Discrete time optimization

Keywords: Discrete time optimization for finite and infinite horizon, optimal value function, fundamental equations, Bellman equation

Contents

1 Finite Horizon and Fundamental equations

1.1 Introduction

In this chapter we will discuss optimization problems for discrete time processes. We will consider systems/processes that change at discrete times $t = 0, 1, \ldots, T$.

Suppose that the state of the system at time t is completely described by a real number x_t and assume that the initial state x_0 is given and that from then on the system evolves through time under the influence of a sequence of so-called controls u_t, which can be chosen freely from a given set U, called the control region.

Furthermore, we will always assume that the evolution of the system can be described by a given difference equation

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

It is easy to see, that if we choose values for $u_0, u_1, \ldots, u_{T-1}$ the difference equation gives

$$\begin{aligned} x_1 &= g(0, x_0, u_0) \\ x_2 &= g(1, x_1, u_1) = g(1, g(0, x_0, u_0), u_1) \\ &\ldots \end{aligned}$$

and we can compute recursively the states $x_1, x_2, \ldots, x_T$ in terms of the initial state x_0 and the controls $u_0, u_1, \ldots, u_{T-1}$.

Hence each choice of $(u_0, u_1, \ldots, u_{T-1})$ gives rise to a sequence $(x_0, x_1, \ldots, x_T)$. Let us denote these corresponding pairs by $(\{x_t\}, \{u_t\})$ and call them admissible sequence pairs. Of course, different choices of the control sequence will give different sequences of the states.

Example 1.1 *$x_0 = 0$, $T = 2$, $x_{t+1} = x_t + u_t$ and $u_t \in [-1, 1] = U \subset \mathbb{R}$*

- *controls* $(u_0, u_1) = (0, 0)$

$$\begin{aligned} x_0 &= 1 \\ x_1 &= x_0 + u_0 = 1 + 0 = 1 \\ x_2 &= x_1 + u_1 = 1 + 0 = 1 \end{aligned}$$

- *controls* $(u_0, u_1) = (1, -1)$

$$\begin{aligned} x_0 &= 1 \\ x_1 &= x_0 + u_0 = 1 + 1 = 2 \\ x_2 &= x_1 + u_1 = 2 - 1 = 1 \end{aligned}$$

Different pathes of the system through time have usually a 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 of the system is given by the sum

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

The sum is called the objective function and it represents the sum of utilities obtained at each point $t = 0, 1, \ldots, T$ of time.

Remarks 1.1 *The objective function is sometimes defined by*

$$\sum_{t=0}^{T-1} f(t, x_t, u_t) + S(x_T),$$

where S measures the value associated with the terminal state x_T of the system. This seems to be quite natural, because in the state x_T we are ready and there is (mostly) no decision u_T to make or a decision u_T has no impact. But this objective function is a special case of our definition with $f(T, x_T, u_T) = S(x_T)$.

For each admissible pair $(\{x_t\}, \{u_t\})$ the objective function has a definite value and we can now describe the general problem:

Among all admissible pairs $(\{x_t\}, \{u_t\})$ find one, $(\{x_t^*\}, \{u_t^*\})$, that makes the value of the objective function as large as possible or

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

This pair $(\{x_t^*\}, \{u_t^*\})$ is called optimal pair and the sequence u_t^* is called optimal control.

Example 1.2 *Let x_t be an individual's wealth at time t. At each time $t = 0, 1, \ldots, T$ the individual has to decide the proportion $u_t \in [0, 1] = U$ of x_t to consume, leaving the remaining proportion $(1 - u_t)$ for savings. Assume that the wealth earns interest at rate $\rho - 1 > 0$.*

After $u_t \cdot x_t$ has been withdrawn for consumption, the remaining stock of wealth is $(1 - u_t)x_t$. Because of the interest, this grows to the amount

$$x_{t+1} = \rho(1 - u_t)x_t$$

at the beginning of period $t + 1$.

