Chapter 1

What Is Optimal Control Theory?

Many management science applications involve the control of dynamic systems, i.e., systems that evolve over time. They are called *continuous-time systems* or *discrete-time systems* depending on whether time varies continuously or discretely. We will deal with both kinds of systems in this book, although the main emphasis will be on continuous-time systems.

Optimal control theory is a branch of mathematics developed to find optimal ways to control a dynamic system. The purpose of this book is to give an elementary introduction to the mathematical theory, and then apply it to a wide variety of different situations arising in management science. We have deliberately kept the level of mathematics as simple as possible in order to make the book accessible to a large audience. The only mathematical requirements for this book are elementary calculus, including partial differentiation, some knowledge of vectors and matrices, and elementary ordinary and partial differential equations. The last topic is briefly covered in Appendix A. Chapter 12 on stochastic optimal control also requires some concepts in stochastic calculus, which are introduced at the beginning of that chapter.

The principle management science applications discussed in this book come from the following areas: finance, economics, production and inventory, marketing, maintenance and replacement, and the consumption of natural resources. In each major area we have formulated one or more simple models followed by a more complicated model. The reader may

S. P. Sethi, *Optimal Control Theory*,
https://doi.org/10.1007/978-3-319-98237-3_1

wish at first to cover only the simpler models in each area to get an idea of what could be accomplished with optimal control theory. Later, the reader may wish to go into more depth in one or more of the applied areas.

Examples are worked out in most of the chapters to facilitate the exposition. At the end of each chapter, we have listed exercises that the reader should solve for deeper understanding of the material presented in the chapter. Hints are supplied with some of the exercises. Answers to selected exercises are given in Appendix E.

1.1 Basic Concepts and Definitions

We will use the word *system* as a primitive term in this book. The only property that we require of a system is that it is capable of existing in various *states*. Let the (real) variable $x(t)$ be the *state variable* of the system at time $t \in [0, T]$, where $T > 0$ is a specified time horizon for the system under consideration. For example, $x(t)$ could measure the inventory level at time t, the amount of advertising goodwill at time t, or the amount of unconsumed wealth or natural resources at time t.

We assume that there is a way of controlling the state of the system. Let the (real) variable $u(t)$ be the *control variable* of the system at time t. For example, $u(t)$ could be the production rate at time t, the advertising rate at time t, etc.

Given the values of the state variable $x(t)$ and the control variable $u(t)$ at time t, the *state equation*, a differential equation,

$$\dot{x}(t) = f(x(t), u(t), t), \quad x(0) = x_0, \tag{1.1}$$

specifies the instantaneous rate of change in the state variable, where $\dot{x}(t)$ is a commonly used notation for $dx(t)/dt$, f is a given function of x, u, and t, and x_0 is the initial value of the state variable. If we know the initial value x_0 and the *control trajectory*, i.e., the values of $u(t)$ over the whole time interval $0 \leq t \leq T$, then we can integrate (1.1) to get the *state trajectory*, i.e., the values of $x(t)$ over the same time interval. We want to choose the control trajectory so that the state and control trajectories maximize the *objective functional*, or simply the *objective function*,

$$J = \int_0^T F(x(t), u(t), t)dt + S[x(T), T]. \tag{1.2}$$

In (1.2), F is a given function of x, u, and t, which could measure the benefit minus the cost of advertising, the utility of consumption, the negative of the cost of inventory and production, etc. Also in (1.2), the function S gives the *salvage value* of the ending state $x(T)$ at time T. The salvage value is needed so that the solution will make "good sense" at the end of the horizon.

Usually the control variable $u(t)$ will be constrained. We indicate this as

$$u(t) \in \Omega(t), \quad t \in [0, T], \tag{1.3}$$

where $\Omega(t)$ is the set of feasible values for the control variable at time t.

Optimal control problems involving (1.1), (1.2), and (1.3) will be treated in Chap. 2.

In Chap. 3, we will replace (1.3) by inequality constraints involving control variables. In addition, we will allow these constraints to depend on state variables. These are called mixed inequality constraints and written as

$$g(x(t), u(t), t) \geq 0, \quad t \in [0, T]\,, \tag{1.4}$$

where g is a given function of u, t, and possibly x.

In addition, there may be constraints involving only state variables, but not control variables. These are written as

$$h(x(t), t) \geq 0, \quad t \in [0, T], \tag{1.5}$$

where h is a given function of x and t. Such constraints are the most difficult to deal with, and are known as pure state inequality constraints. Problems involving (1.1), (1.2), (1.4), and (1.5) will be treated in Chap. 4.

Finally, we note that all of the imposed constraints limit the values that the terminal state $x(T)$ may take. We denote this by saying

$$x(T) \in X, \tag{1.6}$$

where X is called the *reachable set* of the state variable at time T. Note that X depends on the initial value x_0. Here X is the set of possible terminal values that can be reached when $x(t)$ and $u(t)$ obey imposed constraints.

Although the above description of the control problem may seem abstract, you will find that in each specific application, the variables and parameters will have specific meanings that make them easy to understand and remember. The examples that follow will illustrate this point.

1.2 Formulation of Simple Control Models

We now formulate three simple models chosen from the areas of production, advertising, and economics. Our only objective here is to identify and interpret in these models each of the variables and functions described in the previous section. The solutions for each of these models will be given in detail in later chapters.

Example 1.1 *A Production-Inventory Model.* The various quantities that define this model are summarized in Table 1.1 for easy comparison with the other models that follow.

