CHAPTER 1

Optimal Control in Deterministic Case

Our starting point is the discussion on the special, deterministic case of the theory of optimal stochastic control. We start with the analysis of discrete-time systems.

1. Discrete-time systems, finite horizon

The main tool used in this case is the so-called dynamic programming, an algorithm which enables to solve a certain class of problems, by an induction argument which reduces them to simpler sub-problems. Or, to put it in the reverse direction, the approach allows to tackle difficult problems by solving simpler ones first and relating these solutions to the harder context by intrinsic recurrence relations. It plays an important role in computer science, as it can be used to construct effective algorithms of polynomial complexity. Furthermore, the method is utilized in optimal planning problems (e.g. in problems of the optimal distribution of resources available, the theory of inventory management, replacement of equipment, etc.).

The basic idea can be formulated as follows. Suppose that a given system $\mathcal{S}$, taking values in some set E, is controlled with a procedure which consists of N steps, where N is a fixed positive integer. At the beginning, the system is in some state $x_0 \in E$. At the m-th step of the procedure, there is a possibility of applying a class of controls, each of which transforms the state $x_{m-1} \in E$ obtained through the previous operations into some new state $x_m \in E$. Formally, if u_m denotes the control applied at m-th step, then we have the identity

$$x_m = f_m(x_{m-1}, u_m) \tag{1.1}$$

for some f_m (the “transition function associated with m-th step”). An important feature of the method is the absence of after-effects, i.e. the controls selected for a given step may only affect the state of the system at that moment. We also emphasize that the class of controls applicable at m-th step may depend on the value of x_{m-1}: we have $u_m \in U(x_{m-1})$ for some function U. As the result of the whole procedure u_1, u_2, $\ldots$, u_N, the system is converted from the state x_0 into the final state x_N. Now, given the initial state x_0 and a fixed objective function

$$J^N(x_0, u_1, u_2, \ldots, u_N) = \sum_{m=1}^{N} g_m(x_{m-1}, u_m) + r(x_N), \tag{1.2}$$

there are two aspects worth investigating:

• to identify the controls u_1^*, u_2^*, $\ldots$, u_N^* (if exist), which yield the optimal performance, i.e., such that we have

$$J^N(x_0, u_1^*, u_2^*, \ldots, u_N^*) = \sup_{u_1, u_2, \ldots, u_N} J^N(x_0, u_1, u_2, \ldots, u_N).$$

• to compute the explicit value $\sup_{u_1,u_2,\dots,u_N} J^N(x_0,u_1,u_2,\dots,u_N)$.

Sometimes, depending on the context, we might be interesting in the analysis of only one of the above aspects: we will illustrate this on some later examples.

Conventional methods of tackling such problems are often either inapplicable, or involve lengthy and elaborate calculations. Dynamic programming allows to solve the above problem by the following recursive formula.

THEOREM 1.1. *Under the above notation, define measurable functions B_N, B_{N-1}, …, B_0 on E by $B_N(x) = r(x)$ and the Bellman equation*

$$B_{n-1}(x) = \sup_{u\in U(x)} \big(g_n(x,u) + B_n(f_n(x,u))\big), \quad n = N,\, N-1,\, \dots,\, 1. \tag{1.3}$$

Then for any strategy u_1, u_2, …, u_N we have $J^N(x_0,u_1,u_2,\dots,u_N) \le B_0(x_0)$. Furthermore, if for any n and any $x \in E$ there is a control $\hat u = \hat u_n(x)$ for which

$$B_{n-1}(x) = g_n(x,\hat u) + B_n(f_n(x,\hat u))\big),$$

then the optimal strategy is given by $u^*_n = \hat u_n(x_{n-1})$, $n = 1,\, 2,\, \dots,\, N$.

PROOF. Consider the subsystem which starts at time n and then evolves according to (1.1) for $m = n+1$, $n+2$, …, N. Introduce the truncated functional

$$J^N_n(x_n,u_{n+1},u_{n+2},\dots,u_N) = \sum_{m=n+1}^{N} g_m(x_{m-1},u_m) + r(x_N).$$

Then, as we will show inductively, we have $J^N_n(x_n,u_{n+1},u_{n+2},\dots,u_N) \le B_n(x_n)$. Indeed, we have $J^N_N(x_N) = r(x_N) = B_N(x_N)$ and the induction step follows from

$$\begin{aligned} &J^N_{n-1}(x_{n-1},u_n,u_{n+1},\dots,u_N)\\ &\qquad = g_n(x_{n-1},u_n) + J^N_n(f_n(x_{n-1},u_n),u_{n+1},\dots,u_N)\\ &\qquad \le g_n(x_{n-1},u_n) + B_n(f_n(x_{n-1},u_n)) \le B_{n-1}(x_{n-1}). \end{aligned}$$

Plugging $n = 1$, we get the first part of the assertion. To get the second part, note that for the above special controls u^*_1, u^*_2, …, u^*_N, all the above inequalities become equalities. □

Several helpful comments are in order.

I. It is worth to rewrite the proof of $J^N(x_0,u_1,u_2,\dots,u_N) \le B_0(x_0)$ in the form

$$\begin{aligned} B_0(x_0) &\ge g_1(x_0,u_1) + B_1(x_1)\\ &\ge g_1(x_0,u_1) + g_2(x_1,u_2)B_2(x_2)\\ &\ge \dots \ge \sum_{m=1}^{n} g_m(x_{m-1},u_m) + B_N(x_N) = J^N(x_0,u_1,u_2,\dots,u_N). \end{aligned}$$

II. It follows from the above proof that the Bellman sequence admits the alternative definition

$$B_n(x) = \sup_{u_{n+1},u_{n+2},\dots,u_N} \Big\{ J^N_n(x_n,u_{n+1},u_{n+2},\dots,u_N) \,\big|\, x_n = x \Big\}$$

for $n = 0,\, 1,\, 2,\, \dots,\, N$. Thus in particular $B_0(x_0)$ is the desired optimal value of the function J^N. As we will see later, writing down the above abstract formula for B_n is a convenient start for the analysis.

III. The above argumentation leads to the following optimality principle: given x_0, if there is an optimal strategy u_1^*, u_2^*, …, u_N^* for J^N, then

· for any n, the strategy u_{n+1}^*, u_{n+2}^*, …, u_N^* must be optimal for the functional $J_n^N(x_n, u_{n+1}, u_{n+2}, \ldots, u_N)$ (where x_n comes from x_0 by applying u_1^*, u_2^*, …, u_n^*);

· for any n, the strategy u_1^*, u_2^*, …, u_n^* must be optimal for the functional

$$J^n(x_0, u_1, u_2, \ldots, u_n) = \sum_{m=1}^{n} g_m(x_{m-1}, u_m) + B_n(x_n).$$

IV. The above discussion concerns the case in which we are interested in the largest value of J^N. Sometimes one wants to minimize J^N: then all the argumentation works, we only need to replace suprema by infima in appropriate places.

V. The form of the Bellman equation (1.3) is strictly connected to the additive form of the functional. As we shall see later, the general approach extends to the case of other functionals, for which the associated Bellman equations look differently.

The above statements give a transparent method of handling the problem. We solve the system (1.3) of functional equations, obtaining $B_0(x_0)$, the value of the optimal performance under the assumption that the system is initially in the state x_0. We also identify the optimal controls u_1^*, u_2^*, …, u_N^*: these are the parameters for which the suprema in (1.3) are attained. Specifically, having determined B_0, B_1, …, B_N (called the Bellman sequence in the sequel), we find u_1^* by the requirement

$$B_0(x_0) = g_1(x_0, u_1^*) + B_1(f_1(x_0, u_1^*)).$$

Let $x_1 = f_1(x_0, u_1^*)$ be the position of the system after the optimal first move. Then we get the control u_2^* from the equation

$$B_1(x_1) = g_2(x_1, u_2^*) + B_2(f_2(x_1, u_2^*)),$$

and so on. Summarizing, the approach rests on solving the initial problem by embedding it into a class of similar sub-problems, the collection of which can be treated, as a whole, by means of the recursive formulas. We should point out, however, that still, in many cases, the analysis of the obtained setting can be technically involved and require plenty of laborious computations.

In what follows, we will see the above reasoning in various disguises. We have purposefully restrained ourselves from the discussion concerning the state space or the class of feasible controls; this would probably complicate the above presentation, as these objects can be multidimensional or change from step to step. Sometimes it will be convenient to enumerate the steps by numbers 1, 2, …, N instead of 0, 1, …, N. Moreover, in many cases it will be more natural to work with the reversed sequence B_N, B_{N-1}, …, B_1 instead of B_1, B_2, …, B_N (then the lower index indicates the number of steps up to the termination of the process). However, the main idea remains essentially unchanged.

Instead of exploring further the abstract description, we continue with the analysis of several examples which will serve as an illustration of the above concepts.

1.1. A warm up. Consider an investor, whose capital at n-th day is equal to x_n, $n = 0, 1, 2, \dots, N$. At m-th day, the investor consumes u_m of the capital and invests the remaining part; the interest rate is equal to $\gamma > 1$. That is, the sequence $(x_n)_{n=0}^N$ is governed by the equation

$$x_m = \gamma(x_{m-1} - u_m), \qquad m = 1, 2, \dots, N,$$

where $u_m \leq x_{m-1}$ for each m. Suppose that the purpose of the investor is to maximize the functional $J(x_0, u_1, u_2, \dots, u_N) = \sum_{m=1}^N u_m$. This can be easily solved by the above approach. Introduce the Bellman sequence

$$B_n(x) = \sup\left\{ \sum_{m=n+1}^N u_m \,:\, x_n = x \right\}.$$

We have $B_N(x) = 0$ and the Bellman equation gives

$$B_{N-1}(x) = \sup_{u \leq x} \left\{ u + B_N(\gamma(x-u)) \right\} = \sup_{u \leq x} u = x.$$

(This is perfectly intuitive: "there is no tomorrow", so the investor spends all the money). Furthermore,

$$B_{N-2}(x) = \sup_{u \leq x} \left\{ u + B_{N-1}(\gamma(x-u)) \right\} = \sup_{u \leq x} \left\{ u + \gamma(x-u) \right\} = \gamma x$$

(with supremum attained at zero),

$$B_{N-3}(x) = \sup_{u \leq x} \left\{ u + B_{N-2}(\gamma(x-u)) \right\} = \sup_{u \leq x} \left\{ u + \gamma^2(x-u) \right\} = \gamma^2 x$$

(with supremum attained at zero), etc.: for any $n = 0, 1, 2, \dots, N-1$ we have $B_n(x) = \gamma^{N-n-1}x$ and the appropriate control is equal to zero. This implies that the investor should save the money for the last day, and then spend all the capital. This answer is obvious: keeping the capital unchanged maximizes the profit obtained via the interest rate.

1.2. Towards analytic applications. For a fixed N, we will compute the quantity

$$\sup\left\{ \frac{a_1}{a_0 + a_1} + \frac{a_2}{a_1 + a_2} + \ldots + \frac{a_N}{a_{N-1} + a_N} \right\},$$

where the supremum is taken over all positive numbers $a_0, a_1, a_2, \dots, a_N$. This problem can be studied with the use of optimal control: consider the strategy u_1, $u_2, \dots, u_N$ given by $u_n = a_n$. Then the sequence $x_0, x_1, x_2, \dots, x_N$ given by $x_n = a_n$ consists of positive numbers and satisfies (1.1) with $f_n(x, u) = u$. Our goal is to maximize the functional

$$J^N(x_0, u_1, u_2, \dots, u_N) = \sum_{m=1}^N \frac{u_m}{x_{m-1} + u_m},$$

which is of the form (1.2), with $g_m(x, u) = u/(x+u)$ and $r(x) = 0$. The Bellman sequence is given by

$$B_n(x) = \sup\left\{ \frac{a_{n+1}}{x_n + a_{n+1}} + \frac{a_{n+2}}{a_{n+1} + a_{n+2}} + \ldots + \frac{a_N}{a_{N-1} + a_N} \,\Big|\, x_n = x \right\},$$

where $x > 0$ and the supremum is taken over all positive numbers a_{n+1}, a_{n+2}, …, a_N. By the Bellman equation, we see that for any $x > 0$ we have $B_N(x) = 0$,

$$\begin{aligned} B_{N-1}(x) &= \sup_{u>0} \{g_N(x,u) + B_N(f_N(x,u))\} \\ &= \sup_{u>0} \left\{ \frac{u}{x+u} + B_N(u) \right\} = \sup_{u>0} \frac{u}{x+u} = 1, \\ B_{N-2}(x) &= \sup_{u>0} \left\{ \frac{u}{x+u} + B_{N-1}(u) \right\} = \sup_{u>0} \left\{ \frac{u}{x+u} + 1 \right\} = 2, \end{aligned}$$

and so on,

$$B_0(x) = \sup_{u>0} \left\{ \frac{u}{x+u} + B_1(u) \right\} = \sup_{u>0} \left\{ \frac{u}{x+u} + N - 1 \right\} = N.$$

This gives the answer to the problem: the supremum is equal to N. Note that the optimal controls do not exist.

1.3. A probabilistic inequality. Next, we will prove that for any $N \geq 1$ and any numbers a_1, a_2, …, $a_N \in [0,1]$ we have

$$(1-a_1)(1-a_2)\ldots(1-a_N) \geq 1 - a_1 - a_2 - \ldots - a_N.$$

The first step is to rewrite the inequality in the form

$$-a_1 - a_2 - \ldots - a_N - (1-a_1)(1-a_2)\ldots(1-a_N) \leq -1.$$

This fits perfectly into the above scheme. We consider the strategy $(u_1, u_2, \ldots, u_N) = (a_1, a_2, \ldots, a_N)$ and define the sequence x_0, x_1, x_2, …, x_N by $x_0 = 1$ and

$$x_n = x_{n-1} \cdot (1 - a_n).$$

Then the sequence takes values in $[0,1]$ and satisfies the evolution equation (1.1) with $f_n(x,u) = x(1-u)$. We need to maximize the functional

$$J(x_0, u_1, u_2, \ldots, u_N) = \sum_{m=1}^{N} (-u_m) - x_N,$$

which is of the form (1.2) with $g_m(x,u) = -u$ and $r(x) = -x$. By the above discussion, we introduce the Bellman sequence by

$$B_n(x) = \sup \left\{ \sum_{m=n+1}^{N} (-u_m) - x_N \,\middle|\, x_n = x \right\}, \qquad x \in [0,1],$$

where the supremum is taken over all u_{n+1}, u_{n+2}, …, $u_N \in [0,1]$. By the very definition, we have $B_N(x) = -x$ and, by the Bellman equation,

$$B_{N-1}(x) = \sup_{u\in[0,1]} \Big(-u + B_N(x(1-u)) \Big) = \sup_{u\in[0,1]} \big(u(x-1) - x \big) = -x,$$

where the supremum is attained for $u = 0$. The remaining functions B_{N-2}, B_{N-3}, …, B_0 are computed identically: we have $B_0(x) = B_1(x) = \ldots = B_N(x) = -x$ for all n and therefore

$$J(x_0, u_1, u_2, \ldots, u_N) \leq B_0(x_0) = -1,$$

which is the claim. Note that equality holds if and only if all the controls are zero: $a_1 = a_2 = \ldots = a_N = 0$.

1.4. Modified approach: AM-GM inequality. Now we will show how dynamic programming yields one of the most fundamental inequalities.

THEOREM 1.2. *For any positive integer N and any nonnegative numbers a_1, a_2, …, a_N we have*

$$\frac{a_1 + a_2 + \ldots + a_N}{N} \geq (a_1 a_2 \ldots a_N)^{1/N}.$$

Equality holds if and only if $a_1 = a_2 = \ldots = a_N$.

PROOF. Although the inequality fits into the scheme developed above (see Remark 1.3 below), it is instructive to discuss a slightly different approach. As previously, we start with rephrasing the desired claim into the problem of the optimization of some value function. This can be done as follows. Fix a nonnegative number x and consider the quantity

$$\sup\Big\{a_1 a_2 \ldots a_N \,:\, a_1,\, a_2,\, \ldots,\, a_N \geq 0,\; a_1 + a_2 + \ldots + a_N = x\Big\}.$$

We need to show that for any x, the above supremum does not exceed $(x/N)^N$. Note that this problem is not of the form discussed at the beginning, since the functional does not have the additive form. However, the general methodology of "decomposing" the problem into simpler, similar sub-problems, which are then connected via induction argument, applies. Again, the controls are the numbers a_1, a_2, …, a_N. The sequence x_0, x_1, x_2, …, x_N is given by $x_0 = x$ and

$$x_n = x_{n-1} - a_n, \qquad n = 1,\, 2,\, \ldots,\, N,$$

so that $x_n = a_{n+1} + a_{n+2} + \ldots + a_N$. The Bellman sequence is given by

$$B_n(x) = \sup\Big\{a_{n+1} a_{n+2} \ldots a_N \,\Big|\, x_n = x\Big\}, \qquad n = 0,\, 1,\, 2,\, \ldots,\, N-1.$$

The dynamic approach rests on writing the system of equations which govern the evolution of the Bellman sequence. By the very definition, we have $B_{N-1}(x) = x$, and the version of Bellman equation is the following: for any $n = 0, 1, \ldots, N-2$ and any $x \geq 0$ we have

$$B_n(x) = \sup_{t \in [0,x]} \big\{B_{n+1}(x-t) \cdot t\big\}. \tag{1.4}$$

Indeed, if a_{n+1}, a_{n+2}, …, a_N are arbitrary nonnegative numbers summing up to x and we denote $t = a_{n+1} \in [0, x]$, then

$$a_{n+1} a_{n+2} \ldots a_N = t \cdot a_{n+2} a_{n+3} \ldots a_N \leq B_{n+1}(x-t) \cdot t,$$

by the definition of B_{n+1} and the fact that $a_{n+2} + a_{n+3} + \ldots + a_N = x - t$. Taking the supremum over all a_{n+1}, a_{n+2}, …, a_N as above, gives the inequality "$\leq$" in (1.4). To get the reverse, we fix $t \in [0, x]$ and nonnegative numbers a_{n+2}, a_{n+3}, …, a_N summing up to $x - t$. Then $t + a_{n+2} + a_{n+3} + \ldots + a_N = x$, so the definition of B_n yields

$$t \cdot a_{n+2} a_{n+3} \ldots a_n \leq B_n(x).$$

Now, taking the supremum over a_{n+2}, a_{n+3}, …, a_N as above gives $tB_{n+1}(x-t) \leq B_n(x)$, and since $t \in [0, x]$ was arbitrary, the identity (1.4) follows.

It remains to solve the recurrence. In general, this might be quite involved, but here the *conjecture* for the formula for B_n is directly encoded in the problem:

$$B_n(x) = (x/(N-n))^{N-n}. \tag{1.5}$$

This conjecture is easily confirmed by induction. Let us briefly present the calculations, as they will be useful in the identification of the optimal controls. For $n = N-1$ the hypothesis is true. Assuming the validity for a fixed $n+1 \in \{2,3,\ldots,N-1\}$, we derive that the expression

$$B_{n+1}(x-t)\cdot t = \left(\frac{x-t}{N-n-1}\right)^{N-n-1}\cdot t,$$

considered as a function of t, attains its maximum for $t = x/(N-n)$ (only). Furthermore, this maximal value is equal to $(x/(N-n))^{N-n}$. This yields (1.5) and the claim follows.

The above calculations encode, for any fixed $x \geq 0$, the optimal controls a_1^*, a_2^*, ..., a_N^* for $B_0(x)$. To see this, assume that $x > 0$ (for $x = 0$ there is nothing to prove). We go back to the above proof of (1.4). We have

$$B_0(x) = \sup_{t\in[0,x]}\big\{B_1(x-t)\cdot t\big\},$$

and the supremum is attained for the unique choice $t = x/N$. This necessarily implies that a_1^* must be equal to x/N and $a_2^* + a_3^* + \ldots + a_N^* = (N-1)x/N$. To get a_2^*, we make use of the following version of the optimality principle. We have

$$a_1^*a_2^*\ldots a_N^* = B_1(x) = B_2\left(x - \frac{x}{N}\right)\cdot\frac{x}{N} = B_2\left(x-\frac{x}{N}\right)\cdot a_1^*,$$

or, equivalently (recall that we have assumed $x > 0$)

$$a_2^*a_3^*\ldots a_N^* = B_2\left(x-\frac{x}{N}\right).$$

This brings us to the same position as above, with the length of the unknown extremal sequence decreased by 1 (and the required sum x replaced by $x - x/N$). Repeating the arguments, we show that $a_2^* = (x-x/N)/(N-1) = x/N$, $a_3^* + a_4^* + \ldots + a_N = x - 2x/N$, and the numbers a_3^*, a_4^*, ..., a_N^* satisfy

$$a_3^*a_4^*\ldots a_N^* = B_3\left(x - \frac{2x}{N}\right),$$

and so on. The procedure can be carried out until we get all the values of a_1^*, a_2^*, ..., a_N^*: one easily checks by induction that $a_1^* = a_2^* = \ldots = a_N^* = x/N$ is the extremal sequence we have searched for. □

REMARK 1.3. There is an alternative approach to the AM-GM estimate. Consider the sequence $x_0 = 1$, $x_1 = a_1$ and

$$x_n = \frac{n^n}{(n-1)^{n-1}}\cdot x_{n-1}\cdot a_n, \qquad n = 1,\, 2,\, \ldots,\, N.$$

Then we have $x_n = n^n a_1 a_2 \ldots a_n$ and the inequality is equivalent to showing that the functional

$$-a_1 - a_2 - \ldots - a_N + x_N^{1/N}$$

is nonpositive. This fits into the scheme developed above.

1.5. A higher-dimensional DP problem. We turn to the case in which Bellman functions depend on more than one variable.

THEOREM 1.4. *For any nonnegative numbers* $a_1, a_2, \ldots, a_N, b_1, b_2, \ldots, b_N$ *we have*

$$(a_1 + b_1)(a_2 + b_2) \ldots (a_N + b_N) \geq ((a_1 a_2 \ldots a_N)^{1/N} + (b_1 b_2 \ldots b_N)^{1/N})^N. \tag{1.6}$$

PROOF. We may assume that all numbers a_i and b_j are non-zero (otherwise, the claim is obvious). For any x, $y > 0$, consider the function

$$B_n(x, y) = \inf \Big\{ (a_1 + b_1)(a_2 + b_2) \ldots (a_n + b_n) \Big\},$$

where the infimum is taken over all sequences $a_1, a_2, \ldots, a_n, b_1, b_2, \ldots, b_n$ of positive numbers such that $a_1 a_2 \ldots a_n = x$ and $b_1 b_2 \ldots b_n = y$. Clearly, we have $B_1(x, y) = x + y$, and Bellman equation becomes

$$B_{n+1}(x, y) = \inf \Big\{ (s + t) B_n(x/s, y/t) \Big\}, \tag{1.7}$$

where the infimum is taken over all s, $t > 0$. One easily shows by induction that the solution to this recurrence is given by $B_n(x, y) = (x^{1/n} + y^{1/n})^n$, and the infimum in (1.7) is attained for s, t such that $y/x = (t/s)^{n+1}$. This establishes the desired inequality; furthermore, we obtain that the optimal controls $a_1^*, a_2^*, \ldots, a_N^*, b_1^*, b_2^*, \ldots, b_N^*$ satisfy

$$\frac{a_1^* a_2^* \ldots a_{n+1}^*}{b_1^* b_2^* \ldots b_{n+1}^*} = \left(\frac{a_{n+1}^*}{b_{n+1}^*} \right)^{1/(n+1)}, \qquad n = 1, 2, \ldots, N-1,$$

i.e., the equality in (1.6) is attained if and only if $a_1/b_1 = a_2/b_2 = \ldots = a_N/b_N$. □

2. Discrete system, infinite horizon

Now we turn our attention to the case in which we are interested in the control evolving in an infinite number of sets. There are essentially two types of approaches, it is best to present them on a concrete example.

THEOREM 1.5. *Prove that for an infinite sequence* $(a_n)_{n=1}^\infty \subset \mathbb{R}_+$ *we have*

$$\prod_{n=1}^\infty a_n \leq \sum_{n=1}^\infty \frac{a_n^{2^n}}{2^n},$$

provided the product makes sense.

PROOF, THE FIRST APPROACH. The idea is to make use of the method for the finite horizon, and then let the horizon go to infinity. More precisely, for any $N \geq 1$, introduce the Bellman functions $B_N : \mathbb{R}_+ \to \mathbb{R}$ by

$$B_N(x) = \sup \left\{ a_1 a_2 \ldots a_N \; : \; \sum_{n=1}^N \frac{a_n^{2^n}}{2^n} = x \right\}.$$

We have $B_1(x) = \sqrt{2x}$ and the Bellman equation reads

$$B_N(x) = \sup \left\{ a B_{N-1} \left(x - \frac{a^{2^N}}{2^N} \right) \right\},$$

where the supremum is taken over all a such that $a^{2^N}/2^N \leq x$. This recurrence is not difficult to solve: after some lengthy, but rather straightforward calculations we compute that

$$B_N(x) = \left(\frac{x}{\frac{1}{2}+\frac{1}{4}+\ldots+\frac{1}{2^N}}\right)^{\frac{1}{2}+\frac{1}{4}+\ldots+\frac{1}{2^N}}.$$

It remains to perform a limiting argument: for any sequence $(a_n)_{n\geq 1}$ as in the statement, we have

$$a_1a_2\ldots a_N \leq B_N\left(\sum_{n=1}^N \frac{a_n^{2^n}}{2^n}\right).$$

This becomes the desired estimate in the limit. □

Sometimes the first approach fails. It happens quite often that the Bellman sequence is extremely difficult (or impossible) to compute explicitly, however, its limit version can be handled efficiently. Here are the details.

PROOF, THE SECOND APPROACH. We introduce the *single* Bellman function $B:\mathbb{R}_+\to\mathbb{R}$ given by

$$B(x) = \sup\left\{\prod_{n=1}^\infty a_n \,:\, \sum_{n=1}^\infty \frac{a_n^{2^n}}{2^n} = x\right\}.$$

Now the key is that the Bellman equation becomes a *functional equation* for B. Take any sequence $(a_n)_{n\geq 1}$ and note that

$$\prod_{n=1}^\infty a_n = a_1\prod_{n=2}^\infty a_n,$$

$$\sum_{n=1}^\infty \frac{a_n^{2^n}}{2^n} = \frac{a_1}{2} + \sum_{n=2}^\infty \frac{a_n^{2^n}}{2^n} = \frac{a_1}{2} + \frac{1}{2}\sum_{n=1}^\infty \frac{(a_{n+1}^2)^{2^n}}{2^n}.$$

Therefore, we see that

$$B(x) = \sup_{a\in[0,\sqrt{2x}]}\left\{a\cdot\sqrt{B(2x-a^2)}\right\}.$$

Now the analysis splits into two steps. First we need to solve the above equation, and then check rigorously that it does yield the desired estimate.

I. Search for B. Note that the above equation does not have unique solution: for example, $B\equiv 0$ satisfies it. To find the formula for B, or rather *guess it*, we look at the abstract definition. Note that the supremum in the definition of B is taken over all sequences $(a_n)_{n\geq 1}$ with $\sum_{n=1}^\infty \frac{a_n^{2^n}}{2^n} = x$. If we multiply a_n by $\lambda^{1/2^n}$, then the total sum multiplies by λ (and hence it is equal to λx). On the other hand, the product multiplies by $\lambda^{1/2+1/4+\cdots} = \lambda$, and hence we must have $B(\lambda x) = \lambda B(x)$. This implies that B should be a linear function: $B(x) = cx$. Let us go back to the Bellman equation:

$$cx = \sup_a a\sqrt{c(2x-a^2)}.$$

We easily check that the supremum is attained for $a=\sqrt{x}$, so $c=1$. This gives us the candidate for the Bellman function: $B(x)=x$.

II. Verification. From the above construction, we know that $B(x) = x$ satisfies the Bellman equation. Therefore, for any sequence $(a_n)_{n\geq 1}$,

$$\begin{aligned}\sum_{n=1}^{\infty} \frac{a_n^{2^n}}{2^n} &= B\left(\frac{a_1^2}{2} + \frac{a_2^4}{4} + \frac{a_3^8}{8} + \dots\right) \\ &\geq a_1 \sqrt{B\left(\frac{a_2^4}{2} + \frac{a_3^8}{4} + \frac{a_4^{16}}{8} + \dots\right)} \\ &\geq a_1 \sqrt{a_2^2 \sqrt{B\left(\frac{a_3^8}{2} + \frac{a_4^{16}}{4} + \frac{a_5^{32}}{8} + \dots\right)}} \\ &= a_1 a_2 \sqrt[4]{B\left(\frac{a_3^8}{2} + \frac{a_4^{16}}{4} + \frac{a_5^{32}}{8} + \dots\right)}\end{aligned}$$

and so on; after a finite (say, N) number of steps, the sequence (a_n) stabilizes at 1, so the arguments inside B become equal to 1. We thus obtain

$$\sum_{n=1}^{\infty} \frac{a_n^{2^n}}{2^n} \geq a_1 a_2 \dots a_N,$$

which is the desired claim. □

We conclude with another instructive example.

THEOREM 1.6. *For any sequence a_1, a_2, …, a_n of numbers belonging to $[0, 1]$, we have*

$$\prod_{k=1}^{n}(1-a_k) \leq 1 - \sum_{k=1}^{n} a_k + \frac{1}{2}\left(\sum_{k=1}^{n} a_k\right)^2.$$

PROOF. We rewrite the inequality in the form

$$\frac{\prod_{k=1}^{n}(1-a_k)}{1 - \sum_{k=1}^{n} a_k + \frac{1}{2}\left(\sum_{k=1}^{n} a_k\right)^2} \leq 1$$

and write down the associated Bellman function

$$B(x,y) = \sup\left\{\frac{x\prod_{k=1}^{n}(1-a_k)}{1 - (y + \sum_{k=1}^{n} a_k) + \frac{1}{2}\left(y + \sum_{k=1}^{n} a_k\right)^2}\right\},$$

the supremum taken over all n. We want to show that $B(1, 0) \leq 1$.

