Lecture Notes on Optimization

Pravin Varaiya

Contents

PREFACE to this edition

Notes on Optimization was published in 1971 as part of the Van Nostrand Reinhold Notes on System Sciences, edited by George L. Turin. Our aim was to publish short, accessible treatments of graduate-level material in inexpensive books (the price of a book in the series was about five dollars). The effort was successful for several years. Van Nostrand Reinhold was then purchased by a conglomerate which cancelled Notes on System Sciences because it was not sufficiently profitable. Books have since become expensive. However, the World Wide Web has again made it possible to publish cheaply.

Notes on Optimization has been out of print for 20 years. However, several people have been using it as a text or as a reference in a course. They have urged me to re-publish it. The idea of making it freely available over the Web was attractive because it reaffirmed the original aim. The only obstacle was to retype the manuscript in LaTex. I thank Kate Klohe for doing just that.

I would appreciate knowing if you find any mistakes in the book, or if you have suggestions for (small) changes that would improve it.

Berkeley, California P.P. Varaiya
September, 1998

PREFACE

These *Notes* were developed for a ten-week course I have taught for the past three years to first-year graduate students of the University of California at Berkeley. My objective has been to present, in a compact and unified manner, the *main* concepts and techniques of mathematical programming and optimal control to students having diverse technical backgrounds. A reasonable knowledge of advanced calculus (up to the Implicit Function Theorem), linear algebra (linear independence, basis, matrix inverse), and linear differential equations (transition matrix, adjoint solution) is sufficient for the reader to follow the *Notes*.

The treatment of the topics presented here is deep. Although the coverage is not encyclopedic, an understanding of this material should enable the reader to follow much of the recent technical literature on nonlinear programming, (deterministic) optimal control, and mathematical economics. The examples and exercises given in the text form an integral part of the *Notes* and most readers will need to attend to them before continuing further. To facilitate the use of these *Notes* as a textbook, I have incurred the cost of some repetition in order to make almost all chapters self-contained. However, Chapter V must be read before Chapter VI, and Chapter VII before Chapter VIII.

The selection of topics, as well as their presentation, has been influenced by many of my students and colleagues, who have read and criticized earlier drafts. I would especially like to acknowledge the help of Professors M. Athans, A. Cohen, C.A. Desoer, J-P. Jacob, E. Polak, and Mr. M. Ripper. I also want to thank Mrs. Billie Vrtiak for her marvelous typing in spite of starting from a not terribly legible handwritten manuscript. Finally, I want to thank Professor G.L. Turin for his encouraging and patient editorship.

Berkeley, California P.P. Varaiya
November, 1971

Chapter 1

INTRODUCTION

In this chapter, we present our model of the optimal decision-making problem, illustrate decision-making situations by a few examples, and briefly introduce two more general models which we cannot discuss further in these *Notes*.

1.1 *The Optimal Decision Problem*

These *Notes* show how to arrive at an optimal decision assuming that complete information is given. The phrase *complete information is given* means that the following requirements are met:

1. The set of all permissible decisions is known, and
2. The cost of each decision is known.

When these conditions are satisfied, the decisions can be ranked according to whether they incur greater or lesser cost. An *optimal decision* is then any decision which incurs the least cost among the set of permissible decisions.

In order to model a decision-making situation in mathematical terms, certain further requirements must be satisfied, namely,

1. The set of all decisions can be adequately represented as a subset of a vector space with each vector representing a decision, and
2. The cost corresponding to these decisions is given by a real-valued function.

Some illustrations will help.

Example 1: The Pot Company (Potco) manufacturers a smoking blend called Acapulco Gold. The blend is made up of tobacco and mary-john leaves. For legal reasons the fraction α of mary-john in the mixture must satisfy $0 < \alpha < \frac{1}{2}$. From extensive market research Potco has determined their expected volume of sales as a function of α and the selling price p. Furthermore, tobacco can be purchased at a fixed price, whereas the cost of mary-john is a function of the amount purchased. If Potco wants to maximize its profits, how much mary-john and tobacco should it purchase, and what price p should it set?

Example 2: Tough University provides "quality" education to undergraduate and graduate students. In an agreement signed with Tough's undergraduates and graduates (TUGs), "quality" is

defined as follows: every year, each u (undergraduate) must take eight courses, one of which is a seminar and the rest of which are lecture courses, whereas each g (graduate) must take two seminars and five lecture courses. A seminar cannot have more than 20 students and a lecture course cannot have more than 40 students. The University has a faculty of 1000. The Weary Old Radicals (WORs) have a contract with the University which stipulates that every junior faculty member (there are 750 of these) shall be required to teach six lecture courses and two seminars each year, whereas every senior faculty member (there are 250 of these) shall teach three lecture courses and three seminars each year. The Regents of Touch rate Tough's President at α points per u and β points per g "processed" by the University. Subject to the agreements with the TUGs and WORs how many u's and g's should the President admit to maximize his rating?

Example 3: (See Figure 1.1.) An engineer is asked to construct a road (broken line) connection point a to point b. The current profile of the ground is given by the solid line. The only requirement is that the final road should not have a slope exceeding 0.001. If it costs \$$c$ per cubic foot to excavate or fill the ground, how should he design the road to meet the specifications at minimum cost?

Example 4: Mr. Shell is the manager of an economy which produces one output, wine. There are two factors of production, capital and labor. If $K(t)$ and $L(t)$ respectively are the capital stock used and the labor employed at time t, then the rate of output of wine $W(t)$ at time is given by the production function

$$W(t) \quad = \quad F(K(t), L(t))$$

As Manager, Mr. Shell allocates some of the output rate $W(t)$ to the consumption rate $C(t)$, and the remainder $I(t)$ to investment in capital goods. (Obviously, W, C, I, and K are being measured in a common currency.) Thus, $W(t) = C(t) + I(t) = (1 - s(t))W(t)$ where $s(t) = I(t)/W(t)$

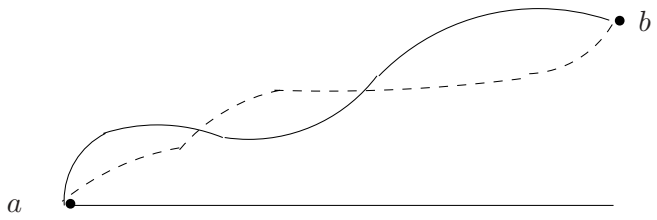


Figure 1.1: Admissable set of example.

$\in [0, 1]$ is the fraction of output which is saved and invested. Suppose that the capital stock decays exponentially with time at a rate $\delta > 0$, so that the net rate of growth of capital is given by the following equation:

$$\begin{aligned} \dot{K}(t) &= \frac{d}{dt}K(t) \\ &= -\delta K(t) + s(t)W(t) \\ &= -\delta K(t) + s(t)F(K(t), L(t)). \end{aligned} \tag{1.1}$$

The labor force is growing at a constant birth rate of $\beta > 0$. Hence,

$$\dot{L}(t) \quad = \quad \beta L(t). \tag{1.2}$$

Suppose that the production function F exhibits constant returns to scale, *i.e.*, $F(\lambda K, \lambda L) = \lambda F(K, L)$ for all $\lambda > 0$. If we define the relevant variable in terms of per capita of labor, $w = W/L, c = C/L, k = K/l$, and if we let $f(k) = F(k, l)$, then we see that $F(K, L) - LF(K/L, 1) = Lf(k)$, whence the consumption per capita of labor becomes $c(t) = (l - s(t))f(k(t))$. Using these definitions and equations (1.1) and (1.2) it is easy to see that $K(t)$ satisfies the differential equation (1.3).

$$\dot{k}(t) \quad = \quad s(t)f(k(t)) - \mu k(t) \tag{1.3}$$

where $\mu = (\delta + \beta)$. The first term of the right-hand side in (3) is the increase in the capital-to-labor ratio due to investment whereas the second terms is the decrease due to depreciation and increase in the labor force.

Suppose there is a planning horizon time T, and at time 0 Mr. Shell starts with capital-to-labor ratio k_o. If "welfare" over the planning period $[0, T]$ is identified with total consumption $\int_0^T c(t)dt$, what should Mr. Shell's savings policy $s(t)$, $0 \leq t \leq T$, be so as to maximize welfare? What savings policy maximizes welfare subject to the additional restriction that the capital-to-labor ratio at time T should be at least k_T? If future consumption is discounted at rate $\alpha > 0$ and if time horizon is ∞, the welfare function becomes $\int_0^\infty e^- \alpha t \, c(t)dt$. What is the optimum policy corresponding to this criterion?

These examples illustrate the kinds of decision-making problems which can be formulated mathematically so as to be amenable to solutions by the theory presented in these *Notes*. We must always remember that a mathematical formulation is inevitably an abstraction and the gain in precision may have occurred at a great loss of realism. For instance, Example 2 is caricature (see also a faintly related but more more elaborate formulation in Bruno [1970]), whereas Example 4 is light-years away from reality. In the latter case, the value of the mathematical exercise is greater the more insensitive are the optimum savings policies with respect to the simplifying assumptions of the mathematical model. (In connection with this example and related models see the critique by Koopmans [1967].)

In the examples above, the set of permissible decisions is represented by the set of all points in some vector space which satisfy certain constraints. Thus, in the first example, a permissible decision is any two-dimensional vector (α, p) satisfying the constraints $0 < \alpha < \frac{1}{2}$ and $0 < p$. In the second example, any vector (u, g) with $u \geq 0, \ g \geq 0$, constrained by the number of faculty and the agreements with the TUGs and WORs is a permissible decision. In the last example, a permissible decision is any real-valued function $s(t), \ 0 \leq t \leq T$, constrained by $0 \leq s(t) \leq 1$. (It is of mathematical but not conceptual interest to note that in this case a decision is represented by a vector in a function space which is infinite-dimensional.) More concisely then, these *Notes* are concerned with optimizing (i.e. maximizing or minimizing) a real-valued function over a vector space subject to constraints. The constraints themselves are presented in terms of functional inequalities or equalities.

At this point, it is important to realize that the distinction between the function which is to be optimized and the functions which describe the constraints, although convenient for presenting the mathematical theory, may be quite artificial in practice. For instance, suppose we have to choose the durations of various traffic lights in a section of a city so as to achieve optimum traffic flow. Let us suppose that we know the transportation needs of all the people in this section. Before we can begin to suggest a design, we need a criterion to determine what is meant by "optimum traffic flow." More abstractly, we need a criterion by which we can compare different decisions, which in this case are different patterns of traffic-light durations. One way of doing this is to assign as cost to each decision the total amount of time taken to make all the trips within this section. An alternative and equally plausible goal may be to minimize the maximum waiting time (that is the total time spent at stop lights) in each trip. Now it may happen that these two objective functions may be inconsistent in the sense that they may give rise to different orderings of the permissible decisions. Indeed, it may be the case that the optimum decision according to the first criterion may be lead to very long waiting times for a few trips, so that this decision is far from optimum according to the second criterion. We can then redefine the problem as minimizing the first cost function (total time for trips) subject to the constraint that the waiting time for any trip is less than some reasonable bound (say one minute). In this way, the second goal (minimum waiting time) has been modified and reintroduced as a constraint. This interchangeability of goal and constraints also appears at a deeper level in much of the mathematical theory. We will see that in most of the results the objective function and the functions describing the constraints are treated in the same manner.

1.2 *Some Other Models of Decision Problems*

Our model of a single decision-maker with complete information can be generalized along two very important directions. In the first place, the hypothesis of complete information can be relaxed by allowing that decision-making occurs in an uncertain environment. In the second place, we can replace the single decision-maker by a group of two or more agents whose collective decision determines the outcome. Since we cannot study these more general models in these *Notes*, we merely point out here some situations where such models arise naturally and give some references.

1.2.1 *Optimization under uncertainty.*

A person wants to invest $1,000 in the stock market. He wants to maximize his capital gains, and at the same time minimize the risk of losing his money. The two objectives are incompatible, since the stock which is likely to have higher gains is also likely to involve greater risk. The situation is different from our previous examples in that the outcome (future stock prices) is uncertain. It is customary to model this uncertainty stochastically. Thus, the investor may assign probability 0.5 to the event that the price of shares in Glamor company increases by $100, probability 0.25 that the price is unchanged, and probability 0.25 that it drops by $100. A similar model is made for all the other stocks that the investor is willing to consider, and a decision problem can be formulated as follows. How should $1,000 be invested so as to maximize the *expected value* of the capital gains subject to the constraint that the probability of losing more than $100 is less than 0.1?

As another example, consider the design of a controller for a chemical process where the decision variable are temperature, input rates of various chemicals, *etc*. Usually there are impurities in the chemicals and disturbances in the heating process which may be regarded as additional inputs of a

random nature and modeled as stochastic processes. After this, just as in the case of the portfolio-selection problem, we can formulate a decision problem in such a way as to take into account these random disturbances.

If the uncertainties are modelled stochastically as in the example above, then in many cases the techniques presented in these *Notes* can be usefully applied to the resulting optimal decision problem. To do justice to these decision-making situations, however, it is necessary to give great attention to the various ways in which the uncertainties can be modelled mathematically. We also need to worry about finding equivalent but simpler formulations. For instance, it is of great significance to know that, given appropriate conditions, an optimal decision problem under uncertainty is equivalent to another optimal decision problem under complete information. (This result, known as the Certainty-Equivalence principle in economics has been extended and baptized the Separation Theorem in the control literature. See Wonham [1968].) Unfortunately, to be able to deal with these models, we need a good background in Statistics and Probability Theory besides the material presented in these *Notes*. We can only refer the reader to the extensive literature on Statistical Decision Theory (Savage [1954], Blackwell and Girshick [1954]) and on Stochastic Optimal Control (Meditch [1969], Kushner [1971]).

1.2.2 *The case of more than one decision-maker.*

Agent Alpha is chasing agent Beta. The place is a large circular field. Alpha is driving a fast, heavy car which does not maneuver easily, whereas Beta is riding a motor scooter, slow but with good maneuverability. What should Alpha do to get as close to Beta as possible? What should Beta do to stay out of Alpha's reach? This situation is fundamentally different from those discussed so far. Here there are two decision-makers with opposing objectives. Each agent does not know what the other is planning to do, yet the effectiveness of his decision depends crucially upon the other's decision, so that optimality cannot be defined as we did earlier. We need a new concept of rational (optimal) decision-making. Situations such as these have been studied extensively and an elaborate structure, known as the Theory of Games, exists which describes and prescribes behavior in these situations. Although the practical impact of this theory is not great, it has proved to be among the most fruitful sources of unifying analytical concepts in the social sciences, notably economics and political science. The best single source for Game Theory is still Luce and Raiffa [1957], whereas the mathematical content of the theory is concisely displayed in Owen [1968]. The control theorist will probably be most interested in Isaacs [1965], and Blaquiere, *et al.*, [1969].

The difficulty caused by the lack of knowledge of the actions of the other decision-making agents arises even if all the agents have the same objective, since a particular decision taken by our agent may be better or worse than another decision depending upon the (unknown) decisions taken by the other agents. It is of crucial importance to invent schemes to coordinate the actions of the individual decision-makers in a consistent manner. Although problems involving many decision-makers are present in any system of large size, the number of results available is pitifully small. (See Mesarovic, *et al.*, [1970] and Marschak and Radner [1971].) In the author's opinion, these problems represent one of the most important and challenging areas of research in decision theory.

Chapter 2

OPTIMIZATION OVER AN OPEN SET

In this chapter we study in detail the first example of Chapter 1. We first establish some notation which will be in force throughout these *Notes*. Then we study our example. This will generalize to a canonical problem, the properties of whose solution are stated as a theorem. Some additional properties are mentioned in the last section.

2.1 *Notation*

2.1.1

All vectors are *column* vectors, with two consistent exceptions mentioned in 2.1.3 and 2.1.5 below and some other minor and convenient exceptions in the text. Prime denotes transpose so that if $x \in R^n$ then x' is the row vector x' = $(x_1, \ldots, x_n)$, and $x = (x_1, \ldots, x_n)'$. Vectors are normally denoted by lower case letters, the *ith* component of a vector $x \in R^n$ is denoted x_i, and different vectors denoted by the same symbol are distinguished by superscripts as in x^j and x^k. 0 denotes both the zero vector and the real number zero, but no confusion will result.

Thus if $x = (x_1, \ldots, x_n)'$ and $y = (y_1, \ldots, y_n)'$ then $x'y = x_1y_1 + \ldots + x_ny_n$ as in ordinary matrix multiplication. If $x \in R^n$ we define $|x| = +\sqrt{x'x}$.

2.1.2

If $x = (x_1, \ldots, x_n)'$ and $y = (y_1, \ldots, y_n)'$ then $x \geq y$ means $x_i \geq y_i$, $i = 1, \ldots, n$. In particular if $x \in R^n$, then $x \geq 0$, if $x_i \geq 0, i = 1, \ldots, n$.

2.1.3

Matrices are normally denoted by capital letters. If A is an $m \times n$ matrix, then A^j denotes the *jth column of A*, and A_i denotes the *ith row of A*. Note that A_i is a *row* vector. A_i^j denotes the entry of A in the ith row and jth column; this entry is sometimes also denoted by the lower case letter a_{ij}, and then we also write $A = \{a_{ij}\}$. I denotes the identity matrix; its size will be clear from the context. If confusion is likely, we write I_n to denote the $n \times n$ identity matrix.

2.1.4

If $f : R^n \to R^m$ is a function, its *ith* component is written $f_i, i = 1, \ldots, m$. Note that $f_i : R^n \to R$. Sometimes we describe a function by specifying a rule to calculate $f(x)$ for every x. In this case we write $f : x \mapsto f(x)$. For example, if A is an $m \times n$ matrix, we can write $F : x \mapsto Ax$ to denote the function $f : R^n \to R^m$ whose value at a point $x \in R^n$ is Ax.

2.1.5

If $f : R^n \mapsto R$ is a differentiable function, the derivative of f at $\hat{x}$ is the *row* vector $((\partial f/\partial x_1)(\hat{x}), \ldots, (\partial f/\partial x_n)(\hat{x}))$. This derivative is denoted by $(\partial f/\partial x)(\hat{x})$ or $f_x(\hat{x})$ or $\partial f/\partial x|_{x=\hat{x}}$ or $f_x|_{x=\hat{x}}$, and if the argument $\hat{x}$ is clear from the context it may be dropped. The *column* vector $(f_x(\hat{x}))'$ is also denoted $\nabla_x f(\hat{x})$, and is called the *gradient* of f at $\hat{x}$. If $f : (x, y) \mapsto f(x, y)$ is a differentiable function from $R^n \times R^m$ into R, the partial derivative of f with respect to x at the point $(\hat{x}, \hat{y})$ is the n-dimensional row vector $f_x(\hat{x}, \hat{y}) = (\partial f/\partial x)(\hat{x}, \hat{y}) = ((\partial f/\partial x_1)(\hat{x}, \hat{y}), \ldots, (\partial f/\partial x_n)(\hat{x}, \hat{y}))$, and similarly $f_y(\hat{x}, \hat{y}) = (\partial f/\partial y)(\hat{x}, \hat{y}) = ((\partial f/\partial y_1)(\hat{x}, \hat{y}), \ldots, (\partial f/\partial y_m)(\hat{x}, \hat{y}))$. Finally, if $f : R^n \to R^m$ is a differentiable function with components $f_1, \ldots, f_m$, then its derivative at $\hat{x}$ is the $m \times n$ matrix

$$\begin{aligned} \frac{\partial f}{\partial x}(\hat{x}) = f_x\hat{x} &= \begin{bmatrix} f_{1x}(\hat{x}) \\ \vdots \\ f_{mx}(\hat{x}) \end{bmatrix} \\ &= \begin{bmatrix} \frac{\partial f_1}{\partial x_1}(\hat{x}) & \cdots & \frac{\partial f_1}{\partial x_n}(\hat{x}) \\ \vdots & & \vdots \\ \frac{\partial f_m}{\partial x_1}(\hat{x}) & \cdots & \frac{\partial f_m}{\partial x_n}(\hat{x}) \end{bmatrix} \end{aligned}$$

2.1.6

If $f : R^n \to R$ is twice differentiable, its second derivative at $\hat{x}$ is the $n \times n$ matrix $(\partial^2 f/\partial x \partial x)(\hat{x}) = f_{xx}(\hat{x})$ where $(f_{xx}(\hat{x}))_i^j = (\partial^2 f/\partial x_j \partial x_i)(\hat{x})$. Thus, in terms of the notation in Section 2.1.5 above, $f_{xx}(\hat{x}) = (\partial/\partial x)(f_x)'(\hat{x})$.

2.2 *Example*

We consider in detail the first example of Chapter 1. Define the following variables and functions:

$$\begin{aligned} \alpha &= \text{fraction of mary-john in proposed mixture,} \\ p &= \text{sale price per pound of mixture,} \\ v &= \text{total amount of mixture produced,} \\ f(\alpha, p) &= \text{expected sales volume (as determined by market research) of mixture as a function of} (\alpha, p). \end{aligned}$$

Since it is not profitable to produce more than can be sold we must have:

$$\begin{aligned} v &= f(\alpha, p), \\ m &= \text{amount (in pounds) of mary-john purchased, and} \\ t &= \text{amount (in pounds) of tobacco purchased.} \end{aligned}$$

Evidently,

$$\begin{aligned} m &= \alpha v, \text{and} \\ t &= (l - \alpha)v. \end{aligned}$$

Let

$$\begin{aligned} P_1(m) &= \text{purchase price of } m \text{ pounds of mary-john, and} \\ P_2 &= \text{purchase price per pound of tobacco.} \end{aligned}$$

Then the total cost as a function of α, p is

$$C(\alpha, p) = P_1(\alpha f(\alpha, p)) + P_2(1 - \alpha)f(\alpha, p).$$

The revenue is

$$R(\alpha, p) = pf(\alpha, p),$$

so that the net profit is

$$N(\alpha, p) = R(\alpha, p) - C(\alpha, p).$$

The set of admissible decisions is Ω, where $\Omega = \{(\alpha, p) | 0 < \alpha < \frac{1}{2}, 0 < p < \infty\}$. Formally, we have the the following decision problem:

$$\begin{array}{ll} \text{Maximize} & N(\alpha, p), \\ \text{subject to} & (\alpha, p) \in \Omega. \end{array}$$

Suppose that $(\alpha^*, p*)$ is an optimal decision, *i.e.,*

$$\begin{array}{lll} (\alpha^*, p^*) \in \Omega & & \text{and} \\ N(\alpha^*, p^*) \geq N(\alpha, p) & \text{for all} & (\alpha, p) \in \Omega. \end{array} \tag{2.1}$$

We are going to establish some properties of (a^*, p^*). First of all we note that Ω is an *open* subset of R^2. Hence there exits $\varepsilon > 0$ such that

$$(\alpha, p) \in \Omega \quad \text{w}\mathit{henever} \quad |(\alpha, p) - (\alpha^*, p^*)| < \varepsilon \tag{2.2}$$

In turn (2.2) implies that for every vector $h = (h_1, h_2)'$ in R^2 there exists $\eta > 0$ (η of course depends on h) such that

$$((\alpha^*, p^*) + \delta(h_1, h_2)) \in \Omega \quad \text{for} \quad 0 \leq \delta \leq \eta \tag{2.3}$$

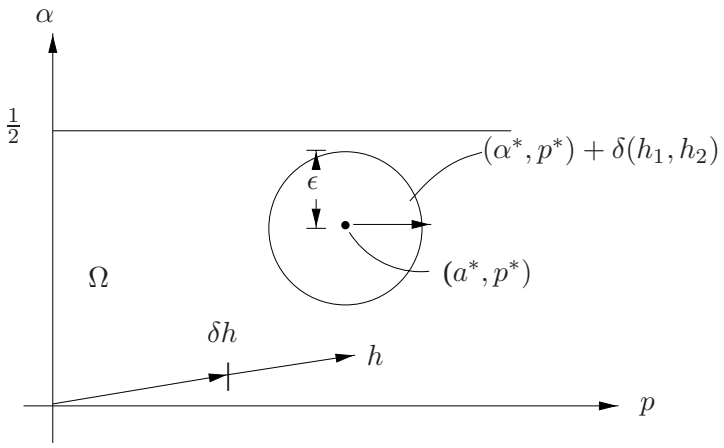


Figure 2.1: Admissable set of example.

Combining (2.3) with (2.1) we obtain (2.4):

$$N(\alpha^*, p^*) \geq N(\alpha^* + \delta h_1, p^* + \delta h_2) \quad \text{for} \quad 0 \leq \delta \leq \eta \tag{2.4}$$

Now we assume that the function N is *differentiable* so that by Taylor's theorem

$$\begin{array}{rl} & N(\alpha^*, p^*) \\ N(\alpha^* + \delta h_1, p^* + \delta h_2) = & +\delta[\frac{\partial N}{\partial \alpha}(\delta^*, p^*)h_1 + \frac{\partial N}{\partial p}(\alpha^*, p^*)h_2] \\ & +o(\delta), \end{array} \tag{2.5}$$

where

$$\tfrac{o\delta}{\delta} \to 0 \quad \text{as} \quad \delta \to 0. \tag{2.6}$$

Substitution of (2.5) into (2.4) yields

$$0 \geq \delta[\tfrac{\partial N}{\partial \alpha}(\alpha^*, p^*)h_1 + \tfrac{\partial N}{\partial p}(\alpha^*, p^*)h_2] + o(\delta).$$

Dividing by $\delta > 0$ gives

$$0 \geq [\tfrac{\partial N}{\partial \alpha}(\alpha^*, p^*)h_1 + \tfrac{\partial N}{\partial p}(\alpha^*, p^*)h_2] + \tfrac{o(\delta)}{\delta}. \tag{2.7}$$

Letting δ approach zero in (2.7), and using (2.6) we get

$$0 \geq [\tfrac{\partial N}{\partial \alpha}(\alpha^*, p^*)h_1 + \tfrac{\partial N}{\partial p}(\alpha^*, p^*)h_2]. \tag{2.8}$$

Thus, using the facts that N is differentiable, (α^*, p^*) is optimal, and δ is open, we have concluded that the inequality (2.9) holds for *every* vector $h \in R^2$. Clearly this is possible only if

$$\tfrac{\partial N}{\partial \alpha}(\alpha^*, p^*) = 0, \quad \tfrac{\partial N}{\partial p}(\alpha^*, p^*) = 0. \tag{2.9}$$

Before evaluating the usefulness of property (2.8), let us prove a direct generalization.