Suppose that the utility of consuming $c_t = u_t \cdot x_t$ is given by $U(t, c_t) = U(t, u_t \cdot x_t)$. The total utility over all periods $t = 0, 1, \ldots, T$ is

$$\sum_{t=0}^{T} U(t, u_t \cdot x_t)$$

The dynamic optimization problem is:

$$\begin{aligned} &\max \sum_{t=0}^{T} U(t, u_t \cdot x_t) \\ &\textit{subject to} \\ &x_{t+1} = \rho(1 - u_t)x_t, \\ &x_0 \textit{ given}, \\ &u_t \in [0, 1] = U. \end{aligned}$$

1.2 The optimal value function

Suppose we have to solve the general problem $(*)$

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

and suppose that at time $t = s$ the state of the system is $x_s = x \in \mathbb{R}$. The best we can do in the remaining periods is to choose $(u_s, u_{s+1}, \ldots, u_{T-1})$ (and perhaps u_T) and thereby $(x_{s+1}, x_{s+2}, \ldots, x_T)$ such that

$$\sum_{t=s}^{T} f(t, x_t, u_t)$$

is maximal with $x_s = x$.

Definition 1.1 *The optimal value function at time s of the problem $(*)$ is*

$$J_s(x) = \max_{u_s, \ldots, u_{T-1} \in U} \sum_{t=s}^{T} f(t, x_t, u_t)$$

where $x_s = x$ and $x_{t+1} = g(t, x_t, u_t)$ for $t > s$.

Important property of J_s Suppose that at time $t = s$ we are in the state $x_s = x$. What is the optimal choice for $u_s = u$?

If we choose $u_s = u$, then at $t = s$

- we obtain the immediate reward $f(s, x, u)$,
- the state of the system at time $t = s + 1$ will be $x_{s+1} = g(s, x, u)$ and
- the highest obtainable value for the reward

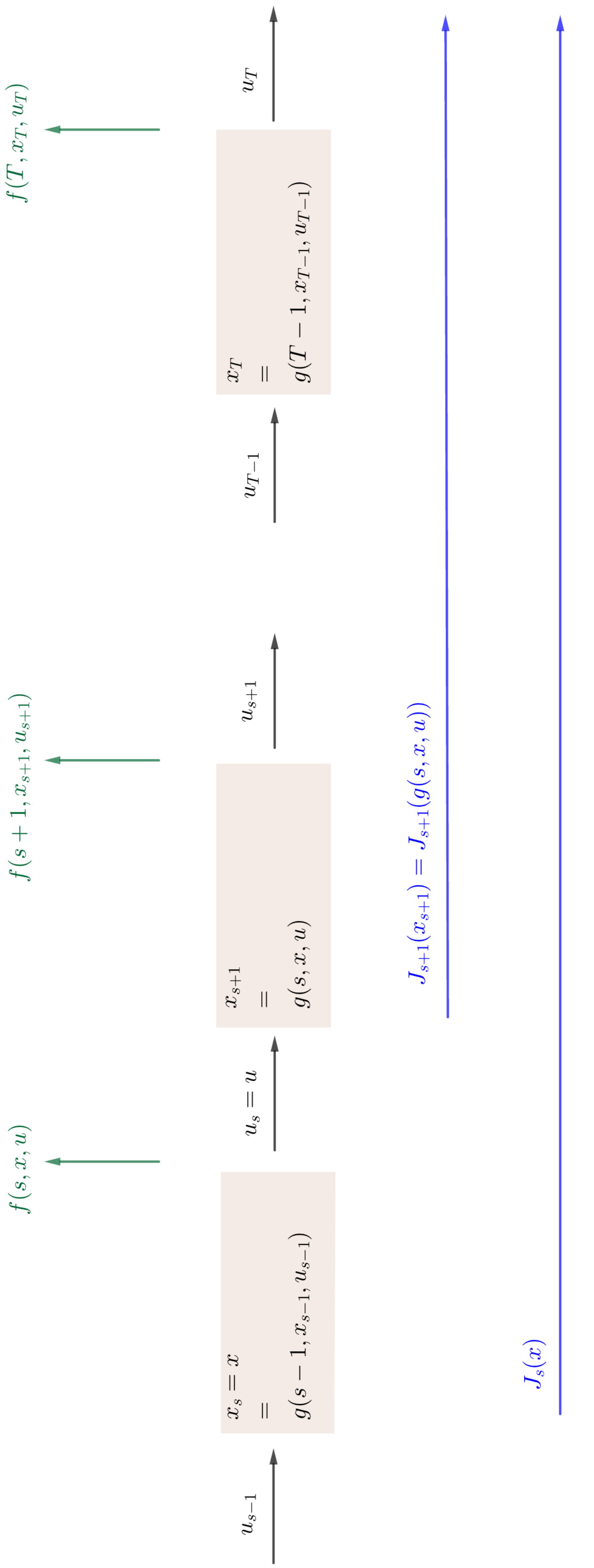
$$\sum_{t=s+1}^{T} f(t, x_t, u_t)$$

starting from the state x_{s+1} is

$$J_{s+1}(x_{s+1}) = J_{s+1}(g(s, x, u)).$$

Hence the best choice of $u = u_s$ at time s must be the value of u that maximizes the sum