The next step is to write the Bellman equation. We see that picking a_1 makes x go to $x(1-a_1)$ and y go to $y + a_1$: hence

$$B(x,y) = \sup_{a>0} B(x(1-a), y+a).$$

Now, as previously, we need to find the solution of this equation, and we start with guessing. By the very definition of B, we have $B(x, y) = xB(1, y) =: x\varphi(y)$. Furthermore, if we plug $a = 0$, the supremum is attained: therefore,

$$-xB_x(x,y) + B_y(x,y) = \frac{d}{da}B(x(1-a), y+a)\bigg|_{a=0} \leq 0.$$

In the language of φ, this means $-x\varphi(y) + x\varphi'(y) \leq 0$, or $\varphi'(y) \leq \varphi(y)$. Assume equality: we obtain $\varphi(y) = Ke^y$. Thus we have constructed the candidate $B(x, y) =$

Kxe^y. It is easy to check that it satisfies the Bellman equation: $xe^y \geq x(1-a)e^{y+a}$ is equivalent to $e^{-a} \geq 1-a$. Now, if we put $K=1$, then $B(1,0)=1$; furthermore, we have $B(x,y) \geq x/(1-y+y^2/2)$ (a simple verification). Therefore,

$$\begin{aligned} 1 = B(1,0) &\geq B(1-a_1, a_1) \geq B((1-a_1)(1-a_2), a_1+a_2) \geq \dots \\ &\geq B((1-a_1)\dots(1-a_N), a_1+\dots+a_N) \geq \frac{\prod_{k=1}^n (1-a_k)}{1-\sum_{k=1}^n a_k + \frac{1}{2}\left(\sum_{k=1}^n a_k\right)^2}. \end{aligned}$$

This gives the desired claim. □

We will continue the discussion on the case of infinite horizon in the next section.

3. Problems

1. Prove the AM-GM inequality following the approach from Remark 1.3.

2. Prove that for any $n \geq 1$ and any real numbers a_1, a_2, …, a_n we have

$$a_1 a_2 \dots a_n \leq \frac{a_1^2}{2} + \frac{a_2^4}{4} + \frac{a_3^8}{8} + \dots + \frac{a_n^{2^n}}{2^n} + \frac{1}{2^n}.$$

3. Show that for any $n \geq 2$ and any positive numbers a_1, a_2, …, a_n we have

$$\sum_{k=1}^n \frac{1}{1+a_k} \geq \min\left\{1, \frac{n}{1+(a_1 a_2 \dots a_n)^{1/n}}\right\}.$$

4. Write down the abstract Bellman functions for the problems below and identify the corresponding Bellman equations.

a) For any positive numbers a_1, a_2, …, a_N we have

$$(1+a_1)(1+a_2)\dots(1+a_N) \geq (1+(a_1 a_2 \dots a_N)^{1/N})^N.$$

b) For any real numbers a_1, a_2, …, a_N we have

$$\sum_{k=1}^n \left(\frac{a_1+a_2+\dots+a_k}{k}\right) \leq \left(2n - \sum_{k=1}^n \frac{1}{k}\right)^{1/2} \left(\sum_{k=1}^n a_k^2\right)^{1/2}.$$

c) Prove that if $a_1 < a_2 < \dots < a_n$ and $b_1 < b_2 < \dots < b_n$, then

$$(a_1+a_2+\dots+a_n)(b_1+b_2+\dots+b_n) < n(a_1b_1 + a_2b_2 + \dots + a_nb_n).$$

d) Prove that if $x_1, x_2, \dots, x_n > 0$ satisfy $x_1+x_2+\dots+x_n \leq \frac{1}{2}$, then

$$(1-x_1)(1-x_2)\dots(1-x_n) \geq \frac{1}{2}.$$

e) Prove that for any a_1, a_2, …, $a_N > 0$ with

$$\frac{1}{a_1} + \frac{1}{a_2} + \dots + \frac{1}{a_N} = N$$

we have

$$a_1 + \frac{a_2^2}{2} + \frac{a_3^3}{3} + \dots + \frac{a_N^N}{N} \geq 1 + \frac{1}{2} + \dots + \frac{1}{N}.$$

f) For any positive integer N and any nonnegative numbers a_1, a_2, …, a_N, b_1, b_2, …, b_N, we have

$$\sum_{n=1}^N n(b_n-1)(a_1 a_2 \dots a_n)^{1/n} + N(a_1 a_2 \dots a_N)^{1/N} \leq \sum_{n=1}^N a_n b_n^n.$$

g) Let $1 < p < \infty$. Then for any positive numbers $a_1, a_2, \dots$ and $\lambda_1, \lambda_2, \dots$ we have the inequality

$$\sum_{n=1}^{\infty} \lambda_n \left(\frac{\lambda_1 a_1 + \lambda_2 a_2 + \ldots + \lambda_n a_n}{\lambda_1 + \lambda_2 + \ldots + \lambda_n} \right)^p \le \left(\frac{p}{p-1} \right)^p \sum_{n=1}^{\infty} \lambda_n a_n^p.$$

h) For any positive numbers $a_1, a_2, \dots,$

$$\sum_{n=1}^{\infty} \left(\frac{a_1 + a_2 + \ldots + a_n}{n} \right)^p \le \left(\frac{p}{p-1} \right)^p \sum_{n=1}^{\infty} a_n^p.$$

5. Prove that for any positive numbers $a_1, a_2, \dots, a_n$ satisfying $a_1 + a_2 + \ldots + a_n < 1$ we have

$$\frac{a_1 a_2 \ldots a_n \big(1 - (a_1 + a_2 + \ldots + a_n)\big)}{(a_1 + a_2 + \ldots + a_n)(1 - a_1)(1 - a_2) \ldots (1 - a_n)} \le \frac{1}{n^{n+1}}.$$

6. Prove that for any positive numbers $a_1, a_2, \dots, a_n$ satisfying $a_1 a_2 \ldots a_n = 1$ we have

$$\frac{1}{n - 1 + a_1} + \frac{1}{n - 1 + a_2} + \ldots + \frac{1}{n - 1 + a_n} \le 1.$$

7. Prove that for any positive numbers $a_1, a_2, \dots, a_n$ we have

$$\left(1 + \frac{1}{a_1}\right)\left(1 + \frac{1}{a_2}\right) \ldots \left(1 + \frac{1}{a_n}\right) \ge \left(1 + \frac{n}{a_1 + a_2 + \ldots + a_n}\right)^n.$$

8. Prove that for any positive numbers $a_1, a_2, \dots, a_n, b_1, b_2, \dots, b_n$ we have

$$\left(\frac{a_1 + a_2 + \ldots + a_n}{b_1 + b_2 + \ldots + b_n} \right)^2 \le \left(\frac{a_1}{b_1} \right)^2 + \left(\frac{a_2}{b_2} \right)^2 + \ldots + \left(\frac{a_n}{b_n} \right)^2.$$

9. Prove that for any positive numbers $a_1, a_2, \dots, a_n$ with $a_1 + a_2 + \ldots + a_n = x$ we have

$$1 + x \le (1 + a_1)(1 + a_2) \ldots (1 + a_n) \le e^x.$$

10. Prove the following form of planar isoperimetry: for any polygon of perimeter p, its area is not bigger than $\dfrac{p^2}{4\pi}$.

CHAPTER 2

Optimal Control, Stochastic Case

1. Finite horizon

We start with a simple model, which will be modified and complicated later on. Suppose that N is a given time horizon and $\xi_1, \xi_2, \ldots, \xi_N$ is a sequence of independent, identically distributed random variables taking values in some set $(S, \mathcal{S})$. Let $(E, \mathcal{E})$ be a fixed measure space (state space), let $(U, \mathcal{U})$ be the set of controls and let $F : E \times U \times S \to E$ be a given measurable function. We assume that X_0 is a random variable taking values in E and the model is evolving according to the equation

$$X_m = F(X_{m-1}, u_m, \xi_m), \qquad m = 1, 2, \ldots, N.$$

Here $u = (u_m)_{m=1}^N$ is a *strategy* (a sequence of *controls*): we assume that for each m, the function $u_m = u_m(x_0, x_1, x_2 \ldots, x_{m-1})$ is a measurable map from E^m to U. Sometimes, to indicate the dependence of X on the strategy, we will use the notation X^u. If needed, we may also assume that the process X starts at time n from some prescribed position X_n, which may be random or non-random.

As in the deterministic case, we introduce the functional

$$J^N(X_0, u) = \mathbb{E}\left\{ \sum_{m=1}^{N} g_m\big(X^u_{m-1}, u_m(X^u_0, X^u_1, \ldots, X^u_{m-1})\big) + r(X^u_N) \right\},$$

which will be subject to optimization (we will restrict ourselves to maximization; for minimization, the argumentation goes along the same lines). Here g_m and r are measurable functions. Here we see the first important difference in comparison to the deterministic case: the functional J^N involves the expectation. Again, there are two problems for the investigation: (i) to compute the value $\sup_u J^N(X_0, u)$ explicitly, and (ii) to identify the controls $u_1^*, u_2^*, \ldots, u_N^*$ which yield this optimal performance.

In the case of finite horizon, the dynamic programming turns out to be the key approach again. For the further discussion, it is convenient to distinguish the auxiliary operator T^u, which acts on bounded functions $h : E \to \mathbb{R}$ by

$$T^u h(x) = \mathbb{E}h(F(x, u, \xi_0)) = \mathbb{E}\big(h(X_n) \,|\, X_{n-1} = x,\, u_n = u\big).$$

This action makes sense also for unbounded functions, even infinite, as long as the expectation makes sense. Observe that if u is an arbitrary strategy, then

$$\mathbb{E}T^{u_n} h(X^u_{n-1}) = \mathbb{E}h(X^u_n). \tag{2.1}$$

THEOREM 2.1. *Under the above notation, introduce the Bellman sequence B_N, $B_{N-1}, \ldots, B_0$ by $B_N(x) = r(x)$ and the Bellman equation*

$$B_{n-1}(x) = \sup_{u \in U(x)} \big(g_n(x, u) + T^u B_n(x)\big), \qquad n = N, N-1, \ldots, 1,$$

where $U(x)$ is the class of controls permitted when the system is at the state x. Then for any strategy u we have $J^N(X_0, u) \leq \mathbb{E}B_0(X_0)$. Furthermore, if there exist measurable maps $\hat{u}_1$, $\hat{u}_2$, $\ldots$, $\hat{u}_N$ such that for all x and n,

$$B_{n-1}(x) = g_n(x, \hat{u}_n(x)) + T^{\hat{u}_n(x)} B_n(x),$$

then the strategy $u^*_m(X_0, X_1, \ldots, X_{m-1}) = \hat{u}_m(X_{m-1})$, $m = 1, 2, \ldots, N$, is optimal and $J^N(X_0, u^*) = \mathbb{E}B_0(X_0)$.

PROOF. Fix an arbitrary strategy u. We have

$$\mathbb{E}B_{n-1}(X^u_{n-1}) \geq \mathbb{E}\left\{g_n(X^u_{n-1}, u_n) + T^{u_n} B_n(X^u_{n-1})\right\} = \mathbb{E}g_n(X^u_{n-1}, u_n) + \mathbb{E}B_n(X^u_n).$$

Consequently,

$$\begin{aligned}\mathbb{E}B_0(X_0) &\geq \mathbb{E}g_1(X^u_0, u_1) + \mathbb{E}B_1(X^u_1) + \mathbb{E}\big(g_1(X^u_0, u_1) + g_2(X^u_1, u_2)\big) + \mathbb{E}B_2(X^u_2) \\ &\geq \ldots \geq \mathbb{E}\left\{\sum_{m=1}^{N} g_m(X^u_{m-1}, u_m)\right\} + \mathbb{E}B_N(X^u_N) = J^N(X_0, u).\end{aligned}$$

The second part of the theorem follows from the fact that if the controls $\hat{u}_1$, $\hat{u}_2$, $\ldots$, $\hat{u}_N$ exist, then all the intermediate estimates above become equalities. □

As in the deterministic case, some comments are in order.

I. It follows directly from the above proof that the Bellman sequence is

$$B_n(x) = \sup_u \mathbb{E}\Bigg\{\sum_{m=n+1}^{N} g_m\big(X^u_{m-1}, u_m(X^u_n, X^u_{n+1}, \ldots, X^u_{m-1})\big) + r(X^u_N)\Bigg\},$$

where it is assumed that $X_n = x$ and the controls u are restricted to the time set $\{n+1, n+2, \ldots, N\}$.

II. We have the following variant of the optimality principle: suppose that u^*_1, u^*_2, $\ldots$, u^*_N is an optimal strategy for

$$J^N(X_0, u) = \mathbb{E}\left\{\sum_{m=1}^{N} g_m\big(X^u_{m-1}, u_m(X^u_0, X^u_1, \ldots, X^u_{m-1})\big) + r(X^u_N)\right\}.$$

Let us split the time set $\{0, 1, 2, \ldots, N\}$ into two parts: $\{0, 1, 2, \ldots, M\}$ and $\{M+1, M+2, \ldots, N\}$. Then u^*_{M+1}, u^*_{M+2}, $\ldots$, u^*_N is optimal for the truncated functional

$$J^N_M(X_M, u) = \mathbb{E}\left\{\sum_{m=M+1}^{N} g_m\big(X^u_{m-1}, u_m(X^u_0, X^u_1, \ldots, X^u_{m-1})\big) + r(X^u_N)\right\}$$

and the optimal value is equal to $\mathbb{E}B_M(X_M)$. The remaining part u^*_1, u^*_2, $\ldots$, u^*_M of the strategy is optimal for the modified functional

$$J^n(X_0, u) = \mathbb{E}\left\{\sum_{m=1}^{M} g_m(X^u_{m-1}, u_m) + B_M(X_M)\right\}.$$

1.1. An example. Consider the following system: X_0, X_1, ..., X_n is given by $X_0 = 1$ and at m-th step, the control $u_m \in \{0,1\}$ acts as follows. If $u_m = 0$, then $X_m = 2X_{m-1}$; if $u_m = 1$, then X_m is equal to 0 or X_m^2 with probability 1/2. We will compute the value $\sup_u \mathbb{E}X_N$ and identify the optimal strategy.

To express the problem in the above language, let $\xi_1, \xi_2, \ldots, \xi_N$ be a sequence of independent random variables with the distribution $\mathbb{P}(\xi_j = 0) = \mathbb{P}(\xi_j = 1) = 1/2$. We let

$$X_m = F(X_{m-1}, u_m, \xi_m), \qquad m = 1, 2, \ldots, N,$$

where $F(x, 0, s) = 2x$ and $F(x, 1, s) = sx^2$. We see that the optimized functional is of the above form, with $g_m = 0$ and $r(x) = x$. Hence, we may make use of the above machinery; note that X takes values in $E = \{0, 2, 4, 8, \ldots\}$, and in the calculations below we assume that x belongs to this set. Take the Bellman sequence $(B_n)_{n=0}^N$, setting $B_N(x) = x$ and

$$B_{n-1}(x) = \max_{u \in \{0,1\}} T^u B_n(x), \qquad n = N, N-1, \ldots, 1,$$

where

$$T^0 h(x) = h(2x), \qquad \text{and} \qquad T^1 h(x) = \frac{1}{2}h(0) + \frac{1}{2}h(x^2).$$

We start the computations. We have $T^0 B_N(x) = 2x$ and $T^1 B_N(x) = x^2/2$, so

$$B_{N-1}(x) = \max\{2x, x^2/2\} = \begin{cases} 2x & \text{if } x \leq 4, \\ x^2/2 & \text{if } x \geq 4. \end{cases}$$

To find B_{N-2} we proceed analogously, checking that

$$T^0 B_{N-1}(x) = B_{N-1}(2x) = \begin{cases} 4x & \text{if } 2x \leq 4, \\ 2x^2 & \text{if } 2x \geq 4 \end{cases}$$

and

$$T^1 B_{N-1}(x) = \frac{1}{2} B_{N-1}(0) + \frac{1}{2} B_{N-1}(x^2) = \begin{cases} x^2 & \text{if } x^2 \leq 4, \\ \frac{x^4}{4} & \text{if } x^2 \geq 4. \end{cases}$$

Hence

$$B_{N-2}(x) = \max\{T^0 B_{N-1}(x), T^1 B_{N-1}(x)\} = \begin{cases} 4x & \text{if } x \leq 2, \\ \frac{x^4}{4} & \text{if } x > 2. \end{cases}$$

Arguing similarly, we compute that

$$B_{n-1}(x) = \max\{T^0 B_n(x), T^1 B_n(x)\} = \begin{cases} 2^{N-n+1}x & \text{if } x \leq 2, \\ \frac{x^{2^{N-n+1}}}{2^{N-n+1}} & \text{if } x > 2, \end{cases}$$

the maximal value coming from $T^0 B_n(x)$ for $x \leq 2$ and from $T^1 B_n(x)$ for $x > 2$. Consequently, we see that $\sup_u \mathbb{E}X_N$ is equal to $\mathbb{E}B_0(X_0) = 2^N$ and the optimal control is given as follows: as long as the process does not exceed 2, use the control 0; otherwise, apply the control 1.

1.2. Another example. Let $\xi_1, \xi_2, \ldots, \xi_N$ be a sequence of independent random variables with the distribution given by $\mathbb{P}(\xi_j = 0) = \mathbb{P}(\xi_j = 8) = 1/2$. Suppose that $(X_n)_{n=0}^N$ is a sequence satisfying $X_0 = 1$ and

$$X_m = u_m \xi_m \cdot X_{m-1}, \qquad m = 1, 2, \ldots, N,$$

where $u_m \in [0, 1]$ is a control at m-th step. Assume that we are interested in the supremum

$$\sup_u J(X_0, u) = \mathbb{E}\Big\{ \sum_{m=1}^N 2^{-m+1} \sqrt{(1-u_m)X_{m-1}} + 2^{-N} \sqrt{2X_N} \Big\}.$$

We start with the definition of the operator T^u. For any function $h : [0, \infty) \to [0, \infty)$, we have

$$T^u h(x) = \mathbb{E} h(F(x, u, \xi_1)) = \mathbb{E} h(u \xi_1 x) = \frac{1}{2} h(0) + \frac{1}{2} h(8ux).$$

We are ready to study the associated Bellman sequence. We let $B_N(x) = 2^{-N}\sqrt{2x}$ and, for any $1 \le n \le N$,

$$B_{n-1}(x) = \sup_{u \in [0,1]} \big(g_n(x, u) + T^u B_n(x) \big) = \sup_{u \in [0,1]} \big(2^{-n+1} \sqrt{(1-u)x} + T^u B_n(x) \big).$$

Hence,

$$\begin{aligned} B_{N-1}(x) &= \sup_{u \in [0,1]} \left(2^{-N+1} \sqrt{(1-u)x} + \frac{1}{2} \cdot 2^{-N} \sqrt{16ux} \right) \\ &= 2^{-N+1} \sqrt{x} \sup_{u \in [0,1]} \big(\sqrt{1-u} + \sqrt{u} \big) = 2^{-N+1} \sqrt{2x}, \end{aligned}$$

the supremum attained for $u = 1/2$. Generally, we prove by a simple induction, using the same calculations as above, that $B_n(x) = 2^{-n}\sqrt{2x}$. Therefore, by the above theorem, $J^N(X_0, u) \le \mathbb{E} B_0(X_0) = \sqrt{2}$. Equality holds if and only if $u_1 = u_2 = \ldots = u_N = 1/2$.

1.3. Yet another example. Let $\eta_1, \eta_2, \ldots, \eta_N$ be a sequence of independent random variables such that η_n has the uniform distribution on $[0, u_n]$, $n = 1, 2, \ldots, N$. The purpose is to find optimal controls u_j which maximize the expectation

$$\mathbb{E}\Big\{ \max\{\eta_1, \eta_2, \ldots, \eta_N\} - \frac{1}{2}(u_1 + u_2 + \ldots + u_N) \Big\}.$$

To analyze this problem, we modify the variables so that they have the same distribution: let $\xi_j = \eta_j / u_j \sim U(0, 1)$. Next, let $X_0 = 0$ and set

$$X_m = \max\{X_{m-1}, u_m \xi_m\}, \qquad m = 1, 2, \ldots, N.$$

Then the problem is to analyze $\sup_u \mathbb{E}\big\{ -\frac{1}{2}(u_1 + u_2 + \ldots + u_N) + X_N \big\}$, which falls into the scope of the above approach. The associated operator T^u is given by

$$T^u h(x) = \mathbb{E} h(\max\{x, u\xi_m\}) = \frac{1}{u} \int_0^u h(\max\{x, s\}) \mathrm{d}s.$$

We are ready to introduce the Bellman sequence. Put $B_N(x) = x$ and

$$B_{n-1}(x) = \sup_u \Big\{ -\frac{u}{2} + T^u B_{n-1}(x) \Big\}.$$

We check that

$$T^u B_N(x) = \mathbb{E}\max\{x, u\xi_1\} = \frac{1}{u}\int_0^u \max\{x,s\}\mathrm{d}s = \begin{cases} x & \text{if } u \le x, \\ \frac{u^2+x^2}{2u} & \text{if } u > x \end{cases}$$

and hence the function

$$-\frac{u}{2} + T^u B_{n-1}(x) = \begin{cases} x - \frac{u}{2} & \text{if } u \le x, \\ \frac{x^2}{2u} & \text{if } u > x \end{cases}$$

is decreasing on $(0,\infty)$. Consequently, the supremum defining $B_{N-1}(x)$ is attained in the limit $u \to 0$ and $B_{N-1}(x) = x$. Repeating the calculations, we obtain that $B_n(x) = x$ for all n, and hence in particular

$$\sup_u \mathbb{E}\bigg\{ -\frac{1}{2}(u_1 + u_2 + \ldots + u_N) + X_N \bigg\} \le \mathbb{E}B_0(X_0) = 0.$$

Clearly, we have equality here: it suffices to take small controls u_1, u_2, …, u_N to obtain quantities as close to zero as we wish.

Remark 2.2. Theorem 2.1 can be generalized in many directions. Let us discuss briefly two possibilities here.

(i) It is easy to extend the approach to the case in which the variables ξ_j do not have the same distribution: this forces the operator T^u to depend also on the number of the step.

(ii) Another important example concerns the case in which $(\xi_n)_{n\ge 0}$ is a time-homogeneous Markov process. Then all one needs is to consider the larger governing process $(X_n, \xi_n)_{n\ge0}$: all the remaining arguments are the same (the role of ξ_0 is to provide the initial distribution). That is, one considers the larger state space $\hat{E} = E \times S$ and introduces the operator T^u acting via

$$T^u h(x,s) = \mathbb{E}(h(X_n,\xi_n) \,|\, X_{n-1} = x,\, u_n = u,\, \xi_n = s).$$

Then the theorem above remains valid, one just needs to replace x with (x,s) in all the relevant places.

1.4. An example. Consider the Markov chain ξ on $\{0,1\}$, starting from 0, with the transition matrix given by

$$P = \begin{bmatrix} 3/4 & 1/4 \\ 1/4 & 3/4 \end{bmatrix}.$$

Let $(X_n)_{n=0}^N$ be given by $X_0 = 1$ and evolving according to

$$X_m = u_m X_{m-1}\xi_m + (1-u_m)X_{m-1}(1-\xi_m),$$

where the controls u_m are arbitrary numbers from $[0,1]$. We will study the quantity

$$\sup_u \mathbb{E}\big(X_N + \xi_N\big).$$

Our starting observation is that we need to enrich the state space: the controlled process must be taken to be $((X_n,\xi_n))_{n=0}^N \in \mathbb{R}\times\{0,1\}$. The Bellman sequence is

$$B_n(x,s) = \sup_u \Big\{\mathbb{E}(X_N+\xi_N) \,:\, (X_n,\xi_n) = (x,s)\Big\}, \qquad n = 0, 1, 2, \ldots, N.$$

We clearly have $B_N(x,s) = x + s$ and the following recurrence holds:

$$B_{n-1}(x,s) = \sup_u T^u B_n(x,s), \qquad n = 1, 2, \ldots, N.$$

Note that

$$\begin{aligned} T^u h(x,s) &= \mathbb{E}\big[h(ux\xi_1 + (1-u)x(1-\xi_1), \xi_1) \,\big|\, \xi_0 = s\big] \\ &= \begin{cases} \frac{3}{4}h((1-u)x,0) + \frac{1}{4}h(ux,1) & \text{if } s=0, \\ \frac{1}{4}h((1-u)x,0) + \frac{3}{4}h(ux,1) & \text{if } s=1. \end{cases} \end{aligned}$$

Thus we compute that

$$T^u B_N(x,0) = \frac{3}{4}(1-u)x + \frac{1}{4}(ux+1), \qquad T^u B_N(x,1) = \frac{1}{4}(1-u)x + \frac{3}{4}(ux+1),$$

and hence

$$B_{N-1}(x,0) = \frac{3}{4}x + \frac{1}{4}, \qquad B_{N-1}(x,1) = \frac{3}{4}x + \frac{3}{4},$$

and the optimal controls are $u_N = 0$ and $u_N = 1$, respectively. A similar calculation shows that

$$B_{N-2}(x,0) = \left(\frac{3}{4}\right)^2 x + \frac{6}{16}, \qquad B_{N-2}(x,1) = \left(\frac{3}{4}\right)^2 x + \frac{10}{16},$$

and the optimal controls are $u_{N-1} = 0$ and $u_{N-1} = 1$, respectively. The above formulas suggest that

$$B_n(x,0) = \left(\frac{3}{4}\right)^{N-n} x + a_n, \qquad B_n(x,1) = \left(\frac{3}{4}\right)^{N-n} x + b_n,$$

for some sequences $(a_n)_{n=0}^N$ and $(b_n)_{n=0}^N$ to be found. Assuming this form for n and arguing as above, we compute that

$$B_{n-1}(x,0) = \left(\frac{3}{4}\right)^{N-n+1} x + \frac{3}{4}a_n + \frac{1}{4}b_n$$

and

$$B_{n-1}(x,1) = \left(\frac{3}{4}\right)^{N-n+1} x + \frac{1}{4}a_n + \frac{3}{4}b_n$$

(with optimal controls equal to 0 and 1, respectively), so that $a_{n-1} = \frac{3}{4}a_n + \frac{1}{4}b_n$ and $b_{n-1} = \frac{1}{4}a_n + \frac{3}{4}b_n$. We have the initial conditions $a_N = 0$ and $b_N = 1$, so solving the above recurrence gives

$$a_n = \frac{1}{2} - \left(\frac{1}{2}\right)^{N-(n-1)}, \qquad b_n = \left(\frac{1}{2}\right)^{N-(n-1)} + \frac{1}{2}.$$

Thus, the supremum we are interested in, is equal to

$$B_0(1,0) = \left(\frac{3}{4}\right)^{N+1} + \frac{1}{2} - \left(\frac{1}{2}\right)^{N+1}.$$

The optimal strategy is as follows: for any n, if we have $(X_n, \xi_n) = (x,s)$, then the control u_{n+1} is equal to s.

REMARK 2.3. As the above examples show, the only relevant information for the solution is carried in the family $(T^u)_{u\in U}$ of operators.

2. Infinite horizon

Now we will turn our attention to the case in which the controlled process consists of infinite number of variables. The first natural approach to this type of problems is to consider the truncated, finite-horizon version, and then let with the size of the horizon to infinity. For the sake of convenience, let us start with applying Theorem 2.1 to the special functional

$$J^N(X_0, u) = \mathbb{E}\left\{\sum_{m=1}^{N} \gamma^{m-1} q(X_{m-1}, u_m) + \gamma^N r(X_N)\right\}, \tag{2.2}$$

where $\gamma \geq 0$ and $q : E \times U \to \mathbb{R}_+$, $r : E \to \mathbb{R}_+$ are measurable functions. Then the assertion can be expressed in a simpler manner. Define the auxiliary operator $\mathcal{A}$ acting via

$$\mathcal{A}h(x) = \sup_{u \in U(x)} \big(q(x,u) + \gamma T^u h(x)\big) \tag{2.3}$$

when J^N is subject to maximization, and

$$\mathcal{A}h(x) = \inf_{u \in U(x)} \big(q(x,u) + \gamma T^u h(x)\big) \tag{2.4}$$

otherwise. In what follows, $\mathcal{A}^j$ will stand for the j-th iteration of $\mathcal{A}$, and we set $\mathcal{A}^0$ to be the identity operator.

THEOREM 2.4. *The Bellman sequence $(B_n)_{n=0}^N$ is given by*

$$B_n(x) = \gamma^n \mathcal{A}^{N-n} r(x), \qquad n = 0, 1, 2, \ldots, N.$$

Furthermore, if $(\hat{u}_n)_{n=1}^N$ is a sequence of measurable functions from E to U such that for any n and any $x \in E$ we have $\hat{u}_n(x) \in U(x)$ and

$$\mathcal{A}^{N-n} r(x) = q(x, \hat{u}_n(x)) + \gamma T^{\hat{u}_n(x)} \mathcal{A}^{N-(n+1)} r(x),$$

then the strategy $(\hat{u}_1(x_0), \hat{u}_2(x_1), \ldots, \hat{u}_N(x_{N-1}))$ is optimal.

We are ready to discuss the optimal control for the infinite horizon. Let us describe the setup. We keep the notation from the beginning of the case of finite horizon; the evolution of the sequence X is given by

$$X_m = F(X_{m-1}, u_m, \xi_m), \qquad m = 1, 2, \ldots.$$

We are interested in the optimization of the functional

$$J(X_0, u) = J^\infty(X_0, u) = \mathbb{E}\sum_{m=1}^{\infty} \gamma^{m-1} q(X^u_{m-1}, u_m(X^u_0, X^u_1, X^u_2, \ldots, X^u_{m-1})),$$

where $q : E \times U \to \mathbb{R}_+$ is measurable and u_n's stand for the strategy. Let us emphasize here that q is a nonnegative function, which does not depend on m. A little thought reveals that this time-homogeneity condition is plausible: without it, one should not expect any regularity of the problem. As we shall see, this will imply the following intuitive fact: the terms of the strategy $u = (u(x_0), u(x_1), \ldots)$ will not depend on time. In contrast to the case of finite horizon, the arguments for minimizing and maximizing J^∞ will be a little different: because of the nonnegativity of q (which is needed in our considerations), the situation is not symmetric.