2.3 *The Main Result and its Consequences*

2.3.1 *Theorem*

.

Let Ω be an open subset of R^n. Let f: $R^n \to R$ be a differentiable function. Let x^* be an optimal solution of the following decision-making problem:

$$\begin{array}{ll} \text{Maximize} & f(x) \\ \text{subject to} & x \in \Omega. \end{array} \tag{2.10}$$

Then

$$\frac{\partial f}{\partial x}(x^*) = 0. \tag{2.11}$$

Proof: Since $x^* \in \Omega$ and Ω is open, there exists $\varepsilon > 0$ such that

$$x \in \Omega \quad \text{whenever } |x - x^*| < \varepsilon. \tag{2.12}$$

In turn, (2.12) implies that for every vector $h \in R^n$ there exits $\eta > 0$ (η depending on h) such that

$$(x^* + \delta h) \in \Omega \quad \text{whenever} \quad 0 \leq \delta \leq \eta. \tag{2.13}$$

Since x^* is optimal, we must then have

$$f(x^*) \geq f(x^* + \delta h) \quad \text{whenever} \quad 0 \leq \delta \leq \eta. \tag{2.14}$$

Since f is differentiable, by Taylor's theorem we have

$$f(x^* + \delta h) = f(x^*) + \frac{\partial f}{\partial x}(x^*)\delta h + o(\delta), \tag{2.15}$$

where

$$\frac{o(\delta)}{\delta} \to 0 \quad \text{as} \quad \delta \to 0 \tag{2.16}$$

Substitution of (2.15) into (2.14) yields

$$0 \geq \delta \frac{\partial f}{\partial x}(x^*)h + o(\delta)$$

and dividing by $\delta > 0$ gives

$$0 \geq \frac{\partial f}{\partial x}(x^*)h + \frac{o(\delta)}{\delta} \tag{2.17}$$

Letting δ approach zero in (2.17) and taking (2.16) into account, we see that

$$0 \geq \frac{\partial f}{\partial x}(x^*)h, \tag{2.18}$$

Since the inequality (2.18) must hold for every $h \in R^n$, we must have

$$0 = \frac{\partial f}{\partial x}(x^*),$$

and the theorem is proved. $\diamondsuit$

Table 2.1

Case	Does there exist an optimal decision for 2.2.1?	At how many points in Ω is 2.2.2 satisfied?	Further Consequences
1	Yes	Exactly one point, say x^*	x^* is the unique optimal
2	Yes	More than one point	
3	No	None	
4	No	Exactly one point	
5	No	More than one point	

2.3.2 Consequences.

Let us evaluate the usefulness of (2.11) and its special case (2.18). Equation (2.11) gives us n equations which must be satisfied at any optimal decision $x^* = (x_1^*, \ldots, x_n^*)'$.
These are

$$\frac{\partial f}{\partial x_1}(x^*) = 0, \quad \frac{\partial f}{\partial x_2}(x^*) = 0, \ldots, \quad \frac{\partial f}{\partial x_n}(x^*) = 0 \tag{2.19}$$

Thus, every optimal decision must be a solution of these n simultaneous equations of n variables, so that the search for an optimal decision from Ω is reduced to searching among the solutions of (2.19). In practice this may be a very difficult problem since these may be nonlinear equations and it may be necessary to use a digital computer. However, in these *Notes* we shall not be overly concerned with numerical solution techniques (but see 2.4.6 below).

The theorem may also have conceptual significance. We return to the example and recall the $N = R - C$. Suppose that R and C are differentiable, in which case (2.18) implies that at every optimal decision (α^*, p^*)

$$\frac{\partial R}{\partial \alpha}(\alpha^*, p^*) = \frac{\partial C}{\partial \alpha}(\alpha^*, p^*), \quad \frac{\partial R}{\partial p}(\alpha^*, p^*) = \frac{\partial C}{\partial p}(\alpha^*, p^*),$$

or, in the language of economic analysis, marginal revenue = marginal cost. We have obtained an important economic insight.

2.4 *Remarks and Extensions*

2.4.1 *A warning.*

Equation (2.11) is only a *necessary* condition for x^* to be optimal. There may exist decisions $\tilde{x} \in \Omega$ such that $f_x(\tilde{x}) = 0$ but $\tilde{x}$ is not optimal. More generally, any one of the five cases in Table 2.1 may occur. The diagrams in Figure 2.1 illustrate these cases. In each case $\Omega = (-1, 1)$.

Note that in the last three figures there is no optimal decision since the limit points -1 and +1 are not in the set of permissible decisions $\Omega = (-1, 1)$. In summary, the theorem does not give us any clues concerning the *existence* of an optimal decision, and it does not give us *sufficient* conditions either.

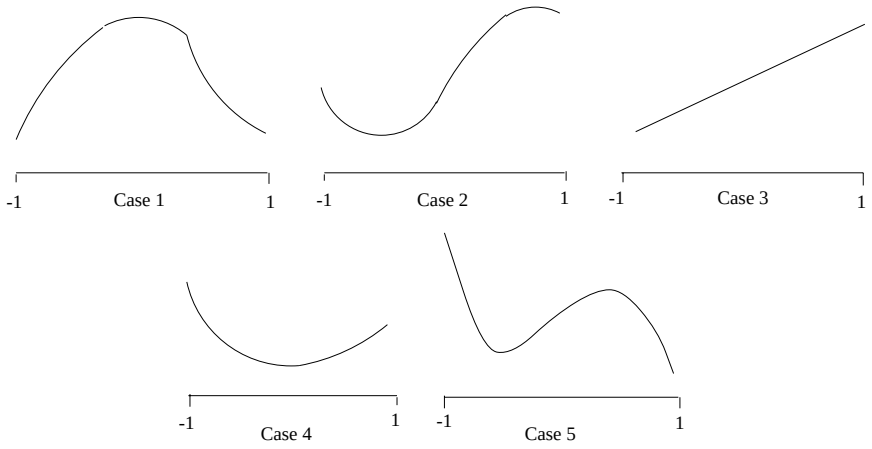


Figure 2.2: Illustration of 4.1.

2.4.2 *Existence.*

If the set of permissible decisions Ω is a closed and bounded subset of R^n, and if f is continuous, then it follows by the Weierstrass Theorem that there exists an optimal decision. But if Ω is *closed* we cannot assert that the derivative of f vanishes at the optimum. Indeed, in the third figure above, if $\Omega = [-1, 1]$, then +1 is the optimal decision but the derivative is positive at that point.

2.4.3 *Local optimum.*

We say that $x^* \in \Omega$ is a locally optimal decision if there exists $\varepsilon > 0$ such that $f(x^*) \geq f(x)$ whenever $x \in \Omega$ and $|x^* - x| \leq \varepsilon$. It is easy to see that the theorem holds *(i.e., 2.11)* for local optima also.

2.4.4 *Second-order conditions.*

Suppose f is twice-differentiable and let $x^* \in \Omega$ be optimal or even locally optimal. Then $f_x(x^*) = 0$, and by Taylor's theorem

$$f(x^* + \delta h) = f(x^*) + \tfrac{1}{2}\delta^2 h' f_{xx}(x^*)h + o(\delta^2), \tag{2.20}$$

where $\frac{o(\delta^2)}{\delta^2} \to 0$ as $\delta \to 0$. Now for $\delta > 0$ sufficiently small $f(x^* + \delta h) \leq f(x^*)$, so that dividing by $\delta^2 > 0$ yields

$$0 \geq \tfrac{1}{2} h' f_{xx}(x^*)h + \tfrac{o(\delta^2)}{\delta^2}$$

and letting δ approach zero we conclude that $h' f_{xx}(x^*)h \leq 0$ for all $h \in R^n$. This means that $f_{xx}(x^*)$ is a negative semi-definite matrix. Thus, if we have a twice differentiable objective function, we get an additional necessary condition.

2.4.5 *Sufficiency for local optimal.*

Suppose at $x^* \in \Omega$, $f_x(x^*) = 0$ and f_{xx} is strictly negative definite. But then from the expansion (2.20) we can conclude that x^* is a local optimum.

2.4.6 *A numerical procedure.*

At any point $\tilde{x} \in \Omega$ the gradient $\bigtriangledown_x f(\tilde{x})$ is a direction along which $f(x)$ increases, *i.e.*, $f(\tilde{x} + \varepsilon \bigtriangledown_x f(\tilde{x})) > f(\tilde{x})$ for all $\varepsilon > 0$ sufficiently small. This observation suggests the following scheme for finding a point $x^* \in \Omega$ which satisfies 2.11. We can formalize the scheme as an algorithm.

Step 1. Pick $x^0 \in \Omega$. Set $i = 0$. Go to Step 2.
Step 2. Calculate $\bigtriangledown_x f(x^i)$. If $\bigtriangledown_x f(x^i) = 0$, stop.
Otherwise let $x^{i+1} = x^i + d_i \bigtriangledown_x f(x^i)$ and go to Step 3.
Step 3. Set $i = i + 1$ and return to Step 2.

The step size d_i can be selected in many ways. For instance, one choice is to take d_i to be an optimal decision for the following problem:

$$\text{Max}\{f(x^i + d \bigtriangledown_x f(x^i)) | d > 0, (x^i + d \bigtriangledown_x f(x^i)) \in \Omega\}.$$

This requires a one-dimensional search. Another choice is to let $d_i = d_{i-1}$ if $f(x^i + d_{i-1} \bigtriangledown_x f(x^i)) > f(x^i)$; otherwise let $d_i = 1/k \; d_{i-1}$ where k is the smallest positive integer such that $f(x^i + 1/k \; d_{i-1} \bigtriangledown_x f(x^i)) > f(x^i)$. To start the process we let $d_{-1} > 0$ be arbitrary.

Exercise: Let f be continuous differentiable. Let $\{d_i\}$ be produced by either of these choices and let
$\{x_i\}$ be the resulting sequence. Then

1. $f(x_{i+1}) > f(x_i)$ if $x_{i+1} \neq x_i, i$
2. if $x^* \in \Omega$ is a limit point of the sequence $\{x_i\}$, $f_x(x^*) = 0$.

For other numerical procedures the reader is referred to Zangwill [1969] or Polak [1971].

Chapter 3

OPTIMIZATION OVER SETS DEFINED BY EQUALITY CONSTRAINTS

We first study a simple example and examine the properties of an optimal decision. This will generalize to a canonical problem, and the properties of its optimal decisions are stated in the form of a theorem. Additional properties are summarized in Section 3 and a numerical scheme is applied to determine the optimal design of resistive networks.

3.1 *Example*

We want to find the rectangle of maximum area inscribed in an ellipse defined by

$$f_1(x,y) = \frac{x^2}{a^2} + \frac{y^2}{b^2} = \alpha. \tag{3.1}$$

The problem can be formalized as follows (see Figure 3.1):

$$\begin{array}{lll} \text{Maximize} & f_0(x,y) & = 4xy \\ \text{subject to} & (x,y) \in \Omega & = \{(x,y) | f_1(x,y) = \alpha\}. \end{array} \tag{3.2}$$

The main difference between problem (3.2) and the decisions studied in the last chapter is that the set of permissible decisions Ω is *not* an open set. Hence, if (x^*, y^*) is an optimal decision we *cannot* assert that $f_0(x^*, y^*) \geq f_0(x, y)$ for all (x, y) in an open set containing (x^*, y^*). Returning to problem (3.2), suppose (x^*, y^*) is an optimal decision. Clearly then either $x^* \neq 0$ or $y^* \neq 0$. Let us suppose $y^* \neq 0$. Then from figure 3.1 it is evident that there exist (i)$\varepsilon > 0$, (ii) an open set V containing (x^*, y^*), and (iii) a differentiable function $g : (x^* - \varepsilon, x^* + \varepsilon) \to V$ such that

$$f_1(x,y) = \alpha \quad \text{and} \quad (x,y) \in V \quad \text{iff} \quad fy = g(x).^{1} \tag{3.3}$$

In particular this implies that $y^* = g(x^*)$, and that $f_1(x, g(x)) = \alpha$ whenever $|x - x^*| < \varepsilon$. Since

[1]Note that $y^* \neq 0$ implies $f_{1y}(x^*, Y^*) \neq 0$, so that this assertion follows from the Implicit Function Theorem. The assertion is false if $y^* = 0$. In the present case let $0 < \varepsilon \leq a - x^*$ and $g(x) = +b[\alpha - (x/a)^2]^{1/2}$.

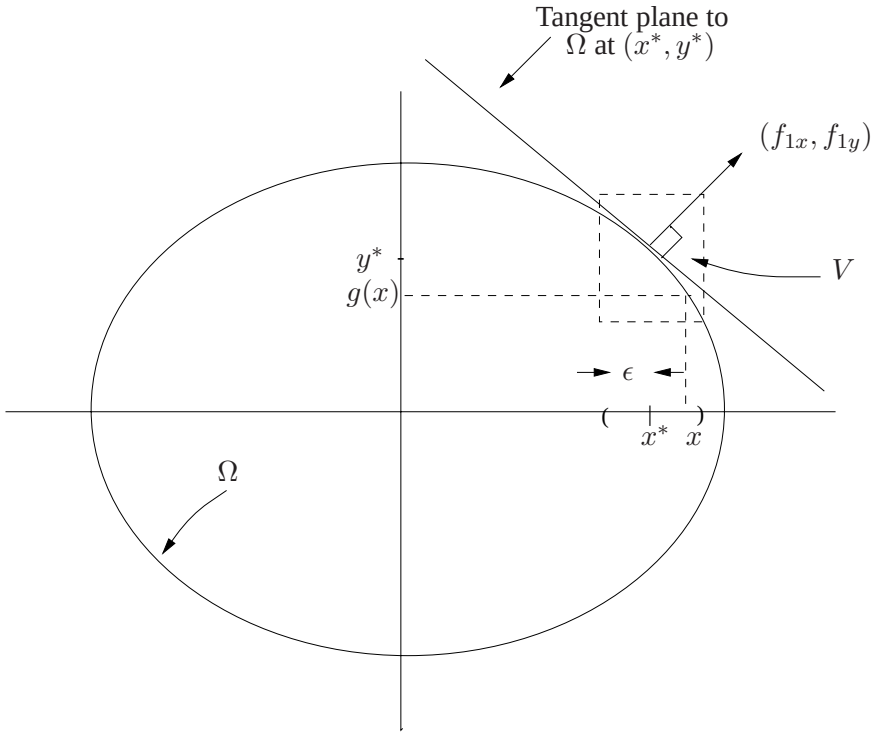


Figure 3.1: Illustration of example.

$(x^*, y^*) = (x^*, g(x^*))$ is optimum for (3.2), it follows that x^* is an optimal solution for (3.4):

$$\begin{array}{ll} \text{Maximize} & \hat{f}_0(x) = f_0(x, g(x)) \\ \text{subject to} & |x - x^*| < \varepsilon. \end{array} \tag{3.4}$$

But the constraint set in (3.4) is an open set (in R^1) and the objective function $\hat{f}_0$ is differentiable, so that by Theorem 2.3.1, $\hat{f}_{0x}(x^*) = 0$, which we can also express as

$$f_{0x}(x^*, y^*) + f_{0y}(x^*, y^*) g_x(x^*) = 0 \tag{3.5}$$

Using the fact that $f_1(x, g(x)) \equiv \alpha$ for $|x - x^*| < \varepsilon$, we see that

$$f_{1x}(x^*, y^*) + f_{1y}(x^*, y^*) g_x(x^*) = 0,$$

and since $f_{1y}(x^*, y^*) \neq 0$ we can evaluate $g_x(x^*)$,

$$g_x(x^*) = -f_{1y}^{-1} f_{1x}(x^*, y^*),$$

and substitute in (3.5) to obtain the condition (3.6):

$$f_{0x} - f_{0y} f_{1y}^{-1} f_{1x} = 0 \ \text{ at } \ (x^*, y^*). \tag{3.6}$$

Thus an optimal decision (x^*, y^*) must satisfy the two equations $f_1(x^*, y^*) = \alpha$ and (3.6). Solving these yields

$$x^* = \ {}^{+}_{-}(\alpha/2)^{1/2} a \ , \ y^* = \ {}^{+}_{-}(\alpha/2)^{1/2} b.$$

Evidently there are two optimal decisions, $(x^*, y^*) = \pm(\alpha/2)^{1/2}(a, b)$, and the maximum area is

$$m(\alpha) = 2\alpha ab. \tag{3.7}$$

The condition (3.6) can be interpreted differently. Define

$$\lambda^* = f_{0y} f_{1y}^{-1}(x^*, y^*). \tag{3.8}$$

Then (3.6) and (3.8) can be rewritten as (3.9):

$$(f_{0x}, f_{0y}) = \lambda^*(f_{1x}, f_{1y}) \quad \text{at} \quad (x^*, y^*) \tag{3.9}$$

In terms of the gradients of f_0, f_1, (3.9) is equivalent to

$$\bigtriangledown f_0(x^*, y^*) = [\bigtriangledown f_1(x^*, y^*)]\lambda^*, \tag{3.10}$$

which means that at an optimal decision the gradient of the objective function f_0 is normal to the plane tangent to the constraint set Ω.

Finally we note that

$$\lambda^* = \frac{\partial m}{\partial \alpha}. \tag{3.11}$$

where $m(\alpha) =$ maximum area.

3.2 *General Case*

3.2.1 *Theorem.*

Let $f_i : R^n \to R, i = 0, 1, \ldots, m \ \ (m < n)$, be continuously differentiable functions and let x^* be an optimal decision of problem (3.12):

$$\begin{array}{ll} \text{Maximize} & f_0(x) \\ \text{subject to} & f_i(x) = \alpha_i, \quad i = 1, \ldots, m. \end{array} \tag{3.12}$$

Suppose that at x^* the derivatives $f_{ix}(x^*), i = 1, \ldots, m$, are *linearly independent.* Then there exists a vector $\lambda^* = (\lambda_1^*, \ldots, \lambda_m^*)'$ such that

$$f_{0x}(x^*) = \lambda_1^* f_{1x}(x^*) + \ldots + \lambda_m^* f_{mx}(x^*) \tag{3.13}$$

Furthermore, let $m(\alpha_1, \ldots, \alpha_m)$ be the maximum value of (3.12) as a function of $\alpha = (\alpha_1, \ldots, \alpha_m)'$. Let $x^*(\alpha)$ be an optimal decision for (3.12). If $x^*(\alpha)$ is a *differentiable* function of α then $m(\alpha)$ is a differentiable function of α, and

$$(\lambda^*)' = \frac{\partial m}{\partial \alpha} \tag{3.14}$$

Proof. Since $f_{ix}(x^*), i = 1, \ldots, m$, are linearly independent, then by re-labeling the coordinates of x if necessary, we can assume that the $m \times m$ matrix $[(\partial f_i/\partial x_j)(x^*)], 1 \le i, j \le m$, is nonsingular. By the Implicit Function Theorem (see Fleming [1965]) it follows that there exist (i) $\varepsilon > 0$, (ii) an

open set V in R^n containing x^*, and (iii) a differentiable function $g : U \to R^m$, where $U = [(x_{m+1}, \ldots, x_n)] \mid |x_{m+\ell} - x^*_{m+\ell}| < \varepsilon, \ell = 1, \ldots, n-m]$, such that

$$f_i(x_1, \ldots, x_n) = \alpha_i, 1 \leq i \leq m, \quad \text{and} \quad (x_1, \ldots, x_n) \in V$$

iff

$$x_j = g_j(x_{m+1}, \ldots, x_n), 1 \leq j \leq m, \quad \text{and} \quad (x_{m+1}, \ldots, x_n) \in U \tag{3.15}$$

(see Figure 3.2).

In particular this implies that $x^*_j = g_j(x^*_{m+1}, \ldots, x^*_n), 1 \leq j \leq m$, and

$$f_i(g(x_{m+1}, \ldots, x_n), x_{m+1}, \ldots, x_n) = \alpha_i \quad , \quad i = 1, \ldots, m. \tag{3.16}$$

For convenience, let us define $w = (x_1, \ldots, x_m)'$, $u = (x_{m+1}, \ldots, x_n)'$ and $f = (f_1, \ldots, f_m)'$. Then, since $x^* = (w^*, u^*) = (g(u^*), u^*)$ is optimal for (3.12), it follows that u^* is an optimal decision for (3.17):

$$\begin{array}{ll} \text{Maximize} & \hat{f}_0(u) = f_0(g(u), u) \\ \text{subject to} & u \in U. \end{array} \tag{3.17}$$

But U is an open subset of R^{n-m} and $\hat{f}_0$ is a differentiable function on U (since f_0 and g are differentiable), so that by Theorem 2.3.1 , $\hat{f}_{0u}(u^*) = 0$, which we can also express using the chain rule for derivatives as

$$\hat{f}_{0u}(u^*) = f_{0w}(x^*)g_u(u^*) + f_{0u}(x^*) = 0. \tag{3.18}$$

Differentiating (3.16) with respect to $u = (x_{m+1}, \ldots, x_n)'$, we see that

$$f_w(x^*)g_u(u^*) + f_u(x^*) = 0,$$

and since the $m \times m$ matrix $f_w(x^*)$ is nonsingular we can evaluate $g_u(u^*)$,

$$g_u(u^*) = -[f_w(x*)]^{-1} f_u(x^*),$$

and substitute in (3.18) to obtain the condition

$$-f_{0w} f_w^{-1} f_u + f_{0u} = 0 \quad \text{at} \quad x^* = (w^*, u^*). \tag{3.19}$$

Next, define the m-dimensional column vector λ^* by

$$(\lambda^*)' = f_{0w} f_w^{-1} | x^*. \tag{3.20}$$

Then (3.19) and (3.20) can be written as (3.21):

$$(f_{0w}(x^*), f_{0u}(x^*)) = (\lambda^*)'(f_w(x^*), f_u(x^*)). \tag{3.21}$$

Since $x = (w, u)$, this is the same as

$$f_{0x}(x^*) = (\lambda^*)' f_x(x^*) = \lambda^*_1 f_{1x}(x^*) + \ldots + \lambda^*_m f_{mx}(x^*),$$

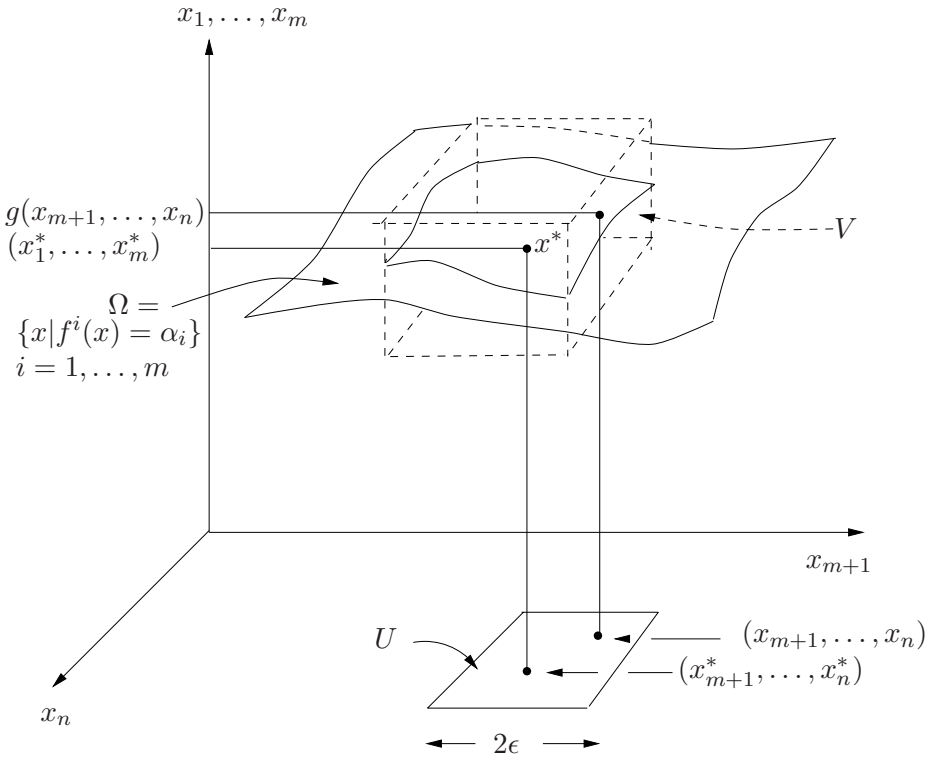


Figure 3.2: Illustration of theorem.

which is equation (3.13).

To prove (3.14), we vary α in a neighborhood of a fixed value, say $\underline{\alpha}$. We define $w^*(\alpha) = (x_1^*(\alpha), \ldots, x_m^*(\alpha))'$ and $u^*(\alpha) = (x_{m+1}^*(\alpha), \ldots, x_(^*\alpha))'$. By hypothesis, f_w is nonsingular at $x^*(\underline{\alpha})$. Since $f(x)$ and $x^*(\alpha)$ are continuously differentiable by hypothesis, it follows that f_w is nonsingular at $x^*(\alpha)$ in a neighborhood of $\underline{\alpha}$, say N. We have the equation

$$f(w^*(\alpha), u^*(\alpha)) = \alpha, \tag{3.22}$$

$$-f_{0w} f_w^{-1} f_u + f_{0u} = 0 \quad \text{at} \quad (w^*(\alpha), u^*(\alpha)), \tag{3.23}$$

for $\alpha \in N$. Also, $m(\alpha) = f_0(x^*(\alpha))$, so that

$$m_\alpha = f_{0w} w_\alpha^* + f_{0u} u_\alpha^* \tag{3.24}$$

Differentiating (3.22) with respect to α gives

$$f_w w_\alpha^* + f_u u_\alpha^* = I,$$

so that

$$w_\alpha^* + f_w^{-1} f_u u_\alpha^* = f_w^{-1},$$

and multiplying on the left by f_{0w} gives

$$f_{0w}w^*_\alpha + f_{0w}f_w^{-1}f_u u^*_\alpha = f_{0w}f_w^{-1}.$$

Using (3.23), this equation can be rewritten as

$$f_{0w}w^*_\alpha + f_{0u}u^*_\alpha = f_{0w}f_w^{-1}. \tag{3.25}$$

In (3.25), if we substitute from (3.20) and (3.24), we obtain (3.14) and the theorem is proved. $\diamondsuit$

3.2.2 *Geometric interpretation.*

The equality constraints of the problem in 3.12 define a $n-m$ dimensional surface

$$\Omega = \{x | f_i(x) = \alpha_i, i = 1, \ldots, m\}.$$

The hypothesis of linear independence of $\{f_{ix}(x^*)|1 \leq i \leq m\}$ guarantees that the tangent plane through Ω at x^* is described by

$$\{h | f_{ix}(x^*)h = 0 \quad , \quad i = 1, \ldots, m\}, \tag{3.26}$$

so that the set of (column vectors orthogonal to this tangent surface is

$$\{\lambda_1 \bigtriangledown_x f_1(x^*) + \ldots + \lambda_m \bigtriangledown_x f_m(x^*) | \lambda_i \in R, i = 1, \ldots, m\}.$$

Condition (3.13) is therefore equivalent to saying that at an optimal decision x^*, the gradient of the objective function $\bigtriangledown_x f_0(x^*)$ is normal to the tangent surface (3.12).

