FURTHER MATHEMATICS FOR ECONOMIC ANALYSIS

KNUT SYDSÆTER
PETER HAMMOND
ATLE SEIERSTAD
ARNE STRØM

FT Prentice Hall
FINANCIAL TIMES

An imprint of Pearson Education
Harlow, England • London • New York • Boston • San Francisco • Toronto • Sydney • Singapore • Hong Kong
Tokyo • Seoul • Taipei • New Delhi • Cape Town • Madrid • Mexico City • Amsterdam • Munich • Paris • Milan

5 Differential Equations I: First-Order Equations in One Variable 185

5.1 Introduction 185
5.2 The Direction is Given: Find the Path! 189
5.3 Separable Equations 190
5.4 First-Order Linear Equations 197
5.5 Exact Equations and Integrating Factors 203
5.6 Transformation of Variables 206
5.7 Qualitative Theory and Stability 209
5.8 Existence and Uniqueness 215

6 Differential Equations II: Second-Order Equations and Systems in the Plane 221

6.1 Introduction 221
6.2 Linear Differential Equations 224
6.3 Constant Coefficients 226
6.4 Stability for Linear Equations 233
6.5 Simultaneous Equations in the Plane 235
6.6 Equilibrium Points for Linear Systems 241
6.7 Phase Plane Analysis 244
6.8 Stability for Nonlinear Systems 249
6.9 Saddle Points 253

7 Differential Equations III: Higher-Order Equations 257

7.1 Linear Differential Equations 257
7.2 Constant Coefficients 261
7.3 Stability of Linear Differential Equations 264
7.4 Systems of Differential Equations 267
7.5 Stability for Nonlinear Systems 271
7.6 Qualitative Theory 276
7.7 A Glimpse at Partial Differential Equations 278

8 Calculus of Variations 285

8.1 The Simplest Problem 285
8.2 The Euler Equation 288
8.3 Why the Euler Equation is Necessary 291
8.4 Optimal Savings 296
8.5 More General Terminal Conditions 298
8.6 Generalizations 303

9 Control Theory: Basic Techniques 305

9.1 The Basic Problem 305
9.2 A Simple Case 307
9.3 Regularity Conditions 311
9.4 The Standard Problem 313
9.5 The Maximum Principle and the Calculus of Variations 321
9.6 Adjoint Variables as Shadow Prices 323
9.7 Sufficient Conditions 329
9.8 Variable Final Time 335
9.9 Current Value Formulations 337
9.10 Scrap Values 340
9.11 Infinite Horizon 347
9.12 Phase Diagrams 352

10 Control Theory with Many Variables 359

10.1 Several Control and State Variables 359
10.2 Some Examples 366
10.3 Infinite Horizon 371
10.4 Existence Theorems and Sensitivity 374
10.5 A Heuristic Proof of the Maximum Principle 378
10.6 Mixed Constraints 380
10.7 Pure State Constraints 384
10.8 Generalizations 387

11 Difference Equations 389

11.1 First-Order Difference Equations 389
11.2 Present Discounted Values 396
11.3 Second-Order Difference Equations 397
11.4 Constant Coefficients 401
11.5 Higher-Order Equations 408
11.6 Systems of Difference Equations 412
11.7 Stability of Nonlinear Difference Equations 416

12 Discrete Time Optimization 421

12.1 Dynamic Programming 421
12.2 The Euler Equation 431
12.3 Infinite Horizon 433
12.4 The Maximum Principle 437

12.5 More Variables 441
12.6 Stochastic Optimization 444
12.7 Infinite Horizon Stationary Problems 453

13 Topology and Separation 459

13.1 Point Set Topology in $\mathbb{R}^n$ 459
13.2 Topology and Convergence 465
13.3 Continuous Functions 469
13.4 Maximum Theorems 475
13.5 Convex Sets 479
13.6 Separation Theorems 484
13.7 Productive Economies and Frobenius's Theorem 489

14 Correspondences and Fixed Points 493

14.1 Correspondences 494
14.2 A General Maximum Theorem 504
14.3 Fixed Points for Contraction Mappings 508
14.4 Brouwer's and Kakutani's Fixed Point Theorems 510
14.5 Equilibrium in a Pure Exchange Economy 515

Appendix A Sets, Completeness and Convergence 519

A.1 Sets and Functions 519
A.2 Least Upper Bound Principle 524
A.3 Sequences of Real Numbers 527
A.4 Infimum and Supremum of Functions 536

Appendix B Trigonometric Functions 539

B.1 Basic Definitions and Results 539
B.2 Differentiating Trigonometric Functions 545
B.3 Complex Numbers 549

Answers to Odd-Numbered Exercises 555

References 585

Index 589

(b) The equation $x = e^x - 3$ also has a positive solution, but this is an unstable equilibrium of $x_{t+1} = e^{x_t} - 3$. Explain how nevertheless we can find the positive solution by rewriting the equation and using the same technique as above.

3. The function f in Fig. 4 is given by $f(x) = -x^2 + 4x - 4/5$. Find the values of the cycle points ξ_1 and ξ_2, and use (5) to determine whether the cycle is stable. It is clear from the figure that the difference equation $x_{t+1} = f(x_t)$ has two equilibrium states. Find these equilibria, show that they are both unstable, and verify the result in Problem 1.

12

DISCRETE TIME OPTIMIZATION

In science, what is capable of proof
must not be believed without a proof.[1]
—R. Dedekind (1887)

This chapter gives a brief introduction to *discrete time dynamic optimization problems*. Th term *dynamic* refers to the fact that the problems involve systems evolving over time. Tim is here measured by the number of whole periods (say weeks, quarters, or years) that hav passed since time 0. So we speak of *discrete* time. In this case it is natural to study dynami systems whose development is governed by difference equations.

If the horizon is finite, then such dynamic problems can be solved, in principle, using classica calculus methods. There are, however, special solution techniques described in the presen chapter that take advantage of the special structure of discrete dynamic optimization problems

Most of the chapter is concerned with dynamic programming. This is a general metho for solving discrete time optimization problems that was formalized by R. Bellman in the lat 1950s. There is also a brief introduction to discrete time control theory. The last two section cover stochastic dynamic programming. (This is the only part of the book that relies on som knowledge of probability theory, though at a basic level.)

12.1 Dynamic Programming

Consider a system that is observed at times $t = 0, 1, \ldots, T$. Suppose the **state** of the syster at time t is characterized by a real number x_t. For example, x_t might be the quantity of grai that is stockpiled at time t. Assume that the initial state x_0 is historically given, and tha from then on the system evolves through time under the influence of a sequence of **control** u_t, which can be chosen freely from a given set U, called the **control region**. For example u_t might be the fraction of grain removed from the stock x_t during period t. The control

[1] There is no ideal English translation of the German original: "Was beweisbar ist, soll in de Wissenschaft nicht ohne Beweis geglaubt werden."

influence the evolution of the system through a difference equation

$$x_{t+1} = g(t, x_t, u_t), \quad x_0 \text{ given}, \quad u_t \in U \tag{1}$$

where g is a *given* function. Thus, we assume that the state of the system at time $t+1$ depends explicitly on the time t, on the state x_t in the preceding period t, and on u_t, the value chosen for the control at time t.

Suppose that we choose values for $u_0, u_1, \ldots, u_{T-1}$. Then (1) gives $x_1 = g(0, x_0, u_0)$. Since x_1 is now known, $x_2 = g(1, x_1, u_1)$, and next $x_3 = g(2, x_2, u_2)$, etc. In this way, (1) can be used to compute successively, or recursively, the values or states $x_1, x_2, \ldots, x_T$ in terms of the initial state, x_0, and the time path of the controls, $u_0, \ldots, u_{T-1}$. Each choice of $(u_0, u_1, \ldots, u_{T-1})$ gives rise to a sequence $(x_1, x_2, \ldots, x_T)$, for instance path 1 in Fig. 1. A different choice of $(u_0, u_1, \ldots, u_{T-1})$ gives another path, such as path 2 in the figure. Such controls u_t that depend only on time, are often called **open-loop controls**.

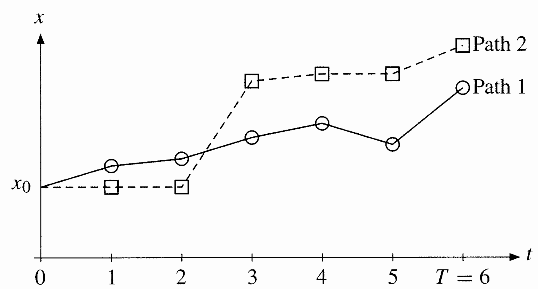

Figure 1 Different evolutions of system (1)

Different paths will usually have different utility or value. Assume that there is a function $f(t, x, u)$ of three variables such that the utility associated with a given path is represented by the sum

$$\sum_{t=0}^{T} f(t, x_t, u_t) \tag{$*$}$$

The sum is called the **objective function**, and it represents the sum of utilities (values) obtained at each point of time.

NOTE 1 The objective function is sometimes specified as $\sum_{t=0}^{T-1} f(t, x_t, u_t) + S(x_T)$, where S measures the net value associated with the terminal period. This is a special case of $(*)$ in which $f(T, x_T, u_T) = S(x_T)$. (S is often called a **scrap value function**.)

Suppose that we choose values for $u_0, u_1, \ldots, u_{T-1}, u_T$, all from the set U. The initial state x_0 is given, and as explained above, (1) gives us $x_1, \ldots, x_T$. Let us denote corresponding pairs $(x_0, \ldots, x_T)$, $(u_0, \ldots, u_T)$ by $(\{x_t\}, \{u_t\})$, and call them **admissible sequence pairs**. For each admissible sequence pair the objective function has a definite value. We shall study the following problem:

Among all admissible sequence pairs $(\{x_t\}, \{u_t\})$ *find one,* $(\{x_t^*\}, \{u_t^*\})$, *that makes the value of the objective function as large as possible.* Such an admissible sequence pair is

called an **optimal pair**, and the corresponding control sequence $\{u_t^*\}_{t=0}^{T}$ is called an **o**
control. Briefly formulated, the problem is this:

$$\max \sum_{t=0}^{T} f(t, x_t, u_t) \quad \text{subject to} \quad x_{t+1} = g(t, x_t, u_t), \ x_0 \text{ given}, \ u_t \in U$$

EXAMPLE 1 Let x_t be an individual's wealth at time t. At each point of time t, the individu to decide the proportion u_t of x_t to consume, leaving the remaining proportion $1 -$ savings. Assume that wealth earns interest at the rate $\rho - 1 > 0$. After $u_t x_t$ is withdra consumption, the stock of wealth is $(1 - u_t)x_t$. Because of interest, this grows to the a $x_{t+1} = \rho(1 - u_t)x_t$ at the beginning of period $t+1$. This goes for $t = 0, \ldots, T - 1$ x_0 a positive constant. Suppose that the utility of consuming $c_t = u_t x_t$ is $U(t, c_t)$, $U(t, c)$ is increasing and concave in c. Then the total utility over periods $t = 0, \ldots$ $\sum_{t=0}^{T} U(t, u_t x_t)$. The problem facing the individual is therefore the following:

$$\max \sum_{t=0}^{T} U(t, u_t x_t) \quad \text{subject to} \quad x_{t+1} = \rho(1 - u_t)x_t, \ t = 0, \ldots, T - 1$$

with x_0 given and with u_t in $[0, 1]$ for $t = 0, \ldots, T$. Note that this is a standard dy optimization problem of the type described above. (See Problems 2, 3, and 8.)

The Value Function and its Properties

Return to the general problem described by (2). In order to find the optimal solutio shall use a method that appears to solve a more general problem.

Suppose that at time $t = s$ the state of the system is x (any given real number best we can do in the remaining periods is to choose $u_s, u_{s+1}, \ldots, u_T$ (and thereb $x_{s+1}, \ldots, x_T$) to maximize $\sum_{t=s}^{T} f(t, x_t, u_t)$ with $x_s = x$. We define the **(optimal) function** for the problem at time s by[2]

$$J_s(x) = \max_{u_s, \ldots, u_T \in U} \sum_{t=s}^{T} f(t, x_t, u_t)$$

where

$$x_s = x \quad \text{and} \quad x_{t+1} = g(t, x_t, u_t) \ \text{for} \ t > s, \quad u_t \in U$$

The controls $u_s^*, \ldots, u_T^*$ that give the maximum value in (3) subject to (4), will d on x. In particular, the first control, u_s^*, will depend on $x \in \mathbb{R}$, $u_s^* = u_s^*(x_s) = u_s^*(x)$, $x = x_s$ is the state at time s. Controls that depend on the state of the system in this wa called **closed-loop controls**, **feedback controls**, or **policies**.

Suppose that we have found the first control $u_s^*(x)$ for *each* $s = 0, 1, \ldots, T$. The have actually found the solution to the original problem (2). In particular, since the st

[2] We assume that the maximum in (3) is attained. This is true if, for example, the functi and g are continuous and U is compact.

$t = 0$ is x_0, the best choice of u_0 is $u_0^*(x_0)$. After $u_0^*(x_0)$ is found, the difference equation in (1) determines the corresponding x_1 as $x_1^* = g(0, x_0, u_0^*(x_0))$. Then $u_1^*(x_1^*)$ is the best choice of u_1 and this choice determines x_2^* by (1). Then again, $u_2^*(x_2^*)$ is the best choice of u_2, and so on.

We now prove an important property of the value function. At the terminal time $t = T$, we have $J_T(x) = \max_{u \in U} f(T, x, u)$. Suppose that at time $t = s$ $(< T)$ we are in state $x_s = x$. What is the optimal choice for u_s? If we choose $u_s = u$, then at time $t = s$ we obtain the immediate reward $f(s, x, u)$, and according to (4), the state at time $s + 1$ will be $x_{s+1} = g(s, x, u)$. The highest total reward obtainable from time $s + 1$ to time T, starting from the state x_{s+1}, is $J_{s+1}(x_{s+1}) = J_{s+1}(g(s, x, u))$ according to definition (3). Hence the best choice of $u = u_s$ at time s must be a value of u that maximizes the sum

$$f(s, x, u) + J_{s+1}(g(s, x, u))$$

This leads to the following general result:

THEOREM 12.1.1 (FUNDAMENTAL EQUATIONS OF DYNAMIC PROGRAMMING)

Let $J_s(x)$ be the value function (3) for the problem

$$\max \sum_{t=0}^{T} f(t, x_t, u_t) \quad \text{subject to} \quad x_{t+1} = g(t, x_t, u_t), \quad u_t \in U \tag{5}$$

with x_0 given. Then $J_s(x)$ satisfies the equations

$$J_s(x) = \max_{u \in U} \left[f(s, x, u) + J_{s+1}(g(s, x, u)) \right], \quad s = 0, 1, \ldots, T - 1 \tag{6}$$

$$J_T(x) = \max_{u \in U} f(T, x, u) \tag{7}$$

NOTE 2 If we minimize rather than maximize the sum in (5), then Theorem 12.1.1 holds with "max" replaced by "min" in (6) and (7), because minimizing f is equivalent to maximizing $-f$.

NOTE 3 Let $\mathcal{X}_t(x_0)$ denote the range of all possible values of the state x_t that can be generated by the difference equation (1) if we start in state x_0 and then go through all possible values of $u_0, \ldots, u_{t-1}$. Of course J_t need only be defined on $\mathcal{X}_t(x_0)$.

Theorem 12.1.1 is the basic tool for solving dynamic optimization problems. It is used as follows: First find the function $J_T(x)$ by using (7). The maximizing value of u depends (usually) on x, and was denoted by $u_T^*(x)$ above. The next step is to use (6) to determine $J_{T-1}(x)$ and the corresponding $u_{T-1}^*(x)$. Then work backwards in this fashion to determine recursively all the value functions $J_T(x), \ldots, J_0(x)$ and the maximizers $u_T^*(x), \ldots, u_0^*(x)$. As explained above, this allows the solution to the original optimization problem to be constructed.

EXAMPLE 2 Use Theorem 12.1.1 to solve the problem

$$\max \sum_{t=0}^{3} (1 + x_t - u_t^2), \quad x_{t+1} = x_t + u_t, \quad t = 0, 1, 2, \quad x_0 = 0, \quad u_t \in \mathbb{R}$$

Solution: Here $T = 3$, $f(t, x, u) = 1 + x - u^2$, and $g(t, x, u) = x + u$. Cor first (7) and note that $J_3(x)$ is the maximum value of $1 + x - u^2$ for $u \in (-\infty, \infty)$. maximum value is obviously attained for $u = 0$. Hence, in the notation introduced ab

$$J_3(x) = 1 + x, \quad \text{with } u_3^*(x) \equiv 0$$

For $s = 2$, the function to be maximized in (6) is $h_2(u) = 1 + x - u^2 + J_3(x + u$ course, $J_3(x + u)$ is obtained by replacing x by $x + u$ in the formula for $J_3(x)$. $h_2(u) = 1 + x - u^2 + 1 + (x + u) = 2 + 2x + u - u^2$. The function $h_2(u)$ is concave and $h_2'(u) = 1 - 2u = 0$ for $u = 1/2$, so this is the optimal choice of u. Then the maxi value of $h_2(u)$ is $h_2(1/2) = 2 + 2x + 1/2 - 1/4 = 9/4 + 2x$. Hence,

$$J_2(x) = \tfrac{9}{4} + 2x, \quad \text{with } u_2^*(x) \equiv \tfrac{1}{2}$$

For $s = 1$, the function to be maximized in (6) is given by $h_1(u) = 1 + x - u^2 + J_2(x +$ $1 + x - u^2 + 9/4 + 2(x + u) = 13/4 + 3x + 2u - u^2$. Because h_1 is concave $h_1'(u) = 2 - 2u = 0$ for $u = 1$, the maximum value of $h_1(u)$ is $13/4 + 3x + 2 -$ $17/4 + 3x$, so

$$J_1(x) = \tfrac{17}{4} + 3x, \quad \text{with } u_1^*(x) \equiv 1$$

Finally, for $s = 0$, the function to be maximized is $h_0(u) = 1 + x - u^2 + J_1(x +$ $1 + x - u^2 + 17/4 + 3(x + u) = 21/4 + 4x + 3u - u^2$. The function h_0 is con and $h_0'(u) = 3 - 2u = 0$ for $u = 3/2$, so the maximum value of $h_0(u)$ is $h_0(3/2$ $21/4 + 4x + 9/2 - 9/4 = 15/2 + 4x$. Thus,

$$J_0(x) = \tfrac{15}{2} + 4x, \quad \text{with } u_0^*(x) \equiv \tfrac{3}{2}$$

In this particular case the optimal choices of the controls are constants, independent of states. The corresponding optimal values of the state variables are $x_1 = x_0 + u_0 =$ $x_2 = x_1 + u_1 = 3/2 + 1 = 5/2$, $x_3 = x_2 + u_2 = 5/2 + 1/2 = 3$. The maximum valu the objective function is 7.5.