Recall the operators $\mathcal{A}$ given by (2.3) and (2.4), recall also that

$$T^u h(x) = \mathbb{E}h(F(x, u, \xi_0))$$

makes sense also for h taking infinite values, as long as the expectation is well-defined. The initial observation is that the operators $\mathcal{A}$ are monotone: for any $h_1, h_2 : E \to \mathbb{R}_+ \cup \{\infty\}$ with $h_1 \le h_2$ we have $\mathcal{A}h_1 \le \mathcal{A}h_2$. Let $\mathsf{0}$ denote the function on E identically equal to zero. Since q is nonnegative, the functional sequence $(\mathcal{A}^n(\mathsf{0}))_{n\ge 0}$ is monotone and hence the pointwise limit

$$b_\infty(x) = \lim_{n\to\infty} \mathcal{A}^n(\mathsf{0})(x), \qquad x \in E,$$

exists. Let us emphasize here that b_∞ is allowed to take infinite values. The use of function $\mathsf{0}$ is clear: the truncation of J^∞ is

$$J^N(X_0, u) = \mathbb{E}\left\{\sum_{m=1}^{N} \gamma^{m-1} q(X_{m-1}, u_m)\right\},$$

which is (2.2) with $r = \mathsf{0}$. The basic results are formulated in two theorems below.

THEOREM 2.5. *Suppose that J^∞ is the cost functional (i.e., it is subject to minimization). Then for any $x \in E$ and any strategy u we have*

$$b_\infty \le \mathcal{A}b_\infty(x), \qquad b_\infty(x) \le J^\infty(x, u).$$

Furthermore, if for any $x \in E$ we have

$$b_\infty(x) = \mathcal{A}b_\infty(x) = q(x, u_\infty(x)) + \gamma T^{u_\infty(x)} b_\infty(x),$$

then the strategy $u^* = (u_\infty(x_0), u_\infty(x_1), \dots)$ is optimal.

PROOF. We have $\mathcal{A}^n(\mathsf{0}) \le b_\infty$ and hence, by the monotonicity of $\mathcal{A}$, we get $\mathcal{A}^{n+1}(\mathsf{0}) \le \mathcal{A}b_\infty$. This yields $b_\infty \le \mathcal{A}b_\infty$, by passing to the limit. Next, using Theorem 2.4, for any strategy u and any starting point x we have

$$\begin{aligned}
&\mathbb{E}\sum_{m=1}^{\infty} \gamma^{m-1} q(X^u_{m-1}, u_m(X^u_0, \dots, X^u_{m-1}))\\
&\qquad \ge \mathbb{E}\sum_{m=1}^{N} \gamma^{m-1} q(X^u_{m-1}, u_m(X^u_0, \dots, X^u_{m-1})) \ge \mathcal{A}^N(\mathsf{0})(x).
\end{aligned}$$

This gives $J^\infty(x, u) \ge \mathcal{A}^N(\mathsf{0})(x)$ and hence, letting $N \to \infty$, we get $b_\infty(x) \le J^\infty(u, x)$. To show the second half of the theorem, note that

$$\begin{aligned}
J^N(x, u^*) &= \mathbb{E}\sum_{m=1}^{N} \gamma^{m-1} q(X^{u^*}_{m-1}, u_\infty(X^{u^*}_{m-1}))\\
&\le \mathbb{E}\left\{\sum_{m=1}^{N} \gamma^{m-1} q(X^{u^*}_{m-1}, u_\infty(X^{u^*}_{m-1})) + \gamma^N b_\infty(X^{u^*}_N)\right\}\\
&= \mathcal{A}^N b_\infty(x) = b_\infty(x),
\end{aligned}$$

since $\mathcal{A}b_\infty = b_\infty$. Letting $N \to \infty$ gives $J^\infty(x, u^*) \le b_\infty$ and hence equality holds (we proved the reverse estimate a moment before). This proves the optimality of the strategy. □

THEOREM 2.6. *Suppose that we are interested in maximizing J^∞. Then for any $x \in E$ and any strategy u we have*

$$b_\infty(x) = \mathcal{A}b_\infty(x), \qquad b_\infty(x) \ge J^\infty(x, u).$$

Furthermore, if for any $x \in E$ we have

$$b_\infty = \mathcal{A}b_\infty(x) = q(x, u_\infty(x)) + \gamma T^{u_\infty(x)} b_\infty(x)$$

for some strategy $u^* = (u_\infty(x_0), u_\infty(x_1), \ldots)$, then this strategy is optimal, provided

$$\lim_{N\to\infty} \mathbb{E}\gamma^N b_\infty(X_N^{u^*}) = 0. \tag{2.5}$$

PROOF. We first show the equality $b_\infty = \mathcal{A}b_\infty$. Note that $b_\infty \le \mathcal{A}b_\infty$, by the monotonicity of $\mathcal{A}$ (see the proof of the previous theorem), and hence it is enough to establish the reverse bound. To this end, note that

$$\begin{aligned}\mathcal{A}^{n+1}(\mathsf{0}) &= \sup_{u\in U(x)} \big(q(x,u) + \gamma T^u \mathcal{A}^n(\mathsf{0})(x)\big)\\ &\ge q(x,u) + \gamma T^u \mathcal{A}^n(\mathsf{0})(x)\end{aligned}$$

for any $x \in E$ and any $u \in U(x)$. Hence, passing to the limit (and using Lebesgue's monotone convergence theorem) we get $b_\infty(x) \ge q(x,u) + \gamma T^u b_\infty(x)$. Take the supremum over u to obtain $b_\infty \ge \mathcal{A}b_\infty$.

Next, we will prove that for any strategy u we have $b_\infty \ge J^\infty(x,u)$. This follows from the finite horizon: for any N and any starting point $x \in E$ we have

$$\mathbb{E}\sum_{m=1}^{N} \gamma^{m-1} q(X_{m-1}^u, u_m(X_0^u, \ldots, X_{m-1}^u)) \le \mathcal{A}^N(\mathsf{0})(x) \le b_\infty(x),$$

and letting $N \to \infty$ gives the claim. Finally, to show the optimality of the strategy u in the statement, it is enough to prove that $J^\infty(x,u) = b_\infty(x)$ for all $x \in E$. Consider the modified functional

$$\tilde{J}^N(x,u) = \mathbb{E}\left\{\sum_{m=1}^{N} \gamma^{m-1} q(X_{m-1}^u, u_m) + \gamma^N b_\infty(X_N^u)\right\}.$$

By (2.5), we have $J^\infty(x,u) = \lim_{N\to\infty} \tilde{J}^N(x,u)$. However, Theorem 2.4 combined with $\mathcal{A}b_\infty = b_\infty$ yields $\tilde{J}^N(x,u) = b_\infty(x)$. Plugging this into the preceding equation gives the assertion. □

REMARK 2.7. The following example shows that the equality $\mathcal{A}b_\infty = b_\infty$ need not hold in general; the example below concerns the case of cost functional. Consider a Markov process on the state space $E = \{(j,k) : k = 1, 2, \ldots, j = 1, 2, \ldots, k\}$ and let $U = \{2, 3, \ldots\}$. The transition probabilities are given by

$$p^u((1,1),(1,u)) = \frac{1}{u}, \qquad p^u((1,1),(1,1)) = 1 - \frac{1}{u},$$

and $p^u((k,k),(k,k)) = 1$, $p^u((j,k),(j+1,k)) = 1$ for $k \ge 2$ and $j \le k-1$. For $x = (j,k) \in E$ and $u \in U$, let

$$q(x,u) = q((j,k),u) = \begin{cases} 1 & \text{if } j < k,\\ 0 & \text{if } j = k.\end{cases}$$

Furthermore, we assume that the discount factor γ is equal to 1. We compute directly that

$$\mathcal{A}^n(\mathsf{0})(1,k) = \min\{n, k-1\} \qquad \text{for } k \ge 2,\ n = 1, 2, \ldots$$

and

$$\mathcal{A}^n(\mathsf{0})(1,1) = 0, \qquad n = 1, 2, \ldots.$$

This implies $b_\infty(1,k) = k-1$ for $k \geq 2$ and $b_\infty(1,1) = 0$. On the other hand,

$$\mathcal{A}b_\infty(1,1) = \inf_{u\geq 2} \frac{u-1}{u} = \frac{1}{2} \neq b_\infty(1,1).$$

To solve the problem, note that for any strategy u we have

$$\begin{aligned} J^\infty((1,1),u) &= \frac{1}{u_0}(u_0-1) + \left(1-\frac{1}{u_0}\right)\cdot\frac{1}{u_1}(u_1-1) + \dots \\ &= \left(1-\frac{1}{u_0}\right) + \left(1-\frac{1}{u_0}\right)\left(1-\frac{1}{u_1}\right) + \dots, \end{aligned}$$

which is minimal (and equal to 1) for $u_0 = u_1 = u_2 = \ldots = 2$.

The second method of handling the infinite horizon is to search directly for the optimal value of the functional

$$J^\infty(X_0,u) = \mathbb{E}\left\{\sum_{m=1}^\infty \gamma^{m-1} q(X^u_{m-1}, u_m(X^u_{m-1}))\right\}.$$

Suppose that we are interested in the maximal value of J^∞. Then we introduce the Bellman function $b(x) = \sup_u J^\infty(x,u)$ and apply the standard recursive argument to obtain the equation

$$b(x) = \mathcal{A}b(x) = \sup_u \big\{q(x,u) + \gamma T^u b(x)\big\} \tag{2.6}$$

(note that this equation appears in both Theorems 2.5 and 2.6 above). Now a natural idea is to search for the solution of this equation; once we find it, we may hope that it is indeed the desired extremal value of $J^\infty(x,u)$, and the optimal controls are those for which equation in (2.6) holds. However, there are some drawbacks: as we have already seen in the deterministic case, the solution to (2.6) may not be unique. In practice, having constructed some solution, we need to prove rigorously that this *candidate* does coincide with the Bellman function in question. Here is the idea. Suppose that some function $\tilde{b}$ satisfies the equation

$$\tilde{b}(x) = \mathcal{A}\tilde{b}(x) = \sup_u \big\{q(x,u) + \gamma T^u \tilde{b}(x)\big\}.$$

Then for any strategy $u = (u_m)_{m\geq 1}$, with $u_m = u_m(X_0, X_1, \ldots, X_{m-1})$ (note that now we allow the dependence of u_m on the past values of the process!) we have

$$\tilde{b}(X^u_{m-1}) \geq q(X^u_{m-1}, u_m)) + \gamma\mathbb{E}(\tilde{b}(X^u_m)|X_{m-1}),$$

which gives

$$\mathbb{E}\Big\{\gamma^{m-1}\tilde{b}(X^u_{m-1}) - \gamma^m \tilde{b}(X^u_m)\Big\} \geq \mathbb{E}\big\{\gamma^{m-1} q(X_{m-1}, u_m)\big\}.$$

Consequently, for any N we obtain

$$\mathbb{E}\tilde{b}(X_0) \geq \gamma^N \mathbb{E}\tilde{b}(X^u_{N-1}) + J^N(X_0,u).$$

In addition, the above inequality may be as close to equality as we wish, by a proper choice of controls. If for any such optimal, or almost-optimal strategy u the first expectation on the right converges to zero as $N \to \infty$ (compare this to (2.5)), we obtain

$$\mathbb{E}\tilde{b}(X_0) \geq J^\infty(X_0,u),$$

and equality is attained asymptotically.

Let us consider concrete examples.

2.1. Example. We will solve the optimal control problem

$$\sup_u \mathbb{E} \sum_{m=1}^{\infty} \sqrt{u_m},$$

where the system is driven by the process $(X_n)_{n\geq 0}$ satisfying $X_0 = 1$ and

$$X_m = (X_{m-1} - u_m)\xi_m, \qquad m = 1,\, 2,\, \ldots.$$

Here $u_m \in [0, X_{m-1}]$ and $\xi_1,\, \xi_2,\, \ldots$ is a sequence of independent, identically distributed random variables with $\mathbb{P}(\xi_m = 0) = \mathbb{P}(\xi_m = 1) = 1/2$.

We will present two possible approaches to this problem.

Method I. Approximation. Consider the associated operator

$$\mathcal{A}h(x) = \sup_{u\in[0,x]} \{\sqrt{u} + \mathbb{E}h((x-u)\xi)\} = \sup_{u\in[0,x]} \left\{\sqrt{u} + \frac{1}{2}h(0) + \frac{1}{2}h(u-x)\right\}.$$

We compute that

$$\mathcal{A}(0)(x) = \sup_{u\in[0,x]} \sqrt{u} = \sqrt{x}$$

and

$$\mathcal{A}^2(0)(x) = \sup_{u\in[0,x]} \left\{\sqrt{u} + \frac{1}{2}\sqrt{x-u}\right\} = \frac{\sqrt{5}}{2}\sqrt{x},$$

with equality attained for $u = 4x/5$. This suggests that $\mathcal{A}^n(0)(x) = a_n\sqrt{x}$ for some deterministic sequence $(a_n)_{n\geq 0}$, with $a_0 = 0$, $a_1 = 1$. Repeating the above argumentation, we check that this is indeed the case and $(a_n)_{n\geq 0}$ satisfies the recurrence

$$a_{n+1} = \sqrt{1 + \frac{a_n^2}{4}}, \qquad n = 0,\, 1,\, 2,\, \ldots.$$

A straightforward analysis shows that this sequence increases to $2/\sqrt{3}$ in the limit; hence, we have $b_\infty(x) = \frac{2}{\sqrt{3}}\sqrt{x}$. It follows from Theorem 2.6 that

$$b_\infty(x) = \mathcal{A}b_\infty(x) = \sup_{u\in[0,x]} \left\{\sqrt{u} + \frac{1}{2}b_\infty(x-u)\right\},$$

and by the above calculation, equality is attained for the control $u = 3x/4$. For this control, we easily check that the condition (2.5) holds (we have $X_N \downarrow 0$ almost surely) and hence $\mathbb{E}b_\infty(1) = 2/\sqrt{3}$ is the desired optimal value of the functional.

Method II. A direct search for the Bellman function. The starting point is the abstract formula

$$b(x) = \sup_u \mathbb{E}\left[\sum_{m=1}^{\infty} \sqrt{u_m} \,:\, X_0 = x\right]$$

and the associated equation

$$b(x) = \mathcal{A}b(x) = \sup_{u\in[0,x]} \left\{\sqrt{u} + \frac{1}{2}b(0) + \frac{1}{2}b(x-u)\right\}. \tag{2.7}$$

By the very definition, we have $b(0) = 0$, so the equation reduces to

$$b(x) = \sup_{u\in[0,x]} \left\{\sqrt{u} + \frac{1}{2}b(x-u)\right\}. \tag{2.8}$$

How to find a solution to this equation? They key is to go back to the abstract definition of b. In principle, there are two possible methods. The first rests on exploitation of some inner homogeneity hidden in b. In our case, observe that there is a one-to-one correspondence between sequences X starting from x and controlled by u, and sequences $\tilde{X}$ starting from λx and controlled by λu: all that is needed, is to divide/multiply the sequence and controls by λ. Consequently, by the very definition of b, we obtain

$$b(\lambda x) = \sqrt{\lambda} b(x),$$

so that $b(x) = \alpha\sqrt{x}$ for some parameter $\alpha \geq 0$ to be found. Plugging this into the Bellman equation (2.7), we see that $\alpha = 2/\sqrt{3}$.

It remains to verify rigorously that b is the desired optimal performance of the functional. We have

$$b(X_{m-1}) \geq \sqrt{u_m} + \mathbb{E}(b(X_m)|X_{m-1}),$$

with equality for $u_m = 3X_{m-1}/4$. Hence $\mathbb{E}b(X_{m-1}) - \mathbb{E}b(X_m) \geq \sqrt{u_m}$, so

$$b(1) = \mathbb{E}b(X_0) \geq \mathbb{E}b(X_N) + \sum_{m=1}^{N} \sqrt{u_m}.$$

Letting $N \to \infty$ we get $b(1) \geq J^\infty(X_0, u)$. It follows from the above considerations that equality holds for the strategy $(3X_0/4, 3X_1/4, 3X_2/4, \ldots)$.

The second method of the search for b is to *guess directly the optimal control, or at least the shape of the optimal control*. In our case, is seems plausible to assume that the optimal controls u_m are given by βX_{m-1} for some (unknown) constant $\beta \in (0,1)$: it is best to obtain X_m by "cutting off" the same proportion of X_{m-1}. For such strategy, we have $X_m = (1-\beta)X_{m-1}\xi_m$ and $X_m = x(1-\beta)^m\xi_1\xi_2\ldots\xi_m$ for each m. Then

$$\mathbb{E}\sqrt{u_m} = (\beta x)^{1/2}(1-\beta)^{(m-1)/2}(\mathbb{E}\sqrt{\xi_1})^{m-1} = (\beta x)^{1/2}\left(\frac{1-\beta}{4}\right)^{(m-1)/2}$$

and the value of the Bellman function at x is

$$\sum_{m=1}^{\infty} (\beta x)^{1/2}\left(\frac{1-\beta}{4}\right)^{(m-1)/2} = \frac{\sqrt{\beta x}}{1-\sqrt{(1-\beta)/4}}.$$

The largest value of this expression is attained for $\beta = 3/4$. Pugging this choice above, we get the conjecture

$$b(x) = \frac{2}{\sqrt{3}}\sqrt{x}.$$

Having obtained the candidate, we need to check that it works: this is done by showing (2.8) and repeating the above argumentation.

2.2. Another example. The following deterministic problem may serve as a warning. We will study the quantity

$$\sup_u \sum_{m=1}^{\infty} 2^{-m} \cdot \frac{X_{m-1}+u_m}{2},$$

where $X_0 = 1$ and $X_m = X_{m-1} - u_m$, $m = 1, 2, \ldots$. Here the controls u_m are arbitrary real numbers. Obviously, the above supremum is equal to infinity: we may take $u_m = -X_{m-1} + 2^m$ for all m, which makes the sum divergent.

On the other hand, the Bellman equation reads

$$b(x) = \sup_{u \in \mathbb{R}} \left\{ \frac{x+u}{2} + \frac{1}{2} b(x-u) \right\}.$$

Note that $b(x) = x$ solves this equation; arguing as above, we obtain

$$2^{1-m} b(X_{m-1}) - 2^{-m} b(X_m) \geq 2^{1-m} \cdot \frac{X_{m-1} + u_m}{2},$$

which leads to

$$1 = b(1) = b(X_0) \geq 2^{-N} b(X_N) + 2J^N(X_0, u) = 2^{-N} X_N + 2J^N(X_0, u).$$

The point is that there is no argument which would allow to discard the term $2^{-N} X_N$ and let N go to infinity; there is no control over the sign of X_n.

For didactic reasons, let us modify the above problem, allowing u_m to belong to $[0, X_{m-1}]$; this will make the sequence (X_n) decreasing and nonnegative. Then the limiting argument clearly works and gives $J^\infty(X_0, u) \leq 1/2$ for any u. Actually, equality holds for any strategy: indeed,

$$\begin{aligned}\sum_{m=1}^{\infty} 2^{-m} \cdot \frac{X_{m-1} + u_m}{2} &= \sum_{m=1}^{\infty} 2^{-m} \cdot \frac{2X_{m-1} - X_m}{2} \\ &= \frac{1}{2} \sum_{m=1}^{\infty} \left(2^{1-m} X_{m-1} - 2^{-m} X_m \right) = \frac{1}{2}.\end{aligned}$$

2.3. An inequality for Rademacher variables. Suppose that ε_1, ε_2, ... is a sequence of independent Rademachers. We will show that for any random variables a_1, a_2, ..., with a_m depending on ε_1, ε_2, ..., ε_{m-1}, we have

$$\mathbb{E} \left| \sum_{n=1}^{N} a_n \varepsilon_n \right| \leq \mathbb{E} \sum_{n=1}^{N} a_n^2 + \frac{1}{4}, \qquad N = 0, 1, 2, \ldots.$$

To this end, we put the problem into the framework developed above. Let $X_m = \sum_{n=1}^{m} a_n \varepsilon_n$; thus, $(X_n)_{n \geq 0}$ is governed by the conditions $X_0 = 0$ and $X_m = X_{m-1} + a_m \varepsilon_m$, and $(a_m)_{m \geq 1}$ is a strategy to be optimized. We rewrite the inequality in the form

$$\inf_a \mathbb{E} \left\{ \sum_{n=1}^{N} a_n^2 - |X_N| \right\} \geq -\frac{1}{4}.$$

Because of the appearance of the term $|X_N|$, we should try to solve this problem using the finite horizon approach. So, fix N and define the associated Bellman sequence

$$B_m(x) = \inf_a \mathbb{E} \left[\sum_{n=m+1}^{N} a_n^2 - |X_N| \,\middle|\, X_m = x \right], \qquad m = 0, 1, 2, \ldots.$$

Then $B_N(x) = -|x|$ and we have the Bellman equation

$$\begin{aligned} B_{n-1}(x) &= \inf_a \left\{ a^2 + \mathbb{E}(B_n(X_n) | X_{n-1} = x) \right\} \\ &= \inf_a \left\{ a^2 + \frac{1}{2} B_n(x-a) + \frac{1}{2} B_n(x+a) \right\}. \end{aligned}$$

The drawback is that the computations become involved: one checks that

$$B_{N-1}(x) = \begin{cases} -\frac{1}{4} & \text{if } |x| \le 1/4, \\ -|x| & \text{if } |x| \ge 1/4, \end{cases}$$

in general, B_n is a piecewise linear function, with quite a complicated formula.

It turns out that the infinite horizon approach works efficiently here. We start with the Bellman equation

$$b(x) = \inf_a \left\{ a^2 + \frac{1}{2} b(x-a) + \frac{1}{2} b(x+a) \right\} \tag{2.9}$$

and search for b (or rather, for a candidate for this function). This search might be informal, as the obtained candidate will be rigorously analyzed later. As in the deterministic setting, we write down the infinitesimal version of this condition. Namely, for any a we must have

$$b(x) \le a^2 + \frac{1}{2} b(x-a) + \frac{1}{2} b(x+a),$$

or equivalently,

$$0 \le 1 + \frac{b(x-a) + b(x+a) - 2b(x)}{2a^2}.$$

Letting $a \to 0$ (and assuming that b is of class C^2), we see that b must satisfy $b''(x) + 2 \ge 0$. Assuming equality and noting the obvious symmetry of the function b, we see that $b(x) = -x^2 + c$ for some constant c. This is our desired candidate.

We start the verification. Regardless of the value of c, the function b evidently satisfies (2.9): the expression in the parentheses is constant as a function of a. Therefore, for any n we have

$$b(X_{n-1}) \le a_n^2 + \mathbb{E}(b(X_n)|X_{n-1}),$$

so $\mathbb{E}b(X_{n-1}) - \mathbb{E}b(X_n) \le \mathbb{E}a_n^2$ and

$$c = b(0) = \mathbb{E}b(X_0) \le \mathbb{E}b(X_N) + \mathbb{E}\sum_{n=1}^N a_n^2.$$

However, we have the pointwise inequality $b(x) = -x^2 + c \le -|x| + \frac{1}{4} + c$, which plugged above yields

$$c \le -\mathbb{E}|X_N| + \mathbb{E}\sum_{n=1}^N a_n^2 + \frac{1}{4} + c.$$

This is the claim.

3. Problems

1. We roll a dice at most four times; at each time we may take the number we have just obtained and stop. Find the strategy which yields the largest expectation.

2. The urn contains N black balls and three white balls. We draw balls without replacement, one at a time, and after each draw we may decide to stop. Having stopped the procedure, we obtain the reward, which is equal to 0 if we picked a white ball during the process; if we only drew black balls, then the reward is equal to the number of black balls. Find the strategy which maximizes the expected reward.

3. We have a four-faced dice, with numbers 0, 1, 2 and 3 on its faces. We roll it N times; before each time, we decide (control) whether the outcome we will obtain will be added, or multiplied by the number collected so far. Find the strategy which yields optimal expected return.

4. A farmer annually produces X_k units of a certain crop and stores $(1-u_{k+1})X_k$ units of his production, where $0 \le u_{k+1} \le 1$, and invests the remaining $u_{k+1}X_k$ units, thus increasing the next year production to a level X_{k+1} given by

$$X_{k+1} = X_k + W_{k+1}u_{k+1}X_k, \qquad k = 0, 1, \ldots, N-1.$$

The scalars W_k are bounded independent random variables with identical probability distributions that depend neither on X_k nor on u_k. Furthermore, $\mathbb{E}W_k = \bar{w} > 0$. The problem is to find the optimal policy that maximizes the total expected product stored over N years:

$$\mathbb{E}\left\{X_N + \sum_{m=1}^{N}(1-u_m)X_{m-1}\right\}.$$

5. Consider the problem

$$\sup \mathbb{E}\left\{-\delta\exp(-\gamma X_N) - \sum_{m=1}^{N}\exp(-\gamma u_m)\right\},$$

where u_n are controls taking values anywhere in $\mathbb{R}$, δ and γ are given positive numbers, and where

$$X_{n+1} = 2X_n - u_{n+1} + V_{n+1}, \qquad X_0 \text{ given.}$$

Here V_{n+1}, $n = 0, 1, 2, ..., N-1$, are identically and independently distributed, and $K := \mathbb{E}\exp(-\gamma V_{n+1}) < \infty$.

6. Solve the problem

$$\sup \mathbb{E}\left\{\sum_{m=1}^{N} 2\sqrt{u_m} + aX_N\right\},$$

where $u_n \ge 0$, $a > 0$, $X_0 > 0$ and $X_{n+1} = X_n - u_{n+1}$ with probability $1/2$, $X_{n+1} = 0$ with probability $1/2$.

7. Solve the problem

$$\sup \mathbb{E}\left\{\sum_{m=1}^{N}(1-u_m)X_{m-1}\right\}, \qquad X_0 = 1,$$

where $u_n \in [0,1]$ and

$$X_{n+1} = X_n + u_{n+1}X_n + V_{n+1},$$

where $V_{n+1} \ge 0$ is exponentially distributed with parameter λ.

8. Consider the problem

$$\sup \mathbb{E}\left\{\sum_{m=1}^{N}\{(1-u_m)X_{m-1}^2 - u_m\} + 2X_N^2\right\},$$

subject to

$$X_{n+1} = u_{n+1}X_nV_{n+1}, \qquad u \in [0,1],$$

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

9. The capital of the investor at time n is equal to X_n. At each time, the investor splits its capital, devoting $c_n \in [0, X_n]$ to consumption and investing the remaining $X_n - c_n$ using two types of instruments. First, he puts $b_n(X_n - c_n)$ into the bank, with the return interest equal to $1+r$; for the remaining $(1-b_n)(X_n - c_n)$, he buys risky shares, for which the return is random and equals $1+\xi_n$. We assume that ξ_1, ξ_2, ..., ξ_N are independent and identically distributed random variables, with $\mathbb{P}(\xi \geq -1) = 1$. Maximize the functional

$$\mathbb{E}\left\{\sum_{n=0}^{N-1} \gamma^n c_n^\alpha + \gamma^N w X_N^\alpha\right\},$$

where γ, $\alpha \in [0,1]$ and $w > 0$ are given parameters.

10. Solve the problem

$$\sup \mathbb{E}\left\{\sum_{m=1}^{\infty} \gamma^m(-u_m^2 - X_{m-1}^2)\right\},$$

where $u_m \in \mathbb{R}$, $\gamma \in (0,1)$ and $(X_n)_{n\geq 0}$ is governed by the recurrence

$$X_m = X_{m-1} + u_m + \xi_m, \qquad m = 1, 2, \ldots.$$

Here ξ_1, ξ_2, ... is a sequence of independent identically distributed random variables with $\mathbb{E}\xi_m = 0$ and $\mathbb{E}\xi_m^2 = \sigma^2$.

11. Solve the problem

$$\sup \mathbb{E}\left\{\sum_{m=1}^{\infty} \beta^m(\ln u_m + \ln X_{m-1})\right\},$$

where $X_0 > 0$, $X_m = (X_{m-1} - u_m)\xi_m$, $u_m \in (0, X_{m-1})$, $\beta \in (0,1)$ and ξ_1, ξ_2, ... is a sequence of independent, identically distributed random variables satisfying $|\mathbb{E} \ln \xi_m| < \infty$.

12. At time n, a salesman has $X_n \in [0,1]$ tons of sugar; let ξ_n denote the demand for the sugar at that time. At time n, the salesman places the order u_{n+1} for the sugar, which is realized at time $n+1$; we assume that $u_{n+1} \in [-\max\{X_n - \xi_n, 0\}, 1 - \max\{X_n - \xi_n, 0\}]$. The cost at time n consists of two parts: the insufficiency cost $\max\{\xi_n - X_n, 0\}$ and the storage cost cX_n. Find the optimal strategy for the finite and infinite horizon, with the discount factor $\gamma \in (0,1)$.

13. Consider a Markov chain $(X_n)_{n\geq 0}$ on $\{0,1\}$, starting from 0. We assume that for each n, the transities are given by

$$\mathbb{P}(X_n = 1 | X_{n-1} = 0) = 1, \quad \mathbb{P}(X_n = 0 | X_{n-1} = 1) = u_n = 1 - \mathbb{P}(X_n = 1 | X_{n-1} = 1.$$

For a given $\gamma \in (0,1)$, find the strategy u which minimizes the cost functionals

$$\mathbb{E}\sum_{n=1}^{\infty} \gamma^{n-1}\sqrt{u_n} 1_{\{X_{n-1}=1\}}, \qquad \mathbb{E}\sum_{n=1}^{\infty} \gamma^{n-1}(\sqrt{u_n}+1) 1_{\{X_{n-1}=1\}}.$$

14. A king moves randomly across the chessboard of dimension $N \times N$; it starts at the upper-left corner and stops at the down-right corner. At each step, we may decided whether the king moves vertically or horizontally (having determined the direction, the king moves, picking the neighboring fields with equal probability). Find the strategy which minimizes the average number of moves.

15. Consider the symmetric random walk X over integers, starting from 0. At each time n, we may perform an action $u_n \in \{0,1\}$; the choice $u_n = 0$ does nothing, while $u_n = 1$ forces the walk to decrease by 1. For a given $\gamma \in (0,1)$, find the strategy which minimizes the functional

$$\sum_{n=1}^{\infty} \gamma^{n-1} \big(u_n + 1_{\{\max_{0 \le k \le n-1} X_k \ge 1\}} \big).$$

CHAPTER 3

Optimal Stopping Theory

1. Martingale approach: description

In the optimal stopping theory, one distinguishes two techniques: the martingale approach and the Markovian approach, the purpose of this section is to describe the first of them.