Table 1.1: The production-inventory model of Example 1.1

State variable	$I(t)$ = Inventory level
Control variable	$P(t)$ = Production rate
State equation	$\dot{I}(t) = P(t) - S(t),\ I(0) = I_0$
Objective function	Maximize $\left\{ J = \int_0^T -[h(I(t)) + c(P(t))]dt \right\}$
State constraint	$I(t) \geq 0$
Control constraints	$0 \leq P_{\min} \leq P(t) \leq P_{\max}$
Terminal condition	$I(T) \geq I_{\min}$
Exogenous functions	$S(t)$ = Demand rate
	$h(I)$ = Inventory holding cost
	$c(P)$ = Production cost
Parameters	T = Terminal time
	$I_{\min}$ = Minimum ending inventory
	$P_{\min}$ = Minimum possible production rate
	$P_{\max}$ = Maximum possible production rate
	I_0 = Initial inventory level

We consider the production and inventory storage of a given good, such as steel, in order to meet an exogenous demand. The state variable $I(t)$ measures the number of tons of steel that we have on hand at time $t \in [0, T]$. There is an exogenous demand rate $S(t)$ tons of steel per day at time $t \in [0, T]$, and we must choose the production rate $P(t)$ tons of steel per day at time $t \in [0, T]$. Given the initial inventory of I_0 tons of steel on hand at $t = 0$, the state equation

$$\dot{I}(t) = P(t) - S(t)$$

describes how the steel inventory changes over time. Since $h(I)$ is the cost of holding inventory I in dollars per day, and $c(P)$ is the cost of producing steel at rate P, also in dollars per day, the objective function is to maximize the negative of the sum of the total holding and production costs over the period of T days. Of course, maximizing the negative sum is the same as minimizing the sum of holding and production costs. The state variable constraint, $I(t) \geq 0$, is imposed so that the demand is satisfied for all t. In other words, *backlogging* of demand is not permitted. (An alternative formulation is to make $h(I)$ become very large when I becomes negative, i.e., to impose a *stockout* penalty cost.) The control constraints keep the production rate $P(t)$ between a specified lower bound $P_{\min}$ and a specified upper bound $P_{\max}$. Finally, the terminal constraint $I(T) \geq I_{\min}$ is imposed so that the terminal inventory is at least $I_{\min}$.

The statement of the problem is lengthy because of the number of variables, functions, and parameters which are involved. However, with the production and inventory interpretations as given, it is not difficult to see the reasons for each condition. In Chap. 6, various versions of this model will be solved in detail. In Sect. 12.2, we will deal with a stochastic version of this model.

Example 1.2 *An Advertising Model.* The various quantities that define this model are summarized in Table 1.2.

We consider a special case of the Nerlove-Arrow advertising model which will be discussed in detail in Chap. 7. The problem is to determine the rate at which to advertise a product at each time t. Here the state variable is *advertising goodwill*, $G(t)$, which measures how well the product is known at time t. We assume that there is a *forgetting coefficient* δ, which measures the rate at which customers tend to forget the product.

To counteract forgetting, advertising is carried out at a rate measured by the control variable $u(t)$. Hence, the state equation is

$$\dot{G}(t) = u(t) - \delta G(t),$$

with $G(0) = G_0 > 0$ specifying the initial goodwill for the product.

Table 1.2: The advertising model of Example 1.2

State variable	$G(t)$ = Advertising goodwill
Control variable	$u(t)$ = Advertising rate
State equation	$\dot{G}(t) = u(t) - \delta G(t), G(0) = G_0$
Objective function	Maximize $\left\{ J = \int_0^\infty [\pi(G(t)) - u(t)]e^{-\rho t} dt \right\}$
State constraint	$\cdots$
Control constraints	$0 \leq u(t) \leq Q$
Terminal condition	$\cdots$
Exogenous function	$\pi(G)$ = Gross profit rate
Parameters	δ = Goodwill decay constant
	ρ = Discount rate
	Q = Upper bound on advertising rate
	G_0 = Initial goodwill level

The objective function J requires special discussion. Note that the integral defining J is from time $t = 0$ to time $t = \infty$; we will later call a problem having an upper time limit of ∞, an *infinite horizon problem*. Because of this upper limit, the integrand of the objective function includes the discount factor $e^{-\rho t}$, where $\rho > 0$ is the (constant) discount rate. Without this discount factor, the integral would (in most cases) diverge to infinity. Hence, we will see that such a discount factor is an essential part of infinite horizon models. The rest of the integrand in the objective function consists of the gross profit rate $\pi(G(t))$, which results from the goodwill level $G(t)$ at time t less the cost of advertising assumed to be proportional to $u(t)$ (proportionality factor = 1); thus $\pi(G(t)) - u(t)$ is the net profit rate at time t. Also $[\pi(G(t)) - u(t)]e^{-\rho t}$ is the net profit rate at time t discounted to time 0, i.e., the present value of the time t profit rate. Hence, J can be interpreted as the total value of discounted future profits, and is the quantity we are trying to maximize.

There are control constraints $0 \le u(t) \le Q$, where Q is the upper bound on the advertising rate. However, there is no state constraint. It can be seen from the state equation and the control constraints that the goodwill $G(t)$ in fact never becomes negative.

You will find it instructive to compare this model with the previous one and note the similarities and differences between the two.

Example 1.3 *A Consumption Model.* Rich Rentier plans to retire at age 65 with a lump sum pension of W_0 dollars. Rich estimates his remaining life span to be T years. He wants to consume his wealth during these T retirement years, beginning at the age of 65, and leave a bequest to his heirs in a way that will maximize his total utility of consumption and bequest.

Since he does not want to take investment risks, Rich plans to put his money into a savings account that pays interest at a continuously compounded rate of r. In order to formulate Rich's optimization problem, let $t = 0$ denote the time when he turns 65 so that his retirement period can be denoted by the interval $[0, T]$. If we let the state variable $W(t)$ denote Rich's wealth and the control variable $C(t) \ge 0$ denote his rate of consumption at time $t \in [0, T]$, it is easy to see that the state equation is

$$\dot{W}(t) = rW(t) - C(t),$$

with the initial condition $W(0) = W_0 > 0$. It is reasonable to require that $W(t) \ge 0$ and $C(t) \ge 0, \ \ t \in [0, T]$. Letting $U(C)$ be the utility function of consumption C and $B(W)$ be the bequest function of leaving a bequest of amount W at time T, we see that the problem can be stated as an optimal control problem with the variables, equations, and constraints shown in Table 1.3.

Note that the objective function has two parts: first the integral of the discounted utility of consumption from time 0 to time T with ρ as the discount rate; and second the bequest function $e^{-\rho T}B(W)$, which measures Rich's discounted utility of leaving an estate W to his heirs at time T. If he has no heirs and does not care about charity, then $B(W) = 0$. However, if he has heirs or a favorite charity to whom he wishes to leave money, then $B(W)$ measures the strength of his desire to leave an estate of amount W. The nonnegativity constraints on state and control variables are obviously natural requirements that must be imposed.

You will be asked to solve this problem in Exercise 2.1 after you have learned the maximum principle in the next chapter. Moreover, a stochastic extension of the consumption problem, known as a consumption/investment problem, will be discussed in Sect. 12.4.