Alternative solution: In simple cases like this, a dynamic optimization problem can be sol quite easily by ordinary calculus methods. By letting $t = 0$, 1, and 2 in the differe equation $x_{t+1} = x_t + u_t$, we get $x_1 = x_0 + u_0 = u_0$, $x_2 = x_1 + u_1 = u_0 + u_1$, $x_3 = x_2 + u_2 = u_0 + u_1 + u_2$. Using these results, the objective function becomes following function of u_0, u_1, u_2, and u_3:

$$\begin{aligned} I &= (1 - u_0^2) + (1 + u_0 - u_1^2) + (1 + u_0 + u_1 - u_2^2) + (1 + u_0 + u_1 + u_2 - u_3^2) \\ &= 4 + 3u_0 - u_0^2 + 2u_1 - u_1^2 + u_2 - u_2^2 - u_3^2 \end{aligned}$$

The problem has been reduced to that of maximizing I with respect to the control variables u_0, u_1, u_2, and u_3. We see that I is a sum of concave functions and so is concave. Hence a stationary point will maximize I. The first-order derivatives of I are

$$\frac{\partial I}{\partial u_0} = 3 - 2u_0, \quad \frac{\partial I}{\partial u_1} = 2 - 2u_1, \quad \frac{\partial I}{\partial u_2} = 1 - 2u_2, \quad \frac{\partial I}{\partial u_3} = -2u_3$$

Equating these partial derivatives to zero yields the unique stationary point $(u_0, u_1, u_2, u_3) = (3/2, 1, 1/2, 0)$, which then solves our problem. We have the same solution as the one we obtained by using Theorem 12.1.1.

In principle, all finite dimensional dynamic programming problems can be solved this way using ordinary calculus, but the method becomes very unwieldy if the horizon T is large.

In the next example the terminal time is an arbitrarily given natural number and the optimal control turns out to depend on the state of the system.

EXAMPLE 3 Solve the following problem

$$\max\Big(\sum_{t=0}^{T-1} -\tfrac{2}{3}u_t x_t + \ln x_T\Big), \quad x_{t+1} = x_t(1 + u_t x_t), \quad x_0 \text{ positive constant}, \quad u_t \ge 0 \quad (*)$$

Solution: Because $x_0 > 0$ and $u_t \ge 0$, we have $x_t > 0$ for all t. Now, $f_0(T, x, u) = \ln x$ is independent of u, so $J_T(x) = \ln x$, and any u_T is optimal. Equation (6) with $s = T - 1$ yields

$$J_{T-1}(x) = \max_{u\ge 0}\big[-\tfrac{2}{3}ux + J_T(x(1+ux))\big] = \max_{u\ge 0}\big[-\tfrac{2}{3}ux + \ln x + \ln(1+ux)\big]$$

The maximum of the concave function $h(u) = -\frac{2}{3}ux + \ln x + \ln(1+ux)$ is where its derivative is 0. This gives $h'(u) = -\frac{2}{3}x + x/(1+ux) = 0$, or (since we can assume $x > 0$), $u = 1/(2x)$. Then $h(1/(2x)) = \ln x - 1/3 + \ln(3/2)$. Hence

$$J_{T-1}(x) = h(1/(2x)) = \ln x + C, \text{ with } C = -1/3 + \ln(3/2), \text{ and } u^*_{T-1}(x) = 1/(2x)$$

The next step is to use (6) for $s = T - 2$:

$$J_{T-2}(x) = \max_{u\ge 0}\big[-\tfrac{2}{3}ux + J_{T-1}(x(1+ux))\big] = \max_{u\ge 0}\big[-\tfrac{2}{3}ux + \ln x + \ln(1+ux) + C\big]$$

Again $u = u^*_{T-2}(x) = 1/(2x)$ gives the maximum because the first-order condition is the same, and we get

$$J_{T-2}(x) = \ln x + 2C, \text{ with } C = -1/3 + \ln(3/2), \text{ and } u^*_{T-2}(x) = 1/(2x)$$

This pattern continues and so, for $k = T, T-1, \dots, 1, 0$, we get

$$J_{T-k}(x) = \ln x + kC, \text{ with } C = -1/3 + \ln(3/2), \text{ and } u^*_{T-k}(x) = 1/(2x)$$

(or $J_t(x) = \ln x + (T-t)C$, $u^*_t = 1/(2x)$). Inserting $u^*_t = 1/(2x^*_t)$ in the difference equation gives $x^*_{t+1} = \left(\frac{3}{2}\right)x^*_t$, so $x^*_t = \left(\frac{3}{2}\right)^t x_0$, with $\bar{u}_t = \left(\frac{2}{3}\right)^t/(2x_0)$ as optimal control values.

NOTE 4 Consider the difference equation $x_{t+1} = g(t, x_t, u_t)$ and any given sequer policies $u_0(x), \dots, u_{T-1}(x)$, all taking values in a control region U. When the initial st is given, the evolution of the state x_t is then uniquely determined by the difference equ

$$x_{t+1} = g(t, x_t, u_t(x_t)), \quad x_0 \text{ given}$$

Let us write down the control values (numbers) $\bar{u}_t = u_t(x_t)$ given by this particular seq of states $\{x_t\}$. Next insert these numbers $\bar{u}_t$ into the difference equation:

$$x_{t+1} = g(t, x_t, \bar{u}_t), \quad x_0 \text{ given}$$

This difference equation has the same solution as equation (∗). Hence, we get the result whether we insert the functions $u_t(x)$, or the numbers $\bar{u}_t$.

Consider, for example, the case $x_{t+1} = x_t + u_t$, and choose $u_t(x) = 2x$ for all t. equation (∗) is $x_{t+1} = x_t + 2x_t = 3x_t$, and with x_0 given, the solution is $x_t = 3^t x_0$ associated controls are $u_t = 2x_t = 2 \cdot 3^t x_0$, and equation (∗∗) is now $x_{t+1} = x_t + 2 \cdot$ This equation is easily seen to have the solution $x_t = 3^t x_0$ as well. (Insert and check.

Now, the dynamic programming method gives us the optimal control functions u Given the initial situation x_0, once we have calculated the values $\bar{u}_t$ using the optimal co functions $u^*_t(x)$, we can forget about these functions: At each point of time, we know it is optimal to use $\bar{u}_t$ as the control variable.

It may nevertheless be useful not to forget entirely the form of each control func Suppose that at time τ, there is an unexpected disturbance to the state x^*_τ obtained fro difference equation, which has the effect of changing the state to $\hat{x}_\tau$. Then $u^*_\tau(\hat{x}_\tau)$ still the optimal control to be used at that time, provided we know that no further disturba will occur.

NOTE 5 Theorem 12.1.1 also holds if the control region is not fixed, but depends on ($U = U(t, x)$. Then the maximization in (2), (3), and (5) is carried out for u_t in $U(t, x$ (6) and (7), the maximization is carried out for $u \in U(s, x)$ and $u \in U(T, x)$, respecti Frequently, the set $U(t, x)$ is determined by a set of inequalities, $\{u : h_i(t, x, u) \le 0$ $1, \dots, i^*\}$. If $U(t, x)$ is empty, then by convention, the maximum over $U(t, x)$ is set e to $-\infty$.

NOTE 6 In the above formulation, the state x and the control u may well be vecto say $\mathbb{R}^n$ and $\mathbb{R}^r$, respectively. Then g must be a vector function as well, and the differ equation is a system of difference equations, one for each component of x. No change then needed in Theorem 12.1.1 (except that we would use boldface letters for x, u, an

EXAMPLE 4 Let x_t denote the value of an investor's assets at the start of period t, and u_t consum during period t. Suppose that assets at the start of period $t+1$ are proportional to sav $x_t - u_t$ in period t, with a factor of proportionality depending on t, i.e.

$$x_{t+1} = a_t(x_t - u_t), \quad a_t \text{ given positive numbers}$$

Assume that the initial assets, x_0, are positive. The utility associated with a level of consumption u during one period is supposed to be $u^{1-\gamma}$, while the utility of the assets at the end of period T is $Ax_T^{1-\gamma}$. Here A is a positive constant and $\gamma \in (0, 1)$. The investor wants to maximize the discounted value of the sum of utility from consumption and terminal assets. Define $\beta = 1/(1+r)$, where r is the rate of discount. Assume that no borrowing is allowed, with $0 < u_t < x_t$. The investor's problem is thus:

$$\max\Big[\sum_{t=0}^{T-1} \beta^t u_t^{1-\gamma} + \beta^T A x_T^{1-\gamma}\Big], \quad x_{t+1} = a_t(x_t - u_t), \quad u_t \in (0, x_t) \tag{i}$$

Solution: Combined with Note 5, and with $U(t, x) = (0, x)$, Theorem 12.1.1 can be applied to the present problem. We then have $f(t, x, u) = \beta^t u^{1-\gamma}$ for $t = 0, 1, \dots, T-1$, whereas $f(T, x, u) = \beta^T A x^{1-\gamma}$. Since this function does not depend on u, (7) yields

$$J_T(x) = \max_{u \in (0,x)} \beta^T A x^{1-\gamma} = \beta^T A x^{1-\gamma} \tag{ii}$$

and any u_T in $(0, x)$ is optimal. Moreover, equation (6) yields

$$J_s(x) = \max_{u \in (0,x)} \Big[\beta^s u^{1-\gamma} + J_{s+1}(a_s(x-u))\Big] \tag{iii}$$

In particular, (ii) gives $J_T(a_{T-1}(x-u)) = \beta^T A a_{T-1}^{1-\gamma}(x-u)^{1-\gamma}$, so

$$J_{T-1}(x) = \beta^{T-1} \max_{u \in (0,x)} \Big[u^{1-\gamma} + \beta A a_{T-1}^{1-\gamma}(x-u)^{1-\gamma}\Big] \tag{iv}$$

Put $g(u) = u^{1-\gamma} + \beta A a_{T-1}^{1-\gamma}(x-u)^{1-\gamma}$ for u in $(0, x)$. Computing $g'(u)$ and solving the equation $g'(u) = 0$ for u yields

$$u_{T-1} = u = x/C_{T-1}^{1/\gamma}, \qquad \text{where} \quad C_{T-1}^{1/\gamma} = 1 + (\beta A a_{T-1}^{1-\gamma})^{1/\gamma} \tag{v}$$

Because $\gamma \in (0, 1)$ and $\beta A a_{T-1}^{1-\gamma} > 0$, g is easily seen to be concave over $(0, x)$. Then the value of u given in (v) does maximize $g(u)$. Now,

$$g\Big(\frac{x}{C_{T-1}^{1/\gamma}}\Big) = x^{1-\gamma} C_{T-1}^{(\gamma-1)/\gamma} + \beta A a_{T-1}^{1-\gamma}\Big(x - \frac{x}{C_{T-1}^{1/\gamma}}\Big)^{1-\gamma}$$

$$= x^{1-\gamma} C_{T-1}^{(\gamma-1)/\gamma} + x^{1-\gamma}(C_{T-1}^{1/\gamma} - 1)^{\gamma} \cdot \frac{(C_{T-1}^{1/\gamma} - 1)^{1-\gamma}}{C_{T-1}^{(1-\gamma)/\gamma}} = x^{1-\gamma} C_{T-1}$$

Hence, by (iv),

$$J_{T-1}(x) = \beta^{T-1} C_{T-1} x^{1-\gamma} \tag{vi}$$

Notice that $J_{T-1}(x)$ has the same form as $J_T(x)$. Proceed by substituting $s = T-2$ in (iii) to get:

$$J_{T-2}(x) = \beta^{T-2} \max_{u \in (0,x)} \Big[u^{1-\gamma} + \beta C_{T-1} a_{T-2}^{1-\gamma}(x-u)^{1-\gamma}\Big]$$

Comparing with (iv), we see that the maximum value is attained for

$$u_{T-2} = u = x/C_{T-2}^{1/\gamma}, \qquad \text{where} \quad C_{T-2}^{1/\gamma} = 1 + \big(\beta C_{T-1} a_{T-2}^{1-\gamma}\big)^{1/\gamma}$$

and that $J_{T-2}(x) = \beta^{T-2} C_{T-2} x^{1-\gamma}$. We can obviously go backwards repeatedly i way, and obtain for every t,

$$J_t(x) = \beta^t C_t x^{1-\gamma}$$

From (ii), $C_T = A$, while C_t for $t < T$ is determined recursively backwards by the follc linear difference equation of the first order in $C_t^{1/\gamma}$:

$$C_t^{1/\gamma} = 1 + \big(\beta C_{t+1} a_t^{1-\gamma}\big)^{1/\gamma} = 1 + \big(\beta a_t^{1-\gamma}\big)^{1/\gamma} C_{t+1}^{1/\gamma}$$

The optimal control is

$$u_t^*(x) = x/C_t^{1/\gamma}, \qquad t < T$$

We find the optimal path by successively inserting $u_0^*, u_1^*, \dots$ into the difference equ (i) for x_t.

Suppose in particular that $a_t = a$ for all t. Then (viii) reduces to

$$C_{t+1}^{1/\gamma} - \frac{1}{\omega} C_t^{1/\gamma} = -\frac{1}{\omega}, \qquad \text{where} \quad \omega = \beta^{1/\gamma} a^{1/\gamma - 1}$$

This is a first-order linear difference equation with constant coefficients. Using C_T = and solving the equation for $C_t^{1/\gamma}$, we obtain

$$C_t^{1/\gamma} = A^{1/\gamma}\omega^{T-t} + \frac{1-\omega^{T-t}}{1-\omega}, \qquad t = T, T-1, \dots, 0$$

PROBLEMS FOR SECTION 12.1

1. (a) Use Theorem 12.1.1 to solve the problem

$$\max \sum_{t=0}^{2}[1 - (x_t^2 + 2u_t^2)], \quad x_{t+1} = x_t - u_t, \quad t = 0, 1$$

where $x_0 = 5$ and $u_t \in \mathbb{R}$. (Compute $J_s(x)$ and $u_s^*(x)$ for $s = 2, 1, 0$.)

(b) Use the difference equation in (∗) to compute x_1 and x_2 in terms of u_0 and u_1 ($x_0 = 5$), and find the sum in (∗) as a function S of u_0, u_1, and u_2. Next, maxin this function as in Example 2.

2. Consider the problem

$$\max_{u_t \in [0,1]} \sum_{t=0}^{T} \Big(\frac{1}{1+r}\Big)^t \sqrt{u_t x_t}, \quad x_{t+1} = \rho(1-u_t)x_t, \quad t = 0, \dots, T-1, \quad x_0 > ($$

where r is the rate of discount. Compute $J_s(x)$ and $u_s^*(x)$ for $s = T, T-1, T-2$

3. (a) Replace the utility function in Problem 2 by $\sum_{t=0}^{T}(1+r)^{-t}u_t x_t$. Compute $J_T(x)$, $u_T^*(x)$, $J_{T-1}(x)$, and $u_{T-1}^*(x)$ for $x \geq 0$.

   (b) Prove that there exist constants P_s (depending on ρ and r) such that $J_s(x) = P_s x$ for $s = 0, 1, \dots, T$.

   (c) Find $J_0(x)$ and optimal values of $u_0, u_1, \dots, u_T$.

4. (a) Compute the value functions $J_T(x)$, $J_{T-1}(x)$, $J_{T-2}(x)$, and the corresponding control functions, $u_T^*(x)$, $u_{T-1}^*(x)$, and $u_{T-2}^*(x)$ for the problem

$$\max_{u_t \in [0,1]} \sum_{t=0}^{T}(3 - u_t)x_t^2, \quad x_{t+1} = u_t x_t,\ t = 0, \dots, T-1,\ x_0 \text{ is given}$$

   (b) Try to find a general expression for $J_{T-n}(x)$ for $n = 0, 1, 2, \dots, T$, and the corresponding optimal controls.

5. Solve the problem:

$$\max_{u_t \in [0,1]} \left[\sum_{t=0}^{T-1}\left(-\tfrac{2}{3}u_t\right) + \ln x_T\right], \quad x_{t+1} = x_t(1+u_t),\quad t = 0, \dots, T-1,\ x_0 > 0 \text{ given}$$

6. (a) Write down the fundamental equations for the problem

$$\max_{u_t \in \mathbb{R}} \sum_{t=0}^{T}(x_t - u_t^2),\ x_{t+1} = 2(x_t + u_t),\ t = 0, 1, \dots, T-1,\ x_0 = 0$$

   (b) Prove that the value function for the problem is given by

$$J_{T-n}(x) = (2^{n+1} - 1)x + \sum_{j=0}^{n}(2^j - 1)^2,\ n = 0, 1, \dots, T$$

   Determine the optimal controls $u_t = u_t^*$ and the maximum value $V = J_0(0)$.

7. Consider the problem

$$\max_{u_t \in \mathbb{R}} \left[\sum_{t=0}^{T-1}(-e^{-\gamma u_t}) - \alpha e^{-\gamma x_T}\right],\ x_{t+1} = 2x_t - u_t,\ t = 0, 1, \dots, T-1,\ x_0 \text{ given}$$

   where α and γ are positive constants.

   (a) Compute $J_T(x)$, $J_{T-1}(x)$, and $J_{T-2}(x)$.

   (b) Prove that $J_t(x)$ can be written in the form

$$J_t(x) = -\alpha_t e^{-\gamma x}$$

   and find a difference equation for α_t.

8. Consider the following special case of Problem 2, where $r = 0$:

$$\max_{u_t \in [0,1]} \sum_{t=0}^{T}\sqrt{u_t x_t},\quad x_{t+1} = \rho(1 - u_t)x_t,\ t = 0, \dots, T-1,\ x_0 > 0$$

   (a) Compute $J_T(x)$, $J_{T-1}(x)$, $J_{T-2}(x)$. (*Hint:* Prove that $\max_{u \in [0,1]}\left[\sqrt{u} + A\sqrt{1-u}\right]$ $\sqrt{1 + A^2}$ with $u = 1/(1 + A^2)$.)

