Optimal Control

Multistage Decision Processes in Economy and Finance

Margaréta Halická
Pavel Brunovský
Pavol Jurča

Bratislava 2015

The book presents a comprehensive exposition of the theory of optimal decision making in several stages. It shows how to use the theory to formulate and solve problems in economics, finance and management. From the reader it requires knowledge of mathematics at the level of mathematical, economic, finance or management study programs. It can be used as a textbook in those programs.

Contents

Preface to English Edition

The present text is a free English translation of the book *Optimálne riadenie: Viacetapové rozhodovacie procesy v ekonómii a financiách* published in 2009 in Slovak. The book has served as a text for the study program Mathematics of Economs and Finance at Comenius University in Bratislava. Thanks to the project *Preparation of the study of mathematics and informatics at Comenius University in English* supported by the Agency of the Ministry of Education, Science, Research and Sport of the Slovak Republic for the Structural Funds of EU we obtained an opportunity to translate the book into English and so, in accord with the intention of the project, to allow the students to study our program in English.

This English edition follows the contents of the Slovak book and uses the same enumeration of theorems and exercises. Nevertheless, the cultural background of the examples has occasionally been altered albeit without effect on the solutions. We have also included corrections of all errors and misprints we learned about during the 6 years of our use of the book.

Discrete time optimal control theory has received much less attention in the past than its continuous time counterpart. Some issues of the former are to our knowledge discussed in the present book for the first time. This is why we believe that this text may be of interest to a wider community of readers.

Bratislava, September 2015 Authors

Preface

A piece of wood floating on the Danube does not have control over the arm of the river delta it ends up. Fish does. It *decides* the arm promising more food or better living conditions for its offsprings. It chooses the best possibility, or, as we will say, it *optimizes.* Living organisms differ from inanimate materia among other things also in their capability to take decisions. The higher their evolutionary level, the more complicated decisions they are able to take.

Man as a species on the highest evolutionary level is most advanced in decision taking. He is distinguished by creativity – he is capable to take decisions in dilemmas he encountered never before. Sometimes, his common sense, intuition, is sufficient. At other instances it is more involved to take the right decision and is virtually impossible without a more thorough examination. Here mathematics becomes useful. Its part dealing with decision making and optimization is called *operations research.* It encompasses e.g. linear, nonlinear and integer programming, network optimization etc.

One of the branches of operations research is devoted to problems in which there is a need to make a sequence of decisions in several *stages* rather than a single one. Decision in one stage affects not only the result but also the constraints for the subsequent decisions. For instance, a rational and perspectively thinking individual in possession of larger financial amount lays aside a part for increasing its value in order to utilize it later. A good example can be found in the Old Testament:

Joseph ordered to store a part of the crop during the seven fertile years as a reserve for the coming infertile ones. An analogy to pension saving is obvious.

The branch of operations research dealing with multistage decision procedures was born in the second part of the last century. It is mostly called *optimal control theory*. This is because, as a rule, the variable representing the decision factor is called *control*. Alternatively, the theory is being called *theory of optimal processes*, *dynamic optimization* or *dynamic programming*. However, the last name is not completely adequate: dynamic programming is only one of the possible approaches to the problem.

The initial impetus to the birth of the theory came from automation and aeronautics in USA and the former USSR. This is also the reason why historically the *continuous* version developed earlier. This theory assumes that decisions are being taken continuously in time. The crucial result, *Pontryagin maximum principle* (L. S. Pontryagin), was developed in the USSR. Parallel to the Pontryagin theory, in the USA an alternative approach to the solution of optimal control problems has been developed. It was motivated largely by economic problems. It is attributed mainly to R. Bellman. Unlike Pontryagin's continuous theory it focuses primarily to decisions in separated *discrete* time instants, stages. It is based on a simple optimality principle and leads to a recursive procedure for the computation of *optimal control*, called *Bellman dynamic programming* method. Later the relation of dynamic programming to the Pontryagin maximum principle and to the calculus of variations has been discovered.

Initially, optimal control theory found its application mainly in engineering disciplines like aeronautics, chemical and electrical engineering, robotics. In the later decades it has found more and more applications in economic theory and computational finance, e. g. in macroeconomic growth theory, microeconomic theory of firm and consumer as well as in the management of investment portfolios. For an individual saving for his pension it is of interest to decide into which fund he should allocate his savings. In turn, those applications gave origin to further requirements on the development of specific components of the theory and have pushed the theory forward.

There is a number of books of various nature on optimal control theory. Some of them focus on rigorous presentation of the theory either in the discrete or in the continuous context, others on applications in various disciplines.

In our book we deal exclusively with *discrete* optimal control problems. As an important component, the book contains a rich spectrum of both solved and unsolved problems principally from the fields of economics and finance. The problems may come from models which are naturally discrete but may also be a discretization of continous ones.

What are the discrete optimal control problems, we learn in Chapter 1. Chapter 2 deals with dynamic programming, which is a useful tool for solving discrete problems numerically. One of its subchapters is devoted to stochastic dynamic programming. Chapter 3 deals with the maximum principle, which has its origin in the continuous theory. Although its validity in the discrete theory is limited, it is still a useful tool for qualitative analysis. The last Chapter 4 consists of appendices summarizing results from nonlinear programming, difference equations and probability needed in Chapters 2 and 3.

The book is primarily meant for a reader who would like to master the basics of discrete optimal control theory in depth. Unlike in many other textbooks presenting concepts and principles in a rather intuitive level, this book strives for mathematical rigor and all crucial conclusions are mathematically justified. Organization of the text allows the reader to skip some of the proofs, or, to restrict himself to conclusions for simpler situations. To master the text the reader needs knowledge of the foundations of mathematical analysis of functions of several variables; knowledge of probability theory and nonlinear programming being an advantage, but not a necessity, all needed concepts and conclusions from those areas are outlined in the appendices.

We strived to illustrate the theory and the tools developed in it on a wide spectrum of specific problems originating in economics, business and finance. Problems of optimal consumption, optimal allocation of funds into several investment projects, optimal investment into the particular funds of the pension saving, or several kinds of optimal renewal problems may serve as examples. As a rule the problems are simplified to such an extent that their solutions are not overly time consuming. In such a way theory can be understood better. Realistic problems can frequently not be solved in such a simple manner, in particular in the presence of uncertainty. Therefore, the book contains sample problems exhibiting how theory can be employed to write a computer code capable to solve more involved problems.

The purpose of some of the examples is to indicate variability of problems which can be solved by the methods discussed in the book. They should help the reader to get an intuition to which kinds of problems the theory can be applied. Simultaneously, in Chapter 1 we discuss in detail how the problems from practice should be formulated in a way allowing us to solve them by the presented theory. Some of the problems appear in several modifications through the text. This helps to compare various solution methods. The optimal consumption problem may serve as an example. Extended space is dedicated to it, partially due to its importance in current dynamic macroeconomic models. A part of the examples is solved completely in the text, solutions of others are left to the reader.

Bratislava, February 2009 Authors

Chapter 1

Introduction to Discrete Optimal Control Theory

In this chapter we specify the subject of discrete optimal control theory. We begin by several examples which allow us to comprehend the general formulation of the standard optimal control problem.

1.1 Examples

Example 1.1. Optimal allocation of resources. An investor owns capital of initial amount $a > 0$ and has two alternatives how to invest it: into oil drills and into real estates. At the beginning of every year he needs to decide how to allocate his disposable capital into those two investment alternatives. Capital of amount y, invested into oil drills yields yearly profit gy with $g > 0$; invested into real estate, it yields profit hy with $h > 0$. Profit is not being invested any more. Invested capital depreciates by yearly rate $0 \leq b < 1$ in the case of oil drills and by the rate $0 \leq c < 1$ in the case of real estate. The sum of those two capitals net of depreciation at the end of each year is the capital to be invested in the next year. The goal of the investor's decisions is to maximize total profit during the planning period of k years.

Label the years of the planning period consecutively by $i =$

$0, \ldots, k-1$. Denote x_i the amount of the disposable capital at the beginning of the i-th year. Let $u_i \in [0,1]$ be the ratio of the amount of capital to be invested into oil drills to its total amount x_i. That is, from the amount x_i to be invested in the i-th year, the part $u_i x_i$ is invested into oil drills and the remainder $(1-u_i)x_i$ into real estate. This decision yields at the end of the year profit $gu_i x_i + h(1-u_i)x_i$, the amount $x_{i+1} = bu_i x_i + c(1-u_i)x_i$ will remain for investment in the $(i+1)$-th year. The goal of the investment decisions is to

$$\text{maximize} \quad \sum_{i=0}^{k-1} [gu_i x_i + h(1-u_i)x_i] \tag{1.1}$$

$$\text{subject to} \quad x_{i+1} = bu_i x_i + c(1-u_i)x_i, \; i = 0, \ldots, k-1, \tag{1.2}$$

$$x_0 = a, \tag{1.3}$$

$$u_i \in [0,1], \quad i = 0, \ldots, k-1, \tag{1.4}$$

where a, b, c, g, h are given constants and maximum is to be found with respect to the variables $u_i, \; i = 0, \ldots, k-1$ and $x_i, \; i = 0, \ldots, k$.

We have obtained a mathematical programming problem of a particular structure characteristic for a multistage decision procedure. The decision procedure (of reallocation and investment of capital in our case) is divided into *stages* (years in our case), consecutively labeled by $i = 0, \ldots, k-1$. The state of the process at the beginning of the i- th stage is represented by the *state variable* x_i (capital assigned for investment). The process is subject to external decisions u_i (defining the allocation of capital between the two investment alternatives) called *control variable*. The values of the state and control variables in the i-th stage determine the revenue of the i-th stage (the i-th summand in the objective function (1.1)) as well as the value of the state variable at the beginning of the next $(i+1)$-th stage (the right hand side of the difference equation (1.2)). Let us note that the decision maximizing total profit need not maximize partial profits of individual stages.

Classification of optimization variables into state and control ones, a certain separability of the objective function as well constraints in the form of a difference equation will be present in the next example as well.

Example 1.2. Optimal consumption. Assume that we own capital the volume of which at the beginning of the i-th month we denote by x_i. During the i-th month the capital x_i yields interest rx_i. At the end of each month we can decide which part of the available capital we consume. Consumption of the amount u_i provides us with (discounted) utility $(\frac{1}{1+\delta})^i \ln u_i$, where $\delta > 0$ being the measure of impatience of consumption in time.[1] We assume that at the beginning of the process we possess the amount $a > 0$ of capital. The goal is to determine consumption during each month in such a way that the value of capital at the end of the planning period of k months reaches a prescribed value $b \geq 0$ and the total discounted utility of consumption during the planning period is maximal.

Mathematically, we end up with the problem

$$\text{maximize} \quad \sum_{i=0}^{k-1} \left(\frac{1}{1+\delta}\right)^i \ln u_i \tag{1.5}$$

$$\text{subject to} \quad x_{i+1} = (1+r)x_i - u_i, \ i = 0, \ldots, k-1, \tag{1.6}$$

$$x_0 = a, \tag{1.7}$$

$$x_k = b, \tag{1.8}$$

where a, b, δ, r, are given constants and the maximum is sought with respect to the variables $u_i, \ i = 0, \ldots, k-1$ and $x_i, \ i = 0, \ldots, k$.
Also in this case the problem can be understood as a k-*stage decision process*, with months as stages. The state variable is represented by x_i; the control variable by u_i - it should be understood as an input by the choice of which we influence the development of the process. The

[1] By the choice of an increasing and concave function $\ln u_i$ as a measure of utility we express that a unit increase of consumption represents for the consumer a smaller increase of utility at a higher consumption value than at a smaller one. By the discount factor we express that the consumer prefers instantaneous consumption to a later one.

behavior of the process during the transition from the i-th stage to the $(i+1)$-th one is described by the *difference equation* (1.6), the right hand side of which again depends on x_i and u_i only – the values of the state and control variables in the i-th stage. *The objective function* sums the discounted utilities of the particular stages, the utility of each particular stage i depending only on x_i and u_i.

As in the preceding example, the objective function (1.5) exhibits a certain separability and the dynamics of the system is described by the difference equation (1.6), completed by the initial condition (1.7). Unlike in Example 1.1, the state variable is subject to a terminal condition (1.8) but we have no constraint on the control variable of type (1.4).

1.2 Standard Form of the Optimal Control Problem

In this subchapter we formulate a general optimization problem which covers Examples 1.1 and 1.2 as special cases.

Assume that we have an object the behavior of which we control in the course of k stages. The state of the system at the beginning of stage i, $i = 0, \dots, k-1$, is described by the state variable $x_i \in X_i$. The behavior of the object in the i-th stage we control by the control variable $u_i \in U_i$, the input to the system. Here X_i is a given set of admissible states and U_i is the set of admissible values of the control variable in the i-th stage. The values of x_i and u_i determine uniquely the value of $x_{i+1} := f_i(x_i, u_i)$, where f_i is a given function. The yield in the i-th stage is determined by the value $f_i^0(x_i, u_i)$ with f_i^0 given. We assume that at the beginning the value of the state variable x_0 is equal to a given value a and we require that x_k, the value of the state variable at the end of the process, belongs to a given set of terminal states C. The goal is to determine in each stage the value of the control variable u_i in such a way that all the conditions are satisfied and the sum of all yields of the particular stages is maximal.

The problem can be written as follows:

$$\begin{aligned}
\text{maximize} \quad & \sum_{i=0}^{k-1} f_i^0(x_i, u_i) && (1.9)\\
\text{subject to} \quad & x_{i+1} = f_i(x_i, u_i),\ i = 0, \ldots, k-1, && (1.10)\\
& x_0 = a, && (1.11)\\
& x_k \in C, && (1.12)\\
& u_i \in U_i, \quad i = 0, \ldots, k-1, && (1.13)\\
& x_i \in X_i, \quad i = 0, \ldots, k-1. && (1.14)
\end{aligned}$$

Maximum is to be found with respect to the variables $u_i,\ i = 0, \ldots, k-1$ and $x_i,\ i = 0, \ldots, k$. In this formulation the variables x_i and u_i seemingly have the same status. However, observe that by the choice of the control variables u_i the variables x_i become uniquely determined as the solution of the difference equation (1.10) with initial value (1.11). Both in the terminology and in the solution methods this circumstance is reflected in the understanding of (1.9)–(1.14) as optimal control problem.

Let us now formulate the problem (1.9)–(1.14) as an optimal control one. To this end we need to define the concepts of control, its response, admissible and optimal control.

By a *control* we call a sequence of values of the control variables $\mathcal{U} = \{u_0, \ldots, u_{k-1}\}$ satisfying $u_i \in U_i$ for all $i = 0, \ldots, k-1$. By the *response* to the control $\mathcal{U}$ for a fixed initial value (1.11) we understand the sequence, $\mathcal{X} = \{x_0, \ldots, x_k\}$, where the values $x_i = x_i(\mathcal{U})$ solve the *state equation* (1.10) with given control $\mathcal{U}$ and *initial condition* (1.11). Obviously, the response to a particular control may, but may not satisfy the constraints (1.12) and (1.14). If it does, then the control is called admissible. This means that an *admissible control* and its response satisfy all the constraints of the problem (1.9)–(1.14). The class of admissible controls we denote by $\mathcal{P}$. The value of the objective function in (1.9) can now be understood as a function of the control $\mathcal{U}$. Hence, we denote it by

$$J(\mathcal{U}) := \sum_{i=0}^{k-1} f_i^0(x_i(\mathcal{U}), u_i).$$

The optimal control problem is to find among all admissible controls $\mathcal{U}$ the one for which the objective function $J(\mathcal{U})$ reaches its maximal value. Such a control is called *optimal*. Formally, we write the problem as follows:

$$\max_{\mathcal{U}\in\mathcal{P}} J(\mathcal{U}).$$

Remark 1.1. Problems with *minimum* instead of maximum can be rewritten in the standard form by replacing the objective function by its negative.

Remark 1.2. As a rule, the variable $i = 0, \ldots, k-1$ labeling the decision stages has commonly the nature of time, therefore we call it accordingly. In addition to the values of control and state variables, the functions f_i, f_i^0 and the sets U_i a X_i (i.e. the data of the problem) may vary with the stages as well. Such problems are called *non-autonomous*. In case the data do not depend on i, the problem is called *autonomous*. To distinguish an autonomous problem from a non-autonomous one will be particularly important in the next chapter. Observe that Example 1.1 leads to an autonomous problem, whereas Example 1.2 leads to a non-autonomous one, because $f_i^0(x_i, u_i) = (1+\delta)^{-i} \ln u_i$ depends on i.

Remark 1.3. Occasionally the constraint (1.13) or (1.14) is absent. In such case we refer to a *problem without control constraints* or a *problem without state constraints*, respectively. The control and state variables can then attain arbitrary values from the natural domains of definition of the functions f_i and f_i^0. Example 1.1 leads to a problem with control constraint (cf. (1.4)) but without state constraints. Example 1.2 leads to a problem without constraints on either control or state. The natural domain of definition of the function ln which appears in (1.5) limits the domain of the control variable to real positive $u_i > 0$. In both the motivating examples the functions f_i a f_i^0 are defined on Euclidean spaces or their positive sectors and hence u_i and x_i have a continuous nature. However, we will encounter also problems in which u_i and x_i will be of discrete nature, either due to constraints (1.13), (1.14) or due to the natural domain of definition of the functions f_i and f_i^0.

Remark 1.4. The constraint (1.12) concerns the terminal stage. In case the problem does not include such a constraint, i. e. there are no restricitions on the terminal value x_k, we refer to such a problem as to a *free endpoint problem.* In case C consists of a single point, we refer to a *fixed endpoint problem.* In all the remaining cases we refer to a *partially restricted endpoint problem.* Obviously, Example 1.1 leads to a free endpoint problem, while Example 1.2 to a fixed endpoint one.

Remark 1.5. In all the motivating examples mentioned in this chapter both the control and the state variables have a single component only, i.e. they are *one-dimensional.* In general, they may be *higher dimensional.* Such a problem can be obtained from Problem 1.4; a number of higher dimensional problems will be encountered in the following chapters.

Remark 1.6. In the standard problem we maximize the sum of yields from the particular stages which leads to an objective function of form (1.9). The optimal control problem in this form is being called a problem in Lagrange form in the literature. However, multistage decision processes may lead to a need to maximize also a function of the terminal state. In such case the objective function has the form $\sum_{i=0}^{k-1} f_i^0(x_i, u_i) + \phi(x_k)$, or only $\phi(x_k)$, where ϕ is a given function, and the corresponding problem is called a problem in *Bolza* form, or in *Mayer* form, respectively.

Remark 1.7. Frequently, particularly in economic problems, the yield $f_i(x_i, u_i)$ from the i-th stage has the form $f_i(x_i, u_i) = \beta^i F_i(x_i, u_i)$, where $F_i(x_i, u_i)$ is the yield of the i-th stage and $\beta \in (0, 1)$ is the coefficient adjusting (discounting) this value to the time of the beginning of the process, i.e., to the time $i = 0$. The factor β can reflect the loss of value caused by inflation, by not having invested the amount to interest bearing deposit with interest $r = \frac{1}{\beta} - 1$ for the i-th stage (cf. Problem 1.17), as well as propensity to prefer earlier consumption, if $\beta = \frac{1}{1+\delta}$, as in the optimal consumption problem in Example 1.2. The problem in which the objective function is in the special form $J = \sum_{i=0}^{k-1} \beta^i F_i(x_i, u_i)$ we call *problem with discount factor*. Since the value of the objective function expressed in this way measures the present value of future yields, it is sometimes called *present value objective function*.

Obviously, problems with discount factor are non-autonomous in the sense of Remark 1.2. If, however, the problem is non-autonomous solely because of discount factor, we call the problem *autonomous with discount factor*. Special treatment of such problems is justified because they share many essential features with autonomous ones. An example of an autonomous problem with discount factor is Example 1.2.

Remark 1.8. In the standard problem (1.9)–(1.14) the number k of stages is fixed. Therefore, we call it a *fixed time problem*. However, there are problems in which k is one of the variables with respect to which maximum is sought. Such problems are called *free time problems*. In free time problems, each finite sequence of control variables of arbitrary length is a control. Economic applications frequently lead to *infinite horizon problems*, in which infinite sequences serve as controls. We encounter free time and infinite horizon problems in the next chapter. Both the Problems 1.1 and 1.2 are fixed time problems.

We conclude this subchapter by a remark on notation and formulation of problems.

Remark 1.9. It is frequently comfortable to call a sequence $\{j, \dots k\}$ an *interval* and denote it by $[j, k]$. Sometimes we will shorten the formulations by replacing the terms *maximize*, resp. *minimize* by *max*, resp. *min* and drop completely the formulation *subject to*. The conditions we either separate from the objective function by a colon or put them in a new row below the objective function.

1.3 Further Examples

Now, we present several examples documenting the variability of problems admitting optimal control formulation. We begin by a simple problem in which the admissible values of both control and state variables belong to finite sets.

Example 1.3. Container transportation. A firm transports containers from the railway station to its storage. It has a lorry which can be sent to the railway station at most once a day and loaded by at most 2 containers. The firm is informed that in the days of the coming week it will receive in a row 1, 3, 1, 2, 1 containers for transportation. A trip to the station costs EUR 70, penalty for an uncollected container costs EUR 50 per day. No container remained from the previous week and at the end of the week all the containers have to be transported. It should be decided at which days the lorry should be ordered so as to make the sum of the costs of the drives and penalties minimal.

In this problem the stages correspond to the days of the week to be labeled by $i = 0, \dots, 4$. We denote by

- x_i the number of containers waiting from previous days for transport from the railway station,
- a_i the number of containers received by the railway station in the morning of i-th day,
- u_i the number of containers we decide to transport from the railway station in the afternoon of i-th day.

Since at most two containers per day can be transported from the railway station, we have $u_i \in \{0, 1, 2\}$. Obviously, the decision $u_i = 0$ means no lorry ride and, thus, no transport cost, whereas the decisions $u_i = 1$ or $u_i = 2$ mean the lorry ride to the station of the cost of 70. After the ride, at the end of day i, $x_{i+1} := x_i + a_i - u_i$ containers remain waiting for transportation. The costs for this day will consists of the penalty $50(a_i + x_i - u_i)$ for the waiting containers and $70\text{sgn}(u_i)$ for the lorry ride. Here $\mathrm{sgn}(u)$ denotes the function which equals 1 for u positive and 0 for $u = 0$.

Wee have obtained the problem

$$\text{minimize} \quad \sum_{i=0}^{4} [70\mathrm{sgn}(u_i) + 50(a_i + x_i - u_i)] \tag{1.15}$$

$$\text{subject to} \quad x_{i+1} = x_i + a_i - u_i, \quad i = 0, \ldots, 4, \tag{1.16}$$

$$x_0 = 0, \quad x_5 = 0, \tag{1.17}$$

$$u_i \in \{0, 1, 2\}, \tag{1.18}$$

$$x_i \geq 0, \quad i = 0, \ldots, 4, \tag{1.19}$$

where the constraint $x_i \geq 0$ expresses the condition that the number of transported containers cannot exceed their actual number at the station, and the condition $x_5 = 0$ means that at the end of the week all containers have to be transported from the station.

This problem represents a fixed time non-autonomous problem with control and state constraints. Note that from the conditions of the problem it follows that the values of the state variable can only be nonnegative integers which are bounded by the number of all received containers from above.

The following two examples are particulary interesting in that the stages, unlike in the previous examples, do not represent time instants. The first example belongs (similarly as Example 1.1) again to the wide class of resource allocation problems. The second one represents an extremely simple case of a wide class of *optimal renewal problems*. It is instructive also because the control variable values have a logical rather than quantitative nature.

Example 1.4. Optimal allocation of funds. There are k investment plans how to use funds of amount S. An amount u_i invested into plan i yields revenue $h_i(u_i)$, $i = 0, \ldots, k-1$. It should be decided how to allocate the funds among the particular investment plans so as to maximize the total revenue, i.e.,

$$\text{maximize} \sum_{i=0}^{k-1} h_i(u_i) \quad \text{subject to} \sum_{i=0}^{k-1} u_i \leq S.$$

Note that while in the previous examples the decision process was clearly divided into time stages, in this problem this is not the case. Therefore, in order to put this problem into the optimal control form we need to replace time by something else: we will understand the simultaneous decision about investment into k plans as a sequence of decisions, first into the first plan, then into the second plan etc.

Before the i-th stage investment decision, $i = 0, \ldots, k-1$, we take into account the amount x_i of funds which is still available for investment. The decision consists in the choice of the amount u_i of funds to be invested into plan i. Afterwards $x_{i+1} = x_i - u_i$ of funds remains for investment into further plans. The condition $x_0 = S$ reflects the total amount of funds and the condition $x_k \geq 0$ secures that in none of stages we invest more than what is currently available.

We obtain the optimal control problem

$$\text{maximize} \quad \sum_{i=0}^{k-1} h_i(u_i) \tag{1.20}$$

$$\text{subject to} \quad x_{i+1} = x_i - u_i, \quad i = 0, \ldots, k-1, \tag{1.21}$$

$$x_0 = S, \tag{1.22}$$

$$u_i \geq 0, \tag{1.23}$$

$$x_k \geq 0. \tag{1.24}$$

This is a fixed-time non-autonomous problem with control constraints, without state constraints, with partially fixed endpoint, where $C = \{x : x \geq 0\}$.

Example 1.5. Optimal machine maintenance. Fruit to be pressed arrives packed in boxes, an open box has to be pressed immediately. The pressing machine clogs, depending on the amount pressed its productivity decreases. More precisely, after each pressed box the pressing time of the subsequent one increases by 20 minutes. Cleaning of the pressing machine takes 30 minutes. Altogether, k boxes have to be pressed. It should be decided whether and, if yes, after which box the machine should be cleaned so that the required number of boxes is processed in the shortest possible time. We assume to start with $a \geq 0$ boxes pressed since the last cleaning and to end with the machine possibly uncleaned.

We introduce the following notations:
i – the sequence number of the box to be processed, $i = 0, \ldots, k-1$,
x_i – the number of processed boxes after the last cleaning,
$u_i = C$ – if the machine is to be cleaned before pressing the i-th box,
$u_i = N$ – otherwise.

In this example the i-th stage represents the part of the process beginning after the $(i-1)$-th box has been pressed and ending with completion of pressing the i-th one. We now describe the time sequence of the particular steps in the i-th stage. At first we have the information x_i about the clog level of the machine. Then, we choose the value of u_i, i.e. we decide whether to clean it ($u_i = C$) or to continue pressing ($u_i = N$). If $u_i = C$ is chosen, then the machine will be cleaned, otherwise not. Then, pressing of the i-th box follows and we remember that after the last cleaning the amount of x_{i+1} boxes has been pressed, where $x_{i+1} = x_i + 1$, if we have not cleaned, $x_{i+1} = 1$ otherwise. We also record the time losses: 30 in the case of cleaning, $20x_i$ otherwise. Obviously $x_0 = a$. Our goal is to find such values of u_i that the total time loss is minimal.

We have obtained the optimal control problem

$$\text{minimize} \qquad \sum_{i=0}^{k-1} f^0(x_i, u_i),$$

$$\text{subject to} \qquad x_{i+1} = \begin{cases} x_i + 1, & \text{if } u_i = N, \\ 1, & \text{if } u_i = C, \end{cases} \qquad i = 0, \ldots, k-1,$$

$$x_0 = a.$$

where

$$f^0(x,u) = \begin{cases} 30, & \text{if } u = C, \\ 20x, & \text{if } u = N. \end{cases}$$

This is a free endpoint autonomous problem without control and state constraints. The control variable has a logical nature, the values of the state variable are positive integers.

1.4 Basic Tasks in Discrete Optimal Control Theory

Obviously, the primary goal of optimal control theory is to develop effective methods to find solutions of problems – optimal controls. However, what should we understand by the solution method? As we will see, even fairly simple problems often do not allow a closed form solution. Existence of powerful computing tools may lead to the idea that by discretization of all continuous quantities the problem could be transformed to optimization of a function with finite number of values which could be solved a systematic and exhaustible search of all candidates. We show that this road is not feasible because of the prohibitively large extent of the resulting problem. The difficulties of this approach as well as possible ways out we demonstrate on the problem of optimal machine maintenance (Example 1.5).

Example 1.6. Optimal machine maintenance from Example 1.5. This problem is naturally discrete, the control $\{u_i\}_{i=0}^{k-1}$ is of logical nature with $u_i \in \{C, N\}$. Thus the total number of admissible controls for this problem is 2^k. The exhaustive search method would therefore be of exponential complexity. Indeed, it would require to compute the value of the objective function for each admissible control. For large values of k this leads to a problem which cannot be mastered even by the most powerful computing technology. It is therefore natural to look for more sophisticated solution procedures. Let us try to do so.

Recall that the exhaustive search method requires a complete computation of the objective function to be repeated in each step, even if e.g. merely the last member u_{k-1} of the sequence $\{u_i\}_{i=0}^{k-1}$ has changed. This shortcoming can be taken care of by computing the objective function "from behind". For instance, when comparing the controls $\{C,\dots,C,C\}$ and $\{C,\dots,C,N\}$ we can use that for both controls the state x_{k-1} will be the same ($x_{k-1}=1$) and in this state it is optimal not to clean the machine. It is therefore sufficient to remember that as long as the machine is in state $x_{k-1}=1$ before pressing the last box, the smallest time loss from pressing it is 20 minutes. This is the case if the machine is not being cleaned, i.e. $u_{k-1}=N$. Analogically it can be concluded that if the machine is in state $x_{k-1}=2$ before pressing the last box, it is optimal to clean ($u_{k-1}=C$) and the smallest time loss being 30 min. We can equally compute the smallest possible time loss (to be denoted by V_{k-1}) before pressing the last box and the optimal value u_{k-1} for all remaining values of x_{k-1}. One can similarly proceed further: if, e.g., the machine is in state $x_{k-2}=1$ and we clean it, the least possible time loss is 30+20 minutes. Indeed, by cleaning, the machine gets into state $x_{k-1}=1$ for which the smallest possible loss is 20 min. In case we do not clean the machine, the smallest losses from pressing the last two boxes will be 20+30 min. We see that in this case there are two optimal controls with values $u_{k-2}=C$ as well as $u_{k-2}=N$.

Solution of the problem for $x_0=1$ and $k=3$ by the exhaustive search method is illustrated in Figure 1.1. The nodes represent the states of the machine (i.e. x_i) and the smallest losses from pressing the remaining boxes. In each node we can decide whether to clean the machine (dashed arrow) or not (full arrow). The number attached to the arrow gives additional time needed to press the next box. By the bold arrow we represent the optimal decision. In this algorithm decisions have to be made in all 2^k-1 nodes. Complexity of the algorithm is therefore exponential, although computation of the value of the objective function is simpler.

However, note that in the last row the decision in the state $x_{k-1}=1$

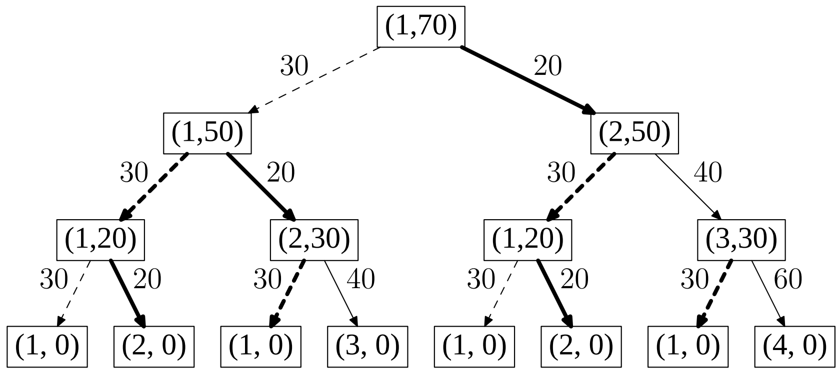

Figure 1.1: Solution of the optimal machine maintenance problem, exponential algorithm

appears twice and, naturally, is in both cases the same. Indeed, if the press is in state x_i before pressing the i-th box, subsequent decisions will not depend on the way it got into that state. As we will see later, this "Markov" property is crucial for the dynamic programming solution method to be explained in the next chapter. In our case it implies that all the nodes with the same state in one row can be merged. This approach to the solution is illustrated in the scheme of Figure 1.2. Note that for each particular i, the value of x_i has to belong to the set $\{1, \ldots, i+1\}$. The number of mutually distinct nodes in which a decision has to be taken is therefore $\frac{(k+1)k}{2}$, the proposed algorithm thus has polynomial complexity. For a large value of k the computation can be significantly less time consuming.

Note that for problems in which the variables x_i a u_i are of continous nature, the exhaustive search algorithm is equivalent to the search of the extreme of a function of k variables, whereas the newly proposed algorithm searches in each stage the minimum of a function of a single variable employing the relation (cf. Problem 1.2)

$$\max_{x_1,x_2}[g(x_1) + h(x_1, x_2)] = \max_{x_1}[g(x_1) + \max_{x_2} h(x_1, x_2)]. \tag{1.25}$$

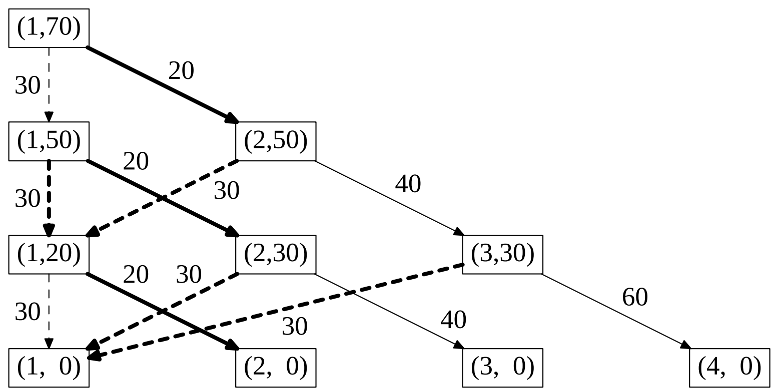

Figure 1.2: Solution of the optimal machine maintenance problem, polynomial algorithm

Discrete optimal control theory attempts to develop methods which would (similarly as in the above example) utilize the special nature of optimal control problems and thus allows to solve the problems effectively.

This is not all, though. As a further goal the theory studies properties of optimal controls. This means that we try to conclude something about the optimal control directly from the nature of the problem without actually knowing its numerical values. This approach is similar to the one of stability theory of linear differential equations: we can completely decide about stability of an equilibrium from the coefficients of the equation without having to compute its solutions. This is important not only because we are rarely able to compute the optimal control but also because frequently it helps to restrict radically the set of candidates.

A reader familiar with nonlinear programming may recognize the analogy with Kuhn-Tucker conditions. This analogy is justified, necessary conditions of optimality we develop in Chapter 3 are in essence Kuhn-Tucker conditions utilizing the specific structure of optimal control problems.

1.5 Exercises

Exercise 1.1. Using an arbitrary method solve the problem from Example 1.1 about optimal allocation of resources for the case $k = 2,\ b = 0.7,\ c = 0.4,\ g = 2,\ h = 2.5$ and justify the correctness of the solution.

Exercise 1.2. Under the assumption of existence of all maxima in (1.25) prove equality (1.25).

Exercise 1.3. Show that an optimal control problem in Mayer form can be transformed to a problem in Lagrange form and vice versa. Hint: for the proof of the reverse implication introduce an additional state variable.

Exercise 1.4. Consider the optimal control problem (1.9)–(1.14) with additional condition

$$\sum_{i=0}^{k-1} g_i(x_i, u_i) = 0.$$

Because of this condition the problem does not have the standard form. By introducing an additional variable reformulate the problem in a standard form. (Hint: see Example 1.4 on the optimal allocation of funds.)

Exercise 1.5. Optimal herd-keeping at the ranch. A rancher Pacho keeps a herd of cattle with a units. If he sells y units at an annual market he gets for them $\phi(y)$; the rest reproduces $b > 1$ times. Pacho plans to ranch for 10 years and naturally wants for these 10 years achieve the greatest possible profit.

Denote x_i the state of the herd before the market in the i-th year and formulate the problem as an optimal control problem in the two variants: in the first one, denote $u_i \in [0, 1]$ the proportion of herd that Pacho brings to the market in the i-th year; in the second one denote u_i the number of units of cattle Pacho brings to the market in the i-th year. Discuss the advantages of each of these two formulation options to this situation.

Exercise 1.6. Compare both formulations of the problem from the previous exercise with Example 1.2. Reformulate the problem from Example 1.2 in a form, where the control expresses a proportion of the capital for consumption. Describe the nature of the problem thus formulated.

Exercise 1.7. Transport operator. Formulate the following problem as an optimal control one. Try to guess its solution.
For the coming (5 days) week, the transport operator expects the arrival of 10, 3, 7, 2, 5 vans in the consecutive days, which should be unloaded. The vans are unloaded by a team which unloads a van in two hours. In case the team works more than 8 hours in one day, for each additional hour it receives extra payment of 15 EUR. For a van which failed to be unloaded in the day of its arrival a penalty of 20 EUR daily has to be paid. It should be decided what number of vans which day should be unloaded in order to minimize the sum of extra payments and penalties. It is assumed that the team works even numbers of hours a day (or, equivalently, that from one day to another no van remains partially loaded).

Exercise 1.8. Formula 1. In the Formula 1 race, the average time passing one round increases depending on the wear of tires. Suppose that the additional time can be expressed in seconds by $0.001n^2$, where n is the number of full rounds from the last tire replacement. Replacement takes 10 seconds. When is it worth to replace the tires? Formulate in a standard form of discrete optimal control problem.

Exercise 1.9. Cash holding optimization. The First Bank of Šariš with headquarters in Prešov has to take care that its branch in Snina has enough cash for withdrawals of its clients. For the next week it received requests for large withdrawals (in thousands of EUR) given by Table 1.1.

In the case of need, cash is supplied by being driven from the headquarters to the branch office before its opening hours. A supply ride costs

Table 1.1: Cash withdrawal requests

Day	Mo	Tu	We	Th	Fr
Withdrawal	200	250	130	420	310

100 EUR independently of the amount of cash driven. For safety reasons it is not allowed to transport more than 1 mil. EUR at once. Holding excessive amount of cash in the branch office overnight leads to losses of the bank because it can not be placed to the interbank market. Assume that the loss is 0,02% per day. Which days should the cash be supplied? Formulate as a standard discrete optimal control problem.

Exercise 1.10. Vans unloading. Tomorrow morning the operator receives 10 vans of deteriorating raw material which has to be processed by 3 days. There are losses in the process of unloading: if 3 vans a day or less are unloaded, the loss is 5%; if 4 resp. 5 vans (maximum possible) a day are unloaded, the loss is 10%, resp. 15% . On the other hand, if not unloaded, the material deteriorates: 10% during the first and 20% during the second night. Formulate as a standard discrete optimal control problem.

Exercise 1.11. Optimal replacement of a car. Let $\psi(x)$ resp. $\varphi(x)$ be the value resp. cost of maintenance of a certain type of car of age x. Assume that during the planning horizon k at the end of which we sell our car the same type of car continues to be produced and its price and maintenance costs does not change. It should be decided when to replace our car by a new one in such a way that the total costs during the planning horizon k are minimal. We further require that we never hold a car for more than 5 years and that the replacement of the car takes place at the beginning of the year. Formulate as a standard discrete optimal control problem.

Exercise 1.12. Optimal scheduling of general repairs. The cost of general repair (GR) of a machine is 1 000 EUR. By using the machine its quality deteriorates which leads to losses due to defective products. Those losses can be estimated by $200(n+1)$ EUR per year, where n is the number of entire years elapsed from the last GR. When should the machine be scheduled for GR if we would like to use it for 5 years and the most recent GR was yesterday? Formulate as a standard discrete optimal control problem.

Exercise 1.13. Harvesting operator. The harvesting operator has to harvest crop from a 10 hectares field in three days. He is able to harvest an area of five hectares per day, but the speed of harvesting increases the grain losses as follows:
3 hectares or less of harvesting per day, the loss is negligible,
4 hectares of harvesting per day, the loss is 5 percent,
5 hectares of harvesting per day, the loss is 15 percent.
However, the losses also increase because the grain matures and drops out from ears, as follows:
if harvested the first day, the loss is 10 percent,
if harvested the second day, the loss is 15 percent,
if harvested the third day, the loss is 20 percent.
Assuming that the two types of daily losses are added, the goal is to determine, what area of the field has to be harvested each day in order that the total losses are minimal. Formulate as a standard discrete optimal control problem.

Exercise 1.14. Optimal scheduling of orange orders. The grocer orders oranges from the wholesaler for the months of September, October and November. The wholesale and retail prices are given by Table 1.2. The grocer estimates monthly demands to be consecutively 2, 3 and 4 tons. In his sale stock the grocer can store at most 3 tons. If he wants to store more he has to rent a store, the price per month and capacity of which is 1000 EUR and 4 tons, respectively. At the begin of September he has a stock of 1 ton, at the end of November his sale stock should be empty. The deliveries from the wholesaler take place at the beginnings of the months in entire tons only. Formulate as a standard discrete optimal control problem.

Table 1.2: Wholesale and retail prices of oranges

	September	October	November
Wholesale price (EUR/kg)	0.5	1	1.5
Retail price (EUR/kg)	1	1.5	2

Exercise 1.15. Formulate the following modification of Example 1.5 into an optimal control problem: at the beginning of the process the machine is clean and at the end it has to be cleaned.

Exercise 1.16. Optimal allocation of time between work and study. Each semester a university student decides how to partition his time between study (e.g. attendance of lectures) and work for earning money. He faces a similar dilemma after graduation as well. His level of education increases his salary in the future but the level of education decreases in time by a constant rate. The goal is to maximize the total discounted earnings during his entire planning period. Formulate as a standard discrete optimal control problem. Try to specify in more detail the nature of the functions of the formulation (increasing, linear, convex) in such a way as to reflect reality most adequately.

Exercise 1.17. In the i-th period ($i > 0$) we receive an income of D units. What is its present value (at the zero period), if we assume that during one period, one unit earns interest rate r? The result of this exercise was used in Remark 1.7.

Chapter 2

Dynamic Programming for Discrete Problems

Life must be lived forward
and understood backwards.
Soren Kierkegaard.

In this chapter we deal with the first approach to optimal control problems called *dynamic programming*. In essence it consists in an imbedding of the problem to be solved into a family of truncated problems. The solution of the original problem is obtained by solving consecutively truncated optimization problems, using the solution of a truncation in the next step. Such a procedure has been employed in Example 1.6 solving the problem of optimal maintenance of a machine from Example 1.5. It results in a formula called *Bellman's dynamic programming equation.* By its use we can achieve essential savings compared to the exhaustive search method.

We first develop the dynamic programming equation for fixed time problems. Then we extend this tool to other types of problems. In the last subsection we discuss problems subject to random effects. Such problems we call stochastic optimal control problems.

2.1 Problems with Fixed Terminal Time

Consider a fixed time problem which differs from the standard one (1.9)–(1.14) only in the labeling of the initial state. That is, we consider the problem

$$\text{maximize} \qquad J(x,\mathcal{U}) := \sum_{i=0}^{k-1} f_i^0(x_i, u_i), \tag{2.1}$$

$$\text{subject to} \qquad x_{i+1} = f_i(x_i, u_i), \quad i = 0, \ldots, k-1, \tag{2.2}$$

$$x_i \in X_i, \quad i = 0, \ldots, k-1, \tag{2.3}$$

$$u_i \in U_i, \quad i = 0, \ldots, k-1, \tag{2.4}$$

$$x_0 = x, \; x \text{ is fixed in } X_0, \tag{2.5}$$

$$x_k \in C. \tag{2.6}$$

Recall that by a *control* we understand a sequence $\mathcal{U} = \{u_0, ..., u_{k-1}\}$, where $u_i \in U_i$, for all $i = 0, \ldots, k-1$. By the *response* $\mathcal{X} = \{x_0, \ldots, x_k\}$ to the control $\mathcal{U}$ we understand the solution of the *state equation* (2.2) for the chosen $\mathcal{U}$ and the given initial state $x_0 = x$ at time 0. In case the response $\mathcal{X} = \{x_0, \ldots, x_k\}$ to the control $\mathcal{U}$ satisfies all constraints, i.e. $x_i \in X_i$, $i = 0, \ldots, k-1$ a $x_k \in C$, the control $\mathcal{U}$ is called *admissible.*

2.1.1 Imbedding of the Problem and the Value Function

Referring to the initial time $j = 0$ and the initial state $x_0 = x$ we denote the problem (2.1)–(2.6) as $D_0(x)$ and call it as *problem of optimal transition from the point x to the set C on the interval* $[0, k]$. We imbed it into the system of problems

$$\mathcal{D} = \{D_j(x) : \quad j \in [0, k-1], \quad x \in X_j\},$$

where $D_j(x)$ is the problem

$$\text{maximize} \qquad J_j(x,\mathcal{U}_j) := \sum_{i=j}^{k-1} f_i^0(x_i, u_i) \tag{2.7}$$

$$\text{subject to} \qquad x_{i+1} = f_i(x_i, u_i), \quad i = j, \ldots, k-1, \tag{2.8}$$

$$x_i \in X_i, \quad i = j, \ldots, k-1, \tag{2.9}$$

$$u_i \in U_i, \quad i = j, \ldots, k-1, \tag{2.10}$$

$$x_j = x, \; x \text{ is fixed in } X_j, \tag{2.11}$$

$$x_k \in C. \tag{2.12}$$

This problem we call *problem of optimal transition from the point x to the set C on the interval $[j,k]$*. A control $\mathcal{U}_j = \{u_j, \ldots, u_{k-1}\}$, $u_i \in U_i$, $(i = j, \ldots, k-1)$ is called admissible in case its response $\mathcal{X}_j = \{x_j, x_{j+1}, \ldots, x_k\}$ with initial point $x_j = x$ satisfies the constraints $x_i \in X_i$ $(i = j, \ldots, k-1)$ and $x_k \in C$. The class of *admissible controls* for the problem $D_j(x)$ we denote by $\mathcal{P}_j(x)$.

For each $j \in [0, k-1]$ and $x \in X_j$ define $\Gamma_j(x)$ as the set of those $u \in U_j$, for which there exist a $\mathcal{U}_j = \{u_j, \ldots, u_{k-1}\} \in \mathcal{P}_j(x)$ such that $u_j = u$. Obviously, for some j, x, the set $\Gamma_j(x)$ can be empty. On the other hand, in case of a free endpoint problem without state constraints, $\Gamma_j(x) = U_j$.

The lemma below establishes a crucial property of admissible controls for the system of problems $\mathcal{D}$. Roughly speaking, it says that a truncation of an admissible control is admissible for the truncated problem and that a concatenation of admissible controls is admissible for the concatenation of the problems. At the same time it provides a formula for a recurrent computation of the objective function.