3.2.3 *Algebraic interpretation.*

Let us again define $w = (x_1, \ldots, x_m)'$ and $u = (x_{m+1}, \ldots, x_n)'$. Suppose that $f_w(\tilde{x})$ is nonsingular at some point $\tilde{x} = (\tilde{w}, \tilde{u})$ in Ω which is not necessarily optimal. Then the Implicit Function Theorem enables us to solve, in a neighborhood of $\tilde{x}$, the m equations $f(w, u) = \alpha$. u can then vary *arbitrarily* in a neighborhood of $\tilde{u}$. As u varies, w must change according to $w = g(u)$ (in order to maintain $f(w, u) = \alpha$), and the objective function changes according to $\hat{f}_0(u) = f_0(g(u), u)$. The derivative of $\hat{f}_0$ at $\tilde{u}$ is

$$\hat{f}_{0u}(\tilde{u}) = f_{0w}g_u + f_{0u\tilde{x}} = -\tilde{\lambda}' f_u(\tilde{x}) + f_{0u}(\tilde{x}),$$

where

$$\tilde{\lambda}' = f_{0w}f_{w\tilde{x}}^{-1}, \tag{3.27}$$

Therefore, *the direction of steepest increase of* $\hat{f}_0$ *at* $\tilde{u}$ *is*

$$\bigtriangledown_u \hat{f}_0(\tilde{u}) = -f'_u(\tilde{x})\tilde{\lambda} + f'_{Ou}(\tilde{x}) \quad , \tag{3.28}$$

and if $\tilde{u}$ is optimal, $\bigtriangledown_u \hat{f}_0(\tilde{u}) = 0$ which, together with (3.27) is equation (3.13). We shall use (3.27) and (3.28) in the last section.

3.3 *Remarks and Extensions*

3.3.1 *The condition of linear independence.*

The necessary condition (3.13) need not hold if the derivatives $f_{ix}(x^*), 1 \leq i \leq m$, are not linearly independent. This can be checked in the following example

$$\begin{array}{l} \text{Minimize} \\ \text{subject to } \sin(x_1^2 + x_2^2) \\ \frac{\pi}{2}(x_1^2 + x_2^2) = 1. \end{array} \tag{3.29}$$

3.3.2 *An alternative condition.*

Keeping the notation of Theorem 3.2.1, define the *Lagrangian function* $L : R^{n+m} \to R$ by $L : (x, \lambda) \mapsto f_0(x) - \sum_{i=1}^m \lambda_i f_i(x)$. The following is a reformulation of 3.12, and its proof is left as an exercise.

Let x^* be optimal for (3.12), and suppose that $f_{ix}(x^*), 1 \leq i \leq m$, are linearly independent. Then there exists $\lambda^* \in R^m$ such that (x^*, λ^*) is a *stationary point* of L, *i.e.*, $L_x(x^*, \lambda^*) = 0$ and $L_\lambda(x^*, \lambda^*) = 0$.

3.3.3 *Second-order conditions.*

Since we can convert the problem (3.12) into a problem of maximizing $\hat{f}_0$ over an open set, all the comments of Section 2.4 will apply to the function $\hat{f}_0$. However, it is useful to translate these remarks in terms of the original function f_0 and f. This is possible because the function g is uniquely specified by (3.16) in a neighborhood of x^*. Furthermore, if f is twice differentiable, so is g (see Fleming [1965]). It follows that if the functions f_i, $0 \leq i \leq m$, are twice continuously differentiable, then so is $\hat{f}_0$, and a necessary condition for x^* to be optimal for (3.12) and (3.13) and the condition that the $(n-m) \times (n-m)$ matrix $\hat{f}_{0uu}(u^*)$ is negative semi-definite. Furthermore, if this matrix is negative definite then x^* is a local optimum. the following exercise expresses $f\hat{f}_{0uu}(u^*)$ in terms of derivatives of the functions f_i.

Exercise: Show that

$$\hat{f}_{0uu}(u^*) = [g_u' \vdots I] \begin{bmatrix} L_{ww} & L_{wu} \\ L_{uw} & L_{uu} \end{bmatrix} \left[\begin{array}{c} g_u \\ \cdots \\ I \end{array} \right] \Bigg|_{(w^*, u^*)}$$

where

$$g_u(u^*) = -[f_w(x^*)]^{-1} f_u(x^*), L(x) = f_0(x) - \sum_{i=1}^m \lambda_i^* f_i(x).$$

3.3.4 *A numerical procedure.*

We assume that the derivatives $f_{ix}(x), 1 \leq i \leq m$, are linearly independent for all x. Then the following algorithm is a straightforward adaptation of the procedure in Section 2.4.6.

Step 1. Find x_0 arbitrary so that $f_i(x^0) = \alpha_i, 1 \leq i \leq m$. Set $k = 0$ and go to Step 2.
Step 2. Find a partition $x = (w, u)$[2] of the variables such that $f_w(x^k)$ is nonsingular. Calculate λ^k by $(\lambda^k)' = f_{0w} f^{-1}_{w(xk)}$, and $\bigtriangledown \hat{f}^k_0(u^k) = -f'_u(x^k)\lambda^k + f'_{0u}(x^k)$. If $\bigtriangledown \hat{f}^k_0(u^k) = 0$, stop. Otherwise go to Step 3.
Step 3. Set $\tilde{u}^k = u^k + d_k \bigtriangledown \hat{f}^k_0(u^k)$. Find $\tilde{w}^k$ such that $f_i(\tilde{w}^k, \tilde{u}^k) = 0, 1 \leq i \leq m$. Set $x^{k+1} = (\tilde{w}^k, \tilde{u}^k)$, set $k = k + 1$, and return to Step 2.
Remarks. As before, the step sizes $d_k > 0$ can be selected various ways. The practical applicability of the algorithm depends upon two crucial factors: the ease with which we can find a partition $x = (w, u)$ so that $f_w(x^k)$ is nonsingular, thus enabling us to calculate λ^k; and the ease with which we can find $\tilde{w}^k$ so that $f(\tilde{w}^k, \tilde{u}^k) = \alpha$. In the next section we apply this algorithm to a practical problem where these two steps can be carried out without too much difficulty.

3.3.5 *Design of resistive networks.*

Consider a network N with $n + 1$ nodes and b branches. We choose one of the nodes as datum and denote by $e = (e_1, \ldots, e_n)'$ the vector of node-to-datum voltages. Orient the network graph and let $v = (v_1, \ldots, v_b)'$ and $j = (j_1, \ldots, j_b)'$ respectively, denote the vectors of branch voltages and branch currents. Let A be the $n \times b$ reduced incidence matrix of the network graph. Then the Kirchhoff current and voltage laws respectively yield the equations

$$Aj = 0 \quad \text{and} \quad A'e = v \tag{3.30}$$

Next we suppose that each branch k contains a (possibly nonlinear)resistive element with the form shown in Figure 3.3, so that

$$j_k - j_{sk} = g_k(v_{rk}) = g_k(v_k - v_{sk}), 1 \leq k \leq b, \tag{3.31}$$

where v_{rk} is the voltage across the resistor. Here j_{sk}, v_{sk} are the source current and voltage in the *kth* branch, and g_k is the characteristic of the resistor. Using the obvious vector notation $j_s \in R^b$, $v_s \in R^b$ for the sources, $v_r \in R^b$ for the resistor voltages, and $g = (g_1, \ldots, g_b)'$, we can rewrite (3.30) as (3.31):

$$j - j_s = g(v - v_s) = g(v_r). \tag{3.32}$$

Although (3.30) implies that the current $(j_k - j_s k)$ through the *kth* resistor depends only on the voltage $v_{rk} = (v_k - v_{sk})$ across itself, no essential simplification is achieved. Hence, in (3.31) we shall assume that g_k is a function of v_r. This allows us to include coupled resistors and voltage-controlled current sources. Furthermore, let us suppose that there are ℓ design parameters $p = (p_1, \ldots, p_\ell)'$ which are under our control, so that (3.31) is replaced by (3.32):

$$j - j_x = g(v_r, p) = g(v - v_s, p). \tag{3.33}$$

[2]This is just a notational convenience. The w variable may consist of any m components of x.

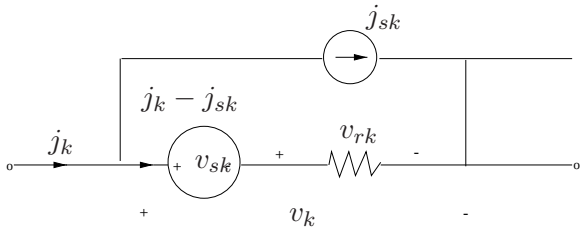


Figure 3.3: The *kth branch.*

If we combine (3.29) and (3.32) we obtain (3.33):

$$Ag(A'e - v_s, p) = i_s, \tag{3.34}$$

where we have defined $i_s = A_{j_s}$. The network design problem can then be stated as finding p, v_s, i_s so as to minimize some specified function $f_0(e, p, v_s, i_s)$. Formally, we have the optimization problem (3.34):

$$\begin{array}{ll} \text{Minimize} & f_0(e, p, v_s, i_s) \\ \text{subject to} & Ag(A'e - v_s, p) - i_s = 0. \end{array} \tag{3.35}$$

We shall apply the algorithm 3.3.4 to this problem. To do this we make the following assumption.

Assumption: (a) f_0 is differentiable. (b) g is differentiable and the $n \times n$ matrix $A(\partial g/\partial v)(v, p)A'$ is nonsingular for all $v \in R^b, p \in R^\ell$. (c) The network N described by (3.33) is determinate ***i.e.***, for every value of (p, v_s, i_s) there is a unique $e = E(p, v_s, i_s)$ satisfying (3.33).

In terms of the notation of 3.3.4, if we let $x = (e, p, v_s, i_s)$, then assumption (b) allows us to identify $w = e$, and $u = (p, v_s, i_s)$. Also let $f(x) = f(e, p, v_s, i_s) = Ag(A'e - v_s, p) - i_s$. Now the crucial part in the algorithm is to obtain λ^k at some point x^k. To this end let $\tilde{x} = (\tilde{e}, \tilde{p}, \tilde{v}_s, \tilde{i}_s)$ be a fixed point. Then the corresponding $\lambda = \tilde{\lambda}$ is given by (see (3.27))

$$\tilde{\lambda}' = f_{0w}(\tilde{x}) f_w^{-1}(\tilde{x}) = f_{0e}(\tilde{x}) f_e^{-1}(\tilde{x}). \tag{3.36}$$

From the definition of f we have

$$f_e(\tilde{x}) = AG(\tilde{v}_r, \tilde{p})A',$$

where $\tilde{v}_r = A'\tilde{e} - \tilde{v}_s$, and $G(\tilde{v}_r, \tilde{p}) = (\partial g/\partial v_r)(\tilde{v}_r, \tilde{p})$. Therefore, $\tilde{\lambda}$ is the solution (unique by assumption (b)) of the following linear equation:

$$AG'(\tilde{v}_r, \tilde{p})A'\tilde{\lambda} = f'_{0e}(\tilde{x}). \tag{3.37}$$

Now (3.36) has the following extremely interesting physical interpretation. If we compare (3.33) with (3.36) we see immediately that $\tilde{\lambda}$ is the node-to-datum response voltages of a *linear* network $N(\tilde{v}_r, \tilde{p})$ driven by the current sources $f'_{0e}(\tilde{x})$. Furthermore, this network has the *same* graph as the original network (since they have the same incidence matrix); moreover, its branch admittance matrix, $G'(\tilde{v}_r, \tilde{p})$, is the transpose of the incremental branch admittance matrix (evaluated at $(\tilde{v}_r, \tilde{p})$) of the original network N. For this reason, $N(\tilde{v}_r, \tilde{p})$ is called the *adjoint network* (of N) at $(\tilde{v}_r, \tilde{p})$.

Once we have obtained $\tilde{\lambda}$ we can obtain $\bigtriangledown_u \hat{f}_0(\tilde{u})$ using (3.28). Elementary calculations yield (3.37):

$$\bigtriangledown_u \hat{f}_0(\tilde{u}) \;=\; \begin{bmatrix} \hat{f}'_{0p}(\tilde{u}) \\ \hat{f}'_{0v_s}(\tilde{u}) \\ \hat{f}'_{0i_s}(\tilde{u}) \end{bmatrix} \;=\; \begin{bmatrix} [\frac{\partial g}{\partial p}(\tilde{v}_r, \tilde{p})]' A' \\ G'(\tilde{v}_r, \tilde{p}) A' \\ -I \end{bmatrix} \tilde{\lambda} \;+\; \begin{bmatrix} f'_{0p}(\tilde{x}) \\ f'_{0v_s}(\tilde{x}) \\ f'_{0i_s}(\tilde{x}) \end{bmatrix} \tag{3.38}$$

We can now state the algorithm.

Step 1. Select $u^0 = (p^0, v_s^0, i_s^0)$ arbitrary. Solve (3.33) to obtain $e^0 = E(p^0, v_s^0, i_s^0)$. Let $k = 0$ and go to Step 2.
Step 2. Calculate $v_r^k = A' e^k - v_s^k$. calculate $f'_{0e}(x^k)$. Calculate the node-to-datum response λ^k of the adjoint network $N(v_r^k, p^k)$ driven by the current source $f'_{0e}(x^k)$. Calculate $\bigtriangledown_u \hat{f}_0(u^k)$ from (3.37). If this gradient is zero, stop. Otherwise go to Step 3.
Step 3. Let $u^{k+1} = (p^{k+1}, v_s^{k+1}, i_s^{k+1}) = u^k - d_k \bigtriangledown_u \hat{f}_0(u^k)$, where $d_k > 0$ is a predetermined step size. [3] Solve (3.33) to obtain $e^{k+1} = (Ep^{k+1}, v_s^{k+1}, i_s^{k+1})$. Set $k = k + 1$ and return to Step 2.
Remark 1. Each iteration from u^k to u^{k+1} requires one linear network analysis step (the computation of λ^k in Step 2), and one nonlinear network analysis step (the computation of e^{k+1} in step 3). This latter step may be very complex.
Remark 2. In practice we can control only some of the components of v_s and i_s, the rest being fixed. The only change this requires in the algorithm is that in Step 3 we set $p^{k+1} = p^k - d_k \hat{f}'_{0p}(u^k)$ just as before, where as $v_{sj}^{k+1} = v_{sj}^k - d_k(\partial \hat{f}_0 / \partial v_{sj})(u^k)$ and $i_{sm}^{k+1} = i_{sm}^k - d_k(\partial \hat{f}_0 / \partial i_{sm})(u^k)$ with j and m ranging only over the controllable components and the rest of the components equal to their specified values.
Remark 3. The interpretation of λ as the response of the adjoint network has been exploited for particular function f_0 in a series of papers (director and Rohrer [1969a], [1969b], [1969c]). Their derivation of the adjoint network does not appear as transparent as the one given here. Although we have used the incidence matrix A to obtain our network equation (3.33), one can use a more general cutset matrix. Similarly, more general representations of the resistive elements may be employed. In every case the "adjoint" network arises from a network interpretation of (3.27),

$$[f_w(\tilde{x})]' \tilde{\lambda} = f_{0w}(\tilde{x}),$$

with the transpose of the matrix giving rise to the adjective "adjoint."
Exercise: [DC biasing of transistor circuits (see Dowell and Rohrer [1971]).] Let N be a transistor circuit, and let (3.33) model the dc behavior of this circuit. Suppose that i_s is fixed, v_{sj} for $j \in J$ are variable, and v_{sj} for $j \notin J$ are fixed. For each choice of v_{sj}, $j \in J$, we obtain the vector e and hence the branch voltage vector $v = A'e$. Some of the components $v_t,\ t \in T$, will correspond to bias voltages for the transistors in the network, and we wish to choose $v_{sj},\ j \in J$, so that v_t is as close as possible to a desired bias voltage $v_t^d, t \in T$. If we choose nonnegative numbers α_t, with relative magnitudes reflecting the importance of the different transistors then we can formulate the criterion

[3] Note the minus sign in the expression $u^k - d_k \bigtriangledown_u \hat{f}_0(u^k)$. Remember we are minimizing f_0, which is equivalent to maximizing $(-f_0)$.

$$f_0(e) = \sum_{t \in T} \alpha_t |v_t - v_t^d|^2.$$

(i) Specialize the algorithm above for this particular case.
(ii) How do the formulas change if the network equations are written using an arbitrary cutset matrix instead of the incidence matrix?

Chapter 4

OPTIMIZATION OVER SETS DEFINED BY INEQUALITY CONSTRAINTS: LINEAR PROGRAMMING

In the first section we study in detail Example 2 of Chapter I, and then we define the general linear programming problem. In the second section we present the duality theory for linear programming and use it to obtain some sensitivity results. In Section 3 we present the Simplex algorithm which is the main procedure used to solve linear programming problems. In section 4 we apply the results of Sections 2 and 3 to study the linear programming theory of competitive economy. Additional miscellaneous comments are collected in the last section. For a detailed and readily accessible treatment of the material presented in this chapter see the companion volume in this Series (Sakarovitch [1971]).

4.1 *The Linear Programming Problem*

4.1.1 *Example.*

Recall Example 2 of Chapter I. Let g and u respectively be the number of graduate and undergraduate students admitted. Then the number of seminars demanded per year is $\frac{2g+u}{20}$, and the number of lecture courses demanded per year is $\frac{5g+7u}{40}$. On the supply side of our accounting, the faculty can offer $2(750) + 3(250) = 2250$ seminars and $6(750) + 3(250) = 5250$ lecture courses. Because of his contractual agreements, the President must satisfy

$$\frac{2g+u}{20} \leq 2250 \text{ or } 2g + u \leq 45,000$$

and

$$\frac{5g+7u}{40} \leq 5250 \text{ or } 5g + 7u \leq 210,000 \ .$$

Since negative g or u is meaningless, there are also the constraints $g \geq 0,\ u \geq 0$. Formally then the President faces the following decision problem:

$$\begin{array}{ll} \text{Maximize} & \alpha g + \beta u \\ \text{subject to} & 2g + u \leq 45,000 \\ & 5g + 7u \leq 210,000 \\ & g \geq 0, u \geq 0\ . \end{array} \tag{4.1}$$

It is convenient to use a more general notation. So let $x = (g, u)',\ c = (\alpha, \beta)',\ b = (45000, 210000, 0, 0)'$ and let A be the 4×2 matrix

$$A = \begin{bmatrix} 2 & 1 \\ 5 & 7 \\ -1 & 0 \\ 0 & -1 \end{bmatrix}.$$

Then (4.1) can be rewritten as (4.2)[1]

$$\begin{array}{ll} \text{Maximize} & c'x \\ \text{subject to} & Ax \leq b\ . \end{array} \tag{4.2}$$

Let $A_i, 1 \leq i \leq 4$, denote the *rows* of A. Then the set Ω of all vectors x which satisfy the constraints in (4.2) is given by $\Omega = \{x | A_i x \leq b_i,\ 1 \leq i \leq 4\}$ and is the polygon $OPQR$ in Figure 4.1.

For each choice x, the President receives the payoff $c'x$. Therefore, the surface of constant payoff k say, is the hyperplane $\pi(k) = \{x | c'x = k\}$. These hyperplanes for different values of k are parallel to one another since they have the same normal c. Furthermore, as k increases $\pi(k)$ moves in the direction c. (Obviously we are assuming in this discussion that $c \neq 0$.) Evidently an optimal decision is any point $x^* \in \Omega$ which lies on a hyperplane $\pi(k)$ which is farthest along the direction c. We can rephrase this by saying that $x* \in \Omega$ is an optimal decision if and only if the plane π^* through x^* does not intersect the interior of Ω, and futhermore at x^* the direction c points away from Ω. From this condition we can immediately draw two very important conclusions: (i) at least one of the vertices of Ω is an optimal decision, and (ii) x^* yields a higher payoff than all points in the cone K^* consisting of all rays starting at x^* and passing through Ω, since K^* lies "below" π^*. The first conclusion is the foundation of the powerful Simplex algorithm which we present in Section 3. Here we pursue consequences of the second conclusion. For the situation depicted in Figure 4.1 we can see that $x^* = Q$ is an optimal decision and the cone K^* is shown in Figure 4.2. Now x^* satisfies $A_x x^* = b_1,\ A_2 x^* = b_2$, and $A_3 x^* < b_3,\ A_4 x^* < b_4$, so that K^* is given by

$$K^* = \{x^* + h | A_1 h \leq 0\ , A_2 h \leq 0\}\ .$$

Since $c'x^* \geq c'y$ for all $y \in K^*$ we conclude that

$$c'h \leq 0 \ \text{ for all } \ h \ \text{ such that } \ A_1 h \leq 0,\ A_2 h \leq 0\ . \tag{4.3}$$

We pause to formulate the generalization of (4.3) as an exercise.

[1] Recall the notation introduced in 1.1.2, so that $x \leq y$ means $x_i \leq y_i$ for all i.

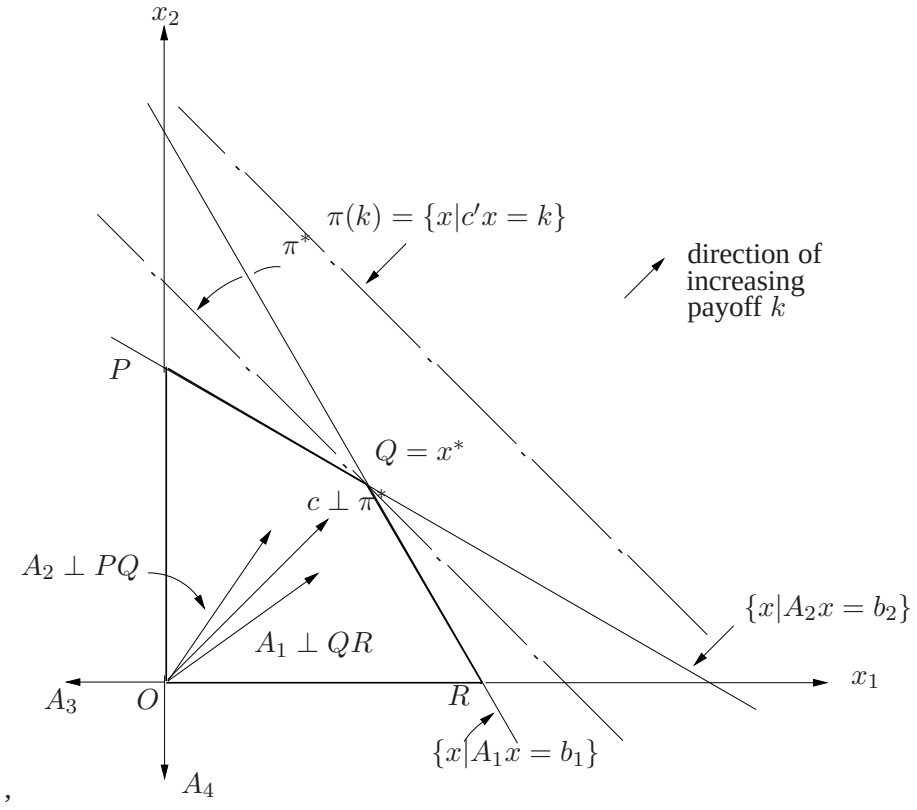


Figure 4.1: $\Omega = OPQR$.

Exercise 1: Let $A_i,\ 1 \leq i \leq k$, be n-dimensional *row* vectors. Let $c \in R^n$, and let $b_i,\ 1 \leq i \leq k$, be real numbers. Consider the problem

$$\begin{aligned} &\text{Maximize } c'x \\ &\text{subject to } A_ix \leq b_i,\ 1 \leq i \leq k\ . \end{aligned}$$

For any x satisfying the constraints, let $I(x) \subset \{1, \ldots, n\}$ be such that $A_i(x) = b_i, i \in I(x), A_ix < b_i, i \notin I(x)$. Suppose x^* satisfies the constraints. Show that x^* is optimal if an only if

$$c'h \leq 0 \text{ for all } h \text{ such that } A_ih \leq 0\ ,\ i \in I(x^*).$$

Returning to our problem, it is clear that (4.3) is satisfied as long as c lies between A_1 and A_2. Mathematically this means that (4.3) is satisfied if and only if there exist $\lambda_1^* \geq 0,\ \lambda_2^* \geq 0$ such that [2]

$$c' = \lambda_1^*, A_1 + \lambda_2^* A_2. \tag{4.4}$$

As c varies, the optimal decision will change. We can see from our analysis that the situation is as follows (see Figure 4.1):

[2] Although this statement is intuitively obvious, its generalization to n dimensions is a deep theorem known as Farkas' lemma (see Section 2).