Table 1.3: The consumption model of Example 1.3

State variable	$W(t) =$ Wealth
Control variable	$C(t) =$ Consumption rate
State equation	$\dot{W}(t) = rW(t) - C(t),\ W(0) = W_0$
Objective function	$\text{Max}\left\{J = \int_0^T U(C(t))e^{-\rho t}dt + B(W(T))e^{-\rho T}\right\}$
State constraint	$W(t) \geq 0$
Control constraint	$C(t) \geq 0$
Terminal condition	$\cdots$
Exogenous	$U(C) =$ Utility of consumption
Functions	$B(W) =$ Bequest function
Parameters	$T =$ Terminal time
	$W_0 =$ Initial wealth
	$\rho =$ Discount rate
	$r =$ Interest rate

1.3 History of Optimal Control Theory

Optimal control theory is an extension of the calculus of variations (see Appendix B), so we discuss the history of the latter first.

The creation of the calculus of variations occurred almost immediately after the formalization of calculus by Newton and Leibniz in the seventeenth century. An important problem in calculus is to find an argument of a function at which the function takes on its maximum or minimum. The extension of this problem posed in the calculus of variations is to find a function which maximizes or minimizes the value of an integral or functional of that function. As might be expected, the extremum problem in the calculus of variations is much harder than the extremum problem in differential calculus. Euler and Lagrange are generally considered to be the founders of the calculus of variations. Newton, Legendre, and the Bernoulli brothers also contributed much to the early development of the field.

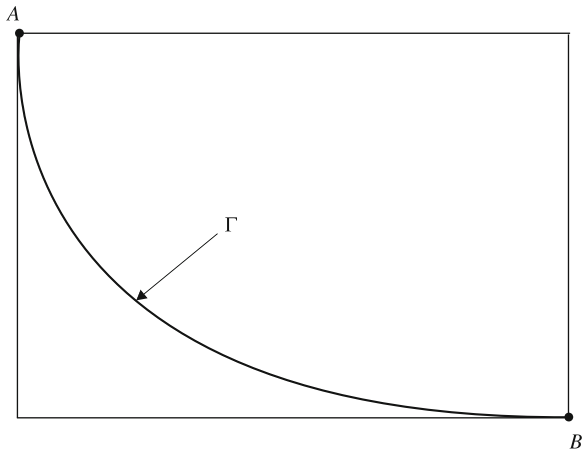

Figure 1.1: The Brachistochrone problem

A celebrated problem first solved using the calculus of variations was the *path of least time* or the *Brachistochrone problem*. The problem is illustrated in Fig. 1.1. It involves finding the shape of a curve Γ connecting the two points A and B in the vertical plane with the property that a bead sliding along the curve under the influence of gravity will move from A to B in the shortest possible time. The problem was posed by Johann Bernoulli in 1696, and it played an important part in the development of calculus of variations. It was solved by Johann Bernoulli, Jakob Bernoulli, Newton, Leibnitz, and L'Hôpital. In Sect. B.4, we provide a solution to the Brachistochrone problem by using what is known as the *Euler-Lagrange equation*, stated in Sect. B.2, and show that the shape of the solution curve is represented by a *cycloid*.

In the nineteenth and early twentieth centuries, many mathematicians contributed to the calculus of variations; these include Hamilton, Jacobi, Bolza, Weierstrass, Carathéodory, and Bliss.

Converting calculus of variations problems into control theory problems requires one more conceptual step—the addition of control variables to the state equations. Isaacs (1965) made such an extension in two-person pursuit-evasion games in the period 1948–1955. Bellman (1957) made a similar extension with the idea of dynamic programming.

Modern control theory began with the publication (in Russian in 1961 and English in 1962) of the book, *The Mathematical Theory of Optimal Processes*, by Pontryagin et al. (1962). Well-known American mathematicians associated with the maximum principle include Valentine, McShane, Hestenes, Berkovitz, and Neustadt. The importance of the book by Pontryagin et al. lies not only in a rigorous formulation of a calculus of variations problem with constrained control variables, but also in the proof of the maximum principle for optimal control problems. See Pesch and Bulirsch (1994) and Pesch and Plail (2009) for historical perspectives on the topics of the calculus of variations, dynamic programming, and optimal control.

The maximum principle permits the *decoupling* of the dynamic problem over time, using what are known as *adjoint variables* or *shadow prices*, into a series of problems, each of which holds at a single instant of time. The optimal solution of the instantaneous problems can be shown to give the optimal solution to the overall problem.

In this book we will be concerned principally with the application of the maximum principle in its various forms to find the solutions of a wide variety of applied problems in management science and economics. It is hoped that the reader, after reading some of these problems and their solutions, will appreciate, as we do, the importance of the maximum principle.

Some important books and surveys of the applications of the maximum principle to management science and economics are Connors and Teichroew (1967), Arrow and Kurz (1970), Hadley and Kemp (1971), Bensoussan et al. (1974), Stöppler (1975), Clark (1976), Sethi (1977a, 1978a), Tapiero (1977, 1988), Wickwire (1977), Bookbinder and Sethi (1980), Lesourne and Leban (1982), Tu (1984), Feichtinger and Hartl (1986), Carlson and Haurie (1987b), Seierstad and Sydsæter (1987), Erickson (2003), Léonard and Long (1992), Kamien and Schwartz (1992), Van Hilten et al. (1993), Feichtinger et al. (1994a), Maimon et al. (1998), Dockner et al. (2000), Caputo (2005), Grass et al. (2008), and Bensoussan (2011). Nevertheless, we have included in our bibliography many works of interest.

1.4 Notation and Concepts Used

In order to make the book readable, we will adopt the following notation which will hold throughout the book. In addition, we will define some important concepts that are required, including those of concave, convex and affine functions, and saddle points.

We use the symbol "=" to mean "is equal to" or "is defined to be equal to" or "is identically equal to" depending on the context. The symbol ":=" means "is defined to be equal to," the symbol "≡" means "is identically equal to," and the symbol "≈" means "is approximately equal to." The double arrow "⇒" means "implies," "∀" means "for all," and "∈" means "is a member of." The symbol □ indicates the end of a proof.

