R_{it}, $i = 1, \ldots, n$, is distributed independently and identically both over time *and* across i. Furthermore, we suppose that $u(c) = [1/(1-\alpha)]c^{1-\alpha}$, where $0 < \alpha < 1$, so that $u'(c) = c^{-\alpha}$. For this case we make the guess that $s_{it} = kA_t$, $i = 1, \ldots, n$, where k is a constant to be determined. Note that we are guessing that k is independent of i, a guess inspired by the independence and identity of the distribution of the R_{it} over time and across i. Substituting this guess into (1.55), using $u'(c) = c^{-\alpha}$ and rearranging, gives

$$k^\alpha = \beta E_t \frac{R_{it}}{\left(\sum_{j=1}^n R_{jt}\right)^\alpha}, \qquad i = 1, \ldots, n.$$

Because R_{it} is independently and identically distributed, the above equation can hold for all $i = 1, \ldots, n$. This result verifies our guess and gives an equation that can be solved for k. (Notice how this example conforms to the preceding one.)

In the present example the household allocates the same constant fraction of wealth to each asset in each period. This result depends on the independence of the R_{it} over time and the independence and identity of the distribution across assets i. We now explore the implications of relaxing the assumption of identical distributions across i while retaining independence across time and assets. Under these new assumptions we guess that the optimal policies will be of the form $s_{it} = k_i A_t$. Substituting this guess into the Euler equation gives

$$u'\left[A_t\left(1 - \sum_{i=1}^n k_i\right)\right] = \beta E_t R_{it} u'\left[\sum_{h=1}^n k_h R_{ht} A_t\left(1 - \sum_{j=1}^n k_j\right)\right],$$
$$i = 1, \ldots, n.$$

Further specializing this example, we take $u(c) = \ln c$. Substituting $u'(c) = c^{-1}$ into the above equations and rearranging gives

$$1 = \beta E_t \frac{R_{it}}{\sum_{j=1}^n k_j R_{jt}}, \qquad i = 1, \ldots, n.$$

This is a system of n equations in the n unknowns $k_1, \ldots, k_n$. For example, when $n = 2$, we have the two equations

$$1 = \beta E_t[k_1 + (k_2 R_{2t}/R_{1t})]^{-1}$$
$$1 = \beta E_t[(k_1 R_{1t}/R_{2t}) + k_2]^{-1},$$

which are to be solved for k_1 and k_2.

Stochastic Optimal Growth

We consider the stochastic growth example of Brock and Mirman (1972), to maximize

$$E_0 \sum_{t=0}^\infty \beta^t \ln c_t, \qquad 0 < \beta < 1,$$

subject to $c_t + k_{t+1} = Ak_t^\alpha\theta_t$, $0 < \alpha < 1$, where $\ln \theta_t$ is an independently and identically distributed random variable with normal distribution with mean zero and variance σ^2. (The stochastic optimal growth model is a workhorse in the literatures on capital asset pricing and business fluctuations; see Brock 1982 and Kydland and Prescott 1982.) A planner is supposed to know (k_t, θ_t) at time t but not to know future values of θ. We define the state of the system as (k_t, θ_t). Bellman's equation becomes

(1.57) $$V(k_t, \theta_t) = \max_{k_{t+1}}\{\ln(Ak_t^\alpha\theta_t - k_{t+1}) + \beta E[V(k_{t+1}, \theta_{t+1})|k_t, \theta_t]\}.$$

The reader is invited to verify the guess that the solution of (1.56) is of the form $V(k, \theta) = E + F \ln k + G \ln \theta$, where E, F, and G are undetermined coefficients, and that the optimal policy rule is $k_{t+1} = A\alpha\beta k_t^\alpha\theta_t$.

1.8 The Stochastic Linear Optimal Regulator Problem

We consider the discounted stochastic linear optimal regulator problem, to maximize

(1.58) $$E_0 \sum_{t=0}^{\infty} \beta^t\{x_t'Rx_t + u_t'Qu_t\}, \qquad 0 < \beta < 1,$$

subject to x_0 given, and the law of motion

(1.59) $$x_{t+1} = Ax_t + Bu_t + \epsilon_{t+1}, \qquad t \geq 0,$$

where ϵ_{t+1} is an $(n \times 1)$ vector of random variables that is independently and identically distributed through time and obeys the normal distribution with mean vector zero and contemporaneous covariance matrix

(1.60) $$E\epsilon_t\epsilon_t' = \Sigma$$

(See Kwakernaak and Sivan 1972 for an extensive study of the continuous-time version of this problem; also see Chow 1981.) The matrixes R, Q, A, and B obey the assumption described in Section 1.5 above.

For this problem the value function turns out to be

(1.61) $$v(x) = x'Px + d,$$

where P is the unique negative semidefinite solution of the discounted algebraic matrix Riccati equation corresponding to (1.49), which is the limit of iterations on (1.49) starting from $P_0 = 0$, and where d is given by

(1.62) $$d = \beta(1-\beta)^{-1}\mathrm{tr}P\,\Sigma$$

where "tr" denotes the trace of a matrix. Furthermore, the optimal policy

continues to be given by $u_t = -Fx_t$, where

(1.63) $$F = \beta(Q + \beta B'P'B)^{-1}B'PA.$$

A notable feature of this solution is that the feedback rule (1.63) is identical with the rule for the corresponding nonstochastic linear optimal regulator problem.

To prove the preceding assertions, we substitute the guess (1.61) into Bellman's equation to obtain

$$v(x) = \max_u\{x^TRx + u^TQu + \beta E[(Ax + Bu + \epsilon)^T \cdot P(Ax + Bu + \epsilon)] + \beta d\},$$

where ϵ is the realization of ϵ_{t+1} when $x_t = x$ and where $E\epsilon|x = 0$. (Both prime ['] and superscript T denote transposition.) The above equation implies

$$\begin{aligned} v(x) = \max_u\{&x^TRx + u^TQu + \beta E\{x^TA^TPAx + x^TA^TPBu \\ &+ x^TA^TP\epsilon + u^TB^TPAx + u^TB^TPBu + u^TB^TP\epsilon \\ &+ \epsilon^TPAx + \epsilon^TPBu + \epsilon^TP\epsilon\} + \beta d\}. \end{aligned}$$

Evaluating the expectations inside the braces and using $E\epsilon|x = 0$ gives

$$\begin{aligned} v(x) = \max\{&x^TRx + u^TQu + \beta 2x^TA^TPBu + \beta x'A'PAx + \\ &+\beta u^TB^TPBu + \beta E\epsilon^TP\epsilon\} + \beta d. \end{aligned}$$

The first-order condition for u is

$$(Q + \beta B^TPB)u = -\beta B^TPAx,$$

which implies (1.63). Using $E\epsilon^TP\epsilon = \mathrm{tr}E\epsilon^TP\epsilon = \mathrm{tr}PE\epsilon\epsilon^T = \mathrm{tr}P\Sigma$, substituting (1.63) into the preceding expression for $v(x)$, and using (1.61) gives

$$P = R + \beta A'PA - \beta^2A'PB(Q + \beta B'PB)^{-1}B'PA,$$

and $$d = \beta(1-\beta)^{-1}\,\mathrm{tr}P\,\Sigma.$$

This step concludes the demonstration of the claims about the optimal value function and the optimal decision rule.