Lemma 2.1. *For every $j \in [0, k-2]$ and $x \in X_j$ one has: $\mathcal{U}_j = \{u_j, u_{j+1}, \ldots, u_{k-1}\} \in \mathcal{P}_j(x)$ if and only if $u_j \in \Gamma_j(x)$ and $\mathcal{U}_{j+1} = \{u_{j+1}, \ldots, u_{k-1}\} \in \mathcal{P}_{j+1}(f_j(x, u_j))$. Moreover, for the objective function one has:*

$$J_j(x,\mathcal{U}_j) = f_j^0(x, u_j) + J_{j+1}(f_j(x, u_j),\mathcal{U}_{j+1}). \tag{2.13}$$

Proof: The first part of the lemma is obvious, the second part, i. e. (2.13), is the consequence of the additive nature of the objective function. □

Thanks to this lemma we can utilize the notation

$$\mathcal{U}_j = \{u_j, \mathcal{U}_{j+1}\}.$$

From the lemma it follows that given $\bar{u}_j \in \Gamma_j(x)$ the set of all admissible controls for the problem $D_{j+1}(f_j(x, \bar{u}_j))$ can be obtained from those admissible controls for the problem $D_j(x)$ the first term of which is $u_j = \bar{u}_j$, by dropping the first term. Formally, we have

Corollary 2.1. *For each $j \in [0, k-2]$, $x \in X_j$ and $\bar{u}_j \in \Gamma_j(x)$ one has:*

$$\mathcal{P}_{j+1}(f_j(x, \bar{u}_j)) = \{\mathcal{U}_{j+1} \ : \ \{\bar{u}_j, \mathcal{U}_{j+1}\} \in \mathcal{P}_j(x)\}. \tag{2.14}$$

Below we assume that the following assumption about the existence of optimal controls is satisfied:

Assumption 2.1. *For each $j \in [0, k-1]$, $x \in X_j$ the following holds: if there is an admissible control for $D_j(x)$, then there exists an optimal control as well, i.e. there exists an $\hat{\mathcal{U}}_j$, for which*

$$\max_{\mathcal{U}_j \in \mathcal{P}_j(x)} J_j(x, \mathcal{U}_j) = J_j(x, \hat{\mathcal{U}}_j).$$

We adopt the convention that the maximum of a function over an empty set is $-\infty$ and define the *value function* as the maximum of the objective function as follows:

Definition 2.1. *For each $j \in [0, k-1]$ define $V_j : X_j \to [-\infty, \infty)$ by*

$$V_j(x) = \max_{\mathcal{U}_j \in \mathcal{P}_j(x)} J_j(x, \mathcal{U}_j).$$

The function V_j is called value function for the problem $D_j := \{D_j(x) \ : \ x \in X_j\}$ and the sequence of functions $V = \{V_0, \ldots, V_{k-1}\}$ is called value function for the system of problems $\mathcal{D}$.

Obviously, because of Assumption 2.1, $V_j(x)$ is finite for those j, x, for which $\mathcal{P}_j(x) \neq \emptyset$. Then $V_j(x) = J_j(x,\hat{\mathcal{U}}_j)$, where $\hat{\mathcal{U}}_j$ is the optimal control from Assumption 2.1. However, because of our convention, the function V is defined for all $j \in [0, k-1]$, $x \in X_j$.

Let us note that Assumption 2.1 is merely technical, it simplifies the justification of the dynamic programming equation. The dynamic programming equation holds also without this assumption, however in this case, the maximum is replaced by supremum in Definition 2.1 and, consequently, in Theorem 2.1 as well.

2.1.2 Dynamic Programming Equation as a Necessary and Sufficient Optimality Condition

From the definition of the value fuction V it follows

$$V_{k-1}(x) = \max_{\mathcal{U}_{k-1} \in \mathcal{P}_{k-1}(x)} J_{k-1}(x,\mathcal{U}_{k-1}) = \max_{u_{k-1} \in \Gamma_{k-1}(x)} f^0_{k-1}(x, u_{k-1}), \tag{2.15}$$

because $\mathcal{U}_{k-1} = \{u_{k-1}\}$. Similarly, for $j = 0, \ldots, k-2$, we obtain

$$\begin{aligned} V_j(x) &= \max_{\mathcal{U}_j \in \mathcal{P}_j(x)} J_j(x,\mathcal{U}_j) \\ &= \max_{\mathcal{U}_j \in \mathcal{P}_j(x)} [f^0_j(x, u_j) + J_{j+1}(f_j(x, u_j),\mathcal{U}_{j+1})] \\ &= \max_{u_j \in \Gamma_j(x)} \max_{\mathcal{U}_{j+1} \in \mathcal{P}_{j+1}(f_j)} [f^0_j(x, u_j) + J_{j+1}(f_j(x, u_j),\mathcal{U}_{j+1})] \\ &= \max_{u_j \in \Gamma_j(x)} [f^0_j(x, u_j) + \max_{\mathcal{U}_{j+1} \in \mathcal{P}_{j+1}(f_j)} J_{j+1}(f_j(x, u_j),\mathcal{U}_{j+1})] \\ &= \max_{u_j \in \Gamma_j(x)} [f^0_j(x, u_j) + V_{j+1}(f_j(x, u_j))], \end{aligned} \tag{2.16}$$

where $\mathcal{P}_{j+1}(f_j) := \mathcal{P}_{j+1}(f_j(x, u_j))$. Here we have consecutively used the definition of the value function in the first, formula (2.13) in the second, formula (2.14) in the third, formula (1.25) in the fourth and, finally, the definition of the value function again in the fifth equation.

In addition, if we define $V_k(x) = 0$ pre $x \in C$, we can synthetize (2.15) a (2.16) into the equality

$$V_j(x) = \max_{u_j \in \Gamma_j(x)} [f_j^0(x, u_j) + V_{j+1}(f_j(x, u_j))], \; j = 0, \dots, k-1, \quad (2.17)$$

to be called *Bellman's dynamic programming equation* (DPE). In the following theorem we will see that Bellman's equation is not only a necessary but also a sufficient condition of optimality.

Theorem 2.1. *The function V is the value function for the system of problems $\mathcal{D}$ if and only if the following identities are satisfied:*

$$V_j(x) = \max_{u \in \Gamma_j(x)} [f_j^0(x, u) + V_{j+1}(f_j(x, u))], \quad (2.18)$$

for $j = 0, \dots, k-1$, $x \in X_j$ and

$$V_k(x) = \begin{cases} 0, & \text{for } x \in C, \\ -\infty, & \text{for } x \notin C. \end{cases} \quad (2.19)$$

Proof: The necessity part has already been proved. The sufficiency part will be proved by induction.

The claim obviouly holds for $j = k-1$: if $V_{k-1}(x)$ solves DPE, then

$$V_{k-1}(x) = \max_{u \in \Gamma_{k-1}(x)} f_{k-1}^0(x, u) = \max_{\mathcal{U}_{k-1} \in \mathcal{P}_{k-1}(x)} J_{k-1}(x, \mathcal{U}_{k-1}),$$

thus V_{k-1} is the value function.

Let now the theorem hold for $i = j+1, \dots, k-1$. We would like to prove that it holds for $i = j, \dots, k-1$ as well. Let V_j, V_{j+1},..., V_k satisfy DPE, $x \in X_j$. From the induction assumption it follows that V_{j+1},...,V_k are the value functions for the corresponding problems. It should be proved that V_j is the value function for D_j. Suppose the contrary, i. e. V_j is not the maximal value of the objective function for $D_j(x)$. Then, there exists an admissible control $\bar{\mathcal{U}}_j$ such that $J_j(x,\bar{\mathcal{U}}_j) > V_j(x)$. This means

$$\begin{aligned} V_j(x) &< J_j(x,\bar{\mathcal{U}}_j) \\ &= f_j^0(x,\bar{u}_j) + J_{j+1}(f_j(x,\bar{u}_j),\bar{\mathcal{U}}_{j+1}) \\ &\leq f_j^0(x,\bar{u}_j) + V_{j+1}(f_j(x,\bar{u}_j)) \\ &\leq \max_{u_j\in\Gamma_j(x)}[f_j^0(x,u_j) + V_{j+1}(f_j(x,u_j))], \end{aligned}$$

the inequality in the third row having been obtained by the use of the induction assumption. The resulting inequality contradicts V_j, $V_{j+1},\ldots$ to satisfy DPE, hence V_j is the value function. □

The dynamic programming equation is a *recurrent relation*. Beginning with the condition for V_k, DPE provides a possibility to compute consecutively the value function V_i for all $i = k-1,...,0$. The following example serves as an illustration.

Example 2.1. By the dynamic programming method we find the value function for the problem:

$$\begin{aligned} &\text{maximize} && \sum_{i=0}^{1}(2x_iu_i - u_i^2 - 2x_i^2) \\ &\text{subject to} && x_{i+1} = x_i + u_i, \quad i = 0,1, \\ & && x_0 = a, \end{aligned}$$

where a is given.

Because this is a free endpoint problem, $V_2(x) = 0$ for all x.
For $j = 1$, from DPE we obtain

$$V_1(x) = \max_u(2xu - u^2 - 2x^2) = -x^2,$$

where in the last equality we have used the fact that for each fixed x the function $2xu - u^2 - 2x^2$ is concave in u and achieves its maximum at $u = x$. We denote this control value by $v_1(x)$. By the symbol $v_1(x)$

we thus denote the solution of DPE for $k = 1$ and given x. We can then write $v_1(x) = u_1 = x$.

For $j = 0$ we obtain

$$\begin{aligned} V_0(x) &= \max_u (2xu - u^2 - 2x^2 + V_1(x+u)) \\ &= \max_u (2xu - u^2 - 2x^2 - (x+u)^2) \\ &= \max_u (-2u^2 - 3x^2) = -3x^2. \end{aligned}$$

Similarly as in the previous case the last equality holds because for every fixed x the function $-2u^2-3x^2$ is concave in u and achieves its maximum at $u = 0$, hence $v_0(x) = 0$, for every x.

Because the functions $V_2(x)$, $V_1(x)$ a $V_0(x)$ have been obtained by solving DPE, by Theorem 2.2 they constitute the value function and by its definition $V_0(a) = -3a^2$ is the maximal value of the objective function of the problem.

Note that solving this problem with the help of DPE, in additon to the value function we have obtained also the functions $v_0(a) = u_0 = 0$ and $v_1(x_1) = u_1 = x_1$ as solutions of the particular maximization problems. By these functions we can generate the control $\hat{\mathcal{U}} = \{0,\ a\}$ and its response $\hat{\mathcal{X}} = \{a,\ a,\ 2a\}$ in such a way that first we determine

$$v_0(a) = 0, \text{ and thus } x_1 = a + v_0(a) = a$$

and then

$$v_1(a) = a, \text{ and thus } x_2 = a + v_1(a) = 2a.$$

By evaluating the value of the objective function in this control and its response we obtain $J(a, \hat{\mathcal{U}}) = (-2a^2) + (2a^2 - a^2 - 2a^2) = -3a^2$. Hence $J(a, \hat{\mathcal{U}}) = V_0(a)$ and it follows that $\hat{\mathcal{U}}$ is the optimal control for this problem.

2.1.3 Principle of Optimality and Feedback Control

In Example 2.1 we have shown that the control generated by the functions $v_i(x)$ is optimal for the given problem. Now we show that this conclusion is not occasional and that one can generate optimal controls by the functions solving the maximization problem of DPE in general. To this end further concepts and auxiliary arguments are needed.

A subset of the set $\Gamma_j(x)$ of those $u \in \Gamma_j(x)$ for which there exists an optimal control $\hat{\mathcal{U}}_j = \{\hat{u}_j, \dots\}$ with $\hat{u}_j = u$, we denote by $\hat{\Gamma}_j(x)$. Note that due to Assumption 2.1 on the existence of an optimal control, one has $\hat{\Gamma}_j(x) \neq \emptyset$ if $\Gamma_j(x) \neq \emptyset$.

The following theorem adapts Lemma 2.1 to optimality. Roughly speaking, it claims that a truncation of an optimal control is optimal and, in addition, that a concatenation of optimal controls is optimal as well. The first part of this claim is called *Bellman's optimality principle.*

Theorem 2.2. *For every $j \in [0, k-2]$ and $x \in X_j$ one has: $\hat{\mathcal{U}}_j = \{\hat{u}_j, \hat{\mathcal{U}}_{j+1}\}$ is an optimal control for the problem $D_j(x)$ if and only if $\hat{\mathcal{U}}_{j+1}$ is an optimal control for the problem $D_{j+1}(f_j(x, \hat{u}_j))$ and $\hat{u}_j \in \hat{\Gamma}_j(x)$.*

Proof: First, we prove Bellman's optimality principle, where the optimality of $\hat{\mathcal{U}}_{j+1}$ will be proved by contradiction. Suppose $\hat{\mathcal{U}}_{j+1}$ is not optimal for the problem $D_{j+1}(f_j(x, \hat{u}_j))$. Then, there exists an admissible control $\bar{\mathcal{U}}_{j+1}$ for the problem $D_{j+1}(f_j(x, \hat{u}_j))$ such that

$$J_{j+1}(f_j(x, \hat{u}_j), \bar{\mathcal{U}}_{j+1}) > J_{j+1}(f_j(x, \hat{u}_j), \hat{\mathcal{U}}_{j+1}).$$

We define another control for the problem $D_j(x)$ by $\bar{\mathcal{U}}_j = \{\hat{u}_j, \bar{\mathcal{U}}_{j+1}\}$. By Lemma 2.1, this control is admissible for $D_j(x)$ and one has

$$\begin{aligned} J_j(x, \bar{\mathcal{U}}_j) &= f_j^0(x, \hat{u}_j) + J_{j+1}(f_j(x, \hat{u}_j), \bar{\mathcal{U}}_{j+1}) \\ &> f_j^0(x, \hat{u}_j) + J_{j+1}(f_j(x, \hat{u}_j), \hat{\mathcal{U}}_{j+1}) \\ &= J_j(x, \hat{\mathcal{U}}_j), \end{aligned}$$

which contradicts the optimality of $\hat{\mathcal{U}}_j$. The relation $\hat{u}_j \in \hat{\Gamma}_j(x)$ is obvious.

Now, we prove the reverse implication. From the assumption $\hat{u}_j \in \hat{\Gamma}_j(x)$ it follows that there exists $\bar{\mathcal{U}}_j = \{\bar{u}_j, \bar{\mathcal{U}}_{j+1}\}$, an optimal control for the problem $D_j(x)$, such that $\bar{u}_j = \hat{u}_j$. From the already proved Bellman's principle it follows that $\bar{\mathcal{U}}_{j+1}$ is optimal for $D_{j+1}(f_j(x, \hat{u}_j))$. That is, for this problem we have two optimal controls $\bar{\mathcal{U}}_{j+1}$ and $\hat{\mathcal{U}}_{j+1}$ and, therefore,

$$\begin{aligned} J_j(x,\hat{\mathcal{U}}_j) &= f_j^0(x, \hat{u}_j) + J_{j+1}(f_j(x, \hat{u}_j),\hat{\mathcal{U}}_{j+1}) \\ &= f_j^0(x, \hat{u}_j) + J_{j+1}(f_j(x, \hat{u}_j),\bar{\mathcal{U}}_{j+1}) \\ &= J_j(x,\bar{\mathcal{U}}_j), \end{aligned}$$

which proves the optimality of $\hat{\mathcal{U}}_j$. □

Corollary 2.2. *For every $j \in [0, k-1]$ and $\hat{x}_j \in X_j$ one has:*

(a) If $\hat{\mathcal{U}}_j$ is an optimal control for the problem $D_j(\hat{x}_j)$ and $\hat{\mathcal{X}}_j$ is its response, then $\hat{u}_i \in \hat{\Gamma}_i(\hat{x}_i)$ for all $i = j, \ldots, k-1$.

(b) Conversely, if for a control $\hat{\mathcal{U}}_j$ and its response $\hat{\mathcal{X}}_j$ one has $\hat{u}_i \in \hat{\Gamma}_i(\hat{x}_i)$ for all $i = j, \ldots, k-1$, then $\hat{\mathcal{U}}_j$, $\hat{\mathcal{X}}_j$ are optimal for the problem $D_j(\hat{x}_j)$.

Proof: Conclusion (a) is an immediate consequence of Bellman's optimality princple. Conclusion (b) will be proved by induction from behind.

For $i = k-1$ conclusion (b) holds trivially. Suppose that it holds for $i = j+1, \ldots, k-1$. We prove validity of (b) for $i = j, \ldots, k-1$. Suppose that $\hat{\mathcal{X}}_j$, $\hat{\mathcal{U}}_j$ satisfy the assumptions of (b), i.e., for $\hat{\mathcal{X}}_j$, $\hat{\mathcal{U}}_j$ one has $\hat{u}_i \in \hat{\Gamma}_i(\hat{x}_i)$, $i = j, \ldots, k-1$. It hence follows that $\hat{\mathcal{U}}_{j+1} = \{\hat{u}_{j+1}, \ldots, \hat{u}_{k-1}\}$, $\hat{\mathcal{X}}_{j+1} = \{\hat{x}_{j+1}, \ldots, \hat{x}_{k-1}\}$ satisfy the assumptions of (b) for $j+1$. Thus, by the induction assumption $\hat{\mathcal{U}}_{j+1}$ is an optimal control for the problem $D_{j+1}(\hat{x}_{j+1})$ and, since $\hat{u}_j \in \hat{\Gamma}_j(\hat{x}_j)$, by the previous theorem the concatenation $\hat{\mathcal{U}}_j = \{\hat{u}_j, \hat{\mathcal{U}}_{j+1}\}$ is optimal for $D_j(x)$. Claim (b) is proved. □

From the corollary it follows that if v_i are arbitrary *selections* from $\hat{\Gamma}_i$, i.e., arbitrary functions v_i such that $v_i(x) \in \hat{\Gamma}_i(x)$, $i = 0, \ldots, k-1$, then an optimal control for the problem $D_0(x_0)$ can be generated by the functions v_i via the formulas

$$u_i = v_i(x_i),$$

recurrently from the given x_0, using the state equation. In other words, optimal responses solve the difference equation

$$x_{i+1} = f_i(x_i, v_i(x_i)), \qquad i = 0, \ldots, k-1.$$

Definition 2.2. *Denote* $\tilde{X}_i = \{x \in X_i : \ \Gamma_i(x) \neq \emptyset\}$. *For every* $i \in [0, k-1]$ *define the function* $v_i : \tilde{X}_i \to U_i$ *which associates with each* $x \in \tilde{X}_i$ *an element from the set* $\hat{\Gamma}_i(x)$. *Then, the sequence of functions* $v = \{v_0, \ldots, v_{k-1}\}$ *is called optimal feedback control (or, alternatively, optimal closed-loop control). In case of need to distinguish, the optimal control in the previous sense (defined in subsection 1.2 by the help of sequence of points* u_i*) is called optimal open-loop control.*

The name *feedback* is justified by the circumstance that v_i determines how the input of the system should react to its current state to keep it in the optimal regime. The value of optimal control in a given time instant is a function of the state of the system in that instant only. In other words, to determine the value of optimal control, the information about its current state is sufficient. No other knowledge of the past, not even of the initial state is needed.

Optimal feedback is the ultimate goal of the solution of the optimal control problem. Compared to open-loop control it has the advantage to be able to react to deviations of the process from its optimal regime. More precisely, in case the system deviates from its optimal regime because of external effects, the closed-loop control will remain to be optimal for the new deviated initial state (see Figure 2.1).
The following theorem provides a hint how to compute the optimal feedback control. It confirms the empirical experience from Example 2.1, by which the functions which solve DPE are closed-loop optimal control. In addition it claims that each optimal feedback control solves DPE.

Theorem 2.3. *Let* V *be the value functions for the system of problems* $\mathcal{D}$ *and let* V_k *satisfy (2.19). Let for each* $j \in [0, k-1]$ *one has* v_j*:* $\tilde{X}_j \ni x \mapsto \Gamma_j(x)$. *Then,* v *is the optimal feedback control if and only the*

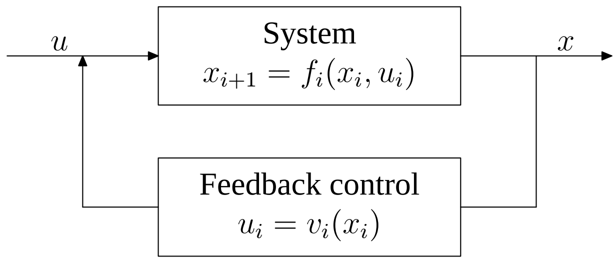

Figure 2.1: The optimal feedback control scheme

following relations hold:

$$V_j(x) = f_j^0(x, v_j(x)) + V_{j+1}(f_j(x, v_j(x))), \tag{2.20}$$

for $j = 0, \ldots, k-1$ *and* $x \in \tilde{X}_j$.

Proof: First, we prove " only if ". Let v be an optimal feedback control and let $j \in [0, k-1]$ and $x \in \tilde{X}_j$. By definition 2.2, $v_j(x) \in \hat{\Gamma}_j(x)$, and thus, there exists an optimal control $\hat{\mathcal{U}}_j = (v_j(x), \hat{\mathcal{U}}_{j+1})$ for the problem $D_j(x)$. From the optimality principle it follows that $\hat{\mathcal{U}}_{j+1}$ is optimal for $D_{j+1}(f_j(x, v_j(x))$. From the definition of the value function, from the optimality of $\hat{\mathcal{U}}_j$ and $\hat{\mathcal{U}}_{j+1}$, as well as from (2.13) we obtain

$$\begin{aligned} V_j(x) &= J_j(x, \hat{\mathcal{U}}_j) = f_j^0(x, v_j(x)) + J_{j+1}(f_j(x, v_j(x)), \hat{\mathcal{U}}_{j+1}) \\ &= f_j^0(x, v_j(x)) + V_{j+1}(f_j(x, v_j(x))), \end{aligned}$$

which is (2.20) we claimed.

Now we prove the reverse implication, i.e., the "if "part. Let $j \in [0, k-1]$ and $x \in \tilde{X}_j$ and let $v_j(x)$ satisfy (2.20). From the definition of the value function it follows that there exists an optimal control $\hat{\mathcal{U}}_{j+1}$ for the problem $D_{j+1}(f_j(x, v_j(x)))$ satisfying

$$V_{j+1}(f_j(x, v_j(x))) = J_{j+1}(f_j(x, v_j(x)), \hat{\mathcal{U}}_{j+1}).$$

The control $\hat{\mathcal{U}}_j = (v_j(x), \hat{\mathcal{U}}_{j+1})$ is admissible for $D_j(x)$ by Lemma 2.1. Obviously,

$$\begin{aligned} V_j(x) &= f_j^0(x, v_j(x)) + V_{j+1}(f_j(x, v_j(x))) \\ &= f_j^0(x, v_j(x)) + J_{j+1}(f_j(x, v_j(x)), \hat{\mathcal{U}}_{j+1}) = J_j(x, \hat{\mathcal{U}}_j), \end{aligned}$$

from which it follows that $\hat{\mathcal{U}}_j$ is an optimal control and hence $v_j(x) \in \hat{\Gamma}_j(x)$ is an optimal feedback control. □

Combining Theorem 2.1 and Theorem 2.3 we obtain

Theorem 2.4. *The function V is the value function and v is an optimal feedback control for the system of problems $\mathcal{D}$ if and only if the following relations hold:*

$$\begin{aligned} V_j(x) &= f_j^0(x, v_j(x)) + V_{j+1}(f_j(x, v_j(x))) && (2.21) \\ &= \max_{u \in \Gamma_j(x)} [f_j^0(x, u) + V_{j+1}(f_j(x, u))], && (2.22) \\ & && (2.23) \end{aligned}$$

for $j = 0, \ldots, k-1$, $x \in \tilde{X}_j$ and

$$V_k(x) = \begin{cases} 0, & \text{for } x \in C, \\ -\infty, & \text{for } x \notin C. \end{cases} \qquad (2.24)$$

2.1.4 Notes on the Dynamic Programming Equation

Remark 2.1. DPE has been developed for the standard *maximum* problem (2.1)–(2.6). However, if the problem to be solved by DPE naturally leads to *minimum* instead of maximum, there is no need to transform it formally to a maximization one by replacing the objective function by its negative. Indeed, in a straightforward way one can verify that DPE is the same for the minimum problem, the only difference being that max in it is replaced by min. The value function is naturally defined as the minimum of the objective function where finite and ∞ where not.

Remark 2.2. Since DPE is a sufficient condition for optimality makes us sure that the solution obtained from it is indeed optimal.

Remark 2.3. In case $X_i,\ U_i$ are finite sets, computation by DPE could be illustrated by a flow chart as follows:

Read $k,\ x_0,\ C$
For $j = 0, \ldots, k-1$ read $X_j,\ U_j,\ f_j(x,u),\ f_j^0(x,u)$
  if $x \in C$ then $V_k(x) := 0$
  if $x \notin C$ then $V_k(x) := -\infty$
    For $j = k-1, \ldots, 0$ do
      For $x \in X_j$ do
        For $u \in U_j$ do
          $f_j^0 := f_j^0(x,u)$
          $f_j := f_j(x,u)$
          $g_j(u) := f_j^0 + V_{j+1}(f_j)$
        $V_j(x) := \max\limits_{u \in U_j} g_j(u)$
        $v_j(x) \in \arg\max\limits_{u \in U_j} g_j(u)$
    Set $x_j := x_0$
    For $j = 0, \ldots, k-1$ do
      $v_j := v_j(x_j)$
      write $j,\ v_j,\ x_j$
      $x_{j+1} := f_j(x_j, v_j)$

Remark 2.4. Assume that the numbers of elements of the sets X_i, U_i are p, q, respectively, and that the target set C is a singleton. Would we search for the optimal control simply by the exhaustive search method, the number of computations of the functions f_i^0, f_i would be of order q^k. For the recurrent computation by the scheme of Remark 2.3 kpq computations are needed. For larger k this means a considerable reduction. Let us note, though, that this reduction is paid off by higher memory requirements. While in the case of the exhaustive search we need to store merely the last best possibility, in the case of the Bellman equation we have to store $v_j(x)$ for all x and j, i.e., kp values. Reduction of the number of operations is more substantial in the case of problems with states represented by vectors with continuous components. Indeed, if we discretize each of the n components of the state vector into N values, we obtain $q = N^n$ which can be large already for small n. The volume of computations grows with n so rapidly that even with the help of dynamic programming it cannot be mastered by highly powerful computing technology.

Remark 2.5. When computing V_j in a discretized problem with continuous nature of state, one frequently needs the value V_{j+1} in a point not included in the table. Then, one has either to interpolate the values V_{j+1} or round off $f_j(x,u)$. When solving continuous problems by discretization, in general we obtain merely an approximate solution of the original problem. By refining the discretization grid we can make the solution more precise at the cost of higher requirements on memory and computing time.

Remark 2.6. In case of a problem with fixed time and discount factor, where $f_i^0(x_i,u_i) = \beta^i F_i(x_i,u_i)$, it is natural to imbed it into a family of problems $\tilde{D}_j(x)$ in which we consider the objective functions of current value (in time j): $\tilde{J}_j = \sum_{i=j}^{k-1} \beta^{i-j} F_i(x_i,u_i)$. $\tilde{J}_j = \beta^{-j} J_j$. The function $\tilde{J}_j$ generates the *current value function* $\tilde{V}_j$, for which one obviously has $\tilde{V}_j = \beta^{-j} V_j$. By substituting these relations to DPE for the problem with discount factor we obtain an equation for the current value function $\tilde{V}$ in the following form:

$$\tilde{V}_j(x) = \max_{u \in \Gamma_j(x)} [F_j(x,u) + \beta \tilde{V}_{j+1}(f_j(x,u))], j = 0, \ldots, k-1, \tag{2.25}$$

$$\tilde{V}_k(x) = \begin{cases} 0, & \text{for } x \in C, \\ -\infty, & \text{for } x \notin C. \end{cases} \tag{2.26}$$

Note that in case of no danger of confusion we denote the current value function by the same symbol as the value function in its original meaning (so-called present value function).

2.1.5 Problem Solving

The dynamic programming equation provides a possibility to compute recurrently the optimal feedback and the value function. We show it on several examples.

Example 2.2. Optimal allocation of resources (solution of the problem from Example 1.1). The problem described in Example 1.1 will be solved for $k = 2$ and $b = 0.7$, $c = 0.4$, $g = 2$, $h = 2.5$. By rearranging the expressions in the difference equation and the objective function we obtain the problem

$$\begin{aligned} \text{maximize} \quad & \sum_{i=0}^{1}[-0.5u_i x_i + 2.5x_i] \\ \text{subject to} \quad & x_{i+1} = 0.3u_i x_i + 0.4x_i, \quad i = 0, 1, \\ & u_i \in [0, 1], \\ & x_0 = a > 0. \end{aligned}$$

First, we examine the data of the problem. It follows from them that for each i the responses to admissible controls satisfy $x_i > 0$. Therefore, it suffices to compute the value function for $x > 0$.

Because this is a free endpoint problem, $V_2(x) = 0$ for all x.

For $j = 1$ and arbitrary $x > 0$ we obtain

$$V_1(x) = \max_{u_1 \in [0,1]}[-0.5u_1 x + 2.5x] = 2.5x,$$

since maximum of a decreasing (because of $x > 0$) linear function is achieved in the left endpoint $u_1 = 0$ of the interval $U = [0, 1]$. This means that $v_1(x) = 0$ for all x.

For $j = 0$ and $x = a$ we obtain

$$\begin{aligned} V_0(a) &= \max_{u_0 \in [0,1]} [f_0^0(a, u_0) + V_1(f_0(a, u_0))] \\ &= \max_{u_0 \in [0,1]} [-0.5u_0 a + 2.5a + 2.5(0.3u_0 a + 0.4a)] \\ &= \max_{u_0 \in [0,1]} [-0.5u_0 a + 2.5a + 0.75u_0 a + a] \\ &= \max_{u_0 \in [0,1]} [+0.25u_0 a + 3.5a] = 3.75a, \end{aligned}$$

the maximum being achieved at $v_0(a) = u_0 = 1$. Therefore, $\hat{\mathcal{U}} = \{1,\ 0\}$; $\hat{\mathcal{X}} = \{a,\ 0.7a,\ 0.28a\}$ and the maximal value of the objective function is $3.75a$.

The optimal control of the considered system during the first two years is thus given by the following rule: one has to invest all the available funds into oil drills during the first year and into the purchase of real estates during the second one.

Example 2.3. Optimal consumption(solution of the problem from Example 1.2 for $b = 0$). As this is an autonomous problem with discount factor, for its solution we employ the modification of DPE for such problems of (2.25) and (2.26). Denote $\alpha = 1 + r$ and $\beta = \frac{1}{1+\delta}$. For zero value of the terminal point ($b = 0$) we then obtain the problem

$$\begin{aligned} &\text{maximize} && \sum_{i=0}^{k-1} \beta^i \ln u_i \\ &\text{subject to} && x_{i+1} = \alpha x_i - u_i, \quad i = 0, \ldots, k-1, \\ &&& x_0 = a, \\ &&& x_k = 0. \end{aligned}$$

For this problem, the relations (2.25) and (2.26) have the form

$$V_i(x) = \max_u [\ln u + \beta V_{i+1}(\alpha x - u)], \quad i = 0, \ldots, k-1,$$

where

$$V_k(x) = \begin{cases} 0, & \text{if } x = 0, \\ -\infty, & \text{if } x \neq 0. \end{cases}$$

For $i = k - 1$ we have

$$V_{k-1}(x) = \max_u [\ln u + \beta V_k(\alpha x - u)],$$

maximum being achieved at $u = \alpha x$ because of the definition of V_k. Therefore,

$$\begin{aligned} V_{k-1}(x) &= \ln \alpha x = \ln x + \ln \alpha, \\ v_{k-1}(x) &= \alpha x. \end{aligned}$$

For $i = k - 2$ we have

$$V_{k-2}(x) = \max_u [\ln u + \beta V_{k-1}(\alpha x - u)] = \max_u [\ln u + \beta \ln(\alpha x - u) + \beta \ln \alpha].$$

The function being concave, its maximum is achieved in the point $u = \dfrac{\alpha x}{1+\beta}$. Therefore

$$\begin{aligned} V_{k-2}(x) &= \ln \frac{\alpha x}{1+\beta} + \beta \ln \frac{\alpha \beta x}{1+\beta} + \beta \ln \alpha \\ &= (1+\beta) \ln x + (1+\beta) \ln \frac{\alpha}{1+\beta} + \beta \ln(\alpha\beta), \\ v_{k-2}(x) &= \frac{\alpha x}{1+\beta}. \end{aligned}$$

For $i = k - 3$ we have

$$\begin{aligned} V_{k-3}(x) &= \max_u [\ln u + \beta V_{k-2}(\alpha x - u)] \\ &= \max_u \Big[\ln u + \beta(1+\beta) \ln(\alpha x - u) \\ &\quad + \beta(1+\beta) \ln \frac{\alpha}{1+\beta} + \beta^2 \ln(\alpha\beta) \Big]. \end{aligned}$$

Maximum is achieved at $u = \dfrac{\alpha x}{1+\beta+\beta^2}$, hence

$$\begin{aligned}
V_{k-3}(x) &= \ln\frac{\alpha x}{1+\beta+\beta^2} + (\beta+\beta^2)\ln\frac{\alpha(\beta+\beta^2)x}{1+\beta+\beta^2} \\
&\quad +(\beta+\beta^2)\ln\frac{\alpha}{1+\beta} + \beta^2\ln(\alpha\beta) \\
&= (1+\beta+\beta^2)\ln x + (1+\beta+\beta^2)\ln\frac{\alpha}{1+\beta+\beta^2} \\
&\quad +(\beta+2\beta^2)\ln(\alpha\beta), \\
v_{k-3}(x) &= \frac{\alpha x}{1+\beta+\beta^2}.
\end{aligned}$$

From the above argument one can conclude that the value function and the optimal feedback will for $\beta < 1$ be of the form

$$\begin{aligned}
V_{k-i}(x) &= \frac{1-\beta^i}{1-\beta}\left(\ln x + \ln\frac{\alpha(1-\beta)}{1-\beta^i}\right) && (2.27) \\
&\quad + \left(\beta + 2\beta^2 + \cdots + (i-1)\beta^{i-1}\right)\ln(\alpha\beta). \\
v_{k-i}(x) &= \frac{\alpha x}{1+\beta+\cdots+\beta^{i-1}} = \frac{\alpha x(1-\beta)}{1-\beta^i}. && (2.28)
\end{aligned}$$

This can be easily proved by induction (see Problem 2.4).

For the above two problems it was possible to compute the optimal value of the objective function and the optimal feedback from the dynamic programming equation in a closed from. Because this is frequently not possible, other approaches to the use of DPE have to be developed.

A method of direct computation of the value function and the optimal feedback for problems with a finite number of state and control values we explain on Example 1.5 on the optimal maintenance of a machine. We have discussed this problem already in the previous chapter. We have proposed a method of its solution which in essence employed the dynamic programming principle. Now, we show how to record the steps of the computation by equation (2.18) in a tabular way.

Example 2.4. Optimal maintenance of a machine (solution of the problem). First, note that this is a free endpoint problem with no constraints on control or state whatsoever. It follows that $V_k(x) = 0$ for all x and that for $i = 0, \dots, k-1$ one has $\Gamma_i(x) = \{\text{C, N}\}$ for all x. Further, for all $i = 1, \dots, k-1$, the admissible values of x are integers from 1 to i. Because in this problem, for each i, the variables x_i and u_i can achieve a finite number of values only, for each of their possible values we can compute the possible values of f_i and f_i^0 and recurrently from $i = k-1$ to $i = 0$ also the values $V_i(x_i)$ a $v_i(x_i)$. These values we record in Table 2.1. Note that in Table 2.1 we display only the values x from 1 to 3, because – as one can verify by direct computation – for larger values of x the functions V_i and v_i do not vary. Having completed the table, we can read from it the optimal value of the objective function and reconstruct the (open-loop) optimal control as well as its response. The optimal value of the objective function for the initial value $a = x_0$ we can find in the column $V_i(x)$ for the values $i = 0$ a $x = a$. Proceeding consecutively from the end of the table upwards we can find the optimal control(s) and their responses.[1]

For simplicity we solve the problem for $k = 5$ and $a = 0$. In such case we obtain the optimal value of the objective function $V_0(0) = 100$ and three optimal controls:

$$\begin{aligned} \mathcal{U}_1 &= \{\text{N,C,N,C,N}\}, & \mathcal{X}_1 &= \{0,1,1,2,1,2\} \\ \mathcal{U}_2 &= \{\text{N,N,C,C,N}\}, & \mathcal{X}_2 &= \{0,1,2,1,1,2\} \\ \mathcal{U}_3 &= \{\text{N,N,C,N,C}\}, & \mathcal{X}_3 &= \{0,1,2,1,2,1\}. \end{aligned}$$

The values for the first of these optimal controls are highlighted in Table 2.1 by frames.

[1]Note that in general there may be more of them.

Table 2.1: Solution of Example 2.4.

i	x	u	$f_i^0(x,u)$	$f_i(x,u)$	$f_i^0 + V_{i+1}(f_i)$	$V_i(x)$	$v_i(x)$
5						0	
4	$\boxed{1}$	C	30	1	$30+0=30$	20	$\boxed{N}$
		$\boxed{N}$	20	$\boxed{2}$	$20+0=20$		
	2	C	30	1	$30+0=30$	30	C
		N	40	3	$40+0=40$		
	3	C	30	1	$30+0=30$	30	C
		N	60	4	$60+0=60$		
3	1	C	30	1	$30+20=50$	50	N, C
		N	20	2	$20+30=50$		
	$\boxed{2}$	$\boxed{C}$	30	$\boxed{1}$	$30+20=50$	50	$\boxed{C}$
		N	40	3	$40+30=70$		
	3	C	30	1	$30+20=50$	50	C
		N	60	4	$60+30=90$		
2	$\boxed{1}$	C	30	1	$30+50=80$	70	$\boxed{N}$
		$\boxed{N}$	20	$\boxed{2}$	$20+50=70$		
	2	C	30	1	$30+50=80$	80	C
		N	40	3	$40+50=90$		
	3	C	30	1	$30+50=80$	80	C
		N	60	4	$60+50=110$		
1	$\boxed{1}$	$\boxed{C}$	30	$\boxed{1}$	$30+70=100$	100	$N, \boxed{C}$
		N	20	2	$20+80=100$		
	2	C	30	1	$30+70=100$	100	C
		N	40	3	$40+80=120$		
	3	C	30	1	$30+70=100$	100	C
		N	60	4	$60+80=140$		
0	$\boxed{0}$	C	30	1	$30+100=130$	$\boxed{100}$	$\boxed{N}$
		$\boxed{N}$	0	$\boxed{1}$	$0+100=100$		

The following example is a problem with a finite number of state and control values as well. Therefore, we employ for its solution the tabular approach. However, unlike the previous example, it is a fixed endpoint problem with control constraints which results to non-admissibility of some of the control values as a consequence.

Example 2.5. Container transportation (solution of Example 1.3). First note that the possible values of state x_i for this problem are non-negative integers bounded by the total number of containers received by the railway station from the 0-th day until the $(i-1)$-th one. Due to the condition $x_5 = 0$ and the constraint $x_i \geq 0$, the sets $\Gamma_i(x)$ will be empty for some values of i and $x \in X_i$, hence $V_i(x) = \infty$. To save space we will not display such values x_i in the table. Consecutively, we will fill in the table from $k = 5$ to $k = 0$. We see that the minimal value of the objective function is 400. Then, in reverse order upwards from the bottom, beginning with the predetermined value $x_0 = 0$ we find the optimal value $\mathcal{U} = \{1, 2, 2, 2, 1\}$ and its response $\mathcal{X} = \{0, 0, 1, 0, 0, 0\}$. The corresponding values in Table 2.2 are highlighted by framing.

Recalling the problem of optimal consumption from Example 1.2 we show how one can employ dynamic programming to solve approximately problems with state or control variables of continuous nature. Two possible approaches are described in the two examples below. A third approach is the subject of Problem 2.9.

Example 2.6. Optimal consumption (tabular solution). In this example we replace continuous variables by discrete ones in such a way that the values of the state variable are integers only. We solve the problem for $r = 0.5$, $k = 3$, $\delta = 0$, $x_0 = 3$ and $x_3 = 2$.

Table 2.2: Solution of Example 2.5.

i	x	u	$f_i^0(x,u)$	$f_i(x,u)$	$f_i^0 + V_{i+1}(f_i)$	$V_i(x)$	$v_i(x)$
5	0					$\boxed{0}$	
4	$\boxed{0}$	0	50	1	$50+\infty=\infty$	70	$\boxed{1}$
		$\boxed{1}$	70	$\boxed{0}$	$70+0=70$		
	1	0	100	2	$100+\infty=\infty$	70	2
		1	120	1	$120+\infty=\infty$		
		2	70	0	$70+0=70$		
3	$\boxed{0}$	0	100	2	$100+\infty=\infty$	140	$\boxed{2}$
		1	120	1	$120+70=190$		
		$\boxed{2}$	70	$\boxed{0}$	$170+70=140$		
	1	0	150	3	$150+\infty=\infty$	190	2
		1	170	2	$170+\infty=\infty$		
		2	120	1	$120+70=190$		
2	0	0	50	1	$50+190=240$	210	1
		1	70	0	$70+140=210$		
	$\boxed{1}$	0	100	2	$100+\infty=\infty$	210	$\boxed{2}$
		1	120	1	$120+190=310$		
		$\boxed{2}$	70	$\boxed{0}$	$70+140=210$		
	2	0	150	3	$150+\infty=\infty$	310	2
		1	170	2	$170+\infty=\infty$		
		2	120	1	$120+190=310$		
1	$\boxed{0}$	0	150	3	$150+\infty=\infty$	330	$\boxed{2}$
		1	170	2	$170+310=480$		
		$\boxed{2}$	120	$\boxed{1}$	$120+210=330$		
	1	0	200	4	$200+\infty=\infty$	480	2
		1	220	3	$220+\infty=\infty$		
		2	170	$\boxed{2}$	$170+310=480$		
0	$\boxed{0}$	0	50	1	$50+480=530$	$\boxed{400}$	$\boxed{1}$
		$\boxed{1}$	70	$\boxed{0}$	$70+330=400$		

Thus, we obtain the problem

$$
\begin{aligned}
\text{maximize} \quad & \sum_{i=0}^{2} \ln u_i \\
\text{subject to} \quad & x_{i+1} = 1.5x_i - u_i, \quad i = 0, 1, 2, \\
& x_0 = 3, \quad x_3 = 2, \\
& u_i \geq 0, \quad x_i \in N.
\end{aligned}
$$

The solution is found in Table 2.3.

Table 2.3: Solution of Example 2.6.

i	x	u	f_i^0	f_i	$f_i^0 + V_{i+1}(f_i)$	V_i	v_i
3	2					$\boxed{0}$	
2	2	1	ln 1	2	0+0=0	0	1
	$\boxed{3}$	$\boxed{2.5}$	ln 2.5	$\boxed{2}$	0+ln 2.5=ln 2.5	ln 2.5	$\boxed{2.5}$
	4	4	ln 4	2	0+ln 4=ln 4	ln 4	4
	5	5.5	ln 5.5	2	0+ln 5.5=ln 5.5	ln 5.5	5.5
1	2	1	ln 1	2	0+0=0	0	1
		0.5	$\ln 0.5$	4	$\ln 0.5 + \ln 4 = \ln 2$		
	$\boxed{3}$	$\boxed{1.5}$	$\ln 1.5$	$\boxed{3}$	$\ln 1.5 + \ln 2.5 = \ln 3.75$	ln 3.75	$\boxed{1.5}$
		2.5	$\ln 2.5$	2	$\ln 2.5 + 0 = \ln 2.5$		
		1	$\ln 1$	5	$0 + \ln 5.5 = \ln 5.5$		
	4	2	$\ln 2$	4	$\ln 2 + \ln 4 = \ln 8$	ln 8	2
		3	$\ln 3$	3	$\ln 3 + \ln 2.5 = \ln 7.5$		
		4	$\ln 4$	2	$\ln 4 + 0 = \ln 4$		
		0.5	$\ln 0.5$	4	$\ln 0.5 + \ln 8 = \ln 4$		
0	$\boxed{3}$	$\boxed{1.5}$	$\ln 1.5$	$\boxed{3}$	$\ln 1.5 + \ln 3.75 = \ln 5.625$	$\boxed{\ln 5.625}$	$\boxed{1.5}$
		2.5	$\ln 2.5$	2	$\ln 2.5 + 0 = \ln 2.5$		

We see that the optimal value of the objective function is $V_0(3) =$

$\ln 5.625$, the optimal control and its response being $\mathcal{U} = \{1.5,\ 1.5,\ 2.5\}$, $\mathcal{X} = \{3, 3, 3, 2\}$ respectively.

Example 2.7. Optimal consumption (solution by computer). Now, we replace the continuous (state and control) variables by discrete ones by an equidistant grid. In such case it may happen that we need the value V_{i+1} in such a grid point x in which it was not computed. Then, for each x we choose the closest value $\tilde{x}$, $V_{i+1}(x) \sim V_{i+1}(\tilde{x})$. We demonstrate the solution procedure for $r = 0.2$, $k = 21$, $\delta = 0$, $x_1 = 3$ and $x_{21} = 0$. The problem then is

$$\begin{aligned} \text{maximize} \quad & \sum_{i=1}^{20} \ln u_i \\ \text{subject to} \quad & x_{i+1} = 1.2x_i - u_i, \quad i = 1, \ldots, 20, \\ & x_1 = 3, \\ & x_{21} = 0. \end{aligned}$$

For the solution we use the MATLAB program.[2]

```
k=20; N=50; x_steps=N; u_steps=N; x_max=25; u_max=7;
x=0:(x_max/x_steps):x_max; u=0:(u_max/u_steps):u_max;
V=zeros(k+1,x_steps+1); v=zeros(k+1,x_steps+1); vopt=zeros(k+1,1);

%Computation of the value function and the optimal feedback
V(:,:)=-inf;
V(k+1,1)=0; %only x_k=0 is feasible
for i=k:-1:1
    for j=1:(x_steps+1)
        for m=1:(u_steps+1)
            F=1.2*x(j)-u(m);
            if (F>=0) & (F<=x_max)
               F=round(F*x_steps/x_max)+1;
```

[2]The program is in MATLAB 6.5.

```
               if log(u(m))+V(i+1,F)>V(i,j)
               % log() is natural logarithm
                  V(i,j)=log(u(m))+V(i+1,F);
                  v(i,j)=m;
               end
            end
        end
    end
end

%Computation of the optimal open-loop control
j=round(3*(x_steps/x_max))+1; xopt(1)=x(j); for i=1:k
    vopt(i)=u(v(i,j));
    F=1.2*x(j)-u(v(i,j));
    xopt(i+1)=F;
    j=round(F*x_steps/x_max)+1;
end
```

Note that in addition to the parameter N determining the discretization fineness of the state and control variables one also has to chose the constraint for the maximal value of the state variable (x_max), although this constraint does not appear in the original formulation of the problem. This value is chosen sufficiently large in order not to affect the results. On Figure 2.2 one can compare the resulting shapes of the state and control variable for various values of the paramater N. In this figure also the exact solution computed by the formula (2.28) for $\beta = 1$ is shown.

2.1.6 Exercises

Exercise 2.1. Prove: if $\mathcal{U} = \{u_0, \ldots, u_j, \ldots, u_{k-1}\}$ is an optimal control for the problem $D_0(x_0)$, then $\mathcal{U}_j = \{u_j, \ldots, u_{k-1}\}$ is an optimal control

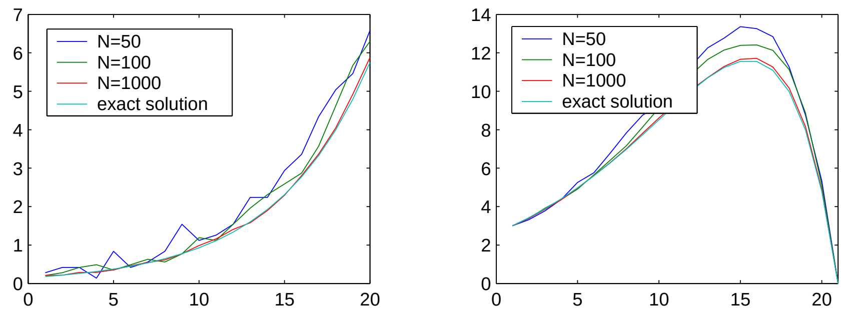

Figure 2.2: Approximative solutions of Example 2.7 by MATLAB (time development of the optimal control (left) and optimal response (right)).

for the problem $D_j(x_j(x_0,\mathcal{U}))$ for every $j \in [0, k-1]$ and thus $u_j \in \hat{\Gamma}(x_j(x_0,\mathcal{U}))$.