Let y be an n-component column vector and z be an m-component row vector, i.e.,

$$y = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} = (y_1, \ldots, y_n)^T \text{ and } z = (z_1, \ldots, z_m),$$

where the superscript ^{T} on a vector (or, a matrix) denotes the transpose of the vector (or, the matrix). At times, when convenient and not confusing, we will use the superscript $'$ for the transpose operation. If y and

z are functions of time t, a scalar, then the time derivatives $\dot{y} := dy/dt$ and $\dot{z} := dz/dt$ are defined as

$$\dot{y} = \frac{dy}{dt} = (\dot{y}_1, \cdots, \dot{y}_n)^T \text{ and } \dot{z} = \frac{dz}{dt} = (\dot{z}_1, \ldots, \dot{z}_m),$$

where $\dot{y}_i$ and $\dot{z}_j$ denote the time derivatives dy_i/dt and dz_j/dt, respectively.

When $n = m$, we can define the inner product

$$zy = \Sigma_{i=1}^n z_i y_i. \tag{1.7}$$

More generally, if

$$A = \{a_{ij}\} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1k} \\ a_{21} & a_{22} & \cdots & a_{2k} \\ \vdots & \vdots & \cdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mk} \end{bmatrix}$$

is an $m \times k$ matrix and $B = \{b_{ij}\}$ is a $k \times n$ matrix, we define the matrix product $C = \{c_{ij}\} = AB$, which is an $m \times n$ matrix with components

$$c_{ij} = \Sigma_{r=1}^k a_{ir} b_{rj}. \tag{1.8}$$

Let E^k denote the k-dimensional Euclidean space. Its elements are k-component vectors, which may be either row or column vectors, depending on the context. Thus in (1.7), $y \in E^n$ is a column vector and $z \in E^m$ is a row vector.

Next, in Sects. 1.4.1–1.4.4, we provide the notation for multivariate differentiation. Needless to say, the functions introduced are assumed to be appropriately differentiable for their derivatives being defined.

1.4.1 Differentiating Vectors and Matrices with Respect To Scalars

Let $f : E^1 \to E^k$ be a k-dimensional function of a scalar variable t. If f is a row vector, then we define

$$\frac{df}{dt} = f_t = (f_{1t}, f_{2t}, \cdots, f_{kt}), \text{ a row vector.}$$

We will also use the notation $f' = (f_1', f_2', \cdots, f_k')$ and $f'(t)$ in place of f_t.

If f is a column vector, then

$$\frac{df}{dt} = f_t = \begin{bmatrix} f_{1t} \\ f_{2t} \\ \vdots \\ f_{kt} \end{bmatrix} = (f_{1t}, f_{2t}, \cdots, f_{kt})^T, \text{a column vector.}$$

Once again, $f(t)$ may also be written as f' or $f'(t)$.

A similar rule applies if a matrix function is differentiated with respect to a scalar.

Example 1.4 Let $f(t) = \begin{bmatrix} t^2 & 2t+3 \\ e^{3t} & 1/t \end{bmatrix}$. Find f_t.

Solution $f_t = \begin{bmatrix} 2t & 2 \\ 3e^{3t} & -1/t^2 \end{bmatrix}$.

1.4.2 Differentiating Scalars with Respect to Vectors

If $F(y, z)$ is a scalar function defined on $E^n \times E^m$ with y an n-dimensional column vector and z an m-dimensional row vector, then the *gradients* F_y and F_z are defined, respectively, as

$$F_y = (F_{y_1}, \cdots, F_{y_n}), \text{ a row vector,} \tag{1.9}$$

and

$$F_z = (F_{z_1}, \cdots, F_{z_m}), \text{ a row vector,} \tag{1.10}$$

where F_{y_i} and F_{z_j} denote the partial derivatives with respect to the subscripted variables.

Thus, we always define the gradient with respect to a row or column vector as a row vector. Alternatively, F_y and F_z are also denoted as $\nabla_y F$ and $\nabla_z F$, respectively. In this notation, if F is a function of y only or z only, then the subscript can be dropped and the gradient of F can be written simply as ∇F.

Example 1.5 Let $F(y,z) = {y_1}^2 y_3 z_2 + 3y_2 \ln z_1 + y_1 y_2$, where $y = (y_1, y_2, y_3)^T$ and $z = (z_1, z_2)$. Obtain F_y and F_z.

Solution $F_y = (F_{y_1}, F_{y_2}, F_{y_3}) = (2y_1 y_3 z_2 + y_2, 3\ln z_1 + y_1, {y_1}^2 z_2)$ and $F_z = (F_{z_1}, F_{z_2}) = (3y_2/z_1, {y_1}^2 y_3)$.

1.4.3 Differentiating Vectors with Respect to Vectors

If $f : E^n \times E^m \to E^k$ is a k-dimensional vector function, f either row or column, i.e.,

$$f = (f_1, \cdots, f_k) \text{ or } f = (f_1, \cdots, f_k)^T,$$

where each component $f_i = f_i(y,z)$ depends on the column vector $y \in E^n$ and the row vector $z \in E^m$, then f_z will denote the $k \times m$ matrix

$$f_z = \begin{bmatrix} \partial f_1/\partial z_1, & \partial f_1/\partial z_2, & \cdots & \partial f_1/\partial z_m \\ \partial f_2/\partial z_1, & \partial f_2/\partial z_2, & \cdots & \partial f_2/\partial z_m \\ \vdots & \vdots & \cdots & \vdots \\ \partial f_k/\partial z_1, & \partial f_k/\partial z_2, & \cdots & \partial f_k/\partial z_m \end{bmatrix} = \{\partial f_i/\partial z_j\}, \tag{1.11}$$

and f_y will denote the $k \times n$ matrix

$$f_y = \begin{bmatrix} \partial f_1/\partial y_1 & \partial f_1/\partial y_2 & \cdots & \partial f_1/\partial y_n \\ \partial f_2/\partial y_1 & \partial f_2/\partial y_2 & \cdots & \partial f_2/\partial y_n \\ \vdots & \vdots & \cdots & \vdots \\ \partial f_k/\partial y_1 & \partial f_k/\partial y_2 & \cdots & \partial f_k/\partial y_n \end{bmatrix} = \{\partial f_i/\partial y_j\}. \tag{1.12}$$

Matrices f_z and f_y are known as *Jacobian* matrices. It should be emphasized that the rule of defining a Jacobian does not depend on the row or column nature of the function or its arguments. Thus,

$$f_z = (f^T)_z = f_{z^T} = (f^T)_{z^T}.$$

Example 1.6 Let $f : E^3 \times E^2 \to E^3$ be defined by $f(y,z) = ({y_1}^2 y_3 z_2 + 3y_2 \ln z_1, z_1 {z_2}^2 y_3, z_1 y_1 + z_2 y_2)^T$ with $y = (y_1, y_2, y_3)^T$ and $z = (z_1, z_2)$. Obtain f_z and f_y.