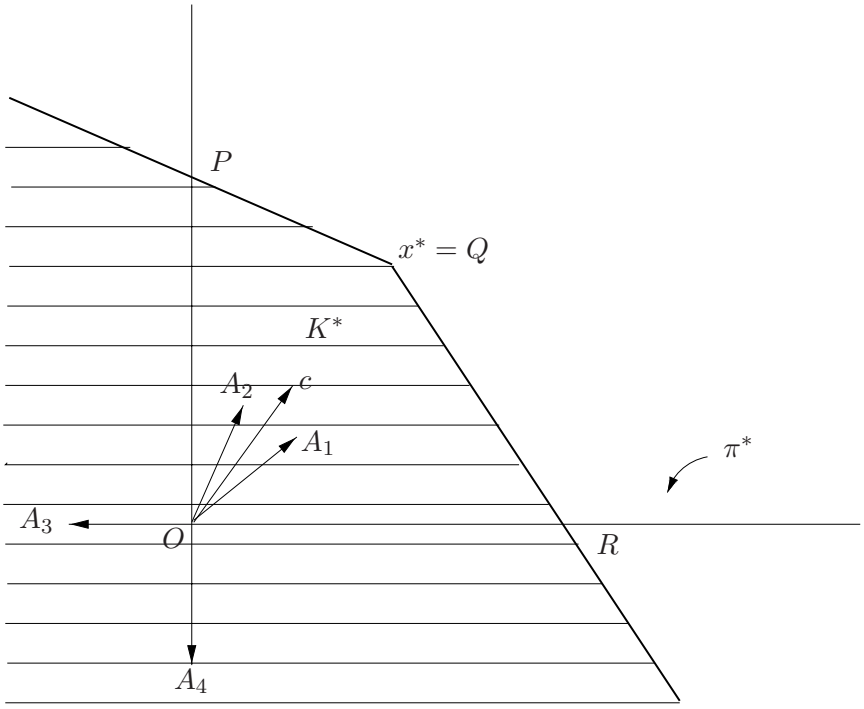


Figure 4.2: K^* is the cone generated by Ω at x^*.

1. $x^* = Q$ is optimal iff c lies between A_1 and A_2 iff $c' = \lambda_1^* A_1 + \lambda_2^* A_2$ for some $\lambda_1^* \geq 0,\ \lambda_2^* \geq 0$,
2. $x^* \in QP$ is optimal iff c lies along A_2 iff $c' = \lambda_2^* A_2$ for some $\lambda_2^* \geq 0$,
3. $x^* = P$ is optimal iff c lies between A_3 and A_2 iff $c' = \lambda_2^* A_2 + \lambda_3^* A_3$ for some $\lambda_2^* \geq 0, \lambda_3^* \geq 0$, *etc.*

These statements can be made in a more elegant way as follows:

$x^* \in \Omega$ is optimal iff there exists $\lambda_i^* \geq 0\ ,\ 1 \leq i \leq 4$, such that

$$\text{(a)}\ \ c' = \sum_{i=1}^{4} \lambda_i^* a_i\ ,\ \ (b)\ \text{ if } A_i\, x^* < b^i \text{ then } \lambda_i^* = 0\ . \tag{4.5}$$

For purposes of application it is useful to separate those constraints which are of the form $x_i \geq 0$, from the rest, and to reformulate (4.5) accordingly We leave this as an exercise.

Exercise 2: Show that (4.5) is equivalent to (4.6), below. (Here $A_i = (a_{i1}, a_{i2})$.) $x^* \in \Omega$ is optimal iff there exist $\lambda_1^* \geq 0\ ,\ \lambda_2^* \geq 0$ such that

$$\begin{array}{l} (a)\ \ c_i \leq \lambda_1^* a_{1i} + \lambda_2^* a_{2i},\ i = 1,\ 2, \\ (b)\ \text{ if } a_{j1} x_1^* + a_{j2} x_2^* < b_j \text{ then } x_j^* = 0, j = 1, 2. \\ (c)\ \text{ if } c_i < \lambda_{1i}^* + \lambda_2^* a_{2i} \text{ then } x_i^* = 0, i = 1, 2. \end{array} \tag{4.6}$$

4.1.2 *Problem formulation.*

A linear programming problem (or LP in brief) is any decision problem of the form 4.7.

$$\begin{array}{l}\text{Maximize } c_1x_1 + c_2x_2 + \ldots + c_nx_n \\ \text{subject to} \\ a_{il}x_1 + a_{i2}x_2 + \ldots + a_{in}x_n \leq \ b_i \ , \ l \leq i \leq k \ , \\ a_{il}x_1 + \ldots\ldots\ldots + a_{in}x_n \ \geq \ b_i \ , \ k+1 \leq i \leq \ell \ , \\ a_{il}x_1 + \ldots\ldots\ldots + a_{in}x_n \ = \ b_i \ , \ \ell+1 \leq i \leq m \ ,\end{array}$$

and

$$\begin{array}{l} x_j \geq 0 \quad , \ 1 \leq j \leq p \ , \\ x_j \geq 0 \quad , \ p+1 \leq j \leq q; \\ x_j \text{ arbitary} \ , \ q+1 \leq j \leq n \ , \end{array} \tag{4.7}$$

where the c_j, a_{ij}, b_i are fixed real numbers.

There are two important special cases:
Case I: (4.7) is of the form (4.8):

$$\begin{array}{l} \text{Maximize } \displaystyle\sum_{j=1}^{n} \ c_jx_j \\ \text{subject to } \displaystyle\sum_{j=1}^{n} \ a_{ij}x_j \leq b_i \ , \quad 1 \leq i \leq m \ , \\ \qquad\qquad x_j \geq 0 \quad , \ 1 \leq j \leq n \end{array} \tag{4.8}$$

Case II: (4.7) is of the form (4.9):

$$\begin{array}{l} \text{Maximize } \displaystyle\sum_{j=1}^{n} \ c_jx_j \\ \text{subject to } \displaystyle\sum_{j=1}^{n} \ a_{ij}x_j = b_i \ , \quad 1 \leq i \leq m \ , \\ \qquad\qquad x_j \geq 0 \quad , \ 1 \leq j \leq n \ . \end{array} \tag{4.9}$$

Although (4.7) appears to be more general than (4.8) and (4.9), such is not the case.

Proposition: Every LP of the form (4.7) can be transformed into an equivalent LP of the form (4.8).
Proof.
Step 1: Replace each inequality constraint $\sum a_{ij}x_j \geq b_i$ by $\sum(-a_{ij})x_j \leq (-b_i)$.
Step 2: Replace each equality constraint $\sum a_{ij}x_j = b_i$ by two inequality constraints:
$\sum a_{ij}x_j \leq b_i, \ \sum(-a_{ij})x_j \leq (-b_i)$.
Step 3: Replace each variable x_j which is constrained $x_j \leq 0$ by a variable $y_j = -x_j$ constrained $y_j \geq 0$ and then replace $a_{ij}x_j$ by $(-a_{ij})y_j$ for every i and c_jx_j by $(-c_j)y_j$.

Step 4: Replace each variable x_j which is not constrained in sign by a pair of variables $y_j - z_j = x_j$ constrained $y_j \geq 0, z_j \geq 0$ and then replace $a_{ij}x_j$ by $a_{ij}y_j + (-a_{ij})z_j$ for every i and $c_j x_j$ by $c_j y_j + (-c_j)z_j$. Evidently the resulting LP has the form (4.8) and is equivalent to the original one. ♢

Proposition: Every LP of the form (4.7) can be transformed into an equivalent LP of the from (4.9)

Proof.

Step 1: Replace each inequality constraint $\sum a_{ij}x_j \leq b_i$ by the equality constraint $\sum a_{ij}x_j + y_i = b_i$ where y_i is an additional variable constrained $y_i \geq 0$.

Step 2: Replace each inequality constraint $\sum a_{ij}x_j \geq b_i$ by the equality constraint $\sum a_{ij}x_j - y_i = b_i$ where y_i is an additional variable constrained by $y_i \geq 0$. (The new variables added in these steps are called *slack* variables.)

Step 3, Step 4: Repeat these steps from the previous proposition. Evidently the new LP has the form (4.9) and is equivalent to the original one. ♢

4.2 *Qualitative Theory of Linear Programming*

4.2.1 *Main results.*

We begin by quoting a fundamental result. For a proof the reader is referred to (Mangasarian [1969]).

Farkas' Lemma. Let $A_i,\ 1 \leq i \leq k$, be n-dimensional *row* vectors. Let $c \in R^n$ be a column vector. The following statements are equivalent:

(i) for all $x \in R^n$, $A_i x \leq 0$ for $1 \leq i \leq k$ implies $c'x \leq 0$,

(ii) there exists $\lambda_1 \geq 0, \ldots, \lambda_k \geq 0$ such that $c' = \sum_{i=1}^{k} \lambda_i A_i$.

An algebraic version of this result is sometimes more convenient.

Farkas' Lemma (algebraic version). Let A be a $k \times n$ matrix. Let $c \in R^n$. The following statements are equivalent.

(i) for all $x \in R^n,\ Ax \leq 0$ implies $c'x \leq 0$,

(ii) there exists $\lambda \geq 0,\ \lambda \in R^k$, such that $A'\lambda = c$.

Using this result it is possible to derive the main results following the intuitive reasoning of (4.1). We leave this development as two exercises and follow a more elegant but less intuitive approach.

Exercise 1: With the same hypothesis and notation of Exercise 1 in 4.1, use the first version of Farkas′ lemma to show that there exist $\lambda_i^* \geq 0$ for $i \in I(x^*)$ such that $\sum_{i \in I(x^*)} \lambda_i^* A_i = c'$.

Exercise 2: Let x^* satisfy the constraints for problem (4.17). Use the previous exercise to show that x^* is optimal iff there exist $\lambda_1^* \geq 0, \ldots, \lambda_m^* \geq 0$ such that

(a) $c_j \leq \sum_{i=1}^{m} \lambda_i^* a_{ij}\ ,\ 1 \leq j \leq n$

(b) if $\sum_{j=1}^{n} a_{ij}x_j^* < b_i$ then $\lambda_i^* = 0\ ,\ 1 \leq i \leq m$ (c) if $\sum_{i=1}^{m} \lambda_i^* a_{ij} > c_j$ then $x_j^* = 0\ ,\ 1 \leq j \leq m$.

In the remaining discussion, $c \in R^n, b \in^n$ are fixed vectors, and $A = \{a_{ij}\}$ is a fixed $m \times n$ matrix, whereas $x \in R^n$ and $\lambda \in R^m$ will be variable. Consider the pair of LPs (4.10) and (4.11)

below. (4.10) is called the *primal* problem and (4.11) is called the *dual* problem.

$$\begin{array}{ll} \text{Maximize} & c_1x_1 + \ldots + c_nx_n \\ \text{subject to} & a_{i1}x_1 + \ldots + a_{in}x_n \leq b_i \ , \quad 1 \leq i \leq m \\ & x_j \geq 0 \ , \quad 1 \leq j \leq n \ . \end{array} \tag{4.10}$$

$$\begin{array}{ll} \text{Maximize} & \lambda_1 b_1 + \ldots + \lambda_m b_m \\ \text{subject to} & \lambda_1 a_{1j} + \ldots + \lambda_m a_{mj} \geq c_j \ , \quad 1 \leq j \leq n \\ & \lambda_i \geq 0 \ , \quad 1 \leq i \leq m \ . \end{array} \tag{4.11}$$

Definition: Let $\Omega_p = \{x \in R^n | Ax \leq b, x \geq 0\}$ be the set of all points satisfying the constraints of the primal problem. Similarly let $\Omega_d = \{\lambda \in R^m | \lambda' A \geq c', \lambda \geq 0\}$. A point $x \in \Omega_p (\lambda \in \Omega_d)$ is said to be a *feasible solution* or *feasible decision* for the primal (dual).

The next result is trivial.

Lemma 1: (Weak duality) Let $x \in \Omega_p, \lambda \in \Omega_d$. Then

$$c'x \leq \lambda' Ax \leq \lambda' b. \tag{4.12}$$

Proof: $x \geq 0$ and $\lambda' A - c' \geq 0$ implies $(\lambda' A - c')x \geq 0$ giving the first inequality. $b - Ax \geq 0$ and $\lambda' \geq 0$ implies $\lambda'(b - Ax) \geq 0$ giving the second inequality. $\diamondsuit$

Corollary 1: If $x^* \in \Omega$ and $\lambda^* \in \Omega_d$ such that $c'x^* = (\lambda^*)'b$, then x^* is optimal for (4.10) and λ^* is optimal for (4.11).

Theorem 1: (Strong duality) Suppose $\Omega_p \neq \phi$ and $\Omega_d \neq \phi$. Then there exists x^* which is optimum for (4.10) and λ^* which is optimum for (4.11). Furthermore, $c'x^* = (\lambda^*)'b$.

Proof: Because of the Corollary 1 it is enough to prove the last statement, *i.e.*, we must show that there exist $x \geq 0, \lambda \geq 0$, such that $Ax \leq b, A'\lambda \geq c$ and $b'\lambda - c'x \leq 0$. By introducing slack variables $y \in R^m, \mu \in R^m, r \in R$, this is equivalent to the existence of $x \geq 0, y \geq 0, \lambda \geq 0, \mu \leq 0, r \leq 0$ such that

$$\left[\begin{array}{c|c|c|c|c} A & I_m & & & \\ \hline & & A' & -I_n & \\ \hline -c' & & b' & & 1 \end{array}\right] \left[\begin{array}{c} x \\ \hline y \\ \hline \lambda \\ \hline \mu \\ \hline r \end{array}\right] = \left[\begin{array}{c} b \\ \hline c \\ \hline 0 \end{array}\right]$$

By the algebraic version of Farkas' Lemma, this is possible only if

$$\begin{array}{l} A'\xi - c\theta \leq 0 \quad , \ \xi \leq 0 \ , \\ Aw = b\theta \leq 0 \ , \ -w \leq 0 \ , \\ \theta \leq 0 \end{array} \tag{4.13}$$

implies

$$b'\xi + c'w \leq 0. \tag{4.14}$$

Case (i): Suppose (w, ξ, θ) satisfies (4.13) and $\theta < 0$. Then $(\xi/\theta) \in \Omega_d$, $(w/-\theta) \in \Omega_p$, so that by Lemma 1 $c'w/(-\theta) \leq b'\xi/\theta$, which is equivalent to (4.14) since $\theta < 0$.
Case (ii): Suppose (w, ξ, θ) satisfies (4.13) and $\theta = 0$, so that $-A'\xi \geq 0$, $-\xi \geq 0$, $Aw \leq 0$, $w \geq 0$. By hypothesis, there exist $x \in \Omega_p$, $\lambda \in \Omega_d$. Hence, $-b'\xi = b'(-\xi) \geq (Ax)'(-\xi) = x'(-A'\xi) \geq 0$, and $c'w \leq (A'\lambda)'w = \lambda'(Aw) \leq 0$. So that $b'\xi + c'w \leq 0$. $\diamondsuit$

The existence part of the above result can be strengthened.

Theorem 2: (i) Suppose $\Omega_p \neq \phi$. Then there exists an optimum decision for the primal LP iff $\Omega_d \neq \phi$.
(ii) Suppose $\Omega_d \neq \phi$. Then there exists an optimum decision for the dual LP iff $\Omega_p \neq \phi$.
Proof Because of the symmetry of the primal and dual it is enough to prove only (i). The sufficiency part of (i) follows from Theorem 1, so that only the necessity remains. Suppose, in contradiction, that $\Omega_d = \phi$. We will show that sup $\{c'x | x \in \Omega_p\} = +\infty$. Now, $\Omega_d = \phi$ means there does not exist $\lambda \geq 0$ such that $A'\lambda \geq c$. Equivalently, there does not exist $\lambda \geq 0,\ \mu \leq 0$ such that

$$\left[\begin{array}{c|c} A' & -I_n \end{array}\right] \left[\begin{array}{c} \lambda \\ - - - \\ \mu \end{array}\right] = \left[\begin{array}{c} c \end{array}\right]$$

By Farkas' Lemma there exists $w \in R^n$ such that $Aw \leq 0$, $-w \leq 0$, and $c'w > 0$. By hypothesis, $\Omega_p \neq \phi$, so there exists $x \geq 0$ such that $Ax \leq b$. but then for any $\theta > 0$, $A(x + \theta w) \leq b$, $(x + \theta w) \geq 0$, so that $(x + \theta w) \in \Omega_p$. Also, $c'(x + \theta w) = c'x + \theta c'w$. Evidently then, sup $\{c'x | x \in \Omega_p\} = +\infty$ so that there is no optimal decision for the primal. $\diamondsuit$

Remark: In Theorem 2(i), the hypothesis that $\Omega_p \neq \phi$ is essential. Consider the following exercise.
Exercise 3: Exhibit a pair of primal and dual problems such that *neither* has a feasible solution.
Theorem 3: (Optimality condition) $x^* \in \Omega_p$ is optimal if and only if there exists $\lambda^* \in \Omega_d$ such that

$$\begin{array}{l} \sum_{j=1}^{m} a_{ij} x_j^* < b_i \text{ implies } \lambda_i^* = 0 \ , \\ \text{and} \\ \sum_{i=1}^{m} \lambda_i^* a_{ij} < c_j \text{ implies } x_j^* = 0 \ . \end{array} \tag{4.15}$$

((4.15) is known as the condition of *complementary slackness*.)

Proof: First of all we note that for $x^* \in \Omega_p$, $\lambda^* \in \Omega_d$, (4.15) is equivalent to (4.16):

$$(\lambda^*)'(Ax^* - b) = 0, \text{ and } (A'\lambda^* - c)'x^* = 0 \ . \tag{4.16}$$

Necessity. Suppose $x^* \in \Omega_p$ is optimal. Then from Theorem 2, $\Omega_d \neq \phi$, so that by Theorem 1 there exists $\lambda^* \in \Omega_d$ such that $c'x^* = (\lambda^*)'b$. By Lemma 1 we always have $c'x^* \leq (\lambda^*)'Ax^* \leq (\lambda^*)'b$ so that we must have $c'x^* = (\lambda^*)'Ax^* = (\lambda^*)'b$. But (4.16) is just an equivalent rearrangement of these two equalities.
Sufficiency. Suppose (4.16) holds for some $x^* \in \Omega_p$, $\lambda^* \in \Omega_d$. The first equality in (4.16) yields $(\lambda^*)'b = (\lambda^*)'Ax^* = (A'\lambda^*)'x^*$, while the second yields $(A'\lambda^*)'x^* = c'x^*$, so that $c'x^* = (\lambda^*)'b$. By Corollary 1, x^* is optimal. $\diamondsuit$

The conditions $x^* \in \Omega_p$, $x^* \in \Omega_d$ in Theorem 3 can be replaced by the weaker $x^* \geq 0$, $\lambda^* \geq 0$ provided we strengthen (4.15) as in the following result, whose proof is left as an exercise.

Theorem 4: (Saddle point) $x^* \geq 0$ is optimal for the primal if and only if there exists $\lambda^* \geq 0$ such that

$$L(x, \lambda^*) \leq L(x^*, \lambda^*) \leq L(x^*, \lambda) \text{ for all } x \geq 0, \text{ and all} \lambda \geq 0, \tag{4.17}$$

where L: $R^n x R^m \to R$ is defined by

$$L(x, \lambda) = c'x - \lambda'(Ax - b) \tag{4.18}$$

Exercise 4: Prove Theorem 4.
Remark. The function L is called the *Lagrangian.* A pair (x^*, λ^*) satisfying (4.17) is said to form a *saddle-point* of L over the set $\{x | x \in R^n, x \geq 0\} \times \{\lambda | \lambda \in R^m, \lambda \geq 0\}$.

4.2.2 *Results for problem (4.9).*

It is possible to derive analogous results for LPs of the form (4.9). We state these results as exercises, indicating how to use the results already obtained. We begin with a pair of LPs:

$$\begin{array}{ll} \text{Maximize} & c_1 x_1 + \ldots + c_n x_n \\ \text{subject to} & a_{il} x_1 + \ldots + a_{in} x_n = b_i \ , \quad 1 \leq i \leq m \ , \\ & \qquad\qquad x_j \geq 0 \ , \quad 1 \leq j \leq n \ . \end{array} \tag{4.19}$$

$$\begin{array}{ll} \text{Minimize} & \lambda_1 b_1 + \ldots + \lambda_m b_m \\ \text{subject to} & \lambda_1 a_{1j} + \ldots + \lambda_m a_{mj} \geq c_j \quad , \quad 1 \leq j \leq n \quad . \end{array} \tag{4.20}$$

Note that in (4.20) the λ_i are unrestricted in sign. Again (4.19) is called the primal and (4.20) the dual. We let Ω_p, Ω_d denote the set of all x, λ satisfying the constraints of (4.19), (4.20) respectively.

Exercise 5: Prove Theorems 1 and 2 with Ω_p and Ω_d interpreted as above. (Hint. Replace (4.19) by the equivalent LP: maximize $c'x$, subject to $Ax \leq b$, $(-A)x \leq (-b)$, $x \geq 0$. This is now of the form (4.10). Apply Theorems 1 and 2.)
Exercise 6: Show that $x^* \in \Omega_p$ is optimal iff there exists $\lambda^* \in \Omega_d$ such that

$$x_j^* > 0 \text{ implies } \sum_{i=1}^{m} \lambda_i^* a_{ij} = c_j \ .$$

Exercise 7: $x^* \geq 0$ is optimal iff there exists $\lambda^* \in R^m$ such that

$$L(x, \lambda^*) \leq L(x^*, \lambda^*) \leq L(x^*, \lambda) \text{ for all } x \geq 0, \ \lambda \in R^m \ .$$

where L is defined in (4.18). (Note that, unlike (4.17), λ is not restricted in sign.)
Exercise 8: Formulate a dual for (4.7), and obtain the result analogous to Exercise 5.

4.2.3 *Sensitivity analysis.*

We investigate how the maximum value of (4.10) or (4.19) changes as the vectors b and c change. The matrix A will remain fixed. Let Ω_p and Ω_d be the sets of feasible solutions for the pair (4.10) and (4.11) *or* for the pair (4.19) and (4.20). We write $\Omega_p(b)$ and $\Omega_d(c)$ to denote the explicit dependence on b and c respectively. Let $B = \{b \in R^m | \Omega_p(b) \neq \phi\}$ and $C = \{c \in R^n | \Omega_d(c) \neq \phi\}$, and for $(b, c) \in B \times C$ define

$$M(b,c) = \max\{c'x | x \in \Omega_p(b)\} = \min\{\lambda' b | \lambda \in \Omega_d(c)\} . \tag{4.21}$$

For $1 \leq i \leq m$, $\varepsilon \in R$, $b \in R^m$ denote

$$b(i,\varepsilon) = (b_1, b_2, \ldots, b_{i-1}, b_i + \varepsilon, b_{i+1}, \ldots, b_m)' ,$$

and for $1 \leq j \leq n$, $\varepsilon \in R$, $c \in R^n$ denote

$$c(j,\varepsilon) = (c_1, c_2, \ldots, c_{j-1}, c_j + \varepsilon, c_{j+1}, \ldots, c_n)' .$$

We define in the usual way the right and left hand partial derivatives of M at a point $(\hat{b}, \hat{c}) \in B \times C$ as follows:

$$\frac{\partial M^+}{\partial b_i}(\hat{b},\hat{c}) = \lim_{\substack{\varepsilon \to 0 \\ \varepsilon > 0}} \frac{1}{\varepsilon}\{M(\hat{b}(i,\varepsilon),\hat{c}) - M(\hat{b},\hat{c})\} ,$$

$$\frac{\partial M^-}{\partial b_i}(\hat{b},\hat{c}) = \lim_{\substack{\varepsilon \to 0 \\ \varepsilon > 0}} \frac{1}{\varepsilon}\{M(\hat{b},\hat{c}) - M(\hat{b}(i,-\varepsilon),\hat{c})\} ,$$

$$\frac{\partial M^+}{\partial c_j}(\hat{b},\hat{c}) = \lim_{\substack{\varepsilon \to 0 \\ \varepsilon > 0}} \frac{1}{\varepsilon}\{M(\hat{b},\hat{c}(j,\varepsilon)) - M(\hat{b},\hat{c}\} ,$$

$$\frac{\partial M^-}{\partial c_j}(\hat{b},\hat{c}) = \lim_{\substack{\varepsilon \to 0 \\ \varepsilon > 0}} \frac{1}{\varepsilon}\{M(\hat{b},\hat{c} - M(\hat{b},\hat{c}(j,-\varepsilon))\} ,$$

Let $\overset{\circ}{B}, \overset{\circ}{C}$ denote the interiors of B, C respectively.

Theorem 5: At each $(\hat{b},\hat{c}) \in \overset{\circ}{B} \times \overset{\circ}{C}$, the partial derivatives given above exist. Furthermore, if $\hat{x} \in \Omega_p(\hat{b})$, $\hat{\lambda} \in \Omega_d(\hat{c})$ are optimal, then

$$\frac{\partial M^+}{\partial b_i}(\hat{b},\hat{c}) \leq \hat{\lambda}_i \leq \frac{\partial M^-}{\partial b_i}(\hat{b},\hat{c}) , \quad 1 \leq i \leq m , \tag{4.22}$$

$$\frac{\partial M^+}{\partial c_j}(\hat{b},\hat{c}) \geq \hat{x}_j \geq \frac{\partial M^-}{\partial c_j}(\hat{b},\hat{c}) \ , \quad 1 \leq j \leq n \ , \tag{4.23}$$

Proof: We first show (4.22), (4.23) assuming that the partial derivatives exist. By strong duality $M(\hat{b},\hat{c}) = \hat{\lambda}'\hat{b}$, and by weak duality $M(\hat{b}(i,\varepsilon),\hat{c}) \leq \hat{\lambda}'\hat{b}(i,\varepsilon)$, so that

$$\begin{array}{rcl} \frac{1}{\varepsilon}\{M(\hat{b}(i,\varepsilon),\hat{c}) - M(\hat{b},\hat{c})\} & \leq & \frac{1}{\varepsilon}\hat{\lambda}'\{\hat{b}(i,\varepsilon) - \hat{b}\}\hat{\lambda}_i, \text{ for } \varepsilon > 0, \\ \frac{1}{\varepsilon}\{M(\hat{b},\hat{c}) - M(\hat{b}(i,-\varepsilon),\hat{c})\} & \geq & \frac{1}{\varepsilon}\hat{\lambda}'\{\hat{b} - \hat{b}(i,-\varepsilon)\} = \hat{\lambda}_i, \text{ for } \varepsilon > 0. \end{array}$$

Taking limits as $\varepsilon \to 0$, $\varepsilon > 0$, gives (4.22).