It is a remarkable feature of this solution that, although through d the objective function (1.60) depends on Σ, the covariance matrix of the "noises" ϵ, the optimal decision rule $u_t = -Fx_t$ is independent of Σ. This is the message of (1.63) and the discounted algebraic Riccati equation for P, which are identical with the formulas derived earlier under certainty. In other words, when expressed in the feedback form $u_t = h(x_t)$, the optimal policy function that solves this problem is independent of the noise statistics of the problem. This feature is called the "certainty equivalence principle" by economists. This is a special property of the optimal linear regulator problem and is due to the quadratic nature of the objective function and the linear nature of the transition equation. Certainty equivalence does not characterize stochastic control problems generally.

For the stochastic optimal linear regulator, substituting the optimal control $u_t = -Fx_t$ into the transition equation gives the stochastic optimal closed-loop system $x_{t+1} = (A - BF)x_t + \epsilon_{t+1}$. Under the condition that $(A - BF)$ is a stable matrix (that is, one whose eigenvalue of maximum absolute value is less than unity in absolute value), the system converges as $t \to \infty$ to a unique stationary probability distribution. The spectral density of the stationary distribution is given by

$$S_x(\omega) = [I - (A - BF)e^{-i\omega}]^{-1}\, \Sigma\, [I - (A - BF)^T e^{+i\omega}]^{-1},$$
$$\omega \epsilon [-\pi, \pi].$$

Here $S_x(\omega)$ is the Fourier transform of the covariogram of x_t,

$$S_x(\omega) = \sum_{\tau=-\infty}^{\infty} C_x(\tau) e^{-i\omega\tau},$$

where $C_x(\tau) = Ex_t x_{t-\tau}$. The covariances $C_x(\tau)$ can be recovered from $S_x(\omega)$ by the inversion formula

$$C_x(\tau) = (1/2\pi) \int_{-\pi}^{\pi} S_x(\omega) e^{+i\omega\tau}\, d\omega.$$

Spectral densities for continuous-time systems are discussed by Kwakernaak and Sivan (1982). For an elementary discussion of discrete-time systems, see Sargent (1986). Also see Sargent (1986, chap. 11) for definitions of the spectral density function and methods of evaluating the above integral.

The preceding discussion shows how the stochastic optimal linear regulator provides a complete description of the theoretical second moments of the stationary distribution of the controlled process x_t. The mapping from (A, B, R, Q) to these theoretical moments that is implicitly described by the above equations is the foundation of econometric methods designed to estimate a wide class of linear rational expectations models (see Hansen and Sargent 1980, 1981). Briefly, these methods use the following procedures for matching observations with theory. A sample of observations for some elements of x_t, $t = 1, \ldots, T$, is assumed to be available. All possible sample second moments of the observations are calculated. How well the theory matches the observations is measured by choosing a metric that gives the distance

between the sample moments and the theoretical moments associated with a given (A, B, R, Q). The metric is chosen not arbitrarily but in order to deliver good statistical properties of the estimates of (A, B, Q, R), consistency and asymptotic efficiency. Then A, B, Q, and R are estimated by choosing them to minimize the metric. For discussions of a good metric, see Hansen and Sargent (1980, 1982). The theory is "tested" by measuring how far the observations deviate from the theory, A, B, Q, and R being set at their best values.

As a simple example of a stochastic linear regulator problem, consider a monopolist who seeks to maximize

$$E_0 \sum_{t=0}^{\infty} \beta^t [P_t Y_t - J_t K_t - (d/2)(K_{t+1} - K_t)^2], \qquad 0 < \beta < 1$$

$$\text{subject to} \quad \begin{aligned} P_t &= A_0 - A_1 Y_t + \theta_t, && A_0, A_1 > 0 \\ Y_t &= f K_t && f > 0 \\ J_{t+1} &= \lambda J_t + \epsilon_{Jt}, && |\lambda| < 1/\sqrt{\beta} \\ \theta_{t+1} &= \mu \theta_t + \epsilon_{\theta t}, && |\mu| < 1/\sqrt{\beta}, \end{aligned}$$

K_0 given, J_t, θ_t, and K_t known at time t. Here P_t is output price, Y_t is output, J_t is the rental rate on capital, and θ_t is a random shock to demand, whereas ϵ_{Jt} and $\epsilon_{\theta t}$ are white noises. The maximization is over a stochastic process for K_{t+1} as a linear function of (K_t, J_t, θ_t).

To map this problem into the stochastic linear regulator problem, we define the state x_t as the vector $(K_t, J_t, u_t, 1)'$, whereas the control u_t is simply $K_{t+1} - K_t$. Then take A, B, Q, and R to be

$$A = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & \lambda & 0 & 0 \\ 0 & 0 & \mu & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \qquad B = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

$$R = \begin{bmatrix} -A_1 f^2 & -1/2 & f/2 & A_0 f/2 \\ -1/2 & 0 & 0 & 0 \\ f/2 & 0 & 0 & 0 \\ A_0 f/2 & 0 & 0 & 0 \end{bmatrix}, \quad Q = -d/2, \quad \epsilon_t = \begin{bmatrix} 0 \\ e_{Jt} \\ \epsilon_{\theta t} \\ 0 \end{bmatrix}$$

The optimal feedback law is an investment schedule of the form $u_t = -Fx_t$ or $(K_{t+1} - K_t) = -F(K_t, J_t, \theta_t, 1)'$.

Problems of the kind exhibited in this example can be formulated as stochastic discrete-time calculus-of-variations problems and can be solved as linear difference equations (see Sargent 1986). The methods involve two steps: (1) factoring the characteristic polynomial associated with the nonstochastic version of the problem in order to obtain the feedback and feedforward parts of the solution; and (2) utilizing the Wiener-Kolmogorov linear least-squares forecasting formula in order to express the feedforward part in terms of information available at the decision date. Notice that the linear regulator problem in effect accomplishes both optimization and prediction —and does so simultaneously via iterations on the matrix Riccati difference equation.

In Exercises 1.6 and 1.7, the reader is asked to take two problems with very large state spaces and to map them into linear regulator problems. These exercises are designed to show the chief advantage of the linear regulator framework: the tractability it retains even in the face of state spaces of very large dimension.

1.9 Dynamic Programming and Lucas's Critique

Recall the following version of the time-invariant, discounted stochastic control problem treated in Section 1.6, namely, to choose a strategy for u_t that maximizes

$$E_0 \sum_{t=0}^{\infty} \beta^t r(x_t, u_t), \qquad 0 < \beta < 1,$$

subject to $x_{t+1} = g(x_t, u_t, \epsilon_{t+1})$, $t \geq 0$, where ϵ_{t+1} is a sequence of independently and identically distributed random variables. We have seen that the solution is a time-invariant policy function $u_t = h(x_t)$ that satisfies the functional equation

$$v(x_t) = r[x_t, h(x_t)] + \beta E(v(g[x_t, h(x_t), \epsilon_{t+1}])|x_t). \tag{1.64}$$

In general, the optimal policy $h(x_t)$ that solves the functional equation (1.64) depends on the return function $r(x_t, u_t)$, on the transition function $g(x_t, u_t, \epsilon_{t+1})$, on the probability distribution of ϵ_{t+1}, and on β. In particular, even with preferences [that is, β and $r(x_t, u_t)$] fixed, the optimal decision rule $u_t = h(x_t)$ depends on the law of motion $g(x_t, u_t, \epsilon_{t+1})$. The implication is that, in dynamic decision problems, it is in general impossible to find a single decision rule $h(x_t)$ that will be invariant with respect to variations in the laws of motion $g(x_t, u_t, \epsilon_{t+1})$. This principle is illustrated in a variety of ways in the examples that we have considered above.