   (b) Show that the optimal control function is $u_s(x) = 1/(1 + \rho + \rho^2 + \cdots + \rho^{T-}$ and find the corresponding $J_s(x)$, $s = 1, 2, \dots, T$.

12.2 The Euler Equation

Economics literature sometimes considers the following "control variable free" formulat of the basic dynamic programming problem (e.g. Stokey et al. (1989))

$$\max \sum_{t=0}^{T} F(t, x_t, x_{t+1}), \quad x_0 \text{ given and } x_1, x_2, \dots, x_{T+1} \text{ vary freely in } \mathbb{R}$$

In this formulation the instantaneous reward $F(t, x_t, x_{t+1})$ at time t depends on t and the values of the state variable in the periods t and $t+1$.

If we define $u_t = x_{t+1}$, then (1) becomes a standard dynamic programming probl with $U = \mathbb{R}$. On the other hand, the dynamic optimization problem (12.1.2) can usu ly be formulated as a problem of the type (1). Suppose, in particular, that the equati $x_{t+1} = g(t, x_t, u_t)$ for each choice of x_t and x_{t+1} has a unique solution u_t in U, w $u_t = \varphi(t, x_t, x_{t+1})$. If we define the function F by $F(t, x_t, x_{t+1}) = f(t, x_t, \varphi(t, x_t, x_{t+}$ for $t < T$, and $F(T, x_T, x_{T+1}) = \max_{u \in U} f(T, x_T, u)$, then problem (12.1.2) is the sa as problem (1). (If there is more than one value of u such that $g(t, x_t, u) = x_{t+1}$, let u_t a value of u that maximizes $f(t, x_t, u)$, i.e. choose the best u that leads from x_t to x_t. Then, in any case, $F(t, x_t, x_{t+1}) = \max\{f(t, x_t, u) : u \in U, x_{t+1} = g(t, x_t, u)\}$.)

Let $\{x_0^*, \dots, x_{T+1}^*\}$ be an optimal solution to problem (1). For each given t, the derivat of the expression in (1) w.r.t. x_{t+1} must be zero. If we define $F(T+1, x_{T+1}, x_{T+2}) =$ then $\{x_0^*, \dots, x_{T+1}^*\}$ satisfies[3]

$$F_2'(t+1, x_{t+1}, x_{t+2}) + F_3'(t, x_t, x_{t+1}) = 0,\ t = 0, 1, \dots, T \quad \textbf{(Euler equation)}$$

This is a second-order difference equation analogous to the Euler equation in the classi calculus of variations. (See Section 8.2.) Note carefully that the partial derivatives in are evaluated at different triples.

[3] Only the two terms $F(t, x_t, x_{t+1}) + F(t+1, x_{t+1}, x_{t+2})$ in the sum in (1) depend on x_t Alternatively, we can require the derivative w.r.t. x_t to be 0. This gives the same equation.

A solution procedure for the Euler equation is as follows: First, for $t = T$, (2) reduces to $F_3'(T, x_T, x_{T+1}) = 0$. This equation is solved for x_{T+1}, yielding the function $x_{T+1} = x_{T+1}(x_T)$. Next, this function is inserted into (2) for $t = T-1$, and (2) is then solved for x_T, yielding $x_T = x_T(x_{T-1})$. Then this function is inserted into (2) for $t = T-2$ and (2) is then solved for x_{T-1} yielding the function $x_{T-1}(x_{T-2})$. In this manner we work backwards until the function $x_1(x_0)$ has been constructed. Since x_0 is given, the value of x_1 is determined, and then x_2 is determined, and so on.

EXAMPLE 1 Write down the Euler equation for the problem

$$\max \left[\sum_{t=0}^{2} [1 + x_t - (x_{t+1} - x_t)^2] + (1 + x_3) \right], \quad x_0 = 0,\ x_1, x_2, x_3 \in \mathbb{R} \qquad (*)$$

and find the solution of the problem. Show that the problem is equivalent to the problem in Example 12.1.2.

Solution: Define $F(t, x_t, x_{t+1}) = 1 + x_t - (x_{t+1} - x_t)^2$ for $t = 0$, 1, and 2, and let $F(3, x_3, x_4) = 1 + x_3$. Then the problem is of the type (1). For $t = 0, 1, 2$, we get $F_2'(t, x_t, x_{t+1}) = 1 + 2(x_{t+1} - x_t)$, and hence $F_2'(t+1, x_{t+1}, x_{t+2}) = 1 + 2(x_{t+2} - x_{t+1})$. Moreover, $F_3'(t, x_t, x_{t+1}) = -2(x_{t+1} - x_t)$, so that the Euler equation for $t = 0, 1$ becomes $1 + 2(x_{t+2} - x_{t+1}) - 2(x_{t+1} - x_t) = 0$, or

$$x_{t+2} - 2x_{t+1} + x_t = -\tfrac{1}{2}, \quad t = 0, 1 \qquad (**)$$

For $t = 2$ the Euler equation is $F_2'(3, x_3, x_4) + F_3'(2, x_2, x_3) = 0$. With $F(3, x_3, x_4) = 1 + x_3$, this gives $1 + (-2)(x_3 - x_2) = 0$, or

$$x_3 - x_2 = \tfrac{1}{2} \qquad (***)$$

Let us solve the problem backwards. As x_4 does not appear in the Euler equation for $t = 3$, there is nothing to determine as regards x_4. The equation $(***)$ gives $x_3 = 1/2 + x_2$. Inserting this into $(**)$ for $t = 1$, gives $1/2 + x_2 - 2x_2 + x_1 = -1/2$, i.e. $x_2 = x_1 + 1$. Inserting this into $(**)$ for $t = 0$, gives $x_1 + 1 - 2x_1 + x_0 = -1/2$, i.e. $x_1 = x_0 + 3/2$. Since $x_0 = 0$, then $x_1 = 3/2$, and so $x_2 = 5/2$ and $x_3 = 3$.

Look back at Example 12.1.2. From the difference equation there we obtained $u_t = x_{t+1} - x_t$, so that if we define $F(t, x_t, x_{t+1}) = 1 + x_t - (x_{t+1} - x_t)^2$ for $t = 0$, 1 and 2, and $F(3, x_3, x_4) = \max_{u \in \mathbb{R}}(1 + x_3 - u^2) = 1 + x_3$, the problem in Example 12.1.2 is equivalent to problem $(*)$. Note how the two approaches yield the same optimal solution.

PROBLEMS FOR SECTION 12.2

1. (a) Transform Problem 12.1.1 to the form (1).

   (b) Derive the corresponding Euler equation, and find its solution. Compare with the answer to Problem 12.1.1.

2. (a) Transform the problem in Example 12.1.3 to the form (1).

   (b) Derive the corresponding Euler equation, and find its solution. Compare with tl answer in Example 12.1.3.

12.3 Infinite Horizon

Economists often study dynamic optimization problems over an infinite horizon. This avoi specifying what happens after the finite horizon is reached. It also avoids having the horiz as an extra exogenous variable that features in the solution. This section considers hc dynamic programming methods can be used to study the following infinite horizon probler

$$\max \sum_{t=0}^{\infty} \beta^t f(x_t, u_t), \quad x_{t+1} = g(x_t, u_t), \quad t = 0, 1, 2, \ldots, \quad x_0 \text{ given}, \quad u_t \in U \subseteq \mathbb{R} \qquad ($$

Here f and g are given functions of two variables, $\beta \in (0, 1)$ is a constant discount fact and x_0 is a given number in $\mathbb{R}$. Having a discount factor $\beta < 1$ plays an important role our subsequent analysis of this problem.

The sequence pair $(\{x_t\}, \{u_t\})$ is called **admissible** provided $u_t \in U$, x_0, and the diffe ence equation in (1) is satisfied for all $t = 0, 1, 2, \ldots$. Note that neither f nor g depen explicitly on t. For this reason, problem (1) is called **autonomous**.

Assume that f satisfies the boundedness condition

$$M_1 \le f(x, u) \le M_2 \text{ for all } (x, u),\ u \in U, \text{ where } M_1 \text{ and } M_2 \text{ are given numbers} \qquad ($$

Because $0 < \beta < 1$, the sum in (1) will always converge. Take any control sequen $\pi = (u_s, u_{s+1}, \ldots)$, where $u_{s+k} \in U$ for $k = 0, 1, \ldots$, and let $x_{t+1} = g(x_t, u_t)$ for $t = s + 1, \ldots$, with $x_s = x$. The total utility (or benefit) obtained during the periods $s, s + \ldots$, is then

$$V_s(x, \pi) = \sum_{t=s}^{\infty} \beta^t f(x_t, u_t) = \beta^s V^s(x, \pi), \quad \text{where} \quad V^s(x, \pi) = \sum_{t=s}^{\infty} \beta^{t-s} f(x_t, u_t) \qquad ($$

Moreover, let

$$J_s(x) = \max_{\pi} V_s(x, \pi) = \beta^s J^s(x), \quad \text{where} \quad J^s(x) = \max_{\pi} V^s(x, \pi) \qquad ($$

and where the maximum is taken over all sequences $\pi = (u_s, u_{s+1}, \ldots)$ with $u_{s+k} \in U$ Thus, $J_s(x)$ is the maximum total utility (or benefit) that can be obtained in all the perio from $t = s$ to ∞, given that the system is in state x at $t = s$. We call $J_s(x)$ the **(optim value function** for problem (1).

[4] The existence of this maximum is discussed later in Note 2.

The function $J^s(x)$ satisfies the following property:

$$J^0(x) = J^s(x) \tag{5}$$

Intuitively, this equality is rather obvious. When maximizing $V^s(x, \pi)$ and $V^0(x, \pi)$, we obtain the same value in both cases, since the future looks exactly the same at time 0 as at time s. Equation (5) implies that

$$J_s(x) = \beta^s J^0(x), \quad s = 0, 1, \ldots \tag{6}$$

Define

$$J(x) = J_0(x) = J^0(x) \tag{7}$$

From (6) it follows that if we know $J_0(x) = J(x)$, then we know $J_s(x)$ for all s. The main result in this section is the following:

THEOREM 12.3.1 (FUNDAMENTAL EQUATION FOR INFINITE HORIZON)

The value function $J_0(x) = J(x)$ in (4) for problem (1) satisfies the equation

$$J(x) = \max_{u \in U}\big[f(x, u) + \beta J(g(x, u))\big] \qquad \textbf{(the Bellman equation)} \tag{8}$$

A rough argument for (8) resembles the argument for Theorem 12.1.1: Suppose we are in state x at time $t = 0$. If we choose the control u, the immediate reward is $\beta^0 f(x, u) = f(x, u)$, and at time $t = 1$ we end up in state $x_1 = g(x, u)$. Choosing an optimal control sequence from $t = 1$ on gives a total reward over all subsequent periods that equals $J_1(g(x, u)) = \beta J(g(x, u))$. Hence, the best choice of u at $t = 0$ is one that maximizes the sum $f(x, u) + \beta J(g(x, u))$. The maximum of this sum is therefore $J(x)$.

Equation (8) is a "functional equation", in which the unknown function $J(x)$ appears on both sides. *Under the boundedness condition (2) and the assumptions that the maximum in (8) is attained and that* $0 < \beta < 1$, *this equation always has one and only one bounded solution* $\hat{J}(x)$, *and this solution is automatically the optimal value function for the problem. The control* $u(x)$ *that maximizes the right-hand side of (8) is the optimal control, which is therefore independent of* t.

In general it is difficult to use equation (8) to find $J(x)$. The problem is that maximizing the right-hand side of (8) requires knowledge of the function $J(x)$.

EXAMPLE 1 Consider the following infinite horizon analogue of problem (i) in Example 12.1.4 in the case $a_t = a$, and where we have introduced a new control v defined by $u = vx$. The former constraint $u \in (0, x)$ is then replaced by $v \in (0, 1)$:

$$\max \sum_{t=0}^{\infty} \beta^t (x_t v_t)^{1-\gamma}, \qquad x_{t+1} = a(1 - v_t)x_t, \quad t = 0, 1, \ldots, \quad v_t \in (0, 1) \tag{i}$$

where a and x_0 are positive constants, $\beta \in (0, 1)$, $\gamma \in (0, 1)$, and $\beta a^{1-\gamma} < 1$. Bec the horizon is infinite, we may think of x_t as the assets of some timeless institution li university, corporation, or government.

In the notation of problem (1), $f(x, v) = (xv)^{1-\gamma}$ and $g(x, v) = a(1 - v)x$. $J_0(x) = J(x)$, equation (8) yields

$$J(x) = \max_{v \in (0,1)}\big[(xv)^{1-\gamma} + \beta J(a(1 - v)x)\big]$$

In the closely related problem in Example 12.1.4, the value function was proportiona $x^{1-\gamma}$. A reasonable guess in the present case is that $J(x) = kx^{1-\gamma}$ for some constant k. try this as a solution. Then, cancelling the factor $x^{1-\gamma}$, (ii) reduces to

$$k = \max_{v \in (0,1)}\big[v^{1-\gamma} + \beta k a^{1-\gamma}(1 - v)^{1-\gamma}\big]$$

Put $\varphi(v) = v^{1-\gamma} + \beta k a^{1-\gamma}(1 - v)^{1-\gamma}$. Then the first-order condition is

$$\varphi'(v) = (1 - \gamma)v^{-\gamma} - \beta(1 - \gamma)k a^{1-\gamma}(1 - v)^{-\gamma} = 0$$

implying that $v^{-\gamma} = \beta k a^{1-\gamma}(1 - v)^{-\gamma}$. Raising each side to the power $-1/\gamma$ and solv for v yields

$$v = \frac{1}{1 + \rho k^{1/\gamma}}, \quad \text{where } \rho = \big(\beta a^{1-\gamma}\big)^{1/\gamma}$$

Note that $v \in (0, 1)$ and it is easy to verify that $\varphi(v)$ is concave. Thus we have shown t if $J(x) = kx^{1-\gamma}$, then the value of v that solves the maximization problem in (iii) is gi by (iv). Then equation (iii) implies that k satisfies the equation

$$k = \frac{1}{(1 + \rho k^{1/\gamma})^{1-\gamma}} + \beta k a^{1-\gamma}\frac{\rho^{1-\gamma}k^{(1-\gamma)/\gamma}}{(1 + \rho k^{1/\gamma})^{1-\gamma}}$$

Recalling that $\beta a^{1-\gamma} = \rho^{\gamma}$, simple algebra reduces this equation to

$$k = \big(1 + \rho k^{1/\gamma}\big)^{\gamma-1}\big[1 + k\rho^{\gamma}\rho^{1-\gamma}k^{(1-\gamma)/\gamma}\big] = \big(1 + \rho k^{1/\gamma}\big)^{\gamma}$$

Raise each side to the power $1/\gamma$, and solve for $k^{1/\gamma}$ to obtain $k^{1/\gamma} = 1/(1 - \rho)$, $k = (1 - \rho)^{-\gamma}$. Then (iv) implies $v = 1 - \rho$. Because $J(x) = kx^{1-\gamma}$, we have

$$J(x) = (1 - \rho)^{-\gamma}x^{1-\gamma}, \quad \text{with} \quad v = 1 - \rho, \quad \rho = (\beta a^{1-\gamma})^{1/\gamma}$$

In this example the boundedness assumption (2) is not valid until one makes a simp transformation of the problem. Define the new state variable $y_t = x_t/a^t$. Then y_t satisfi the equation $y_{t+1} = (1 - v_t)y_t$. The objective function is now $\sum_{t=0}^{\infty}\hat{\beta}^t(y_t v_t)^{1-\gamma}$, whe $\hat{\beta} = \beta a^{1-\gamma}$ and so $0 < \hat{\beta} < 1$. It is easy to verify that the function $\hat{J}(y) = J(ay)$ satisfi the Bellman equation for this new problem (the optimal v is the same). In the new proble the condition in Note 3 below is satisfied with $\bigcup_t \mathcal{X}_t(x_0) \subseteq (0, y_0)$ because the state remains within the interval $(0, y_0)$ for all t, and so $0 < (y_t v_t)^{1-\gamma} < y_0^{1-\gamma}$ for all t and all in (0, 1). Therefore v defined in (iv) *is* optimal and the problem is solved. According to (v), the optimal v_t is constant, $v_t = 1 - \rho$. The corresponding optimal x_t satisfies the difference equation $x_{t+1} = a(1 - v_t)x_t = a\rho x_t$. With $x_0 = x$, the solution is $x_t = x(a\rho)^t$. The value of the objective function in (i) is then

$$\sum_{t=0}^{\infty} \beta^t \big(x(a\rho)^t\big)^{1-\gamma} (1-\rho)^{1-\gamma} = (1-\rho)^{1-\gamma} x^{1-\gamma} \sum_{t=0}^{\infty} \big(\beta(a\rho)^{1-\gamma}\big)^t$$
$$= (1-\rho)^{1-\gamma} x^{1-\gamma} \sum_{t=0}^{\infty} \rho^t = (1-\rho)^{-\gamma} x^{1-\gamma}$$

where we have used the fact that $\beta(a\rho)^{1-\gamma} = (\beta a^{1-\gamma})\rho^{1-\gamma} = \rho^{\gamma}\rho^{1-\gamma} = \rho$ and that $\sum_{t=0}^{\infty} \rho^t = 1/(1-\rho)$. The value of the objective function is therefore precisely equal to $J(x)$, as given by (v).

NOTE 1 As pointed out in Note 12.1.6, the same theory applies without change when x_t, u_t, and g are vector functions. Moreover U may depend on the state, $U = U(x)$ (but not explicitly on time).