Exercise 2.2. Let $\hat{\mathcal{U}} = \{\hat{u}_0, \ldots, \hat{u}_j, \ldots, \hat{u}_{k-1}\}$ be an optimal control for the problem $D_0(x_0)$, let $j \in [0, k-1]$ and let $\bar{\mathcal{U}}_j = \{\bar{u}_j, \ldots, \bar{u}_{k-1}\}$ be an optimal control for $D_j(x_j(x_0,\hat{\mathcal{U}}))$ possibly different to $\hat{\mathcal{U}}_j = \{\hat{u}_j, \ldots, \hat{u}_{k-1}\}$. Prove that then $\bar{\mathcal{U}} := \{\hat{u}_0, \ldots, \hat{u}_{j-1}, \bar{u}_j, \ldots, \bar{u}_{k-1}\}$ is an optimal control for the problem $D_0(x_0)$ as well.

Exercise 2.3. (a) Derive the DPE for the problem with the objective function dependent on the terminal state, i.e., $J(x_0,\mathcal{U}) = \sum_{i=0}^{k-1} f_i^0(x_i, u_i) + \phi(x_k)$.

(b) Derive the DPE for the discounted problem

$$J = \sum_{i=0}^{k-1} \beta^i F_i(x_i, u_i) + \beta^k \phi(x_k)$$

with the objective function depending on the terminal state.

Exercise 2.4. Using mathematical induction prove (2.27) and (2.28).

Exercise 2.5. By the dynamic programming method solve the problem

$$\begin{aligned}
&\min \sum_{i=0}^{1} f_i^0(x_i, u_i),\\
&x_{i+1} = x_i + u_i, \quad i = 0, 1, \quad (k = 2),\\
&x_0 = 1,\\
&u_i \in [-3, 1],
\end{aligned}$$

where $f_0^0(x, u) = x^2 - 0,5u^2 - 2u - 1,5$ and $f_1^0(x, u) = x^2 + u^2 + 4u + 3$.

Exercise 2.6. By the dynamic programming method solve the problem

$$\begin{aligned}
&\max \sum_{i=0}^{2} x_i^2 + 2u_i\\
&x_{i+1} = x_i - u_i, \quad i = 0, 1, 2,\\
&x_0 = 1,\\
&u_i \in [-1, 1],\\
&x_3 \text{ free.}
\end{aligned}$$

Exercise 2.7. In Example 2.5 we dealt with the *container transportation problem*, where we required $x_5 = 0$. Now solve the problem without this requirement, i. e., as a free endpoint problem.

Exercise 2.8. The *optimal machine maintenance problem* was solved in Example 2.4 for $k = 5$ and $a = 0$ by Table 2.1. Now solve the problem for (a) $k = 4$ and (b) $k = 6$. Is it possible to employ Table 2.1? If yes, why and how?

Exercise 2.9. The *optimal consumption problem* was solved in Example 2.6 for $r = 0.5$, $\delta = 0$, $k = 3$, $x_0 = 3$, $x_3 = 3$ by the Table 2.3.
(a) What is the solution of this problem in the case of the initial condition (i) $x_0 = 2$, (ii) $x_0 = 4$? Report only the result, i.e. the optimal control, its response and the optimal value of the objective function.

(b) Solve this problem for $\delta = 1$. Use DPE for the discounted problem represented by (2.25) and (2.26). Compare and interprete the solutions for $\delta = 0$ and $\delta = 1$.

Exercise 2.10. Write down a modification of the solution from Example 2.7, where the value $V_{i+1}(x)$ in a point x not in the table is determined by linear interpolation. Compare the results with the results of Example 2.7.

Exercise 2.11. By DPE solve the *transport operator problem* from Exercise 1.7. Use the following formulation: $(a_1, \ldots, a_5) = (10, 3, 7, 2, 5)$,

$$\min \sum_{i=1}^{5} [(x_i + a_i - u_i)20 + (u_i - 4)^+ 30], \quad \text{where } z^+ = \begin{cases} z, & z \geq 0, \\ 0, & z < 0, \end{cases},$$

$$x_{i+1} = x_i + a_i - u_i, \quad i = 1, \ldots, 5,$$

$$x_1 = 0, \quad x_6 = 0,$$

$$u_i \in \{0, \ldots, 12\}, \quad x_i \geq 0.$$

Exercise 2.12. Solve the *optimal car replacement* problem from Exercise 1.11 for the initial states (i) $x_0 = 0$, (ii) $x_0 = 1$, (iii) $x_0 = 2$ if $k = 10$ for values of the functions ψ a φ in Table 2.4.

Table 2.4: Values of ψ and φ in Exercise 2.12

x	0	1	2	3	4	5
$\psi(x)$	50	43	37	32	28	24
$\varphi(x)$	2	2	4	3	5	

Exercise 2.13. Solve the *general repair* problem of Exercise 1.12 using DPE a) by the table procedure b) analytically.

Exercise 2.14. D'Artagnan. Formulate the following problem as an optimal control one and solve it by dynamic programming:
D'Artagnan hurries on his horse from Paris to London to deliver a confidential message from the queen to Duke of Buckingham as soon as possible. On the way to Calais there are 3 stations where he can change the horse. The distances between the stations (as well as between Paris and the first station and between the last station and Calais) are 50, 30, 40, 50 km beginning from Paris. The speed of the horse decreases with passed distance as follows: the k-th 10 kilometers the horse makes in $(30 + k)$ minutes. The change of the horse takes 13 minutes. Decide at which stations D'Artagnan should change his horse in order to reach Calais in the shortest possible time.

Exercise 2.15. Solve the problem of *optimal scheduling of orange purchase orders* from Exercise 1.14.

Exercise 2.16. In the problem of *optimal resource allocation* from Example 1.1 choose $k = 2$ and find a necessary and sufficient condition for the parameters g, h, b and c so that $\hat{u}^0 = 0$, $\hat{u}^1 = 0$ is the only optimal solution. Interpret this condition economically.

Exercise 2.17. Using DPE find the optimal solution of the following problem:

$$\begin{aligned}
&\max \sum_{i=0}^{3} x_i^2 - u_i^2 \\
&x_{i+1} = x_i + u_i, \quad i = 0, \ldots, 3, \\
&u_i \in [-1, 1], \\
&x_0 = 0.
\end{aligned}$$

Exercise 2.18. Solve the previous problem with the additional constraint $x_4 = 0$.

Exercise 2.19. Management of a harbor. The harbor unloads freight cars. The schedule of the numbers of freight cars which arrive during the coming year is known in advance. The harbor is obliged by a contract to secure unload the cargo and transport it by a boat to a given destination. In case the cargo will not be transported in the day of its arrival in the harbor the latter has to pay a contracted penalty of 100 per car and day. The boat has a capacity of 25 cars, the cost of one (return) trip being 2000 independently of the volume of the load. At the end of the last day of the year the harbor has to be empty. Write a Matlab code allowing to find the optimal control for this problem with help of DPE and answer the following questions:

a) What are the minimal costs of the harbor for transport and penalties for the whole year?

b) The harbor has the possibility to rent a boat with capacity 30 cars for 20 000 EUR a year. Is it worth to do this?

c) A new manager does not allow cargo exceeding 10 cars to stay in the harbor overnight. What will be the consequence of this regulation on the harbor in terms of economic costs ?

2.2 Autonomous Problems with Free Terminal Time. Infinite Horizon Problems

The dynamic programming equation has been derived for the standard problem (2.1)–(2.6), in general a *non-autonomous fixed time* problem defined on an interval $[0, k]$ with k fixed. For this problem, on the intervals $[j, k]$ we have introduced subproblems $D_j(x)$ and for them we have defined functions $V_j(x)$, $v_j(x)$.
In this section we will first deal with the special features of the autonomous problem. Then, we extend the validity of DPE to both the autonomous and the non-autonomous free time problems as well as for

infinite horizon problem. To this end we assume that the data entering the formulation of the problem (2.1)–(2.6), i.e., the functions f_i, f_i^0 and the sets U_i, X_i will be defined for all $i \in [0, \infty)$ and in the case of the autonomous problem will not depend on i. For every $k \in [0, \infty)$ we can first formulate for this data problems with *fixed time* defined on the intervals $[0, k]$ and then their subproblems to be denoted by $D_{j,k}$, $j \in [0, k]$; the corresponding value function we will denote by $V_{j,k}$ and the optimal feedback control by $v_{j,k}$.

2.2.1 Autonomous Problems

For $j < k$ call $D_{j,k}$ *autonomous* problem of optimal transition from x to C on the interval $[j, k]$. That is, $D_{j,k}$ is the problem

$$\text{maximize} \quad J_{j,k}(x, \mathcal{U}_{j,k}) := \sum_{i=j}^{k-1} f^0(x_i, u_i), \tag{2.29}$$

$$\text{subject to} \quad x_{i+1} = f(x_i, u_i), \quad i = j, \dots, k-1, \tag{2.30}$$

$$x_i \in X, \quad i = j, \dots, k-1, \tag{2.31}$$

$$u_i \in U, \quad i = j, \dots, k-1, \tag{2.32}$$

$$x_j = x, \ x \text{ is fixed in } X, \tag{2.33}$$

$$x_k \in C. \tag{2.34}$$

Choose an arbitrary integer $h \geq -j$. By a shift of the time variable $i \to i + h$ for each $i = j, \dots, k-1$ in the problem formulation one obtains from $D_{j,k}(x)$ the problem $D_{j+h,k+h}(x)$. It is obvious that if

$$\bar{\mathcal{U}}_{j,k} = \{\bar{u}_j, \dots, \bar{u}_{k-1}\}$$

is admissible for $D_{j,k}(x)$, then

$$\tilde{\mathcal{U}}_{j+h,k+h} = \{\tilde{u}_{j+h}, \dots, \tilde{u}_{k+h-1}\}, \ \text{where} \ \ \tilde{u}_{i+h} = \bar{u}_i,$$

is admissible for $D_{j+h,k+h}(x)$ and

$$J_{j+h,k+h}(x, \tilde{\mathcal{U}}_{j+h,k+h}) = J_{j,k}(x, \bar{\mathcal{U}}_{j,k}).$$

Now, it follows immediately that

$$V_{j+h,k+h} = V_{j,k}, \quad v_{j+h,k+h} = v_{j,k},$$

which means that the functions $V_{j,k}, v_{j,k}$ depend only on the difference between k and j.

2.2.2 Free Terminal Time Problems

Free terminal time problems are problems in which k is not given in advance. In such problems one optimizes also with respect to k. As controls one considers finite sequences of type $\{u_0, \ldots, u_i, \ldots, u_k\}$, where $u_i \in U_i$ for each $i = 0, ..., k-1$. We define value functions and optimal feedback controls also for such problems, however, since k is no longer the problem parameter, it disappears from the denotation of V and v. By arguments similar to those by which we derived the fixed time DPE one can derive the free time DPE, albeit merely as a necessary condition of optimality. In addition, we do not know for which k should $V_k(x)$ be defined. Therefore, for *non-autonomous free time problems* the practical use of the resulting recurrent relation is limited. Would we like to solve a non-autonomous free time problem by the dynamic programming method, we have to solve the corresponding fixed time problem on $[0, k]$ for each fixed k and select among the solutions the one yielding the maximal value of the objective function.

In case the free time problem is *autonomous*, neither the value function V nor the optimal feedback v depend on j. Due to this circumstance, the recurrent equation turns into the functional equation

$$V(x) = f^0(x, v(x)) + V(f(x, v(x))) = \max_{u \in \Gamma(x)} [f^0(x, u) + V(f(x, u))]. \tag{2.35}$$

Because this equation is no more a recurrent relation, it does not provide a tool to compute V. Moreover, (2.35) is a necessary condition only, its sufficiency does not hold in general.

2.2.3 Infinite Horizon Problems

A similar difficulty is encountered in *infinite horizon autonomous problems.* This is how optimal control problems, most frequently of economic nature, are called if $k = \infty$. As controls we consider in such problems infinite sequences of type $\{u_0, \dots, u_i, \dots\}$, $u_i \in U_i$, the objective function being $J := \sum_{i=0}^{\infty} f_i^0(x_i, u_i)$. In addition to the conditions, an *admissible* control should satisfy in fixed time problems, convergence of the infinite series in the definition of J is required.

Almost as a rule, infinite horizon problems of economic nature are autonomous with discount factor (see definition in Remark 1.7). That is, they are of the form

$$\begin{aligned}
\text{maximize} \quad & J := \sum_{i=0}^{\infty} \beta^i F(x_i, u_i), \\
\text{subject to} \quad & x_{i+1} = f(x_i, u_i), \quad i = 0, 1, \dots, \\
& x_i \in X, \quad i = 0, 1, \dots, \\
& u_i \in U, \quad i = 0, 1, \dots, \\
& x_0 = x, \ x \text{ is fixed from } X, \\
& \lim_{i \to \infty} x_i \in C,
\end{aligned}$$

where $\beta < 1$ is the discount factor. In order to secure convergence of the infinite series in the objective function it is now sufficient to assume F to be bounded.

Similarly as in Remark 2.6 we can imbed this problem into a family of problems $\tilde{D}_j(x)$ with initial condition $x_j = x$, for which we consider the current value objective function

$$\tilde{J}_j := \sum_{i=j}^{\infty} \beta^{i-j} F(x_i, u_i).$$

The objective function $\tilde{J}_j$ is in correspondence with the current value function $\tilde{V}_j$ for which the DPE (2.25) holds for $i = 0, 1, ...$, albeit merely as a necessary condition of optimality. Since the problem is autonomous with discount factor, by a procedure similar to that of Subsection 2.2.1 one can prove that $\tilde{V}_j(x) = \tilde{V}_{j+h}(x)$, and, therefore, the function $\tilde{V}_j$ does not depend on j. If we denote this function by $\tilde{V}$, equation (2.25) for it becomes

$$\tilde{V}(x) = \max_{u \in \Gamma(x)} [F(x,u) + \beta \tilde{V}(f(x,u))]. \tag{2.36}$$

2.2.4 Solution Methods

The equation (2.36) is a *functional equation*, i.e. an equation in which the unknown is a function. Such equation is in general difficult to solve, in our case the solution is even more difficult because of the need to maximize. We now introduce two techniques for the solution of the functional equation (2.36) from which the first one is being used more frequently.

(a) *Approximations in the space of value functions*

  (i) Choose an initial iteration of V, i.e., choose $V(x)$ for every $x \in X$.

  (ii) For the given iteration of V, by the formula

$$v(x) := \arg\max_u [F(x,u) + \beta V(f(x,u)]$$

  determine the values of v in every $x \in X$

  (iii) For the given V and v determine the values of the new iteration V in every $x \in X$ by the formulas

$$V^{(1)}(x) := [F(x, v(x)) + \beta V(f(x, v(x)))],$$
$$V(x) := V^{(1)}(x).$$

  (iv) Repeat (ii) and (iii).

(b) *Approximations in the space of closed-loop controls*

(i) Choose an initial iteration v, i.e., choose $v(x)$, for every $x \in X$.

(ii) For the given iteration v determine V as a solution of the functional equation

$$V(x) = F(x, v(x)) + \beta V(f(x, v(x))), \quad x \in X.$$

(iii) Determine the values of the next iteration of the function v in every $x \in X$ by the formula

$$v(x) := \arg\max_u [F(x, u) + \beta V(f(x, u))].$$

(iv) Repeat (ii) and (iii).

Note that in general there is no guarantee for any of the two techniques to converge or, the limit to be the solution of (2.36) and the actual value function of the problem. However, there are classes of problems for which the method can be fully justified (see Exercises 2.23 and 2.24). The solution methods for (2.36) are extensively discussed in the books [19] and [11].

We illustrate the method of approximations in the value functions space on the optimal consumption problem of Example 1.2. In addition, on this example we explain also the idea of the *method of undetermined coefficients.*

Example 2.8. Let us solve the *infinite horizon optimal consumption problem*

$$\begin{aligned} \text{maximize} \quad & \sum_{i=0}^{\infty} \beta^i \ln u_i \\ \text{subject to} \quad & x_{i+1} = \alpha x_i - u_i, \quad i = 0, 1, \dots, \\ & x_0 = a > 0, \\ & \lim_{k \to \infty} x_k \geq 0. \end{aligned}$$

I: Approximations in the space of value functions. If we choose

$$V^{(0)}(x) = \begin{cases} 0, & \text{ak } x \geq 0, \\ -\infty, & \text{ak } x < 0, \end{cases}$$

then we can compute $V^{(1)}(x)$, $V^{(2)}(x), \ldots$ and $v^{(1)}(x)$, $v^{(2)}(x), \ldots$ similarly as $V_{k-1}(x)$, $V_{k-2}(x), \ldots$ and $v_{k-1}(x)$, $v_{k-2}(x), \ldots$ in Example 2.3, albeit instead of (2.18) we employ the equation for autonomous problems with discount factor (2.36). That is, in this case we obtain

$$\begin{aligned} v^{(i)}(x) &= \frac{\alpha x}{1+\beta+\cdots+\beta^{i-1}}, \\ V^{(i)}(x) &= (1+\beta+\cdots+\beta^{i-1})\ln x \\ &\quad +(1+\beta+\cdots+\beta^{i-1})\ln\frac{\alpha}{1+\beta+\cdots+\beta^{i-1}} \\ &\quad +(\beta+2\beta^2+\cdots+(i-1)\beta^{i-1})\ln(\alpha\beta), \end{aligned}$$

hence

$$v(x) = \lim_{i\to\infty} v^{(i)}(x) = \alpha(1-\beta)x \tag{2.37}$$

$$V(x) = \lim_{i\to\infty} V^{(i)}(x) \tag{2.38}$$

$$= \frac{1}{1-\beta}\ln x + \frac{1}{1-\beta}\ln(\alpha(1-\beta)) + \frac{\beta}{(1-\beta)^2}\ln\alpha\beta. \tag{2.39}$$

II: Method of undetermined coefficients. The above procedure is rather cumbersome, because it requires several iterations of computation of the functions $V^{(i)}$ and $v^{(i)}$, an estimate of the general form of them, a proof that this form is adequate for all $i = 1, 2, \ldots$ and, finally, computation of the limit of the iterations for $i \to \infty$. However, the procedure can be simplified if, after several iterations of the computations of $V^{(i)}$ and $v^{(i)}$ we observe that the value functions $V^{(i)}$ are of the form $V^{(i)}(x) = c_i \ln x + d_i$. This observation allows us to assume that $V(x)$ will be of the form

$$V(x) = c\ln x + d, \tag{2.40}$$

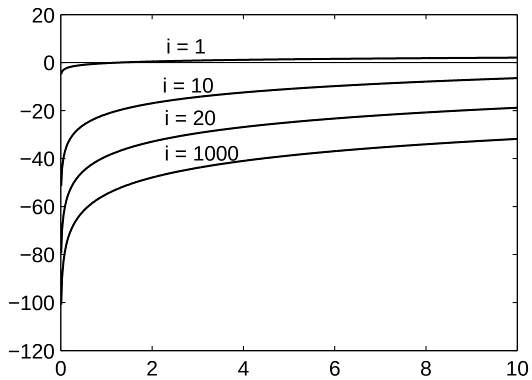

Figure 2.3: Sequence of functions $V^{(i)}$ for $\alpha = 0.8$ and $\beta = 0.9$ in Example 2.8.

as well. Therefore, we will try to find constants c, d such that

$$c \ln x + d = \max_u \left[\ln u + \beta(c \ln(\alpha x - u) + d)\right]. \tag{2.41}$$

holds. ¿From the first order condition for maximum we obtain

$$u = \frac{\alpha x}{1 + \beta c}. \tag{2.42}$$

Substituting into (2.41) we obtain

$$\begin{aligned} c \ln x + d &= \ln \frac{\alpha x}{1 + \beta c} + \beta \left[c \ln \left(\alpha x - \frac{\alpha x}{1 + \beta c} \right) + d \right] \\ &= (1 + \beta c) \ln x + \ln \frac{\alpha}{1 + \beta c} + \beta c \ln \frac{\alpha \beta c}{1 + \beta c} + \beta d. \end{aligned}$$

This equality has to hold for all x, hence

$$c = \frac{1}{1 - \beta}, \tag{2.43}$$

$$d = \frac{1}{1 - \beta} \left(\ln(\alpha(1 - \beta)) + \frac{\beta}{1 - \beta} \ln(\alpha \beta) \right). \tag{2.44}$$

If we substitute (2.43) and (2.44) into (2.40) and (2.42) we obtain the same form of $V(x)$ and $v(x) = u$ as in (2.37) and (2.39).

Note that when solving the problem we did not take into account the limit inequality for the terminal state. However, it is easy to verify that for the obtained solution it is satisfied. Indeed, from the form of the closed-loop optimal control it follows

$$x_{i+1} = \alpha\beta x_i.$$

Due to the assumptions $\alpha > 0$ and $\beta \in (0, 1)$, for all $x_0 = a > 0$ all the values x_i will be positive.

Remark 2.7. In this case we assumed the form of the value function $V(x)$. The procedure of the method of undetermined coefficients could be based alternatively on the assumption on the form of the optimal feedback control $v(x)$. After several iterations of the computation of this function one can observe that the functions $v^{(i)}(x)$ have the form $v^{(i)}(x) = C_i x$. Therefore, one can try to find the function $v(x)$ in the form $v(x) = Cx$ (cf. Exercise 2.20).

2.2.5 Exercises

Exercise 2.20. Solve the problem of Example 2.8 using the method of undetermined coefficients applied to the feedback control. Start with the assumption that the optimal feedback control is of the form $v(x) = Cx$ (cf. Remark 2.7). Hint: First derive the differential equation for $V(x)$ as the first order condition for the maximization of the function on the right-hand side of (2.36). It might seem that this leads to merely one unknown constant (C) and therefore this procedure is simpler compared to the method used in Example 2.8 with two unknown constants (c and d). This is however not true – the second constant arises as the integration one when solving the differential equation for the function V.

Exercise 2.21. Solve *the infinite horizon problem of optimal consumption* with the Cobb-Douglas production function in the form

$$\begin{aligned}
&\max \sum_{i=0}^{\infty} \beta^i \ln u_i \\
&x_{i+1} = x_i^{\alpha} - u_i, \quad i = 0, \ldots, k-1, \\
&x_0 = a, \\
&\lim_{k\to\infty} x_k \geq 0.
\end{aligned}$$

Use the method of undetermined coefficients: assume either the form of the value function ($V(x) = c \ln x + d$) or the form of the optimal feedback control ($v(x) = Cx^{\alpha}$).

Exercise 2.22. Solve the optimal control problem

$$\begin{aligned}
&\min \sum_{i=0}^{\infty} x_i^2 + u_i^2 \\
&x_{i+1} = \alpha x_i - u_i, \quad i = 0, \ldots, k-1, \\
&x_0 = 1
\end{aligned}$$

using the method of undetermined coefficients. Assume that the value function is of the form $V(x) = ax^2$, where $a \geq 0$.

Exercise 2.23. Prove that if the problem in Section 2.2.3 contains neither state nor terminal state constraints and $0 \leq F(x,u) \leq \gamma$, then the sequence of functions $V^{(k)}$ generated by the method of approximation in the space of value functions according to the formula

$$\begin{aligned}
V^{(0)}(x) &= 0, \\
V^{(k)}(x) &= \max_{u \in U} [F(x,u) + \beta V^{(k-1)}(f(x,u))]
\end{aligned}$$

is convergent. Hint: First use mathematical induction to prove that for all x and k one has $V^{(k)}(x) \leq \frac{1-\beta^k}{1-\beta}\gamma$. Then, using mathematical induction again, prove that $V^{(k+1)}(x) \geq V^{(k)}(x)$ and finally use the property that any non-decreasing sequence bounded from above converges.

Exercise 2.24. Prove that if $X = C = \mathbb{R}^n$, U is a compact subset of $\mathbb{R}^m$, the functions f, F are continuous and F is bounded, then the method of approximations in the space of value functions described in Subsection 2.2.4 converges to the solution of (2.36).
Hint: Show that the relation $V \mapsto \mathcal{T}V$ defined by the formula $\mathcal{T}V(x) = \max_u[F(x,u) + \beta V(f(x,u)]$ is a contraction on X in the space of continuous bounded functions. Hence, we can use the Banach fixed-point theorem for contraction mappings. The proof can be carried out in three steps.
Step 1. Prove that if V is continuous then $\mathcal{T}V$ is continuous as well. In addition, use the notation $F(x, v(x)) + \beta V(f(x, v(x))) := \max_u[F(x,u) + \beta V(f(x,u))]$, and apply the inequality

$$\begin{aligned}\mathcal{T}V(x_1) - \mathcal{T}V(x_2) \quad &= F(x_1, v(x_1)) + \beta V(f(x_1, v(x_1)))\\ &-F(x_2, v(x_2)) - \beta V(f(x_2, v(x_2)))\\ &\leq F(x_1, v(x_1)) + \beta V(f(x_1, v(x_1)))\\ &-F(x_2, v(x_1)) - \beta V(f(x_2, v(x_1))),\end{aligned}$$

as well as the fact that interchange of indices yields the analogous inequality for $\mathcal{T}V(x_2) - \mathcal{T}V(x_1)$. The proof of continuity can be completed by using continuity of the functions V, F and f and compactness of U.
Step 2. Prove that if V is bounded then $\mathcal{T}V$ is bounded as well. Use that F is bounded.
Step 3. Prove that $\mathcal{T}V$ is a contraction mapping. Denoting

$$F(x, v_i(x)) + \beta V_i f(x, v_i(x)) := \max_{u \in U}[F(x,u) + \beta V_i(f(x,u))], \;\; i = 1, 2,$$

use the fact that the functions $V_1(x), V_2(x)$ which are continuous on X satisfy

$$\begin{aligned}&\mathcal{T}V_1(x) - \mathcal{T}V_2(x)\\ &= F(x, v_1(x)) + \beta V_1(f(x, v_1(x))) - F(x, v_2(x)) - \beta V_2(f(x, v_2(x)))\\ &\leq F(x, v_1(x)) + \beta V_1(f(x, v_1(x))) - F(x, v_1(x)) - \beta V_2(f(x, v_1(x)))\\ &= \beta(V_1(f(x, v_1(x)) - V_2(f(x, v_1(x))),\end{aligned}$$

and interchange of indices yields

$$\mathcal{T}V_2(x) - \mathcal{T}V_1(x) \leq \beta(V_2(f(x, v_2(x)) - V_1(f(x, v_2(x))).$$

2.3 The Linear-Quadratic Problem

The dynamic programming equation allows us to compute the optimal feedback control and value function recurrently from the end. It rarely leads to a closed-form solution, however. One of the exceptions is the following linear-quadratic problem without constrains. This problem might appear to be artificial but it has an important role as a local approximation of a general problem in a neighborhood of its time independent "steady state" solution.

2.3.1 The Riccati Difference Equation for Fixed Terminal Time Problems

Consider the following optimal control problem

$$\begin{aligned} \text{minimize} \quad & \sum_{i=0}^{k-1} (x_i^T Q_i x_i + u_i^T R_i u_i) \\ \text{subject to} \quad & x_{i+1} = A_i x_i + B_i u_i, \quad i = 0, \ldots, k-1, \\ & x_0 = a, \end{aligned}$$

where $x_i \in \mathbb{R}^n$, $u_i \in \mathbb{R}^m$. Furthermore, a is a given vector, A_i, B_i, Q_i, R_i are given matrices of appropriate types. In addition, Q_i is symmetric positive semidefinite ($Q_i \geq 0$) and R_i is symmetric positive definite ($R_i > 0$). Hence, this problem is a non-autonomous fixed-time problem with a free terminal state and without constraints on control and state variables. This problem is called *linear-quadratic problem* and is denoted by LQ.

Applying the dynamic programming equation to this problem we obtain $V_k(x) = 0$ for all x. Furthermore, for $k-1$ one has

$$\begin{aligned} V_{k-1}(x) &= \min_u [x^T Q_{k-1} x + u^T R_{k-1} u] \\ &= x^T Q_{k-1} x + \min_u [u^T R_{k-1} u] = x^T Q_{k-1} x, \end{aligned}$$

because $R_{k-1} > 0$. This in addition implies $v_{k-1}(x) = 0$.

We can see that for $j = k-1$, V_{k-1} is a positive semidefinite quadratic form. Hence, one could conclude that this is also true for $j < k-1$. This claim can be proved by mathematical induction. In addition, we simultaneously prove that v_j is a linear function of x.

Assume that $V_{j+1}(x) = x^T W_{j+1} x$, where W_{j+1} is a positive semidefinite matrix. Furthermore, assume that the solution has the form $V_j(x) = x^T W_j x$. We obtain

$$\begin{aligned} V_j(x) &= x^T W_j x \\ &= \min_u [x^T Q_j x + u^T R_j u + (A_j x + B_j u)^T W_{j+1} (A_j x + B_j u)] \\ &= x^T Q_j x + x^T A_j^T W_{j+1} A_j x + \min_u [u^T (R_j + B_j^T W_{j+1} B_j) u \\ &\quad + u^T B_j^T W_{j+1} A_j x + x^T A_j^T W_{j+1} B_j u]. \end{aligned}$$

Denote the minimized function in the previous equality by g, i.e.

$$g(u) := [u^T (R_j + B_j^T W_{j+1} B_j) u + u^T B_j^T W_{j+1} A_j x + x^T A_j^T W_{j+1} B_j u].$$

The gradient $\nabla_u g$ and the Hessian $\nabla_u^2 g$ of the function g are as follows:

$$\begin{aligned} \nabla_u g &= 2(R_j + B_j^T W_{j+1} B_j) u + 2 B_j^T W_{j+1} A_j x, \\ \nabla_u^2 g &= 2(R_j + B_j^T W_{j+1} B_j). \end{aligned}$$

From $W_{j+1} \geq 0$ one obtains $B_j^T W_{j+1} B_j \geq 0$, hence $R_j + B_j^T W_{j+1} B_j > 0$. This implies that g is a strictly convex function and it attains its minimum at the point satisfying $\nabla_u g(u) = 0$, i.e.

$$u = -(R_j + B_j^T W_{j+1} B_j)^{-1} B_j^T W_{j+1} A_j x.$$

Hence
$$v_j(x) = H_j x, \tag{2.45}$$
where
$$H_j = -(R_j + B_j^T W_{j+1} B_j)^{-1} B_j^T W_{j+1} A_j. \tag{2.46}$$
Substituting (2.45) and (2.46) into the equality for $V_j(x)$ stated above yields

$$\begin{aligned} x^T W_j x &= x^T[Q_j + A_j^T W_{j+1} A_j + H_j^T (R_j + B_j^T W_{j+1} B_j) H_j \\ &\quad + H_j^T B_j^T W_{j+1} A_j + A_j^T W_{j+1} B_j H_j] x \\ &= x^T[Q_j + A_j^T W_{j+1} A_j \\ &\quad - A_j^T W_{j+1} B_j (R_j + B_j^T W_{j+1} B_j)^{-1} B_j^T W_{j+1} A_j] x. \end{aligned}$$

Hence
$$W_j = Q_j + A_j^T [W_{j+1} - W_{j+1} B_j (R_j + B_j^T W_{j+1} B_j)^{-1} B_j^T W_{j+1}] A_j. \tag{2.47}$$

This implies that $V_j(x) = x^T W_j x$ and $v_j(x) = H_j x$. It can be easily seen from (2.47) that W_j is a symmetric matrix. The function $V_j(x)$ is a value function for the LQ problem, where the objective function has only non-negative values. Hence, $x^T W_j x = V_j(x) \geq 0$ for all $x \in \mathbb{R}^n$, therefore W_j is a positive semidefinite matrix. We have proved the following theorem

Theorem 2.5. *$V_j(x)$ is the value function and $v_j(x)$ is the optimal feedback control to LQ problem if and only if*

$$\begin{aligned} V_j(x) &= x^T W_j x, \quad j = 0, \ldots, k-1, & (2.48) \\ v_j(x) &= H_j x, \quad j = 0, \ldots, k-1, & (2.49) \end{aligned}$$

where W_j, $j = 0, \ldots, k-1$ is a solution to (2.47) with $W_k = 0$ and H_j, $j = 0, \ldots, k-1$, is given by (2.46).

Remark 2.8. Note that the proof of Theorem 2.5 is based on the method of undetermined coefficients. The solution to the dynamic programming equation was assumed to be in a positive semidefinite quadratic form.

2.3.2 The Riccati Equation for Infinite Horizon Problems

Autonomous infinite horizon LQ problem is the following problem:

$$\begin{array}{ll} \text{minimize} & \sum_{i=0}^{\infty}(x_i^T Q x_i + u_i^T R u_i) \\ \text{subject to} & x_{i+1} = Ax_i + Bu_i, \quad i = 0, 1, \dots, \\ & x_0 = a, \end{array}$$

where R, Q are symmetric matrices, $R > 0$, $Q \geq 0$, $U = \mathbb{R}^m$, $X = \mathbb{R}^n$, $C = \mathbb{R}^n$. The dynamic programming equation for such problems represents only a necessary optimality condition. The value function $V(x)$ and the optimal feedback control $v(x)$ do not depend on i and satisfy the functional equation

$$\begin{aligned} V(x) &= x^T Q x + v^T(x) R v(x) + V(Ax + Bv(x)) \\ &= \min_u [x^T Q x + u^T R u + V(Ax + Bu)]. \end{aligned}$$

One can assume that the solution to this equation has the form $V(x) = x^T W x$, where $W \geq 0$ is a solution to so-called *Riccati matrix equation*

$$W = Q + A^T[W - WB(R + B^T WB)^{-1} B^T W]A. \tag{2.50}$$

The equation (2.50) can be solved analytically only in some simple cases. Usually it is necessary to use an appropriate iterative method, e.g.

$$W^{(k+1)} = Q + A^T[W^{(k)} - W^{(k)}B(R + B^T W^{(k)} B)^{-1} B^T W^{(k)}]A,$$

where $W^{(0)} = 0$. Based on some specific assumptions, it can be proved that this method yields W such that $W = \lim_{k\to\infty} W^{(k)}$.

In general,the equation (2.50) might or might not have a positive semidefinite solution. Conditions for existence as well as for determining the number of these solutions are known. Their formulation, however, would require the introduction of some specific concepts which are beyond the scope of this book. An interested reader can find more details in [14].

The application of this theory to LQ problems can be illustrated in the following example.

Example 2.9. Solve the following optimal control problem

$$\begin{aligned} &\text{minimize} && \sum_{i=0}^{k-1}[u_i^2 + x_i^2] \\ &\text{subject to} && x_{i+1} = x_i - u_i, \quad i = 0, \ldots, k-1, \\ & && x_0 = 1. \end{aligned}$$

Find the value function and the optimal feedback control for (a) $k = 5$ and (b) $k = \infty$.

(a) Solution for $k = 5$. The problem is in the LQ form with $A = 1$, $B = -1$, $R = 1$ and $Q = 1$. By using formulas (2.47) and (2.46), we obtain the following recurrence equations

$$W_i = \frac{1 + 2W_{i+1}}{1 + W_{i+1}}, \qquad H_i = \frac{W_{i+1}}{1 + W_{i+1}},$$

from which, starting from $W_5 = 0$, we obtain the values of W_i, H_i from $i = 4$ to $i = 0$ as summarized in Table 2.5.

Table 2.5: Values of W_i, H_i in Example 2.9

i	4	3	2	1	0
W_i	1	$\frac{3}{2}$	$\frac{8}{5}$	$\frac{21}{13}$	$\frac{55}{34}$
H_i	0	$\frac{1}{2}$	$\frac{3}{5}$	$\frac{8}{13}$	$\frac{21}{34}$

Substitution of these values into (2.48) and (2.49) gives the form of $V_i(x)$ and $v_i(x)$ as displayed in Table 2.6.

Table 2.6: Functions $V_i(x)$ and $v_i(x)$ in Example 2.9

i	4	3	2	1	0
$V_i(x)$	x^2	$\frac{3}{2}x^2$	$\frac{8}{5}x^2$	$\frac{21}{13}x^2$	$\frac{55}{34}x^2$
$v_i(x)$	0	$\frac{1}{2}x$	$\frac{3}{5}x$	$\frac{8}{13}x$	$\frac{21}{34}x$

The optimal (open-loop) control and its response for the initial value $x_0 = 1$ calculated by using $v_i(x)$ is given in Table 2.7.

Table 2.7: Optimal control and its response in Example 2.9

i	0	1	2	3	4	5
x_i	1	$\frac{13}{34}$	$\frac{5}{34}$	$\frac{2}{34}$	$\frac{1}{34}$	$\frac{1}{34}$
$v_i(x)$	$\frac{21}{34}x$	$\frac{8}{13}x$	$\frac{3}{2}x$	$\frac{1}{2}x$	0	
u_i	$\frac{21}{34}$	$\frac{8}{34}$	$\frac{3}{34}$	$\frac{1}{34}$	0	
$x_i - u_i$	$\frac{13}{34}$	$\frac{5}{34}$	$\frac{2}{34}$	$\frac{1}{34}$	$\frac{1}{34}$	

The solution is:

$$\hat{\mathcal{U}} = \{u_0, u_1, u_2, u_3, u_4\} = \left\{\frac{21}{34}, \frac{8}{34}, \frac{3}{34}, \frac{1}{34}, 0\right\},$$

$$\hat{\mathcal{X}} = \{x_0, x_1, x_2, x_3, x_4, x_5\} = \left\{1, \frac{13}{34}, \frac{5}{34}, \frac{2}{34}, \frac{1}{34}, \frac{1}{34}\right\}.$$

(b) Solution for $k = \infty$. In this case, we use (2.50), (2.46) and (2.45).

After substituting to (2.50), we obtain the Riccati equation in the form

$$W = \frac{1+2W}{1+W},$$

with a positive solution $W = \frac{1+\sqrt{5}}{2}$ and hence

$$H = \frac{W}{1+W} = \frac{1+\sqrt{5}}{3+\sqrt{5}}, \qquad v(x) = \frac{1+\sqrt{5}}{3+\sqrt{5}}x = \frac{\sqrt{5}-1}{2}x.$$

The optimal response can be obtained by solving the equation

$$x_{i+1} = x_i - u_i = x_i - \frac{\sqrt{5}}{2}x_i + \frac{1}{2}x_i = \left(\frac{3}{2} - \frac{\sqrt{5}}{2}\right) x_i.$$

2.3.3 Exercises

Exercise 2.25. Replace the objective function in the LQ problem analysed in Subsection 2.3.1 by the objective function with a discount factor

$$J = \sum_{i=0}^{k-1} \beta^i [x_i^T Q_i x_i + u_i^T R_i u_i].$$

Prove that in this case, the recurrence formula for H_j and W_j has the following form:

$$H_j = -\beta(R_j + \beta B_j^T W_{j+1} B_j)^{-1} B_j^T W_{j+1} A_j. \tag{2.51}$$

$$W_j = Q_j + \beta A_j^T [W_{j+1} - \beta W_{j+1} B_j (R_j + \beta B_j^T W_{j+1} B_j)^{-1} B_j^T W_{j+1}] A_j. \tag{2.52}$$

Help: use equalities (2.25) and (2.26).

Exercise 2.26. Solve the following optimal control problem:

$$\min \sum_{i=0}^{k-1} [(u_i)^2 + (x_i - u_i)^2],$$
$$x_{i+1} = x_i - u_i, \quad i = 0, \ldots, k-1,$$
$$x_0 = a.$$

Derive the value function and optimal feedback control for (a) $k = 5$, (b) $k = \infty$. Help: use the substitution $\tilde{u}_i = u_i - \frac{1}{2} x_i$.

Exercise 2.27. Assume that households would like to maintain their consumption (c_t) as stable as possible with minimal investments (i_t) into assets (a_t), in spite of instability of their incomes (y_t) modelled by a second order stationary autoregressive process:

$$\min_{c_t} \quad \sum_{t=0}^{\infty} \beta^t \left[(c_t - b)^2 + \gamma i_t^2\right],$$
$$a_{t+1} = a_t + i_t,$$
$$c_t + i_t = r a_t + y_t,$$
$$y_{t+1} = \rho_1 y_t + \rho_2 y_{t-1},$$
$$a_0, y_0, y_{-1} \text{ are given},$$

where $b > 0$, $\gamma > 0$, $r > 0$ and $\beta \in (0, 1)$ are given constants. Formulate this problem in the form of a linear-quadratic problem and write the corresponding Riccati equation.

2.4 Other Optimization Problems

The idea of embedding the optimization problem to a system of problems, allowing a recurrent calculation of the optimal control, has wide applications. It can be applied not only to problems in standard form which have been discussed so far, but also to other types of optimal control problems.

2.4.1 Optimal Control Problems with State Constraints

The method of dynamic programming can be applied also to optimal control problems with constraints of the type $\phi_i(x_i, u_i) = 0$ or $\phi_i(x_i, u_i) \leq 0$. Whereas these constraints may cause difficulties in other methods, in this case they do not pose any major complications. They are simply added to other maximization constraints in each stage of the recurrent process, i.e. they become part of the conditions determining the set $\Gamma_i(x)$. We illustrate this approach in the following example, which is interesting also because both the state and the control variable are two-dimensional.

Example 2.10. Bottleneck problem. Consider a k-stage production process, the amount of available resource in the i-th stage being denoted by x_i. This amount of resources is split into three parts. The first part u_i is used to augment resource, the part v_i is used to increase the production capacity c_i and the remaining part is left as the input to the production process itself. These amounts have to satisfy the condition $u_i \leq c_i$. The aim is to maximize the total amount of resource available to the production process, given that the growth of x_i, c_i satisfies the following equalities: $x_{i+1} = (1+\alpha)u_i, \quad c_{i+1} = c_i + \beta v_i$, where $\alpha, \beta > 0$. We obtain the following optimal control problem:

$$\begin{aligned}
\text{maximize} \quad & \sum_{i=0}^{k-1}(x_i - u_i - v_i) && (2.53)\\
\text{subject to} \quad & x_{i+1} = (1+\alpha)u_i,\ i = 0, \dots, k-1, && (2.54)\\
& c_{i+1} = c_i + \beta v_i, && (2.55)\\
& x_0 = \hat{x}_0 > 0, && (2.56)\\
& c_0 = \hat{c}_0 > 0, && (2.57)\\
& u_i \geq 0,\ v_i \geq 0 \quad i = 0, \dots, k-1, && (2.58)\\
& u_i + v_i \leq x_i, \quad i = 0, \dots, k-1, && (2.59)\\
& u_i \leq c_i, \quad i = 0, \dots, k-1. && (2.60)
\end{aligned}$$

In this problem, both the state and the control variables are two-dimensional: The control variable has components u_i and v_i and the state variable has components x_i and c_i. Besides "pure" constraints on control variables (2.58) encompassed in the standard scheme we have here also "mixed" constraints (2.59) and (2.60) tying together state and control variables. These constraints become binding in case of low values of state variables, when they narrow the set of admissible values of control variables. This property is also expressed in the name of this problem.

The dynamic programming equation for this autonomous problem has the following form:

$$V_i(x,c) = \max_{(u,v)\in\Gamma(x,c)} [x - u - v + V_{i+1}((1+\alpha)u, c + \beta v)],$$

where $i = 0, \dots, k-1$ and

$$\Gamma(x,c) = \{(u,v) : u \geq 0,\ v \geq 0, \quad u + v \leq x,\ u \leq c\}.$$

Since this is a problem with free terminal state, we have

$$V_k(x,c) = 0, \quad \text{for all } x, c.$$

For $i = k-1$ we obtain

$$V_{k-1}(x,c) = \max_{(u,v)\in\Gamma(x,c)} [x - u - v] = x,$$

where the maximum is attained at $u = 0$ and $v = 0$. This implies that in this stage the control variables expressed in the form of optimal feedback control are $u_{k-1}(x,c) = 0, \quad v_{k-1}(x,c) = 0$ for each (x,c).
For $i = k-2$ one has

$$V_{k-2}(x,c) = \max_{(u,v)\in\Gamma(x,c)} [x - u - v + V_{k-1}\left((1+\alpha)u, c + \beta v\right)]$$

$$= \max_{(u,v)\in\Gamma(x,c)} [x - u - v + (1+\alpha)u] = \max_{(u,v)\in\Gamma(x,c)} [x + \alpha u - v].$$

It is obvious that the constrained maximum is attained at $v = 0$ and at the highest possible value of u which still satisfies the constraint $0 \le u \le x$ and $0 \le u \le c$. It means that $u = \min\{x, c\}$, hence

$$V_{k-2}(x, c) = x + \alpha \min\{x, c\}.$$

The optimal feedback control is $v_{k-2}(x, c) = 0$ and $u_{k-2}(x, c) = \min\{x, c\}$.
For $i = k - 3$ we have

$$V_{k-3}(x, c) = \max_{(u,v)\in\Gamma(x,c)} [x - u - v + V_{k-1}((1+\alpha)u, c + \beta v)]$$

$$= \max_{(u,v)\in\Gamma(x,c)} [x - u - v + (1+\alpha)u + \alpha \min\{(1+\alpha)u, c + \beta v\}]$$

$$= \max_{(u,v)\in\Gamma(x,c)} [x + \alpha u - v + \alpha \min\{(1+\alpha)u, c + \beta v\}].$$

By solving this maximization problem we would obtain again the solution in the form of optimal feedback control. We could continue further analogously until $i = 0$.

We can see that including additional constraints in the formulation of the problem does not pose any theoretical problems and the problem can be solved by means of the dynamic programming equation. Too many constraints could, however, cause practical difficulties, mainly when the problem is solved analytically. In our case, the analytical solution for $i = k-3$ is more difficult. Nevertheless, it would be still possible to continue further and derive $V_0(x, c)$. It is obvious that in this example it would be more effective to solve this problem for larger values of k as a linear programming problem using e.g. the simplex method (due to the linearity of the objective function and the constraints on the variables u_i, v_i, x_i, c_i), instead of trying to solve it analytically using the dynamic programming equation.

2.4.2 Optimal Control Problems with Other Types of Objective Functions

So far we have dealt with problems where the objective function was a sum of the income in individual stages, i.e. the Lagrange objective function (according to the terminology in Remark 1.6). In this section we show that the dynamic programming equation can be derived not only for Mayer and Bolza objective functions but also for certain other types of functions that exhibit similar separability of variables.

(a) Objective function depending on terminal state.

Consider the problem which, unlike the standard problem, has the objective function in the Bolza form, i.e.

$$J := \sum_{i=0}^{k-1} f_i^0(x_i, u_i) + \phi(x_k),$$

where ϕ is a given function. For such a problem, the dynamic programming equation can be easily derived:

$$V_i(x) = \max_{u \in \Gamma_i(x)} [f_i^0(x, u) + V_{i+1}(f_i(x, u)],\ i = 0, \dots, k-1, \quad (2.61)$$

$$V_k(x) = \begin{cases} \phi(x), & \text{if } x \in C, \\ -\infty & \text{if } x \notin C. \end{cases} \quad (2.62)$$

As the Mayer objective function is a special case of the Bolza one with $f_i^0 = 0$ for all i, the corresponding dynamic programming equation to problems in Mayer form is given by (2.61) and (2.62) with $f_i^0 = 0$.

(b) Multiplicative type of objective function.