Robert E. Lucas (1976) criticized a range of econometric policy evaluation procedures because they used models that assumed private agents' decision rules to be invariant with respect to the laws of motion that they faced. Those models took as structural (that is, as invariant under interventions) such

private agents' decision rules as consumption functions, investment schedules, portfolio balance schedules, and labor supply schedules. The models were then routinely subjected to hypothetical policy experiments that changed the stochastic processes (or laws of motion, g) of such variables as income flows, tax rates, wage rates, prices, and interest rates, variables that entered private agents' constraints and decision rules. Those hypothetical experiments violated the principle that an optimal decision rule $h(x_t)$ is a function of the law of motion $g(x_t, u_t, \epsilon_{t+1})$.

Lucas's criticism was a particularly telling one because it embraced two of the fundamental ideas that underlay the enterprise of building large-scale Keynesian econometric models. First, there was the idea that for policy experiments it was important to isolate relationships that were structural, that is, invariant with respect to the class of interventions to be studied. Lucas observed that dynamic optimization theory ultimately implied that the key structural equations in Keynesian macroeconometric models, such as the consumption, investment, and portfolio balance schedule, should not be regarded as structural. Second, there was the idea that it was useful to derive the private agents' decision rules from the hypothesis of optimizing behavior in a dynamic context, an idea reflected with increasing sophistication by many works within the Keynesian tradition.

1.10 Dynamic Games and the Time Inconsistency Phenomenon

We now briefly describe the structure of a two-player dynamic, or "differential," game with no randomness. (A more general treatment of the issues described here is to be found in Hansen, Epple, and Roberds 1985; also see Basar and Olsder 1982.) The transition equation of the system is assumed to be

$$x_{t+1} = g(u_{1t}, u_{2t}), \tag{1.65}$$

where x_t is again the state vector, u_{1t} is the control vector of the first player, and u_{2t} is the control vector of the second player. Player i has an objective described by

$$\sum_{t=0}^{T} \beta^t r_i(x_t, u_{1t}, u_{2t}) + \beta^{T+1} V_{0i}(x_{T+1}), \qquad 0 < \beta < 1, \qquad i = 1, 2, \tag{1.66}$$

where $r_i(x_t, u_{1t}, u_{2t})$ is the return function of the ith player. Player i is assumed to maximize (1.66) subject to (1.65) with x_0 given and also subject to some particular assumption about player j's choice of $\{u_{jt}\}$ or about player j's way of choosing it. Alternative particular assumptions about how player i

imagines player j to choose u_{jt} determine the equilibrium concept of the game.

We first describe a particular version of a Nash equilibrium. Assume that player 1 takes player 2's actions $\{u_{2s}\}_{s=0}^{T}$ as given and as beyond player 1's control and that player 2 takes the symmetrical view with respect to player 1's actions. Then player 1's actions solve the following version of the system (1.8) of Euler equations and transition equations

$$\frac{\partial r_1}{\partial u_{1t}}(x_t, u_{1t}, u_{2t}) + \beta \frac{\partial g}{\partial u_{1t}}(u_{1t}, u_{2t}) \frac{\partial r_1}{\partial x_{t+1}}(x_{t+1}, u_{1t+1}, u_{2t+1}) = 0, \quad t = 0, 1, \ldots, T-1 \tag{1.67a}$$

$$x_{t+1} = g(u_{1t}, u_{2t}), \qquad t = 0, \ldots, T, \tag{1.67b}$$

subject to x_0 given and $\{u_{2s}\}_{s=0}^{T}$ known, and the terminal condition

$$\beta^T \frac{\partial r_1}{\partial u_{1T}}(x_T, u_{1T}, u_{2T}) + \beta^{T+1} \frac{\partial g}{\partial u_{1T}}(u_{1T}, u_{2T}) \frac{\partial V_{01}}{\partial x_{T+1}}(u_{T+1}) = 0. \tag{1.67c}$$

Analogously, player 2's actions solve

$$\frac{\partial r_2}{\partial u_{2t}}(x_t, u_{1t}, u_{2t}) + \beta \frac{\partial g}{\partial u_{2t}}(u_{1t}, u_{2t}) \frac{\partial r_2}{\partial x_{t+1}}(x_{t+1}, u_{1t+1}, u_{2t+1}) = 0, \quad t = 0, \ldots, T-1 \tag{1.68a}$$

$$x_{t+1} = g(u_{1t}, u_{2t}), \qquad t = 0, 1, \ldots, T, \tag{1.68b}$$

subject to x_0 given and $\{u_{1s}\}_{s=0}^{T}$ known, and the terminal condition

$$\beta^T \frac{\partial r_2}{\partial u_{2T}}(x_T, u_{1T}, u_{2T}) + \beta^{T+1} \frac{\partial g}{\partial u_{2T}}(u_{1T}, u_{2T}) \frac{\partial V_{02}}{\partial x_{T+1}}(x_{T+1}) = 0. \tag{1.68c}$$

In a Nash equilibrium, Equations (1.67a), (1.67b), (1.67c), (1.68a), and (1.68c) are solved jointly for $\{x_{t+1}, u_{1t}, u_{2t}\}$, $t = 0, 1, \ldots, T$.

We note two features of this equilibrium concept. First, each player's problem continues to be a recursive one, because under the assumption that the other player's actions are given, u_{it} affects returns $r_i(x_s, u_{1s}, u_{2s})$ dated t and later but not earlier. Thus each player's problem satisfies Bellman's principle of optimality. Second, we note that the entire system formed by (1.67a), (1.67b), (1.68a), and the terminal conditions is itself block recursive and can be solved by working backward, starting from date T. Thus the Nash equilibrium itself can be computed by recursive methods.