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



u_{s-1}
$x_s = x$
$=$
$g(s-1, x_{s-1}, u_{s-1})$
$f(s, x, u)$
$u_s = u$
x_{s+1}
$=$
$g(s, x, u)$
$f(s+1, x_{s+1}, u_{s+1})$
u_{s+1}
u_{T-1}
x_T
$=$
$g(T-1, x_{T-1}, u_{T-1})$
$f(T, x_T, u_T)$
u_T
$J_{s+1}(x_{s+1}) = J_{s+1}(g(s, x, u))$
$J_s(x)$

So we get the basic tool for solving discrete dynamic optimization problems.

Theorem 1.1 (Fundamental equations of dynamic programming (finite horizon)) *Let*

$$\begin{aligned} &\max \sum_{t=0}^{T} f(t, x_t, u_t) \\ &\textit{subject to} \\ &x_{t+1} = g(t, x_t, u_t), \\ &x_0 \textit{ given}, \\ &u_t \in U \end{aligned}$$

be a discrete dynamic optimization problem and $J_t(x)$ the optimal value function for $t = 0, 1, \ldots, T$. Then the sequence of value functions $J_0, J_1, \ldots, J_s, J_{s+1}, \ldots, J_{T-1}, J_T$ satisfies the equations:

$$\begin{aligned} J_s(x) &= \max_{u \in U} \left[f(s, x, u) + J_{s+1}(g(s, x, u)) \right] \quad \textit{for } s = 0, 1, \ldots, T-1 \\ J_T(x) &= \max_{u \in U} f(T, x, u) \end{aligned}$$

Example 1.3 *We use the theorem to solve the problem*

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

Here $T = 3$, $f(t,x,u) = 1 + x - u^2$ *and* $g(t,x,u) = x + u$. *We start with* J_3.

- $J_3(x) = \max_{u \in U} (1 + x - u^2) = 1 + x$ *and* $\underline{u_3^*(x) = 0}$
- $$\begin{aligned} J_2(x) &= \max_{u\in U} [\ f(2,x,u) + J_3(g(2,x,u))\ ] \\ &= \max_{u\in U} [\ f(2,x,u) + J_3(x+u)\ ] \\ &= \max_{u\in U} \left[\ 1 + x - u^2 + 1 + (x+u)\ \right] \\ &= \max_{u\in U} \left[\ 2 + 2x + u - u^2\ \right] \\ &= 2.25 + 2x \end{aligned}$$

  because $\frac{d}{du}(2 + 2x + u - u^2) = 1 - 2u = 0$ *if* $\underline{u = u_2^*(x) = 0.5}$ *and this point is a global maximizer of the function (have a look at the second derivative of* $2 + 2x + u - u^2$*).*
- $$\begin{aligned} J_1(x) &= \max_{u\in U} [\ f(1,x,u) + J_2(g(1,x,u))\ ] \\ &= \max_{u\in U} [\ f(1,x,u) + J_2(x+u)\ ] \\ &= \max_{u\in U} \left[\ 1 + x - u^2 + 2.25 + 2(x+u)\ \right] \\ &= \max_{u\in U} \left[\ 14/3 + 3x + 2u - u^2\ \right] \\ &= 4.25 + 3x \end{aligned}$$

  because $\frac{d}{du}(14/3+3x+2u-u^2) = 2-2u = 0$ *if* $\underline{u = u_1^*(x) = 1}$ *and this point is a global maximizer of the function (have a look at the second derivative of* $14/3+3x+2u-u^2$*).*
- $$\begin{aligned} J_0(x) &= \max_{u\in U} [\ f(0,x,u) + J_1(g(0,x,u))\ ] \\ &= \max_{u\in U} [\ f(0,x,u) + J_1(x+u)\ ] \\ &= \max_{u\in U} \left[\ 1 + x - u^2 + 4.25 + 3(x+u)\ \right] \\ &= \max_{u\in U} \left[\ 21 + 4x + 3u - u^2\ \right] \\ &= 7.5 + 4x \end{aligned}$$

  because $\frac{d}{du}(21+4x+3u-u^2) = 3-2u = 0$ *if* $\underline{u = u_0^*(x) = 1.5}$ *and this point is a global maximizer of the function (have a look at the second derivative of* $21+4x+3u-u^2$*).*