Example 2.11. Maximal reliability of a device. Assume that we have a device consisting from k serially connected blocks. In order to function properly, this device requires that each block works without errors. The reliability of an individual block can be increased by connecting additional components in parallel. Denote by $\varphi_i(u)$ the probability of the proper function of the i-th block if it contains u components connected in parallel. The price of a component of the i-th block is denoted by c_i. The aim is to choose the number of components in each block maximizing the probability of correct functioning of the whole device while the total value of these components does not exceed a given value m.

We put $x_1 = 0$ and for $i = 2, \dots, k+1$ we denote by x_i the total price of components from the first block to the $(i-1)$-th block and for $i = 1, \dots, k$ we denote by u_i the number of components in the i-th block. Then, this problem can be formulated as an optimal control problem in the following form:

$$\begin{aligned}
&\max \prod_{i=1}^{k} \varphi_i(u_i), \\
&x_{i+1} = x_i + c_i u_i, \quad i = 1, \dots, k, \\
&x_1 = 0, \\
&x_{k+1} \le m.
\end{aligned}$$

This problem differs from the standard form only in the objective function, which is multiplicative. It can be easily proved that the value function for this problem satisfies the following recurrence formula:

$$V_i(x) = \max_{u \in \mathbb{N}} [\varphi_i(u) V_{i+1}(x + c_i u)], \quad i = 1, \dots, k-1, \tag{2.63}$$

$$V_k(x) = \begin{cases} 1, & \text{if } x \le m, \\ -\infty, & \text{if } x > m, \end{cases} \tag{2.64}$$

where $\mathbb{N}$ is the set of all natural numbers. The equalities (2.63)–(2.64) correspond to the dynamic programming equation. Note that alternatively, this problem could also be rewritten in the standard form by taking logarithm of the objective function.

(c) Objective function is the minimal state value function.

Problem of setting an appropriate countercyclical policy for the national economy can be formulated as maximizing the minimum (annual) national income over k years. Such a problem can then lead to the optimal control problem of the following type:

$$\begin{aligned} \text{maximize} \quad & \min\{x_1, \ldots, x_k\}, \\ \text{subject to} \quad & x_{i+1} = f_i(x_i, u_i), \quad i = 0, \ldots, k-1, \\ & x_0 = a, \\ & u_i \in U_i, \\ & x_i \in X_i, \\ & x_k \in C. \end{aligned} \tag{2.65}$$

For this problem we can derive the following recurrence formula:

$$V_i(x) = \max_{u \in \Gamma_i(x)} \min\{f_i(x,u), V_{i+1}(f_i(x,u)\}, \; i = 0, \ldots, k-1, \tag{2.66}$$

$$V_k(x) = \begin{cases} \infty, & \text{if } x \in C, \\ -\infty, & \text{if } x \notin C, \end{cases} \tag{2.67}$$

which correponds to the dynamic programming equation.

2.4.3 Problems in Other Fields of Optimization

Many problems of other fields of optimization can also be formulated as optimal control problems and solved by the method of dynamic programming.

(a) Some problems of linear and nonlinear programming

Example 2.12. Transportation problem. Consider the transportation problem which is a well-known linear programming problem, in the case of two producers. Assume that the producers produce p_1 resp.

p_2 units of a good which is then transported to n customers, purchasing $q_1, \ldots, q_n$ units of good, where $p_1 + p_2 = q_1 + \cdots + q_n$. Furthermore, assume that c_{ij} denotes the transportation cost of one unit of good from the i-th producer to the j-th customer. We need to determine the amount of the good which should be transported from individual producers to individual customers with the aim to minimize the total transportation costs.

When denoting by x_i^j, $j = 1, 2$, the total amount of the good which is transported from the j-th producer to customers number 1 to $(i-1)$ and by u_i the relative part of delivery which the i-th customers receives from the first producer, we obtain the problem

$$\begin{aligned}
\text{minimize} \quad & J = \sum_{i=1}^{n} q_i [c_{1i} u_i + c_{2i}(1 - u_i)] \\
\text{subject to} \quad & x_{i+1}^1 = x_i^1 + q_i u_i, \quad i = 1, \ldots, n, \\
& x_{i+1}^2 = x_i^2 + q_i(1 - u_i), \quad i = 1, \ldots, n, \\
& x_1^1 = 0, \quad x_1^2 = 0, \\
& u_i \in [0, 1], \\
& x_n^1 = p_1, \quad x_n^2 = p_2.
\end{aligned}$$

This is a non-autonomous optimal control problem with two-dimensional state variable with a fixed terminal time, fixed terminal state and with a constraint on the control variable. In principle, this problem could be solved by dynamic programming. Yet, standard solution methods based on linear programming appear to be much more efficient in this linear case. It cannot be excluded, though, that for some class of nonlinear programming problems which lack a general efficient solution method the dynamic programming approach would lead to a more efficient one.

(b) Integer programming problems

In many of the examples and problems formulated in this book, the control as well as the state variables are discrete or even integers. We have seen that when these discrete variables are bounded, this fact does not complicate the calculation and the problems might by effectively solved by dynamic programming. An important representative of integer programming (or combinatorial optimization) which can be rewritten in the form of an optimal control problem, is so-called *knapsack problem*.

Example 2.13. Knapsack problem. Suppose that an elderly merchant wants to go to market to sell some goods, but she is able to carry not more than W kilograms in her knapsack. There are k types of goods. The number of items of the i-th type $i = 0, \dots,$ is a_i, each of which has a value of c_i and weight W_i. How many items (denoted by U_i) of individual types of goods should she put into her knapsack in order to maximize the total value of all items? The aim is therefore to maximize $\sum_{i=0}^{k-1} u_i c_i$, where $\sum_{i=0}^{k-1} u_i w_i \le W$, $0 \le u_i \le a_i$ and u_i are integers.

In order to formulate this problem as a standard optimal control problem, we introduce a new state variable x_i which expresses the remaining unused weight in the knapsack after determining the number of items for the first i types of goods, i.e. $x_{i+1} = x_i - u_i w_i$, where $x_0 = W$ and $x_k \ge 0$ (cf. Example 1.4 and Exercise 1.4). The formulation of the problem is

$$\text{maximize} \quad \sum_{i=0}^{k-1} u_i c_i$$

$$\begin{aligned} \text{subject to} \quad & x_{i+1} = x_i - u_i w_i, \quad i = 0, \dots, k-1, \\ & x_0 = W, \\ & u_i \in \{0, \dots, a_i\}, \quad i = 0, \dots, k-1, \\ & x_k \ge 0. \end{aligned}$$

We have obtained a non-autonomous problem with constraints on the control variable and terminal state. The problem is non-autonomous because the sets of constrains on the control variable $U_i = \{0, \dots, a_i\}$ depend on i.

(c) Some graph theory problems

The method of dynamic programming can be successfully employed to address many problems which are originally not formulated in the form of optimal control ones. A typical example of such a problem is e.g. the shortest path problem or the travelling salesman problem. Such a problem is solved either by dynamic programming or by other methods, e.g. by integer programming or branch and bound. Although methods based on the principle of dynamic programming applied to these problems require neither complicated terminology nor formulation as an optimal control problem, by their very nature these are optimal control problems as well. We will demonstrate this on the shortest path problem.

Example 2.14. The shortest path problem. Consider a network of paths where nodes $M_1, \dots, M_N$ represent cities. Denote by $d(M_i, M_j) = d_{ij}$ the direct distance from city M_i to city M_j. The aim is to find the shortest path from city $M \in \{M_1, \dots, M_N\}$ to city M_N.

Let us formulate this problem as an optimal control one. We use the following notation:
i - identification number of a city where $i = 0, \dots, k$,
x_i - our position at time i, where $x_i \in \{M_1, \dots, M_N\}$,
u_i - control variable which represents our decision to move from city x_i to city x_{i+1}, i.e., $u_i \in \{M_1, \dots, M_N\} \setminus \{x_i\}$.

We obtain the following optimal control problem:

$$\begin{aligned}
\text{minimize} \quad & \sum_{i=0}^{k-1} d(x_i, u_i) \\
\text{subject to} \quad & x_{i+1} = u_i, \quad i = 1, \ldots, k-1, \quad k \text{ is given}, \\
& x_1 = M, \\
& u_i \in \{M_1, \ldots, M_N\} \setminus \{x_i\}, \\
& x_k = M_N.
\end{aligned}$$

The conditions imply that it is an autonomous problem with free terminal time. For such problems, the dynamic programming equation is only a necessary optimality condition. In this case it is a functional equation, not a recurrence formula, as we have demonstrated in Section 2.2. The value function for the given problem satisfies the functional equation

$$V(x) = \min_{u \in \{M_1, \ldots, M_N\} \setminus \{x\}} [d(x, u) + V(u)].$$

To solve this equation, we can use the method of approximations in the space of value functions described in Section 2.2. According to this method,

$$V^{(0)}(x) = \begin{cases} 0, & \text{if } x = M_N, \\ \infty, & \text{if } x \neq M_N, \end{cases}$$

and for $j \geq 1$

$$V^{(j)}(x) = \begin{cases} \min_{u \in \{M_1, \ldots, M_N\} \setminus \{x\}} [d(x, u) + V^{(j-1)}(x)], & \text{if } x \neq M_N, \\ 0, & \text{if } x = M_N. \end{cases}$$

In addition, note that $V^{(j)}(x)$ represents the minimal length of the path from x to M_N, which does not pass through more than j nodes. Hence, it is obvious that the procedure can be stopped after N steps (i.e. there will be no changes in V in the next steps).

The use of this algorithm will be illustrated on the example of the shortest path in the network of Figure 2.4.

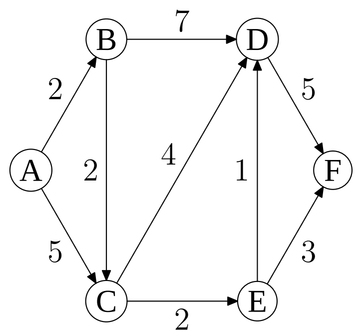

Figure 2.4: Shortest path problem

As this is an acyclic oriented graph, it is sufficient in stage j to only deal with nodes from which there is a direct path to nodes with an assigned finite value $V^{(j-1)}$ in the stage $j-1$.
For $j = 0$ we have

$$\begin{aligned} &V^{(0)}(F) = 0, \\ &V^{(0)}(A) = V^{(0)}(B) = V^{(0)}(C) = V^{(0)}(D) = V^{(0)}(E) = \infty. \end{aligned}$$

For $j = 1$ we have

$$\begin{aligned} &V^{(1)}(D) = d(D,F) + V^{(0)}(F) = 5, \text{ i.e. } v^{(1)}(D) = F, \\ &V^{(1)}(E) = d(E,F) + V^{(0)}(F) = 3, \text{ i.e. } v^{(1)}(E) = F, \\ &V^{(1)}(A) = V^{(1)}(B) = V^{(1)}(C) = \infty, \; V^{(1)}(F) = 0. \end{aligned}$$

For $j = 2$ we obtain

$$\begin{aligned} V^{(2)}(B) &= \min(d(B,D) + V^{(1)}(D), d(B,C) + V^{(1)}(C)) = 12, \\ &\quad \text{i.e. } v^{(2)}(B) = D, \\ V^{(2)}(C) &= \min(d(C,D) + V^{(1)}(D), d(C,E) + V^{(1)}(E)) = 5, \\ &\quad \text{i.e. } v^{(2)}(C) = E, \end{aligned}$$

$$
\begin{aligned}
V^{(2)}(E) &= \min(d(E,D)+V^{(1)}(D), d(E,F)+V^{(1)}(F)) = 3, \\
&\quad \text{i.e. } v^{(2)}(E) = F, \\
V^{(2)}(A) &= \infty,\ V^{(2)}(D) = 5,\ V^{(2)}(F) = 0.
\end{aligned}
$$

For $j = 3$ one has

$$
\begin{aligned}
V^{(3)}(A) &= \min(d(A,B)+V^{(2)}(B), d(A,C)+V^{(2)}(C)) = 10, \\
&\quad \text{i.e. } v^{(3)}(B) = C, \\
V^{(3)}(C) &= \min(d(C,D)+V^{(2)}(D), d(C,E)+V^{(2)}(E)) = 5, \\
&\quad \text{i.e. } v^{(3)}(C) = E, \\
V^{(3)}(D) &= 5,\ V^{(3)}(E) = 3,\ V^{(3)}(F) = 0.
\end{aligned}
$$

For $j = 4$ we have

$$
\begin{aligned}
V^{(4)}(A) &= \min(d(A,B)+V^{(3)}(B), d(A,C)+V^{(3)}(C)) = 9, \\
&\quad \text{i.e. } v^{(4)}(A) = B, \\
V^{(4)}(B) &= 7,\ V^{(4)}(C) = 5,\ V^{(4)}(D) = 5, \\
V^{(4)}(E) &= 3,\ V^{(4)}(F) = 0.
\end{aligned}
$$

The optimal solution is thus the path $A \to B \to C \to E \to F$.

Remark 2.9. This example based on the oriented acyclic graph has been presented only to simplify the explanation of the algoritm. This algorithm, however, works well for undirected graph which may contain cycles as well. In this case, it has to be modified in order to review all the nodes in each stage, and the distance between two unconnected nodes is set as infinite. After this modification, the algorithm again converges in N steps.

2.4.4 Exercises

Exercise 2.28. Derive the dynamic programming equation for the transportation problem in Example 2.12.

Exercise 2.29. Prove that the recurrence formula (2.63)–(2.64) corresponds to the dynamic programming equation for the problem of maximal reliability of a device formulated in Example 2.11.

Exercise 2.30. Assume that in the problem of maximal reliability of a device formulated in Example 2.11, the probability is a linear function of the number of components in one block connected in parallel (i.e. $\phi(u) = \alpha u$, where $\alpha > 0$) and the unit prices of components in each block are the same (i.e. $c_0 = \cdots = c_{k-1}$). By solving the problem using the dynamic programming equation given in (2.63)–(2.64) show that the number of components in each block is the same and its value is $\frac{m}{kc}$ (given that this is an integer).

Exercise 2.31. Consider the following modification of Example 2.11: we minimize the total price of the device while maintaining the reliability of its proper function at least on some given level. Formulate this modification as an optimal control problem.

Exercise 2.32. Prove that the recurrence formula (2.66)–(2.67) corresponds to the dynamic programming equation for Problem (2.65). Derive the dynamic programming equation for the modification of this problem with the objective function $J := \min\{x_0, x_1, \ldots, x_k\}$.

2.5 Stochastic Problems

In many problems the system is exposed to *random external effects*, whose exact values are not known in advance. Such problems are called *stochastic problems* and will be the subject of our interest in this subchapter. We will see that the dynamic programming equation is an effective tool for solving such problems.

2.5.1 Formulation of the Stochastic Optimal Control Problem

Example 2.15. Trading radishes. A merchant buys radishes from a farmer for the price n per bundle and sells them to customers for $p > n$ pre bundle. Daily demands are random, independent of previous days and their probability distributions are known by experience. If some radishes are not sold until the end of the day, they can be sold on the next day but with a loss of s per bundle (for example, representing the storage costs or price decrease due to their deteriorated quality). Let us use the following notation:
x_i - the carry-over from the previous day,
u_i - the number of bundles bought at day i,
z_i - the demand at day i.
Hence the amount of bundles that remained unsold at the end of day i is

$$f(x_i, u_i, z_i) := \begin{cases} x_i + u_i - z_i, & \text{if } x_i + u_i \geq z_i, \\ 0, & \text{if } x_i + u_i < z_i \end{cases} \tag{2.68}$$

and the net daily profit is

$$f^0(x_i, u_i, z_i) := \begin{cases} pz_i - nu_i - s(x_i + u_i - z_i), & \text{if } x_i + u_i \geq z_i, \\ p(x_i + u_i) - nu_i, & \text{if } x_i + u_i < z_i. \end{cases} \tag{2.69}$$

Therefore, the total profit of the merchant for the entire planning period comprising k days is

$$J = \sum_{i=0}^{k-1} f^0(x_i, u_i, z_i).$$

It depends not only on the values of the state variables x_i and the control variables u_i, but also on the values of realizations of the random variables z_i which are not known to the merchant in advance. Hence it makes no sense to solve the problem of choosing the values of control variables u_i, $i = 0, \ldots, k-1$ which maximize J without taking into account the realizations of random variables z_i, $i = 0, \ldots, k-1$. The total profit is in fact a function of these random variables, and thus it does not represent a deterministic number. The distribution of random variables z_i is, however, known. Hence, we can calculate the expected value of J for each possible set of values of the control variable u_i, $i = 0, \dots, k-1$. We can therefore formulate the following problem:

$$\text{maximize} \quad E \sum_{i=0}^{k-1} f^0(x_i, u_i, z_i) \tag{2.70}$$

$$\text{subject to} \quad x_{i+1} = f(x_i, u_i, z_i), \; i = 0, \dots, k-1, \tag{2.71}$$

$$x_0 = 0, \tag{2.72}$$

where the function f^0 in (2.70) is given by (2.69) and the function f is (2.71) given by (2.68).

In general, similarly to the deterministic case, we have a k-stage decision process. The state variable $x_i \in X_i$ describes the state of the system at the beginning of the i-th stage, $i = 0, \dots, k-1$, and the value of the control variable $u_i \in U_i$ has an influence on the behaviour of the system in the i-th stage.

At each stage, the system is exposed to the impact of the random variable $z_i \in Z_i$ whose value (realization) at the time of choosing the control u_i is not known. We assume that the random variables z_i are mutually independent over time and we only know their probability distribution. At the end of the i-th stage, the values x_i and u_i and the realization of the random variable z_i unambiguously determine the value $x_{i+1} = f_i(x_i, u_i, z_i)$ as well as the profit $f_i^0(x_i, u_i, z_i)$ in the i-th stage, where f_i and f_i^0 are given functions. At the beginning, the initial value of the state variable x_0 equals to the given value a. Unlike in the deterministic case, we now assume that the terminal state is always free and that the constraints $x_i \in X_i$ are not binding, i.e. for each $x_i \in X_i$, $u_i \in U_i$ and $z_i \in Z_i$, one has $x_{i+1} = f_i(x_i, u_i, z_i) \in X_{i+1}$.

Since the total profit over all k stages depends on the random variables z_i which we do not know at the time of taking decisions about u_i, the aim is to maximize the expected value of the profit over k stages with respect to the random variables z_i, $i = 0, \ldots, k-1$. Hence we obtain the following problem

$$\text{maximize} \quad E\sum_{i=0}^{k-1} f_i^0(x_i, u_i, z_i) \tag{2.73}$$

$$\text{subject to} \quad x_{i+1} = f_i(x_i, u_i, z_i), \; i = 0, \ldots, k-1, \tag{2.74}$$

$$x_0 = a, \tag{2.75}$$

$$u_i \in U_i, \quad i = 0, \ldots, k-1, \tag{2.76}$$

$$z_i \in Z_i, \quad i = 0, \ldots, k-1. \tag{2.77}$$

In order to formulate (2.73)–(2.77) as a proper optimal control problem, we have to add the definition of the set of admissible controls. As we will see, there are several possibilities.

First we will proceed analogously to the deterministic case: By a control, we will mean a sequence of control variables $\mathcal{U} = \{u_0, \ldots, u_{k-1}\}$, where $u_i \in U_i$ for each $i = 0, \ldots, k-1$. This kind of control is called an *open-loop control* [3]. The sequence of realizations of the random variables will be denoted by $\mathcal{Z} = \{z_0, \ldots, z_{k-1}\}$. By a *response* to the control $\mathcal{U}$ and the realization of the random variable $\mathcal{Z}$ for the given initial state (2.75) we understand the sequence $\mathcal{X} = \{x_0, \ldots, x_k\}$, where for particular i, $x_i = x_i(\mathcal{U}, \mathcal{Z})$ solve (2.74) and (2.75) with the given $\mathcal{U}$ and $\mathcal{Z}$. Since there are no constraints on the state variables and the terminal state, each control $\mathcal{U} = \{u_0, \ldots, u_{k-1}\}$, where $u_i \in U_i$, is admissible. Let us denote the class of admissible controls by $\mathcal{P}$. The value of the objective function in (2.73) can be understood as the value dependent on the choice of the control $\mathcal{U}$ and on the realization $\mathcal{Z} = \{z_0, \ldots, z_{k-1}\}$. Hence, it will be denoted by

$$J(\mathcal{U}, \mathcal{Z}) := \sum_{i=0}^{k-1} f_i^0(x_i(\mathcal{U}, \mathcal{Z}), u_i, z_i).$$

[3] When it is obvious from the context that we deal with an open-loop control, it will be simply called control.

Now we can formulate the problem of choosing a control from all admissible controls $\mathcal{U}$ that makes the *expected value of the objective function* $J(\mathcal{U}, \mathcal{Z})$ attain its maximum value with respect to the multidimensional random variable $\mathcal{Z}$. This control is called *optimal control*. The problem can be formulated as follows:

$$\max_{\mathcal{U} \in \mathcal{P}} EJ(\mathcal{U}, \mathcal{Z}). \tag{2.78}$$

The definition of optimality in the class of open-loop controls is natural, if we intend to find a control which is chosen at the beginning of the process while excluding the possibility of any further adjustments to this control based on information which will become available only later. Recalling Example 2.15, this would mean that the merchant has committed himself to some volumes of daily purchases determined a priori. In the real situation, however, it is common that he has the possibility to choose the purchased volume in individual days based on then already known volume of the carry-over from the previous day. This means that he chooses the control in the i-th stage in the form of an feedback control $u_i = v_i(x_i)$. Intuitively we feel that the possibility to use the feedback control is more favourable for the merchant.

The control will be defined as the sequence of feedback controls. In the theory of stochastic dynamic programming, such a control is called *policy*. An *admissible policy* for (2.73)–(2.77) is the sequence $\mathcal{V} = (v_0, \ldots, v_{k-1})$, where $v_i : X_i \to U_i$ for all $i = 0, ..., k-1$. It is obvious that for a given policy $\mathcal{V} = (v_0, \ldots, v_{k-1})$ and a given realization of the random variable $\mathcal{Z} = \{z_0, \ldots, z_{k-1}\}$, the value of the objective function is uniquely determined

$$\begin{aligned} J(\mathcal{V}, \mathcal{Z}) &:= \sum_{i=0}^{k-1} f_i^0(x_i, v_i(x_i), z_i), \quad \text{where} \\ x_{i+1} &= f_i(x_i, v_i(x_i), z_i), \ i = 0, \ldots, k-1, \\ x_0 &= a. \end{aligned}$$

It makes sense to define an *optimal policy* as an admissible policy which maximizes $EJ(\mathcal{V}, \mathcal{Z})$. Denoting the class of all admissible policies by $\mathcal{S}$, this problem can be formulated as follows:

$$\max_{\mathcal{V}\in\mathcal{S}} EJ(\mathcal{V}, \mathcal{Z}). \tag{2.79}$$

When choosing a control in the policy form, we thus take into account that the evolution of the system over time can be continuously observed.

However, information available in the j-th stage may include not only the immediate state x_j, but also the realizations of random variables $z_0, \ldots, z_{j-1}$. If this enhanced information is used to determine the control u_j, it means that it is determined as a function w_j based on complete information known in the stage j, i.e. $u_j = w_j(x_0, \ldots, x_j, u_0, \ldots u_{j-1}, z_0, \ldots z_{j-1})$. In this framework, the open-loop control is represented by constant functions w_j, and the policy is represented by functions w_j independent on $x_0, \ldots, x_{j-1}, u_0, \ldots u_{j-1}, z_0, \ldots z_{j-1}$. Analogously to the previous case, let us denote the respective sequences of the functions w_j by $\mathcal{W} = (w_0, \ldots, w_{k-1})$ and the set of all admissible sequences by $\mathcal{T}$. This problem is formally written as follows:

$$\max_{\mathcal{W}\in\mathcal{T}} EJ(\mathcal{W}, \mathcal{Z}). \tag{2.80}$$

The following theorem formulates in its first part the *Markovian property* of this problem. This means that the expected value of the objective function does not increase when taking into account, besides the information on the current state, other information known at the decision time. In addition, the second part of this theorem states that use of the information about the value of the current state at the decision time might increase the expected value of the objective function compared to its value obtained if such information is not taken into account.

Theorem 2.6. *Problems (2.78), (2.79) and (2.80) relate with (2.73)–(2.77) as follows:*

$$\max_{\mathcal{W}\in\mathcal{T}} EJ(\mathcal{W}, \mathcal{Z}) = \max_{\mathcal{V}\in\mathcal{S}} EJ(\mathcal{V}, \mathcal{Z}) \geq \max_{\mathcal{U}\in\mathcal{P}} EJ(\mathcal{U}, \mathcal{Z}). \tag{2.81}$$

Proof: The classes of admissible controls for the particular problems satisfy inclusions $\mathcal{T} \supset \mathcal{S} \supset \mathcal{P}$. Hence

$$\max_{\mathcal{W}\in\mathcal{T}} EJ(\mathcal{W},\mathcal{Z}) \geq \max_{\mathcal{V}\in\mathcal{S}} EJ(\mathcal{V},\mathcal{Z}) \geq \max_{\mathcal{U}\in\mathcal{P}} EJ(\mathcal{U},\mathcal{Z}).$$

It remains to prove that the first non-strict inequality holds as an equality. This is implied by the fact that by the choice of the control u_j at time j, we can affect the objective function $J(\mathcal{W}, Z)$ terms $f_i^0(x_i, u_i, z_i)$ merely for $i \geq j$. These terms do not depend on any variables known at time j except x_j. Hence, when choosing the optimal value u_j we do not lose any relevant information by restricting ourselves to functions $u_j = v_j(x_j)$ which do not depend on values $x_0, \ldots, x_{j-1}, u_0, \ldots u_{j-1}, z_0, \ldots z_{j-1}$. □

Remark 2.10. The non-strict inequality in (2.81) is not entirely satisfactory, because it actually only says that the optimal policy is not worse than the optimal open-loop control. It does not however claim that it is better. In general, a better result cannot be expected. Indeed, a deterministic system is a special case of a stochastic one where both types of control lead to the same result, as it is clear from the following remark. We can however expect that the optimal policy yields a better result if the system is random in a non-trivial way. This non-triviality, however, cannot be easily formulated. Therefore, at the end of the following section we only present a simple artificial example where the optimal policy indeed leads to a better result.

Remark 2.11. If the problem (2.73)–(2.77) is deterministic, the inequality in Theorem 2.6 is satisfied as an equality, i.e.

$$\max_{\mathcal{V}\in\mathcal{S}} J(\mathcal{V}) = \max_{\mathcal{U}\in\mathcal{P}} J(\mathcal{U}). \tag{2.82}$$

The proof is implied by the inequality (2.81) and by the fact that for each admissible policy there exists an admissible open-loop control yielding the same value of the objective function as this admissible policy for the given problem (with the given initial state). Note that the term "policy" can be naturally extended to the case of general deterministic problems (with constraints on state variables and on the terminal state) and Property (2.82) remains valid for these problem as well (cf. Exercise 2.33).[4]

Remark 2.12. The stochastic version of the optimal control problem is often more realistic than the deterministic one. The calculation of its solution is, however, more difficult. In addition, the scope of its application is limited only to the cases when the probabilistic distribution of random variables z_i is known.

In the following subsection we show that the optimal policy can be obtained by the method of dynamic programming.

2.5.2 The Dynamic Programming Equation

For any $j \in \{0, \dots, k-1\}$ and $x \in X_j$, we use the following notation: $\mathcal{Z}_j = \{z_j, \dots, z_{k-1}\}$, $\mathcal{V}_j = \{v_j, \dots, v_{k-1}\}$ and

$$\begin{aligned} J_j(x, \mathcal{V}_j, \mathcal{Z}_j) := \quad & \sum_{i=j}^{k-1} f_i^0(x_i, v_i(x_i), z_i), \\ \text{where} \quad & x_{i+1} = f_i(x_i, v_i(x_i), z_i), \; i = j, \dots, k-1, \\ & x_j = x. \end{aligned}$$

By $\mathcal{D}_j(x)$ we will denote the problem

$$\text{maximize} \qquad EJ_j(x, \mathcal{V}_j, \mathcal{Z}_j).$$

This definition requires several more detailed comments:

[4]In the deterministic case an open-loop control was called simply control and a policy was called closed-loop or feedback control (see Subsection 2.1.3). Moreover, from Corollary 2.2 it follows that the optimal closed-loop control generates the optimal open-loop one for the given initial state x_0 and, consequently, the response of the two controls is the same.

Remark 2.13. It is clear from the context that the expected value of the objective function J_j for the problem $\mathcal{D}_j(x)$ is calculated with respect to $\mathcal{Z}_j$. Since we will often calculate the expected value of the expression with respect to different segments of the sequence of random variables $\mathcal{Z} = \{z_0, \ldots, z_{k-1}\}$, we introduce the following notation: We denote by $E_{i,j}$, where $0 \leq i \leq j \leq k-1$ the expected value calculated with respect to the subsequence $\{z_i, \ldots, z_j\}$ of the random variable $\mathcal{Z}$. In case that $i = j$, we replace $E_{i,i}$ by E_i.

Remark 2.14. The maximization is carried out over all admissible policies. Since the problem contains neither any constraints on the control $u_i \in U_i$ nor any other constraints, any sequence of functions $v_i : X_i \to U_i$, where $i = j, \ldots, k-1$, represents an admissible policy.

Remark 2.15. It is sufficient that the functions v_i are defined only for those $x_j \in X_j$, which can be obtained by a realization of the response to the initial state together with the previous segment of the policy. More precisely, let us denote by $X_j(x_0, v_0, \ldots, v_{j-1})$ the set of all possible values of the response in the stage j to the segment of the policy $\mathcal{V}$ and the initial state x_0. Hence it is sufficient to have the function v_j defined only on the set $X_j(x_0, v_0, \ldots, v_{j-1})$.

For simplification, we introduce the following assumption:

Assumption 2.2. *For each $j \in [0, k-1]$ there exists an optimal policy $\mathcal{V}_j$ to the problem $D_j(x)$, where $x \in X_j$. Hence there exists $\hat{\mathcal{V}}_j$ such that*

$$\max_{\mathcal{V}_j} E_{j,k-1} J_j(x, \mathcal{V}_j, \mathcal{Z}_j) = E_{j,k-1} J_j(x, \hat{\mathcal{V}}_j, \mathcal{Z}_j)$$

for each $x \in X_j$.

This assumption allows us to define the *value function* as follows:

Definition 2.3. *For each $j \in [0, k-1]$ we define $V_j : X_j \to R$ as*

$$V_j(x) = \max_{\mathcal{V}_j} E_{j,k-1} J_j(x, \mathcal{V}_j, \mathcal{Z}_j).$$

The function V_j is called the value function for the set of problems $D_j := \{D_j(x) : x \in X_j\}$. In addition, the sequence of functions $V = \{V_0, \ldots, V_{k-1}\}$ is called the value function for the set of problems $\mathcal{D}$.

In the subsection 4.4 Elements of Probability Theory the definition of the expected value is limited to discrete random variables only. This restriction is kept in this subchapter as well, mainly for two reasons. The first one is purely technical – the definition of the expected value for a discrete random variable does not require any further assumptions on the data. The second reason is principal: Using a discrete random variable $\mathcal{Z}$ allows us to obtain equally strong results as in the deterministic case.

Assumption 2.3. *The sets Z_i are finite,*[5] *and hence $\mathcal{Z}_j$ is a discrete multidimensional random variable.*

Theorem 2.7. *Assume that Problem (2.79) satisfies Assumptions 2.2 and 2.3.*
(i) If $\hat{\mathcal{V}} = (\hat{v}_0, \ldots, \hat{v}_{k-1})$ is the optimal policy and $V = (V_0, \ldots, V_{k-1})$ is the value function, then the functions V_j, $\hat{v}_j$, $j = 0, \ldots, k-1$, satisfy for each x and $j = 0, \ldots, k-1$ the dynamic programming equation

$$\begin{aligned} V_j(x) &= \max_{u \in U_i} E_j \left(f_j^0(x, u, z_j) + V_{j+1}(f_j(x, u, z_j)) \right) \\ &= E_j \left(f_j^0(x, \hat{v}_j(x), z_j) + V_{j+1}(f_j(x, \hat{v}_j(x), z_j)) \right), \qquad (2.83) \\ V_k(x) &= 0, \text{ for each } x. \qquad (2.84) \end{aligned}$$

(ii) On the other hand, if for each x and $j = 0, \ldots, k-1$ the functions $\hat{v}_j$ and V_j satisfy (2.83) and (2.84), then $V = (V_0, \ldots, V_{k-1})$ is the value function and $\hat{\mathcal{V}} = (\hat{v}_0, \ldots, \hat{v}_{k-1})$ is the optimal policy.

The proof of this theorem is analogous to the proof of Theorem 2.1 in the deterministic case (see Section 2.5.3).

[5] This assumption can be relaxed to countability of Z_i. Then absolute convergence of all series defining the corresponding expected values has to be assumed, though.

Remark 2.16. Let us note that the theory becomes more complicated in the case of a continuous random variable. This is due to necessity of additional constrains which have to be applied to the original problem in order to ensure the existence of the expected values in both the objective function as well as in the value function (included in the dynamic programming equation). Moreover, in the case of a continuous random variable, Theorem 2.7 may not hold without additional constraints. The reason is that the principle of optimality which is used to prove the second equality in part (i) (the equality (2.83)) may not hold.

Example 2.16. The optimality principle may not hold in the case of continuous random variables. We illustrate this claim on a problem from the collection of counterexamples [7]:

$$\text{maximize} \quad E\sum_{i=0}^{k-1} u_i \tag{2.85}$$

$$\text{subject to} \quad x_{i+1} = z_i, \quad i = 0, \ldots, k-1, \tag{2.86}$$

$$x_0 = 0, \tag{2.87}$$

$$u_i \in \{0, 1\}, \quad i = 0, \ldots, k-1, \tag{2.88}$$

$$z_i \in [0, 1], \quad i = 0, \ldots, k-1, \tag{2.89}$$

where we assume that the random variables z_i are identically distributed on $[0, 1]$, hence the random variable as well as the state variable have a continuous character. This implies that any change of the policy on a measure zero set has no effect on the expected value. Indeed, let us define

$$v_i(x) = 1, \text{ for each } x \text{ and } i = 0, \ldots, k-1, \tag{2.90}$$

$$\tilde{v}_1(x) = \begin{cases} 0, & \text{if } x = 0, \\ 1, & \text{if } x \neq 0. \end{cases} \tag{2.91}$$

It is obvious that both strategies $\mathcal{V} = \{v_0, v_1, v_2, \ldots, v_{k-1}\}$ and $\tilde{\mathcal{V}} = \{v_0, \tilde{v}_1, v_2, \ldots, v_{k-1}\}$ are optimal, yet $\tilde{\mathcal{V}}$ does not satisfy the optimality principle, because the sub-policy $\tilde{\mathcal{V}}_1 = \{\tilde{v}_1, v_2, \ldots, v_{k-1}\}$ which is a subset of $\tilde{\mathcal{V}}$ is not an optimal policy for the sub-problem $\mathcal{D}_1(x_1)$ if $x_1 = 0$.

2.5.3 Proof of Theorem 2.7

Before proceeding to the actual proof of this theorem, we derive some useful auxiliary results and lemmas. Let us denote

$$I_j(x_j, \mathcal{V}_j) := E_{j,k-1} J_j(x_j, \mathcal{V}_j, \mathcal{Z}_j). \tag{2.92}$$

Lemma 2.2. *Denoting an admissible policy for $D_j(x_j)$ by $\mathcal{V}_j = \{v_j, \mathcal{V}_{j+1}\}$, we have*

$$I_j(x_j, \mathcal{V}_j) = E_j[f_j^0(x_j, v_j(x_j), z_j) + I_{j+1}(f_j(x_j, v_j(x_j), z_j), \mathcal{V}_{j+1})]. \tag{2.93}$$

Proof: We can rewrite $\mathcal{I}_j(x_j, \mathcal{V}_j)$ as follows:

$$\begin{aligned} I_j(x_j, \mathcal{V}_j) &= E_{j,k-1} J_j(x_j, \mathcal{V}_j, \mathcal{Z}_j) \\ &= E_j E_{j+1,k-1}[f^0(x_j, v_j(x_j), z_j) + J_{j+1}(f_j, \mathcal{V}_{j+1}, \mathcal{Z}_{j+1})] \\ &= E_j[f^0(x_j, v_j(x_j), z_j) + E_{j+1,k-1} J_{j+1}(f_j, \mathcal{V}_{j+1}, \mathcal{Z}_{j+1})] \\ &= E_j[f_j^0(x_j, v_j(x_j), z_j) + I_{j+1}(f_j, \mathcal{V}_{j+1})], \end{aligned}$$

where $f_j = f_j(x_j, v_j(x_j), z_j)$. We have used the law of iterated expectations for a multidimensional random variable as well as the fact that $f^0(x_j, v_j(x_j), z_j)$ does not depend on $z_{j+1}, \ldots, z_{k-1}$. □

Theorem 2.8. Optimality principle. *For each $j = 0, \ldots, k-1$, the following statement is true: If $\mathcal{V}_j = (v_j, \mathcal{V}_{j+1})$ is an optimal policy for $D_j(x_j)$ and $x_{j+1} \in X_{j+1}(x_j, v_j)$ (i.e. $x_{j+1} = f_j(x_j, v_j(x_j), z_j)$ for some $z_j \in Z_j$), then $\mathcal{V}_{j+1}$ is an optimal policy for Problem $D_{j+1}(x_{j+1})$.*

Proof: We prove that theorem by contradiction. Assume that the theorem does not hold for some j. This means that there exists an optimal policy $\hat{\mathcal{V}}_j = (\hat{v}_j, \hat{\mathcal{V}}_{j+1})$ for $D_j(x_j)$ and $\hat{x}_{j+1} \in X_{j+1}(x_j, \hat{v}_j)$ such that $\hat{\mathcal{V}}_{j+1}$ is not an optimal policy for Problem $D_{j+1}(\hat{x}_{j+1})$.

Assumption 2.2 implies that there exists an optimal policy $\bar{\mathcal{V}}_{j+1}$ for Problem D_{j+1}. Hence

$$I_{j+1}(x_{j+1}, \bar{\mathcal{V}}_{j+1}) \geq I_{j+1}(x_{j+1}, \hat{\mathcal{V}}_{j+1}), \text{ for all } x_{j+1} \tag{2.94}$$

and since $\hat{\mathcal{V}}_{j+1}$ is not an optimal policy for Problem $D_{j+1}(\hat{x}_{j+1})$, inequality (2.94) holds as a strict inequality at $x_{j+1} = \hat{x}_{j+1}$. Hence

$$I_{j+1}(\hat{x}_{j+1}, \bar{\mathcal{V}}_{j+1}) > I_{j+1}(\hat{x}_{j+1}, \hat{\mathcal{V}}_{j+1}). \tag{2.95}$$

Let us now consider a new control

$$\bar{\mathcal{V}}_j := \{\hat{v}_j, \bar{\mathcal{V}}_{j+1}\}$$

and calculate

$$\begin{aligned} I_j(x_j, \bar{\mathcal{V}}_j) &= E_j[f_j^0(x_j, \hat{v}_j(x_j), z_j) + I_{j+1}(x_{j+1}, \bar{\mathcal{V}}_{j+1})] \\ &> E_j[f_j^0(x_j, \hat{v}_j(x_j), z_j) + I_{j+1}(x_{j+1}, \hat{\mathcal{V}}_{j+1})] \\ &= I_j(x_j, \hat{\mathcal{V}}_j), \end{aligned} \tag{2.96}$$

where $x_{j+1} = f_j(x_j, \hat{v}_j(x_j), z_j)$. We have used (2.93) in the first and last equality. In addition, the strict inequality is implied by the strict inequality in (2.95), non-strict equally oriented inequality (2.94) as well as by the fact that Z_i is a countable set. The inequality (2.96) is in contradiction with the optimality of $\hat{\mathcal{V}}_j$ which proves the theorem. □

Now we proceed with the proof of Theorem 2.7 itself.

Proof of Theorem 2.7: We begin with the proof of the first equality in item (i). The definition of V implies

$$V_{k-1}(x) = \max_{\mathcal{V}_{k-1}} E_{k-1} J_{k-1}(x, \mathcal{V}_{k-1}) = \max_{u_{k-1} \in U_{k-1}} E_{k-1} f_{k-1}^0(x, u_{k-1}), \tag{2.97}$$

because $\mathcal{V}_{k-1} = \{v_{k-1}\}$ and because the maximization with respect to all admissible functions v_{k-1} at the state x is analogous to the maximization with respect to all $u_{k-1} \in U_{k-1}$.

Analogously we obtain for $j = 0, \ldots, k-2$

$$\begin{aligned}
V_j(x) &= \max_{\mathcal{V}_j} E_{j,k-1} J_j(x, \mathcal{V}_j, \mathcal{Z}_j) = \max_{\mathcal{V}_j} I_j(x, \mathcal{V}_j) \\
&= \max_{\mathcal{V}_j} E_j[f_j^0(x, v_j(x), z_j) + I_{j+1}(f_j(x, v_j(x), z_j), \mathcal{V}_{j+1})] \\
&= \max_{u_j \in U_j} \max_{\mathcal{V}_{j+1}} E_j[f_j^0(x, u_j, z_j) + I_{j+1}(f_j(x_j, u_j, z_j), \mathcal{V}_{j+1})] \\
&= \max_{u_j \in U_j} \max_{\mathcal{V}_{j+1}} [E_j f_j^0(x, u_j, z_j) + E_j I_{j+1}(f_j(x_j, u_j, z_j), \mathcal{V}_{j+1})] \\
&= \max_{u_j \in U_j} [E_j f_j^0(x, u_j, z_j) + \max_{\mathcal{V}_{j+1}} E_j I_{j+1}(f_j(x, u_j, z_j), \mathcal{V}_{j+1})] \\
&= \max_{u_j \in U_j} [E_j f_j^0(x, u_j, z_j) + E_j \max_{\mathcal{V}_{j+1}} I_{j+1}(f_j(x, u_j, z_j), \mathcal{V}_{j+1})] \\
&= \max_{u_j \in U_j} E_j[f_j^0(x, u_j, z_j) + V_{j+1}(f_j(x, u_j, z_j))]. \qquad (2.98)
\end{aligned}$$

The first equality is the definition of V, the second one is the equation (2.92), the third one is (2.93), the fourth one is implied by the fact that if $\mathcal{V}_j$ is an admissible policy for Problem $D_j(x)$, then $v_j(x) \in U_j$ and $\mathcal{V}_{j+1}$ is an admissible policy for the problem D_{j+1}. The next equality uses the additive property of the expected value and the next one uses the fact that the first term does not depend on $\mathcal{V}_{j+1}$. The sixth equality is explained below in more details. The last equality uses again the additive property of the expected value and the definition of V.

It remains to prove the sixth equality in (2.98), i.e. that

$$\max_{\mathcal{V}_{j+1}} E_j I_{j+1}(f_j(x_j, u_j, z_j), \mathcal{V}_{j+1}) = E_j \max_{\mathcal{V}_{j+1}} I_{j+1}(f_j(x_j, u_j, z_j), \mathcal{V}_{j+1}).$$

Since we have

$$I_{j+1}(f_j(x_j, u_j, z_j), \mathcal{V}_{j+1}) \leq \max_{\mathcal{V}_{j+1}} I_{j+1}(f_j(x_j, u_j, z_j), \mathcal{V}_{j+1}),$$

for all x_{j+1} and $\mathcal{V}_{j+1}$, hence also

$$E_j I_{j+1}(f_j(x_j, u_j, z_j), \mathcal{V}_{j+1}) \leq E_j \max_{\mathcal{V}_{j+1}} I_{j+1}(f_j(x_j, u_j, z_j), \mathcal{V}_{j+1}),$$

for all $\mathcal{V}_{j+1}$ and hence also for

$$\max_{\mathcal{V}_{j+1}} E_j I_{j+1}(f_j(x_j, u_j, z_j), \mathcal{V}_{j+1}) \leq E_j \max_{\mathcal{V}_{j+1}} I_{j+1}(f_j(x, u_j, z_j), \mathcal{V}_{j+1}).$$

Now we will prove the opposite inequality. Obviously for all $\mathcal{V}_{j+1}$ one has

$$\max_{\mathcal{V}_{j+1}} E_j I_{j+1}(f_j(x, u_j, z_j), \mathcal{V}_{j+1}) \geq E_j I_{j+1}(f_j(x, u_j, z_j), \mathcal{V}_{j+1}),$$

and hence also $\mathcal{V}_{j+1} = \hat{\mathcal{V}}_{j+1}$, where $\hat{\mathcal{V}}_{j+1}$ is an optimal policy (due to Assumption 2.2), which maximizes I_{j+1} for each x_{j+1}. Therefore we obtain

$$\begin{aligned} \max_{\mathcal{V}_{j+1}} E_j I_{j+1}(f_j(x, u_j, z_j), \mathcal{V}_{j+1}) &\geq E_j I_{j+1}(f_j(x, u_j, z_j), \hat{\mathcal{V}}_{j+1}) \\ &= E_j \max_{\mathcal{V}_{j+1}} I_{j+1}(f_j(x, u_j, z_j), \mathcal{V}_{j+1}). \end{aligned}$$

Hence we have fully justified the particular steps of the derivation of (2.98). If we define $V_k(x) = 0$, then we can merge equalities (2.97) and (2.98) which proves the first equality in (2.83). The second equality in (2.83) is implied by the optimality principle and by Lemma 2.2:

$$\begin{aligned} V_j(x) &= \max_{\mathcal{V}_j} E_{j,k-1} J_j(x, \mathcal{V}_j, \mathcal{Z}_j) = \max_{\mathcal{V}_j} I_j(x, \mathcal{V}_j) = I_j(x, \hat{\mathcal{V}}_j) \\ &= E_j[f_j^0(x, \hat{v}_j(x), z_j) + I_{j+1}(f_j(x, \hat{v}_j(x), z_j), \hat{\mathcal{V}}_{j+1})] \\ &= E_j[f_j^0(x, \hat{v}_j(x), z_j) + V_{j+1}(f_j(x, \hat{v}_j(x), z_j))]. \end{aligned}$$

Part (i) is proved. We now prove part (ii) by mathematical induction. Obviously, the theorem holds for $j = k - 1$: Indeed, if V_{k-1} and

$\hat{v}_{k-1}$ satisfy (2.83) and (2.84), then

$$\begin{aligned} V_{k-1}(x) &= \max_{u \in U_{k-1}} E_{k-1} f^0_{k-1}(x,u) = E_{k-1} f^0_{k-1}(x, \hat{v}_{k-1}(x)) \\ &= \max_{\mathcal{V}_{k-1}} E_{k-1} J_{k-1}(x, \mathcal{V}_{k-1}) = E_{k-1} J_{k-1}(x, \hat{\mathcal{V}}_{k-1}), \end{aligned}$$

and hence $\hat{v}_{k-1}(x) = u$ is an optimal feedback control and V_{k-1} is the value function. Assume now that the theorem holds for $i = j+1, \ldots, k-1$. We aim to prove that it holds also for $i = j, \ldots, k-1$.