Suppose that $(\Omega, \mathcal{F}, \mathbb{P})$ is a given probability space, equipped with a filtration $(\mathcal{F}_n)_{n\geq 0}$, i.e., a nondecreasing family of sub-σ-algebras of $\mathcal{F}$. Let $X = (X_n)_{n\geq 0}$ be an adapted sequence of random variables, i.e., we assume that for each $n \geq 0$ the random variable X_n is measurable with respect to $\mathcal{F}_n$. Our objective will be to stop this sequence so that the expected return is maximized. The stopping procedure is described by a random variable $\tau : \Omega \to \{0, 1, 2, \ldots\}$, which returns the value of the time when the sequence $(X_n)_{n\geq 0}$ should be stopped. Clearly, a reasonable procedure decides whether to stop the sequence at time n is based on the observations up to time n; this means that for each n we have

$$\{\tau = n\} \in \mathcal{F}_n \qquad \text{for each } n \geq 0.$$

The random variable τ satisfying the above condition will be called a stopping time.

Let us put the discussion into a more precise framework. The optimal stopping problem concerns the study of

$$V_0 = \sup_{\tau} \mathbb{E}X_\tau, \tag{3.1}$$

where the supremum is taken over a certain family of adapted stopping times τ (which depends on the problem). We should point out that the study consists of two parts: (i) to compute the value V_0 as explicitly as possible; (ii) to identify the optimal stopping time τ_* (or the family of almost-optimal stopping times) for which the supremum V_0 is attained.

The first problem we encounter concerns the existence of $\mathbb{E}X_\tau$, to overcome which we need to impose some additional assumptions on X and τ. For example, if

$$\mathbb{E} \sup_{n\geq 0} |X_n| < \infty, \tag{3.2}$$

then the expectation $\mathbb{E}X_\tau$ is well defined for all stopping times τ. Another possibility is to restrict in (3.1) to those τ, for which the expectation exists. One way or another, we should emphasize that in general, this obstacle is just a technicality which is easily removed by some straightforward arguments (which might depend on the problem under the study). For the sake of simplicity and the clarity of the statements, we will assume that the condition (3.2) is satisfied, but it will be evident how to relax this requirement in other contexts.

So, let us assume that the supremum in (3.1) is taken over the family $\mathcal{M}$ of all stopping times τ. A successful treatment of this problem requires the introduction, for each $n \leq N$, the smaller family

$$\mathcal{M}_n^N = \{\tau \in \mathcal{M} : n \leq \tau \leq N\}.$$

We will also write $\mathcal{M}^N = \mathcal{M}_0^N$ and $\mathcal{M}_n = \mathcal{M}_n^\infty$. These families give rise to the related value functions

$$V_n^N = \sup_{\tau \in \mathcal{M}_n^N} \mathbb{E} X_\tau, \tag{3.3}$$

and we will use the notation $V^N = V_0^N$, $V_n = V_n^\infty$ and $V = V_0^\infty$. The primary goal of this section is to present the solution to (3.3) with the use of martingale approach.

Before we proceed, let us mention that the above problem can be put into the general framework of optimal control: we observe the sequence $(X_n)_{n\geq 0}$, term by term, and at each time we can perform two actions (controls): stop the observation and take the last variable, or continue. Perhaps the new feature is the lack of any additional structure of the sequence $(X_n)_{n\geq 0}$ (which was previously expressed in terms of the evolution function F).

1.1. Martingale approach: finite horizon. If $N < \infty$ (the case of "finite horizon"), then the problem (3.3) can be easily solved by means of the backward induction. Indeed, let us fix a nonnegative integer N and try to inspect the value functions as n decreases from N to 0. If $n = N$, then the class $\mathcal{M}_n^N$ consists of one stopping time $\tau \equiv N$ only and hence the optimal gain is equal to X_N (and $V_N^N = \mathbb{E}X_N$). If $n = N-1$, then we have two choices for the stopping time: we can either stop at time $N-1$ or continue and stop at time N. In the first case our gain is X_{N-1}; in the second case we do not know what the random variable X_N will be, so we can only say that on average, we will obtain $\mathbb{E}(X_N|\mathcal{F}_{N-1})$. Therefore, if $X_{N-1} \geq \mathbb{E}(X_N|\mathcal{F}_{N-1})$, we should stop immediately; otherwise, we should continue. For smaller values of n we proceed similarly. More precisely, define recursively the sequence $(B_n^N)_{0\leq n\leq N}$, representing the optimal gains at times 0, 1, 2, ..., N, as follows:

$$\begin{aligned} B_N^N &= X_N, \\ B_n^N &= \max\left\{X_n, \mathbb{E}(B_{n+1}^N|\mathcal{F}_n)\right\}, \qquad n = N-1,\, N-2,\, \ldots. \end{aligned} \tag{3.4}$$

The above discussion also suggests to consider the family of stopping times

$$\tau_n^N = \inf\left\{k \in \{n, n+1, \ldots, N\} \;:\; B_k^N = X_k\right\}, \tag{3.5}$$

for $n = 0, 1, 2, \ldots, N$.

THEOREM 3.1. *Suppose that N is a fixed integer and the sequence $X = (X_k)_{k=n}^N$ satisfies $\mathbb{E}\max_{n\leq k\leq N}|X_k| < \infty$. Consider the optimal stopping problem* (3.3) *and the sequence $(B_k^N)_{k=n}^N$, defined by* (3.4).

(i) The sequence $(B_k^N)_{k=n}^N$ is the smallest supermartingale majorizing $(X_k)_{k=n}^N$. In addition, the stopped sequence $\left(B_{\tau_n^N\wedge k}^N\right)_{k=n}^N$ is a martingale.

(ii) For any $0 \leq n \leq N$ we have, with probability 1,

$$B_n^N \geq \mathbb{E}(X_\tau|\mathcal{F}_n) \qquad \text{for any } \tau \in \mathcal{M}_n^N, \tag{3.6}$$

$$B_n^N = \mathbb{E}(X_{\tau_n^N}|\mathcal{F}_n). \tag{3.7}$$

(iii) The stopping time τ_n^N is optimal in (3.3) *and any other optimal stopping time τ_* satisfies $\tau_* \geq \tau_n^N$ almost surely.*

PROOF. (i) The supermartingale property and the majorization follow directly from the definition of the sequence $(B_n^N)_{n=0}^N$. If $(\bar{B}_n^N)_{n=0}^N$ is another supermartingale majorizing $(X_k)_{k=n}^N$, then the desired inequality $B_k^N \leq \bar{B}_k^N$ almost surely, $k = n, n+1, N+2, \dots, N$, can be proved by backward induction. Indeed, the estimate is trivial for $k = N$ (we have $B_N^N = X_N \leq \bar{B}_N^N$, by the majorization property of $\bar{B}$), and assuming its validity for k, we see that

$$\bar{B}_{k-1}^N \geq \max\{X_{k-1}, \mathbb{E}(\bar{B}_k^N|\mathcal{F}_{k-1})\} \geq \max\{X_{k-1}, \mathbb{E}(B_k^N|\mathcal{F}_{k-1})\} = B_{k-1}^N.$$

So, it remains to prove the martingale property of the stopped process $\left(B_{\tau_n^N \wedge k}^N\right)_{k=n}^N$. We compute directly that

$$\begin{aligned}
\mathbb{E}\left[B_{\tau_n^N \wedge (k+1)}^N \mid \mathcal{F}_k\right] &= \mathbb{E}\left[B_{\tau_n^N \wedge (k+1)}^N 1_{\{\tau_n^N \leq k\}} \mid \mathcal{F}_k\right] + \mathbb{E}\left[B_{\tau_n^N \wedge (k+1)}^N 1_{\{\tau_n^N > k\}} \mid \mathcal{F}_k\right] \\
&= \mathbb{E}\left[B_{\tau_n^N \wedge k}^N 1_{\{\tau_n^N \leq k\}} \mid \mathcal{F}_k\right] + \mathbb{E}\left[B_{k+1}^N 1_{\{\tau_n^N > k\}} \mid \mathcal{F}_k\right] \\
&= B_{\tau_n^N \wedge k}^N 1_{\{\tau_n^N \leq k\}} + 1_{\{\tau_n^N > k\}} \mathbb{E}\left[B_{k+1}^N \mid \mathcal{F}_k\right].
\end{aligned}$$

However, on the set $\{\tau_n^N > k\}$ we have $B_k^N > X_k$ and hence $B_k^N = \mathbb{E}(B_{k+1}^N|\mathcal{F}_k)$. This shows the identity $\mathbb{E}\left[B_{\tau_n^N \wedge (k+1)}^N \mid \mathcal{F}_k\right] = B_{\tau_n^N \wedge k}^N 1_{\{\tau_n^N \leq k\}} + 1_{\{\tau_n^N > k\}} B_k^N = B_{\tau_n^N \wedge k}^N$ and part (i) is established.

(ii) This follows at once from (i) and Doob's optional sampling theorem.

(iii) Taking the expectations in (3.6) and (3.7) gives $\mathbb{E}X_\tau \leq \mathbb{E}B_n^N = \mathbb{E}X_{\tau_n^N}$ for all $\tau \in \mathcal{M}_n^N$, showing that τ_n^N is indeed the optimal stopping time. Suppose that τ_* is another optimal stopping time. Then $B_{\tau_*}^N = X_{\tau_*}$ almost surely, since otherwise we would have

$$\mathbb{E}X_{\tau_*} < \mathbb{E}B_{\tau_*}^N \leq \mathbb{E}B_n^N = \mathbb{E}X_{\tau_n^N},$$

where the second inequality follows from Doob's optional sampling theorem and the supermartingale property of the sequence $(B_k^N)_{k=n}^N$. The contradiction shows that B_{τ_*} and X_{τ_*} must coincide, and clearly τ_n^N is the smallest stopping time which has this property. □

Thus we see that in the case of finite horizon, the solution to the optimal stopping problem is algorithmic. We write down the recursive formula for the Bellman sequence $(B_n^N)_{n=0}^N$ and note that $V^N = \mathbb{E}B_0^N$.

1.2. Martingale approach: infinite horizon. The above method required N to be a finite integer, since we have needed the variable X_N to start the backward recurrence. In the case $N = \infty$ one could try to use approximation-type arguments (of the form $V_n^\infty = \lim_{N\to\infty} V_n^N$), but these do not necessarily work in general, so we will proceed in a different manner. By (3.6) and (3.7) it seems tempting to write

$$B_n^N = \sup_{\tau \in \mathcal{M}_n^N} \mathbb{E}(X_\tau|\mathcal{F}_n).$$

However, two problems arise. The first is that (3.6) and (3.7) hold true on a set of full measure only which might depend on the stopping time, so the above identity might fail to hold. A second obstacle is that the supremum on the right need not be even measurable. To overcome these difficulties, a typical argument in the theory of optimal stopping is to introduce the concept of essential supremum.

DEFINITION 3.1. Let $(Z_\alpha)_{\alpha\in I}$ be a family of random variables. Then there is a countable subset J of I such that the random variable $\overline{Z} = \sup_{\alpha\in J} Z_\alpha$ satisfies the following two properties:

(i) $\mathbb{P}(Z_\alpha \le \overline{Z}) = 1$ for each $\alpha \in I$,
(ii) if $\tilde{Z}$ is another random variable satisfying (i) in the place of $\overline{Z}$, then $\mathbb{P}(\overline{Z} \le \tilde{Z}) = 1$.

The random variable $\overline{Z}$ is called the *essential supremum* of $(Z_\alpha)_{\alpha\in I}$ and is denoted by $\operatorname{ess\,sup}_{\alpha\in I} Z_\alpha$. In addition, if $\{Z_\alpha\ :\ \alpha\in I\}$ is upwards directed in the sense that for any α, $\beta\in I$ there is $\gamma\in I$ such that $\max\{Z_\alpha, Z_\beta\}\le Z_\gamma$, then the countable set $J=\{\alpha_1,\alpha_2,\ldots\}$ can be chosen so that $Z_{\alpha_1}\le Z_{\alpha_2}\le\ldots$ and $\operatorname{ess\,sup}_{\alpha\in I} Z_\alpha = \lim_{n\to\infty} Z_{\alpha_n}$.

Now we see that (3.6) and (3.7) give the identity

$$B_n^N = \operatorname*{ess\,sup}_{\tau\in\mathcal{M}_n^N} \mathbb{E}(X_\tau|\mathcal{F}_n) \tag{3.8}$$

with probability 1. A nice feature of this alternative characterization of the sequence $(B_n^N)_{n=0}^N$ is that it extends naturally to the setting of infinite horizon (i.e., for $N=\infty$) and, as we shall prove now, provides the desired solution.

So, consider the optimal stopping problem (3.3) for $N=\infty$:

$$V_n = \sup_{\tau\ge n} \mathbb{E}X_\tau. \tag{3.9}$$

For $n = 0, 1, 2, \ldots$, introduce the random variable

$$B_n = \operatorname*{ess\,sup}_{\tau\ge n} \mathbb{E}(X_\tau|\mathcal{F}_n) \tag{3.10}$$

and the stopping time

$$\tau_n = \inf\{k\ge n : B_k = X_k\}, \tag{3.11}$$

with the usual convention $\inf\emptyset=\infty$. In the literature, the sequence $(B_n)_{n\ge0}$ is often referred to as the Snell envelope of X.

We will establish the following analogue of Theorem 3.1.

THEOREM 3.2. *Suppose that the sequence* $(X_n)_{n\ge0}$ *satisfies* $\mathbb{E}\sup_{n\ge0}|X_n|<\infty$ *and consider the optimal stopping problem* (3.9). *Then the following statements hold true.*

(i) For any $n\ge0$ *we have the recurrence relation*

$$B_n = \max(X_n, \mathbb{E}(B_{n+1}|\mathcal{F}_n)).$$

(ii) We have $\mathbb{P}(B_n\ge\mathbb{E}(X_\tau|\mathcal{F}_n))=1$ *for all* $\tau\in\mathcal{M}_n$ *and, if the stopping time* τ_n *is finite almost surely, then* $\mathbb{P}(B_n=\mathbb{E}(X_{\tau_n}|\mathcal{F}_n))=1$.

(iii) If $\mathbb{P}(\tau_n<\infty)=1$, *then* τ_n *is optimal in* (3.9). *Furthermore, if* τ_* *is another optimal stopping time for* (3.9), *then* $\tau_n\le\tau_*$ *almost surely.*

(iv) The sequence $(B_k)_{k\ge n}$ *is the smallest supermartingale which majorizes* $(X_k)_{k\ge n}$. *Moreover, the stopped process* $(B_{\tau_n\wedge k})_{k\ge n}$ *is a martingale.*

PROOF. We will only establish (i), the other parts can be shown with the argumentation similar to that used in the proof of Theorem 3.1. We need to show two inequalities to prove the identity. Take $\tau \in \mathcal{M}_n$ and let $\tau' = \tau \vee (n+1)$. Then $\tau' \in \mathcal{M}_{n+1}$ and since $\{\tau \geq n+1\} \in \mathcal{F}_n$, we may write

$$\begin{aligned}
\mathbb{E}(X_\tau|\mathcal{F}_n) &= \mathbb{E}(X_\tau 1_{\{\tau=n\}}|\mathcal{F}_n) + \mathbb{E}(X_\tau 1_{\{\tau\geq n+1\}}|\mathcal{F}_n) \\
&= 1_{\{\tau=n\}} X_n + 1_{\{\tau\geq n+1\}} \mathbb{E}(X_{\tau'}|\mathcal{F}_n) \\
&= 1_{\{\tau=n\}} X_n + 1_{\{\tau\geq n+1\}} \mathbb{E}\big[\mathbb{E}(X_{\tau'}|\mathcal{F}_{n+1})|\mathcal{F}_n\big] \\
&\leq 1_{\{\tau=n\}} X_n + 1_{\{\tau\geq n+1\}} \mathbb{E}(B_{n+1}|\mathcal{F}_n) \\
&\leq \max\{X_n, \mathbb{E}(B_{n+1}|\mathcal{F}_n)\}.
\end{aligned}$$

This proves the inequality "$\leq$". To show the reverse, observe that the family $\{\mathbb{E}(B_\tau|\mathcal{F}_{n+1}) \; : \; \tau \in \mathcal{M}_{n+1}\}$ is upwards directed. Indeed, if α, $\beta \in \mathcal{M}_{n+1}$ and we set $\gamma = \alpha 1_A + \beta 1_{\Omega\setminus A}$, where $A = \{\mathbb{E}(X_\alpha|\mathcal{F}_{n+1}) \geq \mathbb{E}(X_\beta|\mathcal{F}_{n+1})\}$, then γ is a stopping time belonging to $\mathcal{M}_{n+1}$ and

$$\begin{aligned}
\mathbb{E}(X_\gamma|\mathcal{F}_{n+1}) &= \mathbb{E}(X_\alpha 1_A + X_\beta 1_{\Omega\setminus A}|\mathcal{F}_{n+1}) \\
&= 1_A \mathbb{E}(X_\alpha|\mathcal{F}_{n+1}) + 1_{\Omega\setminus A}\mathbb{E}(X_\beta|\mathcal{F}_{n+1}) \\
&= \max\big\{\mathbb{E}(X_\alpha|\mathcal{F}_{n+1}), \mathbb{E}(X_\beta|\mathcal{F}_{n+1})\big\}.
\end{aligned}$$

Therefore, there is a sequence $\{\sigma_k : k \geq 1\}$ in $\mathcal{M}_{n+1}$ such that

$$\operatorname*{ess\,sup}_{\tau\in\mathcal{M}_{n+1}} \mathbb{E}(X_\tau|\mathcal{F}_{n+1}) = \lim_{k\to\infty} \mathbb{E}(X_{\sigma_k}|\mathcal{F}_{n+1})$$

and $\mathbb{E}(X_{\sigma_1}|\mathcal{F}_{n+1}) \leq \mathbb{E}(X_{\sigma_2}|\mathcal{F}_{n+1}) \leq \ldots$ with probability 1. Now we can write, by Lebesgue's monotone convergence theorem,

$$\begin{aligned}
\mathbb{E}(B_{n+1}|\mathcal{F}_n) &= \mathbb{E}\left(\operatorname*{ess\,sup}_{\tau\in\mathcal{M}_{n+1}} \mathbb{E}(X_\tau|\mathcal{F}_{n+1})\middle|\mathcal{F}_n\right) \\
&= \mathbb{E}\left(\lim_{k\to\infty} \mathbb{E}(X_{\sigma_k}|\mathcal{F}_{n+1})\middle|\mathcal{F}_n\right) \\
&= \lim_{k\to\infty} \mathbb{E}\big(\mathbb{E}(X_{\sigma_k}|\mathcal{F}_{n+1})|\mathcal{F}_n\big) = \lim_{k\to\infty} \mathbb{E}(X_{\sigma_k}|\mathcal{F}_n) \leq B_n.
\end{aligned}$$

Since $B_n \geq X_n$ (which can be trivially obtained by considering $\tau \equiv n$ in the definition of B_n), we get the desired identity. □

In the remaining part of this subsection, let us inspect the connection between the contexts of finite and infinite horizons. One easily checks that the random variables B_n^N and τ_n^N do not decrease as we increase N. Consequently, the limits

$$B_n^\infty := \lim_{N\to\infty} B_n^N \qquad \text{and} \qquad \tau_n^\infty := \lim_{N\to\infty} \tau_n^N$$

exist on a set of full measure. Furthermore, we also see that the sequence $(V_n^N)_{N=n}^\infty$ is nondecreasing, so the quantity $V_n^\infty = \lim_{N\to\infty} V_n^\infty$ is well-defined. Now it follows directly from (3.5), (3.8), (3.10) and (3.11) that

$$B_n^\infty \leq B_n \qquad \text{and} \qquad \tau_n^\infty \leq \tau_n \tag{3.12}$$

almost surely. Moreover, we also have

$$V_n^\infty \leq V_n. \tag{3.13}$$

THEOREM 3.3. *Suppose that* $\mathbb{E}\sup_{n\geq 0}|X_n|<\infty$ *and consider the optimal stopping problems* (3.3) *and* (3.9). *Then equalities hold in* (3.12) *and* (3.13).

PROOF. Letting $N\to\infty$ in the recurrence relation (3.4) yields

$$B_n^\infty = \max\{X_n, \mathbb{E}(B_{n+1}^\infty|\mathcal{F}_n)\}, \qquad n=0,\,1,\,2,\,\ldots,$$

by Lebesgue's monotone convergence theorem. Consequently, $(B_n^\infty)_{n\geq 0}$ is an adapted supermartingale dominating $(X_n)_{n\geq 0}$. Thus $B_n^\infty\geq B_n$ for each n, by the fourth part of the preceding theorem. This shows the identity $B^\infty=B$ almost surely, and the remaining equalities follow immediately. □

EXAMPLE 3.4. Strict inequalities may hold in (3.12) and (3.13) if the integrability condition on $\sup_{n\geq 0}|X_n|$ is not imposed. To see this, let $\varepsilon_0,\,\varepsilon_1,\,\varepsilon_2,\,\ldots$ be a sequence of independent Rademacher variables and set $X_n=\varepsilon_0+\varepsilon_1+\ldots+\varepsilon_n$. Then the process $(X_n)_{n\geq 0}$ is a martingale with respect to the natural filtration, so $V_n^N=0$, $B_n^N=X_n$ and $\tau_n^N=n$ for all $0\leq n\leq N<\infty$. Consequently, these identities are preserved in the limit: $V_n^\infty=0$, $B_n^\infty=X_n$ and $\tau_n^\infty=n$ for all n. On the other hand, it is well-known that for any positive integer a, the stopping time $\sigma_n=\inf\{k\geq n: X_k=a\}$ is finite almost surely and hence $V_n\geq \mathbb{E}X_{\sigma_n}=a$. Since a was arbitrary, we get $V_n=\infty$, $B_n=\infty$ and $\tau_n=\infty$ with probability 1.

In comparison to the finite-horizon case, the solution in the infinite case is no longer algorithmic. One typically applies the following procedure, already known to us from the general theory of optimal control. Namely, basing on some structural properties of the problem, or simply by guessing the optimal stopping time, we come up with the candidate $\mathcal{B}_n$ for the Bellman sequence. Then all that needs to be done, is the identity $B_n=\mathcal{B}_n$. The estimate $B_n\geq\mathcal{B}_n$ follows directly from the construction (if $\mathcal{B}_n$ is based on a guess for the optimal stopping time), and the reverse bound is established by checking that $(\mathcal{B}_n)_{n\geq 0}$ is indeed a supermartingale majorant of a given sequence X. It is best to explain this idea on a concrete example.

1.3. An example. Let $\xi_0,\,\xi_1,\,\xi_2,\,\ldots$ be i.i.d. random variables following the $Exp(1)$ law, and let $c>0$ be a fixed constant. We will solve the optimal stopping problem

$$V=\sup_{\tau\in\mathcal{M}}\mathbb{E}\bigg[\max\{\xi_0,\,\xi_1,\,\xi_2,\,\ldots,\,\xi_\tau\}-c\tau\bigg].$$

For the sake of clarity, we will split the reasoning into three separate steps. To put this problem into the general framework of the optimal stopping theory, we set

$$X_n=\max\{\xi_0,\,\xi_1,\,\xi_2,\,\ldots,\,\xi_n\}-cn.$$

Then $V=\sup_{\tau\in\mathcal{M}}\xi_\tau$ and we may proceed.

STEP 1. GUESSING THE OPTIMAL STOPPING RULE AND THE ASSOCIATED EXPECTATION. This is an informal step and it requires some thought and experimentation. It seems reasonable to conjecture that the optimal stopping rule should be of the following threshold type:

$$\tau_a=\inf\{n:\xi_n\geq a\}$$

for some unknown constant a. To find a, let us first compute the corresponding expectation

$$\mathbb{E}\Big[\max\{\xi_0,\,\xi_1,\,\xi_2,\,\ldots,\,\xi_{\tau_a}\} - c\tau_a\Big]. \tag{3.14}$$

Note that τ_a has the geometric distribution with parameter $\mathbb{P}(\xi_0 \geq a)$ and hence

$$\mathbb{E}\tau_a = \frac{\mathbb{P}(\xi_0 < a)}{\mathbb{P}(\xi_0 \geq a)} = \frac{1-e^{-a}}{e^{-a}} = e^a - 1.$$

Furthermore, we have

$$\begin{aligned}\mathbb{E}\max\{\xi_0,\,\xi_1,\,\xi_2,\,\ldots,\,\xi_{\tau_a}\} &= \mathbb{E}\xi_{\tau_a} = \mathbb{E}(\xi_0|\xi_0 \geq a)\\ &= \frac{1}{\mathbb{P}(\xi_0 \geq a)}\int_{\{\xi_0\geq a\}} \xi_0 \mathrm{d}\mathbb{P}\\ &= e^a\left(ae^{-a} + 1\cdot e^{-a}\right) = a+1\end{aligned}$$

and hence the expectation (3.14) equals $a+1-c(e^a-1)$. We want to maximize this expectation (over all possible stopping times, and so, in particular, over τ_a): we easily check that

$$\max_a\Big\{a+1-c(e^a-1)\Big\} = \begin{cases}1 & \text{if } c \geq 1 \text{ (maximum attained at } a_* = 0),\\ c - \ln c & \text{if } c < 1 \text{ (maximum attained at } a_* = -\ln c).\end{cases}$$

Let us denote the right-hand side by $\tilde{V}$. This is the candidate for the value of our optimal stopping problem.

Step 2. Guessing the Snell envelope. Actually, the computation from the previous step easily yields the corresponding candidate for the Snell envelope. From the general theory, we know that

$$B_n = \operatorname*{esssup}_{\tau\geq n} \mathbb{E}(X_\tau|\mathcal{F}_n).$$

Take $\tau = \tau_{a_*} \vee n$ (the additional maximum with n is to enforce the estimate $\tau \geq n$): by the above computations, for this special stopping time, we have

$$\mathbb{E}(G_\tau|\mathcal{F}_n) = \begin{cases}\max\{\xi_0,\xi_1,\ldots,\xi_n\} - cn & \text{if } \max\{\xi_0,\xi_1,\ldots,\xi_n\} \geq a_*,\\ \tilde{V} - c(n+1) & \text{if } \max\{\xi_0,\xi_1,\ldots,\xi_n\} < a_*.\end{cases}$$

Let us denote the right-hand side by $\tilde{B}_n$: this is our candidate for the Snell envelope of $(G_n)_{n\geq 0}$. By the very definition, we have $\tilde{B}_n \leq B_n$.

Step 3. Verification of the properties of $\tilde{B}$. Now we will check that $(\tilde{B}_n)_{n\geq 0}$ is a supermartingale majorizing $(G_n)_{n\geq 0}$. Then by the general theory we will obtain the reverse estimate $\tilde{B}_n \geq B_n$, which will show that $\tilde{B}$ coincides with the Snell envelope and the stopping time τ_{a_*} is optimal.

We start with the majorization. On the set $\{\max\{\xi_0,\xi_1,\ldots,\xi_n\} \geq a_*\}$ we have $\tilde{B}_n = G_n$. On the other hand, on $\{\max\{\xi_0,\xi_1,\ldots,\xi_n\} < a_*\}$ (which is nonempty iff $a_* > 0$, i.e., $c < 1$), the majorization is equivalent to

$$\tilde{V} - c(n+1) \geq \max\{\xi_0,\xi_1,\ldots,\xi_n\} - cn$$

or $-\ln c \geq \max\{\xi_0,\xi_1,\ldots,\xi_n\}$: but this is trivial, since $-\ln c = a_*$.

It remains to check the supermartingale property of $\tilde{B}_n$:

$$\mathbb{E}(\tilde{B}_{n+1}|\mathcal{F}_n) \leq \tilde{B}_n, \qquad n = 0,\,1,\,2,\,\ldots. \tag{3.15}$$

On the set $\{\max\{\xi_0, \xi_1, \dots, \xi_n\} \geq a_*\} \in \mathcal{F}_n$ we automatically have the bound $\max\{\xi_0, \xi_1, \dots, \xi_{n+1}\} \geq a_*$ and hence

$$\mathbb{E}(\tilde{B}_{n+1}|\mathcal{F}_n) = \mathbb{E}\Big(\max\{\xi_0, \xi_1, \dots, \xi_{n+1}\} - c(n+1)|\mathcal{F}_n\Big).$$

Now, consider the random variable $\xi = \max\{\xi_0, \xi_1, \dots, \xi_n\}$. But for any $a > a_*$,

$$\mathbb{E}\Big(\max\{a, \xi_{n+1}\} - c(n+1)\Big) \leq a - cn$$

(this is equivalent to $e^{-a} \leq c$ and holds true, since $e^{-a_*} \leq c$). Hence

$$\begin{aligned}\mathbb{E}\Big(\max\{\xi_0, \xi_1, \dots, \xi_{n+1}\} - c(n+1)|\mathcal{F}_n\Big) &= \mathbb{E}\Big(\max\{a, \xi_{n+1}\} - c(n+1)\Big)\Big|_{a=\xi} \\ &\leq \xi - c_n = \tilde{B}_n.\end{aligned}$$

Next, we analyze (3.15) on the set $\{\max\{\xi_0, \xi_1, \dots, \xi_n\} < a_*\}$ (which is nonempty iff $a_* > 0$, i.e., $c < 1$). On this set, we have the identity

$$\tilde{B}_{n+1} = \begin{cases} \xi_{n+1} - c(n+1) & \text{if } \xi_{n+1} \geq a_*, \\ \tilde{V} - c(n+2) & \text{if } \xi_{n+1} < a_*, \end{cases}$$

which is independent of $\mathcal{F}_n$. Consequently, the conditional expectation $\mathbb{E}(\tilde{B}_{n+1}|\mathcal{F}_n)$ is equal to the average of the right-hand side above, i.e., to

$$\begin{aligned}&\mathbb{E}\Big[(\xi_{n+1} - c(n+1))1_{\{\xi_{n+1} \geq a_*\}} + (\tilde{V} - c(n+2))1_{\{\xi_{n+1} < a_*\}}\Big] \\ &= \mathbb{E}\xi_{n+1}1_{\{\xi_{n+1} \geq a_*\}} - c(n+1) + (\tilde{V} - c)\mathbb{P}(\xi_{n+1} < a_*) \\ &= e^{-a_*}(a_* + 1) - c(n+1) + (\tilde{V} - c)(1 - e^{-a_*}) = V - c(n+1) = \tilde{B}_n,\end{aligned}$$

where we have used the identity $a_* = -\ln c$.