We now turn to a particular dominant-player game. We continue to assume that player 1 regards player 2's actions as given and beyond player 1's control. Player 2, however, is now assumed to understand that his controls influence agent 1's simply by virtue of the fact that agent 1's actions solve the Euler equation (1.67a), taking u_{2s} as given. Let us represent (1.67a), (1.67c) in the implicit form

$$\phi(x_t, x_{t+1}, u_{1t}, u_{1t+1}, u_{2t}, u_{2t+1}) = 0, \qquad t = 0, 1, \ldots, T-1 \tag{1.69}$$
$$\frac{\partial r_1}{\partial u_{1T}}(x_T, u_{1T}, u_{2T}) + \beta \frac{\partial g}{\partial u_{1T}}(u_{1T}, u_{2T}) \frac{\partial V_{01}}{\partial x_{T+1}}(x_{T+1}) = 0.$$

The decisions u_{1s} of the follower agent 1 are determined by (1.69) and by the relevant terminal condition, given x_0 and the actions u_{2s} of the leader. The leader is imagined to choose $\{u_{2s}, u_{1s}, s \geq 0\}$ to maximize (1.66) with $i = 2$, subject to both (1.65) and (1.69). We can represent the leader's problem as being to choose $u_{20}, \ldots, u_{2T}, u_{10}, \ldots, u_{1T}$, to maximize the Lagrangian

$$\begin{aligned} L = &\sum_{t=0}^{T} \beta^t r_2(x_t, u_{1t}, u_{2t}) + \beta^{T+1} V_{02}(x_{T+1}) \\ &+ \sum_{t=0}^{T} \beta^t \lambda_t' [g(u_{1t}, u_{2t}) - x_{t+1}] \\ &+ \sum_{t=0}^{T-1} \beta^t \mu_t' [\phi(x_t, x_{t+1}, u_{1t}, u_{1t+1}, u_{2t}, u_{2t+1})] \\ &+ \theta' \left[\frac{\partial r_1}{\partial u_{1T}}(x_T, u_{1T}, u_{2T}) + \beta \frac{\partial g}{\partial u_{1T}}(u_{1T}, u_{2T}) \frac{\partial V_{01}}{\partial x_{T+1}}(x_{T+1}) \right], \end{aligned} \tag{1.70}$$

where λ_t, $t = 0, \ldots, T$; μ_t, $t = 0, \ldots, T-1$; and θ are each vectors of Lagrange multipliers. The maximization is performed with x_0 taken as given. The first-order necessary condition for the maximization of (1.70) with respect to u_{2t} is

$$\begin{aligned} &\beta \frac{\partial r_2(x_t, u_{1t}, u_{2t})}{\partial u_{2t}} + \beta \frac{\partial g(u_{1t}, u_{2t})}{\partial u_{2t}} \lambda_t \\ &+ \beta \frac{\partial \phi}{\partial u_{2t}}(x_t, x_{t+1}, u_{1t}, u_{1t+1}, u_{2t}, u_{2t+1}) \mu_t \\ &+ \frac{\partial \phi}{\partial u_{2t}}(x_{t-1}, x_t, u_{1t-1}, u_{1t}, u_{2t-1}, u_{2t}) \mu_{t-1} = 0, \\ &t = 0, 1, \ldots, T-1. \end{aligned} \tag{1.71}$$

The reader is invited to obtain the remainder of the first-order necessary conditions and to analyze the resulting system of difference equations for determining $\{x_t, u_{1t}, u_{2t}\}$ via the solution of the dominant player's maximum problem. In addition, the reader is asked to verify that this system of equations is simultaneous and not block recursive (make a table for the system analogous to Table 1.1). This feature of the system and the nature of its cause can readily be seen by comparing (1.71) with (1.68a). In (1.71) the effects of u_{2t} for $t \geq 1$ on values of u_{1s} for $s < t$ and on *past* values of x_s and $r_2(x_s, u_{1s}, u_{2s})$ for $s < t$ are taken into account. Because the follower agent 1's actions at $s < t$ depend on u_{2t}, the leader's problem fails to be recursive. Accordingly, Bellman's principle of optimality may fail to characterize the leader's problem. The failure of the dominant player's problem to satisfy the principle of optimality is often called the time inconsistency of optimal plans.[4] Consequently the optimal plan is not generally self-enforcing in the sense described in Section 1.3 above.

Dynamic games occur in a variety of contexts in dynamic macroeconomics, industrial organization, and public finance. Early examples in macroeconomics were given by Kydland and Prescott (1977) and by Calvo (1978).

The following is a version of Calvo's (1978) example of a system in which private agents' responses to the government tax strategy confront the government with a nonrecursive problem in choosing a tax strategy. This example departs in some details from the preceding framework but exhibits the same essential nonrecursivity of the dominant player's (the government's) problem. The economy is one in which a representative private agent chooses c_t and m_{t+1} sequences to maximize

$$\sum_{t=0}^{\infty} \beta^t u(c_t, m_{t+1}/p_t), \qquad 0 < \beta < 1, \tag{1.72}$$

subject to

$$c_t + \tau_t + m_{t+1}/p_t = y(\tau_t) + m_t/p_t, \qquad m_0 > 0 \text{ given}, \tag{1.73}$$

where $u(c_t, m_{t+1}/p_t) = \ln c_t + \gamma \ln(m_{t+1}/p_t), \gamma > 0$. Here c_t is consumption of a single nonstorable good at time t, m_{t+1} is currency carried over from time t to $(t+1)$, and p_t is the price level at time t. The government imposes a distorting tax (or subsidy) of τ_t at time t. Following Calvo, we represent the distortion by simply positing that output at t is given by a diminishing function of τ_t, $y(\tau_t)$ where $y' < 0$. The private agent maximizes (1.72) by choosing sequences c_t, m_{t+1}, where $t \geq 0$, taking as given the sequences τ_t, p_t.

4. For a solution to fail to satisfy the principle of optimality, it seems to be sufficient that there exist no way of reformulating the problem (say, by redefining variables) so that it becomes recursive. Some problems appear not to be recursive when written in one way but can be transformed into equivalent recursive ones.

The government seeks to maximize the utility of the representative agent (1.72), subject to the constraints

$$c_t + g_t = y(\tau_t) \tag{1.74}$$

$$g_t = \tau_t + (m_{t+1} - m_t)/p_t, \tag{1.75}$$

where $g_t \geq 0$, $t \geq 0$, is an exogenously given sequence of government purchases. Equation (1.74) is the economy's resource constraint, whereas (1.75) is the government's budget constraint. The government takes $\{g_t, t \geq 0\}$ as given and chooses sequences of $\{m_{t+1}, \tau_t, t \geq 0\}$ to maximize (1.72). In performing this maximization, the government is assumed to take as given that the economy is in equilibrium and that private agents are solving their optimum problem. Technically, this is an example of a "team" dynamic game, because the government and the private agent share the same objective function. Private agents are assumed to regard $\{p_t, \tau_t\}_{t=0}^{\infty}$ as given sequences.

An "equilibrium" is defined as a collection of sequences for $(m_{t+1}, \tau_t, p_t, c_t)$ that solve the optimum problems of both the private agent and the government.

It can be verified that a private agent's problem is a recursive one. If we let $y(\tau) + m/p$ be the state at t, Bellman's equation for this problem can be expressed as

$$v[y(\tau) + m/p] = \max_{c,m'}\{u(c, m'/p) + \beta v[y(\tau') + m'/p']\},$$

$$c + \tau + m'/p \leq y(\tau) + m/p.$$

By our usual methods the Euler equation for this problem can be rearranged to imply the difference equation

$$\frac{1}{c_t p_t} = \beta \frac{1}{c_{t+1} p_{t+1}} + \frac{\gamma}{m_{t+1}}. \tag{1.76}$$

We now use (1.76) and the equilibrium condition $[c_t = y(\tau_t) - g_t]$ to solve for p_t as a function of the (m_{t+1}, τ_t) process chosen by the government. Substituting the equilibrium condition into (1.76), regarding (1.76) as determining p_t as a function of the m_{t+j} sequence, and solving the difference equation (1.76) forward produce

$$1/p_t = \gamma[y(\tau_t) - g_t] \sum_{j=0}^{\infty} \beta^j \frac{1}{m_{t+j+1}}. \tag{1.77}$$

Equation (1.77) gives the equilibrium price level as a function of the sequences (τ_t, m_{t+1}) chosen by the government to finance the exogenously given g_t expenditure sequence. Equation (1.77) embodies the results of the Cagan effect (see Cagan 1956) by means of which expectations of future settings of the currency stock influence the current price level.