Assume that V_j, $V_{j+1}, \ldots$, V_k satisfy the dynamic programming equation, $x \in X_j$. The induction hypothesis implies that $V_{j+1}, \ldots, V_k$ are value functions for the respective problems. We need to prove that V_j is a value function for Problem D_j. Assume the contradiction, i.e. that V_j is not the maximal value of the objective function for $D_j(x)$. Then there exists a policy $\bar{\mathcal{V}}_j$ such that

$$E_{j,k-1} J_j(x, \bar{\mathcal{V}}_j, \mathcal{Z}_j) > V_j(x).$$

It means that

$$\begin{aligned} V_j(x) &< E_{j,k-1} J_j(x, \bar{\mathcal{V}}_j, \mathcal{Z}_j) = I_j(x, \bar{\mathcal{V}}_j) \\ &= E_j[f^0_j(x, \bar{v}_j(x)) + I_{j+1}(f_j(x, \bar{v}_j(x), z_j)), \bar{\mathcal{V}}_{j+1})] \\ &\le E_j[f^0_j(x, \bar{v}_j(x)) + V_{j+1}(f_j(x, \bar{v}_j(x)))] \\ &\le \max_{u_j \in U_j} E_j[f^0_j(x, u_j) + V_{j+1}(f_j(x, u_j))], \end{aligned}$$

where the first non-strict inequality is obtained using the induction hypothesis. Hence we have a contradiction with the assumption that V_j, $V_{j+1}, \ldots$, satisfy the dynamic programming equation. Therefore, V_j is a value function.

Assume now that the functions $\hat{v}_j, \hat{v}_{j+1}, \ldots, \hat{v}_{k-1}$ satisfy the dynamic programming equation. Then $\hat{\mathcal{V}}_{j+1} = (\hat{v}_{j+1}, \ldots \hat{v}_{k-1})$ is an optimal policy, due to the induction hypothesis. We need to prove that $\hat{\mathcal{V}}_j = (\hat{v}_j, \hat{\mathcal{V}}_{j+1})$ is an optimal policy. One has

$$\begin{aligned} I_j(\hat{\mathcal{V}}_j) &= E_j[f^0_j(x, \hat{v}_j(x)) + I_{j+1}(f_j(x, \hat{v}_j(x), z_j), \hat{\mathcal{V}}_{j+1})] \\ &= E_j[f^0_j(x, \hat{v}_j(x)) + V_{j+1}(f_j(x, \hat{v}_j(x)))] = V_j(x), \end{aligned} \tag{2.99}$$

where we have used equality (2.93), the fact that $\hat{\mathcal{V}}_{j+1}$ is an optimal policy and V_{j+1} is a value function and finally that $\hat{v}_j$ satisfies the dynamic programming equation. Since V_j is a value function, (2.99) proves the optimality of the policy $\hat{\mathcal{V}}_j$. □

2.5.4 Problem Solving

Example 2.17. Example where $\min_{\mathcal{V}} EJ(\mathcal{V},\mathcal{Z}) < \min_{\mathcal{U}} EJ(\mathcal{U},\mathcal{Z})$. The formulation of the problem is as follows:

$$\begin{aligned}
\text{minimize} \quad & E\left[\sum_{i=0}^{1}(x_i - u_i + z_i)^2\right] \\
\text{subject to} \quad & x_{i+1} = x_i - u_i + z_i, \quad i = 0, 1, \\
& x_0 = 0, \\
& z_i = \begin{cases} 1, & \text{with probability } 1/2, \\ 0, & \text{with probability } 1/2, \end{cases} \quad i = 0, 1.
\end{aligned}$$

We will solve this problem in two versions: We find both (a) the program control as well as (b) the optimal feedback control.

(a) Let us calculate the value of the objective function for the program control. Obviously

$$\begin{aligned}
J(\mathcal{U},\mathcal{Z}) &= (x_0 - u_0 + z_0)^2 + (x_1 - u_1 + z_1)^2 \\
&= (-u_0 + z_0)^2 + (-u_0 + z_0 - u_1 + z_1)^2 \\
&= (2u_0^2 + 2u_0u_1 + u_1^2) + (-4u_0z_0 + 2z_0^2 - 2z_0u_1) \\
&+ (-2u_0z_1 - 2u_1z_1 + z_1^2) + 2z_0z_1.
\end{aligned}$$

From the probability distribution of the random variables z_0 and z_1 one obtains

$$\begin{aligned} EJ(\mathcal{U},\mathcal{Z}) &= 2u_0^2 + 2u_0u_1 + u_1^2 + \frac{1}{4}\underbrace{0}_{z_0=0,z_1=0} \\ &\quad + \frac{1}{4}\underbrace{(-2u_0 - 2u_1 + 1)}_{z_0=0,z_1=1} + \frac{1}{4}\underbrace{(-4u_0 + 2 - 2u_1)}_{z_0=1,z_1=0} \\ &\quad + \frac{1}{4}\underbrace{(-4u_0 + 2 - 2u_1 - 2u_0 - 2u_1 + 1 + 2)}_{z_0=1,z_1=1} \\ &= 2u_0^2 + 2u_0u_1 + u_1^2 - 3u_0 - 2u_1 + 2. \end{aligned}$$

Minimum of this convex function is attained at a point $(\hat{u}_0, \hat{u}_1)$ in which

$$\frac{\partial EJ}{\partial u_0} = \frac{\partial EJ}{\partial u_1} = 0,$$

i.e.

$$\begin{aligned} 4\hat{u}_0 + 2\hat{u}_1 - 3 &= 0, \\ 2\hat{u}_0 + 2\hat{u}_1 - 2 &= 0, \end{aligned}$$

which implies $\hat{u}_0 = \frac{1}{2}$, $\hat{u}_1 = \frac{1}{2}$ and $EJ(\hat{\mathcal{U}}, \mathcal{Z}) = \frac{3}{4}$.

(b) Now we use the dynamic programming equation to calculate the value function. We have $V_2(x) = 0$ for any x. Furthermore,

$$\begin{aligned} V_1(x) &= \min_u E_{z_1}\left[(x - u + z_1)^2 + 0\right] \\ &= \min_u \left[\frac{1}{2}\underbrace{(x-u)^2}_{z_1=0} + \frac{1}{2}\underbrace{(x-u+1)^2}_{z_1=1}\right] \\ &= x^2 + x + \frac{1}{2} + \min_u \left[u - \frac{2x+1}{2}\right]^2 - \frac{(2x+1)^2}{4} \\ &= x^2 + x + \frac{1}{2} - \frac{(2x+1)^2}{4} = \frac{1}{4}, \end{aligned}$$

where the minimum is attained at $u = \frac{1}{2}+x$. Hence we have $v_1(x) = \frac{1}{2}+x$ and $V_1(x) = \frac{1}{4}$ for any x. Using the formula for V_1, one obtains

$$\begin{aligned} V_0(x) &= \min_u E_{z_0}\left[(x-u+z_0)^2 + V_1(x-u+z_0)\right] \\ &= \min_u E_{z_0}\left[(x-u+z_0)^2 + \frac{1}{4}\right] \\ &= \frac{1}{4} + \min_u \left[\frac{1}{2}\underbrace{(x-u)^2}_{z_0=0} + \frac{1}{2}\underbrace{(x-u+1)^2}_{z_0=1}\right]. \end{aligned}$$

The term in the brackets is the same as the one in the previous case and hence again $v_0(x) = \frac{1}{2}+x$, but $V_0(x) = \frac{1}{4}+\frac{1}{4} = \frac{1}{2}$, for any x. Therefore one has

$$\min_{\mathcal{V}} EJ(\mathcal{V},\mathcal{Z}) = V_0(x) = \frac{1}{2} < \frac{3}{4} = \min_{\mathcal{U}} EJ(\mathcal{U},\mathcal{Z}),$$

which is a strict inequality in this case.

Example 2.18. **Trading radishes** (solution to the problem formulated in Example 2.15). The problem will be solved numerically using MATLAB. We suppose that the unit purchase price is $n = 0.50$ EUR, unit selling price is $p = 0.70$ EUR and the loss due to sales delayed to the next day after delivering radishes is $s = 0.10$ EUR or $s = 0.30$ EUR, respectively. The daily demand for radishes is uniformly distributed on the interval [0.30]. Unlike in the original formulation of this problem, we assume that the merchant cannot store more than 100 bundles of radishes overnight. However, it is only a technical assumption that facilitates the search for a solution, which should not affect the result. Furthermore, we assume that the merchant wants to sell radishes during the period of 20 days. We can formulate the problem as follows:

$$\begin{aligned} \text{maximize} \quad & J = E\sum_{i=0}^{19} 0.7\min(x_i+u_i;z_i) - 0.5u_i - 0.1x_i \\ \text{subject to} \quad & x_{i+1} = \min(\max(x_i+u_i-z_i;0);100), i=0,...,19, \\ & x_0 = 0. \end{aligned}$$

Since the highest possible demand is 30 bundles of radishes, clearly the merchant will not order more than 30 bundles in any day. The problem can then be solved using MATLAB as follows (solution for $s = 0.3$):

```
k=20; xmax=100; umax=30; z=0:1:30;
V=zeros(k+1,xmax+1); v=zeros(k+1,xmax+1);
V(:,:)=-inf; V(k+1,:)=0;
for i=k:-1:1
   for x=0:xmax
       for u=0:umax
           f=min(max(x+u-z,0),xmax);
           f0=0.7*min(x+u,z)-0.5*u-0.3*x;
           if 1/31*sum(f0+V(i+1,f+1))>V(i,x+1)
              V(i,x+1)=1/31*sum(f0+V(i+1,f+1));
              v(i,x+1)=u;
           end
       end
   end
end
```

Note that in this case, the variables f and f_0 are 31-dimensional vectors, based on the dimension of the vector z. The optimal feedback control is stored in the matrix v. Its values for each day are in the particular values of the state variable shown in Figure 2.5.
Notice that the number of purchased bundles of radishes decreases with the increasing number of bundles carried-over from the previous day. In addition, the volume of purchases are smaller when the overnight storage loss is higher. At the end of the planning horizon, the purchases volumes decrease because they may not be sold at all.

When changing the distribution of the random variables z_i ($i = 0, \ldots, k-1$), the value of the optimal feedback control can be changed, even in the case when the expected value of variables z_i remains unchanged. Figure 2.6 compares three different distributions of z_i:

- discrete uniform distribution on the interval $[0, 30]$,

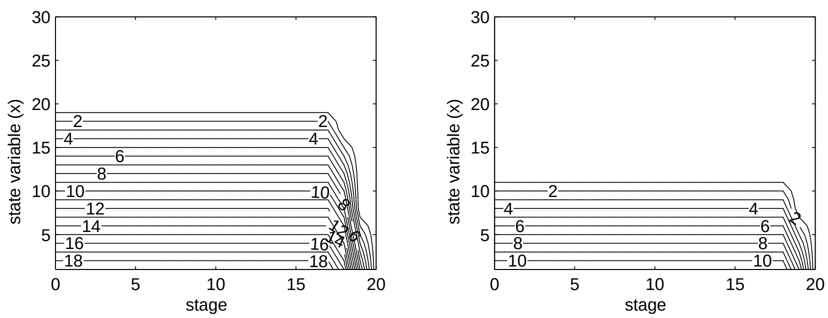

Figure 2.5: Optimal feedback control for the particular stages and values of the state variable in case $s = 0.1$ (left-hand side) and $s = 0.3$ (right-hand side).

- normal distribution with expected value 15 and standard deviation 5,
- z_i attains a deterministic value 15 (i.e. its standard deviation is zero).

By Figure 2.6 the change depends on the parameter s (the loss due to the carry-over to the next day). In case $z_i = 15$, the merchant always orders precisely as many bundles of radishes as required to maintain their number at 15 (including unsold bundles from the previous day). If s is small, the number of ordered bundles increases with standard deviation of the distribution of the random variable z_i. The reason is that the likelihood of higher demand increases in this case as well, and the losses in case of low demand are relatively small. Conversely, if s is large, the number of ordered bundles decreases with increasing standard deviation.

The MATLAB program reflects the change in the demand distribution mainly via the variable z, which is generated using the inverse cumulative distribution function applied to a uniformly distributed vector. In this case, however, the value will not be integer, therefore the

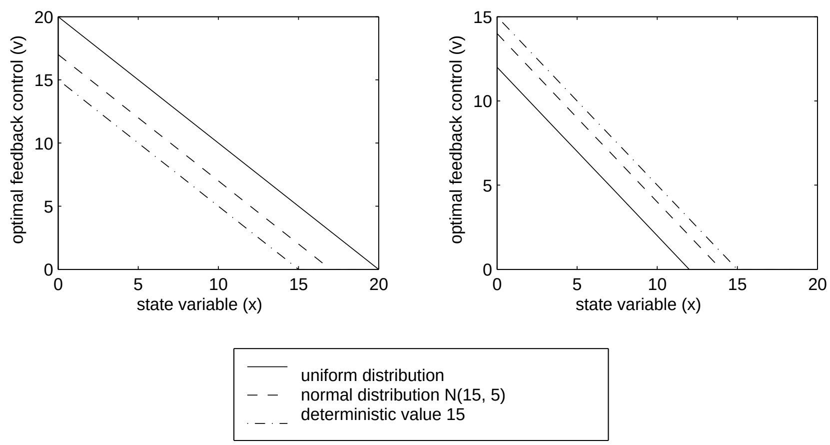

Figure 2.6: Optimal feedback control at time $i = 10$ for different distributions of the random variable z_i in the case $s = 0.1$ (right-hand side) and $s = 0.3$ (left-hand side), resp.

calculation of the new value of the state variable requires rounding off or interpolation. Hence, in order to obtain an accuracy comparable to the case of uniform distribution, partitioning of both the state and control variables has to be refined.

Decrease of the optimal feedback control at times $i = 18$ and $i = 19$ indicated in Figure 2.5 is implied by the fact that the radishes which remained unsold at the end of the sale period do not generate any profits. Hence, we can change the formulation of the problem by assuming that these bundles can be repurchased for their residual value EUR 0.15 pre bundle. The objective function has the form

$$J = E\left(\sum_{i=0}^{19} 0.7\min(x_i + u_i; z_i) - 0.5u_i - 0.1x_i + 0.15x_{20}\right).$$

As indicated by Figure 2.7, the decrease in orders during the last days is less significant in this case. The solution is obtained by replacing the command V(k+1,:)=0; by the command V(k+1,:)=0:0.15:(0.15*xmax);.

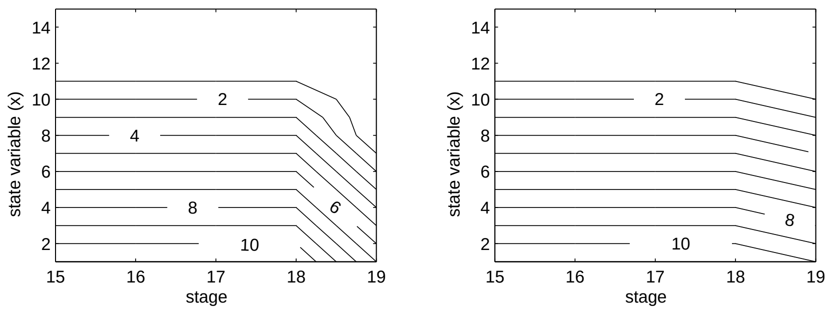

Figure 2.7: Optimal feedback control at stages from 15 to 19 and $s = 0.3$ in the case that the repurchase price of unsold radishes is zero (left-hand side) or EUR 15 (right-hand side), resp.

A possible extension of this problem would also be to consider different distributions of the demands in particular days. In this case, it would be reasonable to expect that, unlike in the autonomous case, the optimal feedback control would not stay unchanged for a fixed value of the state variable at different stages.

Example 2.19. Gambler's problem. A gambler received an offer to take part in the following game: At the beginning of the game, he receives 3 chips. Then three betting rounds follow where he can bet any amount of chips which he currently has at his disposal. The probability of a win is 2/3. In case of win, he gets back his stake plus an additional amount equal to his stake. Otherwise, he loses his stake and gets nothing. If he succeeds to have at least 5 chips after passing these three rounds, he gets EUR p, otherwise gets nothing. What is his optimal betting policy?

This problem can be formulated in the form of a stochastic optimal control problem as follows:

$$\begin{aligned} \text{maximize} \quad & J = E\, p\, h(x_3 - 5) \\ \text{subject to} \quad & x_{i+1} = x_i + z_i u_i, \quad i = 0, 1, 2, \\ & x_0 = 3, \quad 0 \le u_i \le x_i, \text{ where} \end{aligned}$$

$$h(x) = \begin{cases} 1, \text{ if } x \ge 0, \\ 0, \text{ if } x < 0 \end{cases} \qquad \text{and}$$

$$z_i = \begin{cases} 1 & \text{with probability } \frac{2}{3}, \\ -1 & \text{with probability } \frac{1}{3}. \end{cases}$$

The dynamic programming equation for this problem is

$$\begin{aligned} & V_j(x) = \max_{u \ge 0,\ u \le x,\ u \in N} EV_{j+1}(x + z_j u), \quad i = 0, 1, 2, \\ & V_3(x) = p\, h(x - 5). \end{aligned}$$

The solution for $p = 1$ is presented in Table 2.4. The optimal policy is illustrated on Chart 2.8.

Table 2.8: Solution to the Example 2.19.

i	x	u	$f_i^0 + V_{i+1}(F_i)$	$V_i(x)$	$v_i(x)$
3	≤ 4			0	
	≥ 5			1	
2	0	0	$\frac{2}{3}V_3(0+0) + \frac{1}{3}V_3(0-0) = 0$	0	0
	1	0	$\frac{2}{3}V_3(1+0) + \frac{1}{3}V_3(1-0) = 0$	0	0,1
		1	$\frac{2}{3}V_3(1+1) + \frac{1}{3}V_3(1-1) = 0$		
	2	0	$\frac{2}{3}V_3(2+0) + \frac{1}{3}V_3(2-0) = 0$	0	0,1,2
		1	$\frac{2}{3}V_3(2+1) + \frac{1}{3}V_3(2-1) = 0$		
		2	$\frac{2}{3}V_3(2+2) + \frac{1}{3}V_3(2-2) = 0$		
	3	0	$\frac{2}{3}V_3(3+0) + \frac{1}{3}V_3(3-0) = 0$	$\frac{2}{3}$	2,3
		1	$\frac{2}{3}V_3(3+1) + \frac{1}{3}V_3(3-1) = 0$		
		2	$\frac{2}{3}V_3(3+2) + \frac{1}{3}V_3(3-2) = \frac{2}{3}$		
		3	$\frac{2}{3}V_3(3+3) + \frac{1}{3}V_3(3-3) = \frac{2}{3}$		
	4	0	$\frac{2}{3}V_3(4+0) + \frac{1}{3}V_3(4-0) = 0$	$\frac{2}{3}$	1,2,3,4
		1	$\frac{2}{3}V_3(4+1) + \frac{1}{3}V_3(4-1) = \frac{2}{3}$		
		2	$\frac{2}{3}V_3(4+2) + \frac{1}{3}V_3(4-2) = \frac{2}{3}$		
		3	$\frac{2}{3}V_3(4+3) + \frac{1}{3}V_3(4-3) = \frac{2}{3}$		
		4	$\frac{2}{3}V_3(4+4) + \frac{1}{3}V_3(4-4) = \frac{2}{3}$		
	5	0	$\frac{2}{3}V_3(5+0) + \frac{1}{3}V_3(5-0) = 1$	1	0
		1	$\frac{2}{3}V_3(5+1) + \frac{1}{3}V_3(5-1) = \frac{2}{3}$		
		2	$\frac{2}{3}V_3(5+2) + \frac{1}{3}V_3(5-2) = \frac{2}{3}$		
		3	$\frac{2}{3}V_3(5+3) + \frac{1}{3}V_3(5-3) = \frac{2}{3}$		
		4	$\frac{2}{3}V_3(5+4) + \frac{1}{3}V_3(5-4) = \frac{2}{3}$		
		5	$\frac{2}{3}V_3(5+5) + \frac{1}{3}V_3(5-5) = \frac{3}{2}$		

i	x	u	$f_i^0 + V_{i+1}(F_i)$	$V_i(x)$	$v_i(x)$
1	0	0	$\frac{2}{3}V_2(0+0) + \frac{1}{3}V_2(0-0) = 0$	0	0
	1	0	$\frac{2}{3}V_2(1+0) + \frac{1}{3}V_2(1-0) = 0$	0	0,1
		1	$\frac{2}{3}V_2(1+1) + \frac{1}{3}V_2(1-1) = 0$		
	2	0	$\frac{2}{3}V_2(2+0) + \frac{1}{3}V_2(2-0) = 0$	$\frac{4}{9}$	1,2
		1	$\frac{2}{3}V_2(2+1) + \frac{1}{3}V_2(2-1) = \frac{4}{9}$		
		2	$\frac{2}{3}V_2(2+2) + \frac{1}{3}V_2(2-2) = \frac{4}{9}$		
	3	0	$\frac{2}{3}V_2(3+0) + \frac{1}{3}V_2(3-0) = \frac{2}{3}$	$\frac{2}{3}$	0,2,3
		1	$\frac{2}{3}V_2(3+1) + \frac{1}{3}V_2(3-1) = \frac{4}{9}$		
		2	$\frac{2}{3}V_2(3+2) + \frac{1}{3}V_2(3-2) = \frac{2}{3}$		
		3	$\frac{2}{3}V_2(3+3) + \frac{1}{3}V_2(3-3) = \frac{2}{3}$		
	4	0	$\frac{2}{3}V_2(4+0) + \frac{1}{3}V_2(4-0) = \frac{2}{3}$	$\frac{8}{9}$	1
		1	$\frac{2}{3}V_2(4+1) + \frac{1}{3}V_2(4-1) = \frac{8}{9}$		
		2	$\frac{2}{3}V_2(4+2) + \frac{1}{3}V_2(4-2) = \frac{2}{3}$		
		3	$\frac{2}{3}V_2(4+3) + \frac{1}{3}V_2(4-3) = \frac{2}{3}$		
		4	$\frac{2}{3}V_2(4+4) + \frac{1}{3}V_2(4-4) = \frac{2}{3}$		
	5	0	$\frac{2}{3}V_2(5+0) + \frac{1}{3}V_2(5-0) = 1$	1	0
		1	$\frac{2}{3}V_2(5+1) + \frac{1}{3}V_2(5-1) = \frac{2}{3}$		
		2	$\frac{2}{3}V_2(5+2) + \frac{1}{3}V_2(5-2) = \frac{2}{3}$		
		3	$\frac{2}{3}V_2(5+3) + \frac{1}{3}V_2(5-3) = \frac{2}{3}$		
		4	$\frac{2}{3}V_2(5+4) + \frac{1}{3}V_2(5-4) = \frac{2}{3}$		
		5	$\frac{2}{3}V_2(5+5) + \frac{1}{3}V_2(5-5) = \frac{3}{2}$		
0	3	0	$\frac{2}{3}V_1(3+0) + \frac{1}{3}V_1(3-0) = \frac{2}{3}$	$\frac{20}{27}$	1
		1	$\frac{2}{3}V_1(3+1) + \frac{1}{3}V_1(3-1) = \frac{20}{27}$		
		2	$\frac{2}{3}V_1(3+2) + \frac{1}{3}V_1(3-2) = \frac{2}{3}$		
		3	$\frac{2}{3}V_1(3+3) + \frac{1}{3}V_1(3-3) = \frac{2}{3}$		

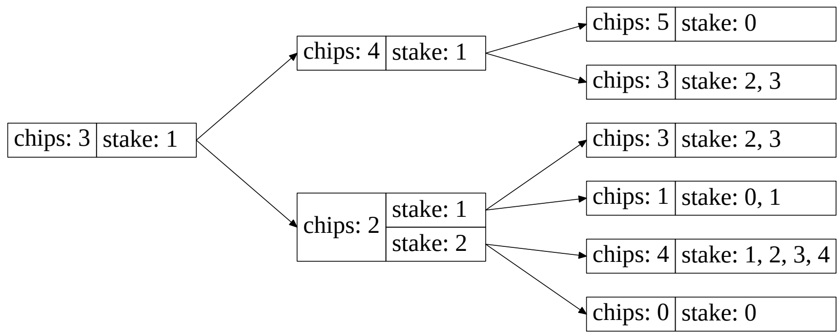

Figure 2.8: Optimal policy in the Example 2.19.

Example 2.20. Optimal selling time of shares. Assume that we possess shares of a certain firm and their total amount is $x_0 > 0$. Assume further that the price of the shares follows a stochastic process given by

$$x_{i+1} = ax_i + b + z_i, \quad i = 0, \ldots, k-1,$$

where $a \in (0, 1)$, $b > 0$ and z_i are independent, identically distributed discrete random variables with mean 0. These shares have to be sold in a single transaction at least before the k-th day. The cash obtained can be deposited in a bank account with a daily interest rate r. The aim is to maximize the amount of wealth on the k-th day.

This problem can be formulated as an optimal control problem in the following form:

$$\begin{aligned} \text{maximize} \quad & E[x_k] \\ \text{subject to} \quad & x_{i+1} = f(x_i, y_i, u_i, z_i), \quad i = 0, \ldots, k-1, \\ & y_{i+1} = y_i - u_i, \quad i = 0, \ldots, k-1, \\ & x_0 \text{ given}, \end{aligned}$$

$$y_0 = 1, \quad y_i \geq 0, \quad i = 0, \dots, k-1,$$
$$u_i \in \{0, 1\}, \quad i = 0, \dots, k-1,$$

where

$$f(x, y, u, z) = \begin{cases} ax + b + z, & \text{if } y = 1 \text{ and } u = 0, \\ (1+r)x, & \text{otherwise.} \end{cases}$$

The state variable y_i indicates whether we still possess the shares at the beginning of the i-th day ($y_i = 1$), or whether they have been already sold off ($y_i = 0$). The control variable u_i indicates whether the shares were sold during the i-th day ($u_i = 1$) or not ($u_i = 0$). The constraint on the non-negativity of the variable y_i together with the state equation ensure that the shares cannot be sold ($u_i = 0$) once we do not own them ($y_i = 0$).

We demonstrate that for each day $i = 0, \dots, k-1$ there exists a threshold c_i for the price of shares indicating whether it is optimal to sell the shares (if their price is above this threshold) or to retain them (if it is below this threshold). We also show that these thresholds represent a decreasing sequence.

The dynamic programming equation for this problem has the following form:

$$V_j(x, y) = \max_{\{u \in \{0,1\} | y-u \geq 0\}} E\big[V_{j+1}(f(x, y, u, z_j), y - u)\big],$$

where $V_k(x, y) = x$. Obviously, for each $j = 1, \dots, k-1$ one has

$$V_j(x, 0) = V_{j+1}((1+r)x, 0),$$

because for $y = 0$ we have $f(x, 0, u, z_j) = (1+r)x$. Using the boundary condition for V_k we obtain

$$V_j(x, 0) = (1+r)^{k-j} x \quad \text{and} \quad v_j(x, 0) = 0.$$

On the other hand, for $y = 1$ the dynamic programming equation is

$$\begin{aligned} V_j(x,1) &= \max_{u\in\{0,1\}} \big\{ \underbrace{E[V_{j+1}((1+r)x,0)]}_{u=1}, \underbrace{E[V_{j+1}(ax+b+z_j,1)]}_{u=0} \big\} \\ &= \max \big\{ (1+r)^{k-j}x, E[V_{j+1}(ax+b+z_j,1)] \big\}. \end{aligned}$$

Using the notation $\tilde{V}_j(x) = (1+r)^{-(k-j)}V_j(x)$, the last equation multiplied by $(1+r)^{-(k-j)}$ can be rewritten to

$$\tilde{V}_j(x,1) = \max \left\{ x, \frac{1}{1+r} E[\tilde{V}_{j+1}(ax+b+z_j,1)] \right\}.$$

For $j = k-1$ we have

$$\tilde{V}_{k-1}(x,1) = \max \{x, h_{k-1}(x)\}, \quad \text{where} \quad h_{k-1}(x) = \frac{ax+b}{1+r}$$

and we have used that $V_k(x,y) = x$. Given that $\frac{a}{1+r} < 1$, this yields

$$\tilde{V}_{k-1}(x,1) = \begin{cases} x, & \text{if } x \geq c_{k-1}, \\ h_{k-1}(x), & \text{if } x \leq c_{k-1}, \end{cases}$$

and

$$v_{k-1}(x,1) = \begin{cases} 1, & \text{if } x \geq c_{k-1}, \\ 0, & \text{if } x \leq c_{k-1}, \end{cases}$$

where

$$c_{k-1} = \frac{b}{1+r-a}.$$

Furthermore, for $j = k-2$ one has

$$\begin{aligned} \tilde{V}_{k-2}(x,1) &= \max \left\{ x, \frac{1}{1+r} E\left[\max \left\{ ax+b+z_{k-2}, \frac{a(ax+b+z_{k-2})+b}{1+r} \right\} \right] \right\} \\ &= \max \left\{ x, E\left[\max \left\{ \frac{ax+b+z_{k-2}}{1+r}, \frac{a^2x+a(b+z_{k-2})+b}{(1+r)^2} \right\} \right] \right\} \\ &= \max\{x, h_{k-2}(x)\} \end{aligned}$$

where

$$h_{k-2}(x) = E\left[\max\left\{\frac{ax+b+z_{k-2}}{1+r}, \frac{a^2x+a(b+z_{k-2})+b}{(1+r)^2}\right\}\right].$$

Given the assumptions on the random variable z_{k-2}, this can be rewritten as

$$h_{k-2}(x) = \frac{1}{N}\sum_{n=0}^{N}\max\left\{\frac{ax+b+z_{k-2}}{1+r}, \frac{a^2x+a(b+z_{k-2})+b}{(1+r)^2}\right\}.$$

Note that each of the summands is a maximum of two linear functions with derivatives $\frac{a}{1+r}$ and $\frac{a^2}{(1+r)^2}$, respectively. Consequently, each summand is an increasing, convex, piecewise linear function with derivative equal to $\frac{a}{1+r} < 1$ for sufficiently large x. Obviously, this is also true for their sum. Hence, there exists a c_{k-2} such that

$$\tilde{V}_{k-2}(x,1) = \begin{cases} x, & \text{if } x \geq c_{k-2}, \\ h_{k-2}(x), & \text{if } x \leq c_{k-2}, \end{cases}$$

and

$$v_{k-2}(x,1) = \begin{cases} 1, & \text{if } x \geq c_{k-2}, \\ 0, & \text{if } x \leq c_{k-2}. \end{cases}$$

Using mathematical induction, it can be proved that

$$\tilde{V}_j(x,1) = \max\{x, h_j(x)\}, \tag{2.100}$$

where again $h_j(x)$ is an increasing, convex, piecewise linear function with derivative equal to $\frac{a}{1+r} < 1$ for sufficiently large x. Hence there exists a c_j such that

$$\tilde{V}_j(x,1) = \begin{cases} x, & \text{if } x > c_j, \\ h_j(x), & \text{if } x < c_j. \end{cases}$$

In addition, convexity of $\tilde{V}_j(x,1)$ implies that this c_j is unique. The optimal policy for the j-th day is therefore

$$\begin{array}{ll} \text{sell the shares} & (v_j(x,1)=1),\ \text{if } x > c_j, \\ \text{retain the shares} & (v_j(x,1)=0),\ \text{if } x < c_j. \end{array}$$

Note that the state equation in this example does not represent a standard model of the dynamics of equity prices. On the other hand, the approach presented in this example can be applied in quantitative finance when pricing American options (i.e. options which can be exercised before their expiration date) by means of dynamic programming.

Example 2.21. Optimal consumption as a stochastic problem. The optimal consumption problem on infinite time horizon will now be extended by considering stochastic disturbances to the production function (for example production or technological shocks):

$$\begin{array}{rl} \text{maximize} & \displaystyle\sum_{i=0}^{\infty} \beta^i \ln u_i \\ \text{subject to} & x_{i+1} = e^{z_i}\, \alpha x_i - u_i, \quad i = 0, 1, \ldots, \\ & x_0 = a, \\ & \displaystyle\lim_{k\to\infty} x_k \geq 0, \end{array}$$

where z_i are independent and identically distributed discrete random variables with mean 0. Assume that the value z_i is known at stage i.[6]

The dynamic programming equation for this problem has the following form:

$$V(x) = \max_u [\ln u + \beta E V(e^z\, \alpha x - u)].$$

It is easy to verify that the functions

$$v(x) = (1-\beta)\, e^z \alpha x$$

[6]It is a very simplified example representing the class of models which are typically called as real business cycle models.

and

$$V(x) = \frac{1}{1-\beta}\ln x + \frac{1}{1-\beta}\ln(\alpha(1-\beta)) + \frac{\beta}{(1-\beta)^2}\ln\alpha\beta + \frac{1}{1-\beta}z$$

satisfy the dynamic programming equation. Substituting the optimal feedback control into the state equation yields

$$x_{i+1} = e^{z_i}\,\alpha\beta x_i,$$

i.e.

$$\ln x_{i+1} = \ln(\alpha\beta) + \ln x_i + z_i.$$

We can see that $\ln x_i$ follows a first order autoregressive process. This process is albeit non-stationary, because the value of the coefficient corresponding to $\ln x_i$ is one (i.e. the process has a unit root). It means that any shocks have a persistent impact on capital as well as consumption. This non-stationarity is caused by the linearity of the production function. The Cobb-Douglas production function would yield a stationary autoregressive process.

2.5.5 Exercises

Exercise 2.33. Modify the definition of admissible policies from this subchapter for the case of a deterministic problem with constraints (2.1)–(2.6) and prove (2.82) for this problem.

Exercise 2.34. Assume that you have a EUR ($a > 0$) you want to bet. In case of win, your profit is equal to the stake, otherwise you lose your stake. The probability of a win is $\frac{3}{4}$. The stake need not be integer but the value of the stake cannot exceed your current wealth. The number of rounds is k. The objective is to maximize the expected utility from your final wealth. Your utility function is $U(x) = \sqrt{x}$.

a) Formulate this problem as a stochastic optimal control problem.

b) By means of mathematical induction prove that the value function has the form $V_i(x) = c_i\sqrt{x}$. Using this result, show that your optimal policy in all stages is to bet $\frac{4}{5}$ of the money that you have at your disposal at that time.

Exercise 2.35. Assets and liquidity management in a bank. A commercial bank has to solve the following problem at the beginning of each week: Money from clients can be invested either in assets with lower liquidity (with a maturity of one week) but higher yield (4 %), or can be deposited in the central bank, where it is anytime immediately at the bank's disposal, but the yield is only 3 %. If these funds are available to the bank at the time of a demand for a new loan, they can be granted as a loan to a client with the interest rate of 7 %. This is however dependent on the current demand for loans which is random. If the loan is not granted immediately to the client when demanded due to an insufficient amount of liquid funds, the client goes to another bank. Based on the prediction of cash flows, the bank expects the inflow of EUR 1 million regularly on a weekly basis. This problem can be formulated as follows:

$$\min E\left[\sum_{i=0}^{3} 0.01(x_i + u_i + 1 - z_i)^+ + 0.03(x_i + u_i + 1 - z_i)^-\right],$$

$$x_{i+1} = x_i + u_i + 1 - z_i, \quad i = 0, 1, 2,$$

$$x_0 = 0,$$

$$z_i = \begin{cases} 1{,}5 & \text{with probability } 1/3, \\ 1{,}0 & \text{with probability } 1/3, \\ 0{,}5 & \text{with probability } 1/3, \end{cases}$$

where x_i denotes the current amount of liquid funds deposited within the central bank at the beginning of the i-th week, u_i represents the increase of these funds during the i-th week and z_i denotes the demand for loans in the i-th week. The objective function represents the opportunity costs. Find the optimal policy using the dynamic programming equation.

Exercise 2.36. Choice of a pension fund. A pension saver is deciding, whether he should invest his money to a conservative fund or to a dynamic fund. The yield of the conservative fund is 5% p.a. without any risk. In the dynamic fund, the yield is 15% p.a. with probability 1/2 and zero with probability 1/2. The saver wants to maximize the expected utility from his investment in two-year horizon, where the utility function of consumption c is $-e^{-c}$. This problem can be formulated as follows:

$$\begin{aligned}
&\max E\Big[-e^{-y_2-z_2}\Big],\\
&y_{i+1} = 1.05(y_i+z_i)u_i, \quad i=0,1,\\
&z_{i+1} = r_i(y_i+z_i)(1-u_i), \quad i=0,1,\\
&u_i \in [0,1],\\
&y_0+z_0 = a > 0,\\
&r_i = \begin{cases} 1.15 & \text{with probability } 1/2,\\ 1.00 & \text{with probability } 1/2. \end{cases}
\end{aligned}$$

where the amount of money invested in the conservative and dynamic fund is denoted by y and z, respectively. Find the optimal policy. Show that under certain conditions, this optimal policy follows the rule: "In the first period, invest everything in the dynamic fund. If its yield is 0%, keep the money in this fund during the second period as well. If the yield is 15%, do not take more risk and put all the money in the conservative fund." Use that by introducing a new variable $x_i = y_i + z_i$, the problem can be transformed in the form:

$$\begin{aligned}
&\max E\Big[-e^{-x_2}\Big],\\
&x_{i+1} = 1.05x_iu_i + r_ix_i(1-u_i), \quad i=0,1,\\
&u_i \in [0,1],\\
&x_0 = a > 0,\\
&r_i = \begin{cases} 1.15 & \text{with probability } 1/2,\\ 1.00 & \text{with probability } 1/2. \end{cases}
\end{aligned}$$

Exercise 2.37. Consider the following modification of Exercise 2.36:

$$
\begin{aligned}
&\max E[\ln x_2],\\
&x_{i+1} = 1.05 x_i u_i + r_i x_i (1-u_i), \quad i = 0,1,\\
&u_i \in [0,1],\\
&x_0 = a > 0,\\
&r_i = \begin{cases} 1.15 & \text{with probability } 1/2,\\ 1.00 & \text{with probability } 1/2. \end{cases}
\end{aligned}
$$

Decide what portion of the money should be optimally invested in the conservative fund (i.e. the fund yielding 5%) at the beginning of the second period. Does the result depend on the yield in the first period?

Exercise 2.38. Solve the following problem using the dynamic programming equation:

$$
\begin{aligned}
&\max E\left[x_3^2 + \sum_{i=1}^{2}(x_i^2 + u_i^2)\right],\\
&x_{i+1} = x_i + u_i + z_i, \quad i = 1,2,\\
&u_i \in \{-1,0,1\},\\
&x_1 = 1,\\
&\text{for } x_i > 0: \quad z_i = \begin{cases} -1 & \text{with probability } 1/2,\\ 0 & \text{with probability } 1/2,\\ 1 & \text{with probability } 0, \end{cases}\\
&\text{for } x_i = 0: \quad z_i = \begin{cases} -1 & \text{with probability } 1/2,\\ 0 & \text{with probability } 0,\\ 1 & \text{with probability } 1/2, \end{cases}\\
&\text{for } x_i < 0: \quad z_i = \begin{cases} -1 & \text{with probability } 0,\\ 0 & \text{with probability } 1/2,\\ 1 & \text{with probability } 1/2. \end{cases}
\end{aligned}
$$

Exercise 2.39. The stochastic linear-quadratic problem on a finite time horizon has the following form:

$$\begin{aligned} \text{minimize} \quad & E\left[\sum_{i=0}^{k-1} \beta^i \left(x_i^T Q_i x_i + u_i^T R_i u_i\right)\right], \\ \text{subject to} \quad & x_{i+1} = Ax_i + Bu_i + z_i, \quad i = 0, \ldots, k-1, \end{aligned}$$

where $x_i \in \mathbb{R}^n$, $u_i \in \mathbb{R}^m$, Q_i and R_i are $n \times n$ and $m \times m$ symmetric matrices, respectively and $R_i > 0$, $Q_i \geq 0$. In addition, A_i and B_i are $n \times n$ and $n \times m$ matrices, respectively and $z_i \in \mathbb{R}^n$ are independent vectors of discrete random variables identically distributed with mean 0 and covariance matrix Σ. Prove that the optimal feedback control and the value function for this problem have the following form:

$$\begin{aligned} v_i(x_i) &= -\beta(R_i + \beta B_i^T W_{i+1} B_i)^{-1} B_i^T W_{i+1} A_i x_i \\ V_i(x_i) &= x_i^T W_i x_i + d_i. \end{aligned}$$

In addition, $W_k = 0$ and for $i = k-1, ..., 0$ one has

$$\begin{aligned} W_i &= Q_i + \beta A_i^T [W_{i+1} - \beta W_{i+1} B_i (R_i + \beta B_i^T W_{i+1} B_i)^{-1} B_i^T W_{i+1}] A_i, \\ d_i &= \frac{\beta}{1-\beta} E(z_i^T W_{i+1} z_i) = \frac{\beta}{1-\beta} \mathrm{tr}(W_{i+1} \Sigma), \end{aligned}$$

where tr denotes the trace of a matrix.

Note that despite the stochastic nature of this problem, the optimal feedback control is independent of Σ and it is the same as in the case of the respective deterministic linear-quadratic problem. This independence of the optimal feedback control on Σ is called *equivalence principle*. The underlying assumption is that the objective function is quadratic and convex in both variables, the state equation is linear and stochastic variables are independent over time. For more general types of such problems the equivalence principle might not hold, though.

Chapter 3

Maximum Principle for Discrete Problems

The farther backward you can look,
the farther forward you can see.
Winston Churchill.

In this chapter, we will discuss an alternative approach for solving optimal control problems. It will be called *variational* for several reasons.

The essential feature of this approach consists in the control optimality test. It verifies whether small variations of the tested control do not lead to an increase in the objective function, which is of course a necessary optimality condition. A charm of this approach is that by carefully exhausting of all possibilities we obtain enough conditions for determining optimal control in the sense that the number of optimization parameters equals the number of conditions.

The first-order optimization conditions in many fields of optimization could be considered as prototypes for such conditions. Such are e.g. the zero value of partial derivation of functions of several variables in the free extreme problem, the Lagrange conditions for the constrained extreme, or the Kuhn-Tucker conditions in nonlinear programming. In infinite dimensional case they correspond to the variational conditions, from which we adopt the term.

The above analogy suggests that the approach uses finer tools of mathematical analysis than before. For example, it is necessary that functions entering the problems are differentiable and hence defined on open subsets of vector spaces. Therefore, we restrict ourselves to the problem in which both the control and the state variables are from Euclidean spaces.

In such a way a wide class of optimal control problems with a finite number of time steps can be reduced to nonlinear programming problems. Necessary conditions for them are then actually only transcriptions of the Kuhn-Tucker conditions to a form reflecting the features of the recurrent nature of optimal control problems.

So, why are we talking about the maximum principle? The reason is historical. The continuous optimal control problem preceded its discrete version. There, the Pontrjagin maximum principle was derived as a basic variational optimality condition. Discrete problems then took over some components, such as the Hamilton function or adjoint variables. These concepts have proved useful and illustrative despite the fact that a complete analogy of Pontrjagin maximum principle in the discrete theory does not apply. The reason is that continuous time allows variations of arbitrarily small duration. In discrete time this is not possible.

At the first sight it is not clear that the variation approach is associated with dynamic programming which we have discussed so far. We will show, however, that there is a relation, and many objects and concepts from one theory have a natural interpretation in the other.

3.1 Notations and Formulation of the Problem

For reasons mentioned in the introduction to this chapter we will work in Euclidean spaces: the n-dimensional space will be denoted as $\mathbb{R}^n$. Its elements, n-dimensional vectors, will be understood as column vectors.

That is, for $x \in \mathbb{R}^n$,

$$x = \begin{pmatrix} x^1 \\ \vdots \\ x^n \end{pmatrix} = (x^1, \ldots, x^n)^T,$$

where ^{T} as superscript of a matrix or a vector is the symbol of transposition. Note that components of vectors will be marked by superscripts, since the subscripts are reserved for time stages.

Similarly, the functions will be defined on several dimensional Euclidean spaces with values either in the one-dimensional or the multidimensional Euclidean space. If $f : \mathbb{R}^n \to \mathbb{R}^m$, then $f = (f^1, \ldots, f^m)^T$. If such a function is continuously differentiable of r-th order, we write $f \in \mathcal{C}^r$. Under the matrix of its first derivatives we understand

$$\frac{\partial f}{\partial x} = \begin{pmatrix} \frac{\partial f^1}{\partial x^1} & \cdots & \frac{\partial f^1}{\partial x^n} \\ \vdots & & \vdots \\ \frac{\partial f^m}{\partial x^1} & \cdots & \frac{\partial f^m}{\partial x^n} \end{pmatrix}.$$

The problem which we addressed in this chapter narrows the standard problem (1.9)–(1.14) in several ways. In particular, we require that $x \in \mathbb{R}^n$ and $u \in \mathbb{R}^m$ - it is for this approach necessary. Also, certain assumptions on the sets U_i and C are needed - we assume that they are given by inequalities or equalities. Finally, we assume that the problem has no constraints on state variables. This assumption is not necessary but greatly simplifies the derivation of necessary conditions of optimality. So, we consider the problem

$$\text{maximize} \quad J := \sum_{i=0}^{k-1} f_i^0(x_i, u_i) \tag{3.1}$$

$$\text{subject to} \quad x_{i+1} = f_i(x_i, u_i), \quad i = 0, \ldots, k-1, \tag{3.2}$$

$$x_0 = a, \tag{3.3}$$

$$x_k \in C = \{x \ : \ g(x) = 0\}, \tag{3.4}$$

$$u_i \in U_i = \{u \ : \ p_i(u) \le 0\}, \ i = 0, \ldots, k-1, \tag{3.5}$$

where $f_i : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^n$, $f_i^0 : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}$, $p_i : \mathbb{R}^m \to \mathbb{R}^{m_i}$, $i = 0, \ldots, k-1$ a $g : \mathbb{R}^n \to \mathbb{R}^l$ are continuously differentiable functions. The letter l stands for the number of the components of constraint g and the letter m_i for the number of components of the constraint p_i, for $i = 0, \ldots, k-1$.

Since the problem (3.1)–(3.5) is a special case of the standard problem treated in the first two chapters, the terminology as well as the results derived there for problems with fixed terminal time apply. Due to the continuous nature of the variables the problem (3.1)–(3.5) can be seen as a nonlinear programming problem and formulated as a standard nonlinear programming problem (see section 4.1.1) in two ways.

1. way: We start directly from the description of the problem (3.1)–(3.5) and as the (multidimensional) optimization variable we chose both the vectors $u_i,\ i = 0, \ldots, k-1$, and the vectors $x_i,\ i = 1, \ldots, k$. These vectors together form a vector $z \in \mathbb{R}^{km+kn}$. Writing the problem as a standard non-linear programming one we obtain the problem

$$\text{maximize} \quad J(z) := \sum_{i=0}^{k-1} f_i^0(x_i, u_i) \tag{3.6}$$

$$\text{subject to} \quad f_i(x_i, u_i) - x_{i+1} = 0,\ i = 0, \ldots, k-1, \tag{3.7}$$

$$g(x_k) = 0, \tag{3.8}$$

$$p_i(u_i) \leq 0, \quad i = 0, \ldots, k-1, \tag{3.9}$$

where $x_0 = a$.

2. way: We take into account the nature of the optimal control problem and as the (multidimensional) optimization variable we chose the vectors $u_i,\ i = 0, \ldots, k-1$ only. Then the control $\mathcal{U} = \{u_0, \ldots, u_{k-1}\}$ is considered not as a sequence, but as a km-dimensional vector. The values $x_i,\ i = 1, \ldots, k$ of the response of the control $\mathcal{U}$ will be considered as dependent variables determined by the functions $x_i = x_i(u_0, \ldots, u_{i-1})$ defined by (3.2) and (3.3). The dependence of x_i on the part $(u_0, \ldots, u_{i-1})$ of the control $\mathcal{U}$ will be written as $x_i(\mathcal{U})$ for all i. So, we obtain a nonlinear programming problem in the space $\mathbb{R}^{km}$

$$\text{maximize} \quad J(\mathcal{U}) := \sum_{i=0}^{k-1} f_i^0(x_i(\mathcal{U}), u_i) \tag{3.10}$$

$$\text{subject to} \quad p_i(u_i) \leq 0, \quad i = 0, \ldots, k-1, \tag{3.11}$$

$$g(x_k(\mathcal{U})) = 0, \tag{3.12}$$

where $x_i(\mathcal{U})$, $i = 1, \ldots, k$, is the solution of the difference equation (3.2) with initial condition (3.3) for the control $\mathcal{U}$.