This completes the proof of (3.15) and finishes the analysis of the optimal stopping problem.

2. Markovian approach

Throughout this section, we assume that $X = (X_0, X_1, X_2, \dots)$ is a Markov family defined on $(\Omega, \mathcal{F}, (\mathcal{F}_n)_{n\geq 0}, (P_x)_{x\in E})$, taking values in some topological space $(E, \mathcal{B}(E))$. For the sake of simplicity, we will assume that $E = \mathbb{R}^d$ for some $d \geq 1$, though the reasoning remains essentially the same for other topological spaces. As usual, we assume that for each $x \in E$, we have $X_0 \equiv x$ $\mathbb{P}_x$-almost surely. Let us also introduce the transition operator T of X, which acts by the formula

$$Tf(x) = \mathbb{E}_x f(X_1) \qquad \text{for } x \in E,$$

on the class I of all measurable functions $f : E \to \mathbb{R}$ such that $f(X_1)$ is $\mathbb{P}_x$-integrable for all $x \in E$.

Suppose that N is a nonnegative integer and let $G : E \to \mathbb{R}$ be a measurable function satisfying

$$\mathbb{E}_x \left(\sup_{0\leq n\leq N} |G(X_n)| \right) < \infty \qquad \text{for all } x \in E. \tag{3.16}$$

Consider the associated finite-horizon optimal stopping problem

$$V^N(x) = \sup \mathbb{E}_x G(X_\tau), \tag{3.17}$$

where $x \in E$ and the supremum is taken over all $\tau \in \mathcal{M}^N$. Obviously, if we define $G_n = G(X_n)$ for $n = 0, 1, 2, \ldots$, then for each separate x this problem is of the form considered in the preceding sections (with $\mathbb{P}$ and $\mathbb{E}$ replaced by $\mathbb{P}_x$ and $\mathbb{E}_x$). However, the joint study of the whole family of optimal stopping problems depending on the initial value x enables the exploitation of the additional Markovian structure of the sequence $(X_n)_{n\geq 0}$.

For a given x, let us consider the random variables B_n^N and the stopping times τ_n^N, $n = 0, 1, 2, \ldots, N$, defined by (3.4) and (3.5). We also introduce the sets

$$C_n = \{x \in E \,:\, V^{N-n}(x) > G(x)\},$$
$$D_n = \{x \in E \,:\, V^{N-n}(x) = G(x)\},$$

for $n = 0, 1, 2, \ldots, N$; we will call these the continuation and stopping regions, respectively. Finally, define the stopping time

$$\tau_D = \inf\{0 \leq n \leq N \,:\, X_n \in D_n\}.$$

Since $V^0 = G$, by the very definition (3.17), we see that $X_N \in D_N$ and hence the stopping time τ_D is finite (it does not exceed N).

Theorem 3.5. *Assume that the function G satisfies the integrability condition* (3.16) *and consider the optimal stopping problem* (3.17).

(i) For any $n = 0, 1, 2, \ldots, N$ we have $B_n^N = V^{N-n}(X_n)$.

(ii) The function $x \mapsto V^n(x)$ satisfies the Wald-Bellman equation

$$V^n(x) = \max\{G(x), TV^{n-1}(x)\}, \qquad x \in E, \tag{3.18}$$

for $n = 1, 2, \ldots, N$.

(iii) The stopping time τ_D is optimal in (3.17). *If τ_* is another optimal stopping time, then $\tau_D \leq \tau_*$ $\mathbb{P}_x$-almost surely for all $x \in E$.*

(iv) For each $x \in E$, the sequence $(V^{N-n}(X_n))_{n=0}^N$ is the smallest $\mathbb{P}_x$-supermartingale majorizing $(G(X_n))_{n=0}^N$, and the stopped sequence $\big(V^{N-n\wedge\tau_D}(X_{n\wedge\tau_D})\big)_{n=0}^N$ is a $\mathbb{P}_x$-martingale.

Proof. We only need to establish (i) and (ii); the remaining parts follow at once from Theorem 3.1. To verify (i), recall that

$$B_n^N = \mathbb{E}_x\left[G(X_{\tau_n^N})|\mathcal{F}_n\right]$$

for all $n = 0, 1, 2, \ldots, N$. This shows the claim for $n = 0$, by the very definition of $V^N(x)$. On the other hand, for $n \geq 1$ we apply the Markov property to get

$$B_n^N = \mathbb{E}_y\left[G(X_{\tau_0^{N-n}})\right]\bigg|_{y=X_n} = V^{N-n}(y)\big|_{y=X_n} = V^{N-n}(X_n).$$

(ii) We apply the definition of the sequence $(B_n^N)_{n=0}^N$ and part (i) to obtain that $\mathbb{P}_x$-almost surely,

$$\begin{aligned} V^N(x) = V^N(X_0) = B_0^N &= \max\{G(X_0), \mathbb{E}_x(B_1^N|\mathcal{F}_0)\} \\ &= \max\{G(x), \mathbb{E}_x(V^{N-1}(X_1)|\mathcal{F}_0)\} \\ &= \max\{G(x), TV^{N-1}(x)\}. \end{aligned}$$

□

Part (ii) above gives the following iterative method of solving (3.17). Define the operator Q acting on $f \in I$ by the formula

$$Qf(x) = \max\{G(x), Tf(x)\}, \qquad x \in E.$$

COROLLARY 3.6. *We have $V^N(x) = Q^N G(x)$ for all $x \in E$ and all integers N.*

Let us illustrate the above considerations by analyzing the following simple example.

EXAMPLE 3.7. Let $(S_n)_{n\geq 0}$ be a symmetric random walk over the space $E = \{-2,-1,0,1,2\}$ stopped at $\{-2,2\}$. Clearly, $(S_n)_{n\geq 0}$ is a Markov family on E. Set $G(x) = x^2(x+2)$ and consider the optimal stopping problem

$$V^N(x) = \sup_{\tau\leq N} \mathbb{E}_x G(S_\tau), \qquad x \in E.$$

To treat the problem successfully, we compute the sequence V^0, V^1, V^2, V^3, $\ldots$. Directly from (3.18), we have

$$\begin{aligned} V^n(x) &= \max\{G(x), TV^{n-1}(x)\} \\ &= \begin{cases} \max\{G(x), V^{n-1}(x)\} & \text{if } x \in \{-2,2\}, \\ \max\left\{G(x), \frac{1}{2}\big(V^{n-1}(x-1) + V^{n-1}(x+1)\big)\right\} & \text{if } x \in \{-1,0,1\}. \end{cases} \end{aligned}$$

For notational simplicity, let us identify a function $f : E \to \mathbb{R}$ with the sequence of its values $f(-2)$, $f(-1)$, $f(0)$, $f(1)$, $f(2)$. Using the above recurrence, we compute that

$$\begin{array}{rrrrrr} V^0 = G: & 0, & 1, & 0, & 3, & 16, \\ V^1: & 0, & 1, & 2, & 8, & 16, \\ V^2: & 0, & 1, & 4\frac{1}{2}, & 9, & 16, \\ V^3: & 0, & 2\frac{1}{4}, & 5 & 10\frac{1}{4}, & 16, \\ V^4: & 0, & 2\frac{1}{2}, & 6\frac{1}{4}, & 10\frac{1}{2}, & 16, \\ \ldots & & & & & \end{array}$$

and so on. Suppose that we want to solve the problem

$$V^4(x) = \sup_{\tau\leq 4} \mathbb{E}_x G(S_\tau), \qquad x \in E.$$

The value function V^4 has been derived above; to describe the optimal stopping strategy, let us write down the continuation and stopping regions C_i and D_i, $i = 0, 1, 2, 3, 4$. Directly from the above formulas for V^i, we see that

$$\begin{aligned} C_0 &= \{-1,0,1\}, & D_0 &= \{-2,2\}, \\ C_1 &= \{-1,0,1\}, & D_1 &= \{-2,2\}, \\ C_2 &= \{0,1\}, & D_2 &= \{-2,-1,2\}, \\ C_3 &= \{0,1\}, & D_3 &= \{-2,-1,2\}, \\ C_4 &= \emptyset & D_4 &= \{-2,-1,0,1,2\}. \end{aligned}$$

The optimal strategy is to wait for the first step n at which we visit the corresponding stopping set D_n; then we stop the process ultimately.

We turn our attention to the case of infinite horizon, i.e., we consider the optimal stopping problem (or rather a family of optimal stopping problems)

$$V(x) = \sup \mathbb{E}_x G(X_\tau), \qquad x \in E, \tag{3.19}$$

where the supremum is taken over the class $\mathcal{M}$ of all adapted stopping times. Recall that the class I consists of all measurable functions $f : E \to \mathbb{R}$ such that $f(X_1)$ is $\mathbb{P}_x$-integrable for all $x \in E$. The following notion will be crucial in our further considerations.

DEFINITION 3.2. The function $f \in I$ is called *superharmonic* (or *excessive*) if we have

$$Tf(x) \leq f(x) \qquad \text{for all } x \in E.$$

We have the following simple observation.

LEMMA 3.8. *The function $f \in I$ is superharmonic if and only if $(f(X_n))_{n\geq 0}$ is a supermartingale under each $\mathbb{P}_x$, $x \in E$.*

PROOF. If f is superharmonic, then by Markov property,

$$\mathbb{E}_x(f(X_{n+1})|\mathcal{F}_n) = \mathbb{E}_y f(X_1)\big|_{y=X_n} = Tf(X_n) \leq f(X_n),$$

for each n. To show the reverse implication, observe that if $(f(X_n))_{n\geq 0}$ is a supermartingale under each $\mathbb{P}_x$, then in particular

$$Tf(x) = \mathbb{E}_x(f(X_1)|\mathcal{F}_0) \leq f(x). \quad \square$$

To formulate the main theorem, we introduce the corresponding continuation set C and stopping set D by

$$C = \{x \in E : V(x) > G(x)\},$$
$$D = \{x \in E : V(x) = G(x)\}.$$

Moreover, we define the stopping time $\tau_D = \inf\{n : X_n \in D\}$. In contrast to the case of finite horizon, this stopping time need not be finite (which will force us to impose some additional assumptions: see the statement below).

THEOREM 3.9. *Consider the optimal stopping problem* (3.19) *and assume that*

$$\mathbb{E}_x \sup_{n\geq 0} |G(X_n)| < \infty, \qquad x \in E. \tag{3.20}$$

Then the following holds.

(i) The function V satisfies the Wald-Bellman equation

$$V(x) = \max\{G(x), TV(x)\}. \tag{3.21}$$

(ii) If τ_D is finite $\mathbb{P}_x$-almost surely for all $x \in E$, then τ_D is the optimal stopping time. If τ_* is another optimal stopping time, then $\tau_* \geq \tau_D$ $\mathbb{P}_x$-almost surely.

(iii) The value function V is the smallest superharmonic function which majorizes the gain function G on E.

(iv) The stopped sequence $(V(X_{\tau_D \wedge n}))_{n\geq 0}$ is a $\mathbb{P}_x$-martingale for every $x \in E$.

PROOF. This follows immediately from the case of finite horizon and the limit Theorem 3.3. $\square$

Let us make here an important comment on the uniqueness of the solutions to the Wald-Bellman equations (3.18) and (3.21). Clearly, in the case of finite horizon there is only one solution: indeed, the starting function V^0 coincides with G and the formula (3.18) produces a unique sequence $V^1, V^2, \ldots, V^N$. In the case of infinite horizon, the situation is less transparent. For instance, if G is a constant function, say, $G \equiv c$, then any constant function $V \equiv c'$ for some $c' \geq c$ satisfies the Wald-Bellman equation. However, any solution to (3.21) is a superharmonic function majorizing G, so part (iii) of Theorem 3.9 immediately yields the following "minimality principle".

COROLLARY 3.10. *The value function V is the minimal solution to* (3.21).

EXAMPLE 3.11. Let us provide solution to the infinite-horizon version of Example 3.7. Under the notation used there, we study the optimal stopping problem

$$V(x) = \sup_{\tau \in \mathcal{M}} \mathbb{E}G(S_\tau), \qquad x \in E.$$

The function G is bounded, so the integrability assumption of Theorem 3.9 is satisfied. Thus, we know that V is the least superharmonic function which majorizes G: here the superharmonicity means that

$$V(x) \geq \frac{1}{2}\big(V(x-1) + V(x+1)\big), \qquad \text{for } x \in \{-1, 0, 1\}.$$

In other words, we search for the smallest concave function on $\{-2,-1,0,1,2\}$ majorizing the function G. One easily checks that the function $x \mapsto 4(x+2)$ is concave (since it is linear), majorizes G and coincides with G at the endpoints ± 2. Thus it is the smallest majorant of G and hence it must be equal to the value function V.

EXAMPLE 3.12. Let ξ_1, ξ_2, $\ldots$ be a sequence of i.i.d. random variables with the distribution given by $\mathbb{P}(\xi_i = 1) = p$, $\mathbb{P}(\xi_i = -1) = q$, where $p+q = 1$ and $p < q$. For a given integer x, define $S_n = x + \xi_1 + \xi_2 + \ldots + \xi_n$, $n = 0, 1, 2, \ldots$. Then the sequences $(S_n)_{n\geq 0}$ (with varying x) form a Markov family. Consider the optimal stopping problem

$$V(x) = \sup_{\tau \in \mathcal{M}} \mathbb{E}_x S_\tau^+, \qquad x \in E.$$

One easily checks the integrability assumption (3.20) (with $G(x) = x^+$) is satisfied. This follows from the well-known fact that

$$\mathbb{P}\Big(\sup_{n\geq 0}(\xi_1 + \xi_2 + \ldots + \xi_n) \geq k\Big) = \left(\frac{p}{q}\right)^k, \qquad k = 0, 1, 2, \ldots. \tag{3.22}$$

Thus, we need to find the least superharmonic majorant of G: V is the least function on $\mathbb{Z}$ satisfying

$$V(x) = \max\left\{x^+, pV(x+1) + qV(x-1)\right\}, \qquad x \in \mathbb{Z}.$$

To identify this object, let us try to inspect the properties of the continuation set C and the stopping region D. A little thought suggests that these sets should be of the form $C = \{\ldots, b-2, b-1\}$, $D = \{b, b+1, \ldots\}$ for some positive integer b (possibly infinite). While this is more or less clear by some intuitive argumentation, we should point out here that this can also be shown rigorously. Indeed, pick $x \in \mathbb{Z}_-$ and take the stopping time $\tau \equiv -x+1$. Then $V(x) \geq \mathbb{E}_x S_\tau^+ = p^{-x+1} > 0 = G(x)$, so in particular C contains all nonpositive integers. Furthermore, if $x > 0$ lies in C, then so does $x-1$. To see this, note that for any $a \in \mathbb{Z}$ we have

$$(x+a)^+ - x^+ \leq (x-1+a)^+ - (x-1)^+$$

(which is equivalent to the trivial bound $(x+a)^+ \leq (x-1+a)^+ + 1$) and hence for any stopping time τ, if we plug $a = \xi_1 + \xi_2 + \ldots + \xi_\tau$,

$$\mathbb{E}_x S_\tau^+ - G(x) \leq \mathbb{E}_{x-1} S_\tau^+ - G(x-1).$$

This yields

$$0 < V(x) - G(x) \leq V(x-1) - G(x-1) \tag{3.23}$$

and thus $x-1 \in C$, as we have claimed. This shows that C and D are of the form postulated above and hence, by the general theory,

$$V(x) = \begin{cases} x & \text{if } x \geq b, \\ pV(x+1) + qV(x-1) & \text{if } x < b. \end{cases}$$

Let us first identify V on C. Solving the linear recurrence, we check that

$$V(x) = \alpha + \beta \left(\frac{q}{p}\right)^x, \qquad x < b,$$

for some constants α, $\beta \in \mathbb{R}$. It follows from (3.22) that $V(x) \to 0$ as $x \to -\infty$ (simply use the estimate $\mathbb{E}_x S_\tau^+ \leq \mathbb{E}_x \sup_{n \geq 0} S_n^+$): this implies $\alpha = 0$ and $\beta \geq 0$. This also shows that $b < \infty$. Indeed, otherwise $V(x)$ would explode exponentially as $x \to \infty$, but on the other hand, by (3.23), for $x > 0$ we would have

$$V(x) \leq G(x) + V(0) - G(0) = x + V(0) - G(0).$$

It remains to find β and the boundary b. First, exploiting the Wald-Bellman equation, we see that $V(b-1) = pV(b) + qV(b-2)$. This implies $V(b) = \beta(q/p)^b$ and hence

$$V(x) = b \left(\frac{q}{p}\right)^{x-b} \qquad \text{for } x \leq b.$$

Secondly, again by Wald-Bellman equation, we see that $V(b) \geq pV(b+1) + qV(b-1)$, which is equivalent to $b \geq p/(q-p)$. Finally, observe that if $x > b$, then

$$V(x) = x = px + qx > p(x+1) + q(x-1) = pV(x+1) + qV(x-1).$$

Therefore, if b satisfies the inequality $b \geq p/(q-p)$, then the function

$$\mathcal{V}(x) = \begin{cases} x & \text{if } x \geq b, \\ b\,(q/p)^{x-b} & \text{if } x < b. \end{cases}$$

is excessive. Let us now check for which b the inequality $\mathcal{V} \geq G$ holds. This majorization is clear on $\{b, b+1, b+2, \ldots\}$. Since the function $x \mapsto (q/p)^{x-b}$ is nonnegative, convex and coincides with G at $x = b$, it suffices to check whether it is bigger than G at $x = b-1$. The latter bound is equivalent to $b < q/(q-p) = p/(q-p) + 1$. This actually *forces* us to take $b = \lceil p/(q-p) \rceil$: this is the only choice for the parameter such that the resulting function $\mathcal{V}$ is superharmonic and majorizes G. Summarizing, we have shown that

$$V(x) = \begin{cases} x & \text{if } x \geq \lceil p/(q-p) \rceil, \\ \lceil p/(q-p) \rceil \,(q/p)^{x - \lceil p/(q-p) \rceil} & \text{if } x < \lceil p/(q-p) \rceil. \end{cases}$$

Observe that by (3.22), the stopping time

$$\tau = \inf \left\{ n \,:\, S_n \geq \lceil p/(q-p) \rceil \right\}$$

is infinite with positive probability. Therefore, there is no optimal stopping time τ^* which would be finite $\mathbb{P}_x$-almost surely for all x. Hence, the value function is attained asymptotically at the stopping times

$$\tau^{(M)} = \inf \left\{ n : S_n \notin \left[M, \lceil p/(q-p) \rceil\right] \right\},$$

as $M \to -\infty$.

We proceed to the analysis of further examples, which will be useful in our later considerations.

EXAMPLE 3.13. Let X_1, X_2, ... be a sequence of independent and identically distributed random variables satisfying $\mathbb{P}(X_i = -1) = \frac{2}{3} = 1 - \mathbb{P}(X_i = 2)$, and set $S_0 = 0$, $S_n = X_1 + X_2 + \ldots + X_n$ for $n = 1, 2, \ldots$. We will solve the optimal stopping problem

$$V = \sup_{\tau \in L^1} \mathbb{E}\left(|S_\tau| - \frac{1}{2}\tau\right). \tag{3.24}$$

It is convenient to split the analysis into several separate steps.

Step 1. Dimension reduction. At the first glance, the problem is two-dimensional, i.e., it involves the stopping of the two-dimensional process (S_n, n). It is possible to reduce the dimension to one, by the following simple observation. Namely, note that the processes $(S_n)_{n \geq 0}$ and $(S_n^2 - 2n)_{n \geq 0}$ are martingales. This is evident for the first process, to check the second we compute that

$$\begin{aligned}\mathbb{E}\left[S_{n+1}^2 - 2(n+1)\middle|\mathcal{F}_n\right] &= S_n^2 - 2n + \mathbb{E}(2S_nX_{n+1} + X_{n+1}^2 - 2|\mathcal{F}_n) \\ &= S_n^2 - 2n + \mathbb{E}(X_{n+1}^2 - 2) = S_n^2 - 2n,\end{aligned}$$

since $\mathbb{E}(X_{n+1}|\mathcal{F}_n) = \mathbb{E}X_{n+1} = 0$ and $\mathbb{E}X_{n+1}^2 = 2$. Consequently, by Doob's optional sampling theorem, for any $\tau \in L^1$ we have

$$\mathbb{E}S_{\tau\wedge n}^2 = 2\mathbb{E}(\tau \wedge n). \tag{3.25}$$

The right-hand side is uniformly bounded in n and converges to $\mathbb{E}\tau$, by Lebesgue's monotone convergence theorem. Therefore, the martingale $(S_{\tau\wedge n})_{n\geq 0}$ is bounded in L^2, and hence also L^2-convergent. Thus, letting $n \to \infty$ in (3.25) gives $\mathbb{E}S_\tau^2 = 2\mathbb{E}\tau$ and hence we may rewrite (3.24) as

$$V = \sup_{\tau \in L^1} \mathbb{E}\left(|S_\tau| - \frac{S_\tau^2}{4}\right).$$

Now the right-hand side depends on the process S only.

Step 2. General theory. To put the above problem into the general framework, we set $G(x) = |x| - x^2/4$ and note that the problem reads

$$V = \sup_{\tau \in L^1} \mathbb{E}G(S_\tau).$$

The process S extends to a Markov family on $E = \mathbb{Z}$, with the transity function determined by $p_{n,n-1} = 2/3$ and $p_{n,n+2} = 1/3$ for all $n \in \mathbb{Z}$. Let $V : \mathbb{Z} \to \mathbb{R}$ be given by

$$V(x) = \sup_{\tau \in L^1} \mathbb{E}_x G(S_\tau).$$

In order to apply the general theory, we need to check the condition $\mathbb{E}\sup_{n\geq 0} G(S_n) < \infty$. This is a little technical, so we take the opportunity to present a different approach. Namely, we will construct the least excessive majorant of the function G, and then exploit its properties to solve rigorously the problem under consideration.

Step 3. On the search of the excessive majorants. Let U be the least excessive majorant of G. In our setting, excessiveness amounts to

$$U(n) \geq \frac{2}{3}U(n-1) + \frac{1}{3}U(n+2) \qquad \text{for all } n \in \mathbb{Z}.$$

So, if a function is concave on $\mathbb{Z}$, then it is automatically excessive; as we shall see during the analysis, the reverse implication does not hold in general (which might be a little surprising at the first glance). A quick look at the graph of G shows that the function

$$H(x) = \begin{cases} G(x) & \text{if } |x| \geq 2, \\ 1 & \text{if } |x| \leq 2 \end{cases}$$

is concave and majorizes G. Thus, we have $H \geq U$ on $\mathbb{Z}$. On the other hand, we have $U(x) \geq G(x) = H(x)$ for $|x| \geq 2$: this shows that $U(x) = G(x)$ for $|x| \geq 2$. Next, observe that

$$\begin{aligned} U(1) &\geq \frac{2}{3}U(0) + \frac{1}{3}U(3) = \frac{2}{3}U(0) + \frac{1}{3} \cdot \frac{3}{4}, \\ U(0) &\geq \frac{2}{3}U(-1) + \frac{1}{3}U(2) = \frac{2}{3}U(-1) + \frac{1}{3}, \\ U(-1) &\geq \frac{2}{3}U(-2) + \frac{1}{3}U(1) = \frac{2}{3} + \frac{1}{3}U(1). \end{aligned}$$

Combining these estimates, we get

$$\begin{aligned} U(1) &\geq \frac{2}{3}U(0) + \frac{1}{4} \geq \frac{2}{3}\left(\frac{2}{3}U(-1) + \frac{1}{3}\right) + \frac{1}{4} \\ &\geq \frac{4}{9}\left(\frac{2}{3} + \frac{1}{3}U(1)\right) + \frac{17}{36} = \frac{4}{27}U(1) + \frac{83}{108}, \end{aligned}$$

or $U(1) \geq 83/92$. Plugging this above, we get $U(-1) \geq 267/276$ and $U(0) \geq 45/46$. Assuming equalities, we obtain the excessive function U; furthermore, one verifies directly that such U majorizes G at -1, 0, 1. Hence U is an excessive majorant of G.

Step 4. Coming back to the optimal stopping problem. Suppose that τ is an arbitrary and integrable stopping time. The function U constructed above is excessive, so the process $(U(S_n))_{n\geq 0}$ is a $\mathbb{P}_0$-supermartingale. Furthermore, since $U \geq G$, Doob's optional sampling theorem gives

$$\mathbb{E}G(S_{\tau\wedge n}) \leq \mathbb{E}U(S_{\tau\wedge n}) = \mathbb{E}U(0) = \frac{45}{46},$$

that is,

$$\mathbb{E}\left(|S_{\tau\wedge n}| - \frac{S_{\tau\wedge n}^2}{4}\right) \leq \frac{45}{46}.$$

However, as we have proved in Step 1 above, the process $(S_{\tau\wedge n})_{n\geq 0}$ is an L^2-bounded supermartingale. Thus, letting $n \to \infty$ yields

$$\mathbb{E}\left(|S_\tau| - \frac{1}{2}\tau\right) \leq \frac{45}{46}.$$

Directly from the analysis in Step 3, the equality is attained for the stopping time $\tau = \inf\{n : |S_n| \geq 2\}$.

The next example is more "continuous" in nature.

EXAMPLE 3.14. Suppose that $\xi_1, \xi_2, \ldots$ and $\varepsilon_1, \varepsilon_2, \ldots$ are independent random variables, such that ξ_n has exponential law with parameter $\lambda > 0$ and $\mathbb{P}(\varepsilon_n = 1) = p = 1 - \mathbb{P}(\varepsilon_n = 0)$, $n = 1, 2, \ldots$. For a given parameter $x > 0$, we define the sequence $(X_n)_{n\geq 0}$ by $X_0 \equiv x$ and $X_{n+1} = \varepsilon_{n+1}(X_n + \xi_{n+1})$, $n = 0, 1, 2, \ldots$. Finally, we fix $c > 0$ and consider the optimal stopping problem

$$V = \sup_{\tau} \mathbb{E}(X_\tau - c\tau), \tag{3.26}$$

where the supremum is taken over all integrable stopping times τ.

Step 1. We start by putting the problem into the general framework developed above. It is easy to see that $(X_n)_{n\geq 0}$ is a time-homogeneous Markov process with the transition function given by

$$P(x, 0) = 1 - p \qquad \text{and} \qquad P(x, x + A) = p \int_A \lambda e^{-\lambda a} \mathrm{d}a \quad \text{ for } A \subseteq [0, \infty).$$

Actually, for the problem (3.26) we need to consider the space-time version (n, X_n), which is also a Markov process (this time on the state space $\mathbb{N} \times [0, \infty)$). We extend it to the Markov family and, as usual, denote the corresponding initial probabilities by $\mathbb{P}_{n,x}$. For any (n, x), we consider the auxiliary optimal stopping problem

$$V(n, x) = \sup_{\tau} \mathbb{E}_{n,x} G(\tau, X_\tau),$$

where the supremum is taken over all finite stopping times τ and $G(n, x) = x - cn$.

Step 2. Now we will reduce the dimension of the problem, by observing a certain homogeneity-type condition on V. Namely, the identity $G(n, x) = G(0, x) - cn$ immediately gives

$$V(n, x) = \sup_{\tau} \mathbb{E}_{n,x} G(\tau, X_\tau) = \sup_{\tau} \mathbb{E}_{0,x} G(\tau, X_\tau) - cn = V(0, x) - cn$$

and hence it is enough to identify the function $f(x) := V(0, x)$.

Step 3. We introduce the continuation and the stopping sets C and D by

$$C = \{(n, x) : V(n, x) > G(n, x)\}, \qquad D = \{(n, x) : V(n, x) = G(n, x)\}.$$

Here is some initial analysis of the stopping domain D. Namely, we will show that if $(0, x) \in D$ and $x' > x$, then necessarily $(0, x') \in D$. To this end, note that for any stopping time τ, we have

$$\mathbb{E}_{0,x'}(X_\tau - c\tau) - \mathbb{E}_{0,x}(X_\tau - c\tau) \leq x' - x.$$

Indeed, one can couple the trajectories of X under $\mathbb{P}_{(0,x)}$ and $\mathbb{P}_{(0,x')}$ in such a way that their difference is equal to $x - x'$, until they both drop to zero and coincide from that time. Since $x' - x = G(0, x') - G(0, x)$, we obtain

$$V(0, x') - G(0, x') \leq V(0, x) - G(0, x) = 0,$$

which gives $(0, x') \in D$. By the homogeneity established in Step 2, this gives the aforementioned property of the set D. Actually, we see that D must be of the form

$$D = \{(n, x) : x \geq b\},$$

for some unknown parameter b (to be found).