Solution.

$$f_z = \begin{bmatrix} 3y_2/z_1 & {y_1}^2 y_3 \\ {z_2}^2 y_3 & 2z_1 z_2 y_3 \\ y_1 & y_2 \end{bmatrix},$$

$$f_y = \begin{bmatrix} 2y_1 y_3 z_2 & 3\ln z_1 & {y_1}^2 z_2 \\ 0 & 0 & z_1 {z_2}^2 \\ z_1 & z_2 & 0 \end{bmatrix}.$$

Applying the rule (1.11) to F_y in (1.9), we obtain $F_{yz} = (F_y)_z$ to be the $n \times m$ matrix

$$F_{yz} = \begin{bmatrix} F_{y_1 z_1} & F_{y_1 z_2} & \cdots & F_{y_1 z_m} \\ F_{y_2 z_1} & F_{y_2 z_2} & \cdots & F_{y_2 z_m} \\ \vdots & \vdots & \cdots & \vdots \\ F_{y_n z_1} & F_{y_n z_2} & \cdots & F_{y_n z_m} \end{bmatrix} = \left\{ \frac{\partial^2 F}{\partial y_i \partial z_j} \right\}. \tag{1.13}$$

Applying the rule (1.12) to F_z in (1.10), we obtain $F_{zy} = (F_z)_y$ to be the $m \times n$ matrix

$$F_{zy} = \begin{bmatrix} F_{z_1 y_1} & F_{z_1 y_2} & \cdots & F_{z_1 y_n} \\ F_{z_2 y_1} & F_{z_2 y_2} & \cdots & F_{z_2 y_n} \\ \vdots & \vdots & \cdots & \vdots \\ F_{z_m y_1} & F_{z_m y_2} & \cdots & F_{z_m y_n} \end{bmatrix} = \left\{ \frac{\partial^2 F}{\partial z_i \partial y_j} \right\}. \tag{1.14}$$

Note that if $F(y, z)$ is twice continuously differentiable, then we also have $F_{zy} = (F_{yz})^T$.

Example 1.7 Obtain F_{yz} and F_{zy} for $F(y, z)$ specified in Example 1.5. Since the given $F(y, z)$ is twice continuously differentiable, check also that $F_{zy} = (F_{yz})^T$.

Solution. Applying rule (1.11) to F_y obtained in Example 1.5 and rule (1.12) to F_z obtained in Example 1.5, we have, respectively,

$$F_{yz} = \begin{bmatrix} 0 & 2y_1y_3 \\ 3/z_1 & 0 \\ 0 & {y_1}^2 \end{bmatrix} \text{ and } F_{zy} = \begin{bmatrix} 0 & 3/z_1 & 0 \\ 2y_1y_3 & 0 & {y_1}^2 \end{bmatrix}.$$

Also, it is easily seen from these matrices that $F_{zy} = (F_{yz})^T$.

1.4.4 Product Rule for Differentiation

Let g be an n-component row vector function and f be an n-component column vector function of an n-component vector x. Then in Exercise 1.9, you are asked to show that

$$(gf)_x = gf_x + f^T g_x = gf_x + f^T(g^T)_x. \tag{1.15}$$

In Exercise 1.10, you are asked to show further that with $g = F_x$, where $x \in E^n$ and the function $F : E^n \to E^1$ is twice continuously differentiable so that $F_{xx} = (F_{xx})^T$, called the Hessian, then

$$(gf)_x = (F_x f)_x = F_x f_x + f^T F_{xx} = F_x f_x + (F_{xx} f)^T. \tag{1.16}$$

The latter result will be used in Chap. 2 for the derivation of (2.25).

Many mathematical expressions in this book will be vector equations or inequalities involving vectors and vector functions. Since scalars are a special case of vectors, these expressions hold just as well for scalar equations or inequalities involving scalars and scalar functions. In fact, it may be a good idea to read them as scalar expressions on the first reading. Then in the second and further readings, the extension to vector form will be easier.

1.4.5 Miscellany

The *norm* of an m-component row or column vector z is defined to be

$$\parallel z \parallel = \sqrt{z_1^2 + \cdots + z_m^2}. \tag{1.17}$$

The norm of a vector is commonly used to define a *neighborhood* N_{z_0} of a point, e.g.,

$$N_{z_0} = \{z | \parallel z - z_0 \parallel < \varepsilon\}, \tag{1.18}$$

where $\varepsilon > 0$ is a small positive real number.

We will occasionally make use of the so-called "little-o" notation $o(z)$. A function $F(z) : E^m \to E^1$ is said to be of the order $o(z)$, if

$$\lim_{\|z\| \to 0} \frac{F(z)}{\| z \|} = 0.$$

The most common use of this notation will be to collect higher order terms in a series expansion.

In the continuous-time models discussed in this book, we generally will use $x(t)$ to denote the state (column) vector, $u(t)$ to denote the control (column) vector, and $\lambda(t)$ to denote the adjoint (row) vector. Whenever there is no possibility of confusion, we will suppress the time indicator (t) from these vectors and write them as x, u, and λ, respectively. When talking about *optimal* state and control vectors, we put an asterisk "*" as a superscript, i.e., as x^* and u^*, respectively, whereas u will refer to an admissible control with x as the corresponding state. No asterisk, however, needs to be put on the adjoint vector λ as it is only defined along an optimal path.

Thus, the values of the control, state and adjoint variables at time t along an optimal path will be written as $u^*(t), x^*(t)$, and $\lambda(t)$. When the control is expressed in terms of the state, it is called a *feedback control.* With an abuse of notation, we will express it as $u(x)$, or $u(x, t)$ if an explicit time dependence is required. Likewise, the optimal feedback control will be denoted as $u^*(x)$ or $u^*(x, t)$.

We also use the simplified notation $x'(t)$ to mean $(x(t))'$, the transpose of $x(t)$. Likewise, for a matrix $A(t)$, we use $A'(t)$ to mean $(A(t))'$ or the transpose of $A(t)$, and $A^{-1}(t)$ to mean $(A(t))^{-1}$ or the inverse of $A(t)$, when the inverse exists.