Two alternative ways of transcription of (3.1)–(3.5) yield two different nonlinear programming problems of different dimensions. The former yields the problem (3.6)–(3.9) of high dimension, but simple structure. The latter yields the problem (3.10)–(3.12) of lower dimension but a more complex structure due to the composite functions $f_i^0(x_i(\mathcal{U}), u_i)$ and $g(x_k(\mathcal{U}))$. To both of these problems we will apply Theorem 4.2 from Appendix in the next subsection. To do this the following *constraint qualification* has to be imposed on the functions defining the problem.

Definition 3.1. *Let for each $i = 0, \ldots, k-1$ the vectors*

$$\frac{\partial p_i^j}{\partial u_i}(\hat{u}_i), \quad j \in I_i(\hat{u}_i)$$

be linearly independent. Then we say that problem (3.1)–(3.5) satisfies the constraint qualification (PR) along the control $\hat{\mathcal{U}}$. Here by $I_i(\hat{u}_i)$ we understand the index set of all those $j \in \{1, \ldots, m_i\}$, for which the j-th component of the constraint p_i at the point $\hat{u}_i$ is binding, i.e. $p_i^j(\hat{u}_i) = 0$.

Remark 3.1. It is not necessary that the functions f_i, f_i^0 etc. are defined on the whole space. This assumption is made only to simplify the formulation, it is sufficient that the functions f_i, f_i^0 etc. are defined on some sufficiently large open domain.

Remark 3.2. Sometimes it is convenient to reformulate the problem into the so-called *geometric form*. We extend the state space to $\mathbb{R}^{n+1}$ by adding the zero component x^0, denote $\tilde{x} = (x^0, x)$ and in the extended state space we consider the system

$$\tilde{x}_{i+1} = \tilde{f}_i(x_i, u_i),$$

where

$$\tilde{f}_i(x_i, u_i) = (f_i^0(x_i, u_i), f_i(x_i, u_i)),$$

with the initial state

$$\tilde{a} = (0, a)$$

and the target set

$$\tilde{C} = \mathbb{R} \times C = \{(x^0, x) : \quad x^0 \in \mathbb{R}, \quad x \in C\}.$$

Since

$$x_k^0 = \sum_{i=0}^{k-1} f_i^0(x_i, u_i),$$

we can equivalently reformulate this problem into the problem of maximization of x_k^0 subject to the constraints for the extended system. Hence, the variable x_i^0 can be interpreted as a measure of the accumulated values of the objective function.

3.2 A Necessary Condition of Optimality

The main objective of this chapter is to derive necessary optimality conditions for problem (3.1)–(3.5). For the derivation we use the corollary of *the John theorem* in nonlinear programming formulated and proved as Theorem 4.2 in the Appendix. Theorem 4.2 can be applied to the problem (3.6)–(3.9) or to the problem (3.10)–(3.12). In both the cases we obtain the same result but the interpretations of the so-called *adjoint variable* are different (see Remarks 3.10 and 3.11).

3.2.1 Derivation and Formulation of Necessary Conditions

In this subsection we will start from the nonlinear programming problem of form (3.6)–(3.9). This means that we choose the maximum of the objective function with respect to both u_i, $i = 0, \ldots, k-1$ and x_i, $i = 1, \ldots, k$ considered as independent variables. Conditions (3.7) and (3.8) are the equality constraints and (3.9) represents inequality constraints. The Lagrange function corresponding to (3.6)–(3.9) has the form (see (4.3)):

$$\begin{aligned} L = \ & \psi^0 \sum_{i=0}^{k-1} f_i^0(x_i, u_i) + \sum_{i=0}^{k-1} \psi_{i+1}^T (f_i(x_i, u_i) - x_{i+1}) + \chi^T g(x_k) \\ & + \sum_{i=0}^{k-1} \lambda_i^T p_i(u_i), \end{aligned}$$

where $\psi^0 \in \mathbb{R}$, $\psi^0 \geq 0$, is the multiplier of the objective function, $\psi_{i+1} \in \mathbb{R}^n$, $i = 0, \ldots, k-1$, are multipliers of the equality constraints in (3.7), $\chi \in \mathbb{R}^l$ is the multiplier of the terminal state equality constraint (3.8) and $\lambda_i \in \mathbb{R}^{m_i}$, $\lambda_i \leq 0$, $i = 0, \ldots, k-1$, are the multipliers of the inequality constraints in (3.9) defining the sets U_i. The corresponding derivatives of the Lagrange function are of the form:

$$\frac{\partial L}{\partial x_i} = \psi^0 \frac{\partial f_i^0}{\partial x_i}(x_i, u_i) + \psi_{i+1}^T \frac{\partial f_i}{\partial x_i}(x_i, u_i) - \psi_i^T, i = 1, ..., k-1, \quad (3.13)$$

$$\frac{\partial L}{\partial x_k} = \chi^T \frac{\partial g}{\partial x_k}(x_k) - \psi_k^T, \quad (3.14)$$

$$\frac{\partial L}{\partial u_i} = \psi^0 \frac{\partial f_i^0}{\partial u_i}(x_i, u_i) + \psi_{i+1}^T \frac{\partial f_i}{\partial u_i}(x_i, u_i) + \lambda_i^T \frac{\partial p_i}{\partial u_i}(u_i), \quad (3.15)$$

where $i = 0, \ldots, k-1$. Note the special structure of (3.13) and (3.14). Indeed, equaling (3.13) and (3.14) to zero and applying the transposition operation we obtain

$$\psi_i = \psi^0 \frac{\partial f_i^{0T}}{\partial x_i}(x_i, u_i) + \frac{\partial f_i^T}{\partial x_i}(x_i, u_i)\psi_{i+1}, \; i = 1, \ldots, k-1, \qquad (3.16)$$

$$\psi_k = \frac{\partial g^T}{\partial x_k}(x_k)\chi. \qquad (3.17)$$

The equation (3.16) represents a backward difference equation with terminal condition (3.17). It allows to compute ψ_i, $i = k, ..., 1$ (for given u_i, x_i, ψ^0, χ) recurrently starting from ψ_k. The equation (3.16) is called *adjoint*, (3.17) is called *transversality condition*, and the multipliers ψ_i *adjoint variables.*

Theorem 3.1. Necessary optimality conditions of variational type. *Let $\hat{\mathcal{U}} = \{\hat{u}_0, \ldots, \hat{u}_{k-1}\}$ be an optimal control for problem (3.1)–(3.5) and let $\hat{\mathcal{X}} = \{\hat{x}_0, \ldots, \hat{x}_k\}$ be its response. Let the constraint qualification (PR) be satisfied along the control $\hat{\mathcal{U}}$. Then there exists a number $\psi^0 \geq 0$, a vector $\chi \in \mathbb{R}^l$ and vectors $\lambda_i \in \mathbb{R}^{m_i}$ for $i = 0, \ldots, k-1$ such that $(\psi^0, \chi) \neq (0, 0)$[1] and for each $i = 0, \ldots, k-1$ the variational condition*

$$\psi^0 \frac{\partial f_i^0}{\partial u_i}(\hat{x}_i, \hat{u}_i) + \psi_{i+1}^T \frac{\partial f_i}{\partial u_i}(\hat{x}_i, \hat{u}_i) + \lambda_i^T \frac{\partial p_i}{\partial u_i}(\hat{u}_i) = 0 \qquad (3.18)$$

and the complementarity condition

$$\lambda_i^T p_i(\hat{u}_i) = 0 \text{ where } \lambda_i \leq 0, \qquad (3.19)$$

hold. Here the vectors $\psi_{i+1} \in \mathbb{R}^n$, $i = k-1, \ldots, 0$, entering (3.18), satisfiy the adjoint equation (3.16) and the transversality condition (3.17)

[1]By $(\psi^0, \chi) \neq (0, 0)$ we understand that the number ψ^0 and the vector χ do not vanish simultaneously, i.e. either ψ^0 or at least one component of the vector χ does not vanish.

for given $\hat{\mathcal{U}}$, $\hat{\mathcal{X}}$ *and* ψ^0, χ, *i.e. it holds*

$$\psi_i = \frac{\partial f_i^{0T}}{\partial x_i}(\hat{x}_i, \hat{u}_i)\psi^0 + \frac{\partial f_i^T}{\partial x_i}(\hat{x}_i, \hat{u}_i)\psi_{i+1},\ i = 1, \dots, k-1, \tag{3.20}$$

$$\psi_k = \frac{\partial g^T}{\partial x_k}(\hat{x}_k)\chi. \tag{3.21}$$

Proof: Let the assumptions of the theorem be satisfied. Then, according to Theorem 4.2, there exist multipliers $\psi^0 \in \mathbb{R}$, $\psi^0 \geq 0$, $\psi_i \in \mathbb{R}^n$, $i = 1, \dots, k$ and $\chi \in \mathbb{R}^l$ not vanishing simultaneously and multipliers $\lambda_i \in \mathbb{R}^{m_i}$, $\lambda_i \leq 0$, $i = 0, \dots, k-1$, such that the relations (3.18)–(3.21) hold. Note that the variational condition (3.18) corresponds to the condition $\frac{\partial L}{\partial u_i} = 0$ and the condition (3.19) represents the complementarity condition corresponding to the inequality constraint. Further, as shown earlier, the adjoint equation (3.20) follows from the condition $\frac{\partial L}{\partial x_i} = 0$ and the transversality condition (3.21) follows from $\frac{\partial L}{\partial x_k} = 0$. It remains to prove that $(\psi^0, \chi) \neq (0, 0)$. If $\psi^0 = 0$ and also $\chi = 0$, then it would follow from (3.21) that $\psi_k = 0$, and from the adjoint equation (3.20) we would deduce that also $\psi_i = 0$ for all $i = 1, \dots, k$. This would be in contradiction with the fact that all these multipliers are not simultaneously zero.

□

3.2.2 Notes

Remark 3.3. Sometimes the conditions of Theorem 3.1 are formulated in terms of the so-called *Hamiltonian*, which is for $i = 0, \dots, k-1$ defined as follows:

$$H_i(x, u, \psi^0, \psi) = \psi^0 f_i^0(x, u) + \psi^T f_i(x, u). \tag{3.22}$$

Then, the variational condition (3.18) can be rewritten as

$$\frac{\partial H_i}{\partial u_i}(\hat{x}_i, \hat{u}_i, \psi^0, \psi_{i+1}) + \lambda_i^T \frac{\partial p_i}{\partial u_i}(\hat{u}_i) = 0, \quad i = 0, \dots, k-1, \tag{3.23}$$

the adjoint equation (3.20) as

$$\psi_i = \frac{\partial H_i^T}{\partial x_i}(\hat{x}_i, \hat{u}_i, \psi^0, \psi_{i+1}), \quad i = 1, ..., k-1 \tag{3.24}$$

and the state equation (3.2) as

$$x_{i+1} = \frac{\partial H_i^T}{\partial \psi_{i+1}}(\hat{x}_i, \hat{u}_i, \psi^0, \psi_{i+1}), \quad i = 0, ..., k-1. \tag{3.25}$$

This therminology and notation is motivated by the continuous theory.

Remark 3.4. Theorem 3.1 holds even in the case when there are no restrictions on control or the terminal state in the problem setting. If there are no restrictions on the control variable, we set $\lambda_i = 0$, $i = 0, ..., k-1$ in the theorem claim. In this case no constraint qualification is needed. If there is no restriction on the terminal state, we set $\chi = 0$ in the theorem claim. Then the transversality condition obtains the form $\psi_k = 0$.

Remark 3.5. Theorem 3.1 can be used for the search of optimal control, because it formally provides enough conditions for its determination (this property was satisfied by Theorem 4.2, from which the conditions were deduced). Theorem 3.1 is, however, only a necessary optimality condition, and hence the conditions can be met also by controls, which are not optimal. However, if the conditions are satisfied by a unique control $\hat{\mathcal{U}}$, the constraint qualification (PR) is satisfied by any admissible control and moreover, if we know that an optimal control exists, then $\hat{\mathcal{U}}$ is an optimal control.

Remark 3.6. If the sets U_i are compact and if there exists at least one admissible control, then it can be proved (see Exercise 3.11) that there exists an optimal control.

Remark 3.7. Note that if Theorem 3.1 is satisfied with multipliers ψ^0, ψ_i, $i = 1, \ldots, k$, χ and λ_i, $i = 0, \ldots, k-1$, then it is also satisfied with multipliers $c\psi^0$, $c\psi_i$, $i = 1, \ldots, k$, $c\chi$ and $c\lambda_i$, $i = 0, \ldots, k-1$, where c is a positive number. It then follows that if we have $\psi^0 \neq 0$, which can often be shown in specific cases, we can put $\psi^0 = 1$. Only in this way Theorem 3.1 is useful for determining the optimal solutions.

Remark 3.8. In some cases we can conclude $\psi^0 = 1$ already from the specific type of the functions f_i, g, p_i. Here are two such cases.

1. The problem has free terminal state, i.e. constraint (3.4) is missing. Then, Theorem 3.1 holds with $\chi = 0$ by Remark 3.4. From $(\psi^0, \chi) \neq 0$ it then follows $\psi^0 \neq 0$, and hence, according to the preceding remark we can put $\psi^0 = 1$.

2. All functions f_i, g, p_i are linear in all variables (then even the constraint qualification (PR) is not needed). In this case the problem (3.6)–(3.9) satisfies the constraint qualification conditions for the Kuhn-Tucker Theorem (see Remark 4.5) and thus $\psi^0 = 1$.

Remark 3.9. It can be easily shown from the transversality condition (see Exercise 3.5) that if any component of the terminal state is free, then the corresponding component of the terminal adjoint variable is zero. And, conversely, if any component of the terminal state is fixed, then the corresponding component of the terminal adjoint variable is free.

Remark 3.10. In nonlinear programming the necessary conditions of optimality are also sufficient for a convex problem. A similar claim holds for an optimal control problem (3.1)–(3.5), where the functions $f_i^0(x,u)$ are concave in (x,u), the sets U_i are convex for all $i = 0, ..., k-1$ and the function $g(x)$ is linear. For such a problem the necessary conditions of the Theorem 3.1 are also sufficient.

Remark 3.11. The same procedure for the derivation of necessary optimality conditions can be used to derive the necessary conditions for the optimal control problems with other kinds of constraints: $q(x_k) \leq 0$, or $h_i(x_i, u_i) \leq 0$, or $h_i(x_i) \leq 0$, where $i = 0, ..., k-1$. The corresponding necessary conditions include gradients of these constraints multiplied by non-positive multipliers as well as the corresponding complementarity conditions.

3.2.3 Alternative Derivation of the Necessary Conditions

In this subsection we present another derivation of the conditions of Theorem 3.1 which consists of problem embedding, the idea being used in the second chapter. The benefit of this procedure will be in interesting interpretation of the adjoint variable. Since this procedure will be applied in a simplified form to a problem without any constraints in subsection 3.4, we recommend to a less motivated reader to skip this subsection.

We will start from the problem formulated in the form (3.10)–(3.12) and apply Theorem 4.2 to it. Clearly, the Lagrange function corresponding to this problem is of the form

$$L = \psi^0 J(\mathcal{U}) + \chi^T g(x_k(\mathcal{U})) + \sum_{i=0}^{k-1} \lambda_i p_i(u_i),$$

where $\psi^0 \geq 0$ is the multiplier of the objective function in (3.10), $\chi \in \mathbb{R}^l$ is the multiplier of the equality constraint (3.12), and $\lambda_i \in \mathbb{R}^{m_i}$, $\lambda_i \leq 0$, $i = 0, ..., k-1$, are the multipliers of the inequality constraints (3.11).

Let $\hat{\mathcal{U}} = \{\hat{u}_0, \dots, \hat{u}_{k-1}\}$ be an optimal control and let the constraint qualification (PR) be satisfied along $\hat{\mathcal{U}}$. Let

$$\hat{\mathcal{X}} = \{\hat{x}_0, \hat{x}_1, \dots, \hat{x}_k\} = \{a, x_1(\hat{\mathcal{U}}) \dots, x_k(\hat{\mathcal{U}})\}$$

be the response to $\hat{\mathcal{U}}$. Then by Theorem 4.2 there exist multipliers $\psi^0 \geq 0$, χ and $\lambda_i \leq 0$ such that $(\psi^0, \chi) \neq (0, 0)$ and for every $i = 0, \dots, k-1$ one has

$$\frac{\partial L}{\partial u_i} = \psi^0 \frac{\partial J}{\partial u_i}(\hat{\mathcal{U}}) + \chi^T \frac{\partial g}{\partial x_k}(\hat{x}_k) \frac{\partial x_k}{\partial u_i}(\hat{\mathcal{U}}) + \lambda_i^T \frac{\partial p_i}{\partial u_i}(\hat{u}_i) = 0, \qquad (3.26)$$

and

$$\lambda_i^T p_i(\hat{u}_i) = 0. \tag{3.27}$$

Our goal is to express (3.26) through relations, in which instead of $\frac{\partial J}{\partial u_i}$ and $\frac{\partial x_k}{\partial u_i}$ only the derivations of functions f_i, f_i^0 and g will occur.

We start with $\frac{\partial x_k}{\partial u_i}$ and denote by $\mathcal{U}_i = \{u_i, \ldots, u_{k-1}\}$ the segment of the control on the interval $[i, k]$. By $\mathcal{U}_k$ we understand an empty set. Further, for any $i = 1, ..., k$, the value of the response $x_k(\mathcal{U})$ is the function of its value x_i in i-th time moment and of the control segment $\mathcal{U}_i$; this function will be denoted as

$$x_{k,i}(x_i, \mathcal{U}_i) = x_k(\mathcal{U}), \qquad i = 0, \ldots, k-1.$$

For $i = k$ define $x_{k,k} = x_k$. From the relation $x_{i+1} = f_i(x_i, u_i)$ we obtain

$$x_{k,i}(x_i, \mathcal{U}_i) = x_{k,i+1}(f_i(x_i, u_i), \mathcal{U}_{i+1}).$$

It follows

$$\frac{\partial x_k}{\partial u_i}(\hat{\mathcal{U}}) = \frac{\partial x_{k,i}}{\partial u_i}(\hat{x}_i, \hat{\mathcal{U}}_i) = \frac{\partial x_{k,i+1}}{\partial x_{i+1}}(\hat{x}_{i+1}, \hat{\mathcal{U}}_{i+1}) \frac{\partial f_i}{\partial u_i}(\hat{x}_i, \hat{u}_i), \tag{3.28}$$

where $i = 0, \ldots, k-1$. Similarly it holds

$$\frac{\partial x_{k,i}}{\partial x_i}(\hat{x}_i, \hat{\mathcal{U}}_i) = \frac{\partial x_{k,i+1}}{\partial x_{i+1}}(\hat{x}_{i+1}, \hat{\mathcal{U}}_{i+1}) \frac{\partial f_i}{\partial x_i}(\hat{x}_i, \hat{u}_i), \tag{3.29}$$

where $i = 1, \ldots, k-1$. Obviously $\frac{\partial x_{k,k}}{\partial x_k} = I$, where I is the identity matrix.

We now derive relations that will help us to find more suitable forms for $\frac{\partial J}{\partial u_i}$. Define $J_k = 0$ and for $i = 0, \ldots, k-1$ denote

$$J_i(x_i, \mathcal{U}_i) = \sum_{j=i}^{k-1} f_j^0(x_j, u_j),$$

where $\{x_i, \ldots, x_k\}$ is the response to the control $\mathcal{U}_i$ and the initial condition x_i on interval $[i, k]$. Obviously

$$J_i(x_i, \mathcal{U}_i) = f_i^0(x_i, u_i) + J_{i+1}(f_i(x_i, u_i), \mathcal{U}_{i+1}).$$

Since $f_j^0(x_j, u_j)$ does not depend on u_i for $j < i$, it holds

$$\begin{aligned} \frac{\partial J}{\partial u_i}(\hat{\mathcal{U}}) = \frac{\partial J_i}{\partial u_i}(\hat{x}_i, \hat{\mathcal{U}}_i) &= \frac{\partial f_i^0}{\partial u_i}(\hat{x}_i, \hat{u}_i) + \frac{\partial J_{i+1}}{\partial u_i}(f_i(\hat{x}_i, \hat{u}_i), \hat{\mathcal{U}}_{i+1}) = \\ &= \frac{\partial f_i^0}{\partial u_i}(\hat{x}_i, \hat{u}_i) + \frac{\partial J_{i+1}}{\partial x_{i+1}}(\hat{x}_{i+1}, \hat{\mathcal{U}}_{i+1}) \frac{\partial f_i}{\partial u_i}(\hat{x}_i, \hat{u}_i), \end{aligned} \tag{3.30}$$

where $i = 0, \ldots, k-1$. Similarly it holds

$$\frac{\partial J_i}{\partial x_i}(\hat{x}_i, \hat{\mathcal{U}}_i) = \frac{\partial f_i^0}{\partial x_i}(\hat{x}_i, \hat{u}_i) + \frac{\partial J_{i+1}}{\partial x_{i+1}}(\hat{x}_{i+1}, \hat{\mathcal{U}}_{i+1}) \frac{\partial f_i}{\partial x_i}(\hat{x}_i, \hat{u}_i), \tag{3.31}$$

where $i = 1, \ldots, k-1$. Obviously $\frac{\partial J_k}{\partial x_k} = 0$.

Substitute (3.28) and (3.30) into (3.26). We obtain (for clarity purposes the arguments of the functions are omitted):

$$\psi^0 \left(\frac{\partial f_i^0}{\partial u_i} + \frac{\partial J_{i+1}}{\partial x_{i+1}} \frac{\partial f_i}{\partial u_i} \right) + \chi^T \frac{\partial g}{\partial x_k} \left(\frac{\partial x_{k,i+1}}{\partial x_{i+1}} \frac{\partial f_i}{\partial u_i} \right) + \lambda_i^T \frac{\partial p_i}{\partial u_i} = 0$$

for each $i = 0, \ldots, k-1$, which can be rewritten as

$$\psi^0 \frac{\partial f_i^0}{\partial u_i} + \left(\psi^0 \frac{\partial J_{i+1}}{\partial x_{i+1}} + \chi^T \frac{\partial g}{\partial x_k} \frac{\partial x_{k,i+1}}{\partial x_{i+1}} \right) \frac{\partial f_i}{\partial u_i} + \lambda_i^T \frac{\partial p_i}{\partial u_i} = 0 \tag{3.32}$$

for each $i = 0, \ldots, k-1$. If we denote the expression in the parenthesis of (3.32) as

$$\psi_{i+1}^T = \psi^0 \frac{\partial J_{i+1}}{\partial x_{i+1}} + \chi^T \frac{\partial g}{\partial x_k} \frac{\partial x_{k,i+1}}{\partial x_{i+1}}, \quad i = 0, \ldots, k-1, \tag{3.33}$$

we obtain from (3.32) that

$$\psi^0 \frac{\partial f_i^0}{\partial u_i}(\hat{x}_i, \hat{u}_i) + \psi_{i+1}^T \frac{\partial f_i}{\partial u_i}(\hat{x}_i, \hat{u}_i) + \lambda_i^T \frac{\partial p_i}{\partial u_i}(\hat{u}_i) = 0, \quad i = 0, \ldots, k-1, \tag{3.34}$$

which is in fact the condition (3.18) from Theorem 3.1. From (3.33) it follows

$$\psi_i^T = \psi^0 \frac{\partial J_i}{\partial x_i} + \chi^T \frac{\partial g}{\partial x_k} \frac{\partial x_{k,i}}{\partial x_i}, \quad i = 1, \ldots, k. \tag{3.35}$$

Since

$$\frac{\partial x_{k,k}}{\partial x_k} = I \quad \text{and} \quad \frac{\partial J_k}{\partial x_k} = 0,$$

from (3.35) we obtain for $i = k$ that

$$\psi_k = \frac{\partial g^T}{\partial x_k}(\hat{x}_k)\chi. \tag{3.36}$$

If we substitute (3.29) and (3.31) onto (3.35), we obtain

$$\begin{aligned} \psi_i^T &= \psi^0 \left(\frac{\partial f_i^0}{\partial x_i} + \frac{\partial J_{i+1}}{\partial x_{i+1}} \frac{\partial f_i}{\partial x_i} \right) + \chi^T \frac{\partial g}{\partial x_k} \frac{\partial x_{k,i+1}}{\partial x_{i+1}} \frac{\partial f_i}{\partial x_i} \\ &= \psi^0 \frac{\partial f_i^0}{\partial x_i} + \left(\psi^0 \frac{\partial J_{i+1}}{\partial x_{i+1}} + \chi^T \frac{\partial g}{\partial x_k} \frac{\partial x_{k,i+1}}{\partial x_{i+1}} \right) \frac{\partial f_i}{\partial x_i} \\ &= \psi^0 \frac{\partial f_i^0}{\partial x_i} + \psi_{i+1}^T \frac{\partial f_i}{\partial x_i}, \quad i = 1, \ldots, k-1. \end{aligned}$$

To sum up, the n-dimensional vector ψ_i satisfies the backward difference equation

$$\psi_i = \frac{\partial f_i^{0T}}{\partial x_i}(\hat{x}_i, \hat{u}_i)\psi^0 + \frac{\partial f_i^T}{\partial x_i}(\hat{x}_i, \hat{u}_i)\psi_{i+1}, \quad i = 1, \ldots, k-1 \tag{3.37}$$

with the terminal condition (3.36), which is the adjoint equation (3.20) and the transversality condition (3.21) from Theorem 3.1.

Remark 3.12. The benefit of this procedure for the derivation of necessary conditions optimality formulated in Theorem 3.1 is an interesting interpretation of the adjoint variable ψ_i given by (3.35). While in the previous procedure ψ_i was the Lagrange multiplier corresponding to the state equation, this time ψ_i is the derivation of $\psi^0 J_i(x_i, \mathcal{U}_i) + \chi^T g(x_k(x_i, \mathcal{U}_i))$ with respect to x_i. In addition, this procedure reflects the idea of embedding the problem from the second chapter.

Remark 3.13. If the problem is without the terminal state constraint, that means, if there are no equality constraints in the formulation (3.10)–(3.12), then the constraint qualification (PR) represents the constraint qualification for the Kuhn–Tucker Theorem as well, and hence $\psi^0 = 1$. In this case the interpretation of the adjoint variable reads

$$\psi_i^T = \frac{\partial J_i}{\partial x_i}, \quad i = 1, \ldots, k.$$

3.3 Maximum Principle

As mentioned in introduction to this chapter, in the continuous theory, which was studied first, the so-called Pontrjagin maximum principle holds under very general conditions. Its discrete version differs from Theorem 3.1 in that the variational condition (3.18) and the complementarity condition (3.19) are replaced by the following *maximum condition*

$$\psi^0 f_i^0(\hat{x}_i, \hat{u}_i) + \psi_{i+1}^T f_i(\hat{x}_i, \hat{u}_i) = \max_{u_i \in U_i} (\psi^0 f_i^0(\hat{x}_i, u_i) + \psi_{i+1}^T f_i(\hat{x}_i, u_i)). \tag{3.38}$$

Maximum condition (3.38) is used much more comfortable than (3.18) and (3.19). Initially, several authors believed that (3.38) holds even in the discrete case under quite general assumptions. Only later it was shown that the condition (3.38) holds only subject to additional assumptions, such as those formulated in this subsection.

Theorem 3.2. Maximum principle. *Let the problem (3.1)–(3.5) satisfy the following additional assumptions:*

(a) f_i are linear in u_i, i.e. $f_i(x_i, u_i) = c_i(x_i) + d_i(x_i)u_i$;

(b) f_i^0 are concave in u_i;

(c) U_i are convex.

Let $\hat{\mathcal{U}} = \{\hat{u}_0, \ldots, \hat{u}_{k-1}\}$ be an optimal control and $\hat{\mathcal{X}} = \{\hat{x}_0, \ldots, \hat{x}_k\}$ be its response. Let the constraint qualifications (PR) be satisfied in the control $\hat{\mathcal{U}}$. Then there exists a number $\psi^0 \geq 0$ and a vector $\chi \in \mathbb{R}^l$ such *that* $(\psi^0, \chi) \neq (0,0)$ *and for each* $i = 0, \ldots, k-1$ *the maximum condition (3.38) holds. The vectors* $\psi_{i+1} \in \mathbb{R}^n$, $i = k-1, \ldots, 0$ *in (3.38) represent the solution of the adjoint equation (3.20) and the transverzality condition (3.21).*

Proof: By Theorem 3.1 there exists $(\psi^0, \chi) \neq (0,0)$, $\psi^0 \geq 0$, such that $\hat{\mathcal{U}}$ and $\hat{\mathcal{X}}$ satisfy the variational condition (3.18) and the complementarity condition (3.19) for each $i = 0, \ldots, k-1$. Here ψ_{i+1}, $i = 1, \ldots, k$ represent the solution of the adjoint equation (3.20) and the transversality condition (3.21). The statement of the theorem now follows from the application of Theorem 4.3 to the pair of conditions (3.18) and (3.19) at each particular time $i \in [0, k-1]$. Assumptions of Theorem 4.3 hold, since at any fixed $\hat{x}_i$ and ψ_{i+1} the function $u_i \mapsto \psi^0 f_i^0(\hat{x}_i, u_i) + \psi_{i+1}^T f_i(\hat{x}_i, u_i)$ is concave. □

The following example shows that in general the maximum principle does not hold. It also shows how to use Theorem 3.1.

Example 3.1. Maximum principle does not hold in general. The problem is to

$$
\begin{aligned}
\text{maximize} \quad & J = \sum_{i=0}^{k-1} (u_i^2 - 2x_i^2) \\
\text{subject to} \quad & x_{i+1} = u_i, \quad i = 0, \ldots, k-1, \\
& x_0 = 0, \\
& u_i \in [-1, 1], \quad i = 0, \ldots, k-1.
\end{aligned}
$$

Obviously, $f_i = u_i$ is linear, $U_i = U = [-1, 1]$ is convex, but the function $f_i^0 = u_i^2 - 2x_i^2$ is not concave in u_i. Therefore only Theorem 3.1 can be used. We first describe the set U by inequalities as follows

$$U = \{u : \ -u \leq 1, \ u \leq 1\}.$$

This description of U satisfies the constraint qualification (PR) in any admissible solution. Since the terminal state of this problem is free, by

Remark 3.8 we have $\psi^0 = 1$. The adjoint equation is

$$\psi_i = -4\hat{x}_i, \quad i = 1, \ldots, k-1,$$

and the transversality condition is

$$\psi_k = 0.$$

The variational condition (3.18) and complementarity conditions (3.19) are of the form

$$2\hat{u}_i + \psi_{i+1} - \lambda_i^1 + \lambda_i^2 = 0, \quad \lambda_i^1(-\hat{u}_i - 1) = 0, \quad \lambda_i^2(\hat{u}_i - 1) = 0, \tag{3.39}$$

for each $i = 0, \ldots, k-1$, where $\lambda_i^1 \leq 0$ and $\lambda_i^2 \leq 0$. Note that if $\hat{u}_i \in (-1, 1)$, then there is no binding constraint and hence the complementarity conditions yield $\lambda_i^1 = 0$ and $\lambda_i^2 = 0$; if $\hat{u}_i = 1$, the first constraint is non-binding, so $\lambda_i^1 = 0$; if $\hat{u}_i = -1$, then the second constraint is nonactive, hence $\lambda_i^2 = 0$.

For $i = k-1$ the condition (3.39) yields

$$2\hat{u}_{k-1} - \lambda_{k-1}^1 + \lambda_{k-1}^2 = 0. \tag{3.40}$$

Let us analyze possible solutions of this condition.

If $\hat{u}_{k-1} \in (-1, 1)$, then $\lambda_{k-1}^1 = \lambda_{k-1}^2 = 0$ and (3.40) holds for $\hat{u}_{k-1} = 0$.

If $\hat{u}_{k-1} = 1$, then $\lambda_{k-1}^1 = 0$ and (3.40) holds with $\lambda_{k-1}^2 = -2 \leq 0$.

If $\hat{u}_{k-1} = -1$, then $\lambda_{k-1}^2 = 0$ and (3.40) holds with $\lambda_{k-1}^1 = -2 \leq 0$.

Therefore, only the following three possibilities

$$\hat{u}_{k-1} = \begin{cases} -1 \\ 0 \\ 1 \end{cases}$$

can occur at $i = k-1$. We now investigate the cases $i = 0, \ldots, k-2$. Substituting the adjoint equation into the first condition of (3.39), we obtain

$$2\hat{u}_i - 4\hat{x}_{i+1} - \lambda_i^1 + \lambda_i^2 = 0. \tag{3.41}$$

By substituting the state equation into (3.41), we have

$$2\hat{u}_i - 4\hat{u}_i - \lambda_i^1 + \lambda_i^2 = -2\hat{u}_i - \lambda_i^1 + \lambda_i^2 = 0, \quad \text{i.e.}$$

$$2\hat{u}_i + \lambda_i^1 - \lambda_i^2 = 0. \tag{3.42}$$

Let us now analyze each possibility.

If $\hat{u}_i \in (-1, 1)$, then $\lambda_i^1 = 0$ and $\lambda_i^2 = 0$, hence (3.42) holds for $\hat{u}_i = 0$.
If $\hat{u}_i = 1$, then $\lambda_i^1 = 0$, hence (3.42) does not hold for any $\lambda_i^2 \leq 0$.
If $\hat{u}_i = -1$, then $\lambda_i^2 = 0$, hence (3.42) does not hold for any $\lambda_i^1 \leq 0$.

Overall, we identified three candidates for the optimal control:

$$\begin{aligned}
\mathcal{U}_1 &= \{\underbrace{0, 0, \ldots, 0}_{k-1}, -1\}, & \mathcal{X}_1 &= \{\underbrace{0, 0, \ldots, 0}_{k}, -1\}, & J &= 1,\\
\mathcal{U}_2 &= \{\underbrace{0, 0, \ldots, 0}_{k-1}, 1\}, & \mathcal{X}_2 &= \{\underbrace{0, 0, \ldots, 0}_{k}, 1\}, & J &= 1,\\
\mathcal{U}_3 &= \{\underbrace{0, 0, \ldots, 0}_{k}\}, & \mathcal{X}_3 &= \{\underbrace{0, 0, \ldots, 0}_{k+1}\}, & J &= 0.
\end{aligned}$$

By Remark 3.6 there exists an optimal control for this problem, and therefore it must be among these candidates. A comparison of the objective function values shows that optimal are the first two controls.

What would we obtain by applying maximum principle to this problem? The maximum principle (3.38) reads

$$(u_i^2 - 2\hat{x}_i^2) + \psi_{i+1} u_i \to \max_{u_i \in [-1,1]}.$$

We are looking for the maximum of a strictly convex function restricted to the closed interval $[-1, 1]$. The maximum can be achieved only in the boundary points of the interval, i.e. $u_i = \pm 1$. That is, if a control satisfies the maximum principle condition, then it takes only the values $u_i = \pm 1$. It follows that the optimal controls $\mathcal{U}_1$ and $\mathcal{U}_2$ do not satisfy the maximum principle condition, since they achieve also values other than ± 1.

3.4 Necessary Conditions for the Problem without Constraints

In this subsection we will discuss the features of the problem without constraints on control, state variables and the terminal state. For such problems the conditions of Theorem 3.1 have a particularly simple form and even their derivation procedure analogous to that of subsection 3.2.3 is quite simple. Simplicity of this problem allows us to understand the link between the dynamic programming equation and the necessary condition from Theorem 3.1 as well as to economically interpret the adjoint variable.

3.4.1 Formulation and Derivation of Necessary Conditions

Consider the problem

$$\text{maximize} \quad J = \sum_{i=0}^{k-1} f_i^0(x_i, u_i) \tag{3.43}$$

$$\text{subject to} \quad x_{i+1} = f_i(x_i, u_i), \quad i = 0, \ldots, k-1, \tag{3.44}$$

$$x_0 = a, \tag{3.45}$$

where $f_i : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^n$, $f_i^0 : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}$ are continuously differentiable functions. Obviously, this problem represents a special case of the problem (3.1)–(3.5), and hence all the results of the preceding subsections can be applied. Theorem 3.1 takes the following form:

Theorem 3.3. The necessary optimality conditions. *Let $\hat{\mathcal{U}} = \{\hat{u}_0, \ldots, \hat{u}_{k-1}\}$ be an optimal control for the problem (3.43)–(3.45), let $\hat{\mathcal{X}} = \{\hat{x}_0, \ldots, \hat{x}_k\}$ be its response. Let the sequence $\{\psi_1, \ldots, \psi_k\}$ be the solution of the adjoint equation and the transversality condition:*

$$\psi_i = \frac{\partial f_i^{0T}}{\partial x_i}(\hat{x}_i, \hat{u}_i) + \frac{\partial f_i^T}{\partial x_i}(\hat{x}_i, \hat{u}_i)\psi_{i+1}, \; i = 1, ..., k-1, \tag{3.46}$$

$$\psi_k = 0. \tag{3.47}$$

Then for each $i = 0, \ldots, k-1$ *one has*

$$\frac{\partial f_i^0}{\partial u_i}(\hat{x}_i, \hat{u}_i) + \psi_{i+1}^T \frac{\partial f_i}{\partial u_i}(\hat{x}_i, \hat{u}_i) = 0. \tag{3.48}$$

The theorem can be formally deduced from Theorem 3.1 for $\chi = 0$ and $\lambda_i = 0$ for $i = 0, ..., k-1$. Now, however, we provide a direct proof based on the idea of the proof from subsection 3.2.3. For the problem (3.43)–(3.45) it is very simple and also allows us to obtain the results needed for the interpretation of the adjoint variable.

Proof: We start with the problem

$$\text{maximize } J(\mathcal{U}) := \sum_{i=0}^{k-1} f_i^0(x_i(\mathcal{U}), u_i),$$

where $x_i(\mathcal{U})$, $i = 1, \ldots, k$, is the solution of the difference equation (3.44) and the initial condition (3.45) for the control $\mathcal{U}$. Let $\hat{\mathcal{U}} = \{\hat{u}_0, \ldots, \hat{u}_{k-1}\}$ and $\hat{\mathcal{X}} = \{\hat{x}_0, \ldots, \hat{x}_k\} = \{x_0(\hat{\mathcal{U}}), \ldots, x_k(\hat{\mathcal{U}})\}$ be the optimal control and its response. Then, by the necessary conditions for the maximum of the function $J(\mathcal{U})$ it holds that

$$\frac{\partial J}{\partial u_i}(\hat{\mathcal{U}}) = 0, \quad i = 0, ..., k-1. \tag{3.49}$$

Define $J_k := 0$ and denote

$$J_i(x_i, \mathcal{U}_i) := \sum_{j=i}^{k-1} f_j^0(x_j, u_j), \quad i = 0, \ldots, k-1,$$

where $\mathcal{U}_i$ is a segment of the control $\mathcal{U}$ on the interval $[i, k]$ and $\{x_i, \ldots, x_k\}$ is the response to the control $\mathcal{U}_i$ and the initial condition x_i. Obviously

$$J_i(x_i, \mathcal{U}_i) = f_i^0(x_i, u_i) + J_{i+1}(f_i(x_i, u_i), \mathcal{U}_{i+1}).$$

Since $f_j^0(x_j, u_j)$ do not depend on u_i for $j < i$, we obtain from (3.49) and from the foregoing relation

$$\begin{aligned} 0 &= \frac{\partial J}{\partial u_i}(\hat{\mathcal{U}}) = \frac{\partial J_i}{\partial u_i}(\hat{x}_i, \hat{\mathcal{U}}_i) \\ &= \frac{\partial f_i^0}{\partial u_i}(\hat{x}_i, \hat{u}_i) + \frac{\partial J_{i+1}}{\partial x_{i+1}}(\hat{x}_{i+1}, \hat{\mathcal{U}}_{i+1})\frac{\partial f_i}{\partial u_i}(\hat{x}_i, \hat{u}_i), \end{aligned} \tag{3.50}$$

for $i = 0, \ldots, k-1$. Similarly it holds

$$\frac{\partial J_i}{\partial x_i}(\hat{x}_i, \hat{\mathcal{U}}_i) = \frac{\partial f_i^0}{\partial x_i}(\hat{x}_i, \hat{u}_i) + \frac{\partial J_{i+1}}{\partial x_{i+1}}(\hat{x}_{i+1}, \hat{\mathcal{U}}_{i+1})\frac{\partial f_i}{\partial x_i}(\hat{x}_i, \hat{u}_i), \tag{3.51}$$

for $i = 1, \ldots, k-1$. Obviously $\frac{\partial J_k}{\partial x_k} = 0$. Denote

$$\frac{\partial J_{i+1}}{\partial x_{i+1}}(\hat{x}_{i+1}, \hat{\mathcal{U}}_{i+1}) = \psi_{i+1}^T, \quad i = 0, \ldots, k-1, \tag{3.52}$$

which means that

$$\psi_i = \frac{\partial J_i}{\partial x_i}(\hat{x}_i, \hat{\mathcal{U}}_i)^T, \quad i = 1, \ldots, k. \tag{3.53}$$

If we substitute (3.52) into (3.50), we obtain

$$\frac{\partial f_i^0}{\partial u_i}(\hat{x}_i, \hat{u}_i) + \psi_{i+1}^T\frac{\partial f_i}{\partial u_i}(\hat{x}_i, \hat{u}_i) = 0, \quad i = 0, \ldots, k-1, \tag{3.54}$$

which is actually condition (3.48) from Theorem 3.3. Since $\frac{\partial J_k}{\partial x_k} = 0$, from (3.52) for $i = k$, we obtain

$$\psi_k = 0,$$

which is the transversality condition (3.47). If we substitute (3.52) into (3.51), we obtain

$$\psi_i^T = \frac{\partial f_i^0}{\partial x_i}(\hat{x}_i, \hat{u}_i) + \psi_{i+1}^T\frac{\partial f_i}{\partial x_i}(\hat{x}_i, \hat{u}_i), \quad i = 1, \ldots, k-1,$$

which is the transposition of the adjoint equation (3.46). □

3.4.2 Relation of the Dynamic Programming Equation to the Necessary Conditions

In the second chapter we discussed the dynamic programming equation, which, according to Theorem 2.4, represents a necessary and sufficient optimality condition and for the problem (3.43)–(3.45) it takes the special form

$$\begin{aligned} V_j(x) &= \max_{u \in \mathbb{R}^m} [f_j^0(x,u) + V_{j+1}(f_j(x,u))] \qquad (3.55) \\ &= f_j^0(x, v_j(x)) + V_{j+1}(f_j(x, v_j(x))), j = 0, \ldots, k-1, \\ V_k(x) &= 0, \text{ for all } x \in \mathbb{R}^n. \qquad (3.56) \end{aligned}$$

On the other hand, in the previous subsection, the necessary optimality condition of variational type were formulated for the problem (3.43)–(3.45) in Theorem 3.3. It is evident that among these, seemingly quite different formulations, must be some connection. We show that under certain technical assumptions, it is possible to derive the conditions of Theorem 3.3 from the dynamic programming equation. This derivation provides us an additional interpretation of the adjoint variable.

Let $\hat{\mathcal{U}} = \{\hat{u}_0, ..., \hat{u}_{k-1}\}$, $\hat{\mathcal{X}} = \{\hat{x}_0, ..., \hat{x}_k\}$ be an optimal control and its response for the problem (3.43)–(3.45). Let $V_i(x)$, $v_i(x)$ be the value function and an optimal feedback control for this problem. Assume that for each $i = 1, ..., k$ the function $V_i(x)$ is defined and continuously differentiable in a neighborhood of $\hat{x}_i$. Obviously, for each $i = 0, ..., k-1$ we have

$$\hat{x}_{i+1} = f_i(\hat{x}_i, \hat{u}_i), \quad \hat{u}_i = v_i(\hat{x}_i)$$

and by (3.55) we obtain

$$\begin{aligned} V_i(\hat{x}_i) &= \max_u [f_i^0(\hat{x}_i, u) + V_{i+1}(f_i(\hat{x}_i, u))] \\ &= f_i^0(\hat{x}_i, \hat{u}_i) + V_{i+1}(f_i(\hat{x}_i, \hat{u}_i)). \qquad (3.57) \end{aligned}$$

We can apply the first order necessary optimality condition on the maximization problem contained in (3.57) at the point $u_i = \hat{u}_i$. We obtain

$$\frac{\partial f_i^0}{\partial u_i}(\hat{x}_i, \hat{u}_i) + \frac{\partial V_{i+1}}{\partial x_{i+1}}(\hat{x}_{i+1})\frac{\partial f_i}{\partial u_i}(\hat{x}_i, \hat{u}_i) = 0. \tag{3.58}$$

For each $i = 0, ..., k-1$ denote

$$\frac{\partial V_{i+1}}{\partial x_{i+1}}(\hat{x}_{i+1}) = \psi_{i+1}^T, \tag{3.59}$$

which actually means that for each $i = 1, ..., k$, we have

$$\psi_i = \frac{\partial V_i^T}{\partial x_i}(\hat{x}_i). \tag{3.60}$$

If we substitute (3.59) into (3.58), we obtain the variational condition (3.48). It remains to show that ψ_{i+1} defined by (3.60) satisfies the adjoint equation (3.46) and the transversality condition (3.47).

By (3.56) we have $V_k(x) = 0$ for any x, hence (3.59) yields the transversality condition (3.47) for $i = k-1$, i.e. $\psi_k = 0$. We now choose an arbitrary $i \in [1, k-1]$ and consider the dynamic programming equation in an arbitrary x_i and in $\hat{u}_i = v_i(\hat{x}_i)$. Since x_i is arbitrary and, therefore, $\hat{u}_i$ may not maximize (3.57) for $x = x_i$, we have

$$-V_i(x_i) + f_i^0(x_i, \hat{u}_i) + V_{i+1}(f_i(x_i, \hat{u}_i)) \leq 0. \tag{3.61}$$

By (3.57) in $x_i = \hat{x}_i$ we have

$$-V_i(\hat{x}_i) + f_i^0(\hat{x}_i, \hat{u}_i) + V_{i+1}(f_i(\hat{x}_i, \hat{u}_i)) = 0. \tag{3.62}$$

Combining (3.61) and (3.62) we obtain

$$\begin{aligned} &-V_i(\hat{x}_i) + f_i^0(\hat{x}_i, \hat{u}_i) + V_{i+1}(f_i(\hat{x}_i, \hat{u}_i)) \\ &= \max_{x_i}[-V_i(x_i) + f_i^0(x_i, \hat{u}_i) + V_{i+1}(f_i(x_i, \hat{u}_i))]. \end{aligned} \tag{3.63}$$

We can apply the first order necessary optimal conditions to the maximization problem contained in (3.63) at the point $x_i = \hat{x}_i$. We obtain

$$\begin{aligned} 0 &= -\frac{\partial V_i}{\partial x_i}(\hat{x}_i) + \frac{\partial f_i^0}{\partial x_i}(\hat{x}_i, \hat{u}_i) + \frac{\partial V_{i+1}}{\partial x_i}(f_i(\hat{x}_i, \hat{u}_i)) \\ &= -\frac{\partial V_i}{\partial x_i}(\hat{x}_i) + \frac{\partial f_i^0}{\partial x_i}(\hat{x}_i, \hat{u}_i) + \frac{\partial V_{i+1}}{\partial x_{i+1}}(\hat{x}_{i+1})\frac{\partial f_i}{\partial x_i}(\hat{x}_i, \hat{u}_i). \end{aligned} \tag{3.64}$$

Substituting (3.59) at ψ_i and ψ_{i+1} into (3.64) we obtain

$$0 = -\psi_i^T + \frac{\partial f_i^0}{\partial x_i}(\hat{x}_i, \hat{u}_i) + \psi_{i+1}^T \frac{\partial f_i}{\partial x_i}(\hat{x}_i, \hat{u}_i),$$

which is the transposition of the adjoint equation (3.46).