NOTE 2 Whenever we wrote "max" above, it was implicitly assumed that the maximum exists. Of course, without further conditions on the system, this may not be true. Under assumptions (2), the same conditions as in the finite horizon case (f and g are continuous and U is compact) ensure that the maxima in (4) and (8) do exist. Meanwhile, we prove that (8) has a unique solution, which must therefore be the optimal value function. This is done using the result in Section A.4 on iterated suprema.

$$J_0(x_0) = \sup_{u_0, u_1, \ldots} \sum_{t=0}^{\infty} \beta^t f(x_t, u_t) = \sup_{u_0 \in U} \Big[f(x_0, u_0) + \sup_{u_1, u_2, \ldots} \sum_{t=1}^{\infty} \beta^t f(x_t, u_t)\Big]$$
$$= \sup_{u_0 \in U} \big[f(x_0, u_0) + J_1(g(x_0, u_0))\big] = \sup_{u_0 \in U} \big[f(x_0, u_0) + \beta J_0(g(x_0, u_0))\big]$$

Next, the contraction mapping theorem 14.3.1 is used to prove that the Bellman equation (8) has a unique solution. Indeed, define the operator T on the space of bounded functions $I(x)$ so that $T(I)(x) = \sup_u[f(x, u) + \beta I(g(x, u))]$ for all x. For any bounded functions $\tilde{J}$ and $\bar{J}$, define $d(\tilde{J}, \bar{J}) = \sup_z |\tilde{J}(z) - \bar{J}(z)|$. Then

$$T(\tilde{J})(x) = \sup_u\big[f(x, u) + \beta \bar{J}(g(x, u)) + \beta\big(\tilde{J}(g(x, u)) - \bar{J}(g(x, u))\big)\big]$$
$$\le \sup_u\big[f(x, u) + \beta \bar{J}(g(x, u)) + \beta d(\tilde{J}, \bar{J})\big] = T(\bar{J})(x) + \beta d(\tilde{J}, \bar{J})$$

Symmetrically, $T(\bar{J})(x) \le T(\tilde{J})(x) + \beta d(\tilde{J}, \bar{J})$. So $|T(\tilde{J})(x) - T(\bar{J})(x)| \le \beta d(\tilde{J}, \bar{J})$ for all x. This verifies that T is a contraction mapping, and the proof is complete.

Finally, it is easily seen that the control $u = u(x)$ yielding maximum in the Bellman equation is optimal: Defining $J^u(x_0) = \sum_{t=0}^{\infty} f(x_t, u(x_t))$, we have

$$J^u(x_0) = f(x_0, u(x_0)) + \beta \sum_{t=1}^{\infty} \beta^{t-1} f(x_t, u(x_t)) = f(x_0, u(x_0)) + \beta J^u(g(x_0, u(x_0)))$$

(The sum from $t = 1$ to infinity, is similar to the sum defining J^u, but the sequence x_t in the former sum starts at $g(x_0, u(x_0))$, hence the term $J^u(g(x_0, u(x_0)))$.) Since also $J(x_0) = f(x_0, u(x_0)) +$

$\beta J(g(x_0, u(x_0)))$, then from the last equalities, we get $J^u(x_0) - J(x_0) = \beta[J^u(x_0) - J(x_0)]$, impl
$J^u(x_0) - J(x_0) = 0$.

Suppose we replace max with sup in (12.1.3), (12.1.5), (12.1.6), (12.1.7), (12.3.4), and (12.
Then equations (12.1.6), (12.1.7), and (12.3.8) still hold, even if no maximum exists. More
(12.3.8) still has a unique solution, which is the optimal value function. Only if the suprem
(12.1.6), (12.1.7), and (12.3.8) are attained by a closed-loop control do optimal controls exist,
these optimal controls are those closed-loop controls.

NOTE 3 It suffices to assume that the boundedness condition (2) holds for all x in $\mathcal{X}(x_0$
$\bigcup_{t=0}^{\infty} \mathcal{X}_t(x_0)$, where $\mathcal{X}_t(x_0)$ is defined in Note 12.1.3.

PROBLEMS FOR SECTION 12.3

1. Consider the following problem with $\beta \in (0, 1)$:

$$\max_{u_t \in (-\infty, \infty)} \sum_{t=0}^{\infty} \beta^t \left(-\tfrac{2}{3}x_t^2 - u_t^2\right), \qquad x_{t+1} = x_t + u_t, \quad t = 0, 1, \ldots, \quad x_0 \text{ give}$$

   (a) Suppose that $J(x) = -\alpha x^2$. Find a third degree equation for α. Find the associ
   value of u^*. (Disregard condition (2).)

   (b) Given a start value x_0. By looking at the objective function, show that it is reason
   to assume that $|x_t| \le |x_{t-1}|$ and that $u_t \le |x_{t-1}|$. Does (2) then apply?

2. Consider the problem

$$\max_{u_t \in (0, \infty)} \left[\sum_{t=0}^{\infty} \beta^t \left(-e^{-u_t} - \tfrac{1}{2}e^{-x_t}\right)\right], \quad x_{t+1} = 2x_t - u_t, \; t = 0, 1, \ldots, \; x_0 \text{ is give}$$

   where $\beta \in (0, 1)$. Suppose that $J(x) = -\alpha e^{-x}, \alpha > 0$. Determine α. Disreg
   condition (2).

12.4 The Maximum Principle

Dynamic programming is the most frequently used method for solving discrete time dyna
optimization problems. An alternative solution technique is based on the so called maxim
principle. The actual calculations needed are often rather similar. However, when there
terminal restrictions on the state variables, the maximum principle is often preferable. T
corresponding principle for optimization problems in continuous time is studied in m
detail in Chapters 9 and 10, because for such problems it is the most important method.

Consider first the discrete time dynamic optimization problem with one state, one control variable and a free right-hand side:

$$\max_{u_t \in U \subseteq \mathbb{R}} \sum_{t=0}^{T} f(t, x_t, u_t), \; x_{t+1} = g(t, x_t, u_t), \; t = 0, \ldots, T-1, \; x_0 \text{ is given}, \; x_T \text{ free} \quad (1)$$

Here we assume that the control region U is convex, i.e. an interval. The state variable x_t evolves from the initial state x_0 according to the law of motion (1), with u_t as a control that is chosen at each $t = 0, \ldots, T$. Define the **Hamiltonian** by

$$H(t, x, u, p) = \begin{cases} f(t, x, u) + pg(t, x, u) & \text{for } t < T \\ f(t, x, u) & \text{for } t = T \end{cases} \quad (2)$$

where p is called an **adjoint function (or co-state variable)**.

THEOREM 12.4.1 (THE MAXIMUM PRINCIPLE. NECESSARY CONDITIONS)

Suppose $(\{x_t^*\}, \{u_t^*\})$ is an optimal sequence pair for problem (1), and let H be defined by (2). Then there exist numbers p_t, with $p_T = 0$, such that for all $t = 0, \ldots, T$,

$$H'_u(t, x_t^*, u_t^*, p_t)(u - u_t^*) \leq 0 \quad \text{for all } u \in U \quad (3)$$

(Note that if u_t^* is an interior point of U, (3) implies that $H'_u(t, x_t^*, u_t^*, p_t) = 0$.) Furthermore, p_t is a solution to the difference equation

$$p_{t-1} = H'_x(t, x_t^*, u_t^*, p_t), \quad t = 1, \ldots, T \quad (4)$$

NOTE 1 For a proof see Arkin and Evstigneev (1987). A closer analogy with the continuous time maximum principle comes from writing the equation of motion as $x_{t+1} - x_t = g(t, x_t, u_t)$. If we redefine the Hamiltonian accordingly, then (4) is replaced by $p_t - p_{t-1} = -H'_x(t, x_t^*, u_t^*, p_t)$, which corresponds to equation (9.2.5).

Sufficient conditions are given in following theorem. The proof is similar to the proof of the corresponding theorem in continuous time.

THEOREM 12.4.2 (SUFFICIENT CONDITIONS)

Suppose that the sequence triple $(\{x_t^*\}, \{u_t^*\}, \{p_t\})$ satisfies all the conditions in Theorem 12.4.1, and suppose further that $H(t, x, u, p_t)$ is concave with respect to (x, u) for every t. Then the sequence triple $(\{x_t^*\}, \{u_t^*\}, \{p_t\})$ is optimal.

NOTE 2 Suppose that admissible pairs are also required to satisfy $(x_t, u_t) \in A_t$, $t = 0, \ldots, T$, where A_t is a convex set for all t. Then Theorem 12.4.2 is still valid, and H need only be concave in A_t.

NOTE 3 If U is compact and f and g are continuous, there will always exist an optir solution. (This result can be proved by using the extreme value theorem.)

EXAMPLE 1 Apply Theorem 12.4.2 to the problem in Example 12.1.2,

$$\max \sum_{t=0}^{3} (1 + x_t - u_t^2), \quad x_{t+1} = x_t + u_t, \; x_0 = 0, \; t = 0, 1, 2, \; u_t \in \mathbb{R}$$

Solution: For $t < 3$, the Hamiltonian is $H = 1 + x - u^2 + p(x + u)$, so $H'_u = -2u +$ and $H'_x = 1 + p$. For $t = 3$, $H = 1 + x - u^2$, so $H'_u = -2u$ and $H'_x = 1$. Note that t Hamiltonian is concave in (x, u). The control region is open, so (3) implies that $(H'_u)^* =$ i.e. $-2u_t^* - p_t = 0$ for $t = 0, 1, 2$, and $-2u_3^* = 0$ for $t = 3$. Thus $u_0^* = \frac{1}{2}p_0$, $u_1^* = \frac{1}{2}$ and $u_2^* = \frac{1}{2}p_2$,

The difference equation (4) for p_t is $p_{t-1} = 1 + p_t$ for $t = 1, 2$, and so $p_0 = 1 +$ $p_1 = 1 + p_2$. Moreover, (5) yields $p_2 = 1 + p_3$, and because x_3 is free, $p_3 = 0$. It follo that $p_2 = 1$, $p_1 = 1 + p_2 = 2$, and $p_0 = 1 + p_1 = 3$. This results in the following optir choices for the controls, $u_0^* = 3/2$, $u_1^* = 1$, $u_2^* = 1/2$, and $u_3^* = 0$, which is the same resu as in Example 12.1.2.

EXAMPLE 2 Consider an oil field in which $x_0 > 0$ units of extractable oil remain at time $t =$ Let $u_t \geq 0$ be the quantity of oil extracted in period t, and let x_t be the remaining sto at time t. Then $u_t = x_t - x_{t+1}$. Let $C(t, x_t, u_t)$ denote the cost of extracting u_t units period t when the stock is x_t. Let p be the price per unit of oil and let r be the discount rat with $\beta = 1/(1 + r) \in (0, 1)$ the corresponding discount factor. If T is the fixed end of t planning period, the problem of maximizing total discounted profit can be written as

$$\max_{u_t \geq 0} \sum_{t=0}^{T} \beta^t [pu_t - C(t, x_t, u_t)], \quad t = 0, 1, \ldots, T-1, \quad x_{t+1} = x_t - u_t, \quad x_0 > 0 \quad ($$

assuming also that

$$u_t \leq x_t, \qquad t = 0, 1, \ldots, T \qquad (\text{i}$$

because the amount extracted cannot exceed the stock.

Because of restriction (ii), this is not a dynamic optimization problem of the type de scribed by (1). However, if we define a new control v_t by $u_t = v_t x_t$, then restriction (i combined with $u_t \geq 0$ reduces to the control restriction $v_t \in [0, 1]$, and we have a stanc ard dynamic optimization problem. Assuming that $C(t, x, u) = u^2/x$, $0 < p < 1$, an $\beta \in (0, 1)$, apply the maximum principle to the problem

$$\max_{v_t \in [0,1]} \sum_{t=0}^{T} \beta^t (pv_t x_t - v_t^2 x_t), \quad x_{t+1} = x_t(1 - v_t), \quad x_0 > 0, \quad v_t \in [0, 1] \qquad (\text{iii}$$

with x_T free.

Solution: We denote the adjoint function by λ_t. We know that $\lambda_T = 0$. The Hamiltonian is $H = \beta^t(pvx - v^2x) + \lambda x(1-v)$. (This is valid also for $t = T$, because then $\lambda = \lambda_T = 0$.) Then $H'_v = \beta^t(px - 2vx) - \lambda x$ and $H'_x = \beta^t(pv - v^2) + \lambda(1 - v)$. So (3) implies that, for $(\{x_t^*\}, \{v_t^*\})$ to solve the problem, there must exist numbers λ_t, with $\lambda_T = 0$, such that, for all $t = 0, \dots, T$,

$$[\beta^t x_t^*(p - 2v_t^*) - \lambda_t x_t^*](v - v_t^*) \le 0 \quad \text{for all } v \text{ in } [0, 1] \tag{iv}$$

For $t = T$, with $\lambda_T = 0$, this condition reduces to

$$\beta^T x_T^*(p - 2v_T^*)(v - v_T^*) \le 0 \quad \text{for all } v \text{ in } [0, 1] \tag{v}$$

Having $v_T^* = 0$ would imply that $pv \le 0$ for all v in $[0, 1]$, which is impossible because $p > 0$. Suppose instead that $v_T^* = 1$. Then (v) reduces to $\beta^T x_T^*(p - 2)(v - 1) \le 0$ for all v in $[0, 1]$, which is impossible because $p - 2 < 0$ (put $v = 0$). Hence, $v_T^* \in (0, 1)$. For $t = T$, condition (v) then reduces to $\beta^T x_T^*(p - 2v_T^*) = 0$, and so

$$v_T^* = \tfrac{1}{2}p \tag{vi}$$

According to (4), for $t = 1, \dots, T$,

$$\lambda_{t-1} = \beta^t v_t^*(p - v_t^*) + \lambda_t(1 - v_t^*) \tag{vii}$$

For $t = T$, because $\lambda_T = 0$ and $v_T^* = \frac{1}{2}p$, this equation reduces to

$$\lambda_{T-1} = \beta^T v_T^*(p - v_T^*) = \tfrac{1}{4}p^2\beta^T \tag{viii}$$

For $t = T - 1$, the term within square brackets in (iv) is

$$\beta^{T-1}x_{T-1}^*(p - 2v_{T-1}^*) - \lambda_{T-1}x_{T-1}^* = \beta^{T-1}x_{T-1}^*[p(1 - \tfrac{1}{4}\beta p) - 2v_{T-1}^*] \tag{ix}$$

Because $0 < p < 1$ and $\beta \in (0, 1)$, one has $1 > \frac{1}{4}\beta p$. It follows that both $v_{T-1}^* = 0$ and $v_{T-1}^* = 1$ are impossible as optimal choices in (iv), so $v_{T-1}^* \in (0, 1)$ can only be the maximizer in (iv) provided the square bracket in the last line of (ix) is 0. Hence

$$v_{T-1}^* = \tfrac{1}{2}p(1 - \tfrac{1}{4}\beta p)$$

Let us now go k periods backwards in time. Define $q_{T-k} = \lambda_{T-k}/\beta^{T-k}$. We prove by backward induction that at each time $T - k$ we have an interior maximum point v_{T-k}^* in (iv). Then $v_{T-k}^* = \frac{1}{2}(p - q_{T-k})$, which belongs to (0,1) if $q_{T-k} \in (2 - p, p)$. Using (vii) and the definition of q_{T-k}, we find that $q_{T-(k+1)} = F(q_{T-k})$ where $F(q) = \beta\big[\frac{1}{4}(p - q)^2 + q\big] \ge 0$. Note that $q \mapsto F(q)$ is a strictly convex function, and by the assumptions on the parameters, we have $0 < F(q) \le \max\{F(0), F(p)\} = \max\{\beta p^2/4, \beta p\} < p$ for all q in $[0, p]$. Because $q_T = 0$, it follows that q_{T-k} does belong to $(0, p)$ for all $k \ge 1$. Thus the solution of the problem is given by $v_{T-k}^* = (1/2)(p - q_{T-k})$, where q_{T-k} is determined by $q_{T-(k+1)} = \beta^{-(T-k)}\lambda_{T-(k+1)} = F(q_{T-k})$, with $q_T = 0$.

PROBLEMS FOR SECTION 12.4

1. Consider Problem 12.1.1.

   (a) Write down the Hamiltonian, condition (3), and the difference equation for p_t.

   (b) Use the maximum principle to find a unique solution candidate.

   (c) Solve the problem by using Theorem 12.4.2.

2. (Boltyanski) Consider the problem

$$\max_{u_t \in [-1,1]} \sum_{t=0}^{T}(u_t^2 - 2x_t^2) \quad \text{s.t.} \quad x_{t+1} = u_t, \quad t = 0, 1, \dots, T-1, \quad x_0 = 0$$

   (a) Prove that $u_t^* = 0$ for $t = 0, 1, \dots, T-1$, and $u_T^* = 1$ (or -1) are optimal contro
   (Express the objective function as a function of $u_0, u_1, \dots, u_T$ only.)

   (b) Verify that the conditions in Theorem 12.4.2 are satisfied.

   (c) Verify that u_t^* does not maximize $H(t, x_t^*, u, p_t)$ for $u \in [-1, 1]$.

12.5 More Variables

Consider the following problem with n state and r control variables:

$$\max \sum_{t=0}^{T} f(t, \mathbf{x}_t, \mathbf{u}_t), \quad \mathbf{x}_{t+1} = \mathbf{g}(t, \mathbf{x}_t, \mathbf{u}_t), \quad \mathbf{x}_0 \text{ is given}, \quad u_t \in U \subseteq \mathbb{R}^r \qquad (1$$

Here $\mathbf{x}_t$ is a state variable in $\mathbb{R}^n$ that evolves from the initial state $\mathbf{x}_0$ according to the law o motion (1), with $\mathbf{u}_t$ as a control that is chosen at each $t = 0, \dots, T$. We put $\mathbf{x}_t = (x_t^1, \dots, x_t^n$ $\mathbf{u}_t = (u_t^1, \dots, u_t^r)$, and $\mathbf{g} = (g^1, \dots, g^n)$. We assume that the control region U is convex.

The terminal conditions are

$$\begin{aligned} &\text{(a)} \quad x_T^i = \bar{x}^i \quad \text{for } i = 1, \dots, l \\ &\text{(b)} \quad x_T^i \ge \bar{x}^i \quad \text{for } i = l+1, \dots, m \\ &\text{(c)} \quad x_T^i \text{ free} \quad \text{for } i = m+1, \dots, n \end{aligned} \qquad (2$$

Define the **Hamiltonian** by

$$H(t, \mathbf{x}, \mathbf{u}, \mathbf{p}) = \begin{cases} q_0 f(t, \mathbf{x}, \mathbf{u}) + \sum_{i=1}^{n} p^i g_i(t, \mathbf{x}, \mathbf{u}) & \text{for } t < T \\ f(t, \mathbf{x}, \mathbf{u}) & \text{for } t = T \end{cases}$$

where $\mathbf{p} = (p^1, \dots, p^n)$ is called an **adjoint function (or co-state variable)**. (For a proo of the following theorem, see Arkin and Evstigneev (1987).