On the other hand, $M(\hat{b},\hat{c}) = \hat{c}'\hat{x}$, and $M(\hat{b},\hat{c}(j,\varepsilon)) \geq (\hat{c}(j,\varepsilon))'\hat{x}$, so that

$$\begin{array}{c} \frac{1}{\varepsilon}\{M(\hat{b},\hat{c}(j,\varepsilon)) - M(\hat{b},\hat{c})\} \geq \frac{1}{\varepsilon}\{\hat{c}(j,\varepsilon)' - \hat{c}\}'\hat{x} = \hat{x}_j, \text{ for } \varepsilon > 0, \\ \frac{1}{\varepsilon}\{M(\hat{b},\hat{c}) - M(\hat{b},\hat{c}(j,-\varepsilon))\} \leq \frac{1}{\varepsilon}\{\hat{c} - \hat{c}(j,-\varepsilon)\}'\hat{x} = \hat{x}_j, \text{ for } \varepsilon > 0, \end{array}$$

which give (4.23) as $\varepsilon \to 0$, $\varepsilon > 0$.

Finally, the existence of the right and left partial derivatives follows from Exercises 8, 9 below. $\diamondsuit$

We recall some fundamental definitions from convex analysis.

Definition: $X \subset R^n$ is said to be *convex* if $x, y \in X$ and $0 \leq \theta \leq 1$ implies $(\theta x + (1-\theta)y) \in X$.

Definition: Let $X \subset R^n$ and $f : X \to R$. (i) f is said to be *convex* if X is convex, and $x, y \in X$, $0 \leq \theta \leq 1$ implies $f(\theta x + (1-\theta)y) \leq \theta f(x) + (1-\theta)f(y)$. (ii) f is said to be *concave* if $-f$ is convex, *i.e.*, $x, y \in X$, $0 \leq \theta \leq 1$ implies $f(\theta x + (1-\theta)y) \geq \theta f(x) + (1-\theta)f(y)$.

Exercise 8: (a) Show that Ω_p, Ω_d, and the sets $B \subset R^m$, $C \subset R^n$ defined above are convex sets. (b) Show that for fixed $c \in C$, $M(\cdot, c) : B \to R$ is concave and for fixed $b \in B$, $M(b, \cdot) : C \to R$ is convex.

Exercise 9: Let $X \subset R^n$, and $f : X \to R$ be convex. Show that at each point $\hat{x}$ in the interior of X, the left and right hand partial derivatives of f exist. (Hint: First show that for $\varepsilon_2 > \varepsilon_1 > 0 > \delta_1 > \delta_2$,$(1/\varepsilon_2)\{f(\hat{x}(i,\varepsilon_2)) - f(\hat{x})\} \geq (1/\varepsilon_1)\{f(\hat{x}(i,\varepsilon_1)) - f(\hat{x}))\} \geq (1/\delta_1)\{f(\hat{x}(i,\delta_1)) - f(\hat{x}))\} \geq (1/\delta_2)\{f(\hat{x}(i,\delta_2)) - f(\hat{x})\}$. Then the result follows immediately.)

Remark 1: Clearly if $(\partial M/\partial b_i)(\hat{b})$ exists, then we have equality in (4.22), and then this result compares with 3.14).

Remark 2: We can also show without difficulty that $M(\cdot, c)$ and $M(b, \cdot)$ are piecewise linear (more accurately, linear plus constant) functions on B and C respectively. This is useful in some computational problems.

Remark 3: The variables of the dual problem are called Lagrange variables or dual variables or shadow-prices. The reason behind the last name will be clear in Section 4.

4.3 *The Simplex Algorithm*

4.3.1 *Preliminaries*

We now present the celebrated Simplex algorithm for finding an optimum solution to any LP of the form (4.24):

$$\begin{array}{lrl} \text{Maximize} & c_1x_1 + \ldots + c_nx_n & \\ \text{subject to} & a_{il}x_1 + \ldots + a_{in}x_n = b_i \ , & 1 \leq i \leq m \\ & x_j \geq 0 \ , & 1 \leq j \leq n \ . \end{array} \tag{4.24}$$

As mentioned in 4.1 the algorithm rests upon the observations that if an optimal exists, then at least one vertex of the feasible set Ω_p is an optimal solution. Since Ω_p has only finitely many vertices (see Corollary 1 below), we only have to investigate a finite set. The practicability of this investigation depends on the ease with which we can characterize the vertices of Ω_p. This is done in Lemma 1.

In the following we let A^j denote the *jth column* of A, *i.e.*, $A^j = (a_{1j}, \ldots, a_{mj})'$. We begin with a precise definition of a vertex.

Definition: $x \in \Omega_p$ is said to be a *vertex* of Ω_p if $x = \lambda y + (1-\lambda)z$, with y, z in Ω_p and $0 < \lambda < 1$, implies $x = y = z$.
Definition: For $x \in \Omega_p$, let $I(x) = \{j | x_j > 0\}$.
Lemma 1: Let $x \in \Omega_p$. Then x is a vertex of Ω_p iff $\{A^j | j \in I(x)\}$ is a linearly independent set.
Exercise 1: Prove Lemma 1.
Corollary 1: Ω_p has at most $\sum_{j=1}^{m} \frac{n!}{(n-j)!}$ vertices.
Lemma 2: Let x^* be an optimal decision of (4.24). Then there is a vertex z^* of Ω_p which is optimal.
Proof: If $\{A^j | j \in I(x^*)\}$ is linearly independent, let $z^* = x^*$ and we are done. Hence suppose $\{A^j | j \in I(x^*)\}$ is linearly dependent so that there exist γ_j, not all zero, such that

$$\sum_{j \in I(x^*)} \gamma_j A^j = 0 \ .$$

For $\theta \in R$ define $z(\theta) \in R^n$ by

$$z_j(\theta) = \begin{cases} x_j^* = \theta\gamma_j \ , & j \in I(x^*) \\ x_j^* = 0 \quad , & j \notin I(x^*) \ . \end{cases}$$

$$Az(\theta) = \sum_{j \in I(x^*)} z_j(\theta) A^j = \sum_{j \in I(x^*)} x_j^* A^j + \theta \sum_{j \in I(x^*)} \gamma_j A^j$$

$$= b + \theta \cdot 0 = b \ .$$

Since $x_j^* > 0$ for $j \in I(x^*)$, it follows that $z(\theta) \geq 0$ when

$$|\theta| \leq \min \left\{ \frac{x_j^*}{|\gamma_j|} \ \middle| \ j \in I(x^*) \right\} = \theta^* \text{ say} \ .$$

Hence $z(\theta) \in \Omega_p$ whenever $|\theta| \leq \theta^*$. Since x^* is optimal we must have

$$c'x^* \geq c'z(\theta) = c'x^* + \theta \sum_{j \in I(x^*)} c_j y_j \text{ for } -^*\theta \leq \theta \leq \theta^* \ .$$

Since θ can take on positive and negative values, the inequality above can hold on if $\sum_{J \in I(x^*)} c_j \gamma_j = 0$, and then $c'x^* = c'z(\theta)$, so that $z(\theta)$ is also an optimal solution for $|\theta| \leq \theta^*$. But from the definition of $z(\theta)$ it is easy to see that we can pick θ_0 with $|\theta_0| = \theta^*$ such that $z_j(\theta_0) = x_j^* + \theta_0 \gamma_j = 0$ for at least one $j = j_0$ in $I(x^*)$. Then,

$$I(z(\theta_0)) \subset I(x^*) - \{j_0\} \ .$$

Again, if $\{A^j | j \in I(z(\theta_0))\}$ is linearly independent, then we let $z^* = z(\theta_0)$ and we are done. Otherwise we repeat the procedure above with $z(\theta_0)$. Clearly, in a finite number of steps we will find an optimal decision z^* which is also vertex. $\diamondsuit$

At this point we abandon the geometric term "vertex" and how to established LP terminology.

Definition: (i) z is said to be a *basic feasible solution* if $z \in \Omega_p$, and $\{A^j | j \in I(z)\}$ is linearly independent. The set $I(z)$ is then called the *basis at z*, and x_j, $j \in I(z)$, are called the *basic variables at z*. x_j, $j \notin I(z)$ are called the *non-basic variables at z*.

Definition: A basic feasible solution z is said to be *non-degenerate* if $I(z)$ has m elements.

Notation: Let z be a non-degenerate basic feasible solution, and let $j_1 < j_2 < \ldots < j_m$ constitute $I(z)$. Let $D(z)$ denote the $m \times m$ non-singular matrix $D(z) = [A^{j_1} \vdots A^{j_2} \vdots \ldots \vdots A^{j_m}]$, let $c(z)$ denote the m-dimensional column vector $c(z) = (c_{j_1}, \ldots, c_{j_m})'$ and define $\lambda(z)$ by $\lambda'(z) = c'(z)[D(z)]^{-1}$. We call $\lambda(z)$ the *shadow-price vector* at z.

Lemma 3: Let z be a non-degenerate basic feasible solution. Then z is optimal if and only if

$$\lambda'(z)A \geq c_j \text{ , for all , } j \notin I(z) \text{ .} \tag{4.25}$$

Proof: By Exercise 6 of Section 2.2, z is optimal iff there exists λ such that

$$\lambda' A^j = c_j \text{ , for , } j \in I(z) \text{ ,} \tag{4.26}$$

$$\lambda' A^j \geq c_j \text{ , for , } j \notin I(z) \text{ ,} \tag{4.27}$$

But since z is non-degenerate, (4.26) holds iff $\lambda = \lambda(z)$ and then (4.27) is the same as (4.25). $\diamondsuit$

4.3.2 *The Simplex Algorithm.*

The algorithm is divided into two parts: In Phase I we determine if Ω_p is empty or not, and if not, we obtain a basic feasible solution. Phase II starts with a basic feasible solution and determines if it is optimal or not, and if not obtains another basic feasible solution with a higher value. Iterating on this procedure, in a finite number of steps, either we obtain an optimum solution or we discover that no optimum exists, *i.e.*, $\sup \{c'x | x \in \Omega_p\} = +\infty$. We shall discuss Phase II first.

We make the following simplifying assumption. We will comment on it later.

Assumption of non-degeneracy. Every basic feasible solution is non-degenerate.

Phase II:

Step 1. Let z^0 be a basic feasible solution obtained from Phase I or by any other means. Set $k = 0$ and go to Step 2.

Step 2. Calculate $[D(z^k)]^{-1}$,$c(z^k)$, and the shadow-price vector $\lambda'(z^k) = c'(z^k)[D(z^k)]^{-1}$. For each $j \notin I(z^k)$ calculate $c_j - \lambda'(z^k)A^j$. If all these numbers are ≤ 0, stop, because z^k is optimal by Lemma 3. Otherwise pick any $\hat{j} \notin I(z^k)$ such that $c_{\hat{j}} - \lambda'(z^k)A^{\hat{j}} > 0$ and go to Step 3.

Step 3. Let $I(z^k)$ consist of $j_1 < j_2 < \ldots < j_m$. Compute the vector $\gamma^k = (\gamma^k_{j_1}, \ldots \gamma^k_{j_m})' = [D(z^k)]^{-1}A^{\hat{j}}$. If $\gamma^k \leq 0$, stop, because by Lemma 4 below, there is no finite optimum. Otherwise go to Step 4.

Step 4. Compute θ = min $\{(z^k_j \gamma^k_j) | j \in i(z), \gamma^k_j > 0\}$. Evidently $0 < \theta < \infty$. Define z^{k+1} by

$$z_j^{k+1} = \begin{cases} z_j^k - \theta\gamma_j^k & , \quad j \in I(z) \\ \theta & , \quad j = \hat{j} \\ z_j^k = 0 & , \quad j \neq \hat{j} \qquad \text{and } j \notin I(z) \ . \end{cases} \tag{4.28}$$

By Lemma 5 below, z^{k+1} is a basic feasible solution with $c'z^{k+1} > c'z^k$. Set $k = k+1$ and return to Step 2.
Lemma 4: If $\gamma^k \leq 0$, sup $\{c'x | x \in \Omega_p\} = \infty$.
Proof: Define $z(\theta)$ by

$$z_j(\theta) = \begin{cases} z_j - \theta\gamma_j^k & , \quad j \in I(z) \\ \theta & , \quad j = \hat{j} \\ z_j = 0 & , \quad j \notin I(z) \quad \text{and } j \neq \hat{j} \ . \end{cases} \tag{4.29}$$

First of all, since $\gamma^k \leq 0$ it follows that $z(\theta) \geq 0$ for $\theta \geq 0$. Next, $Az(\theta) = Az - \theta \sum_{j \in I(z)} \gamma_j^k A^j + \theta A^{\hat{j}} = Az$ by definition of γ^k. Hence, $z(\theta) \in \Omega_p$ for $\theta \geq 0$. Finally,

$$\begin{aligned} c'z(\theta) &= c'z - \theta c'(z^k)\gamma^k + \theta c_{\hat{j}} \\ &= c'z + \theta\{c_{\hat{j}} - c'(z^k)[D(z^k)]^{-1}A^{\hat{j}}\} \\ &= c'z + \theta\{c_{\hat{j}} - \lambda'(z^k)A^{\hat{j}}\}i \ . \end{aligned} \tag{4.30}$$

But from step 2 $\{c\hat{j} - \lambda'(z^k)A^{\hat{j}}\} > 0$, so that $c'z(\theta) \to \infty$ as $\theta \to \infty$. ♢

Lemma 5: z^{k+1} is a basic feasible solution and $c'z^{k+1} > c'z^k$.
Proof: Let $\tilde{j} \in I(z^k)$ be such that $\gamma_{\tilde{j}}^k > 0$ and $z_{\tilde{j}}^k = \theta\gamma_{\tilde{j}}^k$. Then from (4.28) we see that $z_{\tilde{j}}^{k+1} = 0$, hence

$$I(z^{k+1}) \subset (I(z) - \{\tilde{j}\}) \bigcup \{\hat{j}\} \ , \tag{4.31}$$

so that it is enough to prove that $A^{\tilde{j}}$ is independent of $\{A^j | j \in I(z),\ j \neq \tilde{j}\}$. But if this is not the case, we must have $\gamma_{\tilde{j}}^k = 0$, giving a contradiction. Finally if we compare (4.28) and (4.29), we see from (4.30) that

$$c'z_{k+1} - c'z_k = \theta\{c_{\hat{j}} - \gamma'(z^k)A^{\hat{j}}\} \ ,$$

which is positive from Step 2. ♢

Corollary 2: In a finite number of steps Phase II will obtain an optimal solution or will determine that sup$\{c'x | x \in \Omega_p\} = \infty$.
Corollary 3: Suppose Phase II terminates at an optimal basic feasible solution z^*. Then $\gamma(z^*)$ is an optimal solution of the dual of (4.24).
Exercise 2: Prove Corollaries 2 and 3.
Remark 1: By the non-degeneracy assumption, $I(z^{k+1})$ has m elements, so that in (4.31) we must have equality. We see then that $D(z^{k+1})$ is obtained from $D(z^k)$ by replacing the column A^j by

the column $A^{\hat{j}}$. More precisely if $D(z^k) = [A^{j_1}\vdots\ldots\vdots A^{j_{i-1}}\vdots A^{\tilde{j}}\vdots A^{j_{i+1}}\vdots\ldots\vdots A^{j_m}]$ and if $j_k < \hat{j} < j_{k+1}$ then $D(z^{k+1}) = [A^{j_1}\vdots\ldots\vdots A^{j_{i-1}}\vdots A^{j_{i+1}}\vdots\ldots\vdots A^{j_k}\vdots A^{\hat{j}}\vdots A^{j_{k+1}}\vdots\ldots\vdots A^{j_m}]$. Let E be the matrix $E = [A^{j_1}\vdots\ldots\vdots A^{j_{i-1}}\vdots A^{\hat{j}}\vdots A^{j_{i+1}}\vdots\ldots\vdots A^{j_m}]$. Then $[D(z^{k+1})]^{-1} = P\,E^{-1}$ where the matrix P permutes the columns of $D(z^{k+1})$ such that $E = D(z^{k+1})P$. Next, if $A^{\hat{j}} = \sum_{\ell=1}^{m} \gamma_{j_\ell} A^{j_\ell}$, it is easy to check that $E^{-1} = M[D(z^k)]^{-1}$ where

$$M = \begin{bmatrix} 1 & & & & \frac{-\gamma_{j_1}}{\gamma_{\hat{j}}} & & & \\ & 1 & & & & & & \\ & & \ddots & & & & & \\ & & & 1 & & & & \\ & & & & \frac{1}{\gamma_{\hat{j}}} & & & \\ & & & & & 1 & & \\ & & & & & & \ddots & \\ & & & & \frac{-\gamma_{j_m}}{\gamma_{\hat{j}}} & & & 1 \end{bmatrix}$$

$\uparrow$

ith column

Then $[D(z^{k+1})]^{-1} = PM[D(z^k)]^{-1}$, so that these inverses can be easily computed.

Remark 2: The similarity between Step 2 of Phase II and Step 2 of the algorithm in 3.3.4 is striking. The basic variables at z^k correspond to the variables w^k and non-basic variables correspond to u^k. For each $j \notin I(z^k)$ we can interpret the number $c_j - \lambda'(z^k)A_j$ to be the net increase in the objective value per unit increase in the *jth* component of z^k. This net increase is due to the direct increase c_j minus the indirect decrease $\lambda'(z^k)A_j$ due to the compensating changes in the basic variables necessary to maintain feasibility. The analogous quantity in 3.3.4 is $(\partial f_0/\partial u_j)(x^k) - (\lambda^k)'(\partial f/\partial u_j)(x^k)$.
Remark 3: By eliminating any dependent equations in (4.24) we can guarantee that the matrix A has rank n. Hence at any degenerate basic feasible solution z^k we can always find $\bar{I}(z^k) \supset I(z^k)$ such that $\bar{I}(z^k)$ has m elements and $\{A_j | j \in \bar{I}(z^k)\}$ is a linearly independent set. We can apply Phase II using $\bar{I}(z^k)$ instead of $I(z^k)$. But then in Step 4 it may turn out that $\theta = 0$ so that $z^{k+1} = z^k$. The reason for this is that $\bar{I}(z^k)$ is not unique, so that we have to try various alternatives for $\bar{I}(z^k)$ until we find one for which $\theta > 0$. In this way the non-degeneracy assumption can be eliminated. For details see (Canon, *et al.*, [1970]).

We now describe how to obtain an initial basic feasible solution.

Phase I:
Step I. by multiplying some of the equality constraints in (4.24) by -1 if necessary, we can assume that $b \geq 0$. Replace the LP (4.24) by the LP (4.32) involving the variables x and y:

$$\begin{aligned} \text{Maximize} \quad & -\sum_{i=1}^{m} y_i \\ \text{subject to} \quad & a_{il}x_1 + \ldots + a_{in}x_n + y_i = b_i\ ,\ 1 \leq i \leq m\ , \\ & x_j \geq 0\ ,\ y_i \geq 0\ ,\ 1 \leq j \leq n\ , 1 \leq i \leq m\ . \end{aligned} \tag{4.32}$$

Go to step 2.

Step 2. Note that $(x^0, y^0) = (0, b)$ is a basic feasible solution of (4.32). Apply phase II to (4.32) starting with this solution. Phase II must terminate in an optimum based feasible solution (x^*, y^*) since the value of the objective function in (4.32) lies between $-\sum_{i=1}^{m} b_i$ and 0. Go to Step 3.

Step 3. If $y^* = 0,\ x^*$ is a basic feasible solution for (4.24). If $y^* \neq 0$, by Exercise 3 below, (4.24) has no feasible solution.
Exercise 3: Show that (4.24) has a feasible solution iff $y^* = 0$.

4.4 *LP Theory of a Firm in a Competitive Economy*

4.4.1 *Activity analysis of the firm.*

We think of a firm as a system which transforms input into outputs. There are m kinds of inputs and k kinds of outputs. Inputs are usually classified into raw materials such as iron ore, crude oil, or raw cotton; intermediate products such as steel, chemicals, or textiles; capital goods [3] such as machines of various kinds, or factory buildings, office equipment, or computers; finally various kinds of labor services. The firm's outputs themselves may be raw materials (if it is a mining company) or intermediate products (if it is a steel mill) or capital goods (if it manufactures lathes) or finished goods (if it makes shirts or bakes cookies) which go directly to the consumer. Labor is not usually considered an output since slavery is not practiced; however, it may be considered an output in a "closed," dynamic Malthusian framework where the increase in labor is a function of the output. (See the von Neumann model in (Nikaido [1968]), p. 141.)

Within the firm, this transformation can be conducted in different ways, *i.e.*, different combinations of inputs can be used to produce the same combination of outputs, since human labor can do the same job as some machines and machines can replace other kinds of machines, *etc.* This *substitutability among inputs* is a fundamental concept in economics. We formalize it by specifying which transformation possibilities are available to the firm.

By an *input vector* we mean any m-dimensional vector $r = (r_1, \ldots, r_m)'$ with $r \geq 0$, and by an *output vector* we mean any k-dimensional vector $y = (y_1, \ldots, y_k)'$ with $y \geq 0$. We now make three basic assumptions about the firm.
(i) The transformation of inputs into outputs is organized into a finite number, say n, of processes or *activities*.
(ii) Each activity combines the k inputs in *fixed* proportions into the m outputs in *fixed* proportions. Furthermore, each activity can be conducted at any non-negative intensity or *level*. Precisely, the *jth* activity is characterized completely by two vectors $A^j = (a_{1j}, a_{2j}, \ldots, a_{mj})'$ and $B^j = (b_{ij}, \ldots, b_{kj})'$ so that if it is conducted at a level $x_j \geq 0$, then it combines (transforms) the input vector $(a_{1j}x_j, \ldots, a_{mj}x_j)' = x_jA^j$ into the output vector $(b_{1j}x_j, \ldots, b_{kj}x_j)' = x_jB^j$. Let A be the $m \times n$ matrix $[A^1 \vdots \ldots \vdots A^n]$ and B be the $k \times n$ matrix $B = [B^1 \vdots \ldots \vdots B^n]$.

[3] It is more accurate to think of the services of capital goods rather than these goods themselves as inputs. It is these services which are consumed in the transformation into outputs.

(iii) If the firm conducts all the activities simultaneously with the *jth* activity at level $x_j \geq 0, 1 \leq j \leq n$, then it transforms the input vector $x_1A^1 + \ldots + x_nA^n$ into the output vector $x_1B^1 + \ldots + x_nB^n$.

With these assumptions we know all the transformations technically possible as soon as we specify the matrices A and B. Which of these possible transformations will actually take place depends upon their relative profitability and availability of inputs. We study this next.

4.4.2 *Short-term behavior.*

In the short-term, the firm cannot change the amount available to it of some of the inputs such as capital equipment, certain kinds of labor, and perhaps some raw materials. Let us suppose that these inputs are $1, 2, \ldots, \ell$ and they are available in the amounts $r_1^*, \ldots, r_\ell^*$, whereas the supply of the remaining inputs can be varied. We assume that the firm is operating in a competitive economy which means that the unit prices $p = (p_1, \ldots, p_k)'$ of the outputs, and $q = (q_1, \ldots, q_m)'$ of the inputs is fixed. Then the manager of the firm, if he is maximizing the firm's profits, faces the following decision problem:

$$\begin{aligned} &\text{Maximize} \quad p'y - \sum_{j=\ell+1}^{m} q_j r_j \\ &\text{subject to} \quad y = Bx, \\ &\qquad a_{i1}x_1 + \ldots + a_{in}x_n \leq r_i^* \ , \ 1 \leq i \leq \ell \ , \\ &\qquad a_{i1}x_1 + \ldots + a_{in}x_n \leq r_i \ , \ \ell + 1 \leq i \leq m \ , \\ &\qquad x_j \geq 0, \ 1 \leq j \leq n; \ r_i \geq 0 \ , \ \ell + 1 \leq i \leq m \ . \end{aligned} \tag{4.33}$$

The decision variables are the activity levels $x_1, \ldots, x_n$, and the short-term input supplies $r_{\ell+1}, \ldots, r_m$. The coefficients of B and A are the fixed *technical coefficients* of the firm, the r_i^* are the fixed short-term supplies, whereas the p_i, q_j are prices determined by the whole economy, which the firm accepts as given. Under realistic conditions (4.33) has an optimal solution, say, $x_1^*, \ldots, x_n^*, r_{\ell+1}^*, \ldots, r_m^*$.

4.4.3 *Long-term equilibrium behavior.*

In the long run the supplies of the first ℓ inputs are also variable and the firm can change these supplies from $r_1^*, \ldots, r_\ell^*$ by buying or selling these inputs at the market price $q_1, \ldots, q_\ell$. Whether the firm will actually change these inputs will depend upon whether it is profitable to do so, and in turn this depends upon the prices p, q. We say that the prices (p^*, q^*) and a set of input supplies $r^* = (r_1^*, \ldots, r_m^*)$ are in (long-term) *equilibrium* if the firm has no profit incentive to change r^* under the prices (p^*, q^*).