Remark 3.14. It follows from (3.60) and (3.53) that both the gradients of the value function $V_i(x)$ as well as the gradients with respect to x of the objective function $J_i(x, \mathcal{U}_i)$ for the problem $D_i(x)$, where $i = 1, ..., k-1$, are solution of the same difference equation (3.46) and the same initial condition (3.47). The uniqueness of the solution for (3.46), (3.47) implies that

$$\frac{\partial V_i^T}{\partial x_i}(\hat{x}_i) = \frac{\partial J_i^T}{\partial x_i}(\hat{x}_i, \hat{\mathcal{U}}_i), \quad i = 1, \ldots, k, \tag{3.65}$$

what is actually an expression of the *envelope theorem* for problems $D_i(x)$ corresponding to (3.43)–(3.45). The envelope theorem is treated in more details in subsection 4.2.

3.4.3 Economic Interpretation of the Adjoint Variable

It follows from the proof of Theorem 3.1 that the adjoint variable can be interpreted as the Lagrange multiplier corresponding to the difference state equation. On the other hand, in the previous subsection the relation (3.60) was derived for the adjoint variable, from which it follows that the adjoint variable is the gradient of the value function (with respect to the state variable). This relation allows us economically interpret the optimality conditions in the problems with economic context.

First, we derive that the adjoint variable has a price dimension. In fact, if the objective function has a dimension of a certain economic value (such as utility, profit, revenue, expenses), i.e. dimensionally represents $[J] = $ [price].[amount] and the state variable has a dimension of amount $[x] = $ [amount], then the adjoint variable has a dimension of price, since

$$[\psi] = \left[\frac{\partial V}{\partial x}\right] = \frac{[\text{price}][\text{amount}]}{[\text{amount}]} = [\text{price}].$$

Using the Taylor theorem we obtain

$$V_i(\hat{x}_i + \Delta x) - V_i(\hat{x}_i) = \frac{\partial V_i}{\partial x}(\hat{x}_i)\Delta x + o(\Delta x) \approx \psi_i^T \Delta x.$$

If Δx represents a unit increase of capital, then ψ_i expresses the linear part of contribution to the value function, i.e., the marginal value of the one unit of capital stock at the i-th stage. Each ψ_i shows how valuable it is to increase the state variable x_i by a small unit, provided that in the remaining stages we behave optimally. Here ψ_i is referred to as *shadow price*, because it does not represent the market price prevailing at the capital market, but only a corporate, internal, accounting price, assessing an additional capital unit.

Note that the maximum condition can be economically interpreted as well. Actually, H_i is the sum of the profit for the i- th stage and the value of the capital stock at the end of the i-th stage:

$$H_i = f_i^0 + \psi_{i+1}^T f_i = f_i^0 + \psi_{i+1}^T x_{i+1},$$

that should be maximized at each stage.

3.4.4 The Euler Equation

The Euler, or Euler–Lagrange equation is a necessary optimality condition for the functional extremum in calculus of variations. Although the equation was derived by Euler and Lagrange for continuous problems, the equation name is also used for necessary optimality conditions in context of certain discrete optimization problems. In this subsection we formulate such problems, derive the corresponding Euler equation and show connections with the discrete optimal control problems.

Consider the problem

$$\text{maximize} \quad \sum_{i=0}^{k-1} G_i(x_i, x_{i+1}), \tag{3.66}$$

$$\text{where} \quad x_0 = a. \tag{3.67}$$

Here we assume that $G_i : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ are continuously differentiable and $a \in \mathbb{R}^n$ is a given vector. The problem (3.66)–(3.67) will be called the *discrete variational problem*, or the discrete problem of calculus of variations. It is clear that the objective function

$$J(x_1, ..., x_k) := \sum_{i=0}^{k-1} G_i(x_i, x_{i+1})$$

for the discrete variational problem is a function of k n-dimensional variables, and hence at the point $(\hat{x}_1, ..., \hat{x}_k)$, where it achieves an extremum, the following necessary optimality conditions hold:

$$\frac{\partial J}{\partial x_i} = \frac{\partial G_{i-1}}{\partial x_i}(\hat{x}_{i-1}, \hat{x}_i) + \frac{\partial G_i}{\partial x_i}(\hat{x}_i, \hat{x}_{i+1}) = 0, i = 1, ..., k-1, \tag{3.68}$$

$$\frac{\partial J}{\partial x_k} = \frac{\partial G_{k-1}}{\partial x_k}(\hat{x}_{k-1}, \hat{x}_k) = 0. \tag{3.69}$$

In the economic literature equation (3.68) is known as *Euler equation* for the problem (3.66)–(3.67). It represents a second order difference equation which together with the initial condition (3.67) and the condition (3.69) allows to find candidates for the optimal problem solution.

Note that the discrete variational problem can be converted to the discrete optimal control problem

$$\text{maximize} \quad \sum_{i=0}^{k-1} G_i(x_i, u_i) \tag{3.70}$$

$$\text{subject to} \quad x_{i+1} = u_i, \quad i = 0, ..., k-1, \tag{3.71}$$

$$x_0 = a, \tag{3.72}$$

which represents a problem without control constraints and with a special state equation (3.71).

We now show, how the necessary optimality conditions for the optimal control problem (3.70)–(3.72) are related to the Euler equation. Applying Theorem 3.3 to (3.70)–(3.72) and substituting $x_{i+1} = u_i$ we obtain that the solution ψ_i, $i = 1, ..., k$, of the adjoint equation and the transversality condition

$$\psi_i = \frac{\partial G_i^T}{\partial x_i}(\hat{x}_i, \hat{x}_{i+1}), \tag{3.73}$$

$$\psi_k = 0, \tag{3.74}$$

together with the optimal response satisfy

$$\frac{\partial G_i}{\partial x_{i+1}}(\hat{x}_i, \hat{x}_{i+1}) + \psi_{i+1}^T = 0, \quad i = 0, ..., k-1.$$

¿From the last equation we have

$$\psi_i = -\frac{\partial G_{i-1}^T}{\partial x_i}(\hat{x}_{i-1}, \hat{x}_i), \quad i = 1, ..., k. \tag{3.75}$$

Substituting (3.75) to (3.73) and (3.74) we obtain the Euler equation (3.68) and the condition (3.69).

Remark 3.15. The fact that the discrete variational problem leads to an optimal control problem without control constraints is essential for the problems of calculus of variations. On the other hand, the fact that there is no terminal state constraint is not important in this case. Indeed, even for the variational problems with the additional condition $g(x_k) = 0$, the Euler equation (3.68) holds, however, the condition (3.69) is replaced by

$$\frac{\partial G_{k-1}}{\partial x_k}(\hat{x}_{k-1}, \hat{x}_k) + \chi^T \frac{\partial g}{\partial x_k}(\hat{x}_k) = 0.$$

Remark 3.16. Above, we have shown that *every* discrete problem of calculus of variations can be converted to a problem of optimal control. Conversely, only in special cases it is possible to convert the optimal control problem *without control constraints* to a discrete problem of calculus of variations. For example, if from the equation

$$x_{i+1} = f_i(x_i, u_i)$$

it is possible to express a unique

$$u_i = h_i(x_i, x_{i+1}),$$

then the optimal control problem can be replaced by the problem of calculus of variations with the objective function

$$G_i(x_i, x_{i+1}) = f_i(x_i, h_i(x_i, x_{i+1})).$$

For this purpose it is usually necessary that dim $u_i =$ dim x_i and therefore this replacement is mainly performed in one dimensional problems.

Remark 3.17. The Euler equation for the problem (3.66)–(3.67) is sometimes derived from the dynamic programming equation, which, for this problem, is of the form

$$V_i(x_i) = \max_{x_{i+1}}[G_i(x_i, x_{i+1}) + V_{i+1}(x_{i+1})], \ i = 0, ..., k-1, \tag{3.76}$$

$$V_k(x_k) = 0, \quad \text{for all } x_k. \tag{3.77}$$

Necessary optimality conditions are applied to the maximization condition in (3.76) and then the envelope theorem is used in order to eliminate the gradients of the value function from the resulting expression (see Excercise 3.14).

3.5 Infinite Horizon Problems

As mentioned earlier in subsection 2.2.3, in applications we need to deal with problems with infinite number of stages, i.e. infinite horizon problems. In this case, the control is not a finite sentence and therefore the results of nonlinear programming cannot be directly applied. We show, however, that under certain restrictive assumptions we can derive conditions from the finite horizon ones which look very similar to those for the former.

We restrict ourselves to the problem without control constraints

$$\text{maximize} \quad \sum_{i=0}^{\infty} f_i^0(x_i, u_i) \tag{3.78}$$

$$\text{subject to} \quad x_{i+1} = f_i(x_i, u_i), \quad i = 0, 1, \dots, \tag{3.79}$$

$$x_0 = a, \tag{3.80}$$

$$\lim_{k\to\infty} \quad x_k \in C = \{x \ : \ g(x) = 0\}. \tag{3.81}$$

We assume that the functions f_i^0, f_i and g satisfy the same assumptions as in the problem (3.1)–(3.5). Since the objective function is the infinite series in (3.78), as mentioned in subsection 2.2.3, for the definition of admissible control we have to require that the series converges. Thanks to this requirement the qualities of admissible controls (measured by the objective function value) can be mutually compared. A necessary condition for convergence of the series is

$$\lim_{i\to\infty} f_i^0(\hat{x}_i, \hat{u}_i) = 0.$$

Theorem 3.4. *Let $\hat{\mathcal{U}} = \{\hat{u}_0, \hat{u}_1 \dots\}$ be an optimal control for the problem (3.78)-(3.81) and let $\hat{\mathcal{X}} = \{\hat{x}_0, \hat{x}_1 \dots\}$ be its response. Let the matrix $\frac{\partial f_i}{\partial x_i}(\hat{x}_i, \hat{u}_i)$ be regular for each i. Then there exists a $\psi^0 \geq 0$ and a solution $\Psi = \{\psi_1, \psi_2, \dots\}$ of the adjoint equation*

$$\psi_i = \frac{\partial f_i^{0T}}{\partial x_i}(\hat{x}_i, \hat{u}_i)\psi^0 + \frac{\partial f_i^T}{\partial x_i}(\hat{x}_i, \hat{u}_i)\psi_{i+1}, \ i = 1, 2..., \tag{3.82}$$

such that $(\psi^0, \psi_1) \neq 0$ *and for all* $i = 0, 1, \dots$ *one has*

$$\psi^0 \frac{\partial f_i^0}{\partial u_i}(\hat{x}_i, \hat{u}_i) + \psi_{i+1}^T \frac{\partial f_i}{\partial u_i}(\hat{x}_i, \hat{u}_i) = 0. \tag{3.83}$$

Remark 3.18. Note that the necessary conditions formulated in Theorem 3.4 do not contain any transversality condition. This means that the validity of Remark 3.5, according to which the necessary optimality conditions from Theorem 3.1 formally provide enough conditions for determination of optimal solutions, cannot be extended to the infinity horizon problem. In fact, if the terminal limit state of the response is not uniquely determined by the condition (3.81), then the missing conditions for it are not completed by the transversality condition. This deficiency is sometimes replaced by a requirement of convergence to the steady-state of the system of state and adjoint equations.

Remark 3.19. Since Theorem 3.4 does not contain the transversality condition, there is no vector χ. The condition $(\psi^0, \chi) \neq (0, 0)$ from Theorem 3.1 is replaced by the condition $(\psi^0, \psi_1) \neq (0, 0)$ in Theorem 3.4. Note that the condition (ψ^0, χ) from Theorem 3.1 is particulary important. Without it any admissible control would satisfy the conditions of the theorem with zero ψ^0 and zero solution of the adjoint equation. In this way, however, the necessary optimality conditions would be worthless. This role of condition $(\psi^0, \chi) \neq (0, 0)$ is played by $(\psi^0, \psi_1) \neq 0$ in Theorem 3.4. Due to the regularity assumption of the matrices $\frac{\partial f_i}{\partial x_i}(\hat{x}_i, \hat{u}_i)$, we can equivalently write the adjoint equation (3.16) in the form

$$\psi_{i+1} = \left[\frac{\partial f_i^T}{\partial x_i}(\hat{x}_i, \hat{u}_i) \right]^{-1} \left[-\frac{\partial f_i^{0T}}{\partial x_i}(\hat{x}_i, \hat{u}_i)\psi^0 + \psi_i \right], \; i = 1, 2, \dots \tag{3.84}$$

The equation (3.84) represents a recursive rule which uniquely determines its solution for given (ψ^0, ψ_1). In the case $\psi^0 = 0$, this solution is non-trivial.

In the proof of the theorem we use the following auxiliary result, which in fact is a small modification of the Bellman optimality principle (Theorem 2.2).

Lemma 3.1. *Let* $\hat{\mathcal{U}} = \{\hat{u}_0, \hat{u}_1 \ldots\}$ *be an optimal solution for the problem (3.78)-(3.81), let* $\hat{\mathcal{X}} = \{\hat{x}_0, \hat{x}_1 \ldots\}$ *be its response. Then for each* $k = 1, 2, \ldots,$ *the control* $\hat{\mathcal{U}}_{0,k-1} := \{\hat{u}_0, \hat{u}_1 \ldots, \hat{u}_{k-1}\}$ *is optimal for the problem*

$$\text{maximize} \quad \sum_{i=0}^{k-1} f_i^0(x_i, u_i) \tag{3.85}$$

$$\text{subject to} \quad x_{i+1} = f_i(x_i, u_i), \quad i = 0, , \ldots, k-1 \tag{3.86}$$

$$x_0 = a, \tag{3.87}$$

$$x_k = \hat{x}_k, \tag{3.88}$$

where f_i^0*,* f_i *and* a *are the data of the problem (3.78)-(3.81).*

Proof is very similar to the proof of Theorem 2.2. Therefore, we present its idea only and leave the details to the reader as an exercise.

If the claim of the lemma were not true, there would be other admissible control $\bar{\mathcal{U}}_{0,k-1}$ for the problem (3.85)-(3.88) with a higher value of the objective function as $\hat{\mathcal{U}}_{0,k-1}$. Now, the control, the first k elements of which were taken from the control $\bar{\mathcal{U}}_{0,k-1}$ and the other from the control $\hat{\mathcal{U}}$, would give a higher value to the objective function (3.78) than the control $\hat{\mathcal{U}}$. □

Proof of Theorem 3.4: Since by Lemma 3.1 each segment $\hat{\mathcal{U}}_{0,k}$ of the control $\hat{\mathcal{U}}$ is an optimal control for the finite horizon problem (3.85)–(3.88), the necessary optimality conditions from Theorem 3.1 are satisfied. This means that for each $k = 1, 2, \ldots$ there exists a number $\psi^{0(k)}$, a vector $\chi^{(k)}$ and a sequence $\Psi^{(k)} = \{\psi_1^{(k)}, \ldots, \psi_k^{(k)}\}$ such that $(\psi^{0(k)}, \chi^{(k)}) \neq 0$, $\psi^0 \geq 0$ and the conditions (3.18)–(3.21) are satisfied with $\lambda_i = 0$. As it is a problem with fixed terminal state, the matrix

$\frac{\partial g^T}{\partial x_k}(\hat{x}_k)$ is equal to the identity matrix, and hence $\psi_k^{(k)} = \chi^{(k)}$. Therefore

$$(\psi^{0(k)}, \psi_k^{(k)}) \neq 0. \tag{3.89}$$

Obviously, we also have $(\psi^{0(k)}, \psi_1^{(k)}) \neq 0$. Namely, if $(\psi^{0(k)}, \psi_1^{(k)}) = 0$, then from equation (3.84) by induction we obtain $\psi_i^{(k)} = 0$ for all $i = 1, ..., k$, which is in contradiction with (3.89).

By Remark 3.7 we can assume without loss of generality that $\|(\psi^{0(k)}, \psi_1^{(k)})\| = 1$ for each k. Since the sequence $\{(\psi^{0(k)}, \psi_1^{(k)})\}_{k=1}^{\infty}$ is bounded, we can choose a convergent subsequence from it. Denote a limit point of the subsequence as (ψ^0, ψ_1). Obviously $(\psi^0, \psi_1) \neq 0$. Applying a limit approach to the necessary optimality conditions, we obtain the theorem claim. □

The variational necessary optimality conditions for an autonomous infinite horizon problem do not exhibit remarkable specific features compared to the general problem. Worth noting, however, is the autonomous discounted problem. By transformation of adjoint variables we can change the corresponding non-autonomous adjoint equation and the variational condition to autonomous ones, which then allows an analysis of trajectories behavior through phase portraits. Recall that by the autonomous discounted problem we understand a problem, in which the functions $f_i \equiv f$ do not depend on i and $f_i^0 = \beta^i F$, where $0 < \beta < 1$.

In this case the adjoint equation is

$$\psi_i = \beta^i \frac{\partial F^T}{\partial x_i}(\hat{x}_i, \hat{u}_i)\psi^0 + \frac{\partial f^T}{\partial x_i}(\hat{x}_i, \hat{u}_i)\psi_{i+1}, \; i = 1, 2, \ldots \tag{3.90}$$

and the variational equation (3.83) is

$$\psi^0 \beta^i \frac{\partial F}{\partial u_i}(\hat{x}_i, \hat{u}_i) + \psi_{i+1}^T \frac{\partial f}{\partial u_i}(\hat{x}_i, \hat{u}_i) = 0, \quad i = 0, 1, \ldots. \tag{3.91}$$

In terms of transformed adjoint variables $\tilde{\psi}_i = \beta^{-i}\psi_i$, equations (3.90), (3.91) obtain the forms

$$\tilde{\psi}_i = \frac{\partial F^T}{\partial x_i}(\hat{x}_i, \hat{u}_i)\psi^0 + \beta \frac{\partial f^T}{\partial x_i}(\hat{x}_i, \hat{u}_i)\tilde{\psi}_{i+1}, \; i = 1, 2, \ldots \tag{3.92}$$

$$\psi^0 \frac{\partial F}{\partial u_i}(\hat{x}_i, \hat{u}_i) + \beta \tilde{\psi}_{i+1}^T \frac{\partial f}{\partial u_i}(\hat{x}_i, \hat{u}_i) = 0, \quad i = 0, 1, \dots . \tag{3.93}$$

respectively.

Remark 3.20. The necessary optimality conditions, derived in this subsection for the problem without control constrains, remain valid also for problems with constraints, provided the optimal control values are interior points of the sets U_i. Note that most applications leading to infinite horizon problems do satisfy this assumption.

3.6 Problem Solving

Example 3.2. Optimal mine extraction. The owner of a mine containing a tons of ore, has to decide how to optimally allocate extraction into particular years. He would like to extract the mine within k years. The annual revenues from selling u tons is $\sqrt{u}$ and the unit extraction cost is c. The problem is to maximize the total discounted profit from ore mining.
This problem can be formulated as an optimal control problem

$$\begin{aligned} \text{maximize} \quad & \sum_{i=0}^{k-1} \beta^i (\sqrt{u_i} - cu_i) \\ \text{subject to} \quad & x_{i+1} = x_i - u_i, \quad i = 0, \dots, k-1, \\ & x_0 = a, \\ & x_k = 0, \end{aligned}$$

where u_i denotes the amount of extracted ore in the i-th year and x_i denotes the amount of ore available for extraction at the beginning of the i-th year. Obviously $c, a > 0$ and the constraints $u_i \geq 0$ are implicitly contained in the definition domain of the objective function.

The Hamiltonian for this problem is of the form

$$H_i(x_i, u_i, \psi^0, \psi_{i+1}) = \psi^0 \beta^i (\sqrt{u_i} - cu_i) + \psi_{i+1}(x_i - u_i).$$

The maximum principle for this problem holds in the following form: Let $\hat{\mathcal{U}} = \{\hat{u}_0, \ldots, \hat{u}_{k-1}\}$ and $\hat{\mathcal{X}} = \{\hat{x}_0, \ldots, \hat{x}_k\}$ be optimal control and its response. Then, there exist $(\psi^0, \chi) \neq (0, 0)$, $\psi^0 \geq 0$ and $\{\psi_1, \ldots, \psi_k\}$ such that the maximum condition

$$\max_{u \geq 0}[\psi^0 \beta^i(\sqrt{u_i} - cu_i) + \psi_{i+1}(\hat{x}_i - u_i)] = \psi^0 \beta^i(\sqrt{\hat{u}_i} - c\hat{u}_i) + \psi_{i+1}(\hat{x}_i - \hat{u}_i),$$

the adjoint equation

$$\psi_i = \frac{\partial H_i}{\partial x_i}(\hat{x}_i, \hat{u}_i, \psi^0, \psi_{i+1}) = \psi_{i+1}$$

and the transversality condition

$$\psi_k = \chi$$

hold. Note that from the adjoint equation and the transversality condition it follows that $\psi_i = \chi$ for each i. Further, if $\psi^0 = 0$, then $\chi \neq 0$ and the maximum condition would provide either no solution (if $\chi < 0$), or $\hat{u}_i = 0$ for all i. However, such a control would not be admissible, since its response would not satisfy the terminal condition. Because of this we can set $\psi^0 = 1$. Due to a barrier effect of the function $\sqrt{u_i}$ at 0, it is now clear that if there exists a solution $\hat{u}$ of the maximum condition, then necessarily $\hat{u} > 0$ and

$$0 = \frac{\partial H_i}{\partial u_i}(\hat{x}_i, \hat{u}_i, 1, \psi_{i+1}) = \left(\frac{\beta^i}{2\sqrt{\hat{u}_i}} - c\right) - \psi_{i+1},$$

from which it follows

$$\frac{\beta^i}{2\sqrt{\hat{u}_i}} = c + \chi =: b.$$

This means that the value of extraction is such that in each year the discounted marginal profits are equal. From the last expression we can express the control

$$\hat{u}_i = \frac{\beta^{2i}}{4b^2}.$$

The constant b can be determined from the state equation and the terminal condition, where we obtain

$$\sum_{i=0}^{k-1} \frac{\beta^{2i}}{4b^2} = a,$$

and hence

$$b = \sqrt{\frac{4a(1-\beta^2)}{1-\beta^{2k}}}.$$

Remark 3.21. Note that the formulation of this problem leads to a special case of the problem from Example 1.4.

Example 3.3. Optimal consumption. In this example we again solve Example 1.2 about the optimal consumption

$$\text{maximize} \quad \sum_{i=0}^{k-1} \left(\frac{1}{1+\delta}\right)^i \ln u_i \tag{3.94}$$

$$\text{subject to} \quad x_{i+1} = (1+r)x_i - u_i, \; i = 0, \ldots, k-1, \tag{3.95}$$

$$x_0 = a, \tag{3.96}$$

$$x_k = b, \tag{3.97}$$

where a, b, r, δ are given nonnegative constants where in addition $r, a > 0$. This time the example will be solved using Theorem 3.3, the assumptions of linearity of f_i, convexity of U and concavity of f_i^0 being satisfied.

First notice that the problem does not have a solution for all combinations of data a and b. In fact, should we consume nothing, i.e. all the values of u_i were zero, we would achieve the value $\bar{x}_k = (1+r)^k a$ at the end of the proces. Since the form of the objective function implies that the only permitted control values are $u_i > 0$, it is clear that the value $\bar{x}_k$

represents an upper bound for those b, for which the admissible control exists. In the following we therefore solve the problem for $b < (1+r)^k a$.

Since there is no control constraint and the difference equation is linear, by Remark 3.8 we can set $\psi^0 = 1$. The adjoint equation is of the form

$$\psi_i = (1+r)\psi_{i+1}, \quad i = 1, \ldots, k-1 \tag{3.98}$$

and since the problem has fixed terminal state, the transversality condition

$$\psi_k = \chi$$

does not provide any useful information. The maximum condition is

$$\left(\frac{1}{1+\delta}\right)^i \ln u_i + \psi_{i+1}((r+1)x_i - u_i) \to \max_{u_i > 0}, \quad i = 0, \ldots, k-1,$$

which means that we are looking for a free extremum of a convex function. Necessary and sufficient optimality condition for the solution of the maximum condition is:

$$\left(\frac{1}{1+\delta}\right)^i \frac{1}{u_i} - \psi_{i+1} = 0,$$

from which we obtain

$$\psi_{i+1} = \left(\frac{1}{1+\delta}\right)^i \frac{1}{u_i}. \tag{3.99}$$

The adjoint equation yields

$$\psi_{i+1} = \left(\frac{1}{1+r}\right)^i \psi_1. \tag{3.100}$$

Substitution of (3.100) to (3.99) gives

$$\left(\frac{1}{1+r}\right)^i \psi_1 = \left(\frac{1}{1+\delta}\right)^i \frac{1}{u_i},$$

from which we obtain

$$u_i = \left(\frac{1+r}{1+\delta}\right)^i \frac{1}{\psi_1}, \tag{3.101}$$

where $\psi_1 > 0$, because $u_i > 0$.

Already at this stage of the problem solution we can derive an important qualitative information which shows how an increase/decrease of consumption over time depends on the ratio of constants δ and r. Namely, from the relation (3.101) it is seen that

if $r > \delta$, then u_i increases with i,
if $r < \delta$, then u_i decreases with i,
if $r = \delta$, then u_i is constant.

This result is apparent: if the interest rate is higher than the discount factor, it is worth to postpone consumption, otherwise not. To the complete identification of the control we now just calculate the value of ψ_1. For this purpose we use the initial and terminal condition for x. Substituting (3.101) in (3.95) we get

$$x_{i+1} = (1+r)x_i - \left(\frac{1+r}{1+\delta}\right)^i \frac{1}{\psi_1}. \tag{3.102}$$

The solution of the equation (3.102) with the initial condition (3.96) we obtain by the formula (4.10)

$$\begin{aligned} x_{i+1} &= (1+r)^{i+1}a - \sum_{s=0}^{i}(1+r)^{i-s}\left(\frac{1+r}{1+\delta}\right)^s \frac{1}{\psi_1} \\ &= (1+r)^{i+1}a - (1+r)^i \left(\sum_{s=0}^{i}\left(\frac{1}{1+\delta}\right)^s\right)\frac{1}{\psi_1}. \end{aligned}$$

If $\delta = 0$, then

$$x_{i+1} = (1+r)^{i+1}a - (i+1)(1+r)^i \frac{1}{\psi_1},$$

if $\delta > 0$, then

$$x_{i+1} = (1+r)^{i+1}a - (1+r)^i \frac{(1+\delta)^{i+1} - 1}{\delta(1+\delta)^i} \frac{1}{\psi_1}. \tag{3.103}$$

The value of $\frac{1}{\psi_1}$ can be easily determined by substituting (3.103) to the terminal condition (3.97). So, we have a unique candidate for the optimal control. The fact that this candidate is indeed optimal follows by Remark 3.10 from the linear-concave nature of this problem.

In the previous example we saw that the necessary optimality conditions allowed to derive some qualitative characteristics of optimal solutions even if we did not know the specific values of the constants entering the problem formulation. In the following example, which is a generalization of the problem from the previous example, we will see that these characteristics can also be derived in the case of very general production function and utility function.

Example 3.4. General optimal consumption problem. The problem is

$$\begin{aligned} \text{maximize} \quad & J = \sum_{i=0}^{k-1} \beta^i \, U(u_i) \\ \text{subject to} \quad & x_{i+1} = f(x_i) - u_i, \quad i = 0, \dots, k-1, \\ & x_0 = a, \\ & x_k = 0. \end{aligned}$$

We assume that the functions f and U are both continuously differentiable, increasing, concave, and satisfy

$$\lim_{x \to 0^+} f'(x) = \infty, \qquad \lim_{x \to \infty} f'(x) = 0,$$

$$\lim_{u \to 0^+} U'(u) = \infty, \qquad \lim_{u \to \infty} U'(u) = 0,$$

and $f(0) = 0$. With the help of the necessary conditions, similarly as in the previous example, we can derive the following relation

$$\frac{U'(u_i)}{U'(u_{i-1})} = \frac{1}{\beta f'(x_i)}. \tag{3.104}$$

This relation says that if the parameter β or the marginal production increase, the proportion of marginal utilities in two adjacent periods decreases. As a consequence, by concavity of U, we obtain that the proportion of optimal consumptions at stages i and $i-1$ increases. This result is in accord with economic theory: the parameter β expresses the weight assigned to the utility from future consumption and if for example the value of β increases, then we can expect an increase of the future consumption at expense of the present.

Example 3.5. Optimal investment. A product is produced by two industries A and B. Every unit invested into industry A (B) brings at the end of each subsequent year 2 (1) kg of product and 0.5 (0.9) unit of profit. For a given initial amount of units a^1 (a^2) invested to A (B) and some free unit amount a^3, we would like to determine how at the end of each year to divide the free units (generated by profit), to maximize the total production in five years.

Introduce the notations:

x_i^1 – amount of units invested to A up to the end of the i-th year,

x_i^2 – amount of units invested to B up to the end of the i-th year,

x_i^3 – free units available at the end of the i-th year;

u_i – the fraction of the free units invested to A in the i-th year.

We obtain the following three dimensional optimal control problem given by the difference equations

$$\begin{aligned} x^1_{i+1} &= x^1_i + u_i x^3_i, \quad i = 0, \ldots, 4 \\ x^2_{i+1} &= x^2_i + (1 - u_i)x^3_i, \\ x^3_{i+1} &= 0.5x^1_{i+1} + 0.9x^2_{i+1} \\ &= 0.5(x^1_i + u_i x^3_i) + 0.9(x^2_i + (1 - u_i)x^3_i) \\ &= x^3_i + 0.5x^1_i + 0.9x^2_i - 0.1x^4_i - 0.4u^3_i, \end{aligned}$$

the initial conditions

$$x^1_0 = a^1 \geq 0, \quad x^2_0 = a^2 \geq 0, \quad x^3_0 = a^3 \geq 0,$$

the target set $C = \mathbb{R}^3$, the control constraints

$$0 \leq u_i \leq 1, \quad i = 0, \ldots, 4,$$

and the objective function of the form

$$\max J := \sum_{i=1}^{5} (2x^1_i + x^2_i)$$

$$\begin{aligned} &= \sum_{i=0}^{4} \left(2(x^1_i + u_i x^3_i) + x^2_i + (1 - u_i)x^3_i\right) \\ &= \sum_{i=0}^{4} (2x^1_i + x^2_i + x^3_i + u_i x^3_i). \end{aligned}$$

For the given problem we obtain the adjoint equation in the form

$$\begin{aligned} \psi^1_i &= \psi^1_{i+1} + 2\psi^0 + 0.5\psi^3_{i+1}, \\ \psi^2_i &= \psi^2_{i+1} + \psi^0 + 0.9\psi^3_{i+1}, \\ \psi^3_i &= \psi^3_{i+1} + \psi^0(1 + u_i) + \psi^1_{i+1}u_i \\ &\quad + \psi^2_{i+1}(1 - u_i) + (-0.1 - 0.4u_i)\psi^3_{i+1}, \end{aligned}$$

for $i = 1, \ldots, 4$ and the transversality condition

$$\psi_5 = 0.$$

As in the previous example we can put $\psi^0 = 1$. Since the conditions of linearity and concavity are satisfied, we can use the maximum condition, which obtains the form:

$$\begin{aligned}&(1 + \psi^1_{i+1} - \psi^2_{i+1} - 0.4\psi^3_{i+1})\hat{x}^3_i \hat{u}_i \\ &= \max_{0 \le u_i \le 1}(1 + \psi^1_{i+1} - \psi^2_{i+1} - 0.4\psi^3_{i+1})\hat{x}^3_i u_i.\end{aligned}$$

¿From the state equation and the initial condition it follows that $\hat{x}^3_i > 0$ for each i, and hence the maximum condition yields

$$\hat{u}_i = \begin{cases} 1, & \text{if} \quad \psi^2_{i+1} + 0.4\psi^3_{i+1} - \psi^1_{i+1} < 1, \\ 0, & \text{if} \quad \psi^2_{i+1} + 0.4\psi^3_{i+1} - \psi^1_{i+1} > 1, \\ \text{unspecified}, & \text{if} \quad \psi^2_{i+1} + 0.4\psi^3_{i+1} - \psi^1_{i+1} = 1. \end{cases}$$

This solution together with the adjoint equation and the transversality conditions allow to compute ψ_i and u_i, for $i = 4, 3, 2, 1, 0$, and then from the state equation to compute x_i, $i = 1, \ldots, 5$.

Let $a^1 = 0, \quad a^2 = 0, \quad a^3 = 1$. Since $\psi^0 = 1$, then from the adjoint equation we have:

$$\begin{aligned}\psi^1_i &= \psi^1_{i+1} + 2 + 0.5\psi^3_{i+1}, \\ \psi^2_i &= \psi^2_{i+1} + 1 + 0.9\psi^3_{i+1}, \\ \psi^3_i &= \begin{cases} 0.5\psi^3_{i+1} + 2 + \psi^1_{i+1}, & \text{if} \quad \psi^2_{i+1} + 0.4\psi^3_{i+1} - \psi^1_{i+1} < 1, \\ 0.9\psi^3_{i+1} + 1 + \psi^2_{i+1}, & \text{if} \quad \psi^2_{i+1} + 0.4\psi^3_{i+1} - \psi^1_{i+1} > 1. \end{cases}\end{aligned}$$

The solution of this difference system backwards from $\psi_5 = 0$ is presented in Table 3.1.

The optimal response can be obtained solving the state equation recurrently from the initial condition $a = (0, 0, 1)^T$ at $\hat{u}_0 = \hat{u}_1 = 0$, $\hat{u}_2 = \hat{u}_3 = \hat{u}_4 = 1$ (Table 3.2).

Table 3.1: The solution to the adjoint equation and the maximum condition for Example 3.5

i	5	4	3	2	1
ψ_i^1	0	2	5	9.5	16.25
ψ_i^2	0	1	3.8	9.3	18.85
ψ_i^3	0	2	5	9.5	18.85
$\psi_i^2 + 0.4\psi_i^3 - \psi_i^1$	0	-0.2	0.8	3.6	10.14
u_{i-1}	1	1	1	0	0

Table 3.2: Solution of the response for Example 3.5

i	0	1	2	3	4	5
$\hat{x}_i^1$	0	0	0	1.71	4.275	8.1125
$\hat{x}_i^2$	0	1	1.9	1.9	1.9	1.9
$\hat{x}_i^3$	1	0.9	1.71	2.565	3.8475	5.77125

Again we receive the unique candidate for the optimal control and since an optimal control exists (see Remark 3.6), the received control is indeed optimal.

Example 3.6. The linear problem of optimal control. This is a problem in which the state equation, the constraints and the objective function are linear. Hence the problem is

$$\begin{aligned}
\text{maximize} \quad & J = \sum_{i=0}^{k-1} c_i^T x_i + d_i^T u_i \\
\text{subject to} \quad & x_{i+1} = A_i x_i + B_i u_i, \quad i = 0, \ldots, k-1, \\
& F_i u_i \leq w_i, \quad i = 0, \ldots, k-1, \\
& x_0 = a, \\
& G x_k = h,
\end{aligned}$$

where for each $i = 0, \ldots, k-1$ the matrices A_i and B_i are of the types $n \times n$ and $n \times m$, respectively, the matrices F_i, G and the vectors c_i, d_i w_i, h are of appropriate dimensions. Since everything is linear, it is possible to use Theorem 3.2, and we can set $\psi^0 = 1$. The system of adjoint equations does not depend on x_i and u_i:

$$\psi_i = A_i^T \psi_{i+1} + c_i, \quad i = 1, \ldots, k-1.$$

The transversality condition is

$$\psi_k = G^T \chi.$$

The Hamiltonian is of the form

$$H_i = c_i^T x + d_i^T u + \psi_{i+1}^T (A_i x_i + B_i u_i)$$

and the maximum condition reads:

$$d_i^T \hat{u}_i + \psi_{i+1} B_i \hat{u}_i = \max_{u_i \in U_i} (d_i^T u_i + \psi_{i+1}^T B_i u_i), \quad i = 0, \ldots, k-1. \quad (3.105)$$

Assume in particular

$$U_i = \{u_i \ : \ \alpha_i^j \leq u_i^j \leq \beta_i^j, \ j = 1, \ldots, m\},$$

which means

$$U_i = U_i^1 \times \cdots \times U_i^m, \text{ where } U_i^j = [\alpha_i^j, \beta_i^j]. \quad (3.106)$$

Then, the condition (3.105) falls apart into m conditions

$$(d_i^j + (\psi_{i+1}^T B_i)^j) u_i^j \to \max_{u_i^j \in U_i^j},$$

from which it follows

$$\hat{u}_i^j = \begin{cases} \alpha_i^j, & \text{if } d_i^j + \psi_{i+1}^T B_i^j < 0, \\ \beta_i^j, & \text{if } d_i^j + \psi_{i+1}^T B_i^j > 0, \\ \text{unspecified}, & \text{if } d_i^j + \psi_{i+1}^T B_i^j = 0, \end{cases}$$

where B_i^j is the j-th column of the matrix B_i, $j = 1, \ldots, m$, $i = 0, \ldots, k-1$.

So we get an interesting qualitative result: except of an exceptional case $d_i^j + \psi_{i+1}^T B_i^j = 0$, the control $\hat{u}_i$ takes only the extreme values.

Example 3.7. The Linear-quadratic regulator. In this problem the state equation and the control constraints are linear of the same form as in the previous example. This time we however assume free terminal state and we minimize the quadratic convex objective function:

$$\sum_{i=0}^{k-1} \frac{1}{2}(x_i^T Q_i x_i + u_i^T R_i u_i),$$

where $Q_i,\ R_i$ are symmetric matrices, $Q_i \geq 0, R_i > 0$. The adjoint equation is

$$\psi_i = A_i^T \psi_{i+1} + \psi^0 Q_i \hat{x}_i$$

and the maximum condition is

$$\frac{1}{2}\psi^0 \hat{u}_i^T R_i \hat{u}_i + \psi_{i+1}^T B_i \hat{u}_i = \max_{u_i \in U_i} \left[\frac{1}{2}\psi^0 u_i R_i^T u_i + \psi_{i+1}^T B_i u_i \right]. \quad (3.107)$$

Because of the free terminal state we can set $\psi^0 = -1$. If we choose particular U_i as in the previous example and $R_i = I$, then the condition (3.107) can be rewritten into m conditions

$$-\frac{1}{2}(u_i^j)^2 + \psi_{i+1}^T B_i^j u_i^j \to \max_{u_i^j \in [\alpha_i^j, \beta_i^j]}, \quad j = 1, \ldots, m$$

from which it follows

$$\hat{u}_i^j = \begin{cases} \alpha_i^j, & \text{if} \quad \alpha_i^j > \psi_{i+1}^T B_i^j \\ \psi_{i+1}^T B_i^j, & \text{if} \quad \alpha_i^j \le \psi_{i+1}^T B_i^j \le \beta_i^j \\ \beta_i^j, & \text{if} \quad \beta_i^j < \psi_{i+1}^T B_i^j \end{cases}.$$

We again obtain an interesting qualitative result: $\hat{u}_i^j$ is equal to value from the interval $[\alpha_i^j, \beta_i^j]$, nearest to $\psi_{i+1}^T B_i^j$.

Example 3.8. Optimal consumption as an infinite horizon problem. We again return to the problem from Example 3.3, but this time it will be treated as an infinite horizon problem. That is, we solve the problem

$$\text{maximize} \quad \sum_{i=0}^{\infty} \left(\frac{1}{1+\delta}\right)^i \ln u_i \tag{3.108}$$

$$\text{subject to} \quad x_{i+1} = (1+r)x_i - u_i, \ i = 0, 1, \ldots, \tag{3.109}$$

$$x_0 = a, \tag{3.110}$$

$$\lim_{i \to \infty} x_i = b, \tag{3.111}$$

here again a, b, r, δ are given nonnegative constants, where moreover $r > 0, \ a > 0$.

In our case the condition $\frac{\partial f}{\partial x} = 1 + r \neq 0$ holds, and so we can apply the necessary optimality conditions from Theorem 3.4. For particular indices i, these conditions are of the same form as the conditions for the problem from Example 3.3. This means that by solving these conditions we obtain the control given by the relation (3.101) for $i = 0, 1, \ldots,$ i.e.

$$u_i = \left(\frac{1+r}{1+\delta}\right)^i \frac{1}{\psi_1}, \tag{3.112}$$

where $\psi_1 > 0$. Is such a control admissible for the infinite horizon problem? More precisely, for which values of the parameters a, b, r, δ, and yet undetermined positive constant ψ_1, such control is admissible for the problem?

To answer this question we must first investigate convergence of the infinite series in the objective function for such control, i.e. convergence of the series

$$\sum_{i=0}^{\infty}\left(\frac{1}{1+\delta}\right)^i \ln\left[\left(\frac{1+r}{1+\delta}\right)^i \frac{1}{\psi_1}\right], \tag{3.113}$$

where $\psi_1 > 0$. Note that in the case $\delta = 0$ the series diverges (it does not satisfy the necessary convergence condition, since $\lim_{i\to\infty} \ln\left(\frac{(1+r)^i}{\psi_1}\right) = \infty$). This means that if $\delta = 0$, the problem does not have optimal solutions. In the case $\delta > 0$, the series (3.113) always converges, for example by the d'Alambert criterium. So, we further investigate only the case $\delta > 0$.

Our goal is to find (if it exists) $\psi_1 > 0$ such that the response on the control (3.112) with the initial condition (3.110) satisfies also the terminal condition (3.111). Since the response on the control (3.112) is given by the same relation as in the finite horizon case, we obtain from (3.103) that

$$x_i = (1+r)^i a - (1+r)^{i-1}\frac{(1+\delta)^i - 1}{\delta(1+\delta)^{i-1}}\frac{1}{\psi_1}. \tag{3.114}$$

We now try to determine ψ_1 so that the limit condition (3.111) holds. Denote $\psi_1^{(k)}$ the value of ψ_1, for which $x_k = b$ holds. From the relation (3.114) we obtain

$$\frac{1}{\psi_1^{(k)}} = \left[\frac{(1+r)^k a - b}{(1+\delta)^k - 1}\right]\delta\left(\frac{1+r}{1+\delta}\right)^{k-1}. \tag{3.115}$$

By the limit approach in (3.115) we get the following relation for ψ_1:

$$\frac{1}{\psi_1} = \lim_{k\to\infty}\frac{1}{\psi_1^{(k)}} = \delta a\left(\frac{1+r}{1+\delta}\right). \tag{3.116}$$

Substituting into (3.114), after adjustments we obtain

$$x_i = a\left(\frac{1+r}{1+\delta}\right)^i. \qquad (3.117)$$

Let us now analyze the possibilities of the limit behavior of x_i. We can see that if $r > \delta$, then x_i diverge, and hence the optimal control for the problem (3.108)–(3.111) does not exist no matter what the value of b is. In the case $r = \delta$, the response x_i is constant, and hence there exists the optimal control only if $a = b$. In the case $r < \delta$, the response x_i converges to zero, and hence the problem has the solution only if $b = 0$. Note that if we substitute (3.116) to (3.112) we obtain

$$u_i = a\delta\left(\frac{1+r}{1+\delta}\right)^{i+1} \qquad (3.118)$$

and if we compare the last relation to (3.117) we get

$$u_i = \delta\left(\frac{1+r}{1+\delta}\right)x_i. \qquad (3.119)$$

We find that the problem (3.108)–(3.111) has an optimal solution only if $b = 0$ and $r < \delta$, or if $b = a$ and $r = \delta$ and then the optimal u_i is a constant multiple of x_i. How to explain these facts?

We can approach the problem (3.108)–(3.111) from another side. Set the values ψ_i and ψ_{i+1} given by the formula (3.99) to the adjoint equation (3.98). By adjusting and shifting indices we get the difference equation for the control

$$u_{i+1} = \left(\frac{1+r}{1+\delta}\right)u_i. \quad i = 0, 1, \ldots.$$

This equation together with the state equation (3.109) form the following system of linear equations

$$\begin{pmatrix} x_{i+1} \\ u_{i+1} \end{pmatrix} = \begin{pmatrix} 1+r & -1 \\ 0 & \frac{1+r}{1+\delta} \end{pmatrix}\begin{pmatrix} x_i \\ u_i \end{pmatrix}. \qquad (3.120)$$

The matrix of the system (3.120) has eigenvalues

$$\lambda_1 = \frac{1+r}{1+\delta} \quad \text{and} \quad \lambda_2 = (1+r)$$

with eigenvectors

$$v_1 = \begin{pmatrix} 1 \\ \delta\frac{1+r}{1+\delta} \end{pmatrix} \quad \text{and} \quad v_2 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}.$$

The general solution of the equation (3.120) is of the form

$$\begin{pmatrix} x_i \\ u_i \end{pmatrix} = c_1 \begin{pmatrix} 1 \\ \delta\frac{1+r}{1+\delta} \end{pmatrix} \left(\frac{1+r}{1+\delta}\right)^i + c_2 \begin{pmatrix} 1 \\ 0 \end{pmatrix} (1+r)^i,$$

where c_1, c_2 are given by the initial point (x_0, u_0).

In the case $r \neq \delta$, the system (3.120) has a unique equilibrium, namely $\hat{x} = 0$ and $\hat{u} = 0$. We can see that if $r > \delta$, then both eigenvalues are greater than 1, and hence the equilibrium forms an unstable node (see Figure 3.1 in the middle). This explains the nonexistence of the optimal control for the problem with (3.111) in the case $r > \delta$. In fact, all the responses diverge to infinity.

If $r < \delta$, then $\lambda_1 < 1 < \lambda_2$, which means that the equilibrium is a saddle (see Figure 3.1 in the left). In order the trajectory for the point (x_0, u_0) to converge, one must have $c_2 = 0$. In other words, the point must lie on the *stable path*, which is a straight line given by the equation $u = \delta(1+r)/(1+\delta)x$. This explains the existence of optimal control only for $b = 0$.

In the case $r = \delta$, the problem exhibits an infinite number of steady states, all unstable. Each pair $\hat{x}, \hat{u}$ such that $\hat{u} = r\hat{x}$ is a steady state. Trajectories lie on straight lines parallel to the axis x (see Fig. 3.1 on the right). This means that the solutions we found above for $a = b$ correspond to steady states of the system. All solutions for other pairs of initial and terminal points diverge.

The following example represents a simplified deterministic version of so-called real business cycle (RBC) model. Such models allow to analyze national economic responses to technological and other shocks. The

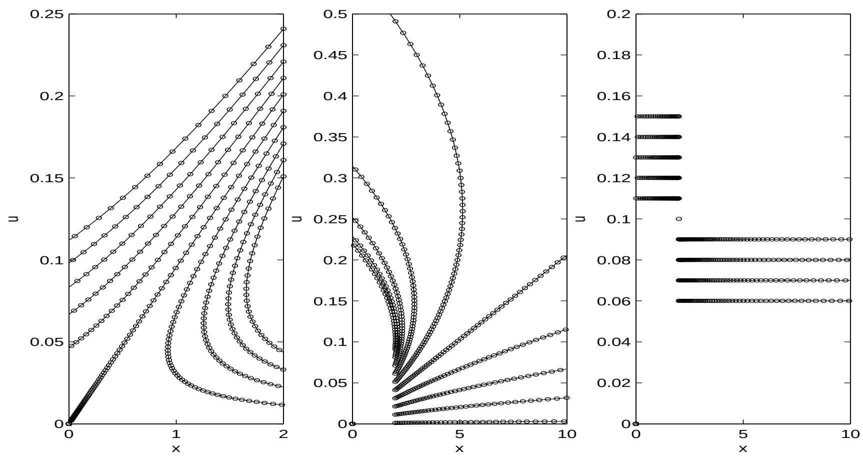

Figure 3.1: Trajectories for the response and control in Example 3.8 for $x_0 = 2$, for different initial values of u_0 and for $r < \delta$ (in the left), $r > \delta$ (in the middle) and $r = \delta$ (in the right).