THEOREM 12.5.1 (THE MAXIMUM PRINCIPLE AND SUFFICIENCY)

Suppose that $(\{\mathbf{x}_t^*\}, \{\mathbf{u}_t^*\})$ is an optimal path for problem (1)–(2). Then there exist vectors $\mathbf{p}_t$ in $\mathbb{R}^n$ and a number q_0, with $(q_0, \mathbf{p}_T) \neq (0, \mathbf{0})$ and with $q_0 = 0$ or 1, such that for $t = 0, \dots, T$,

$$\sum_{i=1}^{r} H'_{u^i}(t, \mathbf{x}_t^*, \mathbf{u}_t^*, \mathbf{p}_t)\,(u_t^i - (u_t^*)^i) \leq 0 \quad \text{for all } \mathbf{u} \in U \tag{3}$$

Also, the vector $\mathbf{p}_t = (p_t^1, \dots, p_t^n)$ is a solution of

$$p_{t-1}^i = H'_{x^i}(t, \mathbf{x}_t^*, \mathbf{u}_t^*, \mathbf{p}_t), \qquad t = 1, \dots, T-1 \tag{4}$$

Moreover,

$$p_{T-1}^i = q_0 \frac{\partial f(T, \mathbf{x}_T^*, \mathbf{u}_T^*)}{\partial x_T^i} + p_T^i \tag{5}$$

where the vector $\mathbf{p}_T = (p_T^1, \dots, p_T^n)$ satisfies

$$\begin{array}{lll} (a') & p_T^i \text{ no conditions} & i = 1, \dots, l \\ (b') & p_T^i \geq 0 \ (p_T^i = 0 \text{ if } x_T^{*i} > \bar{x}^i) & i = l+1, \dots, m \\ (c') & p_T^i = 0 & i = m+1, \dots, n \end{array} \tag{6}$$

If the conditions above are satisfied with $q_0 = 1$ and $H(t, \mathbf{x}, \mathbf{u}, \mathbf{p})$ is concave in $(\mathbf{x}, \mathbf{u})$, then $(\{\mathbf{x}_t^*\}, \{\mathbf{u}_t^*\})$ is optimal.

NOTE 1 If $m = 0$ (so that there are no restrictions on the terminal state $\mathbf{x}_T$), then $\mathbf{p}_T = \mathbf{0}$ and it follows from Theorem 12.5.1 that $q_0 = 1$.

NOTE 2 If $\mathbf{u}_t^*$ is an interior point of U, then (3) implies that $H'_{u^i}(t, \mathbf{x}_t^*, \mathbf{u}_t^*, \mathbf{p}_t) = 0$ for all $i = 1, \dots, r$.

Infinite Horizon

We consider briefly the following infinite horizon version of problem (1)–(2),

$$\max \sum_{t=0}^{\infty} f(t, \mathbf{x}_t, \mathbf{u}_t), \qquad \mathbf{x}_t \in \mathbb{R}^n, \quad \mathbf{u}_t \in U \subseteq \mathbb{R}^r, \quad U \text{ convex} \tag{7}$$

where we maximize over all sequences $(\mathbf{x}_t, \mathbf{u}_t)$ satisfying

$$\mathbf{x}_{t+1} = \mathbf{g}(t, \mathbf{x}_t, \mathbf{u}_t), \quad t = 1, 2, \dots, \qquad \mathbf{x}_0 \text{ given} \tag{8}$$

and the terminal conditions

$$\begin{array}{lll} (a) & \lim_{T\to\infty} x_i(T) = \hat{x}_i, & i = 1, \dots, m' \\ (b) & \underline{\lim}_{T\to\infty} x_i(T) \geq \hat{x}_i, & i = m'+1, \dots, m \end{array} \tag{9}$$

Note that f and $\mathbf{g} = (g_1, \dots, g_n)$ can now depend explicitly on t. Assume that the su
(7) exists for all admissible sequences. The functions f and $\mathbf{g}$ are assumed to be C^1
respect to all x_i and u_j.

We do no more that state a sufficient condition for such problems:[6]

THEOREM 12.5.2 (SUFFICIENT CONDITIONS)

Suppose that the sequence $(\{\mathbf{x}_t^*\}, \{\mathbf{u}_t^*\}, \{\mathbf{p}_t\})$ satisfies the conditions (3)–(6) and for $q_0 = 1$. Suppose further that the Hamiltonian $H(t, \mathbf{x}, \mathbf{u}, \mathbf{p}_t)$ is concave in $(\mathbf{x}, \mathbf{u})$ for every t. Then $(\{\mathbf{x}_t^*\}, \{\mathbf{u}_t^*\})$ is optimal provided that the following transversality condition is satisfied: For all admissible sequences $(\{\mathbf{x}_t\}, \{\mathbf{u}_t\})$,

$$\underline{\lim}_{t\to\infty} \mathbf{p}_t(\mathbf{x}_t - \mathbf{x}_t^*) \geq 0$$

NOTE 3 Suppose there are additional conditions for a sequence $\{\mathbf{x}_t, \mathbf{u}_t\}$ to be admissi
Then (10) needs only to be tested for such sequences.

PROBLEMS FOR SECTION 12.5

1. Consider the problem

$$\max_{u,v\in\mathbb{R}} \sum_{t=0}^{2} [1 + x_t - y_t - 2u_t^2 - v_t^2] \quad \text{s.t.} \quad \begin{cases} x_{t+1} = x_t - u_t, & x_0 = 5 \\ y_{t+1} = y_t + v_t, & y_0 = 2 \end{cases}, \quad t = 0,$$

(a) Solve the problem by using the difference equations to express the objective funct
I as a function only of u_0, u_1, u_2, v_0, v_1, and v_2, and then optimize.

(b) Solve the problem by using dynamic programming. (Find $J_2(x, y)$, $J_1(x, y)$,
$J_0(x, y)$ and the corresponding optimal controls.)

(c) Solve the problem by using Theorem 12.5.1.

2. Solve the problem

$$\max \sum_{t=0}^{T} (-x_t^2 - u_t^2) \quad \text{subject to} \quad x_{t+1} = y_t, \; y_{t+1} = y_t + u_t, \; t = 0, 1, \dots, T -$$

where $x_0 = x^0$ and $y_0 = y^0$ are given numbers and $u_t \in \mathbb{R}$.

[6] For the definition of $\underline{\lim}$ see Section 10.3.

3. Solve the problem

$$\max \sum_{t=0}^{\infty} \beta^t (x_t - u_t)] \quad \text{subject to} \quad x_{t+1} = u_t, \quad x_0 > 0, \quad u_t > 0$$

where $\beta \in (0, 1)$. Verify that $x_t^* > u_t^*$ for all t.

12.6 Stochastic Optimization

What is the best way of controlling a dynamic system subject to random disturbances? Stochastic dynamic programming is a central tool for tackling this problem.

In deterministic dynamic programming the state develops according to a difference equation $\mathbf{x}_{t+1} = \mathbf{g}(t, \mathbf{x}_t, \mathbf{u}_t)$, controlled by appropriate choices of the control variables $\mathbf{u}_t$. In this section, the state $\mathbf{x}_t$ is influenced by random disturbances, so that $\mathbf{x}_t$ is a stochastic variable. Following common practice, we often use capital letters instead of lower case letters for stochastic quantities, e.g. $\mathbf{X}_t$ instead of $\mathbf{x}_t$. We assume that $\mathbf{x}_t$ belongs to $\mathbb{R}^n$, that $\mathbf{u}_t$ is required to belong to a given subset U of $\mathbb{R}^r$, and that $t = 0, \dots, T$.

Suppose now that the state equation takes the new form

$$\mathbf{X}_{t+1} = \mathbf{g}(t, \mathbf{X}_t, \mathbf{u}_t, \mathbf{V}_{t+1}), \quad \mathbf{X}_0 = \mathbf{x}_0, \mathbf{V}_0 = \mathbf{v}_0, \text{ with } \mathbf{x}_0 \text{ and } \mathbf{v}_0 \text{ given}, \quad \mathbf{u}_t \in U \tag{1}$$

We consider two cases. In the first, $\mathbf{V}_{t+1}$ is a random variable that takes values in a finite set $\mathcal{V}$. It is assumed that the probability that $\mathbf{V}_{t+1} = \mathbf{v} \in \mathcal{V}$ may depend on the outcome $\mathbf{v}_t$ at time t, as well as explicitly on time t. Then we consider the **conditional** probability that $\mathbf{V}_{t+1} = v$, given v_t, which is denoted by $P_t(v|v_t)$. In the second case, $\mathbf{V}_{t+1}$ may take values anywhere in a Euclidean space. Then the distribution of $\mathbf{V}_{t+1}$ is assumed to be described by a conditional density $p_t(\mathbf{v}|\mathbf{v}_t)$ that is a continuous function of $\mathbf{v}$ and $\mathbf{v}_t$ together.

EXAMPLE 1 Suppose that $Z_1, Z_2, \dots$ are independently distributed stochastic variables which take positive values with specified probabilities independent of both the state and the control. Thus, at each time $t = 0, \dots, T$, either there is a discrete distribution $P_t(Z_t)$, or a continuous density function $p_t(z_t)$. The state X_t is assumed to evolve according to the stochastic difference equation

$$X_{t+1} = Z_{t+1}(X_t - u_t), \qquad u_t \in [0, \infty) \tag{i}$$

Here u_t is consumption, $X_t - u_t$ is investment, and Z_{t+1} is the return per invested dollar. Moreover, the utility of the terminal state x_T is $S(T, x_T, u) = \beta^T B x_T^{1-\gamma}$ and the utility of the current consumption is $\beta^t u_t^{1-\gamma}$ for $t < T$, where β is a discount factor and $0 < \gamma < 1$. The paths of the state x_t and of the control u_t are now uncertain (stochastic). The objective function to be maximized is the sum of expected discounted utility, given by

$$E\Big[\sum_{t=0}^{T-1} \beta^t u_t^{1-\gamma} + \beta^T B X_T^{1-\gamma}\Big] \tag{ii}$$

This problem will be studied in Example 3. In the discrete variable case, the expec will be a sum, but in the continuous variable case it will be an integral.

Consider first a two-stage decision problem with one state and one control variable. As that one wants to maximize the objective function

$$E\big[f(0, X_0, u_0) + f(1, X_1, u_1)\big] = f(0, X_0, u_0) + Ef(1, X_1, u_1)$$

where E denotes expectation and f is some given function. Here the initial state X_0 and an initial outcome v_0 are given, while X_1 is determined by the difference equatio i.e. $X_1 = g(0, x_0, u_0, V_1)$. We can find the maximum by first maximizing with resp u_1, and then with respect to u_0. When choosing u_1, we simply maximize $f(1, X_1$ assuming that X_1 is known before the maximization is carried out. The maximum u_1^* becomes a function $u_1^*(X_1)$ of X_1. Insert this function instead of u_1 in the obje function, and replace the two occurrences of X_1 by $g(0, x_0, u_0, V_1)$. Then u_0 occurs in terms of the objective function. A maximizing value of u_0 is then chosen, taking both occurrences into account.

To see why it matters that we can observe X_1 before choosing u_1, the following exa is illuminating: Consider the simple two stage decision problem with $f(0, X_0, u_0)$ $f(1, X_1, u_1) = X_1 u_1$, and $X_1 = V_1$, where V_1 takes the values 1 and -1 with probabi $1/2$, and where u must equal one of the two values 1 and -1. Then $E[X_1 u_1] = 0$ have to choose u_1 before observing X_1, hence a constant u_1. But if we can first observ then we can let u_1 depend on X_1. If we choose $u_1 = u_1(X_1) = X_1$, then $E[X_1 u_1]$ which yields a better value of the objective. In all that follows we shall assume that fact, both X_t and V_t), can be observed before choosing u_t.

Let us turn to the general problem. The process determined by (1) and the values c random variables $\mathbf{V}_1, \mathbf{V}_2, \dots$ is to be controlled in the best possible manner by approp choices of the variables $\mathbf{u}_t$. The objective function is now the expectation

$$E\Big[\sum_{t=0}^{T} f(t, \mathbf{X}_t, \mathbf{u}_t(\mathbf{X}_t, \mathbf{V}_t))\Big]$$

Here several things have to be explained. Each control $\mathbf{u}_t$, $t = 0, 1, 2, \dots, T$ shoul a function $\mathbf{u}_t(\mathbf{x}_t, \mathbf{v}_t)$ of the current state $\mathbf{x}_t$ and the outcome $\mathbf{v}_t$. Such functions are c "policies", or more specifically **Markov policies** or **Markov controls**. For many stoch optimization problems, including those studied here, this is the natural class of polici consider in order to achieve an optimum. The policies that occur in (2) are of this type. letter E, as before, denotes expectation. To compute it requires specifying the probabil that lie behind the calculation of the expectation. Recall that in the discrete random vari case the probability for the events $\mathbf{V}_1 = \mathbf{v}_1$ and $\mathbf{V}_2 = \mathbf{v}_2$ to occur jointly, given $\mathbf{V}_0 =$ equals the conditional probability for $\mathbf{V}_2 = \mathbf{v}_2$ to occur given $\mathbf{V}_1 = \mathbf{v}_1$, times the probab for $\mathbf{V}_1 = \mathbf{v}_1$ to occur given $\mathbf{V}_0 = \mathbf{v}_0$. That is, the joint probability equals $P_1(\mathbf{v}_2 \mid \mathbf{v}_1)$ ti $P_0(\mathbf{v}_1 \mid \mathbf{v}_0)$. Similarly, the probability of the joint event $\mathbf{V}_1 = \mathbf{v}_1, \mathbf{V}_2 = \mathbf{v}_2, \dots, \mathbf{V}_t =$ given by

$$P^t(\mathbf{v}_1, \dots, \mathbf{v}_t) = P_0(\mathbf{v}_1 \mid \mathbf{v}_0) \cdot P_1(\mathbf{v}_2 \mid \mathbf{v}_1) \cdots P_{t-1}(\mathbf{v}_t \mid \mathbf{v}_{t-1})$$

In the continuous random variable case, the same formula is valid in determining the joint density $p^t(\mathbf{v}_1, \dots, \mathbf{v}_t)$ provided each P_t is replaced by p_t.

Now, given the policies $\mathbf{u}_t(\mathbf{x}_t, \mathbf{v}_t)$, the sequence $\mathbf{X}_t$, $t = 1, \dots, T$ in (2) is the solution of (1) when $\mathbf{V}_1, \dots, \mathbf{V}_t$ and $\mathbf{u}_t = \mathbf{u}_t(\mathbf{X}_t, \mathbf{V}_t)$, $t = 0, \dots, T-1$ are inserted successively. Hence, $\mathbf{X}_t$ depends on $\mathbf{V}_1, \dots, \mathbf{V}_t$ and, for each t, the expectation $Ef(t, \mathbf{X}_t, \mathbf{u}_t(\mathbf{X}_t, \mathbf{V}_t))$ is calculated by means of the probabilities (or densities) specified in (3). We can write (2) as $\sum_{t=0}^{T} Ef(t, \mathbf{X}_t, \mathbf{u}_t(\mathbf{X}_t, \mathbf{V}_t))$, so we have now explained how the expectation in (2) is calculated.

Though not always necessary, we shall assume that f and $\mathbf{g}$ are continuous in $\mathbf{x}$, $\mathbf{u}$ (or in $\mathbf{x}$, $\mathbf{u}$, and $\mathbf{v}$ in the continuous random variable case).

The optimization problem is to find a sequence of policies $\mathbf{u}_0^*(\mathbf{x}_0, \mathbf{v}_0), \dots, \mathbf{u}_T^*(\mathbf{x}_T, \mathbf{v}_T)$, that gives the expression in (2) the largest possible value.

We now define

$$J(t, \mathbf{x}_t, \mathbf{v}_t) = \max E\left[\sum_{s=t}^{T} f(s, \mathbf{X}_s, \mathbf{u}_s(\mathbf{X}_s, \mathbf{V}_s)) \mid \mathbf{x}_t, \mathbf{v}_t\right] \tag{4}$$

where the maximum is taken over all policy sequences $\mathbf{u}_s = \mathbf{u}_s(\mathbf{x}_s, \mathbf{v}_s)$, $s = t, \dots, T$, given $\mathbf{v}_t$ and given that we "start" equation (1) in state $\mathbf{x}_t$ at time t, as indicated by "$\mid \mathbf{x}_t, \mathbf{v}_t$" in (4). The computation of the expectation in (4) is now based on conditional probabilities of the form $P(\mathbf{v}_{t+1}, \dots, \mathbf{v}_s \mid \mathbf{v}_t) = P_t(\mathbf{v}_{t+1} \mid \mathbf{v}_t) \cdots P_{s-1}(\mathbf{v}_s \mid \mathbf{v}_{s-1})$ in the discrete case, and conditional densities in the continuous case.

The central tool in solving optimization problems of the type (1)–(2) is the following **dynamic programming equation** or **optimality equation**:

$$J(t-1, \mathbf{x}_{t-1}, \mathbf{v}_{t-1}) = \max_{\mathbf{u}_{t-1}} \left\{ f(t-1, \mathbf{x}_{t-1}, \mathbf{u}_{t-1}) + E\left[J(t, \mathbf{X}_t, \mathbf{V}_t) \mid \mathbf{x}_{t-1}, \mathbf{v}_{t-1}\right]\right\} \tag{5}$$

where $\mathbf{X}_t = \mathbf{g}(t-1, \mathbf{x}_{t-1}, \mathbf{u}_{t-1}, \mathbf{V}_t)$. The "$\mathbf{x}_{t-1}$" in the symbol "$\mid \mathbf{x}_{t-1}, \mathbf{v}_{t-1}$" is just a reminder that inserting this value of $\mathbf{X}_t$ makes the expectation depend on $\mathbf{X}_{t-1}$, as well as on $\mathbf{V}_{t-1}$. After this insertion, equation (5) becomes

$$J(t-1, \mathbf{x}_{t-1}, \mathbf{v}_{t-1}) = \max_{\mathbf{u}_{t-1}} \left\{ f(t-1, \mathbf{x}_{t-1}, \mathbf{u}_{t-1}) + E\left[J(t-1, \mathbf{g}(t, \mathbf{x}_{t-1}, \mathbf{u}_{t-1}, \mathbf{V}_t), \mathbf{V}_t) \mid \mathbf{v}_{t-1}\right]\right\}$$

Moreover, when $t = T$ we have

$$J(T, \mathbf{x}_T, \mathbf{v}_T) = J(T, \mathbf{x}_T) = \max_{\mathbf{u}_T} f(T, \mathbf{x}_T, \mathbf{u}_T) \tag{6}$$

As in the deterministic case, first (6) is used to find $\mathbf{u}_T^*(\mathbf{x}_T, \mathbf{v}_T)$. Then (5) is used repeatedly to find $\mathbf{u}_{T-1}^*(\mathbf{x}_{T-1}, \mathbf{v}_{T-1})$, $\mathbf{u}_{T-2}^*(\mathbf{x}_{T-2}, \mathbf{v}_{T-2})$, etc.