The *norm* of an m-dimensional row or column vector function $z(t)$, $t \in [0, T]$, is defined to be

$$\| z \| = \left[\Sigma_{j=1}^{m} \int_0^T z_j^2(\tau) d\tau \right]^{\frac{1}{2}}. \tag{1.19}$$

In Chap. 4 and some other chapters, we will encounter functions of time with jumps. For such functions, it is useful to have the concepts of *left* and *right limits.* With $\varepsilon > 0$, these are defined, respectively, for a function $x(t)$ as

$$x(T^-) = \lim_{\tau \uparrow T} x(\tau) = \lim_{\varepsilon \to 0} x(T - \varepsilon) \text{ and } x(T^+) = \lim_{\tau \downarrow T} x(\tau) = \lim_{\varepsilon \to 0} x(T + \varepsilon). \tag{1.20}$$

These limits are illustrated for a function $x(t)$ graphed in Fig. 1.2. Here,

$$x(0) = 1,\ x(0^+) = 2,$$

$$x(1^-) = 3,\ x(1^+) = x(1) = 4,$$

$$x(2^-) = 3,\ x(2) = 2,\ x(2^+) = 1,$$

$$x(3^-) = 2,\ x(3) = 3.$$

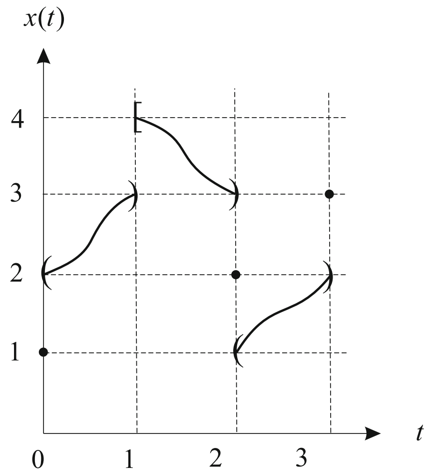

Figure 1.2: Illustration of left and right limits

In the discrete-time models introduced in Chap. 8 and applied in Chap. 9, we use x^k, u^k, and λ^k to denote state, control, and adjoint vectors, respectively, at time k, $k = 0, 1, 2, \ldots, T$. We also denote the difference operator by

$$\Delta x^k := x^{k+1} - x^k.$$

As in the continuous-time case, the optimal values of the state variable x^k and the control variable u^k will have an asterisk as a superscript; thus, x^{k*} and u^{k*} denote the corresponding quantities along an optimal path. Once again, the adjoint variable λ^k along an optimal path will not have an asterisk.

In order to specify the optimal control for linear control problems, we will introduce a special notation, called the *bang function*, as

$$\text{bang}[b_1, b_2; W] = \begin{cases} b_1 & \text{if} \quad W < 0, \\ \text{arbitrary} & \text{if} \quad W = 0, \\ b_2 & \text{if} \quad W > 0. \end{cases} \tag{1.21}$$

In order to specify the optimal control for linear-quadratic problems, we define another special function, called the *sat function*, as

$$\text{sat}[y_1, y_2; W] = \begin{cases} y_1 & \text{if} \quad W < y_1, \\ W & \text{if} \quad y_1 \le W \le y_2, \\ y_2 & \text{if} \quad W > y_2. \end{cases} \tag{1.22}$$

The word “sat” is short for the word “saturation.” The latter name comes from an electrical engineering application to *saturated amplifiers.*

In several applications to be discussed, we will need the concept of impulse control, which is sometimes needed in cases when an unbounded control can be applied for a very short time. An example is the advertising model in Table 1.2 when $Q = \infty$. We apply unbounded control for a short time in order to cause a jump discontinuity in the state variable. For the example in Table 1.2, this might mean an intense advertising campaign (a media blitz) in order to increase advertising goodwill by a finite amount in a very short time. The impulse function defined below is required to evaluate the integral in the objective function, which measures the cost of the intense advertising campaign.

Suppose we want to apply an impulse control at time t to change the state variable from $x(t) = x_1$ to the value x_2 “immediately” after t, i.e., $x(t^+) = x_2$. To compute its contribution to the objective function (1.2), we use the following procedure: given $\varepsilon > 0$ and a constant control $u(\varepsilon)$, integrate (1.1) from t to $t + \varepsilon$ with $x(t) = x_1$ and choose $u(\varepsilon)$ so that $x(t + \varepsilon) = x_2$; this gives the trajectory $x(\tau; \varepsilon, u(\varepsilon))$ for $\tau \in [t, t + \varepsilon]$. We can now compute

$$\text{imp}(x_1, x_2; t) = \lim_{\varepsilon \to 0} \int_t^{t+\varepsilon} F(x, u, \tau) d\tau. \tag{1.23}$$

If the impulse is applied only at time t, then we can calculate (1.2) as

$$J = \int_0^t F(x,u,\tau)d\tau + \text{imp}(x_1,x_2;t) + \int_t^T F(x,u,\tau)d\tau + S[x(T),T]. \tag{1.24}$$

If there are several instants at which impulses are applied, then this procedure is easily extended. Examples of the use of (1.24) occur in Chaps. 5 and 6. We frequently omit t in (1.23) when the impulse function is independent of t.

1.4.6 Convex Set and Convex Hull

A set $D \subset E^n$ is a *convex set* if for each pair of points $y, z \in D$, the entire line segment joining these two points is also in D, i.e.,

$$py + (1-p)z \in D, \text{ for each } p \in [0,1].$$

Given $x^i \in E^n, i = 1,2,\ldots,l$, we define $y \in E^n$ to be a *convex combination* of $x^i \in E^n$, if there exists $p_i \geq 0$ such that

$$\sum_{i=1}^{l} p_i = 1 \text{ and } y = \sum_{i=1}^{l} p_i x^i.$$

The *convex hull* of a set $D \subset E^n$ is

$$\text{co}D := \left\{ \sum_{i=1}^{l} p_i x^i : \sum_{i=1}^{l} p_i = 1,\ p_i \geq 0,\ x^i \in D,\ i = 1,2,\ldots,l \right\}.$$

In other words, co$D$ is the set of all convex combinations of points in D.