We now consider the government's problem, which is to maximize (1.72) subject to (1.74), (1.75), and (1.77) by choosing sequences for (τ_t, m_{t+1}). In maximizing subject to (1.77), the government is taking into account the fact that private agents behave in a way that makes the current price level a function of future m_{t+j}. If we substitute (1.74) and (1.77) into (1.72), the government's problem can be formulated as being the maximization of

$$\sum_{t=0}^{\infty} \beta^t u\left(y(\tau_t) - g_t,\ \gamma m_{t+1}[y(\tau_t) - g_t] \sum_{j=0}^{\infty} \beta^j \frac{1}{m_{t+j+1}}\right), \tag{1.78}$$

subject to

$$g_t = \tau_t + (m_{t+1} - m_t)/p_t, \tag{1.79}$$

by choosing sequences for τ_t and m_{t+1}, $t \geq 0$. This problem is evidently not a recursive one, because future values of the control m_{t+j+1} influence the government's return $u(c_t, m_{t+1}/p_t)$ at date t. The reason is that private agents respond to the government's choice of an m_{t+1+j} sequence by making the current price level respond to future settings of m_{t+1}.[5]

A consequence of the failure of the government's problem to be recursive is that the government's problem cannot be solved sequentially using the method of dynamic programming. As a result, the government's optimal plan will lack the self-enforcing character of dynamic programming solutions described above. In particular, if the government were imagined to reopen its planning process and to consider optimizing the "remainder" of (1.72) namely $\Sigma_{t=s}^{\infty} \beta^t u(c_t, m_{t+1}/p_t)$, $s > 0$, by choosing $\{m_{t+1}, \tau_t \text{ for } t \geq s\}$ subject to (1.74), (1.75), and (1.77), then in general the government would want to depart from the original plan for $(m_{t+1}, \tau_t, t \geq 0)$. This incentive would emerge because the government would want to neglect the effect of m_{t+1} for $t > s$ on p_v for $0 \leq v < s$ in setting m_{t+1} for $t > s$, an effect that was taken into account in the original plan for $\{m_{t+1}, \tau_t, t \geq 0\}$. This lack of a self-enforcement incentive is known as the time inconsistency problem.[6]

5. Hansen, Epple, and Roberds (1985) explicitly calculate a solution of the dominant player's problem for a class of linear-quadratic games. Their method is applied to a simple optimal taxation example in Sargent (1986). In these setups dynamic inconsistency of the optimal policy for the dominant player is evident from the time-varying form of its decision rule.

6. Lucas and Stokey (1983) study two versions of an optimal taxation problem. In their model without currency, they show that there exists a plan for restructuring the government debt each period that, if followed, renders the optimal tax plan time consistent. This finding can be regarded as providing a decentralization scheme between a tax authority and a debt-management authority that is capable of supporting an optimal plan in a self-enforcing way.

The time inconsistency phenomenon that is illustrated in Calvo (1978) also surfaces in a wide variety of other dynamic optimal taxation examples.

1.11 Conclusions

Recursive dynamic optimization is a main tool of macroeconomic modeling today. In this chapter most of the examples have been models of single agents (Section 1.10 on dynamic games is the exception). In the next three chapters, models will be constructed in which dynamic programming is used to compute and study dynamic general equilibrium models. In these models the single agent who is solving the problem is a fictitious social planner. Solving the planner's problem will be the instrument for computing sequences of equilibrium prices and quantities.

The reader is urged to tackle the exercises below. They provide the practice required for comfort with the notions of states and controls and for recognition of problems whose structure obeys the special conditions required to apply dynamic programming.

Exercises

Exercise 1.1. **Brock-Mirman (1972)**

Consider the Brock-Mirman problem of maximizing

$$\sum_{t=0}^{\infty} \beta^t \ln c_t, \qquad 0 < \beta < 1,$$

subject to $c_t + k_{t+1} \leq Ak_t^{\alpha}$, $0 < \alpha < 1$, k_0 given. Let $v(k)$ be the optimal value function. Use recursions on Bellman's equation (1.27), starting from $v_0(k) \equiv 0$ to show that

$$v(k) = (1-\beta)^{-1}\left[\ln A(1-\alpha\beta) + \frac{\alpha\beta}{1-\alpha\beta}\ln A\beta\alpha\right] + \frac{\alpha}{1-\alpha\beta}\ln k.$$

Exercise 1.2. **Howard Policy-Improvement Algorithm**

Consider the Brock-Mirman problem: to maximize

$$E_0 \sum_{t=0}^{\infty} \beta^t \ln c_t,$$

subject to $c_t + k_{t+1} \leq Ak_t^{\alpha}\theta_t$, k_0 given, $A > 0$, $1 > \alpha > 0$, where $\{\theta_t\}$ is an i.i.d. sequence with $\ln \theta_t$ distributed according to a normal distribution with mean zero and variance σ^2.

Consider the following algorithm. Guess at a policy of the form $k_{t+1} =$

$h_0(Ak_t^\alpha\theta_t)$ for any constant $h_0 \in (0, 1)$. Then form

$$J_0(k_0, \theta_0) = E_0 \sum_{t=0}^{\infty} \beta^t \ln(Ak_t^\alpha\theta_t - h_0 Ak_t^\alpha\theta_t).$$

Next choose a new policy h_1 by maximizing

$$\ln(Ak^\alpha\theta - k') + \beta E J_0(k', \theta'),$$

where $k' = h_1 Ak^\alpha\theta$. Then form

$$J_1(k_0, \theta_0) = E_0 \sum_{t=0}^{\infty} \beta^t \ln(Ak_t^\alpha\theta_t - h_1 Ak_t^\alpha\theta_t).$$

Continue iterating on this scheme until successive h_j have converged.

Show that, for the present example, this algorithm converges to the optimal policy function in one step.

√ *Exercise 1.3.* **Levhari and Srinivasan (1969)**

Assume that

$$u(c) = \frac{1}{1-\alpha} c^{1-\alpha}, \qquad \alpha > 0.$$

Assume that R_t is independently and identically distributed and is such that $ER_t^{1-\alpha} < 1/\beta$. Consider the problem

$$\max E \sum_{t=0}^{\infty} \beta^t u(c_t), \qquad 0 < \beta < 1,$$

subject to $A_{t+1} \leq R_t(A_t - c_t)$, $A_0 > 0$ given. It is assumed that c_t must be chosen before R_t is observed. Show that the optimal policy function takes the form $c_t = \lambda A_t$ and give an explicit formula for λ.