Step 4. It follows from the general theory that V is the least excessive majorant of the function G. Here the excessiveness means

$$V(n, x) \geq (1 - p)V(n + 1, 0) + p \int_0^\infty V(n + 1, x + a) \cdot \lambda e^{-\lambda a} \mathrm{d}a,$$

or, by the homogeneity proved in Step 2 above,

$$f(x) - cn \geq (1-p)(f(0) - c(n+1)) + p \int_0^\infty (f(x+a) - c(n+1)) \cdot \lambda e^{-\lambda a} \mathrm{d}a.$$

This is equivalent to the inequality

$$f(x) + c \geq (1-p)f(0) + p \int_0^\infty f(x+a) \cdot \lambda e^{-\lambda a} \mathrm{d}a. \tag{3.27}$$

This allows us to write the corresponding system of requirements: the so-called boundary value problem for V (or rather, for f). Namely,

$$f(x) = x \qquad \text{for } x \geq b, \tag{3.28}$$

$$f(x) + c = (1-p)f(0) + p \int_0^\infty f(x+a) \cdot \lambda e^{-\lambda a} \mathrm{d}a \qquad \text{for } x < b. \tag{3.29}$$

Step 5. Let us try to find the candidate for the solution to the above system. A direct differentiation of (3.29) yields

$$f'(x) = p \int_0^\infty f'(x+a) \cdot \lambda e^{-\lambda a} \mathrm{d}a = -p\lambda f(x) + p\lambda \int_0^\infty f(x+a) \cdot \lambda e^{-\lambda a} \mathrm{d}a,$$

where the last passage follows from the integration by parts. We apply (3.29) again, obtaining

$$f'(x) = -p\lambda f(x) + \lambda(f(x) + c - (1-p)f(0)) = \lambda(1-p)\Big[f(x) + \frac{c}{1-p} - f(0)\Big].$$

This can be solved directly: we get

$$f(x) = f(0) - \frac{c}{1-p} + \alpha e^{\lambda(1-p)x} \qquad x < b,$$

for some parameter α. Plugging $x = 0$, we get $\alpha = c/(1-p)$ and hence

$$f(x) = \begin{cases} f(0) + \frac{c}{1-p}\left(e^{\lambda(1-p)x} - 1\right) & \text{if } x < b, \\ x & \text{if } x \geq b. \end{cases}$$

It is plausible to conjecture that f is continuous at b: this implies

$$f(0) = b - \frac{c}{1-p}\left(e^{\lambda(1-p)b} - 1\right)$$

and we finally obtain

$$f(x) = \begin{cases} b + \frac{c}{1-p}\left(e^{\lambda(1-p)x} - e^{\lambda(1-p)b}\right) & \text{if } x < b, \\ x & \text{if } x \geq b. \end{cases}$$

Step 6. It remains to find the boundary point b. To accomplish this, we verify the excessiveness inequality (3.27) on $[b, \infty)$. Since $f(x) = x$ on this interval, the inequality reads

$$x + c \geq (1-p)\left(b - \frac{c}{1-p}\left(e^{\lambda(1-p)b} - 1\right)\right) + p\left(x + \frac{1}{\lambda}\right).$$

This requirement is most restrictive for smallest x, i.e., for $x = b$. For this particular choice, the estimate becomes

$$\exp(\lambda(1-p)b) \geq \frac{p}{\lambda c}. \tag{3.30}$$

If $p/(\lambda c) \le 1$, then this estimate holds for all $b \ge 0$. In other words, the identity function $f(x) = x$ leads to the excessive function $V(n, x) = x - cn$. In this case, we have $V = G$ and the optimal stopping rule is to stop instantaneously.

If $p/(\lambda c) > 1$, then assuming equality in (3.30), we obtain

$$b = \frac{1}{\lambda(1-p)} \ln \frac{p}{\lambda c}$$

and

$$f(x) = \begin{cases} \frac{1}{\lambda(1-p)} \ln \frac{p}{\lambda c} + \frac{c}{1-p} e^{\lambda(1-p)x} - \frac{p}{1-p} & \text{if } x < b, \\ x & \text{if } x \ge b. \end{cases}$$

Step 7. We need to emphasize that the analysis in Step 6 was informal: we obtained the *candidate* for V (or rather, for f) under a number of additional assumptions. It remains to check that V satisfies all the necessary requirements. Namely, it actually follows from the above analysis that V is excessive. The majorization $V \ge G$ is equivalent to the estimate $f(x) \ge x$ on $[0, b]$; since both sides are equal for $x = b$ and f is strictly convex on $[0, b]$, it is enough to verify that $f'(b-) \le 1$. This is equivalent to the estimate

$$c\lambda \cdot \frac{p}{\lambda c} \le 1,$$

which holds trivially. Finally, let us show that V is the smallest excessive majorant. Assume conversely that this is not the case: the optimal function $\tilde{V}$ satisfies $\tilde{V}(n, x) < V(n, x)$ for some (n, x). Then $\tilde{V}$ satisfies the identity $\tilde{V}(n, y) = \tilde{f}(y) - cn$ for some function $\tilde{f}$ on $[0, \infty)$, and we have $\tilde{f}(x) < f(x)$. We have $\tilde{f}(y) \ge y$ for all y, and hence we must necessarily have $x < b$ By the excessiveness condition (3.27), we get

$$\tilde{f}(x) - f(x) \ge p \int_0^\infty (\tilde{f}(x+a) - f(x+a)) \cdot \lambda e^{-\lambda a} \mathrm{d}a.$$

Now suppose that $x_0 = \sup\{t : \tilde{f}(t) - f(t) \le \tilde{f}(x) - f(x)\}$: clearly x_0 is finite, since $\tilde{f}$ and f coincide on $[b, \infty)$. Then $\tilde{f}(u) - f(u) > \tilde{f}(x) - f(x)$ on (x_0, ∞) and hence we obtain

$$\tilde{f}(x) - f(x) \ge p \int_0^\infty (\tilde{f}(x_0+a) - f(x_0+a)) \cdot \lambda e^{-\lambda a} \mathrm{d}a \ge p(\tilde{f}(x) - f(x)),$$

which implies $\tilde{f}(x) > f(x)$, a contradiction.

We conclude the analysis with the comment that the optimal stopping rule is given by $\tau_* = \inf\{n : X_n \ge b\}$.

3. Problems

1. Let G_1, G_2, ... be a sequence of independent random variables, each of which has the uniform distribution on $[0, 1]$. Solve the optimal stopping problems

$$V^N = \sup_{\tau \in \mathcal{M}^N} \mathbb{E} G_\tau \qquad \text{and} \qquad V_0 = \sup_{\tau \in \mathcal{M}} \mathbb{E} G_\tau,$$

where N is an arbitrary integer.

2. We flip a coin at most five times, at each point we may decide whether to stop or not (in particular, we are allowed to stop at the very beginning, without flipping the coin even once). Having stopped, we look at the outcomes we have obtained. We get 1 if there are no heads and get 2 if we obtained at least three heads. What is the strategy which yields the largest expected gain?

3. Solve the optimal stopping problem

$$\sup_{0\leq\tau\leq N} \mathbb{E}\big(Y_\tau^2 - \tau\big),$$

where $Y_0 = 10$, $Y_{n+1} = Y_n + \xi_{n+1}$, and

$$\mathbb{P}(\xi_n = 1) = 1/(4n) = 1 - \mathbb{P}(\xi_n = 0), \qquad n = 1,\, 2,\, \ldots,\, N.$$

4. Solve the secretary problem.

5. Solve the optimal stopping problem $V = \sup_{\tau\in\mathcal{M}} \mathbb{E}\big(\alpha^\tau X_\tau\big)$, where $\alpha \in (0,1)$, $X_0 = 0$,

$$X_{n+1} = (X_n + 1)V_{n+1}, \qquad n = 0,\, 1,\, 2,\, \ldots,$$

and V_1, V_2, ... is a sequence of independent random variables with the distribution $\mathbb{P}(V_j = 0) = \mathbb{P}(V_j = 1) = 1/2$.

6. We flip a coin infinitely many times. For $n \geq 1$, let $G_n = n2^n/(n+1)$ of there were no tails in the first n flips, and $G_n = 0$ otherwise. Solve the optimal stopping problem

$$V = \sup_{\tau\in\mathcal{M}} \mathbb{E}G_\tau.$$

7. Suppose that $(G_n)_{n\geq 0}$ is an adapted sequence of random variables satisfying $\mathbb{E}\sup_{n\geq 0}|G_n| < \infty$ and let $(B_n)_{n\geq 0}$ be the associated Snell envelope. Prove the identity

$$\sup_{\tau\in\mathcal{M}} \mathbb{E}G_\tau = \sup_{\tau\in\mathcal{M}} \mathbb{E}B_\tau.$$

8. We toss a fair coin and for each $n \geq 0$, we denote by X_n the length of the current sequence of consecutive tails after n flips:

$$\overbrace{\ldots H\underbrace{TT\ldots T}_{X_n}}^{n \text{ flips}}.$$

Solve the optimal stopping problem

$$V = \sup_{\tau\in\mathcal{M}} \mathbb{E}\left(X_\tau - \frac{1}{16}\tau\right).$$

9. A pawn moves over the set $\{1, 2, \ldots, n\}$, according to the following rules. If at some time it is located at the point k, then at the next step it jumps, independently from its evolution in the past, to one of the points k, $k+1$, ..., n (each choice has the same probability). Let X_j be the location of the pawn at the time j. Assuming that $X_0 = 1$, describe the stopping time τ which maximizes the expectation $\mathbb{E}G(X_\tau)$, where $G(x) = x1_{\{x<n\}}$.

10. Let ε_1, ε_2, ... be the sequence of independent Rademacher variables and set $S_0 \equiv 0$ and $S_n = \varepsilon_1 + \varepsilon_2 + \ldots + \varepsilon_n$ for $n = 1$, 2, ... Find the smallest constant C such that for any stopping time τ adapted to the natural filtration of X we have

$$\mathbb{E}S_\tau^4 \leq C\mathbb{E}\tau^2.$$

CHAPTER 4

Deterministic continuous case

Now we turn our attention to the case in which the observed process is indexed by nonnegative numbers; then the role of the evolution is played by an appropriate differential equation.

We start with the necessary background. Assume that the state of the system is described by a vector in $\mathbb{R}^n$, and hence the observed process is $x = (x(t))_{t\geq 0} \subset \mathbb{R}^n$. Next, we will assume that the controls are given as a function $u = (u(t))_{t\geq 0}$ with values in $\mathbb{R}^m$ (or more generally, in some fixed subset U of $\mathbb{R}^m$): this function will be assumed to be piecewise continuous. The evolution equation is given by

$$\dot{x} = f(t, x(t), u(t)),$$

with the initial condition $x(0) = x_0$. Here $f : \mathbb{R}^{n+m+1} \to \mathbb{R}^n$ is a continuous function. As in the discrete-time case, we may consider the more general situation in which the starting time is an arbitrary number $s \geq 0$ and the initial position is equal to y. Sometimes, to emphasize the fact that x is obtained via the control u and starts at s from y, we will use the notation $x^u(\cdot; s, y)$.

We are ready to formulate the optimal control problem. Let $T > 0$ be a fixed parameter (the horizon) and let $q : \mathbb{R}^{n+m+1} \to \mathbb{R}$, $r : \mathbb{R}^n \to \mathbb{R}$ be two given functions. Suppose we are interested in the supremum

$$B := \sup_u J(x_0, u), \tag{4.1}$$

where the functional J is given by

$$J(x_0, u) = \int_0^T q(s, x^u(s), u(s))\mathrm{d}s + r(x(T)).$$

As usual, two problems arise:

· Find the value of (4.1).

· find the optimal control u^* maximizing (4.1).

As previously, an effective way of handling the above questions is to extend the problem to the case of arbitrary initial positions. That is, we assume that the starting time 0 is replaced by t and the initial position is y, and set

$$B(t, y) = \sup_u \left\{ \int_t^T q(s, x^u(s; t, y), u(s))\mathrm{d}s + r(x(T; t, y)) \right\},$$

where the supremum is taken over all controls u on $[t, T]$. Then we have the following two facts.

Lemma 4.1. *Fix (t, y) and an arbitrary control u on $[t, T]$. Then the function $s \mapsto B(s, x^u(s; t, y)) + \int_t^s q(w, x^u(w; t, y), u(w))dw$ is nonincreasing on $[t, T]$.*

PROOF. Pick arbitrary two points w, $w' \in [t,T]$ satisfying $w < w'$ and let $\varepsilon > 0$. Furthermore, set $z = x^u(w';t,y)$. Then, by the very definition of $B(w',z)$, there is a control $\tilde{u} : [w',T] \to \mathbb{R}^k$ for which

$$B(w',z) < \int_{w'}^{T} q(s, x^{\tilde{u}}(s;w',z), \tilde{u}(s))\mathrm{d}s + r(x^{\tilde{u}}(T;w',z)) + \varepsilon.$$

Let us modify the control u given in the statement according to $\tilde{u}$: set

$$\hat{u}(s) = \begin{cases} u(s) & \text{if } t \le s < w', \\ \tilde{u}(s) & \text{if } w' \le s \le T. \end{cases}$$

Treating $(w, x^u(w))$ as the initial position, we get, again by the definition of B,

$$\begin{aligned} B(w, x^u(w)) &\ge \int_w^T q(s, x^{\hat{u}}(s;w,x^u(w)), u(s))\mathrm{d}s + r(x^{\hat{u}}(T;w,x^u(w))) \\ &\ge \int_w^{w'} q(s, x^{\hat{u}}(s;w,x^u(w)), u(s))\mathrm{d}s + B(w',z) - \varepsilon \\ &= \int_w^{w'} q(s, x^{u}(s;w,x^u(w)), u(s))\mathrm{d}s + B(w',x^u(w')) - \varepsilon. \end{aligned}$$

Since ε was arbitrary, the claim follows. □

LEMMA 4.2. *The function $t \mapsto B(t, x^*(t)) + \int_0^t q(s, x^*(s), u^*(s))ds$ is constant.*

PROOF. By the previous lemma, we have

$$\begin{aligned} B(0, x_0) &\ge B(t, x^*(t)) + \int_0^t q(s, x^*(s), u^*(s))\mathrm{d}s \\ &\ge B(T, x^*(T)) + \int_0^T q(s, x^*(s), u^*(s))\mathrm{d}s \\ &= r(x^*(T)) + \int_0^T q(s, x^*(s), u^*(s))\mathrm{d}s. \end{aligned}$$

It remains to note that the first and the last expressions are equal, because of the optimality of x^*. Hence equalities hold throughout above. □

Observe that both lemmas remain valid if we assume that the controls belong to some specific subsets of $\mathbb{R}^k$; the only condition we need is that the „modification procedure" $(u, \tilde{u}) \mapsto \hat{u}$ exploited above does not lead outside the class of permitted controls. It is easy to check that the statement below also remain valid.

We note an important consequence, which follows from the two lemmas above by a direct differentiation. In what follows, B_y is the gradient of B with respect to the variables y_1, y_2, …, y_n.

COROLLARY 4.3. *Suppose that the value function B is of class C^1. Then we have*

$$B_s(s,y) + B_y(s,y) \cdot f(s,y,v) + q(s,y,v) \le 0 \qquad \textit{for all } v \in U. \tag{4.2}$$

Furthermore, if x^*, u^* is an optimal pair for (s,y), then B satisfies the Hamilton-Jacobi-Bellman equation

$$B_s(s, x^*(s)) + B_y(s, x^*(s)) \cdot f(s, x^*(s), u^*(s+)) + q(s, x^*(s), u^*(s+)) = 0.$$

We also record here the following statement, known as the maximum principle. Here f_x is the matrix of dimension $n \times n$, whose columns are the partial derivatives of f with respect to x_1, x_2, ..., x_n.

THEOREM 4.4. *Consider the above optimal control problem and suppose that the value function B is of class C^2, q is of class C^1. Then there exists a C^1 function $p : (0, \infty) \to \mathbb{R}^n$, such that for almost all $t > 0$ (except for a finite number of points),*
(i) $u^*(t)$ maximizes $u \mapsto p(t) \cdot f(t, x^*(t), u) + q(t, x^*(t), u)$;
(ii) we have $\dot{p}(t) = -p(t) \cdot f_x(t, x^*(t), u^*(t)) - q_x(t, x^*(t), u^*(t))$.

PROOF. Let $p(s) = B_y(s, x^*(s))$: note that this is a vector-valued, C^1 function. Then (i) follows at once from the previous corollary. To show the second part, fix $t \in [0, T]$ and apply (4.2) to obtain

$$B_s(t, y) + B_y(t, y) \cdot f(t, y, u^*(t+)) + q(t, y, u^*(t+)) \le 0$$

for all y, with equality for $y = x^*(t)$. Consequently, differentiating with respect to y at $x^*(t)$, we get

$$\begin{aligned} B_{sy}(t, x^*(t)) + B_{yy}(t, x^*(t)) f(t, x^*(t), u^*(t+)) + B_y(t, x^*(t)) f_x(t, x^*(t), u^*(t+)) \\ + q_y(t, y, u^*(t)) = 0. \end{aligned}$$

It remains to note that $p(t) \cdot f(t, x^*(t), u^*(t)) = B_y(t, x^*(t)) \cdot f_x(t, x^*(t), u^*(t+))$ and

$$\begin{aligned} \dot{p}(t) &= B_{sy}(t, x^*(t)) + B_{yy}(t, x^*(t)) \dot{x^*}(t) \\ &= B_{sy}(t, x^*(t)) + B_{yy}(t, x^*(t)) f(t, x^*(t), u^*(t)) \end{aligned}$$

for almost all t. □

REMARK 4.5. There is a nice interpretation of (ii). Consider the associated Hamiltonian $H : [0, \infty) \times \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^n \to \mathbb{R}$, given by

$$H(t, x, u, p) = p \cdot f(t, x, u) + q(t, x, u).$$

Then the optimal solution satisfies $\dot{x}^* = H_p(t, x^*, u^*, p)$ and $\dot{p} = -H_x(t, x^*, u^*, p)$.

EXAMPLE 4.6. Suppose that a particle with given initial position and velocity $x_1(0)$, $x_2(0)$ is to be brought to rest at position 0 in minimal time. This is to be done using the control force u satisfying $|u| \le 1$, with dynamics $\dot{x}_1 = x_2$ and $\dot{x}_2 = u$. That is, in the matrix form, we have

$$\dot{x}(t) = \frac{d}{dt} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} x_2 \\ u \end{pmatrix} = f(t, x, u)$$

and we want to maximize the functional

$$J(x_0, u) = -\int_0^T 1 \mathrm{d}s,$$

where T is the first time the function x reaches $(0, 0)$. The problem is to identify T and the optimal control.

In the light of the above considerations, the Hamiltonian equals

$$H(t, x, u, p) = p \cdot (x_2, u) - 1 = p_1 x_2 + p_2 u - 1.$$

By Theorem 4.4, the optimal control $u^*(t)$ maximizes $u \mapsto H(t, x^*(t), u, p(t))$, so $u = \operatorname{sgn}(p_2(t))$. By the second part of this theorem, we get $\dot{p}_1 = 0$, $\dot{p}_2 = -p_1$. This implies that $p_1 \equiv \alpha$ and $p_2(t) = -p_1 t +$const; let us write $p_2(t) = -\alpha(t - T) + \beta$.

Therefore, the solution is of the following form: there is at most one change of sign of p_2 on the optimal path; u is maximal in one direction and then possibly maximal in the other. Adding/subtracting constant if necessary, we may and assume that $H = 0$; then computing H at time T (and recalling that $x_2(T) = 0$), we obtain $|\beta| = 1$. Suppose that $\beta = 1$ (the case $\beta = -1$ is considered similarly).

· If $\alpha \geq -1/T$, then $p_2 > 0$ for all t, so $u_2 = 1$, $x_2(t) = t - T$ and $x_1(t) = (t-T)^2/2$. We have $x_1(0) = T^2/2$ and $x_2(0) = -T$, so (x_1, x_2) moves along the „left half" of the parabola $x = y^2$.

· If $\alpha < -1/T$, then $p_2 < 0$ for $t < T + 1/\alpha$ and $p_2 > 0$ for $t > T + 1/\alpha$, so the optimal control is

$$u^*(t) = \begin{cases} -1 & \text{for } [0, T+1/\alpha), \\ 1 & \text{for } [T+1/\alpha, T]. \end{cases}$$

Now, since $\dot{x}_2 = u$ and $x_2(T) = 0$, we check that

$$x_2(t) = \begin{cases} -t + T + 2/\alpha & \text{for } [0, T+1/\alpha), \\ t - T & \text{for } [T+1/\alpha, T], \end{cases}$$

and hence

$$x_1(t) = \begin{cases} -(-t + T + 2/\alpha)^2/2 + 1/\alpha^2 & \text{for } [0, T+1/\alpha), \\ (t-T)^2/2 & \text{for } [T+1/\alpha, T]. \end{cases}$$

See Figure 4.6.

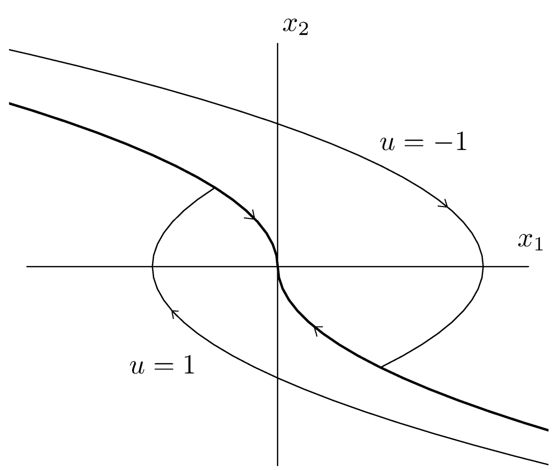

FIGURE 1. The graph of the function $x_1 = -x_2|x_2|/2$ (bold) splits the phase portrait into two parts. The upper region corresponds to the choice $u = -1$, while the lower part corresponds to $u = 1$.

The above reasoning might seem a little informal, so we will present now a rigorous analysis of the problem, basing on the above calculations. We have $x_1(0) = -(x_2(0))^2/2 + 1/\alpha^2$, which allows us to compute α and T. We have

$$\frac{1}{\alpha} = -\sqrt{x_1(0) + \frac{x_2^2(0)}{2}}, \qquad T = x_2(0) + 2\sqrt{x_1(0) + \frac{x_2^2(0)}{2}}.$$

The same computations would apply if the starting time was equal to t and the initial position was $x_1 = x_1(t)$, $x_2 = x_2(t)$ (simply replace 0 with t and T with $T - t$). Therefore, we are led to the following candidate for the Bellman function:

$$\tilde{B}(t, x_1, x_2) = \begin{cases} -x_2 - 2\sqrt{x_1 + x_2^2/2} & \text{in the upper region,} \\ x_2 - 2\sqrt{-x_1 + x_2^2/2} & \text{in the lower region.} \end{cases}$$

Now we check that "the Bellman property": we fix arbitrary $x_1(0)$, $x_2(0)$ and some control u, and prove that the function $\zeta(t) = \tilde{B}(t, x_1, x_2) - t$ is nonincreasing on $[0, T]$, where T is the first time (x_1, x_2) reaches $(0, 0)$. This is straightforward: for example, in the upper region, we compute that

$$\zeta'(t) = -\dot{x}_2 - 1 - \frac{\dot{x}_1 + x_2\dot{x}_2}{\sqrt{x_1 + x_2^2/2}} = -(u + 1)\left(1 + \frac{x_2}{\sqrt{x_1 + x_2^2/2}}\right).$$

Since $u \in [-1, 1]$, this is nonpositive. Indeed, if $x_2 \geq 0$ there is nothing to prove, and if $x_2 < 0$, then we use the fact that $x_1 > x_2^2/2$. Consequently, we have $\zeta(T) \leq \zeta(0)$, which is equivalent to $-T \leq \tilde{B}(0, x_1(0), x_2(0))$. Equality is obtained for the examples studied above. Actually, the same analysis can be carried out if the starting position is (t, x_1, x_2) instead of $(0, x_1(0), x_2(0))$: then we obtain $B(t, x_1, x_2) \leq \tilde{B}(t, x_1, x_2)$, and the reverse follows from the fact that $\tilde{B}$ is based on the concrete examples.

Now we will show how the optimal control theory leads to Hardy inequalities.

EXAMPLE 4.7. Let $1 < \alpha < \infty$ be a fixed constant. We will identify the best constant C_α in the estimate

$$\int_0^\infty \left|\frac{1}{t}\int_0^t f(s)\mathrm{d}s\right|^\alpha \mathrm{d}t \leq C_\alpha \int_0^\infty |f(t)|^\alpha \mathrm{d}t.$$

Note that we may restrict ourselves to nonnegative f: the passage from f to $|f|$ does not change the right-hand side and may only increase the left (making the claim harder). By standard density arguments, we may and do assume that f is piecewise continuous. In addition, by a straightforward limiting argument, it is enough to study the localized estimate

$$\int_0^T \left|\frac{1}{t}\int_0^t f(s)\mathrm{d}s\right|^\alpha \mathrm{d}t \leq C_\alpha \int_0^T |f(t)|^\alpha \mathrm{d}t.$$

Step 1. Abstract Bellman function. Let us first put the problem into an appropriate framework. We rewrite the estimate in the form

$$\sup\left\{\int_0^T \left(\frac{1}{t}\int_0^t f\right)^\alpha - C_\alpha f^\alpha(t)\mathrm{d}t\right\} \leq 0,$$

the supremum taken over all f as above. This suggests that the control should be the function f; furthermore, looking under the integral, it seems natural to take $\int_0^t f$ as the controlled process. Thus, we are ready to introduce the associated Bellman function

$$B(s, y) = \sup\left\{\int_0^s \left(\left(\frac{1}{t}\int_0^t f\right)^\alpha - Cf^\alpha(t)\right)\mathrm{d}t \ : \ \int_0^s f = y\right\},$$

where $C > 0$ is a fixed constant. Our goal is to identify the least choice of C for which $B \le 0$.

Step 2. Homogeneity. The function B depends on two variables; however, as we will see in a moment, it satisfies two homogeity-type conditions, which actually reduces the problem of finding B to a single real number. First, suppose that f is an arbitrary function on $[0, s]$ satisfying $\int_0^s f = y$. Then obviously $\int_0^s \lambda f = \lambda y$ for any constant $\lambda > 0$ and thus

$$B(s, \lambda y) \ge \lambda^\alpha \int_0^s \left(\left(\frac{1}{t} \int_0^t f \right)^\alpha - C f^\alpha(t) \right) \mathrm{d}t.$$

Taking the supremum over all f as above, we get

$$B(s, \lambda y) \ge \lambda^\alpha B(s, y), \qquad \lambda > 0. \tag{4.3}$$

Now plug $y := \lambda^{-1} y$ to obtain

$$B(s, y) \ge \lambda^\alpha B(s, \lambda^{-1} y).$$

Substituting $\lambda := \lambda^{-1}$, we obtain the reverse to (4.3): thus, equality holds here.

The second homogeneity is a little more involved. Namely, suppose again that f is an arbitrary function on $[0, s]$ satisfying $\int_0^s f = y$. Take $\mu > 0$ and consider the dilated function $\tilde{f}(t) = f(t/\mu)$ on $[0, \mu s]$. Then

$$\int_0^{\mu s} \tilde{f} \mathrm{d}t = \int_0^{\mu s} f(t/\mu) \mathrm{d}t = \mu \int_0^s f = \mu y,$$

so that

$$B(\mu s, \mu y) \ge \int_0^{\mu s} \left(\left(\frac{1}{t} \int_0^t \tilde{f} \right)^\alpha - C \tilde{f}^\alpha(t) \right) \mathrm{d}t = \mu \int_0^s \left(\left(\frac{1}{t} \int_0^t f \right)^\alpha - C f^\alpha(t) \right) \mathrm{d}t.$$

Taking the supremum over all f as above, we get

$$B(\mu s, \mu y) \ge \mu B(s, y),$$

and substituting $s := \mu^{-1} s$, $y := \mu^{-1} y$ and then $\mu := \mu^{-1}$ shows that we actually have equality here. Putting all these facts together, we see that

$$B(s, y) = s B(1, y/s) = s \cdot (y/s)^\alpha B(1, 1) = -\kappa y^\alpha s^{1-\alpha}, \tag{4.4}$$

where $\kappa = -B(1, 1)$. In particular, we see that B is of class C^∞ (provided it is finite).

Step 3. Bellman equation. Now we will study the Bellman monotonicity condition. To this end, fix $s > 0$, $\varepsilon > 0$ and any nonnegative, piecewise continuous function f on $[0, s]$ with $\int_0^s f = y$. We extend this function to $\tilde{f} : [0, s + \varepsilon] \to [0, \infty)$ putting $\tilde{f}(t) = a$ for $t \in (s, s + \varepsilon]$; here a is some arbitrary nonnegative number ("control"). Then we have

$$\int_0^{s+\varepsilon} \tilde{f} = y + \varepsilon a$$

and hence

$$\begin{aligned}
&B(s+\varepsilon, y+\varepsilon a)\\
&\geq \int_0^{s+\varepsilon}\left(\left(\frac{1}{t}\int_0^t \tilde{f}\right)^\alpha - C\tilde{f}^\alpha(t)\right)\mathrm{d}t\\
&= \int_0^{s}\left(\left(\frac{1}{t}\int_0^t \tilde{f}\right)^\alpha - C\tilde{f}^\alpha(t)\right)\mathrm{d}t + \int_s^{s+\varepsilon}\left(\left(\frac{1}{t}\int_0^t \tilde{f}\right)^\alpha - C\tilde{f}^\alpha(t)\right)\mathrm{d}t\\
&= \int_0^{s}\left(\left(\frac{1}{t}\int_0^t f\right)^\alpha - Cf^\alpha(t)\right)\mathrm{d}t + \int_s^{s+\varepsilon}\left(\frac{1}{t^\alpha}\left(\int_0^s f + a(t-s)\right)^\alpha - Ca^\alpha\right)\mathrm{d}t\\
&= \int_0^{s}\left(\left(\frac{1}{t}\int_0^t f\right)^\alpha - Cf^\alpha(t)\right)\mathrm{d}t + \int_s^{s+\varepsilon}\left(\frac{1}{t^\alpha}\left(y + a(t-s)\right)^\alpha - Ca^\alpha\right)\mathrm{d}t
\end{aligned}$$

Taking the supremum over all f as above, we get

$$B(s+\varepsilon, y+\varepsilon a) \geq B(s,y) + \int_s^{s+\varepsilon}\left(\frac{1}{t^\alpha}\left(y + a(t-s)\right)^\alpha - Ca^\alpha\right)\mathrm{d}t.$$

Now move $B(s,y)$ to the left, divide throughout by ε and let $\varepsilon \to 0$, obtaining

$$B_s(s,y) + aB_y(s,y) \geq (y/s)^\alpha - Ca^\alpha. \tag{4.5}$$

By (4.4), this simplifies to

$$-\kappa(1-\alpha)(y/s)^\alpha - \kappa a\alpha(y/s)^{\alpha-1} \geq (y/s)^\alpha - Ca^\alpha,$$

or

$$Cb^\alpha - \alpha b\kappa + \kappa(\alpha-1) - 1 \geq 0,$$

where $b = as/y$. This estimate is supposed to hold for any $b \geq 0$; optimizing over this parameter, we see that the left-hand side is minimized for $b = (\kappa/C)^{1/(\alpha-1)}$. Assuming equality for this choice of b, we obtain

$$-\frac{\kappa^{\alpha/(\alpha-1)}}{C^{1/(\alpha-1)}} + \kappa = \frac{1}{\alpha-1}. \tag{4.6}$$

This leads us to the following candidate: if C is a fixed parameter, then we let κ be the solution to the above equation and then we apply (4.4).