1.4.7 Concave and Convex Functions

A real-valued function ψ defined on a convex set $D \subset E^n$, i.e., $\psi : D \to E^1$, is *concave*, if for each pair of points $y, z \in D$ and for all $p \in [0,1]$,

$$\psi(py + (1-p)z) \geq p\psi(y) + (1-p)\psi(z).$$

If the inequalities in the above definition are strict for all $y, z \in D$ with $y \neq z$, and $0 < p < 1$, then ψ is called a *strictly concave function.*

In the single dimensional case of $n = 1$, there is an enlightening geometrical interpretation. Namely, $\psi(x)$ defined on an interval $D = [a, b]$ is concave if, for each pair of points on the graph of $\psi(x)$, the line segment joining these two points lies entirely below or on the graph of $\psi(x)$; see Fig. 1.3.

Reverting back to the n-dimensional case, if ψ is a differentiable function on a convex set $D \subset E^n$, then it is *concave*, if for each pair of points $y, z \in D$,

$$\psi(z) \leq \psi(y) + \psi_x(y)(z - y),$$

where we understand y and z to be column vectors. Furthermore, if the function ψ is twice differentiable, then it is *concave*, if at each point in D, the $n \times n$ symmetric matrix ψ_{xx} is negative semidefinite, i.e., all of its eigenvalues are non-positive.

Finally, if ψ is a concave function, then the negative of the function ψ, i.e., $-\psi : D \to E^1$, is a *convex function.*

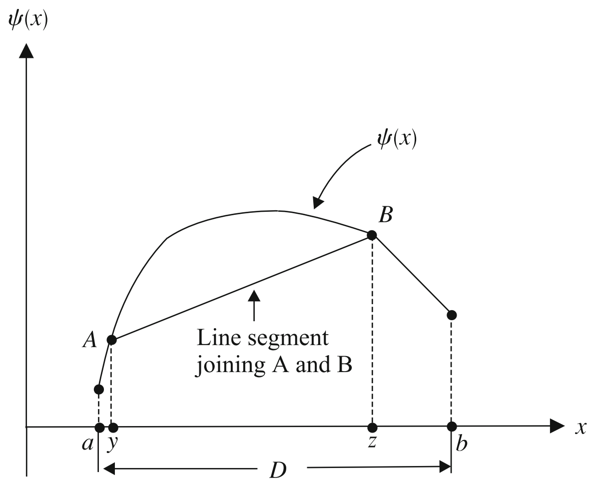

Figure 1.3: A concave function

1.4.8 Affine Function and Homogeneous Function of Degree k

A function $\psi : E^n \to E^1$ is said to be *affine*, if the function $\psi(x) - \psi(0)$ is linear. Thus, ψ can be represented as $\psi(x_1, x_2, \ldots, x_n) = \sum_{i=1}^n a_i x_i + b$, where $a_i,\ i = 1, 2, \ldots, n$, and b are scalar constants.

A function $\psi : E^n \to E^1$ is said to be *homogeneous of degree* k, if $\psi(bx) = b^k \psi(x)$, where $b > 0$ is a scalar constant.

In economics, we often assume that a firm's production function is homogeneous of degree 1, i.e., if all inputs are multiplied by b, then output is multiplied by b. Such a production function is said to exhibit the property of *constant return to scale*. A linear function $\psi(x) = ax = \sum_{i=1}^n a_i x_i$ is a simple example of a homogeneous function of degree 1. Other examples are $\psi(x) = \min\{x_i,\ i = 1, 2, \ldots, n\}$ and $\psi(x) = a(\Pi_{i=1}^n {x_i}^{\alpha_i})^{1/\sum_{i=1}^n \alpha_i}$ with $a > 0$ and $\alpha_i > 0,\ i = 1, 2, \ldots, n$. An important special case of the last example, known as the Cobb-Douglas production function, is $\psi(K, L) = aK^{\alpha_1} L^{\alpha_2}$ with $\alpha_1 + \alpha_2 = 1$, where K and L are factors of production called capital and labor, respectively, and a denotes the total factor productivity.

1.4.9 Saddle Point

An important concept in two-person zero-sum games is that of a saddle point. Let $\psi(x, y)$, a real-valued function defined on the space $E^n \times E^m$, i.e., $\psi : E^n \times E^m \to E^1$, be the payoff of player 1 and $-\psi(x, y)$ be the payoff of player 2, when they make decisions x and y, respectively, in a zero-sum game. A point $(\hat{x}, \hat{y}) \in E^n \times E^m$ is called a *saddle point* of $\psi(x, y)$ or of the game, if

$$\psi(\hat{x}, y) \geq \psi(\hat{x}, \hat{y}) \geq \psi(x, \hat{y}) \text{ for all } x \in E^n \text{ and } y \in E^m.$$

Note that a saddle point may not exist, and even if it exists, it may not be unique. Note also that

$$\psi(\hat{x}, \hat{y}) = \max_x \psi(x, \hat{y}) = \min_y \psi(\hat{x}, y).$$

Intuitively, this could produce a picture like a horse saddle as shown in Fig. 1.4, hence the name saddle point for a point like $(\hat{x}, \hat{y})$. This concept will be used in Sect. 13.1.

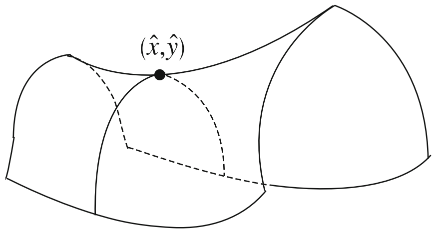

Figure 1.4: An illustration of a saddle point

1.4.10 Linear Independence and Rank of a Matrix

A set of vectors $a_1, a_2, \ldots, a_m$ in E^n, $m \leq n$, is said to be *linearly dependent* if there exist scalars p_i not all zero such that

$$\sum_{i=1}^{m} p_i a_i = 0. \tag{1.25}$$

If (1.25) holds only when $p_1 = p_2 = \cdots = p_m = 0$, then the vectors are said to be *linearly independent*. In particular, if one of the vectors in the set $\{a_1, a_2, \ldots, a_m\}$ is a *null vector*, then the set is linearly dependent.

The *rank* of an $m \times n$ matrix A, written rank(A), is the maximum number of linearly independent rows or, equivalently, the maximum number of linearly independent columns of A. An $m \times n$ matrix is of *full rank* if

$$\text{rank}(A) = \min\{m, n\}.$$

1.5 Plan of the Book

The book has thirteen chapters and five appendices: A, B, C, D, and E, covering a variety of topics which are listed in the table of contents and explained in the prefaces.