Theorem 1: p^*, q^*, r^* are in equilibrium if and only if q^* is an optimal solution of (4.34):

$$\begin{aligned} &\text{Minimize} \ (r^*)'q \\ &\text{subject to} \ A'q \geq B'p^* \\ &\qquad\qquad q \geq 0 \ . \end{aligned} \tag{4.34}$$

Proof: Let $c = B'p^*$. By definition, p^*, q^*, r^* are in equilibrium iff for all fixed $\Delta \in R^m$, $M(\Delta) \leq M(0)$ where $M(\Delta)$ is the maximum value of the LP (4.35):

$$\begin{aligned} &\text{Maximize} \ c'x - (q^*)'\Delta \\ &\text{subject to} \ Ax \leq r^* + \Delta \ , \\ &\qquad\qquad x \geq 0 \ . \end{aligned} \tag{4.35}$$

For $\Delta = 0$, (4.34) becomes the dual of (4.35) so that by the strong duality theorem, $M(0) = (r^*)'q^*$. Hence p^*, q^*, r^* are in equilibrium iff

$$c'x - (q^*)'\Delta \leq M(0) = (r^*)'q^* \ , \tag{4.36}$$

whenever x is feasible for (4.35). By weak duality if x is feasible for (4.35) and q is feasible for (4.34),

$$c'x - (q^*)'\Delta \leq q'(r^* = \Delta) - (q^*)'\Delta \ , \tag{4.37}$$

and, in particular, for $q = q^*$,

$$c'x - (q^*)'\Delta \leq (q^*)'(r^* + \Delta) - (q^*)'\Delta = (q^*)'r^* \qquad \diamondsuit$$

Remark 1: We have shown that $(p^*, q^*, r^*$ are in long-term equilibrium iff q^* is an optimum solution to the dual (namely (4.34)) of (4.38):

$$\begin{aligned} &\text{Maximize } c'x \\ &\text{subject to } Ax \leq r^* \\ &\qquad\qquad x \geq 0 \ . \end{aligned} \tag{4.38}$$

This relation between p^*, q^*, r^* has a very nice economic interpretation. Recall that $c = B'p^*$, *i.e.*, $c_j = p_1^* b_{1j} + p_2^* b_{2j} + \ldots + p_k^* b_{kj}$. Now b_{ij} is the amount of the *ith* output produced by operating the *jth* activity at a unit level $x_j = 1$. Hence, c_j is the revenue per unit level operation of the *jth* activity so that $c'x$ is the revenue when the n activities are operated at levels x. On the other hand if the *jth* activity is operated at level $x_j = 1$, it uses an amount a_{ij} of the *ith* input. If the *ith* input is valued at a_i^*, then the input cost of operating at $x_j = 1$, is $\sum_{i=1}^{m} q_i a_{ij}$, so that the input cost of operating the n activities at levels x is $(A'q^*)' = (q^*)'Ax$. Thus, if x^* is the optimum activity levels for (4.38) then the output revenue is $c'x^*$ and the input cost is $(q^*)'Ax^*$. But from (4.16), $(q^*)'(Ax^* - r^*) = 0$ so that

$$c'x^* = (q^*)'r^* \ , \tag{4.39}$$

i.e., at the optimum activity levels, in equilibrium, total revenues = total cost of input supplies. In fact, we can say even more. From (4.15) we see that if $x^{=}ast_j > 0$ then

$$c_j = \sum_{i=1}^{m} q_i^* a_{ij} \ ,$$

i.e., at the optimum, the revenue of an activity operated at a positive level = input cost of that activity. Also if

$$c_j < \sum_{i=1}^{m} q_i^* a_{ij},$$

then $x_j^* = 0$, *i.e.*, if the revenue of an activity is less than its input cost, then at the optimum it is operated at zero level. Finally, again from (4.15), if an equilibrium the optimum *ith* input supply r_i^* is greater than the optimum demand for the *ith* input,

$$r_i^* > \sum_{j=1}^{n} a_{ij} x_j^*,$$

then $q_i^* = 0$, *i.e.*, the equilibrium price of an input which is in excess supply must be zero, in other words it must be a free good.

Remark 2: Returning to the short-term decision problem (4.33), suppose that $(\lambda_1^*, \ldots, \lambda_\ell^*, \lambda_{\ell+1}^*, \ldots, \lambda_m^*)$ is an optimum solution of the dual of (4.33). Suppose that the market prices of inputs $1, \ldots, \ell$ are $q_1, \ldots, q_\ell$. Let us denote by $M(\Delta_1, \ldots, \Delta_\ell)$ the optimum value of (4.33) when the amounts of the inputs in fixed supply are $r_1^* + \Delta_1, \ldots, r_\ell^* + \Delta_\ell$. Then if $(\partial M/\partial \Delta_i)|_{\Delta=0}$ exists, we can see from (4.22) that it is always profitable to increase the *ith* input by buying some additional amount at price q_i if $\lambda_i^* > q_i$, and conversely it is profitable to sell some of the *ith* input at price q_i if $\lambda_i^* < q_i$. Thus λ_i^* can be interpreted as the firm's internal valuation of the *ith* input or the firm's *imputed or shadow price* of the *ith* input. This interpretation has wide applicability, which we mention briefly. Often engineering design problems can be formulated as LPs of the form (4.10) or (4.19), where some of the coefficients b_i are design parameters. The design procedure is to fix these parameters at some nominal value b_i^*, and carry out the optimization problem. Suppose the resulting optimal dual variables are λ_i^*. then we see (assuming differentiability) that it is worth increasing b_i^* if the unit cost of increasing this parameter is less than λ_i^*, and it is worth decreasing this parameter if the reduction in total cost per unit decrease is greater than λ_i^*.

4.4.4 *Long-term equilibrium of a competitive, capitalist economy.*

The profit-maximizing behavior of the firm presented above is one of the two fundamental building blocks in the equilibrium theory of a competitive, capitalist economy. Unfortunately we cannot present the details here. We shall limit ourselves to a rough sketch. We think of the economy as a feedback process involving firms and consumers. Let us suppose that there are a total of h commodities in the economy including raw materials, intermediate and capital goods, labor, and finished products. By adding zero rows to the matrices (A, B) characterizing a firm we can suppose that all the h commodities are possible inputs and all the h commodities are possible outputs. Of course, for an individual firm most of the inputs and most of the outputs will be zero. the sole purpose for making this change is that we no longer need to distinguish between prices of inputs and prices of outputs. We observe the economy starting at time T. At this time there exists within the economy an inventory of the various commodities which we can represent by a vector $\omega = (\omega_1, \ldots, \omega_h) \geq 0$. ω is that portion of the outputs produced prior to T which have not been consumed up to T. We are assuming that this is a capitalist economy, which means that the ownership of ω is divided among the various consumers $j = 1, \ldots, J$. More precisely, the *jth* consumer owns the vector of commodities $\omega(j) \geq 0$, and $\sum_{j=1}^{J} \omega(j) = \omega$. We are including in $\omega(j)$ the amount of his labor services which consumer j is willing to sell. Now suppose that at time T the prevailing prices of the h commodities are $\lambda = (\lambda_1, \ldots, \lambda_h)' \geq 0$. Next, suppose that the managers of the various firms assume that the prices λ are not going to change for a long period of time. Then, from our previous analysis we know that the manager of the *ith* firm will plan to buy input supplies $r(i) \geq 0,\ r(i) \in R^h$, such

that $(\lambda, r(i))$ is in long term equilibrium, and he will plan to produce an optimum amount, say $y(i)$. Here $i = 1, 2, \ldots, I$, where I is the total number of firms. We know that $r(i)$ and $y(i)$ depend on λ, so that we explicitly write $r(i, \lambda),\ y(i, \lambda)$. We also recall that (see (4.38))

$$\lambda' r(i, \lambda) = \lambda' y(i, \lambda)\ ,\ 1 \leq i \leq I\ . \tag{4.40}$$

Now the *ith* manager can buy $r(i)$ from only two sources: outputs from other firms, and the consumers who collectively own ω. Similarly, the *ith* manager can sell his planned output $y(i)$ either as input supplies to other firms or to the consumers. Thus, the net supply offered for sale to consumers is $S(\lambda)$, where

$$S(\lambda) = \sum_{j=1}^{J} \omega(j) + \sum_{i=1}^{I} y(i, \lambda) - \sum_{i=1}^{i} r(i, \lambda)\ . \tag{4.41}$$

We note two important facts. First of all, from (4.40), (4.41) we immediately conclude that

$$\lambda' S(\lambda) = \sum_{j=1}^{J} \lambda' \omega(j)\ , \tag{4.42}$$

that is the value of the supply offered to consumers is equal to the value of the commodities (and labor) which they own. The second point is that there is no reason to expect that $S(\lambda) \geq 0$.

Now we come to the second building block of equilibrium theory. The value of the *jth* consumer's possessions is $\lambda' \omega(j)$. The theory assumes that he will plan to buy a set of commodities $d(j) = (d_1(j), \ldots, d_h(j)) \geq 0$ so as to maximize his satisfaction subject to the constraint $\lambda' d(j) = \lambda' \omega(j)$. Here also $d(j)$ will depend on λ, so we write $d(j, \lambda)$. If we add up the buying plans of all the consumers we obtain the total demand

$$D(\lambda) = \sum_{j=1}^{J} d(j, \lambda) \geq 0\ , \tag{4.43}$$

which also satisfies

$$\lambda' D(\lambda) = \sum_{j=1}^{J} \lambda' \omega(j)\ . \tag{4.44}$$

The most basic question of equilibrium theory is to determine conditions under which there exists a price vector λ_E such that the economy is in equilibrium, *i.e.*, $S(\lambda_E) = D(\lambda_E)$, because if such an equilibrium price λ_E exists, then at that price the production plans of all the firms and the buying plan of all the consumers can be realized. Unfortunately we must stop at this point since we cannot proceed further without introducing some more convex analysis and the fixed point theorem. For a simple treatment the reader is referred to (Dorfman, Samuelson, and Solow [1958], Chapter 13). For a much more general mathematical treatment see (Nikaido [1968], Chapter V).

4.5 *Miscellaneous Comments*

4.5.1 *Some mathematical tricks.*

It is often the case in practical decision problems that the objective is not well-defined. There may be a number of plausible objective functions. In our LP framework this situation can be formulated as follows. The constraints are given as usual by $Ax \leq b,\ x \geq 0$. However, there are, say, k objective functions $(c^1)'x, \ldots, (c^k)'x$. It is reasonable then to define a single objective function $f_0(x)$ by $f_0(x)$ = minimum $\{(c^1)'x,\ (c^2)'x, \ldots, (c^k)'x\}$, so that we have the decision problem,

$$\begin{array}{l} \text{Maximize } f_0(x) \\ \text{subject to } Ax \leq b,\ x \geq 0\ . \end{array} \tag{4.45}$$

This is *not* a LP since f_0 is *not* linear. However, the following exercise shows how to transform (4.45) into an equivalent LP.

Exercise 1: Show that (4.45) is equivalent to (4.46) below, in the sense that x^* is optimal for (4.45) iff $(x^*, y^*) = (x^*, f_0(x^*))$ is optimal for (4.46).

$$\begin{array}{l} \text{Maximize } y \\ \text{subject to } Ax \leq b,\ x \leq 0 \\ \qquad\qquad y \leq (c^i)'x\ ,\ 1 \leq i \leq k\ . \end{array} \tag{4.46}$$

Exercise 1 will also indicate how to do Exercise 2.
Exercise 2: Obtain an equivalent LP for (4.47):

$$\begin{array}{l} \text{Maximize } \displaystyle\sum_{j=1}^{n} c_i(x_i) \\ \text{subject to } Ax \leq b,\ x \leq 0\ , \end{array} \tag{4.47}$$

where $c_i :\ R \to R$ are concave, piecewise-linear functions of the kind shown in Figure 4.3.

The above-given assumption of the concavity of the c_i is crucial. In the next exercise, the interpretation of "equivalent" is purposely left ambiguous.

Exercise 3: Construct an example of the kind (4.47), where the c_i are piecewise linear (but not concave), and such that there is no equivalent LP.

It turns out however, that even if the c_i are not concave, an elementary modification of the Simplex algorithm can be given to obtain a "local" optimal decision. See (Miller [1963]).

4.5.2 *Scope of linear programming.*

LP is today the single most important optimization technique. This is because many decision problems can be adequately formulated as LPs, and, given the capabilities of modern computers, the Simplex method (together with its variants) is an extremely powerful technique for solving LPs involving thousands of variables. To obtain a feeling for the scope of LP we refer the reader to the book by one of the originators of LP (Dantzig [1963]).

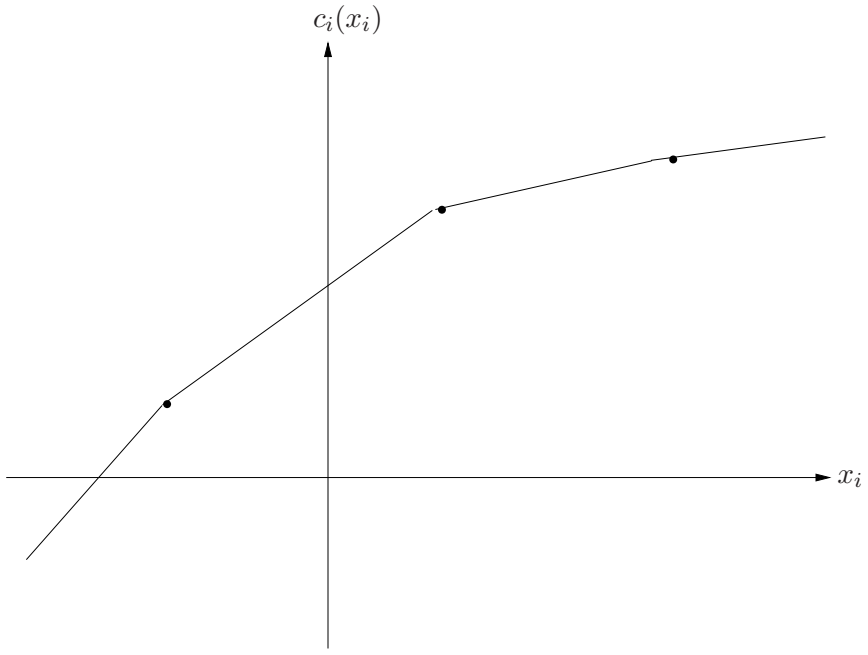


Figure 4.3: A function of the form used in Exercise 2.

Chapter 5

OPTIMIZATION OVER SETS DEFINED BY INEQUALITY CONSTRAINTS: NONLINEAR PROGRAMMING

In many decision-making situations the assumption of linearity of the constraint inequalities in LP is quite restrictive. The linearity of the objective function is not restrictive as shown in the first exercise below. In Section 1 we present the general nonlinear programming problem (NP) and prove the Kuhn-Tucker theorem. Section 2 deals with Duality theory for the case where appropriate convexity conditions are satisfied. Two applications are given. Section 3 is devoted to the important special case of quadratic programming. The last section is devoted to computational considerations.

5.1 *Qualitative Theory of Nonlinear Programming*

5.1.1 *The problem and elementary results.*

The general NP is a decision problem of the form:

$$\begin{array}{l} \text{Maximize } f_0(x) \\ \text{subject to } (x) \leq 0 \ , \ i = 1, \ldots, m, \end{array} \tag{5.1}$$

where $x \in R^n$, $f_i : R^n \to R, \ i = 0, 1, \ldots, m$, are differentiable functions. As in Chapter 4, $x \in R^n$ is said to be a *feasible solution* if it satisfies the constraints of (5.1), and $\Omega \subset R^n$ is the subset of all feasible solutions; $x^* \in \Omega$ is said to be an *optimal decision* or *optimal solution* if $f_0(x^*) \geq f_0(x)$ for $x \in \Omega$. From the discussion in 4.1.2 it is clear that equality constraints and sign constraints on some of the components of x can all be transformed into the form (5.1). The next exercise shows that we could restrict ourselves to objective functions which are linear; however, we will not do this.

Exercise 1: Show that (5.2), with variables $y \in R, x \in R^n$, is equivalent to (5.1):

$$\begin{aligned} &\text{Maximize } y \\ &\text{subject to } f_i(x) \leq 0,\ 1 \leq i \leq m, \text{ and } y - f_0(x) \leq 0\ . \end{aligned} \tag{5.2}$$

Returning to problem (5.1), we are interested in obtaining conditions which any optimal decision must satisfy. The argument parallels very closely that developed in Exercise 1 of 4.1 and Exercise 1 of 4.2. The basic idea is to linearize the functions f_i in a neighborhood of an optimal decision x^*.

Definition: Let x be a feasible solution, and let $I(x) \subset \{1, \ldots, m\}$ be such that $f_i(x) = 0$ for $\imath \in I(x),\ f_i(x) < 0$ for $i \notin I(x)$. (The set $I(x)$ is called the set of *active* constraints at x.)
Definition: (i) Let $x \in \Omega$. A vector $h \in R^n$ is said to be an *admissible direction for Ω at x* if there exists a sequence $x^k,\ k = 1, 2, \ldots$, in Ω and a sequence of numbers $\varepsilon^k,\ k = 1, \ldots$, with $\varepsilon^k > 0$ for all k such that

$$\lim_{k\to\infty} x^k = x\ ,$$

$$\lim_{k\to\infty} \tfrac{1}{\varepsilon^k}(x^k - x) = h\ \ .$$

(ii) Let $C(\Omega, x) = \{h | h$ is an admissible direction for Ω at $x\}$. $C(\Omega, x)$ is called the *tangent cone of Ω at x*. Let $K(\Omega, x) = \{x + h | h \in C(\Omega, x)\}$. (See Figures 5.1 and 5.2 and compare them with Figures 4.1 and 4.2.)

If we take $x^k = x$ and $\varepsilon^k = 1$ for all k, we see that $0 \in C(\Omega, x)$ so that the tangent cone is always nonempty. Two more properties are stated below.

Exercise 2: (i) Show that $C(\Omega, x)$ is a *cone, i.e.,* if $h \in C(\Omega, x)$ and $\theta \geq 0$, then $\theta h \in C(\Omega, x)$. (ii) Show that $C(\Omega, x)$ is a closed subset of R^n. (Hint for (ii): For $m = 1, 2, \ldots$, let h^m and $\{x^{mk},\ \varepsilon^{mk} > 0\}_{k=1}^{\infty}$ be such that $x^{mk} \to x$ and $(1/\varepsilon^{mk})(x^{mk} - x) \to h^m$ as $k \to \infty$. Suppose that $h^m \to h$ as $m \to \infty$. Show that there exist subsequences $\{x^{mk_m}, \varepsilon^{mk_m}\}_{m=1}^{\infty}$ such that $x^{mk_m} \to x$ and $(1/\varepsilon^{mk_m})(x^{mk_m} - x) \to h$ as $m \to \infty$.)

In the definition of $C(\Omega, x)$ we made no use of the particular functional description of Ω. The following elementary result is more interesting in this light and should be compared with (2.18) in Chapter 2 and Exercise 1 of 4.1.

Lemma 1: Suppose $x^* \in \Omega$ is an optimum decision for (5.1).
Then

$$f_{0x}(x^*)h \leq 0 \text{ for all } h \in C(\Omega, x^*)\ \ . \tag{5.3}$$

Proof: Let $x^k \in \Omega,\ \varepsilon^k > 0,\ k = 1, 2, 3, \ldots$, be such that

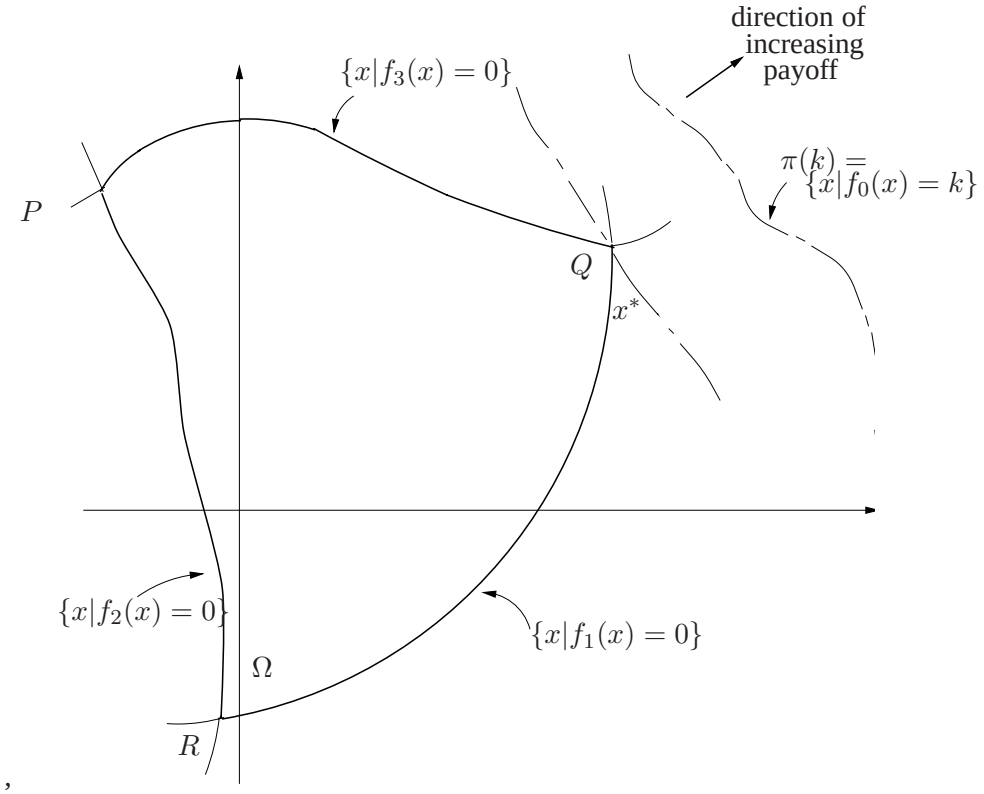


Figure 5.1: $\Omega = PQR$

$$\lim_{k\to\infty} x^k = x^* \quad , \quad \lim_{k\to\infty} \tfrac{1}{\varepsilon^k}(x^k - x^*) = h \quad . \tag{5.4}$$

Note that in particular (5.4) implies

$$\lim_{k\to\infty} \tfrac{1}{\varepsilon^k} \quad |x^k - x^*| \;=\; |h| \quad . \tag{5.5}$$

Since f_0 is differentiable, by Taylor's theorem we have

$$f_0(x^k) = f_0(x^* + (x^k - x^*)) = f_0(x^*) + f_{0x}(x^*)(x^k - x^*) + o(|x^k - x^*|) \; . \tag{5.6}$$

Since $x^k \in \Omega$, and x^* is optimal, we have $f_0(x^k) \leq f_0(x^*)$, so that

$$0 \geq f_{0x}(x^*)\tfrac{(x^k - x^*)}{\varepsilon^k} + \tfrac{o(|x^k - x^*|)}{\varepsilon^k} \quad .$$

Taking limits as $k \to \infty$, using (5.4) and (5.5), we can see that

$$\begin{aligned} 0 \geq & \lim_{k\to\infty} \quad f_{0x}(x^*) \quad \tfrac{(x^k - x^*)}{\varepsilon^k} \quad + \quad \lim_{k\to\infty} \quad \tfrac{o(|x^k - x^*|)}{|x^k - x^*|} \quad \lim_{k\to\infty} \quad \tfrac{|x^k - x^*|}{\varepsilon^k} \\ = & \quad f_{0x}(x^*)h. \quad \diamondsuit \end{aligned}$$

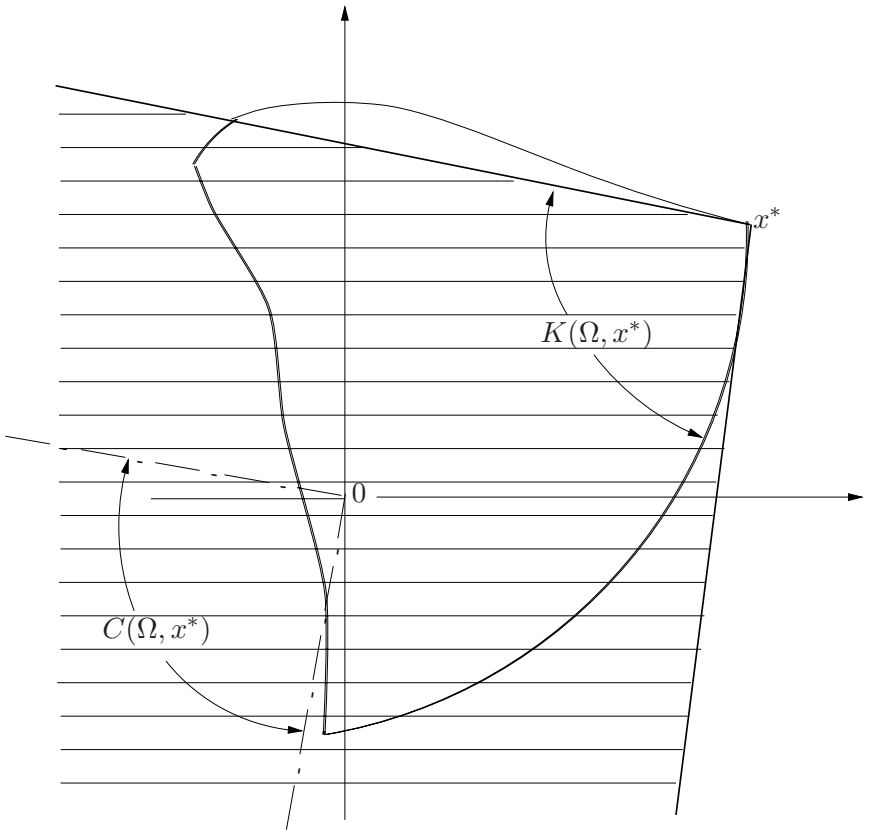


Figure 5.2: $C(\Omega, x^*)$ is the tangent cone of Ω at x^*.

The basic problem that remains is to characterize the set $C(\Omega, x^*)$ in terms of the derivatives of the functions f_i. Then we can apply Farkas' Lemma just as in Exercise 1 of 4.2.

Lemma 2: Let $x^* \in \Omega$. Then

$$C(\Omega, x^*) \subset \{h | f_{ix}(x^*)h \leq 0 \text{ for all } i \in I(x^*)\} \ . \tag{5.7}$$

Proof: Let $h \in R^n$ and $x^k \in \Omega,\ \varepsilon^k > 0,\ k = 1, 2, \ldots$, satisfy (5.4). Since f_i is differentiable, by Taylor's theorem we have

$$f_i(x^k) = f_i(x^*) + f_{ix}(x^*)(x^k - x^*) + o(|x^k - x^*|) \ .$$

Since $x^k \in \Omega,\ f_i(x^k) \leq 0$, and if $i \in I(x^*),\ f_i(x^*) = 0$, so that $f_i(x^k) \leq f_i(x^*)$. Following the proof of Lemma 1 we can conclude that $0 \geq f_{ix}(x^*)h$. $\diamondsuit$

Lemma 2 gives us a partial characterization of $C(\Omega, x^*)$. Unfortunately, in general the inclusion sign in (5.7) cannot be reversed. The main reason for this is that the set $\{f_{ix}(x^*) | i \in I(x^*)\}$ is not in general linearly independent.
Exercise 3: Let $x \in R^2,\ f_1(x_1, x_2) = (x_1 - 1)^3 + x_2$, and $f_2(x_1, x_2) = -x_2$. Let $(x_1^*, x_2^*) = (1, 0)$. Then $I(x^*) = \{1, 2\}$. Show that

$$C(\Omega, x^*) \neq \{h | f_{ix}(x^*)h \leq 0 \ , \ \ i = 1, 2, \}.$$

(Note that $\{f_{1x}(x^*), f_{2x}(x^*)\}$ is not a linearly independent set; see Lemma 4 below.)