Hint. Consider a value function of the general form $v(A) = BA^{1-\alpha}$, for some constant B.

√ *Exercise 1.4.* **Habit Persistence, 1**

Consider the problem of choosing a consumption sequence c_t to maximize

$$\sum_{t=0}^{\infty} \beta^t(\ln c_t + \gamma \ln c_{t-1}), \qquad 0 < \beta < 1, \qquad \gamma > 0,$$

$$\begin{aligned} \text{subject to} \quad & c_t + k_{t+1} \leq Ak_t^\alpha, \\ & A > 0, \\ & 0 < \alpha < 1, \\ & k_0 > 0, \quad \text{and} \quad c_{-1} \text{ given.} \end{aligned}$$

Here c_t is consumption at t, and k_t is capital stock at the beginning of period t. The current utility function $\ln c_t + \gamma \ln c_{t-1}$ is designed to represent habit persistence in consumption.

a. Let $v(k_0, c_{-1})$ be the value of $\Sigma_{t=0}^{\infty} \beta^t(\ln c_t + \gamma \ln c_{t-1})$ for a consumer who begins time 0 with capital stock k_0 and lagged consumption c_{-1} and behaves optimally. Formulate Bellman's functional equation in $v(k, c_{-1})$.

b. Prove that the solution of Bellman's equation is of the form $v(k, c_{-1}) = E + F \ln k + G \ln c_{-1}$ and that the optimal policy is of the form $\ln k_{t+1} = I + H \ln k_t$ where E, F, G, H, and I are constants. Give explicit formulas for the constants E, F, G, H, and I in terms of the parameters A, β, α, and γ.

Exercise 1.5. **Habit Persistence, 2**

Consider the more general version of the preceding problem, to maximize

$$\sum_{t=0}^{\infty} \beta^t u(c_t, c_{t-1}), \qquad 0 < \beta < 1,$$

subject to $c_t + k_{t+1} \leq f(k_t)$, $k_0 > 0$, c_{-1} given, where $u(c_t, c_{t-1})$ is twice continuously differentiable, bounded, increasing in both c_t and c_{t-1}, and concave in (c_t, c_{t-1}), and where $f'(0) = +\infty$, $f' > 0$, $f'' < 0$.

a. Formulate Bellman's functional equation for this problem.

b. Argue that in general, the optimal consumption plan is to set c_t as a function of both k_t and c_{t-1}. What features of the example in the preceding problem combine to make the optimal consumption plan expressible as a function of k_t alone?

Exercise 1.6. **Lucas and Prescott (1971) and Kydland and Prescott (1982) Meet a Linear Regulator**

Consider a linear quadratic version of a Lucas and Prescott (1971) model that has been modified to incorporate a rich time-to-build structure, à la Kydland and Prescott (1982). We first describe the model in terms of lag operators and then show how it can be mapped into a linear regulator problem.

The equilibrium of the model is supposed to solve the following problem:

$$(1) \qquad \max E_0 \sum_{t=0}^{\infty} \beta^t\{(A_0 - A_1 Y_t + v_t)Y_t - J_t i_t - [d(L)K_t][g(L)K_t]\},$$

$$1 > \beta > 0, \qquad A_0 > 0, \qquad A_1 > 0,$$

subject to

$$(2)\qquad \begin{aligned} Y_t &= a(L)K_t \\ K_{t+1} &= (1-\sigma)K_t + z_t^0, \qquad 0<\sigma<1, \\ z_t^l &= z_{t-1}^{l+1}, \qquad l = 0, 1, \ldots, S-1 \\ \alpha(L)J_t &= \epsilon_{Jt} \\ \xi(L)v_t &= \epsilon_{vt} \\ i_t &= \sum_{j=0}^{S-1} \tau_j z_t^j \end{aligned}$$

where

$$(3)\qquad \begin{aligned} a(L) &= a_0 + a_1L + \ldots + a_NL^N \\ d(L) &= d_0 + d_1L + \ldots + d_NL^N \\ g(L) &= g_0 + g_1L + \ldots + g_NL^N \\ \alpha(L) &= 1 - \alpha_1L - \ldots - \alpha_pL^p \\ \xi(L) &= 1 - \xi_1L - \ldots - \xi_qL^q, \end{aligned}$$

where N, M, R, p, and q are all nonnegative and finite. In (2), ϵ_{Jt} and ϵ_{ut} are fundamental white noises for J_t and u_t, respectively. At time t, variables dated t and earlier are observed.

In (1), Y_t denotes output, c_t is investment expenditures, J_t is the price of new capital goods, K_t is the stock of capital, and v_t is a random process disturbing demand. The technology potentially incorporates two sorts of time-to-build delays. First, output Y_t is a distributed lag, $a(L)K_t$, of the capital stock K_t that is in place. As a result, given the capital stock, the one factor of production, it requires time to produce output. Second, time elapses between the moment when investment decisions z_t^s are made at time t and the moment when the machines can be used as capital, z_{t+s}^0, at time $(t+s)$; z_t^j is interpreted as the number of machines in stage j available at time t. Only machines in stage 0 can increase the capital stock.

One interpretation of the parameters τ_j is that they represent the fraction of the total cost of a machine that is incurred when it is in stage j. Total expenditures in this concept, $\Sigma_{j=0}^{S-1} \tau_j z_t^j$, therefore correspond to investment at time j. It is also possible to think of the parameters τ_j as representing payments to another firm that "builds" the machines. In this sense they can reflect financing arrangements. The firm or industry also faces generalized costs of factor adjustment, which are represented by the cost term $[d(L)K_t][g(L)K_t]$.

This problem can be interpreted in a variety of ways. First, it can be interpreted as the solution of a monopoly problem, where the demand curve facing the monopolist is $p_t = A_0 - A_1Y_t + v_t$, where p_t is the output price.

Second, it can be interpreted as the solution of a rational expectations competitive equilibrium where the demand curve is $p_t = A_0 - 2A_1Y_t + v_t$. Third, it can be interpreted as the outcome of a particular kind of Nash equilibrium (see Hansen, Epple, and Roberds 1985).

a. Show how this general problem can be mapped into the structure of the optimal linear regulator problem. Specify the vector of states and controls and the matrixes R, Q, A, and B.

b. Display the solution in feedback form, and show the difference equation that governs the state under the optimal rule.

Exercise 1.7. **Interrelated Factor Demand**

For another illustration of a problem that can readily be mapped into the linear regulator framework, consider the interrelated factor demand problem, to maximize

$$-E\sum_{t=0}^{\infty} \beta^t\{y_t^TFy_t + [G(L)y_t]^T[H(L)y_t] + J_t^Ty_t\}$$

$$\text{where}\quad y_t = \begin{pmatrix} y_{1t} \\ y_{2t} \end{pmatrix}, \qquad J_t = \begin{pmatrix} J_{1t} \\ J_{2t} \end{pmatrix},$$

F is positive semidefinite; $G(L) = G_0 + G_1L + \ldots + G_mL^m$, $H(L) = H_0 + H_1L + \ldots + H_mL^m$, where G_j and H_j are each (2×2) matrixes. It is assumed that $G_0^TH_0$ is positive definite.