In this particular case the optimal controls are constants, independent of the state of the system, generally this is not the case. The corresponding sequence of the states is:

$$\begin{aligned} x_0 = x_0^* &= 0 \\ x_1^* &= x_0 + u_0^* = 1.5 \\ x_2^* &= x_1^* + u_1^* = 2.5 \\ x_3^* &= x_2^* + u_2^* = 3. \end{aligned}$$

Alternative solution *This simple problem can also be solved by ordinary methods. We have to maximize the objective function and use the difference equation to eliminate x_1, x_2 and x_3 (and use $x_0 = 0$):*

$$\begin{aligned} & \sum_{t=0}^{3} (1 + x_t - u_t^2) \\ = \quad & (1 + x_0 - u_0^2) + (1 + x_1 - u_1^2) + (1 + x_2 - u_2^2) + (1 + x_3 - u_3^2) \\ = \quad & (1 + x_0 - u_0^2) + (1 + (x_0 + u_0) - u_1^2) + (1 + (x_1 + u_1) - u_2^2) + (1 + (x_2 + u_2) - u_3^2) \\ = \quad & (1 + x_0 - u_0^2) + (1 + (x_0 + u_0) - u_1^2) + (1 + ((x_0 + u_0) + u_1) - u_2^2) \\ & +(1 + ((x_1 + u_1) + u_2) - u_3^2) \\ = \quad & \ldots \\ = \quad & 4 + 3u_0 - u_0^2 + 2u_1 - u_1^2 + u_2 - u_2^2 - u_3^2 \\ =: \quad & I(u_0, u_1, u_2, u_3) \end{aligned}$$

To maximize the function I, we have to determine the stationary points by solving the system of equations:

$$\nabla I(u_0, u_1, u_2, u_3) = \begin{pmatrix} 3 - 2u_0 \\ 2 - 2u_1 \\ 1 - 2u_2 \\ -2u_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}.$$

We get $(u_0, u_1, u_2, u_3) = (1.5, 1, 0.5, 0)$ and this point is the global maximizer of I, because I is a concave function.

2 Infinite Horizon and Bellmann equations

2.1 Introduction

In this part we will consider and study the following infinite horizon version of the standard problem of discrete dynamic optimization:

$$\max \sum_{t=0}^{\infty} \beta^t f(x_t, u_t)$$
$$\text{subject to}$$
$$x_{t+1} = g(x_t, u_t),$$
$$x_0 \text{ given},$$
$$u_t \in U \subset \mathbb{R}.$$

The number β is a discount factor, this means $0 < \beta < 1$. Because neither f nor g depend expicitely on t, the problem is called autonomous. As before, a sequence pair $(\{x_t\}, \{u_t\})$ is called admissible, if each $u_t \in U$, the initial state of the system is x_0 and the difference equation $x_{t+1} = g(x_t, u_t)$ is satisfied for all $t = 0, 1, \ldots$ For simplicity, we will assume that f satisfies a **boundary condition**

$$M_1 \leq f(x, u) \leq M_2$$

for all (x, u) with $u \in U$, where M_1, M_2 are constants. Because $0 < \beta < 1$, the sum will therefor always converge.

For any given time $s = 0, 1, 2, \ldots$ and any given state $x = x_s$ take any (infinite) control sequence

$$\mathbf{u}_{\geq s} = (u_s, u_{s+1}, \ldots)$$

where $u_t \in U$ for $t = s, s+1, \ldots$

The states generated by this control sequence are given by $x_{t+1} = g(x_t, u_t)$ with $x = x_s$. The discounted sum of the infinite utility sequence at time s, starting from the state x and obtained by applying the control sequence $\mathbf{u}_{\geq s}$, is

$$V_s(x, \mathbf{u}_{\geq s}) = \sum_{t=s}^{\infty} \beta^t f(x_t, u_t) = \beta^s \underbrace{\sum_{t=s}^{\infty} \beta^{t-s} f(x_t, u_t)}_{=:V^s(x, \mathbf{u}_{\geq s})}.$$

We notice the difference between the two functions explicitely:

- V_s measures all benefits from time s on, discounted to the fixed initial time $t = 0$
- V^s measures all benefits from time s on, discounted to the variable time $t = s$.