5.1.2 *Kuhn-Tucker Theorem.*

Definition: Let $x^* \in \Omega$. We say that the *constraint qualification* (CQ) is satisfied at x^* if

$$C(\Omega, x) = \{h | f_{ix}(x^*)h \leq 0 \text{ for all } i \in I(x^*)\},$$

and we say that CQ is satisfied if CQ is satisfied at all $x \in \Omega$. (Note that by Lemma 2 $C(\Omega, x)$ is always a subset of the right-hand side.)

Compare the next result with Exercise 2 of 4.2.

Theorem 1: (Kuhn and Tucker [1951]) Let x^* be an optimum solution of (5.1), and suppose that CQ is satisfied at x^*. Then there exist $\lambda_i^* \geq 0$, for $i \in I(x^*)$, such that

$$f_{0x}(x^*) = \sum_{i \in I(x^*)} \lambda_i^* f_{ix}(x^*) \tag{5.8}$$

Proof: By Lemma 1 and the definition of CQ it follows that $f_{0x}(x^*)h \leq 0$ whenever $f_{ix}(x^*)h \leq 0$ for all $i \in I(x^*)$. By the Farkas' Lemma of 4.2.1 it follows that there exist $\lambda_i^* \geq 0$ for $i \in I(x^*)$ such that (5.8) holds. $\diamondsuit$

In the original formulation of the decision problem we often have equality constraints of the form $r_j(x) = 0$, which get replaced by $r_j(x) \leq 0, -r_j(x) \leq 0$ to give the form (5.1). It is convenient in application to separate the equality constraints from the rest. Theorem 1 can then be expressed as Theorem 2.

Theorem 2: Consider the problem (5.9).

$$\begin{aligned} &\text{Maximize } f_0(x) \\ &\text{subject to } f_i(x) \leq 0 \ , \ i = 1, \ldots, m, \\ &\qquad\qquad r_j(x) = 0 \ , \ j = 1, \ldots, k \ \ . \end{aligned} \tag{5.9}$$

Let x^* be an optimum decision and suppose that CQ is satisfied at x^*. Then there exist $\lambda_i^* \geq 0, \ i = 1, \ldots, m$, and $\mu_j^*, \ j = 1, \ldots, k$ such that

$$f_{0x}(x^*) = \sum_{i=1}^{m} \lambda_i^* f_{ix}(x^*) + \sum_{j=1}^{k} \mu_j^* r_{jx}(x^*) \ , \tag{5.10}$$

and

$$\lambda_i^* = 0 \text{ whenever } f_i(x^*) < 0 \ . \tag{5.11}$$

Exercise 4: Prove Theorem 2.

An alternative form of Theorem 1 will prove useful for computational purposes (see Section 4).
Theorem 3: Consider (5.9), and suppose that CQ is satisfied at an optimal solution x^*. Define $\psi:\ R^n \to R$ by

$$\psi(h) = \max\left\{-f_{0x}(x^*)h, f_1(x^*) + f_{1x}(x^*)h, \ldots, f_m(x^*) + f_{mx}(x^*)h\right\}\ ,$$

and consider the decision problem

$$\begin{aligned} &\text{Minimize } \psi(h) \\ &\text{subject to } -\psi(h) - f_{0x}(x^*)h \le 0, \\ &\quad -\psi(h) + f_i(x^*) + f_{ix}(x^*)h \le 0\ ,\ 1 \le i \le m \\ &\qquad -1 \le h_i \le 1\ ,\ i = 1, \ldots, n\ . \end{aligned} \tag{5.12}$$

Then $h = 0$ is an optimal solution of (5.12).

Exercise 5: Prove Theorem 3. (Note that by Exercise 1 of 4.5, (5.12) can be transformed into a LP.)
Remark: For problem (5.9) define the *Lagrangian function* L:

$$(x_1, \ldots, x_n; \lambda_1, \ldots, \lambda_m; \mu_1, \ldots, \mu_k) \mapsto f_0(x) - \sum_{i=1}^{m} \lambda_i f_i(x) - \sum_{j=1}^{k} \mu_j r_j(x).$$

Then Theorem 2 is equivalent to the following statement: if CQ is satisfied and x^* is optimal, then there exist $\lambda^* \ge 0$ and μ^* such that $L_x(x^*, \lambda^*, \mu^*) = 0$ and $L(x^*, \lambda^*, \mu^*) \le L(x^*, \lambda, \mu)$ for all $\lambda \ge 0, \mu$.

There is a very important special case when the necessary conditions of Theorem 1 are also sufficient. But first we need some elementary properties of convex functions which are stated as an exercise. Some additional properties which we will use later are also collected here.
Recall the definition of convex and concave functions in 4.2.3.

Exercise 6: Let $X \subset R^n$ be convex. Let $h : X \to R$ be a differentiable function. Then
(i) h is convex iff $h(y) \ge h(x) + h_x(x)(y - x)$ for all x, y, in X,
(ii) h is concave iff $h(y) \le h(x) + h_x(x)(y - x)$ for all x, y in X,
(iii) h is concave and convex iff h is *affine, i.e.* $h(x) \equiv \alpha + b'x$ for some fixed $\alpha \in R,\ b \in R^n$.
Suppose that h is twice differentiable. Then
(iv) h is convex iff $h_{xx}(x)$ is positive semidefinite for all x in X,
(v) h is concave iff $h_{xx}(x)$ is negative semidefinite for all x in X,
(vi) h is convex and concave iff $h_{xx}(x) \equiv 0$.
Theorem 4: (Sufficient condition) In (5.1) suppose that f_0 is concave and f_i is convex for $i = 1, \ldots, m$. Then
(i) Ω is a convex subset of R^n, and
(ii) if there exist $x^* \in \Omega,\ \lambda_i^* \ge 0,\ i \in I(x^*)$, satisfying (5.8), then x^* is an optimal solution of (5.1).
Proof:
(i) Let y, z be in Ω so that $f_i(y) \le 0,\ f_i(z) \le 0$ for $i = 1, \ldots, m$. Let $0 \le \theta \le 1$. Since f_i is convex we have

$$f_i(\theta y + (1-\theta)z) \leq \theta f_i(y) + (1-\theta) f_i(z) \leq 0\ ,\ 1 \leq i \leq m,$$

so that $(\theta y + (1-\theta)z) \in \Omega$, hence Ω is convex.
(ii) Let $x \in \Omega$ be arbitrary. Since f_0 is concave, by Exercise 6 we have

$$f_0(x) \leq f_0(x^*) + f_{0x}(x^*)(x - x^*)\ \ ,$$

so that by (5.8)

$$f_0(x) \leq f_0(x^*) + \sum_{i \in I(x^*)} \lambda_i^* f_{ix}(x^*)(x - x^*)\ \ . \tag{5.13}$$

Next, f_i is convex so that again by Exercise 6,

$$f_i(x) \geq f_i(x^*) + f_{ix}(x^*)(x - x^*)\ \ ;$$

but $f_i(x) \leq 0$, and $f_i(x^*) = 0$ for $i \in I(x^*)$, so that

$$f_{ix}(x^*)(x - x^*) \leq 0 \ \text{ for } \ i \in I(x^*)\ . \tag{5.14}$$

Combining (5.14) with the fact that $\lambda_i^* \geq 0$, we conclude from (5.13) that $f_0(x) \leq f_0(x^*)$, so that x^* is optimal. $\diamondsuit$
Exercise 7: Under the hypothesis of Theorem 4, show that the subset Ω^* of Ω, consisting of all the optimal solutions of (5.1), is a convex set.
Exercise 8: A function $h:\ X \to R$ defined on a convex set $X \subset R^n$ is said to be *strictly* convex if $h(\theta y + (1-\theta)z) < \theta h(y) + (1-\theta)h(z)$ whenever $0 < \theta < 1$ and y, z are in X with $y \neq z$. h is said to be strictly concave if $-h$ is strictly convex. Under the hypothesis of Theorem 4, show that an optimal solution to (5.1) is unique (if it exists) if either f_0 is strictly concave or if the $f_i,\ 1 \leq i \leq m$, are strictly convex. (Hint: Show that in (5.13) we have strict inequality if $x \neq x^*$.)

5.1.3 *Sufficient conditions for CQ.*

As stated, it is usually impractical to verify if CQ is satisfied for a particular problem. In this subsection we give two conditions which guarantee CQ. These conditions can often be verified in practice. Recall that a function $g:\ R^n \to R$ is said to be *affine* if $g(x) \equiv \alpha + b'x$ for some fixed $\alpha \in R$ and $b \in R^n$.

We adopt the formulation (5.1) so that

$$\Omega = \{x \in R^n | f_i(x) \leq 0\ \ ,\ \ 1 \leq i \leq m\}\ \ .$$

Lemma 3: Suppose $x^* \in \Omega$ and suppose there exists $h^* \in R^n$ such that for each $i \in I(x^*)$, either $f_{ix}(x^*)h^* < 0$, or $f_{ix}(x^*)h^* = 0$ and f_i is affine. Then CQ is satisfied at x^*.
Proof: Let $h \in R^n$ be such that $f_{ix}(x^*)h \leq 0$ for $i \in I(x^*)$. Let $\delta > 0$. We will first show that $(h + \delta h^*) \in C(\Omega, x^*)$. To this end let $\varepsilon^k > 0,\ k = 1, 2, \ldots$, be a sequence converging to 0 and set $x^k = x^* + \varepsilon^k(h + \delta h^*)$. Clearly x^k converges to x^*, and $(1/\varepsilon^k)(x^k - x^*)$ converges to $(h + \delta h^*)$. Also for $i \in I(x^*)$, if $f_{ix}(x^*)h < 0$, then

$$\begin{aligned} f_i(x^k) &= f_i(x^*) + \varepsilon^k f_{ix}(x^*)(h + \delta h^*) + o(\varepsilon^k |h + \delta h^*|) \\ &\leq \delta \varepsilon^k f_{ix}(x^*)h^* + o(\varepsilon^k |h + \delta h^*|) \\ &< 0 \ \text{ for sufficiently large } k\ , \end{aligned}$$

whereas for $i \in I(x^*)$, if f_i is affine, then

$$f_i(x^k) = f_i(x^*) + \varepsilon^k f_{ix}(x^*)(h + \delta h^*) \leq 0 \text{ for all } k \ .$$

Finally, for $i \notin I(x^*)$ we have $f_i(x^*) < 0$, so that $f_i(x^k) < 0$ for sufficiently large k. Thus we have also shown that $x^k \in \Omega$ for sufficiently large k, and so by definition $(h + \delta h^*) \in C(\Omega, x^*)$. Since $\delta > 0$ can be arbitrarily small, and since $C(\Omega, x^*)$ is a closed set by Exercise 2, it follows that $h \in C(\Omega, x^*)$. $\diamondsuit$

Exercise 9: Suppose $x^* \in \Omega$ and suppose there exists $\hat{x} \in R^n$ such that for each $i \in I(x^*)$, either $f_i(x^*) < 0$ and f_i is convex, or $f_i(\hat{x}) \leq 0$ and f_i is affine. Then CQ is satisfied at x^*. (Hint: Show that $h^* = \hat{x} - x^*$ satisfies the hypothesis of Lemma 3.)

Lemma 4: Suppose $x^* \in \Omega$ and suppose there exists $h^* \in R^n$ such that $f_{ix}(x^*)h^* \leq 0$ for $i \in I(x^*)$, and $\{f_{ix}(x^*) | i \in I(x^*),\ f_{ix}(x^*)h^* = 0\}$ is a linearly independent set. Then CQ is satisfied at x^*.

Proof: Let $h \in R^n$ be such that $f_{ix}(x^*)h \leq 0$ for all $i \in I(x^*)$. Let $\delta > 0$. We will show that $(h + \delta h^*) \in C(\Omega, x^*)$. Let $J_\delta = \{i | i \in I(x^*),\ f_{ix}(x^*)(h + \delta h^*) = 0\}$, consist of p elements. Clearly $J_\delta \subset J = \{i | i \in I(x^*),\ f_i x(x^*)h^* = 0\}$, so that $\{f_{ix}(x^*, u^*) | i \in J_\delta\}$ is linearly independent. By the Implicit Function Theorem, there exist $\rho > 0$, an open set $V \subset R^n$ containing $x^* = (w^*, u^*)$, and a differentiable function $g :\ U \to R^p$, where $U = \{u \in R^{n-p} | |u - u^*| < \rho\}$, such that

$$f_i(w, u) = 0,\ i \in J_\delta, \text{ and } (w, u) \in V$$

iff

$$u \in U, \text{ and } w = g(u) \ .$$

Next we partition h, h^* as $h = (\xi, \eta),\ h^* = (\xi^*, \eta^*)$ corresponding to the partition of $x = (w, u)$. Let $\varepsilon^k > 0, k = 1, 2 \ldots$, be any sequence converging to 0, and set $u^k = u^* + \varepsilon^k(\eta + \delta\eta^*),\ w^k = g(u^k)$, and finally $x^k = (s^k, u^k)$.

We note that u^k converges to u^*, so $w^k = g(u^k)$ converges to $w^* = g(u^*)$. Thus, x^k converges to x^*. Now $(1/\varepsilon^k)(x^k - x^*) = (1/\varepsilon^k)(w^k - w^*, u^k - u^*) = (1/\varepsilon^k)(g(u^k) - g(u^*),\ \varepsilon^k(\eta + \delta\eta^*))$. Since g is differentiable, it follows that $(1/\varepsilon^k)(x^k - x^*)$ converges to $(g_u(u^*)(\eta + \delta\eta^*), \eta + \delta\eta^*)$. But for $i \in J_\delta$ we have

$$0 = f_{ix}(x^*)(h + \delta h^*) = f_{iw}(x^*)(\xi + \delta\xi^*) + f_{iu}(x^*)(\eta + \delta\eta^*) \ . \tag{5.15}$$

Also, for $i \in J_\delta,\ 0 = f_i(g(u), u)$ for $u \in U$ so that $0 = f_{iw}(x^*)g_u(u^*) + f_{iu}(x^*)$, and hence

$$0 = f_{iw}(x^*)g_u(u^*)(\eta + \delta\eta^*) + f_{iu}(x^*)(\eta + \delta\eta^*) \ . \tag{5.16}$$

If we compare (5.15) and (5.16) and recall that $\{f_{iw}(x^*) | i \in J_\delta\}$ is a basis in R^p we can conclude that $(\xi + \delta\xi^*) = g_u(u^*)(\eta + \delta\eta^*)$ so that $(1/\varepsilon^k)(x^k - x^*)$ converges to $(h + h\delta h^*)$.

It remains to show that $x^k \in \Omega$ for sufficiently large k. First of all, for $i \in J_\delta,\ f_i(x^k) = f_i(g(u^k), u^k) = 0$, whereas for $i \notin J_\delta,\ i \in I(x^*)$,

$$\begin{aligned} f_i(x^k) &= f_i(x^*) + f_{ix}(x^*)(x^k - x^*) + o(|x^k - x^*|) \\ &\quad f_i(x^*) + \varepsilon^k f_{ix}(x^*)(h + \delta h^*) + o(\varepsilon^k) + o(|x^k - x^*|), \end{aligned}$$

and since $f_i(x^*) = 0$ whereas $f_{ix}(x^*)(h + \delta h^*) < 0$, we can conclude that $f_i(x^k) < 0$ for sufficiently large k. Thus, $x^k \in \Omega$ for sufficiently large k. Hence, $(h + \delta h^*) \in C(\Omega, x^*)$.

To finish the proof we note that $\delta > 0$ can be made arbitrarily small, and $C(\Omega, x^*)$ is closed by Exercise 2, so that $h \in C(\Omega, x^*)$. $\diamondsuit$

The next lemma applies to the formulation (5.9). Its proof is left as an exercise since it is very similar to the proof of Lemma 4.

Lemma 5: Suppose x^* is feasible for (5.9) and suppose there exists $h^* \in R^n$ such that the set $\{f_{ix}(x^*)|i \in I(x^*), f_{ix}(x^*)h^* = 0\} \bigcup \{r_{jx}(x^*)|j = 1, \ldots, k\}$ is linearly independent, and $f_{ix}(x^*)h^* \leq 0$ for $i \in I(x^*),\ r_{jx}(x^*)h^* = 0$ for $1 \leq j \leq k$. Then CQ is satisfied at x^*.

Exercise 10: Prove Lemma 5

5.2 *Duality Theory*

Duality theory is perhaps the most beautiful part of nonlinear programming. It has resulted in many applications within nonlinear programming, in terms of suggesting important computational algorithms, and it has provided many unifying conceptual insights into economics and management science. We can only present some of the basic results here, and even so some of the proofs are relegated to the Appendix at the end of this Chapter since they depend on advanced material. However, we will give some geometric insight. In 2.3 we give some application of duality theory and in 2.2 we refer to some of the important generalizations. The results in 2.1 should be compared with Theorems 1 and 4 of 4.2.1 and the results in 4.2.3.

It may be useful to note in the following discussion that most of the results do not require differentiability of the various functions.

5.2.1 *Basic results.*

Consider problem (5.17) which we call the *primal* problem:

$$\begin{aligned} &\text{Maximize } f_0(x) \\ &\text{subject to } f_i(x) \leq \hat{b}_i \ , \ \ 1 \leq i \leq m \\ &\qquad\qquad x \in X \ , \end{aligned} \tag{5.17}$$

where $x \in R^n,\ f_i : \ R^n \to R,\ 1 \leq i \leq m$, are given *convex* functions, $f_0 : \ R^n \to R$ is a given *concave* function, X is a given *convex* subset of R^n and $\hat{b} = (\hat{b}_1, \ldots, \hat{b}_m)'$ is a given vector. For convenience, let $f = (f_1, \ldots, f_m)' : \ R^n \to R^m$. We wish to examine the behavior of the maximum value of (5.17) as $\hat{b}$ varies. So we define

$$\Omega(b) = \{x|x \in X, f(x) \leq b\}, \ \ B = \{b|\Omega(b) \neq \phi\},$$

and

$$\begin{aligned} M : \ B \to R \bigcup \{+\infty\} \text{ by } M(b) &= \sup\{f_0(x)|x \in X, f(x) \leq b\} \\ &= \sup\{f_0(x)|x \in \Omega(b)\} \ , \end{aligned}$$

so that in particular if x^* is an optimal solution of (5.17) then $M(\hat{b}) = f_0(\hat{x})$. We need to consider the following problem also. Let $\lambda \in R^m,\ \lambda \geq 0$, be fixed.

$$\begin{aligned} &\text{Maximize } \ f_0(x) - \lambda'(f(x) - \hat{b}) \\ &\text{subject to } \ x \in X \ , \end{aligned} \tag{5.18}$$

and define

$$m(\lambda) = \text{sub}\{f_0(x) - \lambda'(f(x) - \hat{b})|x \in X\} \ .$$

Problem (5.19) is called the *dual* problem:

$$\begin{array}{l} \text{Minimize} \ \ m(\lambda) \\ \text{subject to} \ \ \lambda \geq 0 \ \ . \end{array} \tag{5.19}$$

Let m^* = inf $\{m(\lambda)|\lambda \geq 0\}$.
Remark 1: The set X in (5.17) is usually equal to R^n and then, of course, there is no reason to separate it out. However, it is sometimes possible to include some of the constraints in X in such a way that the calculation of $m(\lambda)$ by (5.18) and the solution of the dual problem (5.19) become simple. For example see the problems discussed in Sections 2.3.1 and 2.3.2 below.
Remark 2: It is sometimes useful to know that Lemmas 1 and 2 below hold *without* any convexity conditions on f_0, f, X. Lemma 1 shows that the cost function of the dual problem is convex which is useful information since there are computation techniques which apply to convex cost functions but not to arbitrary nonlinear cost functions. Lemma 2 shows that the optimum value of the dual problem is always an upper bound for the optimum value of the primal.
Lemma 1: $m: \ \ R^n_+ \rightarrow R \bigcup \{+\infty\}$ is a convex function. (Here $R^n_+ = \{\lambda \in R^n | \lambda \geq 0\}$.)

Exercise 1: Prove Lemma 1.

Lemma 2: (Weak duality) If x is feasible for (5.17), *i.e.*, $x \in \Omega(\hat{b})$, and if $\lambda \geq 0$, then

$$f_0(x) \leq M(\hat{b}) \leq m^* \leq m(\lambda) \ \ . \tag{5.20}$$

Proof: Since $f(x) - \hat{b} \leq 0$, and $\lambda \geq 0$, we have $\lambda'(f(x) - \hat{b}) \leq 0$. So,

$$f_0(x) \leq f_0(x) - \lambda'(f(x) - \hat{b}), \text{ for } x \in \Omega(\hat{b}), \ \lambda \geq 0 \ \ .$$

Hence

$$\begin{array}{rl} f_0(x) \leq & \text{sup} \ \{f_0(x)|x \in \Omega(\hat{b})\} = M(\hat{b}) \\ \leq & \text{sup} \ \{f_0(x) - \lambda'(f(x) - \hat{b})|x \in \Omega(\hat{b})\} \text{ and since } \Omega(\hat{b}) \subset X, \\ \leq & \text{sup} \ \{f_0(x) - \lambda'(f(x) - \hat{b})|x \in X\} = m(\lambda) \ \ . \end{array}$$

Thus, we have

$$f_0(x) \leq M(\hat{b}) \leq m(\lambda) \text{ for } x \in \Omega(\hat{b}), \ \ \lambda \geq 0 \ \ ,$$

and since $M(\hat{b})$ is independent of λ, if we take the infimum with respect to $\lambda \geq 0$ in the right-hand inequality we get (5.20). $\diamondsuit$

The basic problem of Duality Theory is to determine conditions under which $M(\hat{b}) = m^*$ in (5.20). We first give a simple sufficiency condition.

Definition: A pair $(\hat{x}, \hat{\lambda})$ with $\hat{x} \in X$, and $\hat{\lambda} \leq 0$ is said to satisfy the *optimality conditions* if

$$\hat{x} \text{ is optimal solution of (5.18) with } \lambda = \hat{\lambda}, \tag{5.21}$$

$$\hat{x} \text{ is feasible for (5.17), } i.e., f_i(\hat{x}) \leq \hat{b}_i \text{ for } i = 1, \ldots, m \ , \tag{5.22}$$

$$\hat{\lambda}_i = 0 \text{ when } f_i(\hat{x}) < \hat{b}_i, \text{ equivalently, } \hat{\lambda}'(f(\hat{x}) - \hat{b}) = 0. \tag{5.23}$$

$\hat{\lambda} \geq 0$ is said to be an *optimal price vector* if there is $\hat{x} \in X$ such that $(\hat{x}, \hat{\lambda})$ satisfy the optimality condition. Note that in this case $\hat{x} \in \Omega(\hat{b})$ by virtue of (5.22).

The next result is equivalent to Theorem 4(ii) of Section 1 if $X = R^n$, and $f_i, \ 0 \leq i \leq m$, are differentiable.

Theorem 1: (Sufficiency) If $(\hat{x}, \hat{\lambda})$ satisfy the optimality conditions, then $\hat{x}$ is an optimal solution to the primal, $\hat{\lambda}$ is an optimal solution to the dual, and $M(\hat{b}) = m^*$.

Proof: Let $x \in \Omega(\hat{b})$, so that $\hat{\lambda}'(f(x) - \hat{b}) \leq 0$. Then

$$\begin{aligned} f_0(x) &\leq f_0(x) - \hat{\lambda}'(f(x) - \hat{b}) \\ &\leq \ \sup\{f_0(x) - \hat{\lambda}'(f(x) - \hat{b}) | x \in X\} \\ &= f_0(\hat{x}) - \hat{\lambda}'(f(\hat{x}) - \hat{b}) \text{ by (5.21)} \\ &= f_0(\hat{x}) \text{ by (5.23)} \end{aligned}$$

so that $\hat{x}$ is optimal for the primal, and hence by definition $f_0(\hat{x}) = M(\hat{b})$. Also

$$\begin{aligned} m(\hat{\lambda}) &= f_0(\hat{x}) - \hat{\lambda}'(f(\hat{x}) - \hat{b}) \\ f_0(\hat{x}) &= M(\hat{b}) \ \ , \end{aligned}$$

so that from Weak Duality $\hat{\lambda}$ is optimal for the dual. $\diamondsuit$

We now proceed to a much more detailed investigation.

Lemma 3: B is a convex subset of R^m, and $M: \ B \to R \bigcup \{+\infty\}$ is a concave function.

Proof: Let $b, \tilde{b}$ belong to B, let $x \in \Omega(b), \ \tilde{x} \in \Omega(\tilde{b})$, let $0 \leq \theta \leq 1$. Then $(\theta x + (1-\theta)\tilde{x}) \in X$ since X is convex, and

$$f_i(\theta x + (1-\theta)\tilde{x}) \leq \theta f_i(x) + (1-\theta) f_i(\tilde{x})$$

since f_i is convex, so that

$$f_i(\theta x + (1-\theta)\tilde{x}) \leq \theta b + (1-\theta)\tilde{b} \ \ , \tag{5.24}$$

hence

$$(\theta x + (1-\theta)\tilde{x}) \in \Omega(\theta b + (1-\theta)\tilde{b})$$

and therefore, B is convex.