Equations (5) and (6) are, essentially, both necessary and sufficient. They are sufficient in the sense that *if* $\mathbf{u}_{t-1}^*(\mathbf{x}_{t-1}, \mathbf{v}_{t-1})$ with $t = 1, \dots, T$ maximizes the right-hand side of (5) (or (6) for $t = T+1$), then $\mathbf{u}_{t-1}^*(\mathbf{x}_{t-1}, \mathbf{v}_{t-1})$, $t = 1, \dots, T+1$, are optimal policies. On the other hand, they are necessary in the sense that, for every $\mathbf{x}_{t-1}$, an optimal control

$\mathbf{u}_{t-1}^*(\mathbf{x}_{t-1}, \mathbf{v}_{t-1})$, $t = 1, \dots, T$, yields a maximum on the right-hand side of (5), a
$t = T+1$, on the right-hand side of (6). To be a little more precise, it is necessary t
optimal control $\mathbf{u}_{t-1}^*(\mathbf{x}_{t-1}, \mathbf{v}_{t-1})$ yields a maximum on the right-hand side of (5) (
for all values of $\mathbf{x}_{t-1}$ and $\mathbf{v}_{t-1}$ that can occur with positive probability (positive prob
density in the continuous case).

The intuitive argument for (5) is as follows: Suppose the system is in state $\mathbf{x}_{t-1}$
given $\mathbf{u}_{t-1}$, the "instantaneous" reward is $f(t-1, \mathbf{x}_{t-1}, \mathbf{u}_{t-1})$. In addition, the m
expected sum of rewards at all later times is $E[J(t, \mathbf{X}_t, \mathbf{V}_t) \mid \mathbf{x}_{t-1}, \mathbf{v}_{t-1}]$ providec
$\mathbf{g}(t-1, \mathbf{x}_{t-1}, \mathbf{u}_{t-1}, \mathbf{V}_t)$. When using $\mathbf{u}_{t-1}$, the total expected maximum value gain
all future time points (now including even $t-1$) is the sum in (5). The largest expect
comes from choosing $\mathbf{u}_{t-1}$ to maximize the right-hand side of (5).

Note that when $P_t(\mathbf{v} \mid \mathbf{v}_t)$ (or $p_t(\mathbf{v} \mid \mathbf{v}_t)$) does not depend on $\mathbf{v}_t$, then $\mathbf{v}_t$ can be d
from the functions $J_t(\mathbf{x}_t, \mathbf{v}_t)$, $\mathbf{u}_t(\mathbf{x}_t, \mathbf{v}_t)$ and in (5) and (6). (Intuitively, this is bec
(5) the conditioning on $\mathbf{v}_{t-1}$ drops out, so $J(t-1, \mathbf{x}_{t-1}, \mathbf{v}_{t-1})$ and the maximizing
$\mathbf{u}_{t-1} = \mathbf{u}_{t-1}(\mathbf{x}_{t-1}, \mathbf{v}_{t-1})$ will not depend on $\mathbf{v}_{t-1}$.) Some examples below empl
simplification.

NOTE 1 The argument above also holds if the control region is a closed set that d
on t and $\mathbf{x}$—for example, if $U = U(t, \mathbf{x}) = \{\mathbf{u} : h_i(t, \mathbf{x}, \mathbf{u}) \le 0\}$ where the function
continuous in $(\mathbf{x}, \mathbf{u})$. Thus it is here required that $\mathbf{u}_t \in U(t, \mathbf{x}_t)$. In this case the cor
in Note 12.1.5 carry over.

EXAMPLE 2 Suppose that a gambler chooses to bet a certain fraction u of his wealth at eve
Because of his skill, he wins this fraction with probability $p \ge 1/2$. Thus, if his
at time $t-1$ is x_{t-1}, then x_t is equal to $x_{t-1} + ux_{t-1}$ with probability p, and x_t i
to $x_{t-1} - ux_{t-1}$ with probability $1-p$. (Formally, $X_t = X_{t-1} + u_{t-1}V_tX_{t-1}$,
$V_t \in \{-1, 1\}$, $\Pr[V_t = 1] = p$, and $\Pr[V_t = -1] = 1-p$.) Suppose that he is g
play T times, and that the utility of terminal wealth x_T is $f(T, x_T) = \ln x_T = J($
(note that $f(T, x_T)$ is independent of u_T). We also have $f(t, x_t) \equiv 0$ for $t < T$.

If the gambler's wealth at time $T-1$ is x_{T-1} and he then bets ux_{T-1}, he w
$\ln(x_{T-1} + ux_{T-1})$ with probability p and $\ln(x_{T-1} - ux_{T-1})$ with probability $q =$
Thus, the expected utility of his terminal wealth is

$$p\ln(x_{T-1} + ux_{T-1}) + q\ln(x_{T-1} - ux_{T-1}) = \ln x_{T-1} + A(u)$$

where $A(u) = p\ln(1+u) + q\ln(1-u)$ (because $p + q = 1$). At time $T-1$ the opt
equation is therefore

$$J(T-1, x_{T-1}) = \ln x_{T-1} + \max_{0 \le u \le 1} A(u)$$

The function $A(u)$ is concave, so the maximum is attained where

$$A'(u) = p\frac{1}{1+u} - q\frac{1}{1-u} = 0$$

This implies $p(1-u) = q(1+u)$, or $p - q = u(p+q) = u$, so $u^*_{T-1} = p - q$. Inserting this expression for u into the right-hand side gives the maximum value. This is $J(T-1, x) = \ln x + B$, where $B = p\ln[1+(p-q)] + q\ln[1-(p-q)] = p\ln(2p) + q\ln(2q) = \ln 2 + p\ln p + q\ln q$.

Starting from x_{T-2}, we end up at $x_{T-1} = x_{T-2} + ux_{T-2}$ with probability p, and then obtain $J(T-1, x_{T-1}) = \ln(x_{T-2} + u_{T-2}x_{T-2}) + B$; with probability q we end up at $x_{T-1} = x_{T-2} - u_{T-2}x_{T-2}$ and obtain $J(T-1, x_{T-1}) = \ln(x_{T-2} - ux_{T-2}) + B$. Therefore,

$$\begin{aligned} J(T-2, x_{T-2}) &= \max_{0\le u\le 1}\left(p\big[\ln(x_{T-2} + ux_{T-2}) + B\big] + q\big[\ln(x_{T-2} - ux_{T-2}) + B\big]\right) \\ &= \ln x_{T-2} + B + \max_{0\le u\le 1} A(u) \end{aligned}$$

Once again, the maximum value in the latter maximization problem is B, with $u = p - q$. Hence

$$J(T-2, x_{T-2}) = \ln x_{T-2} + 2B \qquad (u^*_{T-2} = p - q)$$

Continuing in this manner, for $k = 3, 4, \ldots$ gives

$$J(T-k, x_{T-k}) = \ln x_{T-k} + kB \qquad (u^*_{T-k} = p - q)$$

To conclude, we see that in every round it is optimal for the gambler to bet the same fraction $u = p - q = 2p - 1$ of his wealth. (If the objective function were $f(T, x_T) = x_T$ and $p > 1/2$, it is easy to see that he would bet all his wealth at every stage.)

The *strict concavity* of the utility function $\ln x_T$ means that a decline in wealth reduces utility more than a corresponding rise in wealth increases utility. Therefore the gambler is careful and bets only a fraction each time. This is what economists call risk aversion.

EXAMPLE 3 Solve the problem in Example 1,

$$\max E\left[\sum_{t=0}^{T-1}\beta^t u_t^{1-\gamma} + \beta^T B X_T^{1-\gamma}\right], \qquad X_{t+1} = Z_{t+1}(X_t - u_t), \quad x_0 > 0 \quad u_t \in (0, x_t)$$

where $0 < \gamma < 1$, $0 < \beta < 1$, $B > 0$, and $\{Z_t\}_{t=0}^{T-1}$ is a sequence of independently distributed non-negative random variables with $EZ_t^{1-\gamma} < \infty$ for all t.

Solution: Here $J(T, x_T) = \beta^T B x_T^{1-\gamma}$. To find $J(T-1, x_{T-1})$, we use the optimality equation

$$J(T-1, x_{T-1}) = \max_u \left(\beta^{T-1}u^{1-\gamma} + E\big[\beta^T B(Z_T(x_{T-1} - u))^{1-\gamma}\big]\right) \qquad (*)$$

The expectation must be calculated by using the probability distribution for Z_T. In fact, the expectation term in $(*)$ is equal to $\beta^T BD_T(x_{T-1} - u)^{1-\gamma}$, where $D_t = E\big[Z_t^{1-\gamma}\big]$. Hence, the expression to be maximized in $(*)$ is $\beta^{T-1}u^{1-\gamma} + \beta^T BD_T(x_{T-1} - u)^{1-\gamma}$. Define $C_T = B$, $C_{T-1}^{1/\gamma} = 1 + (\beta BD_T)^{1/\gamma}$, and generally $C_t^{1/\gamma} = 1 + (\beta C_{t+1}D_{t+1})^{1/\gamma}$. The same calculations as in Example 12.1.4 show that, in general, maximizing the expression $(1-\gamma)u_t + \beta C_{t+1}a_t^{1-\gamma}(x-u)^{1-\gamma}$ for $u_t \in (0, x_t)$ gives $u_t = x/C_t^{1/\gamma}$, with maximum value equal to $C_t x^{1-\gamma}$, where $C_t^{1/\gamma} = 1 + (\beta C_{t+1}a_t^{1-\gamma})^{1/\gamma}$. Applying this for $D_t = a_{t-1}^{1-\gamma}$ gives $J_{T-1}(x) = \beta^{T-1}C_{T-1}x^{1-\gamma}$, $u_{T-1} = x/C_{T-1}^{1/\gamma}$, $J_{T-2}(x) = \beta^{T-2}C_{T-2}x^{1-\gamma}$, $u_{T-2} = x/C_{T-2}^{1/\gamma}$, and generally $J_t = \beta^t C_t x^{1-\gamma}$, $u_t = x/C_t^{1/\gamma}$.

We conclude this section with the following formal result:

THEOREM 12.6.1 (SUFFICIENCY OF THE OPTIMALITY EQUATIONS)

The sequence of policies $\boldsymbol{\pi} = \{\mathbf{u}_t(\mathbf{x}_t, \mathbf{v}_t)\}_{t=0}^T$ solves the problem of maximizing (2) subject to (1) if, together with a sequence of functions $\{J(t, \mathbf{x}_t, \mathbf{v}_t)\}_{t=0}^T$, it satisfies the optimality equations (5) (for $t = 1, 2, \ldots, T$) as well as (6).

Proof: Let $\boldsymbol{\pi} = \{\mathbf{u}_t(\mathbf{x}_t, \mathbf{v}_t)\}_{t=0}^T$ be an arbitrary control sequence. Define

$$J^{\pi}(t, \mathbf{x}_t, \mathbf{v}_t) = E\left[\sum_{s=t}^T f(s, \mathbf{X}_s, \mathbf{u}_s(\mathbf{X}_s, \mathbf{V}_s))|\mathbf{x}_t, \mathbf{v}_t\right]$$

which is the conditionally expected value in state $(\mathbf{x}_t, \mathbf{v}_t)$ at time t of following $\boldsymbol{\pi}$ from that s
Trivially, $J^{\pi}(T, \mathbf{x}_T, \mathbf{v}_T) \le J(T, \mathbf{x}_T, \mathbf{v}_T)$, with equality if $\mathbf{u}_T(\mathbf{x}_T, \mathbf{v}_T)$ satisfies (6). By backw
induction, let us prove that $J^{\pi}(t, \mathbf{x}_t, \mathbf{v}_t) \le J(t, \mathbf{x}_t, \mathbf{v}_t)$, with equality if $\boldsymbol{\pi}$ is such that $\mathbf{u}_s(\mathbf{x}_s$
satisfies (5) for $s = t, t+1, \ldots, T-1$, and $\mathbf{u}_T(\mathbf{x}_T, \mathbf{v}_T)$ satisfies (6). As the induction hypoth
assume that this is true for t. Replacing t by $t-1$ in the above definition gives

$$J^{\pi}(t-1, \mathbf{x}_{t-1}, \mathbf{v}_{t-1}) = f(t-1, \mathbf{x}_{t-1}, \mathbf{u}_{t-1}(\mathbf{x}_{t-1}, \mathbf{v}_{t-1})) + E\left[\sum_{s=t}^T f(s, \mathbf{X}_s, \mathbf{u}_s(\mathbf{X}_s, \mathbf{V}_s))|\mathbf{x}_{t-1}, \mathbf{v}\right.$$

But the law of iterated expectations and the induction hypothesis together imply that

$$\begin{aligned} E\left[\sum_{s=t}^T f(s, \mathbf{X}_s, \mathbf{u}_s(\mathbf{X}_s, \mathbf{V}_s))|\mathbf{x}_{t-1}, \mathbf{v}_{t-1}\right] &= E\left[E\left[\sum_{s=t}^T f(s, \mathbf{X}_s, \mathbf{u}_s(\mathbf{X}_s, \mathbf{V}_s))|\mathbf{X}_t, \mathbf{V}_t\right]|\mathbf{x}_{t-1}, \mathbf{v}_t\right. \\ &= E[J^{\pi}(t, \mathbf{X}_t, \mathbf{V}_t)|\mathbf{x}_{t-1}, \mathbf{v}_{t-1}] \le E[J(t, \mathbf{X}_t, \mathbf{V}_t)|\mathbf{x}_{t-1}, \mathbf{v}_i \end{aligned}$$

where $\mathbf{X}_t = \mathbf{g}(t, \mathbf{x}_{t-1}, \mathbf{u}_{t-1}(\mathbf{x}_{t-1}, \mathbf{v}_{t-1}), \mathbf{V}_t)$, with equality if $\mathbf{u}_s(\mathbf{x}_s, \mathbf{v}_s)$ satisfies (5) for s =
$t+1, \ldots, T-1$, and $\mathbf{u}_T(\mathbf{x}_T, \mathbf{v}_T)$ satisfies (6). Hence

$$\begin{aligned} J^{\pi}(t-1, \mathbf{x}_{t-1}, \mathbf{v}_{t-1}) &\le f(t-1, \mathbf{x}_{t-1}, \mathbf{u}_{t-1}(\mathbf{x}_{t-1}, \mathbf{v}_{t-1})) + E[J(t, \mathbf{X}_t, \mathbf{V}_t)|\mathbf{x}_{t-1}, \mathbf{v}_{t-1}] \\ &\le \max_{\mathbf{u}}\{f(t-1, \mathbf{x}_{t-1}, \mathbf{u}) + E[J(t, \mathbf{g}(t, \mathbf{x}_{t-1}, \mathbf{u}, \mathbf{V}_t), \mathbf{V}_t)|\mathbf{v}_{t-1}]\} \\ &= J(t-1, \mathbf{x}_{t-1}, \mathbf{v}_{t-1}) \end{aligned}$$

with equalities if $\mathbf{u}_s(\mathbf{x}_s, \mathbf{v}_s)$ satisfies (5) for $s = t-1, t, t+1, \ldots, T-1$, and $\mathbf{u}_T(\mathbf{x}_T, \mathbf{v}_T)$ sati
(6). This verifies the induction hypothesis for $t-1$, and so completes the proof.

In the discrete variable case, the above proof is easily adapted to show that a policy $\boldsymbol{\pi}$
optimal only if the optimality equations hold at every time t ($t = 0, 1, 2, \ldots, T$) in any s
$(\mathbf{x}_t, \mathbf{v}_t)$ which is reached with positive probability given $\boldsymbol{\pi}^*$. In the continuous variable c
these necessary conditions become a little bit more complicated: essentially, the optima
equations must hold at every time t ($t = 0, 1, 2, \ldots, T$) in almost every state $(\mathbf{x}_t, \mathbf{v}_t)$ wl
has a positive conditional probability density given $\boldsymbol{\pi}^*$.

The Stochastic Euler Equation

In the formulation of problem (12.2.1) leading to the Euler equation, let the function F in the criterion contain a stochastic variable $V_{t+1} = v$, governed as before, by a conditional probability distribution $P_t(v, |v_t)$ or a conditional density $p_t(v|v_t)$. Hence, consider the problem

$$\max E \sum_{t=0}^{T} F(t, X_t, X_{t+1}, V_{t+1}), \quad x_0 \text{ given}, \ x_t, \ldots, x_{T+1} \text{ free}$$

Now, we allow x_t to be a function of v_t and x_{t-1}. Hence we decide the value of x_t after observing v_t. Then the Euler equations takes the form

$$F'_3(T, x_T, x_{T+1}(x_T, v_{T+1}), v_{T+1}) = 0 \qquad (*)$$

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

$$E[F'_2(t+1, x_{t+1}, x_{t+2}(x_{t+1}, V_{t+2}), V_{t+2})|v_{t+1}] + F'_3(t, x_t, x_{t+1}(x_t, v_{t+1}), v_{t+1}) = 0 \ (**)$$

First, $(*)$ is solved for x_{T+1}, yielding the function $x_{T+1} = x_{T+1}(x_T, v_{T+1})$. Next, this function is inserted into $(**)$ for $t = T-1$, and $(**)$ is then solved for x_T, yielding $x_T = x_T(x_{T-1}, v_T)$. Then this function is inserted into $(**)$ for $t = T-2$ and $(**)$ is then solved for x_{T-1} yielding the function $x_{T-1}(x_{T-2}, v_{T-1})$. In this manner we work backwards until the function $x_1(x_0, v_1)$ has been constructed. Since x_0 is given, the value of x_1 is determined once we have observed v_1. Then the value of $x_2 = x_2(x_1, v_2)$ is determined once we have observed v_2 and so on.