Here J_t denotes a (2×1) vector of factor costs, y_t denotes a (2×1) vector of factors of production, and $[G(L)y_t]^TH(L)y_t$ denotes generalized costs of adjustment. The maximization is subject to a Markov law for J of the form $J_{t+1} = \alpha_1J_t + \ldots + \alpha_{p+1}J_{t-p} + \epsilon_{t+1}$, where ϵ_{t+1} is a (2×1) vector white noise. At time 0, y_{t-j-1} and J_{t-j}, $j \geq 0$, are taken as given.

a. Specify this problem as a linear regulator by defining the states and controls x_t, u_t as well as the matrixes A, B, Q, and R.

Hint. It is easier to map the problem into the following more general version of the linear regulator problem.

$$\max E\sum_{t=0}^{\infty} \beta^t\left\{(x_t^T, u_t^T)\begin{bmatrix} \overline{R} & \overline{W} \\ \overline{W}^T & \overline{Q} \end{bmatrix}\begin{pmatrix} x_t \\ u_t \end{pmatrix}\right\},$$

subject to $x_{t+1} = \overline{A}x_t + \overline{B}u_t + \epsilon_{t+1}$.

The following well-known argument indicates that there is no loss of generality—when Q is nonsingular—in restricting ourselves to the case in which $W = 0$. Simply note that the previous problem is equivalent to the following

$$\max E \sum_{t=0}^{\infty} \beta^t \{x_t^T(\overline{R} - \overline{W}\overline{Q}^{-1}\overline{W}^T)x_t + v_t^T\overline{Q}v_t\},$$

subject to $x_{t+1} = (\overline{A} - \overline{B}\overline{Q}^{-1}\overline{W}^T)x_t + \overline{B}v_t + \epsilon_{t+1}$, where $v_t = \overline{Q}^{-1}\overline{W}^T x_t + u_t$. Therefore, defining R, Q, A, and B by

$$\begin{aligned} R &= \overline{R} - \overline{W}\overline{Q}^{-1}\overline{W}^T \\ Q &= \overline{Q} \\ A &= \overline{A} - \overline{B}\overline{Q}^{-1}\overline{W}^T \\ B &= \overline{B} \end{aligned}$$

gives us the standard version of the problem.

Exercise 1.8. **Two-Sector Growth Models**

a. Consider the following two-sector model of optimal growth. A social planner seeks to maximize the utility of the representative agent given by $\Sigma_{t=0}^{\infty}\beta^t u(c_t, l_t)$, where c_t is consumption of good 1 at t, whereas l_t is leisure at t.

Sector 1 produces consumption goods using capital, k_{1t}, and labor, n_{1t}, according to the production function $c_t \leq f_1(k_{1t}, n_{1t})$. Sector 2 produces the capital good according to the production function $k_{t+1} \leq f_2(k_{2t}, n_{2t})$. Total employment, $n_t = n_{1t} + n_{2t}$, and leisure, l_t, is constrained by the endowment of time, $\bar{l}$, and satisfies $l_t + n_t \leq \bar{l}$. The sum of the amounts of capital used in each sector cannot exceed the initial capital in the economy, that is, $k_{1t} + k_{2t} \leq k_t$, $k_0 > 0$ given. Formulate this problem as a dynamic programming problem. Display the functional equation that the value function satisfies, and clearly specify the state and control variables.

b. Consider another economy that is similar to the previous one except for the fact that capital is sector specific. The economy starts period t with given amounts of capital k_{1t} and k_{2t} that must be used in sectors 1 and 2, respectively. During this period the capital-good sector produces capital that is specific to each sector according to the transformation curve $g(k_{1t+1}, k_{2t+1}) \leq f_2(k_{2t}, n_{2t})$. Display the Bellman's equation associated with the planner's problem. Specify which variables you choose as states and which as controls.

Exercise 1.9. **Learning to Enjoy Spare Time**

A worker's instantaneous utility, $u(\cdot)$, depends on the amount of market-produced goods consumed, c_{1t}, and also on the amount of home-produced goods, c_{2t} (for example, entertainment, leisure). In order to acquire market-produced goods, the worker must allocate some amount of time, l_{1t}, to market activities that pay a salary of w_t, measured in terms of consumption good. The worker takes wages as given and beyond the worker's control. There is no borrowing or lending. It is known that the market wage evolves according to the law of motion $w_{t+1} = h(w_t)$.

The quantity of home-produced goods depends on the stock of "expertise" that the worker has at the beginning of the period, which we label a_t. This stock of "expertise" depreciates at the rate δ and can be increased by allocating time to nonmarket activities. To summarize the problem, the individual agent maximizes

$$\sum_{t=0}^{\infty} \beta^t u(c_{1t}, c_{2t}), \qquad 0 < \beta < 1,$$

subject to

$c_{1t} \leq w_t l_{1t}$	[budget constraint]
$c_{2t} \leq f(a_t)$	[production function of the home-produced good]
$a_{t+1} \leq (1-\delta)a_t + l_{2t}$	[law of motion of the stock of expertise]
$l_{1t} + l_{2t} \leq \bar{l}$	[restriction on the uses of time]
$w_{t+1} = h(w_t)$	[law of motion for the wage rate]
$a_0 > 0$	[given].

It is assumed that $u(\cdot)$ and $f(\cdot)$ are bounded and continuous. Formulate this problem as a dynamic programming problem.

Exercise 1.10. **Investment with Adjustment Costs**

A firm maximizes present value of cash flow, with future earnings discounted at the rate β. Income at time t is given by sales, $p_t \cdot q_t$, where p_t is the price of good, and q_t is the quantity produced. The firm behaves competitively and therefore takes prices as given. It knows that prices evolve according to a law of motion given by $p_{t+1} = f(p_t)$.

Total or gross production depends on the amounts of capital, k_t, and labor, n_t, and on the square of the difference between current ratio of sales to investment, x_t, and the previous-period ratio. This last feature captures the notion that changes in the ratio of sales to investment require some reallocation of resources within the firm and consequently reduce the level of efficiency. It is assumed that the wage rate is constant and equal to w. Capital depreciates at the rate δ. The firm's problem is

$$\max \sum_{t=0}^{\infty} \beta^t \{p_t q_t - w n_t\}, \qquad 0 < \beta < 1,$$

$$\text{subject to} \quad q_t + x_t \le g\left[k_t, n_t, \left(\frac{q_t}{x_t} - \frac{q_{t-1}}{x_{t-1}}\right)^2\right]$$
$$k_{t+1} \le (1-\delta)k_t + x_t, \qquad 0 < \delta < 1$$
$$p_{t+1} = f(p_t)$$
$$k_0 > 0, \quad \frac{q_{-1}}{x_{-1}} > 0 \quad \text{given.}$$

We assume that $g(\cdot)$ is bounded, increasing in the first two arguments and decreasing in the third. Formulate the firm's problem recursively, that is, formulate Bellman's functional equation for this problem. Identify the state and the controls, and indicate the laws of motion of the state variables.