Step 4. Choice for C. The above analysis leads to the following question: for which C there is a solution to (4.6)? The left-hand side converges to $-\infty$ as $\kappa \to \infty$, so we actually ask for those C, for which the maximum of the left-hand side is at least $1/(\alpha-1)$. A direct differentiation shows that the maximum is attained for $\kappa = C \cdot \left(\frac{\alpha-1}{\alpha}\right)^{\alpha-1}$, and it is equal to

$$C\left(\frac{\alpha-1}{\alpha}\right)^{\alpha-1}\cdot\frac{1}{\alpha}.$$

In other words, the solution to (4.6) exists if and only if $C \geq \left(\frac{\alpha}{\alpha-1}\right)^\alpha$. Now we guess that the optimal constant is $C_\alpha = \left(\frac{\alpha}{\alpha-1}\right)^\alpha$. Then we are forced to take $\kappa = \alpha/(\alpha-1)$. Reviewing all the above calculations, we see that the function

$$\tilde{B}(s,y) = -\frac{\alpha}{\alpha-1}y^\alpha s^{1-\alpha}$$

satisfies (4.5), with equality for

$$a = by/s = \frac{(\alpha-1)y}{s\alpha}. \tag{4.7}$$

Step 5. Verification. Now we pick an arbitrary nonnegative, piecewise continuous function f on $[0, s]$ satisfying $\int_0^s f = y$. By a direct differentiation, the function

$$u \mapsto \tilde{B}\left(u, \int_0^u f\right) + \int_u^s \left(\left(\frac{1}{t}\int_0^t f\right)^\alpha - Cf^\alpha(t)\right)\mathrm{d}t \tag{4.8}$$

is nondecreasing: this is precisely (4.5), with $s = u$, $y = \int_0^u f$ and $a = f(u)$. Thus

$$\tilde{B}\left(u, \int_0^u f\right) + \int_u^s \left(\left(\frac{1}{t}\int_0^t f\right)^\alpha - C_\alpha f^\alpha(t)\right)\mathrm{d}t \leq \tilde{B}\left(s, \int_0^s f\right) \leq 0$$

and letting $u \to 0$ we get

$$\int_0^s \left(\left(\frac{1}{t}\int_0^t f\right)^\alpha - C_\alpha f^\alpha(t)\right)\mathrm{d}t \leq 0,$$

so the inequality holds with the constant C_α. To see that this constant is optimal, we inspect (4.7) and (4.8): the function in (4.8) will be constant if f satisfies

$$f(u) = \frac{\alpha-1}{\alpha}\cdot\frac{1}{u}\int_0^u f$$

for all u. This happens for $f(u) = cu^{-1/\alpha}$ (c is some constant), but this function does not have the appropriate integrability: $\int_0^s f^\alpha = \infty$. Anyhow, we might expect that the optimality of C_α will be obtained asymptotically. Take $\beta > \alpha^{-1}$ and consider the function $f(u) = u^{-\beta}$. Then $f \in L^\alpha(0, s)$, we have $\frac{1}{t}\int_0^t f = (1-\beta)^{-1}f$, and

$$\frac{\int_0^s \left(\frac{1}{t}\int_0^t f\right)^\alpha \mathrm{d}t}{\int_0^s f^\alpha \mathrm{d}t} = (1-\beta)^{-\alpha} \to C_\alpha,$$

as $\beta \to \alpha^{-1}$. This proves the desired sharpness.

EXAMPLE 4.8. Now we will present an enhanced analysis of the estimate

$$\int_0^\infty \left|\frac{1}{t}\int_0^t f(s)\mathrm{d}s\right|^\alpha \mathrm{d}t \leq C_\alpha \int_0^\infty |f(t)|^\alpha \mathrm{d}t,$$

which is much more flexible. As previously, we start with noting that we may assume that f is nonnegative: the passage from f to $|f|$ does not change the right-hand side and may only increase the left (making the claim harder). We will restrict ourselves to piecewise continuous f.

Step 1. An abstract Bellman function. The first step of the analysis is to investigate the more general estimate

$$\int_0^T \left(\frac{1}{t}\int_0^t f(s)\mathrm{d}s\right)^\alpha \mathrm{d}t \leq C_\alpha \int_0^T f(t)^\alpha \mathrm{d}t.$$

Instead of formal putting this problem into the appropriate framework, we prefer to proceed with the direct analysis. For a given f as above, consider the functions

$$x_1(t) = \int_0^t f(s)\mathrm{d}s, \qquad x_2(t) = \int_0^t f(s)^\alpha \mathrm{d}s, \qquad t \in [0,T],$$

and let

$$B(t, y_1, y_2) = \sup\left\{ \int_0^t s^{-\alpha} x_1^\alpha(s)\mathrm{d}s \,:\, x_1(t) = y_1,\, x_2(t) = y_2 \right\}.$$

Note that $y_1 \leq y_2^{1/\alpha} t^{1-1/\alpha}$: B is defined on a non-trivial subdomain of $\mathbb{R}^3_+$. Actually, if $y_1 = y_2^{1/\alpha} t^{1-1/\alpha}$, then there is only one function satisfying the integral conditions: $f \equiv y_1/t$. Then $s \mapsto \frac{1}{s}\int_0^s f \equiv y_1/t$ and therefore, $B(t, y_1, y_2) = y_1^\alpha t^{1-\alpha}$.

We will find the explicit formula for B.

Step 2. Homogeneity. Now we will present the dimension reduction: the function B depends on three variables, but enjoys certain homogeneity, which allows for a significant simplification. First, suppose that f is an arbitrary function on $[0,t]$ satisfying $\int_0^t f = y_1$, $\int_0^t f^\alpha = y_2$. Then obviously $\int_0^t \lambda f = \lambda y_1$, $\int_0^t (\lambda f)^\alpha = \lambda^\alpha y_2$ for any constant $\lambda > 0$ and thus

$$B(t, \lambda y_1, \lambda^\alpha y_2) \geq \lambda^\alpha \int_0^t \left(\frac{1}{s}\int_0^s f\right)^\alpha \mathrm{d}s.$$

Taking the supremum over all f as above, we get

$$B(t, \lambda y_1, \lambda^\alpha y_2) \geq \lambda^\alpha B(t, y_1, y_2), \qquad \lambda > 0. \tag{4.9}$$

Now plug $y_1 := \lambda^{-1} y_1$, $y_2 := \lambda^{-\alpha} y_2$ to obtain

$$B(t, y_1, y_2) \geq \lambda^\alpha B(t, \lambda^{-1} y_1, \lambda^{-\alpha} y_2).$$

Substituting $\lambda := \lambda^{-1}$, we obtain the reverse to (4.9): thus, equality holds here.

The second homogeneity is a little more involved. Namely, suppose again that f is an arbitrary function on $[0,t]$ satisfying $\int_0^t f = y_1$, $\int_0^t f^\alpha = y_2$. Take $\mu > 0$ and consider the dilated function $\tilde{f}(s) = f(s/\mu)$ on $[0, \mu t]$. Then

$$\int_0^{\mu t} \tilde{f}\mathrm{d}s = \int_0^{\mu t} f(s/\mu)\mathrm{d}s = \mu \int_0^t f = \mu y_1, \qquad \int_0^{\mu t} \tilde{f}^\alpha \mathrm{d}s = \mu \int_0^t f^\alpha = \mu y_2,$$

so that

$$\begin{aligned} B(\mu t, \mu y_1, \mu y_2) &\geq \int_0^{\mu t} \left(\frac{1}{s}\int_0^s \tilde{f}\right)^\alpha \mathrm{d}s \\ &= \mu \int_0^t \left(\frac{1}{\mu s}\int_0^{\mu s} \tilde{f}\right)^\alpha \mathrm{d}s = \mu \int_0^t \left(\frac{1}{s}\int_0^s f\right)^\alpha \mathrm{d}s. \end{aligned}$$

Taking the supremum over all f as above, we get

$$B(\mu t, \mu y_1, \mu y_2) \geq \mu B(t, y_1, y_2),$$

and substituting $t := \mu^{-1} t$, $y_1 := \mu^{-1} y_1$, $y_2 := \mu^{-1} y_2$ and then $\mu := \mu^{-1}$ shows that we actually have equality here. Putting all these facts together, we see that

$$B(t, y_1, y_2) = tB\left(1, \frac{y_1}{t}, \frac{y_2}{t}\right) = t\left(\frac{y_1}{t}\right)^\alpha B\left(1, 1, \frac{y_2}{t} \cdot \frac{t^\alpha}{y_1^\alpha}\right) = \frac{y_1^\alpha}{t^{\alpha-1}} B\left(1, 1, \frac{y_2 t^{\alpha-1}}{y_1^\alpha}\right),$$

that is, we have

$$B(t, y_1, y_2) = \frac{y_1^\alpha}{t^{\alpha-1}} \varphi\left(\frac{y_2 t^{\alpha-1}}{y_1^\alpha}\right) \tag{4.10}$$

for some unknown function $\varphi : [1, \infty) \to [0, \infty)$ satisfying $\varphi(1) = 1$ (for the domain of φ and the initial condition, see the above discussion on the domain of B).

Step 3. Candidate. What is the Bellman monotonicity property? Pick arbitrary $t > 0$, $\varepsilon > 0$ and any nonnegative, piecewise continuous function f on $[0, t]$ with $\int_0^t f = y_1$, $\int_0^t f^\alpha = y_2$. Extend it to $\tilde{f} : [0, t + \varepsilon] \to [0, \infty)$ putting $\tilde{f}(s) = u$ for $s \in (t, t + \varepsilon]$; here a is some arbitrary nonnegative number ("control"). Then we have

$$\int_0^t \tilde{f} = y_1 + \varepsilon a, \qquad \int_0^t \tilde{f}^\alpha = y_2 + \varepsilon a^\alpha$$

and hence

$$\begin{aligned}
B(t+\varepsilon, y_1 + \varepsilon a, y_2 + \varepsilon a^\alpha) &\geq \int_0^{t+\varepsilon} \left(\frac{1}{s}\int_0^s \tilde{f}\right)^\alpha \mathrm{d}s \\
&= \int_0^t \left(\frac{1}{s}\int_0^s \tilde{f}\right)^\alpha \mathrm{d}s + \int_t^{t+\varepsilon} s^{-\alpha} \left(\int_0^t \tilde{f} + \int_t^s \tilde{f}\right)^\alpha \mathrm{d}s \\
&= \int_0^t \left(\frac{1}{s}\int_0^s f\right)^\alpha \mathrm{d}s + \int_t^{t+\varepsilon} s^{-\alpha} \left(\int_0^t f + a(s-t)\right)^\alpha \mathrm{d}s \\
&= \int_0^t \left(\frac{1}{s}\int_0^s f\right)^\alpha \mathrm{d}s + \int_t^{t+\varepsilon} s^{-\alpha} \left(y_1 + a(s-t)\right)^\alpha \mathrm{d}s.
\end{aligned}$$

Taking the supremum over all f as above, we get

$$B(t+\varepsilon, y_1 + \varepsilon a, y_2 + \varepsilon a^\alpha) \geq B(t, y_1, y_2) + \int_t^{t+\varepsilon} s^{-\alpha} \left(y_1 + a(s-t)\right)^\alpha \mathrm{d}s.$$

Now let us *assume* that B is of class C^1. Because of this assumption, we are no longer permitted to use the letter B: we will write $\tilde{B}$ instead. Putting $\tilde{B}(t, y_1, y_2)$ on the left, dividing throughout by ε and letting $\varepsilon \to 0$ yields

$$\tilde{B}_t(t, y_1, y_2) + a\tilde{B}_{y_1}(t, y_1, y_2) + a^\alpha \tilde{B}_{y_2}(t, y_1, y_2) \geq (y_1/t)^\alpha. \tag{4.11}$$

Now, recall (4.10): we have, for $u = y_2 t^{\alpha-1}/y_1^\alpha$,

$$\tilde{B}_t(t, y_1, y_2) = (1-\alpha)\left(\frac{y_1}{t}\right)^\alpha \left(\tilde{\varphi}(u) - u\tilde{\varphi}'(u)\right),$$

$$\tilde{B}_{y_1}(t, y_1, y_2) = \alpha\left(\frac{y_1}{t}\right)^{\alpha-1} \left(\tilde{\varphi}(u) - u\tilde{\varphi}'(u)\right)$$

and $\tilde{B}_{y_2}(t, y_1, y_2) = \tilde{\varphi}'(u)$. Plugging this above, and substituting $b = at/y_1$ gives

$$\left(\tilde{\varphi}(u) - u\tilde{\varphi}'(u)\right)(1 - \alpha + \alpha b) + b^\alpha \tilde{\varphi}'(u) \geq 1.$$

This inequality must hold for all controls b. We minimize the left-hand side over b: the direct differentiation shows that the minimum is attained for

$$b = \left(\frac{u\tilde{\varphi}'(u) - \tilde{\varphi}(u)}{\tilde{\varphi}'(u)}\right)^{1/(\alpha-1)}.$$

Plugging this above, we obtain

$$(\alpha - 1)\left[1 - \left(\frac{u\tilde{\varphi}'(u) - \tilde{\varphi}(u)}{\tilde{\varphi}'(u)}\right)^{1/(\alpha-1)}\right]\left(u\tilde{\varphi}'(u) - \tilde{\varphi}(u)\right) \geq 1.$$

We assume equality: if we manage to solve this equation, we will claim that the function $\tilde{B}$ obtained via (4.10) is the candidate for the Bellman function.

Step 4. Solution to the differential equation. We will now show that there is an increasing continuous function $\varphi : [1,\infty) \to \mathbb{R}$, satisfying the differential equation

$$(\alpha-1)\left[1-\left(\frac{u\tilde{\varphi}'(u)-\tilde{\varphi}(u)}{\tilde{\varphi}'(u)}\right)^{1/(\alpha-1)}\right](u\tilde{\varphi}'(u)-\tilde{\varphi}(u))=1 \tag{4.12}$$

for $u \in (1,\infty)$ and the initial condition $\varphi(1)=1$. Furthermore,

$$\varphi(u) \leq \left(\frac{\alpha}{\alpha-1}\right)^{\alpha} u \qquad \text{for } u \geq 1. \tag{4.13}$$

Consider the function $\psi : [1,\infty) \to \mathbb{R}$ given by

$$\psi(s) = s\left(1-\frac{1}{\alpha}+\frac{1}{\alpha s}\right)^{\alpha}. \tag{4.14}$$

One easily verifies that this function is smooth, strictly increasing and maps $[1,\infty)$ onto itself. Let φ be the inverse to ψ; then $\varphi(1)=1$ and we have

$$s = \left(1-\frac{1}{\alpha}+\frac{1}{\alpha\varphi(s)}\right)^{\alpha}\varphi(s). \tag{4.15}$$

This, by a direct differentiation, yields

$$1 = \varphi'(s)\left(1-\frac{1}{\alpha}+\frac{1}{\alpha\varphi(s)}\right)^{\alpha} - \frac{\varphi'(s)}{\varphi(s)}\left(1-\frac{1}{\alpha}+\frac{1}{\alpha\varphi(s)}\right)^{\alpha-1},$$

or, equivalently,

$$\frac{1}{\varphi'(s)} = \frac{\alpha-1}{\alpha}\left(1-\frac{1}{\alpha}+\frac{1}{\alpha\varphi(s)}\right)^{\alpha-1}\left(1-\frac{1}{\varphi(s)}\right). \tag{4.16}$$

Multiply both sides by $\varphi(s)$ and subtract the obtained equality from (4.15). We get

$$s-\frac{\varphi(s)}{\varphi'(s)} = \left(1-\frac{1}{\alpha}+\frac{1}{\alpha\varphi(s)}\right)^{\alpha-1} \tag{4.17}$$

and hence (4.16) can be rewritten in the form

$$\frac{1}{\varphi'(s)} = (\alpha-1)\left(s-\frac{\varphi(s)}{\varphi'(s)}\right)\left[1-\left(s-\frac{\varphi(s)}{\varphi'(s)}\right)^{1/(\alpha-1)}\right],$$

which is the desired differential equation (4.12). To show (4.13), note that

$$\left(\frac{\varphi(s)}{s}\right)' = \frac{\varphi'(s)s-\varphi(s)}{s^2} \geq 0,$$

where the latter bound comes from (4.17) (and the estimate $\varphi'(s)>0$, which is a consequence of the strict monotonicity of ψ). Thus, (4.13) follows at once from

$$\lim_{s\to\infty}\frac{\varphi(s)}{s} = \lim_{s\to\infty}\frac{s}{\psi(s)} = \left(\frac{\alpha}{\alpha-1}\right)^{\alpha}.$$

This finishes the proof.

Step 5. $B = \tilde{B}$. Pick an arbitrary t and an arbitrary nonnegative piecewise continuous function f on $[0,t]$ with $\int_0^t f = y_1$ and $\int_0^t f^\alpha = y_2$. Then the function

$$s \mapsto \tilde{B}\left(s, \int_0^s f, \int_0^s f^\alpha\right) + \int_s^t \left(\frac{1}{r}\int_0^r f\right)^\alpha \tag{4.18}$$

is nondecreasing, by (4.11). Consequently, for any $s \le t$ we have

$$\int_s^t \left(\frac{1}{r}\int_0^r f\right)^\alpha \le \tilde{B}\left(s, \int_0^s f, \int_0^s f^\alpha\right) + \int_s^t \left(\frac{1}{r}\int_0^r f\right)^\alpha \le \tilde{B}\left(t, \int_0^t f, \int_0^t f^\alpha\right)$$

and hence, by Fatou's lemma, $\int_0^t \left(\frac{1}{s}\int_0^s f\right)^\alpha \le \tilde{B}\left(t, y_1, y_2\right)$. Taking the supremum over f gives $B \le \tilde{B}$. To show the reverse, it is enough to prove that $\tilde{B}(t, y_1, y_2)$ is obtained as the evaluation of the functional on specific functions. One can extract from the above proof that the functions of the form $f(s) = as^b$, for appropriately chosen a and b, do the job.

1. Problems

1. Solve the optimal control problem

$$\sup_u \left\{ \int_0^T \sqrt{u(t)}\mathrm{d}t + \sqrt{x(T)} \right\},$$

where $x(0) = x_0 > 0$ and $\dot{x} = -u < 0$.

2. Solve the problem

$$\sup \left\{ x_2(1) - \frac{1}{2}(x_1(1))^2 \right\},$$

where $\dot{x} = (\dot{x}_1, \dot{x}_2) = (u, u)$ $(u \in \mathbb{R})$ and $(x_1(0), x_2(0)) = (0, 0)$.

3. Consider the system of ODEs

$$\dot{x}_1 = x_1 + u, \qquad \dot{x}_2 = -u^2,$$

with the initial condition $x_1(0) = 1$, $x_2(0) = 0$, where $u \in \mathbb{R}$. Solve the problem

$$\sup_u \left\{ -x_1(T)^2 + x_2(T) \right\}.$$

4. Solve the problem

$$\sup_u \left\{ -\int_0^1 u^2 \mathrm{d}t + x(1) \right\},$$

where $\dot{x} = u > 0$, $x(0) = x_0 > 0$.

5. Solve the problem

$$\sup_u \int_0^1 (1-u)x\mathrm{d}t,$$

where $\dot{x} = u \in [0,1]$ and $x(0) = 1/2$.

6. Prove that for any $f : [0,\infty) \to \mathbb{R}$ we have

$$\int_0^\infty |f(t)|\mathrm{d}t \le \sqrt{\pi}\left(\int_0^\infty f^2(u)\mathrm{d}u\right)^{1/4}\left(\int_0^\infty f^2(u)u^2\mathrm{d}u\right)^{1/4}.$$

CHAPTER 5

Stochastic continuous case: optimal control of diffusions

We turn our attention to the stochastic setting. The rough idea is that the differential equation

$$\dot{x} = f(t, x(t), u(t)), \quad t \geq 0,$$

which in the previous chapter governed the evolution of the deterministic system, may be continuously influenced by stochastic disturbances. This additional probabilistic ingredient is expressed in terms of (stochastic integrals of) Brownian motion. For the sake of completeness of the presentation, we will discuss some basic properties of stochastic calculus and stochastic differential equations, referring the interested reader to the monograph [**6**] for the more detailed and systematic study.

1. Some background on stochastic integration

Assume that $(\Omega, \mathcal{F}, \mathbb{P})$ is a complete probability space: here by completeness we mean that for any $A \in \mathcal{F}$ of probability zero and any $B \subset A$ we have $B \in \mathcal{F}$. We equip the space with a filtration, i.e., a nondecreasing family $(\mathcal{F}_t)_{t\geq 0}$ of sub-σ-fields of $\mathcal{F}$. For technical reasons, we assume that the filtration satisfies the usual conditions, that is, it is right-continuous (we have $\bigcap_{s\geq t} \mathcal{F}_s = \mathcal{F}_t$ for all $t \geq 0$) and $\mathcal{F}_0$ contains all the events of probability zero.

The following process will be fundamental in our further considerations.

DEFINITION 5.1. A real-valued adapted process $W = (W_t)_{t\geq 0}$ is a Wiener process (Brownian motion), if it satisfies the following conditions.

(i) We have $W_0 = 0$ almost surely.

(ii) The process W has independent increments: for any $0 \leq t_0 < t_1 < t_2 < \ldots < t_n$, the variables W_{t_0}, $W_{t_1} - W_{t_0}$, $W_{t_2} - W_{t_1}$, ..., $W_{t_n} - W_{t_{n-1}}$ are independent.

(iii) For any $0 \leq s < t$, the random variable $W_t - W_s$ has the normal distribution of mean zero and variance $t - s$.

(iv) The process W has continuous trajectories.

A Wiener process $\mathbb{W}$ in $\mathbb{R}^n$ is the collection $(W^{(1)}, W^{(2)}, \ldots, W^{(n)})$ of independent one-dimensional Wiener processes $W^{(1)}$, $W^{(2)}$, ..., $W^{(n)}$.

We will describe now some background on stochastic differential equations with respect to Brownian motion. The idea is to modify the equation

$$\dot{x} = b(t, x_t) \tag{5.1}$$

by adding a stochastic component: the equation will take the form

$$\frac{dx}{dt} = b(t, x_t) + \sigma(t, x_t) W_t, \qquad t \geq 0,$$

or rather

$$dx_t = b(t, x_t)dt + \sigma(t, x_t)dW_t, \qquad t \geq 0. \tag{5.2}$$

Here W can be interpreted as the "white noise" inserted into (5.1). One of the main problems is that formally, the right-hand side above does not make sense: with probability one, the trajectories of Brownian motion are nowhere differentiable and hence dW_t is meaningless. To overcome this difficulty, one rewrites (5.2) in the integral form

$$x_t = x_0 + \int_0^t b(s, x_s)\mathrm{d}s + \int_0^t \sigma(s, x_s)\mathrm{d}W_s, \qquad t \geq 0,$$

reducing the problem to that of giving sense to the latter integral. Again, this cannot be done pointwise by means of Stjeltjes integrals: it can be shown that almost surely, Brownian motion has an infinite variation on any subinterval of $[0, \infty)$. It turns out that the problem can be tackled by referring to the L^2 isometry argument. More specifically, suppose that H is a fixed process and assume that we are interested in defining the integral $(H \cdot W)_t = \int_0^t H_s \mathrm{d}W_s$. Assume first that H is "very simple", that is, we have

$$H = \xi 1_{\{0\}} \quad \text{for some } \mathcal{F}_0\text{-measurable random variable } \xi$$

or, for some $0 \leq s \leq t$,

$$H = \xi 1_{(s,t]} \quad \text{for some } \mathcal{F}_s\text{-measurable random variable } \xi.$$

Then for any r, we define

$$\int_0^r H_u \mathrm{d}W_u = 0$$

in the first case, and

$$\int_0^r H_u \mathrm{d}W_u = \xi(W_{r\wedge t} - W_{s\wedge t})$$

in the second case. By linearity, this definition extends to integrals of simple integrands H (linear combinations of very simple integrands). By a beautiful L^2-limiting argument, one can define the stochastic integrals, with respect to X, of *locally bounded predictable processes* H. The key property which enables the extension is the identity

$$\mathbb{E}\left(\int_0^t H_s \mathrm{d}W_s\right)^2 = \mathbb{E}\int_0^t H_s^2 \mathrm{d}s,$$

provided H is a simple integrand.

One of the fundamental statements in the theory of stochastic integration is the following Itô's formula, which will be stated for Brownian motion only. Namely, for any C^2 function on $[0, \infty) \times \mathbb{R}$, we have

$$\begin{aligned} F(t, W_t) &= F(0,0) + \int_0^t F_t(s, W_s)\mathrm{d}s + \int_0^t F_x(s, W_s) \cdot \mathrm{d}W_s + \frac{1}{2}\int_0^t F_{xx}(s, W_s)\mathrm{d}s \\ &= F(0,0) + \int_0^t \left(F_t(s, W_s) + \frac{1}{2}F_{xx}(s, W_s)\right)\mathrm{d}s + \int_0^t F_x(s, W_s) \cdot \mathrm{d}W_s. \end{aligned}$$

In the n-dimensional case, the formula is slightly more complicated: for any C^2 function on $[0,\infty)\times\mathbb{R}^n$, we have

$$\begin{aligned}
&F(t,\mathbb{W}_t)\\
&= F(0,0,\dots,0) + \int_0^t F_t(s,\mathbb{W}_s)\mathrm{d}s + \sum_{k=1}^n \int_0^t F_{x_k}(s,\mathbb{W}_s)\mathrm{d}W_s^{(k)} + \frac{1}{2}\int_0^t \Delta F(s,\mathbb{W}_s)\mathrm{d}s\\
&= F(0,0,\dots,0) + \int_0^t \Big(F_t(s,\mathbb{W}_s) + \frac{1}{2}\Delta F(s,\mathbb{W}_s)\Big)\,\mathrm{d}s + \sum_{k=1}^n \int_0^t F_{x_k}(s,\mathbb{W}_s)\mathrm{d}W_s^{(k)},
\end{aligned}$$

where $\mathbb{W} = (W^{(1)}, W^{(2)}, \dots, W^{(n)})$ and ΔF stands for the Laplacian of F.

We return to the context of the stochastic differential equation

$$dX_t = b(t,X_t)dt + \sigma(t,X_t)dW_t, \qquad t\ge 0, \tag{5.3}$$

with the initial condition $X_0 = x_0$, where x_0 is a given $\mathcal{F}_0$-measurable random variable. Here we assume that b and σ are piecewise and right-continuous in t and Lipschitz in x, uniformly in t (there is a constant K such that for any t, x and y we have $|b(x,t)-b(y,t)|\le K|x-y|$). One of fundamental questions concerns the existence and uniqueness of the solution.

THEOREM 5.1. *Assume that b and σ are piecewise and right continuous in t and Lipschitz continuous in x uniformly in t, and that $\sigma(t,x_0)$ and $b(t,x_0)$ are functions bounded by some constant. Then there is a unique adapted solution X_t of (5.3) on $[0,\infty)$, continuous in t and locally square-integrable.*

Sometimes the processes satisfying (5.3) are called diffusions. The statement of the theorem remains valid if b and σ are assumed to be adapted processes: $b(t,x,\omega)$, $\sigma(t,x,\omega)$. Furthermore, the formulation is still true in the n-dimensional context, in which b is a function with values in $\mathbb{R}^n$, σ is a function which takes values in the set of $n\times n$ matrices and W is replaced by an n-dimensional Brownian motion $\mathbb{W}$.

All the above discussion concerned the case in which the starting time is equal to 0 and the starting position is a given random variable x_0. However, the statements extend immediately to the context in which the starting time is some parameter $s>0$, and the initial variable is equal to x_s (which can be random or deterministic). For the sake of convenience, we distinguish the family of probability measures $\mathbb{P}_{s,x}$ (and the corresponding expectations $\mathbb{E}_{s,x}$), where the index s indicates the starting time, while x stands for the *deterministic* starting position equal to x.

Before we proceed, we mention an importance consequence of Itô's formula. Suppose that f is of class C^2 on $\mathbb{R}_+\times\mathbb{R}^n$ and X satisfies (5.3). Then

$$\begin{aligned}
&\lim_{\varepsilon\to 0}\frac{\mathbb{E}^{s,x}f(s+\varepsilon,X_{s+\varepsilon}) - f(s,x)}{\varepsilon}\\
&= f_t(s,x) + \Big\langle \nabla_x f(s,x), b(s,x)\Big\rangle + \frac{1}{2}\Big\langle D^2 f(s,x)\sigma(s,x),\sigma(s,x)\Big\rangle,
\end{aligned}$$

where ∇_x denotes the gradient with respect to x-variables, $\langle\cdot,\cdot\rangle$ is the scalar product in $\mathbb{R}^n$ and $D^2 f$ is the Hessian matrix of f.

2. Optimal control

Suppose that the n-dimensional diffusion X_t is governed by the controlled stochastic differential equation

$$dX_t = b(t, X_t, u(t, X_t))dt + \sigma(t, X_t, u(t, X_t))dW_t, \qquad X_0 = x_0.$$

Here the control $u(t, X_t)$ is a function taking values in a given set U and is subject to choice. The functions b, σ and x_0 are given, with σ being an $n \times m$ matrix with entries $\sigma_{i,j}(t, x, u)$, and W_t is a vector of m given independent Brownian motions. We assume that the coordinates/entries of b and σ are piecewise and right-continuous in t, and uniformly Lipschitz continuous in (x, u). For any given (s, x) in $[0, T] \times \mathbb{R}^n$, define the functional

$$J(s, x, u) = \mathbb{E}^{s,x} \left[\int_s^T g(t, X_t, u(t, X_t))\mathrm{d}t + r(T, X_T) \right],$$

where g and r are given functions and T is a fixed time horizon (in case $T = \infty$, the function r is assumed to be zero). Suppose we are interested in the problem

$$B := \sup J(0, x_0, u), \tag{5.4}$$

where the supremum is taken over all controls u with values in U. As usual, two questions arise:

· to identify the control u^* (if exists), which yields the optimal performance, i.e., such that we have

$$J(0, x_0, u^*) = \sup_u J(0, x_0, u).$$

· to compute the explicit value of B.