PROBLEMS FOR SECTION 12.6

1. Consider the stochastic dynamic programming problem

$$\max E\Big[-\delta \exp(-\gamma X_T) + \sum_{t=0}^{T-1} -\exp(-\gamma u_t)\Big], \quad X_{t+1} = 2X_t - u_t + V_{t+1}, \quad x_0 \text{ given}$$

where u_t are controls taking values anywhere in $\mathbb{R}$, $\delta > 0$ and $\gamma > 0$. Here V_{t+1}, $t = 0, 1, 2, \ldots, T-1$, are identically and independently distributed. Moreover, $K = E\big[\exp(-\gamma V_{t+1})\big] < \infty$. Show that the optimal value function $J(t, x)$ can be written $J(t, x) = -\alpha_t \exp(-\gamma x)$, and find a backwards difference equation for α_t. What is α_T?

2. (Blanchard and Fischer (1989)) Solve the problem

$$\max E\Big[\sum_{t=0}^{T-1} (1+\theta)^{-t} \ln C_t + k(1+\theta)^{-T} \ln A_T\Big]$$

where w_t and C_t are controls, k and θ are positive constants, and

$$A_{t+1} = (A_t - C_t)[(1 + r_t)w_t + (1 + V_{t+1})(1 - w_t)]$$

where r_t is a given sequence. The stochastic variables V_t are independently and identically distributed.

3. Solve the problem

$$\max E\Big[\sum_{t<T} 2u_t^{1/2} + aX_T\Big], \quad a > 0, \quad x_0 > 0, \quad T \text{ fixed}, \quad u_t \geq 0$$

where $X_{t+1} = X_t - u_t$ with probability 1/2 and $X_{t+1} = 0$ with probability 1/2.

4. Solve the problem

$$\max E \sum_{t=0}^{T-1} -u_t^2 - X_T^2 \quad \text{subject to} \quad X_{t+1} = X_t V_{t+1} + u_t, \ V_{t+1} \in \{0, 1\}$$

with $\Pr[V_{t+1} = 1|V_t = 1] = 3/4$, $\Pr[V_{t+1} = 1|V_t = 0] = 1/4$. (*Hint:* $J(t, x_t, 1) = -a_t x_t^2$, $J(t, x_t, 0) = -b_t x_t^2$.)

5. Solve the dynamic programming problem

$$\max E \sum_{t=0}^{T} (1 - u_t)X_t, \quad X_{t+1} = X_t + u_t X_t + V_{t+1}, \quad x_0 = 1, \quad u_t \in [0, 1]$$

where $V_{t+1} \geq 0$ is exponentially distributed with parameter λ (i.e. the density of \ is $\varphi(v) = \lambda e^{-\lambda v}$).

6. (Hakansson) Let x_t denote capital, y_t income, c_t consumption (a control), and z_t vestment with uncertain return (another control). The balance $x_t - c_t - z_t$ is place a bank, where it earns a return r equal to 1 plus the interest rate. Let the gross rat return on the uncertain investment (i.e. z) be β_t (so β_t equals 1 plus an uncertain net of return). Assume the random variables β_t are independent and identically distribu Then

$$X_{t+1} = (\beta_t - r)z_t + r(X_t - c_t) + y_t$$

Assume that $E\beta_t > r$. Let $K > 0$, $\gamma \in (0, 1)$ be given numbers. The maximiza problem is

$$\max E\Big[\sum_{t=1}^{T-1} (\alpha^t/\gamma)c_t^\gamma + K(\alpha^T/\gamma)X_T^\gamma\Big]$$

where $c_t \geq 0$, $z_t \geq 0$.

(a) Solve the problem, i.e. find the optimal controls. Assume that $c_t > 0$ and x_{t+1} in optimum. (*Hint:* When maximizing w.r.t. (c, z), first maximize w.r.t. z. W maximizing w.r.t. z, use the fact that for an arbitrarily given number $b > 0$, one

$$\max_{z \geq 0} E\big[\{rb + (\beta - r)z\}^\gamma\big] = b^\gamma a, \quad \text{where } a = \max_{s \geq 0} E\big[\{r + (\beta - r)s\}^\gamma\big]$$

Don't try to find a, use it as a known parameter in the solution of the p lem. Formally, we let $w^\gamma = -\infty$ when $w < 0$. Write expressions of the 1 $\{y + r(x - c) + (\beta - r)z\}^\gamma$ as $\{r(\frac{1}{r}y + x - c) + (\beta - r)z\}^\gamma$ when using $(*)$.

(b) Discuss dependence on parameters in the problem, including the distribution (Show $(*)$.

7. Consider the problem

$$\max E\Big[\sum_{t=0}^{T-1}((1-u_t)X_t^2-u_t)+2X_T^2\Big] \quad \text{s. t.} \quad X_{t+1}=u_tX_tV_{t+1}, \quad u_t\in U=[0,1]$$

where $V_{t+1}=2$ with probability 1/4 and $V_{t+1}=0$ with probability 3/4.

(a) Find $J(T,x)$, $J(T-1,x)$, and $J(T-2,x)$. (Note that the maximand will be convex in the control u, so any maximum will be situated at an endpoint of U.)

(b) Find $J(t,x)$ for general t.

8. Solve the problem

$$\max E\Big[\sum_{1\le t\le T-1} u_t^{1/2}+aX_t^{1/2}\Big] \quad \text{subject to} \quad X_{t+1}=(X_t-u_t)V_{t+1}$$

where a and T are given positive numbers, and where $V_{t+1}=0$ with probability 1/2, $V_{t+1}=1$ with probability 1/2. (*Hint:* Try $J(t,x)=2a_tx^{1/2}$, $a_t>0$.)

9. Solve the gambler's problem in Example 2 when $f(T,x_T)=(x_T)^{1-\alpha}/(1-\alpha)$, where $\alpha>0$, $\alpha\ne 1$.

10. (Bertsekas (1976)) A farmer annually produces X_k units of a certain crop and stores $(1-u_k)X_k$ units of his production, where $0\le u_k\le 1$. He invests the remaining u_kX_k units, thus increasing next year's accumulated output to a level X_{k+1} given by

$$X_{k+1}=X_k+w_ku_kX_k, \qquad k=0,1,\dots,N-1$$

The scalars W_k are independent random variables with an identical probability distribution that depends neither on x_k nor u_k. Furthermore, $E[W_k]=\overline{w}>0$. The problem is to find the optimal policy that maximizes the expected output accumulated over N years,

$$E\Big[X_N+\sum_{k=0}^{N-1}(1-u_k)X_k\Big]$$

Show that one optimal control is given by:

(i) If $\overline{w}>1$, then $u_0^*(x_0)=\cdots=u_{N-1}^*(x_{N-1})=1$.

(ii) If $0<\overline{w}<1/N$, then $u_0^*(x_0)=\cdots=u_{N-1}^*(x_{N-1})=0$.

(iii) If $1/N\le\overline{w}\le 1$, then

$$u_0^*(x_0)=\cdots=u_{N-\bar{k}-1}^*(x_{N-\bar{k}-1})=1$$

$$\vdots$$

$$u_{N-\bar{k}}^*(x_{N-\bar{k}})=\cdots=u_{N-1}^*(x_{N-1})=0$$

where $\bar{k}$ is such that $1/(\bar{k}+1)\le\overline{w}<1/\bar{k}$. Note that this control consists of constant functions.

11. Use the stochastic Euler equation to solve the problem

$$\max E\sum_{t=0}^{2}[1-(v_{t+1}+X_{t+1}-X_t)^2+(1+v_3+X_3)], \quad X_0=0,\ X_1,X_2,X_3$$

where all v_t are identically and independently distributed, with $Ev_t=1/2$.

12.7 Infinite Horizon Stationary Problems

We consider an infinite horizon version of the problem in the previous section. Sup that both $P_t(\mathbf{v}_{t+1}\mid\mathbf{v}_t)$ (or $p_t(\mathbf{v}_{t+1}\mid\mathbf{v}_t)$) and $\mathbf{g}$ are independent of t, and that the instan eous reward is $\beta^t f(\mathbf{x},\mathbf{u})$ with $\beta\in(0,1]$. The problem is then often called **stationar autonomous**. We focus on the discrete variable case, which takes the form

$$\max_{\pi} E\sum_{t=0}^{\infty}\beta^t f(\mathbf{X}_t,\mathbf{u}_t(\mathbf{X}_t,\mathbf{V}_t)), \quad \mathbf{u}_t(\mathbf{X}_t,\mathbf{V}_t)\in U, \qquad \Pr[\mathbf{V}_{t+1}=\mathbf{v}\mid\mathbf{v}_t]=P(\mathbf{v}\mid\mathbf{v}_t)$$

where $\mathbf{X}_t$ is governed by the stochastic difference equation

$$\mathbf{X}_{t+1}=\mathbf{g}(\mathbf{X}_t,\mathbf{u}_t(\mathbf{X}_t,\mathbf{V}_t),\mathbf{V}_{t+1})$$

with $\mathbf{X}_0$ and $\mathbf{V}_0$ given. The functions f and $\mathbf{g}$ are continuous. The control functions $\mathbf{u}_t$ values in a fixed control region U. Among all sequences $\pi=(\mathbf{u}_0(\mathbf{x}_0,\mathbf{v}_0),\mathbf{u}_1(\mathbf{x}_1,\mathbf{v}_1),$ of Markov controls, we seek one that maximizes the objective function in (1).

We introduce the following *boundedness condition*:

$$M_1\le f(\mathbf{x},\mathbf{u})\le M_2 \text{ for all } (\mathbf{x},\mathbf{u})\in\mathbb{R}^n\times U, \text{ where } M_1 \text{ and } M_2 \text{ are given numbers}$$

For a given sequence π, let us write

$$J_\pi(s,\mathbf{x}_s,\mathbf{v}_s)=E\Big[\sum_{t=s}^{\infty}\beta^t f(\mathbf{X}_t,\mathbf{u}_t(\mathbf{X}_t,\mathbf{V}_t))\mid\mathbf{x}_s,\mathbf{v}_s\Big]$$

and define $J(s,\mathbf{x}_s,\mathbf{v}_s)=\sup_\pi J_\pi(s,\mathbf{x}_s,\mathbf{v}_s)$. We claim that $J(1,\mathbf{x}_0,\mathbf{v}_0)=\beta J(0,\mathbf{x}_0,$ The intuitive argument is as follows. Let $J_\pi^k(\mathbf{x},\mathbf{v})=\sum_{t=k}^{\infty}E[\beta^{t-k}f(\mathbf{X}_t,\mathbf{u}_t(\mathbf{X}_t,\mathbf{V}_t))\mid$ and let $J^k(\mathbf{x},\mathbf{v})=\sup_\pi J_\pi^k(\mathbf{x},\mathbf{v})$. Then $J^k(\mathbf{x},\mathbf{v})$ is the maximal expected present valu future rewards discounted back to $t=k$, given that the process starts at $(\mathbf{x},\mathbf{v})$ at time $t=$ When starting at $(\mathbf{x},\mathbf{v})$ at time $t=0$, and discounting back to $t=0$, the correspon maximal expected value is $J^0(\mathbf{x},\mathbf{v})=J(0,\mathbf{x},\mathbf{v})$. Because time does not enter explicitl $P(\mathbf{v}\mid\mathbf{v}_t)$, $\mathbf{g}$, or f, the future looks exactly the same at time $t=k$ as it does at time $t=$ Hence $J^k(\mathbf{x},\mathbf{v})=J^0(\mathbf{x},\mathbf{v})$. But $J(k,\mathbf{x}_0,\mathbf{v}_0)=\beta^kJ^k(\mathbf{x}_0,\mathbf{v}_0)$ because, in the defini of $J_\pi(k, \mathbf{x}_0, \mathbf{v}_0)$, we discount back to $t = 0$. Hence $\beta^k J(0, \mathbf{x}_0, \mathbf{v}_0) = J(k, \mathbf{x}_0, \mathbf{v}_0)$, and in particular $\beta J(0, \mathbf{x}_0, \mathbf{v}_0) = J(1, \mathbf{x}_0, \mathbf{v}_0)$.

The heuristic argument for the optimality equation (12.6.5) works just as well in the infinite horizon case. When $t = 0$, if we write $\mathbf{x}$ and $\mathbf{v}$ instead of $\mathbf{x}_0$ and $\mathbf{v}_0$, define $J(\mathbf{x}, \mathbf{v}) = J(0, \mathbf{x}, \mathbf{v})$, then recognize that $J(1, \mathbf{x}, \mathbf{v}) = \beta J(0, \mathbf{x}, \mathbf{v})$, we derive from (12.6.5) the following **optimality equation** or **Bellman equation**

$$J(\mathbf{x}, \mathbf{v}) = \max_{\mathbf{u}} \{ f(\mathbf{x}, \mathbf{u}) + \beta E[J(\mathbf{X}_1, \mathbf{V}_1) \mid \mathbf{x}, \mathbf{v}] \} \tag{5}$$

where $\mathbf{X}_1 = \mathbf{g}(\mathbf{x}, \mathbf{u}, \mathbf{V}_1)$.

Observe that (5) is a "functional equation" which (we hope) determines the unknown function J that occurs on both sides of the equality sign. Once J is known, the optimal Markov control is obtained from the maximization in the optimality equation. The maximization seems to yield an optimal control function $\mathbf{u}(\mathbf{x}, \mathbf{v})$ not dependent on t. This is to be expected: Whether we observe $(\mathbf{x}, \mathbf{v})$ at time 0 or at time t does not matter; the optimal choice of control should be the same in the two situations, because the future looks exactly the same at both these times.

When the boundedness condition (3) is satisfied, it can be shown that the optimal value function is defined and satisfies the optimality equation. Moreover, the optimality equation has a unique bounded solution $J(\mathbf{x}, \mathbf{v})$. (At least this is so when "max" is replaced by "sup" in the Bellman equation.). Furthermore, $J(\mathbf{x}, \mathbf{v})$ is automatically the optimal value function in the problem, and any control $\mathbf{u}(\mathbf{x}, \mathbf{v})$ that maximizes the right-hand side of (5), given the function $J(\mathbf{x}, \mathbf{v})$, is optimal.

NOTE 1 (**Alternative Boundedness Conditions**) Complications arise when the boundedness condition (3) fails. First, the Bellman equation might then have more than one solution, or perhaps none. Even if it has one or more solutions, it might be that none of them is the optimal value function.

We consider two cases where some results can be obtained. In both cases we must allow infinite values for the optimal value function $J(\mathbf{x})$, $+\infty$ in case A, and $-\infty$ in case B. (Of course, $J(\mathbf{x}, \mathbf{v}) \equiv \infty$ and $J(\mathbf{x}, \mathbf{v}) \equiv -\infty$ in a sense satisfy the Bellman equation, being perhaps "false" solutions.)

A *Either* $f(\mathbf{x}, \mathbf{u}) \geq 0$ *for all* $(\mathbf{x}, \mathbf{u}) \in \mathbb{R}^n \times U$ *and* $\beta \in (0, 1]$, *or for some negative number* γ, $f(\mathbf{x}, \mathbf{u}) \geq \gamma$ *for all* $(\mathbf{x}, \mathbf{u}) \in \mathbb{R}^n \times U$ *and* $\beta \in (0, 1)$.

In this case if $\mathbf{u}(\mathbf{x}, \mathbf{v})$ yields the maximum in the Bellman equation with $J^{\mathbf{u}}(\mathbf{x}, \mathbf{v})$ inserted, then $\mathbf{u}(\mathbf{x}, \mathbf{v})$ is optimal. Here $J^{\mathbf{u}}(\mathbf{x}, \mathbf{v})$ is the value function arising from using $\mathbf{u}(\mathbf{x}, \mathbf{v})$ all the time.

Sometimes it is useful to know the fact that if $J^{\mathbf{u}}(s, \mathbf{x}, \mathbf{v}, T)$ is the value function arising from using $\mathbf{u} = \mathbf{u}(\mathbf{x}, \mathbf{v})$ all the time from s until $t = T$, then $J^{\mathbf{u}}(0, \mathbf{x}, \mathbf{v}, T) \to J^{\mathbf{u}}(\mathbf{x}, \mathbf{v})$ as $T \to \infty$.

B *Either* $f(\mathbf{x}, \mathbf{u}) \leq 0$ *for all* $(\mathbf{x}, \mathbf{u}) \in \mathbb{R}^n \times U$ *and* $\beta \in (0, 1]$, *or for some positive number* γ, $f(\mathbf{x}, \mathbf{u}) \leq \gamma$ *for all* $(\mathbf{x}, \mathbf{u}) \in \mathbb{R}^n \times U$ *and* $\beta \in (0, 1)$.

In this case it is known that if $\mathbf{u}(\mathbf{x}, \mathbf{v})$ satisfies the Bellman equation (i.e. yields maximum) with the optimal value function inserted, then $\mathbf{u}(\mathbf{x}, \mathbf{v})$ is optimal. If we are able to prove

that the Bellman equation has one and only one solution $\hat{J}(\mathbf{x}, \mathbf{v}) > -\infty$, then this is th optimal value function $J(\mathbf{x}, \mathbf{v})$ provided we know that $J(\mathbf{x}, \mathbf{v}) > -\infty$. (Recall that $J(\mathbf{x},$ is known to satisfy the Bellman equation, both in case A and B.) Another possibility is th following: Suppose we have solved the finite horizon problem, with horizon T. Assume th U is compact, and that $f(\mathbf{x}, \mathbf{u})$ and $\mathbf{g}(\mathbf{x}, \mathbf{u}, \mathbf{v})$ are continuous in $(\mathbf{x}, \mathbf{u})$. Denote the optim value function in this problem by $J(0, \mathbf{x}, \mathbf{v}, T)$. If we now find the limit $\lim_{T\to\infty} J(0, \mathbf{x}, \mathbf{v}, T$ then this is the optimal value function. (As $T \to \infty$, $J(0, \mathbf{x}, \mathbf{v}, T)$ converges to the optim value function in this case, as well as in the cases A and (3).) For the results in this note s Bertsekas (1976) and Hernández-Lerma and Lasserre (1996).

To sum up, what do we do after we have found a pair $(\mathbf{u}(\mathbf{x}, \mathbf{v}), \hat{J}(\mathbf{x}, \mathbf{v}))$ satisfying th Bellman equation? In case A, we try to check if $J^{\mathbf{u}} = \hat{J}$ holds. If it does, $(\mathbf{u}(\mathbf{x}, \mathbf{v}), \hat{J}(\mathbf{x}, \mathbf{v}$ is optimal ($\hat{J}$ is then the optimal value function). In case B with $\hat{J} > -\infty$, we either try show that $\hat{J}$ is the only solution greater than $-\infty$ satisfying the Bellman equation and th $J > -\infty$, or we try to check that $J(0, \mathbf{x}, \mathbf{v}, T) \to \hat{J}(\mathbf{x}, \mathbf{v})$ as $T \to \infty$. If either of the tests comes out positive, $(\mathbf{u}(\mathbf{x}, \mathbf{v}), \hat{J}(\mathbf{x}, \mathbf{v}))$ is optimal.