Exercise 1.11. **Investment with Signal Extraction**

Consider a firm that maximizes expected present value of dividends. It is assumed that the price of the good produced by the firm is constant and equal to one. Production requires the use of a single input: capital that is firm specific. Total production, $f(k_t)$, is divided between sales, q_t, and investment, x_t. Revenue from sales is taxed at the rate τ_t. At time t, τ_t is known, as is z_t—a variable that is related to τ_{t+1} by the function $\tau_{t+1} = g(z_t, \epsilon_{t+1})$, where ϵ_{t+1} is an i.i.d. random variable that is not observed at t but whose distribution is known to the firm. Notice that, given z_t, the function g induces a conditional distribution of τ_{t+1} that we denote $F(\tau_{t+1}, z_t)$. The stochastic process $\{z_t\}$ is Markov with transition function $H(z', z) \equiv \text{prob}(z_{t+1} \le z' | z_t = z)$. The capital stock depreciates at the rate τ. The problem faced by the firm is

$$\max E_0 \sum_{t=0}^{\infty} \beta^t (1-\tau_t) q_t, \qquad 0 < \beta < 1,$$

$$\text{subject to} \quad q_t + x_t \le f(k_t)$$
$$k_{t+1} \le (1-\delta)k_t + x_t, \qquad k_0 \text{ given}, \qquad 0 < \delta < 1.$$

It is assumed that $f(k)$ is increasing, concave, and bounded.

Formulate the firm's problem as a dynamic programming problem (that is, display Bellman's equation).

References and Suggested Readings

Basar, Tamer, and Geert Jan Olsder. 1982. *Dynamic Noncooperative Game Theory.* New York: Academic Press.

Bellman, Richard. 1957. *Dynamic Programming.* Princeton, N.J.: Princeton University Press.

Bellman, Richard, and Stuart E. Dreyfus. 1962. *Applied Dynamic Programming.* Princeton, N.J.: Princeton University Press.

Benveniste, Lawrence, and Jose Scheinkman. 1979. On the differentiability of the value function in dynamic models of economics. *Econometrica* 47(3): 727–732.

Bertsekas, Dimitri. 1976. *Dynamic Programming and Stochastic Control.* New York: Academic Press. (Esp. chaps. 2, 6.)

Bertsekas, Dimitri, and Steven E. Shreve. 1978. *Stochastic Optimal Control: The Discrete Time Case.* New York: Academic Press.

Brock, William A. 1982. Asset prices in a production economy. In *The Economics of Information and Uncertainty,* ed. J. J. McCall, pp. 1–43. Chicago: University of Chicago Press.

Brock, William A., and Leonard Mirman. 1972. Optimal economic growth and uncertainty: the discounted case. *Journal of Economic Theory* 4(3):479–513.

Cagan, Phillip. 1956. The monetary dynamics of hyperinflation. In *Studies in the Quantity Theory of Money,* ed. Milton Friedman, pp. 25–117. Chicago: University of Chicago Press.

Calvo, Guillermo A. 1978. On the time consistency of optimal policy in a monetary economy. *Econometrica* 46(6):1411–1428.

Cass, David. 1965. Optimum growth in an aggregative model of capital accumulation. *Review of Economic Studies* 32(3):233–240.

Chow, Gregory. 1981. *Econometric Analysis by Control Methods.* New York: Wiley, 1981.

Hansen, Lars P., Dennis Epple, and Will Roberds. 1985. Linear-quadratic duopoly models of resource depletion. In *Energy, Foresight, and Strategy,* ed. Thomas J. Sargent, pp. 101–142. Washington, D.C.: Resources for the Future.

Hansen, Lars P., and Thomas J. Sargent. 1980. Formulating and estimating dynamic linear rational expectations models. *Journal of Economic Dynamics and Control* 2(1):7–46.

——— 1981. Linear rational expectations models for dynamically interrelated variables. In *Rational Expectations and Econometric Practice,* ed. R. E. Lucas, Jr., and T. J. Sargent, pp. 127–156. Minneapolis: University of Minnesota Press.

——— 1982. Instrumental variables procedures for estimating linear rational expectations models. *Journal of Monetary Economics* 9(3):263–296.

Koopmans, T. C. 1963. On the concept of optimal economic growth. Cowles Foundation Discussion Paper. Yale University, New Haven.

Kwakernaak, Huibert, and Raphael Sivan. 1972. *Linear Optimal Control Systems.* New York: Wiley, 1972.

Kydland, Finn E., and Edward C. Prescott. 1977. Rules rather than discretion: the inconsistency of optimal plans. *Journal of Political Economy* 85(3):473–491.

——— 1982. Time to build and aggregate fluctuations. *Econometrica* 50(6):1345–1371.

Levhari, D., and T. N. Srinivasan. 1969. Optimal savings under uncertainty. *Review of Economic Studies* 36(2):153–163.

Lucas, Robert E., Jr. 1976. Econometric policy evaluation: a critique. In *The Phillips Curve and Labor Markets,* ed. K. Brunner and A. H. Meltzer, pp. 19–46. Amsterdam: North-Holland, 1976.

——— 1977. Class notes on dynamic programming. University of Chicago. Unpublished.

——— 1978. Asset prices in an exchange economy. *Econometrica* 46(6):1426–1445.

Lucas, Robert E., Jr., and Edward C. Prescott. 1971. Investment under uncertainty. *Econometrica* 39(5):659–681.
Lucas, Robert E., Jr., Edward C. Prescott, and Nancy L. Stokey. *Recursive Methods for Economic Dynamics.* Forthcoming.
Lucas, Robert E., Jr., and Nancy Stokey. 1983. Optimal monetary and fiscal policy in an economy without capital. *Journal of Monetary Economics* 12(1):55–94.
Sargent, Thomas J. 1981. Lecture notes on filtering, control, and rational expectations. University of Minnesota, Minneapolis. Unpublished.
——— 1979. *Macroeconomic Theory.* New York: Academic Press.
——— 1986. *Macroeconomic Theory.* 2nd ed. New York: Academic Press.

2 Search

This chapter applies dynamic programming in contexts in which there is a choice between only two actions, as distinguished from the situations studied in the previous chapter, in which the control was typically permitted to take on a continuum of values. The two actions are whether to accept or reject a take-it-or-leave-it offer. In particular, we shall study a variety of search problems, including several contexts in which a buyer or seller is confronted with a probability distribution of prices or characteristics of a job or good from which additional offers can be drawn at a fixed cost per offer. Given the worker's perception of the probability distribution of offers, the worker must devise a strategy for deciding how many offers to solicit before deciding to accept one.

The theory of search was pioneered by Stigler and McCall. It is interesting to macroeconomists because it provides a tool for studying the phenomenon of seemingly unemployed resources. We observe unemployed workers and pieces of capital and variations over time in aggregates of these variables. To explain these observations, search theory puts sellers of labor or capital in a setting in which they rationally *choose* to reject available offers and to remain unemployed in return for the opportunity to wait for better prospective offers in the future. We want to use the theory to study how workers' choices would respond to variations in the rate of unemployment compensation, the perceived riskiness of wage distributions, the quality of information about jobs, and the "technology" for sampling the wage distribution.

The present chapter aims to provide an introduction to the techniques used in the search literature and a sampling of search models. The chapter is