In (5.4), we consider only the controls of the form $u(t, x)$, so-called Markov controls, taking values in U. It is possible to allow more general control functions $u(t, \omega)$ that are dependent on past values of $\mathbb{W}$; however, in most cases the supremum is achieved in the Markovian case.

As previously, the successful treatment of the above problem rests on extending it to general starting positions. Define the associated Bellman (or value) function $B : [0, T] \times \mathbb{R}^n \to \mathbb{R}$ by the formula

$$B(s, x) = \sup_u J(s, x, u).$$

Clearly, B satisfies the initial (or rather terminal) condition $B(T, x) = r(T, x)$ for all x. For $t < T$, the values of $B(t, x)$ are governed by an appropriate version of Hamilton-Jacobi-Bellman equation (HJB). Let us present a rough reasoning which leads to this statement. Fix a point (s, x) and a small positive number ε, and take a look at the definition of $B(s, x)$. Assume that B is of class C^2 and consider the following control. First, on the time interval $[s, s + \varepsilon]$, we take $u \equiv a$, where a is an arbitrarily chosen element of U: hence we have

$$X^u_{s+\varepsilon} = x + \int_s^{s+\varepsilon} b(t, X_t, a)\mathrm{d}t + \int_s^{s+\varepsilon} \sigma(t, X_t, a)\mathrm{d}W_t.$$

Second, depending on where the diffusion is located at time $s + \varepsilon$, we apply the corresponding (almost) optimal control coming from $B(s + \varepsilon, X^u_{s+\varepsilon})$. The control

u we have just constructed leads to the inequality

$$\begin{aligned}B(s,x) &\geq J(s,x,u)\\ &= \mathbb{E}^{s,x}\left[\int_s^{s+\varepsilon} g(t,X_t,a)\mathrm{d}t + \int_{s+\varepsilon}^T g(t,X_t,u(t,X_t))\mathrm{d}t + r(T,X_T)\right],\end{aligned}$$

or

$$B(s,x) \geq \mathbb{E}^{s,x}\left[\int_s^{s+\varepsilon} g(t,X_t,a)\mathrm{d}t\right] + \mathbb{E}^{s,x}B(s+\varepsilon, X_{s+\varepsilon}).$$

We move $B(s,x)$ to the left, divide both sides by ε and let $\varepsilon \to 0$, obtaining

$$\begin{aligned}0 &\geq g(s,x,a) + \mathbb{L}B(s,x)\\ &= g(s,x,a) + B_s(s,x) + B_x(s,x)b(s,x,a) + \frac{1}{2}B_{xx}(s,x)\sigma^2(s,x,a).\end{aligned}$$

Furthermore, we might hope that if a is chosen in an optimal manner, then equality will hold above. This is the aforementioned Hamilton-Jacobi-Bellman equation

$$B_s(s,x) + \sup_{a\in U}\left\{g(s,x,a) + B_x(s,x)b(s,x,a) + \frac{1}{2}B_{xx}(s,x)\sigma^2(s,x,a)\right\} = 0.$$

In the n-dimensional context, the formula is analogous, though a little more complicated:

$$\begin{aligned}B_s(s,x) + \sup_{a\in U}\Big\{&g(s,x,a) + \langle \nabla_x B(s,x), b(s,x,a)\rangle\\ &+ \frac{1}{2}\Big\langle \Delta_x B(s,x)\sigma(s,x,a), \sigma(s,x,a)\Big\rangle\Big\} = 0,\end{aligned} \tag{5.5}$$

where $\langle\cdot,\cdot\rangle$ is the scalar product in $\mathbb{R}^n$.

More precisely, we have the following two statements.

Theorem 5.2 (Necessary Conditions). *Suppose that B is twice continuously differentiable on $(0,T)\times\mathbb{R}^n$ and continuous on $[0,T]\times\mathbb{R}^n$. Then B satisfies the HJB equation (5.5) in $(0,T)\times\mathbb{R}^n$, together with $B(T,x) = r(T,x)$. Moreover, if u^* is an optimal control, then $a = u^*(s,x)$ maximizes the right-hand side of (5.5).*

Theorem 5.3 (Sufficient Conditions). *Suppose that $\tilde{B}$ is a function that is continuous on $[0,T]\times\mathbb{R}^n$, is twice continuously differentiable in $(0,T)\times\mathbb{R}^n$, and satisfies (5.5) on $(0,T)\times\mathbb{R}^n$, together with the boundary condition $\tilde{B}(T,x) = r(T,x)$ for all x. Assume that for every pair (t,x), $u_0(t,x)$ is the value of $u\in U$ that yields the maximum in the right hand side of (5.5). Then u_0 is the optimal control.*

Example 5.4. Consider the problem

$$\sup_u \mathbb{E}\left[\int_0^T -u^2\mathrm{d}t - X_T^2\right],$$

where $dX_t = udt + \sigma dW_t$, $X_0 = x_0$; here $\sigma > 0$ and $x_0\in\mathbb{R}$ are given numbers.

We introduce the Bellman function

$$B(s,y) = \sup_u \mathbb{E}\left[\int_s^T -u^2\mathrm{d}t - X_T^2\right],$$

where X satisfies the same SPDE as above, with $X_s = y$. The Bellman equation takes the form

$$B_s(t,x) + \sup_{a\in\mathbb{R}} \left\{ -a^2 + B_x(t,x)u + \frac{1}{2}B_{xx}(t,x)\sigma^2 \right\} = 0.$$

The supremum is attained for $a = B_x(t,x)/2$, plugging this control above gives

$$0 = B_s(t,x) + \frac{B_x(t,x)^2}{4} + \frac{B_{xx}(t,x)\sigma^2}{2}. \tag{5.6}$$

Furthermore, the terminal condition reads $B(T,x) = -x^2$. To solve the above PDE, we search for some homogeneity-type property of B which enables the reduction to the ordinary differential equation. Suppose that u is an arbitrary control and X is the process starting from y. Next, pick a different starting point $\tilde{y}$ and change the control to $\tilde{u} = u + c$, where c is a parameter. Then we have

$$\tilde{X}_t = \tilde{y} + \int_s^T \tilde{u}\mathrm{d}t + \sigma W_T = \tilde{y} - y + c(T-s) + X_T.$$

Hence

$$\begin{aligned}
B(s,\tilde{y}) &\geq \mathbb{E}\left[\int_s^T -\tilde{u}^2\mathrm{d}t - \tilde{X}_T^2\right] \\
&= \mathbb{E}\bigg[\int_s^T -u^2\mathrm{d}t - 2c\int_s^T u\mathrm{d}t - c^2(T-s) \\
&\qquad - X_T^2 - 2(\tilde{y} - y + c(T-s))\left(y + \int_s^T u\mathrm{d}t\right) - (\tilde{y} - y + c(T-s))^2\bigg].
\end{aligned}$$

Now suppose that $\tilde{y} - y + c(T-s) = -c$. Then the integral $\int_s^T u$ cancels out and we obtain

$$B(s,\tilde{y}) \geq \mathbb{E}\left[\int_s^T -u^2\mathrm{d}t - X_T^2\right] - c^2(T-s) + 2cy - c^2.$$

Hence, taking the supremum over u and recalling that $\tilde{y} - y + c(T-s) = -c$, we get the estimate

$$B(s, y - c(T-s+1)) \geq B(s,y) - c^2(T-s+1) + 2cy \tag{5.7}$$

for all c, $y \in \mathbb{R}$. Now substitute $\bar{y} = y - c(T-s+1)$ and $\bar{c} = -c$. Plugging this above, we see that

$$B(s,\bar{y}) \geq B(s, \bar{y} - \bar{c}(T-s+1)) - \bar{c}^2(T-s+1) - 2\bar{c}(\bar{y} - \bar{c}(T-s+1)),$$

which is equivalent to

$$B(s, \bar{y} - \bar{c}(T-s+1)) \leq B(s,\bar{y}) - \bar{c}^2(T-s+1) + 2\bar{c}\bar{y}.$$

But $\bar{y}$ and $\bar{c}$ can take arbitrary real values; thus equality holds in (5.7), which is the desired homogeneity. Consequently, taking $c = y/(T-s+1)$, we obtain

$$B(s,y) = B(s,0) - \frac{y^2}{T-s+1} = \varphi(s) - \frac{y^2}{T-s+1}.$$

Plugging this into (5.6), we get

$$\varphi'(s) + \frac{\sigma^2}{T-s+1} = 0,$$

so that $\varphi(s) = \sigma^2 \ln(T - s + 1) + C$ for some constant C. This yields

$$B(s, y) = \sigma^2 \ln(T - s + 1) - \frac{y^2}{T - s + 1} + C,$$

and the terminal condition $B(T, y) = -y^2$ gives $C = 0$. Strictly speaking, the function B we obtained is the *candidate* for the Bellman function; however, since it is of class C^2 and satisfies the HJB equation, Theorem 5.3 implies that it does coincide with the abstract Bellman function. Note that the optimal control at time t is equal to $u^*(t) = -X_t/(T - t + 1)$.

The methodology we described above works perfectly for the infinite-horizon problems.

EXAMPLE 5.5. Consider the problem

$$\sup_u \mathbb{E} \left\{ -\int_0^\infty e^{-\rho t}(X_t^2 + u_t^2)\mathrm{d}t \right\},$$

where $dX_t = u\mathrm{d}t + X_t \mathrm{d}W_t$ and $X_0 = 1$. Here ρ is a fixed positive constant. As usual, we start with writing down the abstract Bellman function corresponding to the problem:

$$B(s, y) = \sup_u \mathbb{E} \left\{ -\int_s^\infty e^{-\rho t}(X_t^2 + u_t^2)\mathrm{d}t \right\},$$

where $X_s = y$ and X is governed by the same stochastic differential equation as above. Now we will establish two homogeneity properties of B. First, suppose that $X_s = y$ and X satisfies the above SPDE; then $\tilde{X}_t = X_{t+s}$. $t \geq 0$, satisfies $\tilde{X}_0 = y$ and $d\tilde{X}_t = dX_{t+s} = u(t+s)\mathrm{d}t + X_{t+s}\mathrm{d}W_{t+s} = u(t+s)\mathrm{d}t + \tilde{X}_t \mathrm{d}\tilde{W}_t$, where $\tilde{W}_t = W_{t+s} - W_s$ is a Brownian motion. Since in addition we have

$$\begin{aligned}\mathbb{E}\int_s^\infty e^{-\rho t}(X_t^2 + u_t^2)\mathrm{d}t &= \mathbb{E}\int_0^\infty e^{-\rho(t+s)}(X_{t+s}^2 + u_{t+s}^2)\mathrm{d}t \\ &= e^{-\rho s}\mathbb{E}\int_0^\infty e^{-\rho t}(\tilde{X}_t^2 + u_{t+s}^2)\mathrm{d}t,\end{aligned}$$

we infer that $B(s, y) = e^{-\rho s}B(0, y)$. Next, observe that if X is a process induced by $X_0 = y$ and the control u, then λX is a process induced by $\lambda X_0 = \lambda y$ and λu: this follows at once from the linearity of the stochastic differential equation. Thus

$$B(0, \lambda y) = \sup_u \mathbb{E} \left\{ -\int_0^\infty e^{-\rho t}(\lambda^2 X_t^2 + \lambda^2 u_t^2)\mathrm{d}t \right\} = \lambda^2 B(0, y).$$

Putting the above homogeneity properties together, we see that $B(s, y) = \gamma e^{-\rho s}y^2$ for some unknown constant γ; in particular, B is of class C^2. To compute γ, we derive the corresponding HJB equation. We proceed as usual: for a fixed starting position y, we consider the control u which is equal to a on some small interval $[0, \varepsilon]$, and is the (almost) optimal control induced by $B(\varepsilon, X_\varepsilon)$ on (ε, ∞). Then

$$B(0, y) \geq -\int_0^\varepsilon e^{-\rho t}(X_t^2 + u_t^2)\mathrm{d}t + \mathbb{E}^{0,y}B(\varepsilon, X_\varepsilon).$$

Putting all the terms on the right, dividing by ε and letting $\varepsilon \to 0$ yields

$$\mathbb{L}_{(t,X)}B(0, y) - (y^2 + a^2) \leq 0,$$

or

$$B_s(0, y) + B_x(0, y)a + \frac{1}{2}B_{xx}(0, y)y^2 - (y^2 + a^2) \leq 0.$$

Plugging the formula for B, we obtain

$$\gamma\left(-\rho y^2 + 2ya + y^2\right) - y^2 - a^2 \leq 0.$$

Now we optimize over a and assume equality: the left-hand side is the largest for $a = \gamma y$ and hence we get $y^2\big[\gamma(-\rho+\gamma+1) - 1\big] = 0$. This equality must hold for all y, and thus the expression in the square brackets must be zero. There are two solutions

$$\gamma_{\pm} = \frac{\rho - 1 \pm \sqrt{(1-\rho)^2 + 4}}{2},$$

satisfying $\gamma_+ > 0$ and $\gamma_- < 0$. But by the very definition of B, it is clear that $B \leq 0$; thus we are forced to take γ_- and hence

$$B(s,y) = \frac{\rho - 1 - \sqrt{(1-\rho)^2 + 4}}{2} \cdot e^{-\rho s} y^2.$$

Finally, note that the optimal control is $u_t^* = \gamma_- X_t$, $t \geq 0$.

It should however be emphasized that in general, Bellman functions need not be of class C^2 and hence the application of the HJB may not make sense.

EXAMPLE 5.6. Let

$$F(x) = \begin{cases} 0 & \text{if } x \leq 0, \\ x^2 & \text{if } x \in [0,1], \\ 1 & \text{if } x \geq 1. \end{cases}$$

Consider the optimal control problem $\sup_u \mathbb{E}F(X_T)$, where $X_0 = 0$ and $dX_t = u_t dW_t$; here $u \in \mathbb{R}$ are arbitrary. We introduce the Bellman function

$$B(s,y) = \sup_u \mathbb{E}F(X_T),$$

where $X_s = y$ and X satisfies the same SPDE as above. The terminal condition reads $B(T,y) = F(y)$. Furthermore, if $B \in C^2$, then the HJB equation takes the form

$$B_s(s,y) + \sup_{a\in\mathbb{R}} \left\{\frac{1}{2} B_{yy}(s,y) a^2\right\} = 0.$$

This implies that B_{yy} must be nonpositive; then the optimal control a is zero and we obtain $B_s(s,y) = 0$ for all $s \in [0,T]$ and $y \in \mathbb{R}$. Consequently, we deduce that $B(s,y) = B(T,y) = F(y)$ for all $s \in [0,T]$ and $y \in \mathbb{R}$. But this contradicts the condition $B_{yy} \leq 0$: the function F is not concave. The problem here lies in the fact that B is not of class C^2, and that the optimal control does not exist. To see this, we solve find the Bellman function "by hand". Clearly, if $y \geq 1$, then it is optimal to take $u \equiv 0$: then the process X is constant and $\mathbb{E}F(X_T) = F(y) = 1$ is obviously optimal. Similarly, if at some time t the process X reaches 1, then from that time on it is optimal to take the zero control. On the other hand, if $y < 1$, let us consider $u = a$ on some small time interval $[s, s+\varepsilon]$, for some big positive number a. Then $X_t = s + a(W_t - W_s)$. It follows from the law of iterated logarithm for Brownian motion that if a is chosen sufficiently large, then X reaches 1 with probability as close to 1 as we wish. Putting the above considerations together, we see that $B(s,y) = 1$ for all $s < T$ and $y \in \mathbb{R}$. Furthermore, for $y < 1$ the optimal control does not exist: it would be optimal to take $u^*(t,y) = \infty$ for $y < 1$ and $u^*(t,y) = 0$ for $y \geq 1$, which makes no sense.

3. Problems

1. Solve the problem

$$\sup \left\{ \frac{1}{2} \int_0^T [-u^2(t)e^{-X_t}]\mathrm{d}t + e^{X_T} \right\},$$

where $dX_t = u(t)e^{-X_t}dt + \sigma dW_t$, $X_0 = x_0$; here $\sigma > 0$ and x_0 are fixed numbers.

2. Let X_t denote the wealth of a person at time t. Suppose that at each time, the person has two possible investments to choose. The first, risky asset Y is assumed to satisfy the SPDE

$$dY_t = aY_t dt + \sigma Y_t dW_t, \qquad t \geq 0,$$

for some given parameters a, $\sigma > 0$. The second, risk-free asset Z satisfies

$$dZ_t = bZ_t dW_t,$$

for some $b \in (0, a)$. At each instant t the person can choose how big fraction $u(t) \in [0, 1]$ of the wealth will be invested in the risky asset (the remaining part is invested in Z). Find the control u which optimizes the functional $\mathbb{E}\sqrt{X_T}$, where T is a fixed finite horizon.

3. Solve the problem

$$\sup_u \mathbb{P}(X_T \geq 1),$$

where $dX_t = u(t)dW_t$, $u(t) \in [-1, 1]$, $X_0 = x_0$.

CHAPTER 6

Towards harmonic analysis: Buckley's inequality

The purpose of this short chapter is to show that the methods of optimal control can be applied successfully in some problems of harmonic analysis. The presentation is based on notes [**8**].

Let I be a subinterval of the real line $\mathbb{R}$. A weight w on I is a positive, integrable function. For any subinterval $J \subseteq I$, we will use the notation

$$\langle w \rangle_J = \frac{1}{|J|} \int_J w$$

for the average of w over J. Here and below, the integration will be taken with respect to Lebesgue's measure and $|J|$ is the length of J.

For I as above and $\delta \geq 1$, let

$$A_\infty(I, \delta) := \Big\{ w \, : \, \langle w \rangle_J \leq \delta e^{\langle w \rangle_J} \text{ for all } J \subseteq I \Big\}$$

be the δ-ball in the Muckenhoupt class A_∞. Let $\mathcal{D}_I$ stand for the class of all dyadic subintervals of I. Let $A_\infty^d(I, \delta)$ be the dyadic version of $A_\infty(I, \delta)$, in the definition of which only $J \in D_I$ are considered.

In 1991, S. Buckley proved the following result.

THEOREM 6.1. *There exists a constant $c = c(\delta)$ such that*

$$\sum_{J \in D_I} |J| \left(\frac{\langle w \rangle_{J_+} - \langle w \rangle_{J_-}}{\langle w \rangle_J} \right)^2 \leq c(\delta)|I|$$

for any weight $w \in A_\infty^d(I, \delta)$.

Here $I_\pm$ stand for the left/right half of the interval I. Our primary goal is to prove the above estimate with the constant $c(\delta) = 8 \log \delta$.

The Bellman function corresponding to this problem is given as follows. Fix $\delta \geq 1$ and define

$$B(x_1, x_2) = \sup \left\{ \frac{1}{|I|} \sum_{J \in D_I} |J| \left(\frac{\langle w \rangle_{J_+} - \langle w \rangle_{J_-}}{\langle w \rangle_J} \right)^2 \, : \, \langle w \rangle_J = x_1, \, \langle \log w \rangle_J = x_2 \right\},$$

where the supremum is taken over all $w \in A_\infty(I, \delta)$. Formally, this function is defined on the domain

$$\Omega_\delta = \Big\{ x = (x_1, x_2) \, : \, \log \frac{x_1}{\delta} \leq x_2 \leq \log x_1 \Big\}.$$

To see this, note that the right-hand side is guaranteed by Jensen's inequality and the left is due to the A_∞ condition. Note that B does not depend on the interval I, which can be seen by applying an appropriate affine transformation.

Our goal is to find the explicit formula for B. To this end, we could proceed as usual: find a certain candidate for this object, and then show that it coincides with B. We will follow this path, but obtain a suboptimal candidate; the constant $8\log\delta$ will not be optimal (we will show that $B \le \tilde{B}$ only).

We turn our attention to the properties of B. First, by the very definition of B, we have the following "initial condition":

$$B(x_1, \log x_1) = 0.$$

There is only one weight w which satisfies $\langle w\rangle_I = x_1$ and $\langle \log w\rangle_I = \log x_1$: the constant one, equal to x_1 almost everywhere. For this weight, the sum in the definition of B vanishes.

Here is a version of the Bellman equation, sometimes called "the main inequality" in the literature.

LEMMA 6.2. *For every pair $x^\pm$ of points from Ω_δ such that $(x_+ + x_-)/2 \in \Omega_\delta$, we have*

$$B(x) \ge \frac{B(x^-)+B(x^+)}{2} + \left(\frac{x_1^- - x_1^+}{x_1}\right)^2.$$

PROOF. Pick arbitrary weights $w^\pm \in A_\infty(I_\pm,\delta)$ with $\langle w^\pm\rangle_{I_\pm} = x_1^\pm$ and $\langle w^\pm\rangle_{I_\pm} = x_2^\pm$. Then their concatenation $w : I \to \mathbb{R}$ belongs to $A_infty(I,\delta)$ and satisfies $\langle w\rangle_I = x_1$ and $\langle w\rangle_I = x_2$. Consequently, by the definition of the Bellman function,

$$\begin{aligned}
&B(x_1,x_2)\\
&\ge \frac{1}{|I|}\sum_{J\in D_I}|J|\left(\frac{\langle w\rangle_{J_+}-\langle w\rangle_{J_-}}{\langle w\rangle_J}\right)^2\\
&= \frac{1}{2|I_-|}\sum_{J\in D_{I_-}}|J|\left(\frac{\langle w\rangle_{J_+}-\langle w\rangle_{J_-}}{\langle w\rangle_J}\right)^2 + \frac{1}{2|I_+|}\sum_{J\in D_{I_+}}|J|\left(\frac{\langle w\rangle_{J_+}-\langle w\rangle_{J_-}}{\langle w\rangle_J}\right)^2\\
&\quad + \left(\frac{\langle w\rangle_{I_+}-\langle w\rangle_{I_-}}{\langle w\rangle_I}\right)^2.
\end{aligned}$$

Taking the supremum over $w^\pm$, we get the claim. $\square$

Furthermore, B enojoys the following homogeneity condition.

LEMMA 6.3. *We have $B(\lambda x_1, \log\lambda + x_2) = B(x_1,x_2)$ for any $\lambda > 0$ and any $(x_1,x_2)\in\Omega_\delta$. In particular, $B(x_1,x_2) = B(x_1e^{-x_2},0) = \varphi(x_1e^{-x_2})$.*

PROOF. Take an arbitrary weight $w\in A_\infty(I,\delta)$ with $\langle w\rangle_I = x_1$ and $\langle \log w\rangle_I = x_2$. Then for any $\lambda>0$, the weight $\tilde{w} = \lambda w$ also belongs to $A_\infty(I,\delta)$ and satisfies $\langle\tilde{w}\rangle_I = \lambda x_1$ and $\langle\log\tilde{w}\rangle_I = \log\lambda + x_2$. Consequently, by the very definition of B,

$$B(\lambda x_1, \log\lambda + x_2) \ge \frac{1}{|I|}\sum_{J\in D_I}|J|\left(\frac{\langle\tilde{w}\rangle_{J_+}-\langle\tilde{w}\rangle_{J_-}}{\langle\tilde{w}\rangle_J}\right)^2 = \frac{1}{|I|}\sum_{J\in D_I}|J|\left(\frac{\langle w\rangle_{J_+}-\langle w\rangle_{J_-}}{\langle w\rangle_J}\right)^2.$$

Taking the supremum over w, we get $B(\lambda x_1,\log\lambda + x_2)\ge B(x_1,x_2)$, setting $x_1 := \lambda x_1$, $x_2 := \log\lambda + x_2$ and $\lambda := \lambda^{-1}$, we obtain the reverse bound. $\square$

Step 1. A candidate for B. We write down an infinitesimal version of the Bellman equation. Namely, pick an arbitrary $x = (x_1,x_2)$ belonging to interior of Ω_δ and fix h, $k\in\mathbb{R}$. Then for ε sufficiently close to 0, we have $x^\pm = (x_1\pm\varepsilon h, x_2\pm\varepsilon k)\in$

Ω_δ. Plug x, x^-, x^+ into the Bellman equation, put everything on the left, divide throughout by ε^2 and let $\varepsilon \to 0$, obtaining

$$\frac{1}{2}\tilde{B}_{x_1x_1}(x)h^2 + \tilde{B}_{x_1x_2}(x)hk + \frac{1}{2}\tilde{B}_{x_2x_2}(x)k^2 + 4\left(\frac{h}{x_1}\right)^2 \le 0.$$

Since h, k were chosen arbitrarily, we see that the Hessian matrix

$$\begin{pmatrix} B_{x_1x_1} + \frac{8}{x_1^2} & B_{x_1x_2} \\ B_{x_1x_2} & B_{x_2x_2} \end{pmatrix}$$

must be semipositive-definite. The homogeneity allows us to rewrite this in the form

$$\begin{pmatrix} e^{-2x_2}\left(g''(s) + \frac{8}{s^2}\right) & -e^{-x_2}(sg'(s))' \\ -e^{-x_2}(sg'(s))' & s(sg'(s))' \end{pmatrix} \le 0, \tag{6.1}$$

where $s = x_1e^{-x_2}$. This is equivalent to saying that $(sg')' \le 0$ and the determinant is nonnegative. Now, as usual, we assume the degeneracy condition: the determinant must vanish. This assumption is plausible: if we are interested in sharp estimates, then there should be weights for which the Bellman equation holds (that is, we have equality in the main inequality). If we compute the determinant, we are led to the equation

$$\left(g'(s) - \frac{8}{s}\right)(sg')' = 0.$$

The general solution to this equation is $g(s) = c\log s + c_1$; since $g(1) = 0$ (the initial condition), we are forced to take $c_1 = 0$. To find c, we go back to the inequality (6.1). We must have $e^{-2x_2}(g''(s) + 8/s^2) \le 0$, equivalently, $-c + 8 \le 0$. Thus, it is natural to assume that $c = 8$. We have obtained the candidate

$$g(s) = 8\log s \quad \text{and} \quad \tilde{B}(x_1, x_2) = 8(\log x_1 - x_2).$$

Step 2. We will show that $B \le \tilde{B}$. First, we check that $\tilde{B}$ does satisfy the main inequality. We have

$$\begin{aligned}
&\tilde{B}(x) - \frac{\tilde{B}(x^-) + \tilde{B}(x^+)}{2} - \left(\frac{x_1^+ - x_1^-}{x_1}\right)^2 \\
&= 8\log x_1 - 8x_2 - 4\log(x_1^-x_1^+) + 4(x_2^- + x_2^+) - \left(\frac{x_1^+ - x_1^-}{x_1}\right)^2 \\
&= 4\log\frac{x_1^2}{(x_1+h)(x_1-h)} - 4\left(\frac{h}{x_1}\right)^2 \\
&= -4\left[\log\left(1 - \left(\frac{h}{x_1}\right)^2\right) + \left(\frac{h}{x_1}\right)^2\right] \ge 0.
\end{aligned}$$

This inequality enables us to write an appropriate induction step. Pick an arbitrary weight $w \in A_\infty(I, \delta)$ and, for $J \in D_I$, denote $x^J = (\langle w\rangle_J, \langle \log w\rangle_J)$. The main inequality yields

$$|J|\tilde{B}(x^J) \ge |J_-|\tilde{B}(x^{J_-}) + |J_+|\tilde{B}(x^{J_+})| + |J|\left(\frac{x_1^{J_-} - x_1^{J_+}}{x_1^J}\right)^2.$$

We will use this inequality repeatedly, to I; then to $I_\pm$; then to $I_{\pm\pm}$, etc. Let D_I^n denote the class of all dyadic subintervals of I of length $2^{-n}|I|$. We obtain

$$|I|\tilde{B}(x^I) \geq \sum_{J\in D_I^n} |J|\tilde{B}(x^J)| + \sum_{k=0}^{n-1} \sum_{J\in D_I^k} |J| \left(\frac{x_1^{J_-} - x_1^{J_+}}{x_1^J}\right)^2 .$$

But $\tilde{B}$ is nonnegative; we thus obtain

$$\sum_{k=0}^{n-1} \sum_{J\in D_I^k} |J| \left(\frac{x_1^{J_-} - x_1^{J_+}}{x_1^J}\right)^2 \leq |I|\tilde{B}(x^I).$$

Letting $n \to \infty$ are get

$$\sum_{J\in D_I} |J| \left(\frac{x_1^{J_-} - x_1^{J_+}}{x_1^J}\right)^2 \leq |I|\tilde{B}(x^I).$$

Dividing by $|I|$ and taking the supremum over w, we get $B(x^I) \leq \tilde{B}(x^I)$. In particular, $B(x^I) \leq \varphi(x_1 e^{-x_2}) \leq 8\log\delta$.

Bibliography

[1] Bellman, R. Dynamic programming. Reprint of the 1957 edition. Princeton Landmarks in Mathematics. Princeton University Press, Princeton, NJ, 2010.

[2] P. D. Bertsekas, S. E. Shreve, Stochastic optimal control. The discrete time case. Mathematics in Science and Engineering, 139. Academic Press, Inc. [Harcourt Brace Jovanovich, Publishers], New York-London, 1978.

[3] Carleman, T. *Sur les fonctions quasi-analytiques.* Conférences faites au cinquieme congres des mathématiciens Scandinaves, Helsinki (1923), 181–196.

[4] Carlson, F. *Une inegalité*, Ark. Math. Astron. Fys. **25** (1934), 1–5.

[5] Hardy, G. H.; Littlewood, J. E. and Pólya, G. *Inequalities.* Reprint of the 1952 edition. Cambridge Mathematical Library. Cambridge University Press, Cambridge, 1988.

[6] D. Revuz and M. Yor, Continuous martingales and Brownian motion. Third edition. Grundlehren der mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences], 293. Springer-Verlag, Berlin, 1999.

[7] A. Seierstad, Stochastic control in discrete and continuous time. Springer, New York, 2009. xii+291 pp.

[8] V. Vasyunin, Cincinnati lectures on Bellman functions, https://arxiv.org/abs/1508.07668.