In any given chapter, say Chap. 7, sections are numbered consecutively as 7.1, 7.2, 7.3, etc. Subsections are numbered consecutively within each section, i.e., 7.2.1, 7.2.2, 7.2.3, etc. Mathematical expressions are numbered consecutively by chapter as (7.1), (7.2), (7.3), etc. Theorems are also numbered consecutively by chapter as Theorem 7.1, Theorem 7.2, Theorem 7.3, etc. Similarly, definitions, examples, exercises, figures, propositions, remarks, and tables are numbered consecutively by chapter. These elements will be referenced throughout the book by use of their designated numbers. The same scheme is used in the appendices, thus, sections in Appendix B, for example, are numbered as B.1, B.2, B.3, etc.

Exercises for Chapter 1

E 1.1 In Example 1.1, let the functions and parameters of the production- inventory model be given by:

$$h(I) = 10I,\ c(P) = 20P,\ T = 10,\ I_0 = 1,000$$

$$P_{\min} = 600,\ P_{\max} = 1200,\ I_{\min} = 800,\ S(t) = 900 + 10t.$$

(a) Set $P(t) = 1000$ for $0 \le t \le 10$. Determine whether this control is feasible; if it is feasible, compute the value J of the objective function.

(b) If $P(t) = 800$, show that the terminal constraint is violated and hence the control is infeasible.

(c) If $P(t) = P_{\min}$ for $0 \le t \le 6$ and $P(t) = P_{\max}$ for $6 < t \le 10$, show that the control is infeasible because the state constraint is violated.

E 1.2 In Example 1.1, suppose there is a cost associated with changing the rate of production. One way to formulate this problem is to let the control variable $u(t)$ denote the rate of change of the production rate $P(t)$, having a cost cu^2 associated with such changes, where $c > 0$. Formulate the new problem.

Hint: Let $P(t)$ be an additional state variable.

E 1.3 For the advertising model in Example 1.2, let $\pi(G) = 2\sqrt{G}$, $\delta = 0.05$, $\rho = 0.2$, $Q = 2$, and $G_0 = 16$. Set $u(t) = 0.8$ for $t \ge 0$, and show that $G(t)$ is constant for all t. Compute the value J of the objective function.

E 1.4 In Example 1.2, suppose G measures the number of people who know about the product. Hence, if A is the total population, then $A - G$ is the number of people who do not know about the product. If $u(t)$ measures the advertising rate at time t, assume that $u(A - G)$ is the corresponding rate of increase of G due to this advertising. Formulate the new model.

E 1.5 Rich Rentier in Example 1.3 has initial wealth $W_0 = \$1,000,000$. Assume $B = 0, \rho = 0.1, r = 0.15$, and assume that Rich expects to live for exactly 20 years.

(a) What is the maximum constant consumption level that Rich can afford during his remaining life?

(b) If Rich's utility function is $U(C) = \ln C$, what is the present value of the total utility in part (a)?

(c) Suppose Rich sets aside \$100,000 to start the Rentier Foundation. What is the maximum constant grant level that the foundation can support if it is to last forever?

E 1.6 Suppose Rich in Exercise 1.5 takes on a part-time job, which yields an income of $y(t)$ at time t. Assume $y(t) = 10,000e^{-0.05t}$ and that he has a bequest function $B(W) = 0.5 \ln W$.

(a) Reformulate this new optimal control problem.

(b) If Rich (no longer a *rentier*) consumes at the constant rate found in Exercise 1.5(a), find his terminal wealth and his new total utility.

E 1.7 Consider the following educational policy question. Let $S(t)$ denote the total number of scientists at time t, and let δ be the retirement rate of scientists. Let $E(t)$ be the number of teaching scientists and $R(t)$ be the number of research scientists, so that $S(t) = E(t) + R(t)$. Assume $\gamma E(t)$ is the number of newly graduated scientists at time t, of which the policy allocates $u\gamma E(t)$ to the pool of teachers, where $0 \leq u \leq 1$. The remaining graduates are added to the pool of researchers. The government has a target of maximizing the function $\alpha E(T) + \beta R(T)$ at a given future time T, where α and β are positive constants. Formulate the optimal control problem for the government.

E 1.8 For $F(x, y)$ defined in Example 1.5, obtain the matrices F_{xx} and F_{yy}.

E 1.9 Let $x \in E^m$, g be an n-component row vector function of x, and f be an n-component column vector function of x. Use the ordinary product rule of calculus for functions of scalars to derive the formula

$$(gf)_x = gf_x + f^T(g^T)_x = gf_x + f^T g_x.$$

E 1.10 Let F be a scalar function of $x \in E^n$ and f as defined in Exercise 1.9. Assume F to be twice continuously differentiable. Show that

$$(F_x f)_x = F_x f_x + f^T F_{xx} = F_x f_x + f^T (F_{xx})^T = F_x f_x + (F_{xx} f)^T.$$

Hint: Set the gradient $F_x = g$, a row vector, and then use Exercise 1.9 to derive the first equality. Note in connection with the second equality that the function F being twice continuously differentiable implies that $F_{xx} = (F_{xx})^T$.

E 1.11 For F_y obtained in Example 1.5 and f defined in Example 1.6, obtain $(F_y f)_y$ and verify the relation shown in Exercise 1.10.

E 1.12 Use the bang function defined in (1.21) to sketch the optimal control

$$u^*(t) = \text{bang}[-1, 1; W(t)] \text{ for } 0 \le t \le 5,$$

when

(a) $W(t) = t - 2$

(b) $W(t) = t^2 - 4t + 3$

(c) $W(t) = \sin \pi t$.

E 1.13 Use the sat function defined in (1.22) to sketch the optimal control

$$u^*(t) = \text{sat}[2, 3; W(t)] \text{ for } 0 \le t \le 5,$$

when

(a) $W(t) = 4 - t$

(b) $W(t) = 2 + t^2$

(c) $W(t) = 4 - 4e^{-t}$.

E 1.14 Evaluate the function $\text{imp}(G_1, G_2; t)$ for the advertising model of Table 1.2 when $G_2 > G_1$, $Q = \infty$, and $\pi(G) = pG$, where p is a constant.