RBC models are the starting point for the DGSE (dynamic stochastic general equilibrium) models that currently are a popular tool for economic analysis and forecasting.

Example 3.9. Optimal consumption and optimal amount of labor. We extend the problem of optimal consumption studied in Example 3.8 to a problem with two control variables. We assume that, in addition to consumption, an agent chooses the amount of labor, which is, together with capital, an input factor to the production function. In this case, however, unlike in Example 3.8, we cannot find the optimal solution analytically. Therefore, we are content with finding a stationary point and we analyze local dynamics of the optimal solution using the log-linearization of the differential equations in a neighborhood of this point.

Suppose that in an economy there is a number of identical, infinitely long living households. The households decide how much time they spend working (h_i), how much time they retain as free time (l_i) and what amount of the produced output they consume (c_i). We assume that $h_i + l_i = 1$. Doing so, the households maximize their discounted lifetime utility from consumption and free time, i.e.,

$$\max_{c_i, h_i} \sum_{i=0}^{\infty} \beta^i (\ln c_i + \ln(1 - h_i)), \quad \text{where } \beta \in (0, 1).$$

The amount of production is determined by the Cobb-Douglas production function, which depends on the amount of labor and the amount of capital (k_i). Output produced can be either consumed or used to increase the capital in the next period.

$$k_{i+1} = k_i + f(k_i, h_i) - c_i, \quad \text{where } f(k, h) = k^\alpha h^{1-\alpha} \text{ and } \alpha \in (0, 1). \tag{3.121}$$

The initial level $k_0 > 0$ of capital is given.

The Hamiltonian for the given problem is of the form

$$H(k_i, c_i, h_i, \psi_{i+1}) = \beta^i (\ln c_i + \ln(1 - h_i)) + \psi_{i+1}(k_i + k_i^\alpha h_i^{1-\alpha} - c_i).$$

We use Theorem 3.1 to formulate the necessary conditions. Accordingly, if (c_i, h_i) is an optimal control and k_i is its response, then there exists a sequence of ψ_i such that the variational conditions holds in the form

$$\frac{\beta^i}{c_i} - \psi_{i+1} = 0 \tag{3.122}$$

and

$$-\frac{\beta^i}{1 - h_i} + \psi_{i+1} \frac{\partial f}{\partial h}(k_i, h_i) = 0, \tag{3.123}$$

where ψ_i is the solution of the adjoint equation

$$\psi_i = \psi_{i+1} \left(\frac{\partial f}{\partial k}(k_i, h_i) + 1 \right). \tag{3.124}$$

¿From (3.122) and (3.124) we get

$$\frac{c_i}{c_{i-1}} = \beta \left(\frac{\partial f}{\partial k}(k_i, h_i) + 1 \right) \tag{3.125}$$

and from (3.122) and (3.123) we obtain

$$c_i = (1 - h_i) \frac{\partial f}{\partial h}(k_i, h_i). \tag{3.126}$$

¿From (3.121), (3.122) and (3.123) for the steady state $(\hat{k}, \hat{c}, \hat{h})$ we obtain the following relations

$$\begin{aligned} \hat{c} &= \hat{k}^{\alpha} \hat{h}^{1-\alpha}, \\ 1 &= \beta \left(\alpha \left(\frac{\hat{k}}{\hat{h}} \right)^{\alpha - 1} + 1 \right), \\ \hat{c} &= (1 - \hat{h})(1 - \alpha) \left(\frac{\hat{k}}{\hat{h}} \right)^{\alpha}. \end{aligned}$$

Their solution is

$$\hat{h} = A, \qquad \hat{k} = AB, \qquad \hat{c} = AB^{\alpha}, \tag{3.127}$$

where

$$A = \frac{1-\alpha}{2-\alpha} \qquad \text{and} \qquad B = \left(\frac{1-\beta}{\beta\alpha} \right)^{\frac{1}{\alpha-1}}.$$

The equations (3.121), (3.125) and (3.126) describing model dynamics are nonlinear and their solutions cannot be expressed in closed form. Therefore, we first log-linearize [2] the equation at the equilibrium point.

[2] Under the log-linearization of the function $f(x)$ in a neighborhood of the point $\hat{x}$ we understand the approximation $f(x) \approx f(\hat{x}) + f'(\hat{x})\hat{x}\bar{x}$ where $\bar{x} = \ln x - \ln \hat{x}$. This approximation is based on the approximation of the function f by the first order Taylor polynomial around the point $\hat{x}$ and the approximation of $\ln \frac{x}{\hat{x}} \approx \frac{x}{\hat{x}} - 1$, which results from the approximation of the function $\ln(x)$ by the first order Taylor polynomial in the neighborhood of the point 1.

Log-linearizing the equation (3.121) we obtain

$$\hat{k} + \hat{k}\bar{k}_{i+1} = \hat{k} + \hat{k}\bar{k}_i + f(\hat{k}, \hat{h}) + \frac{\partial f}{\partial k}(\hat{k}, \hat{h})\hat{k}\bar{k}_i + \frac{\partial f}{\partial h}(\hat{k}, \hat{h})\hat{h}\bar{h}_i - \hat{c} - \hat{c}\bar{c}_i,$$

where $\bar{k}_i = \ln k_i - \ln \hat{k}_i$, $\bar{c}_i = \ln c_i - \ln \hat{c}_i$ and $\bar{h}_i = \ln h_i - \ln \hat{h}_i$. Modifying and using (3.127) we obtain

$$\bar{k}_{i+1} = \bar{k}_i + B^{\alpha-1}(\alpha\bar{k}_i + (1-\alpha)\bar{h}_i - \bar{c}_i). \tag{3.128}$$

We modify the equation (3.125) by log-linearization to the form

$$\bar{c}_i - \bar{c}_{i-1} = (1-\beta)(1-\alpha)(\bar{h}_i - \bar{k}_i) \tag{3.129}$$

and finally we adapt the equation (3.126) into the form

$$c_i = \alpha(\bar{k}_i - \bar{h}_i) - \frac{A}{1-A}\bar{h}_i = \alpha(\bar{k}_i - \bar{h}_i) - (1-\alpha)\bar{h}_i.$$

¿From the last equation we can express $\bar{h}_i$ in the form

$$\bar{h}_i = \alpha\bar{k}_i - c_i.$$

After substituting this into (3.128) and (3.129) we obtain the system of linear equations

$$V_1 \begin{pmatrix} \bar{k}_{i+1} \\ \bar{c}_{i+1} \end{pmatrix} = V_2 \begin{pmatrix} \bar{k}_i \\ \bar{c}_i \end{pmatrix},$$

where V_1 and V_2 are 2×2 matrices of the form

$$V_1 = \begin{pmatrix} 1 & 0 \\ (1-\alpha)^2(1-\beta) & 1 + (1-\beta)(1-\alpha) \end{pmatrix},$$

$$V_2 = \begin{pmatrix} 1 + \frac{(1-\beta)(2-\alpha)}{\beta} & \frac{(1-\beta)(\alpha-2)}{\alpha\beta} \\ 0 & 1 \end{pmatrix}.$$

Since the matrix V_1 is triangular, its diagonal elements are its eigenvalues. Both eigenvalues are nonzero and hence the matrix V_1 is regular. The system can be therefore adjusted to the form

$$\begin{pmatrix} \bar{k}_{i+1} \\ \bar{c}_{i+1} \end{pmatrix} = V \begin{pmatrix} \bar{k}_i \\ \bar{c}_i \end{pmatrix},$$

where $V = V_1^{-1} V_2$. The solution to the system is

$$\begin{pmatrix} \bar{k}_i \\ \bar{c}_i \end{pmatrix} = d_1 v_1 \lambda_1^i + d_2 v_2 \lambda_2^i,$$

where λ_1 and λ_2 are the eigenvalues of the matrix V; v_1 and v_2 are the corresponding eigenvectors and d_1 and d_2 are constants determined in such a way that the solution starts at a given point $(\bar{k}_0, \bar{c}_0)$.

To illustrate the solution we simulate the evolution of deviations of the variables from the steady state in the case of an initial deviation of 10%. Note that due to the relation $\ln \frac{x}{\hat{x}} \approx \frac{x-\hat{x}}{\hat{x}}$ for $x \to \hat{x}$ the variables $\hat{k}_i$, $\hat{c}_i$ and $\hat{h}_i$ express approximately one percentage deviation rate from the steady state.

Parameters of the model are chosen as follows: $\alpha = 0.33$, $\beta = 0.98$. For these parameters values the eigenvectors of the matrix V are approximately $\lambda_1 \approx 1.049 > 1$ and $\lambda_2 \approx 0.972 < 1$. This means that the steady state is a saddle[3]. We obtain a stable saddle path for $d_1 = 0$. Simulations of the percentage deviations of variables k_i, c_i and h_i from the equilibrium are shown in Figure 3.2.

[3] A more detailed analysis shows that the steady state is a saddle for any choice of the parameters $\alpha \in (0, 1)$ and $\beta \in (0, 1)$.

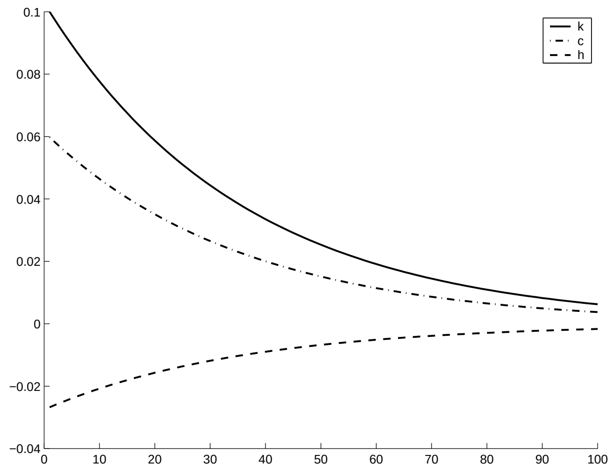

Figure 3.2: Simulations of the percentage deviations of variables k_i, c_i and h_i from the equilibrium for $\bar{k}_i = 0.1$.

3.7 Exercises

Exercise 3.1. Using Theorem 3.1 or Theorem 3.2 solve the problem

$$\begin{aligned}
&\max \sum_{i=0}^{k-1}(x_i + 2.5u_i)\\
&x_{i+1} = x_i - u_i, \quad i = 0, \dots, k-1,\\
&x_0 = a,\\
&u_i \in [-1, 1], \quad i = 0, \dots, k-1.
\end{aligned}$$

Exercise 3.2. Write down the necessary optimality conditions given by Theorem 3.2 or Theorem 3.1 for the following optimal control problem and use them to express the possible value of the control as a function of the state and adjoint variable:

$$
\begin{aligned}
&\min \ \frac{1}{2}\sum_{i=0}^{k-1}[(x_i^1)^2 + (u_i)^2] \\
&x_{i+1}^1 = x_i^1 + u_i, \quad x_0^1 = 1, \quad i = 0, \dots, k-1, \\
&x_{i+1}^2 = x_i^2 + 2x_i^1 + x_i^1 u_i, \quad x_0^2 = 0, \quad i = 0, \dots, k-1, \\
&-1 \le u_i \le 1, \\
&C = \{x_k : x_k^1 = 0\}.
\end{aligned}
$$

Exercise 3.3. By the help of Theorem 3.2 or Theorem 3.1 solve the optimal allocation resources problem formulated in Exercise 1.1. Compare with the solution of Example 2.2.

Exercise 3.4. By the help of the maximum principle solve the LQ problem from the Example 2.9 for $k = 5$.

Exercise 3.5. Prove the claim of Remark 3.9. In particular show that if $C = \{x : x^1 = a^1, \dots, x^l = a^l\}$, where $l \in \{1, \dots, n\}$, then $\psi_k^{l+1} = \dots = \psi_k^n = 0$ and ψ_k^i is free for $i = 1, \dots l$. Hint: Use the transversality condition and justify that the relations $\psi_k^i = \chi^i$ for $i = 1, \dots l$ can be interpreted as ψ_k^i free.

Exercise 3.6. (a) Show that if $C = \{(x_k^1, x_k^2) \in \mathbb{R}^2 | \quad x_k^1 = x_k^2\}$, then the transversality conditions can be written as $\psi_k^1 + \psi_k^2 = 0$.
(b) Find the transversality condition for $C = \{(x_k^1, x_k^2, x_k^3) \in \mathbb{R}^3 | \quad x_k^2 = x_k^3\}$.

Exercise 3.7. Justify in detail that if $\psi^0 = 1$, then Theorem 3.1 or Theorem 3.2 provides formally enough conditions to determine a control. This means that there is the same number of conditions as unknowns. Hint: the unknows are the values of the control, response, adjoint variables and multipliers.

Exercise 3.8. Adapt the formulation and the proof of Theorem 3.1 for the case of the problem (3.1)–(3.5) having the set of terminal states C given by both equalities and inequalities, i.e., $C = \{x : g(x) = 0, \quad q(x) \leq 0\}$.

Exercise 3.9. Adapt the formulation and the proof of Theorem 3.1 for the case, when the problem (3.1)–(3.5) has also state variables constrains $h_i(x_i, u_i) \leq 0$, or $h(x_i) \leq 0$, where $i = 0, \ldots, k-1$.

Exercise 3.10. Adapt the formulation and the proof of Theorem 3.1 for the case when the initial condition is given by an equation $h(x_0) = 0$. Hint: consider J as a function of the variables $(x_0, \mathcal{U})$ together with other equality constraint $h(x_0) = 0$.

Exercise 3.11. Prove the claim from Remark 3.6. To this aim use the Weierstrass theorem, according to which a continuous function on a compact set reaches its minimum.

Exercise 3.12. Derive formulae (3.104) from Example 3.4.

Exercise 3.13. Prove the claim from Remark 3.12.

Exercise 3.14. Derive the Euler equation for the problem (3.66)–(3.67) from the dynamic programming equation (3.76)–(3.77). Hint: Apply the necessary optimality condition to the maximum condition and then use the envelope theorem.

Exercise 3.15. Prove Lemma 3.1 in detail.

Exercise 3.16. Solve the problem from Example 3.2 on the infinite time horizon. Show that the optimal control satisfies the difference equation $u_{i+1} = \beta^2 u_i$ and analyze the phase portrait of the system of equations consisting of the control and the state equation.

Exercise 3.17. Prove that for any choice of parameters $\alpha \in (0, 1)$ and $\beta \in (0, 1)$ the steady state from Example 3.9 is a saddle. Hint: it suffices to prove that for any value of $\alpha \in (0, 1)$ and $\beta \in (0, 1)$ the relation $det(V_2 - V_1) = 0$ does not hold, which is equivalent to the fact that $\lambda = 1$ is not an eigenvalue of the matrix V.

Chapter 4

Appendix

4.1 Maximum Principle in Nonlinear Programming

Necessary optimality conditions for a nonlinear programming problem in the form of the Kuhn-Tucker theorem require that certain regularity assumptions (sometimes also called constraint qualification) are satisfied. Such assumptions are quite difficult to verify if they should be applied to optimal control problems. Therefore, we use a more general version of necessary conditions in the third chapter requiring weaker assumptions of regularity. This version of the necessary conditions used in Section 3 will be now derived from the John theorem.

4.1.1 The Standard Nonlinear Programming Problem

By a nonlinear programming problem (NP) we understand the problem

$$\begin{aligned} \text{maximize} \quad & f^0(x) \\ \text{subject to} \quad & f(x) = 0, \\ & s(x) \le 0, \end{aligned}$$

where $f^0 : \mathbb{R}^n \to \mathbb{R}$, $f : \mathbb{R}^n \to \mathbb{R}^\rho$ and $s : \mathbb{R}^n \to \mathbb{R}^\sigma$. We assume that $f^0,\ f,\ s \in \mathbb{C}^1$. If we denote

$$R := \{x : f(x) = 0\}, \quad S := \{x : s(x) \leq 0\},$$

then the problem (NP) can be rewritten as follows:

$$\max_{x \in R \cap S} f^0(x).$$

The function f^0 is called *objective function*, the constraint $f(x) = 0$ is called *equality constraint* and the constraint $s(x) \leq 0$ *inequality constraint*. As *active constraint at a point* $\hat{x} \in R \cap S$ we call a component s^i of the inequality constraint s for which $s^i(\hat{x}) = 0$ holds. We choose $\hat{x} \in R \cap S$ and define $I(\hat{x})$ as the set of the active constraints indices of s in $\hat{x}$, i.e.

$$I(\hat{x}) := \{i : s^i(\hat{x}) = 0\}.$$

We say that *the problem (NP) satisfies the regularity assumption (PR) at the point* $\hat{x} \in R \cap S$, if the vectors $\frac{\partial s^i}{\partial x}(\hat{x})$, $i \in I(\hat{x})$, are linearly independent. Note that this assumption applies only to the inequality constraints and, therefore, it does not represent the regularity assumption of the Kuhn-Tucker theorem for the problem (NP).

4.1.2 The Necessary Optimality Condition for Nonlinear Programming Problems

In this section the necessary optimality conditions for the problem (NP) are formulated in a form suitable for discrete optimal control theory. To this end the John theorem, which is a certain generalization of the Kuhn-Tucker theorem, will be used. The John theorem is not as well known as the fundamental Kuhn-Tucker one. This is why we present its formulation here in detail.

Theorem 4.1. The John theorem. *Let $\hat{x}$ be an optimal solution for the problem (NP). Then there exist multipliers $\psi^0 \geq 0$, $\psi \in \mathbb{R}^\rho$ and*

$\lambda \in \mathbb{R}^\sigma$, $(\psi^0, \psi, \lambda) \neq (0,0,0)$, such that both the variational condition

$$\psi^0 \frac{\partial f^0}{\partial x}(\hat{x}) + \psi^T \frac{\partial f}{\partial x}(\hat{x}) + \lambda^T \frac{\partial s}{\partial x}(\hat{x}) = 0 \tag{4.1}$$

and the complementarity condition

$$\lambda^T s(\hat{x}) = 0, \ \textit{where } \lambda \leq 0, \tag{4.2}$$

are satisfied.

The proof of this theorem can be found for example in the book [8].

With the problem (NP) we associate the *Lagrange function*

$$L(x, \psi^0, \psi, \lambda) = \psi^0 f^0(x) + \psi^T f(x) + \lambda^T s(x), \tag{4.3}$$

where $\psi^0 \in \mathbb{R}$, $\psi^0 \geq 0$, $\psi \in \mathbb{R}^\rho$ and $\lambda \in \mathbb{R}^\sigma$, $\lambda \leq 0$. Then the variational condition (4.1) can be written in the form

$$\frac{\partial L}{\partial x}(\hat{x}, \psi^0, \psi, \lambda) = 0. \tag{4.4}$$

Here ψ^0 is called *the objective function multiplier*, ψ *the equality constraint multiplier* and λ *the inequality constraint multiplier.*

Notice that the John theorem holds without any regularity assumption. However, while in the Kuhn-Tucker theorem, we have $\psi^0 = 1$, in the John theorem we only have $\psi^0 \geq 0$ plus a condition that not all multipliers ψ^0, ψ and λ simultaneously vanish. In the following theorem we formulate necessary optimality conditions under the regularity assumption (PR). With this assumption we obtain that merely the multipliers ψ^0 and ψ does not simultaneously vanish. Such a modification will be more suitable for our purposes.

Theorem 4.2. Corollary of the John theorem. *Let $\hat{x}$ be an optimal solution for the problem (NP) and let the regularity assumption (PR) be satisfied at the point $\hat{x}$. Then there exist multipliers $\psi^0 \geq 0$, $\psi \in \mathbb{R}^\rho$ and $\lambda \in \mathbb{R}^\sigma$, where $(\psi^0, \psi) \neq 0$ such that both the variational condition (4.1) and the complementarity condition (4.2) are satisfied.*

Proof: From the John theorem it follows that there exist multipliers $\psi^0 \geq 0$, $\psi \in \mathbb{R}^\rho$ and $\lambda \in \mathbb{R}^\sigma$, $(\psi^0, \psi, \lambda) \neq 0$ such that (4.1) and (4.2) hold. We have to prove $(\psi^0, \psi) \neq 0$.

Note that the complementarity condition implies $\lambda^i = 0$ for $i \notin I(\hat{x})$. The condition (4.1) hence yields

$$\sum_{i=0}^{\rho} \psi^i \frac{\partial f^i}{\partial x}(\hat{x}) + \sum_{i \in I(\hat{x})} \lambda^i \frac{\partial s^i}{\partial x}(\hat{x}) = 0. \tag{4.5}$$

If $(\psi^0, \psi) = 0$ were true, then (4.5) would yield

$$\sum_{i \in I(\hat{x})} \lambda^i \frac{\partial s^i}{\partial x}(\hat{x}) = 0. \tag{4.6}$$

Then from (4.6) and from the regularity assumption (PR) also $\lambda^i = 0$ for $i \in I(\hat{x})$ would follow. Hence, we would have not only $(\psi^0, \psi) = 0$, but also $\lambda = 0$, which would be a contradiction with $(\psi^0, \psi, \lambda) \neq 0$. Therefore $(\psi^0, \psi) \neq 0$ and the theorem is proved. □

Remark 4.1. The John theorem, from which we have derived Theorem 4.2, is usually formulated with opposite signs in ψ^0 and λ, i.e. with $\psi^0 \leq 0$ and $\lambda \geq 0$. This fact of course does not change the formulation of Theorem 4.2, because the formulations can be obtained from each other by multiplication of the equality (4.1) by -1. Important is that for the maximization problem and for the given orientation of the inequality in the formulation of the problem (NP), the signs of ψ^0 and λ must be different. We have chosen the positive sign for ψ^0, since this is standard in the optimal control problems.

Remark 4.2. The problem (NP) is formulated as a *maximization problem*. In the case of a *minimization problem*, the only thing that has to be changed in Theorem 4.2 is the sign of the multiplier ψ^0, i.e. $\psi^0 \leq 0$. That means, for the minimization problem the signs of all components of the vector λ are the same as the sign of the objective function multiplier.

Remark 4.3. If the claim of Theorem 4.1 or Theorem 4.2 is true with some (ψ^0, ψ, λ), then it is true also with $c(\psi^0, \psi, \lambda)$, for any $c > 0$. Hence if (ψ^0, ψ, λ) is such that $\psi^0 > 0$ (i.e. $\psi^0 \neq 0$), then, without loss of generality, we can take $\psi^0 = 1$.

Remark 4.4. Theorem 4.1 or Theorem 4.2 is of practical importance only when we can guarantee that $\psi^0 \neq 0$. Such a case occurs whenever the problem satisfies the regularity assumptions required by the Kuhn-Tucker theorem. Those are satisfied, for example, if s and f are linear or if the vectors $\frac{\partial s^i}{\partial x}(\hat{x})$, $i \in I(\hat{x})$, $\frac{\partial f^j}{\partial x}(\hat{x})$, $j = 1, \ldots, \rho$, are linearly independent. In other cases, the possibility of $\psi^0 = 0$ can be usually excluded by the help of the condition $(\psi^0, \psi) \neq 0$.

Remark 4.5. Both Theorem 4.1 and Theorem 4.2 are true even in the case when the problem (NP) has solely inequality or solely equality constraints. In that case we formally put $\psi = 0$ or $\lambda = 0$ in the claims of the theorems. When the problem does not include the equality constraints, then the regularity assumption (PR) represents the regularity assumption of the Kuhn-Tucker theorem, and hence we can put $\psi^0 = 1$.

Remark 4.6. In the case $\psi^0 = 1$, Theorem 4.1 and Theorem 4.2 provide formally enough conditions to determine $\hat{x}$.

4.1.3 Maximum Principle in Nonlinear Programming

Conditions (4.1) and (4.2) from Theorem 4.2 can be hard to verified. In a special case, as seen from the following theorem, it is possible to replace these conditions by another, stronger one that is easier to use. In the optimal control theory, it has a wide use. That special case is the problem (NP) satisfying following additional assumptions:

f is a linear function, i.e. $f(x) = Px + q$,

S is a convex set and

f^0 is a concave function.

We will refer to this problem as to (CP).

Theorem 4.3. *If for the problem (CP) there are* $(\psi^0, \psi) \neq 0$*,* $\psi^0 \geq 0$*,* $\lambda \leq 0$ *and* $\hat{x} \in S$ *such that the conditions (4.1) and (4.2) are satisfied, then also the following maximum condition is true:*

$$\psi^0 f^0(\hat{x}) + \psi^T f(\hat{x}) = \max_{x \in S}(\psi^0 f^0(x) + \psi^T f(x)). \tag{4.7}$$

Proof: Note that we maximize the concave function $\psi^0 f^0(x) + \psi^T f(x)$ on the convex set S in (4.7). For such a problem the Kuhn-Tucker conditions are sufficient for optimality. However, the conditions (4.1) and (4.2) from the assumption of the theorem are exactly those Kuhn-Tucker conditions for the problem in (4.7) at the point $\hat{x}$ and hence the theorem is proved. □

By combining Theorem 4.2 and 4.3 we obtain the following theorem.

Theorem 4.4. *Let* $\hat{x}$ *be an optimal solution to the problem (CP) and let the regularity assumption (PR) be satisfied at* $\hat{x}$*. Then, there exist* $(\psi^0, \psi) \neq 0$*,* $\psi^0 \geq 0$ *such that the maximum condition (4.7) holds.*

Remark 4.7. The condition (4.7) from Theorem 4.4 is preserved as maximum condition even if maximization in the problem (NP) is changed to minimization. Only what needs to be changed is the sign of ψ^0 (instead of $\psi^0 \geq 0$, we have $\psi^0 \leq 0$).

Remark 4.8. If in the problem (CP) both the equality and inequality constraints are linear, then the claims of Theorem 4.2 and Theorem 4.3 are true with $\psi^0 = 1$ (see Remark 4.5). Moreover, in this case the claims of these theorems can be reversed. That is, if $\hat{x} \in S \cap R$ satisfies (4.7) with $\psi^0 = 1$ and some ψ, then $\hat{x}$ is an optimal solution to the problem (CP).

4.2 The Envelope Theorem

The envelope theorem is a result often used in mathematics of economics. It claims that the derivative of the value function, as a function of a parameter of the optimization problem, is equal to the partial derivative of the objective function with respect to the parameter at the optimal value. We referred to the envelope theorem in the third chapter (see Remark 3.14 and Exercise 3.14) and the idea of the proof of the envelope theorem was used in the derivation of necessary optimality conditions from the dynamic programming equations in Subsection 3.4.2. Now, in this subsection is time to formulate and prove this theorem.

Theorem 4.5. *Let $f : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}$ be continuously differentiable and let the function*

$$V(x) := \max_u f(x, u)$$

be defined and continuously differentiable on an open subset $O \subset \mathbb{R}^n$. Let $\hat{x} \in O$ and let $\hat{u} \in arg \max_u f(\hat{x}, u)$. Then

$$\frac{\partial V}{\partial x}(\hat{x}) = \frac{\partial f}{\partial x}(\hat{x}, \hat{u}).$$

Proof: From the definition of V and $\hat{u}$ we obtain that

$$V(\hat{x}) - f(\hat{x}, \hat{u}) = 0$$

and for all $x \in O$ we have

$$V(x) - f(x, \hat{u}) \geq 0.$$

This means that the function $V(x) - f(x, \hat{u})$ attains its minimum at the point $x = \hat{x}$ on O, and hence at this point the necessary optimality conditions are satisfied. This means that

$$\frac{\partial V}{\partial x}(\hat{x}) - \frac{\partial f}{\partial x}(\hat{x}, \hat{u}) = 0,$$

from which the claim of the theorem follows. □

Remark 4.9. If there exists a continuously differentiable function $v : O \to \mathbb{R}^m$ such that for all $x \in O$ the relation

$$V(x) = f(x, v(x)) \tag{4.8}$$

is satisfied, then Theorem 4.5 can be proved by a differentiation of the relation (4.8) with respect to x. Indeed, from (4.8) we obtain

$$\frac{\partial V}{\partial x}(x) = \frac{\partial f}{\partial x}(x, v(x)) + \frac{\partial f}{\partial u}(x, v(x))\frac{\partial v}{\partial x}(x) \tag{4.9}$$

for all $x \in O$. Since

$$v(x) \in arg\max_{u} f(x, u), \quad x \in O,$$

we have

$$\frac{\partial f}{\partial u}(x, v(x)) = 0,$$

which, being substituted into (4.9), yields the claim of the theorem.

Remark 4.10. Note that in Subsection 3.4.2 we derived the adjoint equation from the dynamic programming equation using the idea of the first proof of Theorem 4.5. However, we could use an alternative procedure described in Remark 4.9. To this aim we must assume both $V \in \mathbb{C}^1$ and the existence of $v \in \mathbb{C}^1$.

4.3 Systems of Difference Equations

This section summarizes the most important knowledge related to certain types of difference equations and systems of such equations, necessary for the solution of examples and exercises in this book.

(a) One-dimensional linear difference equations. The solution of equations in the form

$$x_{i+1} = ax_i + b_i, \qquad i = 0, 1, 2, \dots,$$

where x_0 is given, is based on the variation of parameters formula

$$x_{i+1} = a^{i+1}x_0 + \sum_{s=0}^{i} a^{i-s}b_s. \tag{4.10}$$

This formula can be easily proved by mathematical induction.

(b) Two-dimensional linear autonomous systems. These are systems of difference equations of the form

$$x_{i+1} = Ax_i, \qquad i = 0, 1, 2, \dots, \tag{4.11}$$

where A is a 2×2 matrix. For our purposes, it is sufficient to deal with the case where A has real and mutually distinct eigenvalues $\lambda_1 \neq \lambda_2$. The corresponding eigenvectors are denoted by v_1 and v_2. It can be proved that any solution to the system (4.11) is given by the formula

$$x_i = c_1 v_1 \lambda_1^i + c_2 v_2 \lambda_2^i, \tag{4.12}$$

where the coefficients c_1 and c_2 are uniquely determined by the choice of the initial value x_0.

A *steady state* of the system (4.11) is an $\hat{x}$ such that $\hat{x} = A\hat{x}$. It is obvious that if $\lambda_1 \neq 0$ and $\lambda_2 \neq 0$, then $\hat{x} = 0$ is the only steady state.

Formula (4.12) implies that $x_i \to 0$ for $i \to \infty$ (or is bounded, respectively) if and only if all coefficients corresponding to eigenvalues with absolute values ≥ 1 (or > 1, respectively) vanish. This condition is trivially satisfied if there is no such eigenvalue. Hence x is called the *asymptotic stable* (or *stable*, resp.) steady state if for $i = 1, 2$ we have $|\lambda_i| < 1$ ($|\lambda_i| \leq 1$, resp.). Otherwise, this steady state is *unstable*.

Systems (4.11) can be classified according to the behaviour of their trajectories. The terminology is based on the analogy to the two-dimensional systems of differential equations [18].

(i) If $0 < |\lambda_1| < |\lambda_2| < 1$, the steady state 0 is called *stable node*. All solutions converge to this solution.

(ii) If $1 < |\lambda_1| < |\lambda_2|$, the steady state 0 is called *unstable node*. All solutions with $x_0 \neq 0$ diverge from this solution.

(iii) If $|\lambda_1| < 1 < |\lambda_2|$, then only the solutions with $c_2 = 0$, i.e. $x_0 = c_1 v_1$ converge to zero. All other solutions are unbounded. This case is called *saddle* and the line $x = cv_1,\ c \in (-\infty, \infty)$ is called *stable path*. It contains all converging trajectories.

(iv) If $\lambda_1 = 1$, then each point on the line $c_2 = 0$, i.e. $\hat{x} = cv_1,\ c \in (-\infty, \infty)$, is a steady state. This is stable if $|\lambda_2| < 1$ and unstable if $|\lambda_2| > 1$.

(c) Non-linear autonomous two-dimensional systems. The set of steady states of the system of non-linear difference equations

$$x_{i+1} = f(x_i), \qquad i = 0, 1, 2, \ldots, \tag{4.13}$$

where $f : \mathbb{R}^2 \to \mathbb{R}^2$ is a continuously differentiable function, is the set of *fixed points* of f, i.e. $\hat{x}$ such that $f(\hat{x}) = \hat{x}$. Stability, (local) asymptotic stability and instability of a steady state $\hat{x}$ are defined analogously to the linear case but they are considered in a sufficiently small neighbourhood of $\hat{x}$. When determining the asymptotical properties of the solutions, the eigenvalues of the matrix A are replaced by the eigenvalues of the Jacobian $\frac{\partial f}{\partial x}(\hat{x})$. In particular, $\hat{x}$ is locally asymptoticaly stable if the absolute values of both eigenvalues of matrix $\frac{\partial f}{\partial x}(\hat{x})$ are < 1. On the hand, it is unstable if at least one of these eigenvalues is > 1 in absolute value.

If the eigenvalues of the matrix $\frac{\partial f}{\partial x}(\hat{x})$ satisfy $|\lambda_1| < 1 < |\lambda_2|$, then $\hat{x}$ is a *saddle* in the sense that there is a unique *stable path* through $\hat{x}$. This path is a smooth curve which passes through $\hat{x}$ with tangential direction given by the eigenvector of the eigenvalue λ_1.

4.4 Elements of Probability Theory

Given that the theory of continuous random variables requires knowledge of measure theory that is beyond the scope of this book, this appendix only briefly summarizes some definitions and results of probability theory for discrete random variables. This knowledge is necessary to understand Subchapter 2.5 on stochastic dynamic programming.

Consider a non-empty countable set Ω called *set of elementary events* or *sample space*. Assume that $\mathcal{F}$ is a given *σ-algebra* on the set Ω (i.e. $\mathcal{F}$ is a collection of subsets of Ω, which includes the empty set, the complement of each set from $\mathcal{F}$ and the union of any countable collection of sets from $\mathcal{F}$). The set $\mathcal{F}$ is called a *set of events*, where an event is a collection of possible outcomes included in the sample space Ω. When defining discrete random variables, the set $\mathcal{F}$ is usually defined as the set of all subsets of the sample space Ω (denoted by 2^{Ω}), although this may not be the case in general.

On the set of events $\mathcal{F}$, we can define a function $P : \mathcal{F} \to \mathbb{R}$, which assigns to any event from $\mathcal{F}$ the probability of occurrence of this event. This function satisfies three axioms:

(i) $P(A) \geq 0$ for every $A \in \mathcal{F}$ (i.e. the probability of any event is non-negative)

(ii) $P(\Omega) = 1$ (i.e. probability that one of the elementary events included in Ω occurs is equal to 1),

(iii) $P(A \cup B) = P(A) + P(B)$ for arbitrary disjoint sets $A \in \mathcal{F}$ and $B \in \mathcal{F}$.

The triple $(\Omega, \mathcal{F}, P)$ is called a *(discrete) probability space*. On this space, we can define a *discrete random variable*. Discrete random variable is any real function $Z : \Omega \to \mathbb{R}$ satisfying

$$\{\omega \in \Omega \mid Z(\omega) \leq z\} \in \mathcal{F} \text{ for each } z \in \mathbb{R}.$$

In the special case mentioned above where the set $\mathcal{F}$ contains all subsets of the space Ω (i.e. $\mathcal{F} = 2^{\Omega}$), the discrete random variable is any function defined on the space Ω. In the following text, we will suppose that the assumption $\mathcal{F} = 2^{\Omega}$ is satisfied. Note that in this case, the discrete nature of the space Ω implies that for any random variable Z there exists a countable set of values $\{z^1, z^2, \dots\}$, where $z^i \in \mathbb{R}$ for all $i = 1, 2, \dots$ such that

$$P(Z = z^i) := P(\{\omega \in \Omega \mid Z(\omega) = z^i\}) > 0 \quad \text{and} \quad \sum_i P(Z = z^i) = 1.$$

The *expected value of a one-dimensional random variable* Z is denoted by EZ and it is defined as

$$EZ = \sum_i z^i P(Z = z^i)$$

assuming $\sum_i |z^i| P(Z = z^i) < \infty$ (i.e. this series converges absolutely which allows us to change the order of the summands). Analogously, for any function f we can define $Ef(Z)$ as

$$Ef(Z) = \sum_i f(z^i) P(Z = z^i)$$

assuming $\sum_i |f(z^i)| P(Z = z^i) < \infty$.

If $Z_1, Z_2, \ldots, Z_n : \Omega \to \mathbb{R}$ are random variables then the vector function $Z : \Omega \to \mathbb{R}^n$ defined by

$$Z(\omega) = \big(Z_1(\omega), Z_2(\omega), \ldots, Z_n(\omega)\big)^T$$

is called *n-dimensional random variable*. The *expected value of multidimensional random variable* $Z = (Z_1, \ldots, Z_n)^T$ is defined as

$$EZ = (EZ_1, \ldots, EZ_n)$$

assuming that all expected values EZ_i exist. The expected value $Ef(Z)$ is defined analogously to the one-dimensional case:

$$Ef(Z) = \sum_i f(z^i) P(Z = z^i),$$

assuming that random variable Z attains values from the countable set of vectors $\{z^1, z^2, \ldots\}$ with a non-zero probability.

Note that in general, $Ef(Z) \neq f(E(Z))$. However, any discrete random variables $Z_1, Z_2, \ldots$ and real numbers $a_1, a_2, \ldots$ satisfy

(i) $E(a_1 Z_1 + a_2) = a_1 EZ_1 + a_2$, (linearity),

(ii) $E(\sum_i a_i Z_i) = \sum_i a_i E Z_i$, (additivity).

If Z_1, Z_2 are discrete random variables attaining values $\{z_1^1, z_1^2, \dots\}$ and $\{z_2^1, z_2^2, \dots\}$, resp., then the functions $f(z_1^i, Z_2)$ for fixed i and $f(Z_1, z_2^j)$ for fixed j are random variables as well. Expected values of these random variables can be calculated as

$$E_{Z_2} f(z_1^i, Z_2) = \sum_j f(z_1^i, z_2^j)\, P(Z_2 = z_2^j),$$

$$E_{Z_1} f(Z_1, z_2^j) = \sum_i f(z_1^i, z_2^j)\, P(Z_1 = z_1^i).$$

Note that while $Ef(Z_1, Z_2)$ is a deterministic number (expected value of the random variable $f(Z_1, Z_2)$), $E_{Z_2} f(z_1^i, Z_2)$ and $E_{Z_1} f(Z_1, z_2^j)$ are functions of the random variables Z_1 and Z_2, respectively. Hence, they can be understood as random variables as well. Analogously, these considerations can be generalized to any countable set of random variables.

Random variables $Z_1, Z_2, \dots$ (of any dimension) are called *independent*, if

$$P(Z_1 = z_1, Z_2 = z_2, \dots) = P(Z_1 = z_1) P(Z_2 = z_2) \dots$$

for any $z_1, z_2, \dots$. If Z_1, Z_2 are independent (one-dimensional or multidimensional) random variables attaining values $\{z_1^i\}_{i \in I}$ and $\{z_2^j\}_{j \in J}$, resp., where index sets I and J are countable, we can formulate so-called *law of iterated expectations* as follows:

$$Ef(Z_1, Z_2) = E_{Z_1} E_{Z_2} f(Z_1, Z_2).$$

Proof:

$$\begin{aligned}
E_{Z_1}E_{Z_2}f(Z_1,Z_2) &= E_{Z_1}\sum_{j\in J} f(Z_1,z_2^j)P(Z_2=z_2^j)\\
&= \sum_{i\in I}\left(\sum_{j\in J} f(z_1^i,z_2^j)P(Z_2=z_2^j)\right)P(Z_1=z_1^i)\\
&= \sum_{i\in I,\, j\in J} f(z_1^i,z_2^j)P(Z_2=z_2^j)P(Z_1=z_1^i)\\
&= \sum_{i\in I,\, j\in J} f(z_1^i,z_2^j)P(Z_1=z_1^i, Z_2=z_2^j)\\
&= Ef(Z_1,Z_2),
\end{aligned}$$

where we have used the independence of random variables Z_1, Z_2 in the fourth equality.

The consequence of the above mentioned additivity of expected value is that

$$E(f(Z_1)+g(Z_1,Z_2)) = E_{Z_1}(f(Z_1)+E_{Z_2}g(Z_1,Z_2)).$$

This is an important equality for deriving the dynamic programming equation for stochastic problems.

Bibliography

[1] R. Bellman: *Dynamic Programming*, Princeton University Press, Princeton, N.J., 1957

[2] R. Bellman, S. Dreyfus: *Applied Dynamic Programming*, Princeton University Press, Princeton, N.J., 1962

[3] D. P. Bertsekas: *Dynamic Programming and Optimal Control*, Athena Scientific, Belmont, Massachusetts, 2005

[4] D. P. Bertsekas, S. E. Shreve: *Stochastic Optimal Control: The Discrete-Time Case*, Athena Scientific, Belmont, Massachusetts, 1996

[5] P. Brunovský: *Matematická teória optimálneho riadenia*, ALFA, Bratislava, 1980

[6] G. Feichtinger, R. F. Hartl: *Optimale Kontrolle ökonomische Prozesse*, W.Gruyter, Berlin, New York, 1986

[7] H.J. Greenberg: *Myths and Counterexamples in Mathematical Programming*, webpage INFORMS Computing Society Mathematical Programming Glossary, http://glossary.computing.society.informs.org/myths/, 2009

[8] M. Hamala: *Nelineárne programovanie*, ALFA, Bratislava, 1972

[9] A. C. Chiang: *Elements of Dynamic Optimization*, McGraw-Hill, Inc. New York, 1996

[10] G. C. Chow: *Analysis and Control of Dynamic Economic Systems*, John Wiley&Sons, New York, London, Sydney, Toronto, 1975 1989.

[11] K. L. Judd: *Numerical Methods in Economics*, The MIT Press Cambridge, Massachusetts, London, 1998

[12] M. I. Kamien, N. L. Schwartz: *Dynamic Optimization. The Calculus of Variations and Optimal Control in Economics and Management*, Elsevier, 1995

[13] A. Kaufmann, R. Cruon: *Dynamické programovanie*, ALFA, Bratislava, 1969

[14] V. Kučera: *The discrete Riccati equation of optimal control*, Kybernetika 8, 1972, 430-447

[15] D. Leonard, N. Long: *Optimal control theory and static optimization in economics*, Cambridge University Press, 1992

[16] L. Ljungquist, T. J. Sargent: *Recursive Macroeconomic Theory*, The MIT Press, Cambridge, 2000

[17] P. Mandl: Základy optimalizace vícestupňových soustav, *Kybernetika 8*, 1972

[18] M. Medveď: *Dynamické systémy*, Veda, Bratislava, 1988

[19] N. L. Stokey, R. E. Lucas, Jr., E Prescott: *Recursive Methods in Economic Dynamics*, Hardvard University Press, Cambridge, Massachusetts, and London, England, 1989

[20] S. P. Sethi, G. L. Thompson: *Optimal Control Theory. Applications to Management Science and Economics*, Kluwer Academic Publishers, Boston, Dordrecht, London, 2000

[21] M. Vlach: *Optimální řízení regulovaných systémů*, STTL, Praha, 1975

[22] D. A. Wisner, R. Chattergy: *Introduction to Nonlinear Optimization. A Problem Solving Approach.* Elsevier North Holland, New York, Amsterdam, Oxford, 1978

Symbols and Notation

$\mathbb{R}$	set of real numbers
$\mathbb{Z}$	set of integers
$\mathbb{N}$	set of natural numbers
$\mathbb{C}^r$	set of r times continuously differentiable functions
$\mathbb{R}^n$	n-dimensional Euclidean space
sgn	signum
tr	trace of a matrix
a^+	positive part of number a; $a^+ = \max(a, 0)$
a^-	negative part of number a; $a^- = \max(-a, 0)$
u_i	control variable in the stage i
$\mathcal{U}$	control on $[0, k]$; $\mathcal{U} = \{u_0, \dots, u_{k-1}\}$
$\mathcal{U}_j$	control on $[j, k]$; $\mathcal{U} = \{u_j, \dots, u_{k-1}\}$
U_i	set of admissible values of the control variable in the i-th stage
x_i	state variable in the stage i
$\mathcal{X}$	response; $\mathcal{X} = \{x_0, \dots, x_{k-1}\}$
X_i	set of admissible values of the state variable in the i-th stage
v_i	optimal feedback control in the stage i
V_i	value function in the stage i
$D_j(x)$	the problem of optimal transition from the state x to the terminal state on $[j, k]$

$\mathcal{P}_j(x)$	set of admissible controls for problem $D_j(x)$
$\Gamma_j(x)$	set of $u \in U_j$ for which there exists $\mathcal{U}_j \in \mathcal{P}_j(x)$ such that $u_j = u$
$\hat{\Gamma}_j(x)$	set of $u \in U_j$ such that there exists an optimal control $\mathcal{U}_j$ for the problem $D_j(x)$ satisfying $u_j = u$
E	expected value
z_i	random variable in the stage i
Z_i	admissible values of random variable in the stage i
$\mathcal{Z}$	sequence of realizations of random variable on $[0,k]$; $\mathcal{Z} = \{z_0, \dots, z_{k-1}\}$
$\mathcal{Z}_j$	sequence of realizations of random variable on $[j,k]$; $\mathcal{Z}_j = \{z_j, \dots, z_{k-1}\}$
$\mathcal{V}$	policy on $[0,k]$; $\mathcal{V} = \{v_0, \dots, v_{k-1}\}$
$\mathcal{V}_j$	policy on $[j,k]$; $\mathcal{V}_j = \{v_j, \dots, v_{k-1}\}$
$\nabla_u g$	gradient of function g with respect to (w.r.t.) u, i.e. column vector of first partial derivatives g w.r.t. u
$\nabla_u^2 g$	Hessian of function g w.r.t. u, i.e. matrix of second partial derivatives g w.r.t. u
$R > 0$	symmetric matrix R is positively definite
$Q \geq 0$	symmetric matrix Q is positively semi-definite
$x \in \mathbb{R}^n$	n-dimensional column vector
x^T	row vector which is a transpose of x
A^T	matrix which is a transpose of A
x^i	i-th component of vector x, i.e. if $x \in \mathbb{R}^n$, then $x = (x^1, \dots, x^n)^T$,
$\frac{\partial f}{\partial x}$	matrix of partial derivatives of function f, i.e. if $f : \mathbb{R}^n \to \mathbb{R}^m$, then $\frac{\partial f}{\partial x} = \begin{pmatrix} \frac{\partial f^1}{\partial x^1} & \cdots & \frac{\partial f^1}{\partial x^n} \\ \vdots & & \vdots \\ \frac{\partial f^m}{\partial x^1} & \cdots & \frac{\partial f^m}{\partial x^n} \end{pmatrix}$.
$\frac{\partial f^T}{\partial x}$	transposed matrix of partial derivatives of function f, i.e. $\frac{\partial f^T}{\partial x} = \left(\frac{\partial f}{\partial x}\right)^T$
f'	derivative of function f

Index

Authors:

Doc. RNDr. Margaréta Halická, CSc.,
Faculty of Mathematics, Physics and Informatics,
Comenius University in Bratislava,
halicka@fmph.uniba.sk

Prof. RNDr. Pavel Brunovský, DrSc.,
Faculty of Mathematics, Physics and Informatics,
Comenius University in Bratislava,
brunovsky@fmph.uniba.sk

Mgr. Ing. Pavol Jurča, PhD
National Bank of Slovakia,
pavol.jurca@nbs.sk