Now let

$$J_s(x) \; = \; \max_{\mathbf{u}_{\geq s}} \; V_s(x, \mathbf{u}_{\geq s}) \; = \; \beta^s \underbrace{\max_{\mathbf{u}_{\geq s}} \; V^s(x, \mathbf{u}_{\geq s})}_{=: J^s(x)}$$

Thus, $J_s(x)$ is the maximal (to time $t = 0$) discounted utility that can be obtained over all periods from $t = s$ to ∞, given that the system starts in the state x at time $t = s$.

Definition 2.1 *$J_s(x)$ is called the* <u>*optimal value function*</u> *of the problem.*

Lemma 2.1 *For all $s \geq 0$ we have $J^s(x) \; = \; J^0(x)$.*

Proof: Because the problem is autonomous and we start in the same state x, the future looks exactly the same at either time 0 or time s. This is not true if the horizon is finite. So finding

$$J^s(x) \; = \; \max_{\mathbf{u}_{\geq s}} \; V^s(x, \mathbf{u}_{\geq s}) \quad \text{or} \quad J^0(x) \; = \; \max_{\mathbf{u}_{\geq 0}} \; V^0(x, \mathbf{u}_{\geq 0})$$

requires solving the same optimization problem. □

Of course, we have

$$J_s(x) \; = \; \beta^s \; J^s(x) \; = \; \beta^s \; J^0(x)$$

and we define

$$J(x) \; := \; J_0(x) \; = \; J^0(x).$$

We see that if we know $J(x) = J_0(x)$, then we know $J_s(x)$ for all s. The following main result is the analogous to the fundamental equations for finite horizon problems.

2.2 The Bellman equation

Theorem 2.1 (Fundamental equations of dynamic programming (infinite horizon)) *Let*

$$\max \sum_{t=0}^{\infty} \beta^t f(x_t, u_t)$$

$$\text{subject to}$$

$$x_{t+1} = g(x_t, u_t),$$

$$x_0 \text{ given},$$

$$u_t \in U \subset \mathbb{R}$$

be a discrete dynamic optimization problem. The value function $J(x)$ of this problem satisfies the so-called *Bellman equation:*

$$\boxed{J(x) = \max_{u \in U} [\, f(x,u) + \beta J(g(x,u)) \,].}$$

The Bellman equation is a so-called functional equation, because the unknown function J appears on both sides of the equation. It is not clear, that such an equation has a (unique) solution. You may understand that it is very difficult to use the equation to solve the optimization problem, because maximizing the right-hand side to get J on the left-hand side requires the function J. Hence it may be necessary to guess the structure of J.

Example 2.1 *Consider the following problem*

$$\max \sum_{t=0}^{\infty} \beta^t \left(-\frac{2}{3}x_t^2 - u_t^2 \right)$$

$$\textit{subject to}$$

$$\begin{aligned} x_{t+1} &= x_t + u_t, \\ &x_0 \textit{ given}, \\ &u_t \in \mathbb{R}. \end{aligned}$$

We have $f(x,u) = -\frac{2}{3}x^2 - u^2$ *and* $g(x,u) = x + u$. *The Bellman equation of the problem is*

$$J(x) = \max_{u \in \mathbb{R}} \left[-\frac{2}{3}x^2 - u^2 + \beta J(x+u) \right].$$

What to do now? How can we find a function J *with this property? We guess (ingeniously) the type of* J *and try* $J(x) = -\alpha x^2$ *for a suitable* $\alpha \in \mathbb{R}$.

We get the equation

$$-\alpha x^2 = \max_{u \in \mathbb{R}} \left[-\frac{2}{3}x^2 - u^2 + \beta(-\alpha)(x+u)^2 \right]$$

and the funtion $h(u) = -\frac{2}{3}x^2 - u^2 - \alpha\beta(x+u)^2$ *realizes the global maximum in*

$$u^*(x) = -\frac{\alpha\beta}{(1-\alpha\beta)}x.$$

Using this value for u *in the Bellman equation we get*

$$-\alpha x^2 = -\frac{2}{3}x^2 - \left(-\frac{\alpha\beta}{(1-\alpha\beta)}x \right)^2 - \alpha\beta \left(x - \frac{\alpha\beta}{(1-\alpha\beta)}x \right)^2$$

We can cancel x^2 *and with Maple I found*

$$\alpha = \frac{5\beta - 3 \pm \sqrt{25\beta^2 - 6\beta + 9}}{6\beta}.$$