NOTE 2 The boundedness condition (3), or the alternatives in Note 1, need only hold f $\mathbf{x}$ in $\mathcal{X}(\mathbf{x}_0) = \bigcup_t \mathcal{X}_t(\mathbf{x}_0)$ for all t, where $\mathcal{X}_t(\mathbf{x}_0)$ is the set of states that can be reach at time t when starting at $\mathbf{x}_0$ at time 0, considering all controls and all outcomes that c occur with positive probability. The conclusions drawn in the case where (3) is satisfied a also valid if the following weaker condition holds: There exist positive constants M, M α, and δ such that for all $\mathbf{x} \in \mathcal{X}(\mathbf{x}_0)$ and $\mathbf{u} \in U$, one has $|f(\mathbf{x}, \mathbf{u})| \leq M^*(1 + \|\mathbf{x}\|^\alpha)$ a $\|\mathbf{g}(\mathbf{x}, \mathbf{u})\| \leq M + \delta\|\mathbf{x}\|$, with $\beta\delta^\alpha < 1$ and $\beta \in (0, 1)$.

EXAMPLE 1 Consider the following stochastic version of Example 12.3.1:

$$\max_{w_t \in (0,1)} E\left[\sum_{t=0}^{\infty} \beta^t X_t^{1-\gamma} w_t^{1-\gamma}\right]$$

$$X_{t+1} = V_{t+1}(1 - w_t)X_t, \quad x_0 \text{ is a positive constant}, \quad 0 < \gamma < 1 \qquad ($$

Here, $V_1, V_2, \ldots$ are identically and independently distributed nonnegative stochastic va ables, with $D = EV^{1-\gamma} < \infty$, where V is any of the V_t. Now, $w \in (0, 1)$ is the control. is assumed that

$$\beta \in (0, 1), \quad \gamma \in (0, 1), \quad \rho = (\beta D)^{1/\gamma} < 1 \qquad (i$$

In the notation of problem (1)–(3), $f(x, w) = x^{1-\gamma} w^{1-\gamma}$ and $g(x, w, V) = V(1 - w)$ The optimality equation (5) yields

$$J(x) = \max_{w \in (0,1)} \left[x^{1-\gamma} w^{1-\gamma} + \beta E J(V(1 - w)x)\right] \qquad (i$$

We guess that $J(x)$ has the form $J(x) = kx^{1-\gamma}$ for some constant k. (The optimal val function had a similar form in the finite horizon version of this problem discussed in t previous section, as well as in the deterministic infinite horizon version of Example 12.3.1.) Then, cancelling the factor $x^{1-\gamma}$, (iv) reduces to

$$k = \max_{w \in (0,1)} \left[w^{1-\gamma} + \beta k D (1-w)^{1-\gamma} \right] \tag{v}$$

where $D = E[V^{1-\gamma}]$. Note that equation (v) is the same as equation (iii) in Example 12.3.1, except that $a^{1-\gamma}$ is replaced by D. It follows that $J(x) = (1-\rho)^{-\gamma} x^{1-\gamma}$, with $w = 1 - \rho$, where $\rho = (\beta D)^{1/\gamma}$.

In this example the boundedness condition (3) is not satisfied for $x \in \bigcup_t \mathcal{X}_t(x_0)$.

One way to tackle this problem is to use the transformation $y_t = x_t / V_t$ with y_t satisfying $y_{t+1} = (1 - w_t) z_t y_t$, $z_{t+1} = V_{t+1}$. Taking the expectation of the objective function inside the sum, and using the so-called monotone convergence theorem, the problem can be transformed into one in which the new discount factor is $\hat{\beta} = \beta E V^{1-\gamma} < 1$ and where $g = g(y, w) = y^{1-\gamma} w^{1-\gamma}$, with g satisfying (3) in $\bigcup_t \mathcal{X}_t(x_0)$. Yet another way out is the following: Let us use A in the note above. Then we need to know that $J^{w^*}(x) = J(x)$. It is fairly easy to carry out the explicit calculation of $J^{w^*}(x)$ ($W = w^*$ as in (vi)), by taking the expectation inside the sum in the objective and summing the resulting geometric series. But we don't need to do that. Noting that $x_t = x_0 \rho^t Z_1 \cdot \ldots \cdot Z_t$, evidently, we must have that $J^{w^*}(x_0) = k x_0^{1-\gamma}$ for some k. We must also have that $J^{w^*}(x_0)$ satisfies the equilibrium optimality equation with $w = w^*$ and the maximization deleted (in the problem where $U = w^*$, w^* is optimal !). But the only value of k which satisfies this equation were found above. Thus the test in A works and w^* as specified is optimal.

Counterexamples

The Bellman equation may have "false" solutions. Two examples will be given.

Consider the problem $\sum_{t=0}^{\infty} \beta^t (1 - u_t)$, subject to $x_{t+1} = (1/\beta)(x_t + u_t)$, $u_t \in [0, 1]$, x_0 given, $\beta \in (0, 1]$. The Bellman equation is satisfied by $J(x) = \gamma + x$, where $\gamma = 1/(1-\beta)$, with any $u = \bar{u} \in [0, 1]$ yielding the maximum in the Bellman equation, let, say, $\bar{u} = 1/2$. (The Bellman equation is then $\gamma + x = \max_u \{1 - u + \beta(\gamma + (1/\beta)(x + u)\} = 1 + \beta\gamma + x$ and γ equals $1 + \beta\gamma$.) Is then $u_t \equiv 1/2$ the optimal control, and $J(x) = \gamma + x$, the optimal value function? The first thing we note is that $J^{\bar{u}}(x)$ is independent of x, so $J^{\bar{u}}(x) \neq \gamma + x$. Neither $u_t \equiv 1/2$ nor $J(x_0) = \gamma + x_0$ are optimal entities, it is trivial that $u_t \equiv 0$ is optimal, with a criterion value independent of x_0 and strictly greater than the criterion value of $u_t \equiv 1/2$.

What about cases where we have $J^{\mathbf{u}}(\mathbf{x}, \mathbf{v}) = J(\mathbf{x}, \mathbf{v})$, where $(\mathbf{u}(\mathbf{x}, \mathbf{v}), J(\mathbf{x}, \mathbf{v}))$ satisfies the Bellman equation? Consider the problem of maximizing $\sum_{t=0}^{\infty} \beta^t x_t (u_t - \alpha)$ subject to $x_{t+1} = x_t u_t$ and $0 \leq u_t \leq \alpha$, where $x_0 > 0$ is given, and α, β are positive constants satisfying $\alpha\beta = 1$, $\beta \in (0, 1]$.

Note that, regardless of which $u_t \in [0, \alpha]$ is chosen in each period, one has $x_t \geq 0$ for all t, so $\mathcal{X}(x_0) \subseteq [0, \infty)$.

The Bellman equation is

$$J(x) = \max_{u \in [0,\alpha]} \{x(u - \alpha) + \beta J(xu)\}$$

We look for solutions of the form $J(x) = \gamma x$, where γ is a constant. The condition for this to be a solution when $x > 0$ is that

$$\gamma = \max_{u \in [0,\alpha]} \{u - \alpha + \beta\gamma u\}$$

and we see that $\gamma = -1/\beta$ works. In this case any $u \in [0, \alpha]$ yields maximum in the Be equation. If we choose the same $u \in [0, \alpha)$ in each period, then $J^u = \sum_{t=0}^{\infty} \beta^t x_t (u - \alpha)$ \ $x_t = u^t x_0$. Hence $J^u(x_0) = x_0(u - \alpha) \sum_{t=0}^{\infty} (\beta u)^t = x_0(u - \alpha)/(1 - \beta u) = -x_0/\beta =$ So the function $J(x) \equiv -x/\beta = \gamma x$ solves the Bellman equation, and is the criterion valu corresponding stationary policy $u_t \equiv$ constant $\in [0, \alpha)$. However, $J(x) = \gamma x$ is not equal criterion value of $u_t \equiv \alpha$.

Now, as x_t is always ≥ 0, $x_t(u_t - \alpha) \leq 0$, so $u_t \equiv \alpha$ is obviously optimal, with criterion $= 0$, and $J(x) \equiv 0$ solves also the Bellman equation, with $u = \alpha$ as a maximizing control.

A necessary condition for optimality of a policy $\mathbf{u}(\mathbf{x}, \mathbf{v})$ is that it satisfies the Bellman equati $J^{\mathbf{u}}$ inserted. (It is necessary that $\mathbf{u}(\mathbf{x}, \mathbf{v})$ satisfies the Bellman equation for the optimal value fur $J(\mathbf{x}, \mathbf{v})$. It is also necessary that $J^{\mathbf{u}}(\mathbf{x}, \mathbf{v}) = J(\mathbf{x}, \mathbf{v})$, hence the assertion follows.) It is a nece and sufficient condition in case A, but not in Case B. In the first example (which is of type A Bellman equation is not satisfied by $\bar{u} = 1/2$, it is sufficient to note that $J^u(x)$ is a constant this. This condition is not sufficient in problems of type B, as we saw in the last example, whic of this type.

The first example was actually of the type (3), so in that case the Bellman equation can unbounded (and hence "false" solutions), in addition to the unique (correct) bounded one.

Iterative Methods

In this subsection, we describe two new methods yielding approximate solutions to in horizon dynamic programming problems. One approximation result has already been tioned: Under certain conditions, $J(0, \mathbf{x}, \mathbf{v}, T) \to J(\mathbf{x}, \mathbf{v})$ when $T \to \infty$. Another me is the **successive approximation method** that can be formulated as follows.

For any real-valued bounded function $h(\mathbf{x}, \mathbf{v})$, for any function $\mathbf{u} = \mathbf{u}(\mathbf{x}, \mathbf{v})$, d $\Psi^{\mathbf{u}}(h)$, which denotes a function of $(\mathbf{x}, \mathbf{v})$, by the formula

$$\Psi^{\mathbf{u}}(h)(\mathbf{x}, \mathbf{v}) = \{f(\mathbf{x}, \mathbf{u}(\mathbf{x}, \mathbf{v})) + \beta E[h(\mathbf{g}(\mathbf{x}, \mathbf{u}(\mathbf{x}, \mathbf{v}), \mathbf{V}), \mathbf{V}) \mid \mathbf{x}, \mathbf{v}]\}$$

Then define the function $\Psi(h)$ by

$$\Psi(h)(\mathbf{x}, \mathbf{v}) = \max_{\mathbf{u}} \Psi^{\mathbf{u}}(h)(\mathbf{x}, \mathbf{v}) = \max_{\mathbf{u} \in U} \left\{ f(\mathbf{x}, \mathbf{u}) + \beta E[h(\mathbf{g}(\mathbf{x}, \mathbf{u}, \mathbf{V}), \mathbf{V}) \mid \mathbf{x}, \mathbf{v}] \right\}$$

Let $\Psi^2(h) = \Psi(\Psi h)$, $\Psi^3(h) = \Psi(\Psi^2(h))$, and so on. Choose in particular $h =$ (calculate successively $\Psi^1(0) = \Psi(0)$, $\Psi^2(0) = \Psi(\Psi^1(0))$, ... Let the control $\mathbf{u}_k$(be the one that yields a maximum at step k. (When $\Psi^{k-1}(0)$ is known, a maximizati carried out to find $\Psi^k(0)$, and we assume that all maxima are attained.) Provided that satisfied, the controls $\mathbf{u}_k(\mathbf{x}, \mathbf{v})$ are approximately optimal for k large.

The second method is called **policy improvement**. It works as follows. Choose an i stationary policy $\mathbf{u}_0 = \mathbf{u}_0(\mathbf{x}, \mathbf{v})$. Calculate $J^{\mathbf{u}_0}(\mathbf{x}, \mathbf{v})$, the expected value of the obje when starting at $\mathbf{x}$ at time 0, and using $\mathbf{u}_0(\mathbf{x}, \mathbf{v})$ all the time. For each $(\mathbf{x}, \mathbf{v})$ find the cc $\mathbf{u}_1(\mathbf{x}, \mathbf{v})$ that yields a maximum when calculating $\Psi(J^{\mathbf{u}_0})(\mathbf{x}, \mathbf{v})$. Next calculate $J^{\mathbf{u}_1}$(: and find the control $\mathbf{u}_2(\mathbf{x}, \mathbf{v})$ that maximizes $\Psi(J^{\mathbf{u}_1})(\mathbf{x}, \mathbf{v})$, and so on. Since for a $J^{\mathbf{u}_i}(\mathbf{x}, \mathbf{v}) = f(\mathbf{x}, \mathbf{u}_i(\mathbf{x}, \mathbf{v})) + E[\beta J^{\mathbf{u}_i}(g(\mathbf{x}, \mathbf{u}_i(\mathbf{x}, \mathbf{v}), \mathbf{V}))|\mathbf{x}, \mathbf{v}]$, it is clear that $\Psi(J^{\mathbf{u}}$ $J^{\mathbf{u}_i}$. Then $\Psi^{\mathbf{u}_{i+1}}(\Psi(J^{\mathbf{u}_i})) \geq \Psi^{\mathbf{u}_{i+1}}(J^{\mathbf{u}_i}) = \Psi(J^{\mathbf{u}_i}) \geq J^{\mathbf{u}_i}$, $\Psi^{\mathbf{u}_{i+1}}(\Psi^{\mathbf{u}_{i+1}}(\Psi(J^{\mathbf{u}_i})$ $\Psi^{\mathbf{u}_{i+1}}(J^{\mathbf{u}_i}) = \Psi(J^{\mathbf{u}_i}) \geq J^{\mathbf{u}_i}$, and generally $(\Psi^{\mathbf{u}_{i+1}})^k(\Psi(J^{\mathbf{u}_i})) \geq J^{\mathbf{u}_i}$. By a contra argument, $(\Psi^{\mathbf{u}_{i+1}})^k(\Psi(J^{\mathbf{u}_i})) \to J^{\mathbf{u}_{i+1}}$ when $k \to \infty$, and it follows that $J^{\mathbf{u}_{i+1}} \geq J$

fact, $J^{\mathbf{u}_i}(\mathbf{x}, \mathbf{v})$ increases monotonically to the optimal value function $J(\mathbf{x}, \mathbf{v})$ when (3) holds, and so for i large, $\mathbf{u}_i(\mathbf{x}, \mathbf{v})$ is approximately optimal. At each step $J^{\mathbf{u}_{i+1}}$ can be calculated approximately by using $(\Psi^{\mathbf{u}_{i+1}})^k(\Psi(J^{\mathbf{u}_i})) \to J^{\mathbf{u}_{i+1}}$.

PROBLEMS FOR SECTION 12.7

1. Consider the problem

$$\max E \sum_{t=0}^{\infty} \beta^t(-u_t^2 - X_t^2), \quad \beta \in (0, 1), \quad u_t \in \mathbb{R}$$

$$X_{t+1} = X_t + u_t + V_t, \quad E(V_{t+1}) = 0, \quad E(V_{t+1}^2) = d$$

(a) Guess that $J(x)$ is of the form $ax^2 + b$, and insert it into (5) to determine a and b.

(b) Solve the corresponding finite horizon problem assuming $J(t, x) = J(t, x, T) = \beta^t(a_t x^2 + b_t)$. (We now sum only up to time T.) Find $J(0, x_0, T)$, let $T \to \infty$ and prove that the solution in (a) is optimal (we are in case B).

2. Solve the problem

$$\max E \sum_{t=0}^{\infty} \alpha^t(\ln u_t + \ln X_t), \quad X_{t+1} = (X_t - u_t)V_{t+1}, \quad x_0 > 0, \quad u_t \in (0, x_t)$$

where $\alpha \in (0, 1)$, $V_t > 0$, and all the V_t are independent and identically distributed with $|E \ln V_t| < \infty$.

TOPOLOGY AND SEPARATION

We could, of course, dismiss the rigorous proof as being superfluous: if a theorem is geometrically obvious why prove it? This was exactly the attitude taken in the eighteenth century. The result, in the nineteenth century, was chaos and confusion: for intuition, unsupported by logic, habitually assumes that everything is much nicer behaved than it really is.
—I. Stewart (1975)

This chapter concentrates on a few topics of a theoretical nature that turn out to be in some parts of economics, notably general equilibrium and its applications to m macroeconomic theory. Section 13.1 takes a closer look at open and closed sets in $\mathbb{R}^n$, to with closely associated concepts such as the neighbourhood of a point, as well as the interi boundary of a set. Next, Sections 13.2 and 13.3 cover the associated concepts of conver compactness, and continuity in $\mathbb{R}^n$. These concepts play an important part in mather analysis. Their systematic study belongs to *general* or *analytic topology*, an important l of mathematics that saw a period of rapid development early in the 20th century. The p definitions and carefully formulated arguments we provide may strike many readers as formal. Their primary purpose is not to give solution methods for concrete problems, equip the reader with the theoretical basis needed to understand why solutions may no exist, as well as their regularity properties when they do exist. In the case of optimi problems, these ideas lead to the versions of the maximum theorem that are the sub Section 13.4.

Another main theme of this chapter is separation theorems, which are useful in both g equilibrium and optimization theory. A discussion of "productive economies" and a disc of Frobenius roots of square matrices wind up the chapter.

13.1 Point Set Topology in $\mathbb{R}^n$

This section begins by reviewing some basic facts concerning the n-dimensional Euc space $\mathbb{R}^n$, whose elements, or points, are n-vectors $\mathbf{x} = (x_1, \dots, x_n)$. The **Euclidea tance** $d(x, y)$ between any two points $\mathbf{x} = (x_1, \dots, x_n)$ and $\mathbf{y} = (y_1, \dots, y_n)$ in $\mathbb{R}^n$ norm $\|\mathbf{x} - \mathbf{y}\|$ of the vector difference between $\mathbf{x}$ and $\mathbf{y}$. (See (1.1.37).) Thus,

$$d(\mathbf{x}, \mathbf{y}) = \|\mathbf{x} - \mathbf{y}\| = \sqrt{(x_1 - y_1)^2 + \cdots + (x_n - y_n)^2}$$