Also, since f_0 is concave,

$$f_0(\theta x + (1-\theta)\tilde{x}) \geq \theta f_0(x) + (1-\theta) f_0(\tilde{x}) \ .$$

Since (5.24) holds for all $x \in \Omega(b)$ and $\tilde{x} \in \Omega(\tilde{b})$ it follows that

$$\begin{aligned} M(\theta b + (1-\theta)\hat{b}) &\geq \sup\{f_0(\theta x + (1-\theta)\tilde{x}) | x \in \Omega(b),\ \tilde{x} \in \Omega(\tilde{b})\} \\ &\geq \sup\{f_0(x) | x \in \Omega(b)\} + (1-\theta)\ \sup\{f_0(\tilde{x}) | \tilde{x} \in \Omega(\tilde{b})\} \\ &= \theta M(b) + (1-\theta) M(\tilde{b}). \end{aligned}$$ $\diamondsuit$

Definition: Let $X \subset R^n$ and let $g:\ X \to R \bigcup \{\infty, -\infty\}$. A vector $\lambda \in R^n$ is said to be a *supergradient (subgradient)* of g at $\hat{x} \in X$ if

$$g(x) \leq g(\hat{x}) + \lambda'(x - \hat{x}) \text{ for } x \in X.$$
$$(g(x) \geq g(\hat{x}) + \lambda'(x - \hat{x}) \text{ for } x \in X.)$$

(See Figure 5-3.)

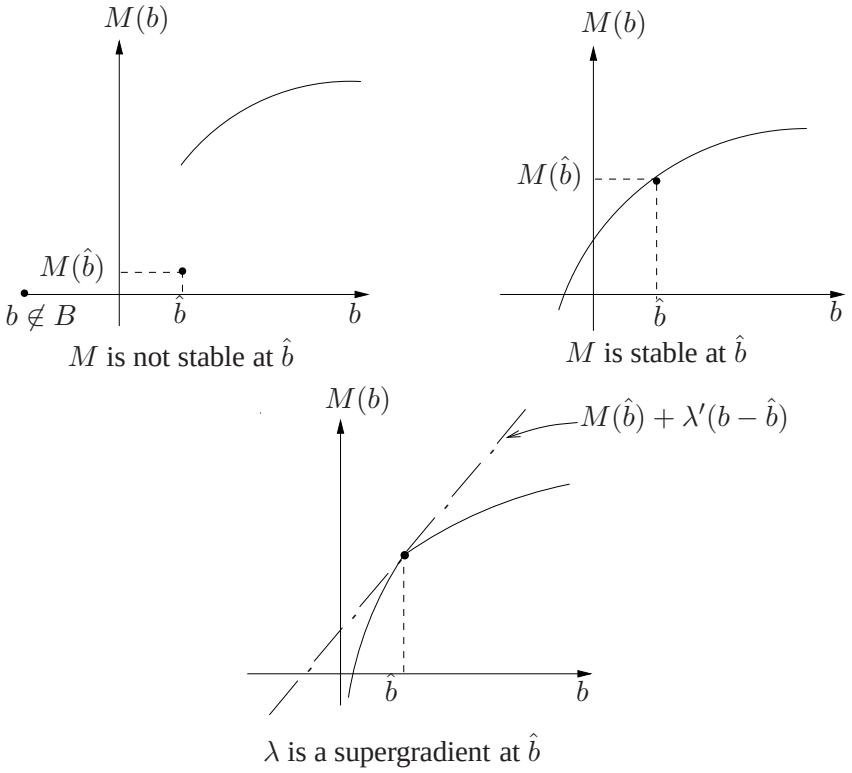


Figure 5.3: Illustration of supergradient of stability.

Definition: The function $M:\ B \to R \bigcup \{\infty\}$ is said to be *stable at* $\hat{b} \in B$ if there exists a real number K such that

$$M(b) \leq M(\hat{b}) + K|b - \hat{b}| \text{ for } b \in B\ .$$

(In words, M is stable at $\hat{b}$ if M does not increase infinitely steeply in a neighborhood of $\hat{b}$. See Figure 5.3.)

A more geometric way of thinking about subgradients is the following. Define the subset $A \subset R^{1+m}$ by

$$A = \{(r,b) | b \in B, \text{ and } r \leq M(b)\} \ .$$

Thus A is the set lying "below" the graph of M. We call A the *hypograph*[1] of M. Since M is concave it follows immediately that A is convex (in fact these are equivalent statements).

Definition: A vector $(\lambda_0, \lambda_1, \ldots, \lambda_m)$ is said to be the normal to a *hyperplane supporting A at a point*$(\hat{r}, \hat{b})$ if

$$\lambda_0 \hat{r} + \sum_{i=1}^{m} \lambda_i \hat{b}_i \geq \lambda_0 r + \sum_{i=1}^{m} \lambda_i b_i \text{ for all } (r,b) \in A \ . \tag{5.25}$$

(In words, A lies below the hyperplane $\hat{\pi} = \{(r,b) | \lambda_0 r + \sum \lambda_i b_i = \lambda_0 \hat{r} + \sum \lambda_i b_i\}$.) The supporting hyperplane is said to be *non-vertical* if $\lambda_0 \neq 0$. See Figure 5.4.

Exercise 2: Show that if $\hat{b} \in B, \ \tilde{b} \geq \hat{b}$, and $\tilde{r} \leq M(\hat{b})$, then $\tilde{b} \in B, \ M(\tilde{b})$, and $(\tilde{r}, \tilde{b}) \in A$.
Exercise 3: Assume that $\hat{b} \in B$, and $M(\hat{b}) < \infty$. Show that (i) if $\lambda = (\lambda_1, \ldots, \lambda_m)'$ is a supergradient of M at $\hat{b}$ then $\lambda \geq 0$, and $(1, -\lambda_1, \ldots, -\lambda_m)'$ defines a non-vertical hyperplane supporting A at $(M(\hat{b}), \hat{b})$, (ii) if $(\lambda_0, -\lambda_1, \ldots, -\lambda_m)'$ defines a hyperplane supporting A at $(M(\hat{b}), \hat{b})$ then $\lambda_0 \geq 0, \ \lambda_i \geq 0$ for $1 \leq i \leq m$; futhermore, if the hyperplane is non-vertical then $((\lambda_1/\lambda_0, \ldots, (\lambda_m/\lambda_0))'$ is a supergradient of M at $\hat{b}$.

We will prove only one part of the next crucial result. The reader who is familiar with the Separation Theorem of convex sets should be able to construct a proof for the second part based on Figure 5.4, or see the Appendix at the end of this Chapter.

Lemma 4: (Gale [1967]) M is stable at $\hat{b}$ iff M has a supergradient at $\hat{b}$. *Proof:* (Sufficiency only) Let λ be a supergradient at $\hat{b}$, then

$$\begin{aligned} M(b) &\leq M(\hat{b}) + \lambda'(b - \hat{b}) \\ &\leq M(\hat{b}) + |\lambda||b - \hat{b}| \ . \end{aligned}$$ $\diamondsuit$

The next two results give important alternative interpretations of supergradients.
Lemma 5: Suppose that $\hat{x}$ is optimal for (5.17). Then $\hat{\lambda}$ is a supergradient of M at $\hat{b}$ iff $\hat{\lambda}$ is an optimal price vector, and then $(\hat{x}, \hat{\lambda})$ satisfy the optimality conditions.
Proof: By hypothesis, $f(\hat{x}) = M(\hat{b}), \ \hat{x} \in X$, and $f(\hat{x}) \leq \hat{b}$. Let $\hat{\lambda}$ be a supergradient of M at $\hat{b}$. By Exercise 2, $(M(\hat{b}), f(\hat{x})) \in A$ and by Exercise 3, $\hat{\lambda} \geq 0$ and

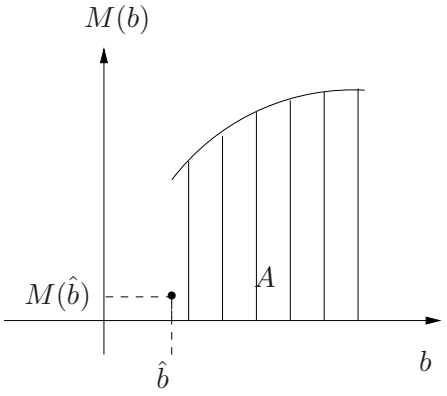
$$M(\hat{b}) - \hat{\lambda}'\hat{b} \geq M(\hat{b}) - \hat{\lambda}' f(\hat{x}) \ ,$$

so that $\hat{\lambda}'(f(\hat{x}) - \hat{b}) \geq 0$. But then $\hat{\lambda}'(\hat{b} - f(\hat{x})) = 0$, giving (5.23). Next let $x \in X$. Then $(f_0(x), f(x)) \in A$, hence again by Exercise 3

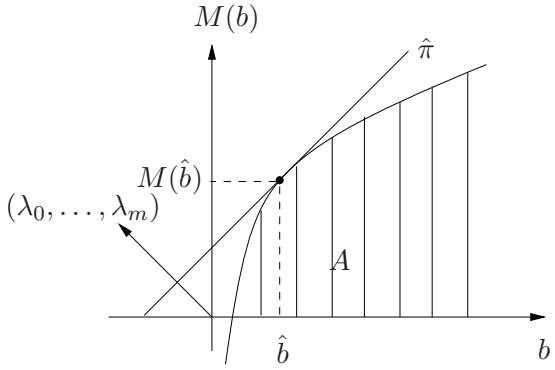
$$M(\hat{b}) - \hat{\lambda}'\hat{b} \geq f_0(x) - \hat{\lambda}' f(x) \ .$$

Since $f_0(\hat{x}) = M(\hat{b})$, and $\hat{\lambda}'(f(\hat{x}) - \hat{b}) = 0$, we can rewrite the inequality above as

[1] From the Greek "hypo" meaning below or under. This neologism contrasts with the *epigraph* of a function which is the set lying above the graph of the function.



No non-vertical hyperplane supporting A at $(M(\hat{b}), \hat{b})$



$\hat{\pi}$ is a non-vertical hyperplane supporting A at $(M(\hat{b}), \hat{b})$

Figure 5.4: Hypograph and supporting hyperplane.

$$f_0(\hat{x}) + \hat{\lambda}'(f(\hat{x}) - \hat{b}) \geq f_0(x) - \hat{\lambda}'(f(x) - \hat{b}) \ ,$$

so that (5.21) holds. It follows that $(\hat{x}, \hat{\lambda})$ satisfy the optimality conditions.

Conversely, suppose $\hat{x} \in X,\ \hat{\lambda} \geq 0$ satisfy (5.21), (5.22), and (5.23). Let $x \in \Omega(b)$, *i.e.*, $x \in X,\ f(x) \leq b$. Then $\hat{\lambda}'(f(x) - b) \leq 0$ so that

$$\begin{aligned} f_0(x) &\leq f_0(x)\ \hat{\lambda}'(f(x) - b) \\ &= f_0(x) - \hat{\lambda}'(f(x) - \hat{b}) + \hat{\lambda}'(b - \hat{b}) \\ &\leq f_0(\hat{x}) - \hat{\lambda}'(f(\hat{x}) - \hat{b}) + \hat{\lambda}'(b - \hat{b}) \qquad \text{by (5.21)} \\ &= f_0(\hat{x}) + \hat{\lambda}'(b - \hat{b}) \qquad\qquad\qquad\ \text{by (5.23)} \\ &= M(\hat{b}) + \hat{\lambda}'(b - \hat{b}) \ . \end{aligned}$$

Hence

$$M(b) = \sup\{f_0(x) | x \in \Omega(b)\} \leq M(\hat{b}) + \hat{\lambda}'(b - \hat{b}) \ ,$$

so that $\hat{\lambda}'$ is a supergradient of M at $\hat{b}$. $\diamondsuit$

Lemma 6: Suppose that $\hat{b} \in B$, and $M(\hat{b}) < \infty$. Then $\hat{\lambda}$ is a supergradient of M at $\hat{b}$ iff $\hat{\lambda}$ is an optimal solution of the dual (5.19) and $m(\hat{\lambda}) = M(\hat{b})$.

Proof: Let $\hat{\lambda}$ be a supergradient of M at $\hat{b}$. Let $x \in X$. By Exercises 2 and 3

$$M(\hat{b}) - \hat{\lambda}'\hat{b} \geq f_0(x) - \hat{\lambda}' f(x)$$

or

$$M(\hat{b}) \geq f_0(x) - \hat{\lambda}'(f(x) - \hat{b}) \quad ,$$

so that

$$M(\hat{b}) \geq \sup\{f_0(x) - \hat{\lambda}'(f(x) - \hat{b}) | x \in X\} = m(\hat{\lambda}) \quad .$$

By weak duality (Lemma 2) it follows that $M(\hat{b}) = m(\hat{\lambda})$ and $\hat{\lambda}$ is optimal for (5.19).

Conversely suppose $\hat{\lambda} \geq 0$, and $m(\hat{\lambda}) = M(\hat{b})$. Then for any $x \in X$

$$M(\hat{b}) \geq f_0(x) - \hat{\lambda}'(f(x) - \hat{b}) \quad ,$$

and if moreover $f(x) \leq b$, then $\hat{\lambda}'(f(x) - b) \leq 0$, so that

$$\begin{aligned} M(\hat{b}) &\geq f_0(x) - \hat{\lambda}'(f(x) - \hat{b}) + \hat{\lambda}'(f(x) - b) \\ &= f_0(x) - \hat{\lambda}' b + \hat{\lambda}'\hat{b} \qquad \text{for } x \in \Omega(b) \quad . \end{aligned}$$

Hence,

$$M(b) = \sup\{f_0(x) | x \in \Omega(b)\} \leq M(\hat{b}) + \hat{\lambda}'(b - \hat{b}) \quad ,$$

so that $\hat{\lambda}$ is a supergradient. ♢

We can now summarize our results as follows.

Theorem 2: (Duality) Suppose $\hat{b} \in B,\ M(\hat{b}) < \infty$, and M is stable at $\hat{b}$. Then

(i) there exists an optimal solution $\hat{\lambda}$ for the dual, and $m(\hat{\lambda}) = M(\hat{b})$,

(ii) $\hat{\lambda}$ is optimal for the dual iff $\hat{\lambda}$ is a supergradient of M at $\hat{b}$,

(iii) if $\hat{\lambda}$ is *any* optimal solution for the dual, then $\hat{x}$ is optimal for the primal iff $(\hat{x}, \hat{\lambda})$ satisfy the optimality conditions of (5.21), (5.22), and (5.23).

Proof: (i) follows from Lemmas 4,6. (ii) is implied by Lemma 6. The "if" part of (iii) follows from Theorem 1, whereas the "only if" part of (iii) follows from Lemma 5. ♢

Corollary 1: Under the hypothesis of Theorem 2, if $\hat{\lambda}$ is an optimal solution to the dual then $(\partial M^+/\partial b_i)(\hat{b}) \leq \hat{\lambda}_i \leq (\partial M^-/\partial b_i)(\hat{b})$.

Exercise 4: Prove Corollary 1. (Hint: See Theorem 5 of 4.2.3.)

5.2.2 *Interpretation and extensions.*

It is easy to see using convexity properties that, if $X = R^n$ and $f_i,\ 0 \leq i \leq m$, are differentiable, then the optimality conditions (5.21), (5.22), and (5.23) are equivalent to the Kuhn-Tucker condition (5.8). Thus the condition of stability of M at $\hat{b}$ plays a similar role to the constraint qualification. However, by Lemmas 4, 6 stability is *equivalent* to the existence of optimal dual variables, whereas CQ is only a *sufficient* condition. In other words if CQ holds at $\hat{x}$ then M is stable at $\hat{b}$. In particular, if $X = R^n$ and the f_i are differentiable, the various conditions of Section 1.3 imply stability. Here we give one sufficient condition which implies stability for the general case.

Lemma 7: If $\hat{b}$ is in the interior of B, in particular if there exists $x \in X$ such that $f_i(x) < \hat{b}_i$ for $1 \leq i \leq m$, then M is stable at $\hat{b}$.

The proof rests on the Separation Theorem for convex sets, and only depends on the fact that M is concave, $M(\hat{b}) < \infty$ without loss of generality, and $\hat{b}$ is the interior of B. For details see the Appendix.

Much of duality theory can be given an economic interpretation similar to that in Section 4.4. Thus, we can think of x as the vector of n activity levels, $f_0(x)$ the corresponding revenue, X as constraints due to physical or long-term limitations, b as the vector of current resource supplies, and finally $f(x)$ the amount of these resources used up at activity levels x. The various convexity conditions are generalizations of the economic hypothesis of non-increasing returns-to-scale. The primal problem (5.17) is the short-term decision problem faced by the firm. Next, if the current resources can be bought or sold at prices $\hat{\lambda} = (\lambda_1, \ldots, \lambda_m)'$, the firm faces the decision problem (5.18). If for a price system $\hat{\lambda}$, an optimal solution of (5.17) also is an optimal solution for (5.18), then we can interpret $\hat{\lambda}$ as a system of *equilibrium* prices just as in 4.2. Assuming the realistic condition $\hat{b} \in B,\ M(\hat{b}) < \infty$ we can see from Theorem 2 and its Corollary 1 that there exists an equilibrium price system iff $(\partial M^+/\partial b_i)(\hat{b}) < \infty,\ 1 \leq i \leq m$; if we interpret $(\partial M^+/\partial b_i)(\hat{b})$ as the marginal revenue of the *ith* resource, we can say that equilibrium prices exist iff marginal productivities of every (variable) resource is finite. These ideas are developed in (Gale [1967]).

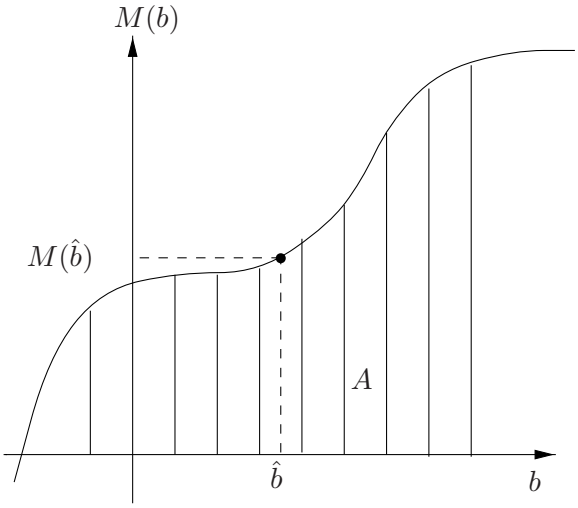


Figure 5.5: If M is not concave there may be no supporting hyperplane at $(M(\hat{b}), \hat{b})$.

Referring to Figure 5.3 or Figure 5.4, and comparing with Figure 5.5 it is evident that if M is not concave or, equivalently, if its hypograph A is not convex, there may be *no* hyperplane supporting A at $(M(\hat{b}), \hat{b})$. This is the reason why duality theory requires the often restrictive convexity hypothesis on X and f_i. It is possible to obtain the duality theorem under conditions slightly weaker than convexity but since these conditions are not easily verifiable we do not pursue this direction any further (see Luenberger [1968]). A much more promising development has recently taken place. The basic idea involved is to consider supporting A at $(M(\hat{b}), \hat{b})$ by (non-vertical) surfaces $\hat{\pi}$ more general than hyperplanes; see Figure 5.6. Instead of (5.18) we would then have more general problem of the form (5.26):

$$\begin{array}{l} \text{Maximize } f_0(x) - F(f(x) - \hat{b}) \\ \text{subject to } x \in X \ , \end{array} \tag{5.26}$$

where $F: R^m \to R$ is chosen so that $\hat{\pi}$ (in Figure 5.6) is the graph of the function $b \mapsto M(\hat{b}) - F(b - \hat{b})$. Usually F is chosen from a class of functions ϕ parameterized by $\mu = (\mu_1, \ldots, \mu_k) \geq 0$. Then for each fixed $\mu \geq 0$ we have (5.27) instead of (5.26):

$$\begin{array}{l} \text{Maximize } f_0(x) - \phi(\mu;\ f(x) - \hat{b}) \\ \text{subject to } x \in X \quad . \end{array} \tag{5.27}$$

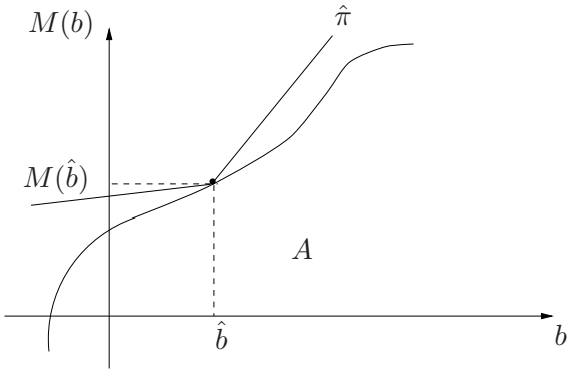


Figure 5.6: The surface $\hat{\pi}$ supports A at $(M(\hat{b}), \hat{b})$.

If we let

$$\psi(\mu) = \text{sup}\{f_0(x) - \phi(\mu; f(x) - \hat{b}) | x \in X\} \quad .$$

then the dual problem is

$$\begin{array}{l} \text{Minimize } \psi(\mu) \\ \text{subject to } \mu \geq 0 \quad , \end{array}$$

in analogy with (5.19).

The economic interpretation of (5.27) would be that if the prevailing (non-uniform) price system is $\phi(\mu; \cdot)$ then the resources $f(x) - \hat{b}$ can be bought (or sold) for the amount $\phi(\mu; f(x) - \hat{b})$. For such an interpretation to make sense we should have $\phi(\mu; b) \geq 0$ for $b \geq 0$, and $\phi(\mu; b) \geq \phi(\mu; \tilde{b})$ whenever $b \geq \tilde{b}$. A relatively unnoticed, but quite interesting development along these lines is presented in (Frank [1969]). Also see (Arrow and Hurwicz [1960]).

For non-economic applications, of course, no such limitation on ϕ is necessary. The following references are pertinent: (Gould [1969]), (Greenberg and Pierskalla [1970]), (Banerjee [1971]). For more details concerning the topics of 2.1 see (Geoffrion [1970a]) and for a mathematically more elegant treatment see (Rockafellar [1970]).

5.2.3 *Applications.*

Decentralized resource allocation.

Parts (i) and (iii) of Theorem 2 make duality theory attractive for computation purposes. In particular from Theorem 2 (iii), if we have an optimal dual solution $\hat{\lambda}$ then the optimal primal solutions are those optimal solutions of (5.18) for $\lambda = \hat{\lambda}$ which also satisfy the feasibility condition (5.22) and

the "complementary slackness" condition (5.23). This is useful because generally speaking (5.18) is easier to solve than (5.17) since (5.18) has fewer constraints.

Consider a decision problem in a large system (*e.g.*, a multi-divisional firm). The system is made up of k sub-systems (divisions), and the decision variable of the *ith* sub-system is a vector $x^i \in R^{n_i},\ 1 \le i \le k$. The sub-system has individual constraints of the form $x^i \in X^i$ where x^i is a convex set. Furthermore, the sub-systems share some resources in common and this limitation is expressed as $f^1(x^1) + \ldots + f^k(x^k) \le \hat{b}$ where $f^i :\ R^{n_i} \to R^m$ are convex functions and $\hat{b} \in R^m$ is the vector of available common resources. Suppose that the objective function of the large system is additive, *i.e.* it is the form $f_0^1(x^1) + \ldots + f_0^k(x^k)$ where $f_0^i :\ R^{n_i} \to R$ are concave functions. Thus we have the decision problem (5.28):

$$\begin{aligned} &\text{Maximize } \sum_{i=1}^{k} f_0^i(x^i) \\ &\text{subject to } x^i \in X^i\ ,\ 1 \le i \le k, \\ &\qquad \sum_{i=1}^{k} f^i(x^i) \le \hat{b}\ . \end{aligned} \tag{5.28}$$

For $\lambda \in R^m,\ \lambda \ge 0$, the problem corresponding to (5.19) is

$$\begin{aligned} &\text{Maximize } f_0^i(x^i) - \lambda' f^i(x^i) - \lambda'(\sum_{i=1}^{k} f^i(x^i) - \hat{b}) \\ &\text{subject to } x^i \in X^i\ ,\ 1 \le i \le k\ , \end{aligned}$$

which decomposes into k separate problems:

$$\begin{aligned} &\text{Maximize } f_0^i(x^i) - \lambda' f^i(x^i) \\ &\text{subject to } x^i \in X_i\ ,\ 1 \le i \le k\ . \end{aligned} \tag{5.29}$$

If we let $m^i(\lambda) = \sup\{f_0^i(x^i) - \lambda' f^i(x^i) | x^i \in X^i\}$, and $m(\lambda) = \sum_{i=1}^{k} m^i(\lambda) + \lambda'\hat{b}$, then the dual problem is

$$\begin{aligned} &\text{Minimize } m(\lambda)\ , \\ &\text{subject to } \lambda \ge 0\ . \end{aligned} \tag{5.30}$$

Note that (5.29) may be much easier to solve than (5.28) because, first of all, (5.29) involves fewer constraints, but perhaps more importantly the decision problems in (5.29) are decentralized whereas in (5.28) all the decision variables $x^1, \ldots, x^k$ are coupled together; in fact, if k is very large it may be practically impossible to solve (5.28) whereas (5.29) may be trivial if the dimensions of x^i are small.

Assuming that (5.28) has an optimal solution and the stability condition is satisfied, we need to find an optimal dual solution so that we can use Theorem 2(iii). For simplicity suppose that the $f_0^i,\ 1 \le i \le k$, are strictly concave, and also suppose that (5.29) has an optimal solution for every $\lambda \ge 0$. Then by Exercise 8 of Section 1, for each $\lambda \ge 0$ there is a unique optimal solution of (5.29), say $x^i(\lambda)$. Consider the following algorithm.

Step 1. Select $\lambda^0 \geq 0$ arbitrary. Set $p = 0$, and go to Step 2.
Step 2. Solve (5.29) for $\lambda = \lambda^p$ and obtain the optimal solution $x^p = (x^1(\lambda^p), \ldots, x^k(\lambda^p))$. Compute $e^p = \sum_{i=1}^{k} f^i(x^i(\lambda^p)) - \hat{b}$. If $e^p \geq 0,\ x^p$ is feasible for (5.28) and can easily be seen to be optimal.
Step 3. Set $\lambda^{p=1}$ according to

$$\lambda_i^{p+1} = \begin{cases} \lambda_i^p & \text{if } e_i^p \geq 0 \\ \lambda_i^p - d^p e_i^p & \text{if } e_i^p < 0 \end{cases}$$

where $d^p > 0$ is chosen *a priori*. Set $p = p + 1$ and return to Step 3.

It can be shown that if the step sizes d^p are chosen properly, x^p will converge to the optimum solution of (5.28). For more detail see (Arrow and Hurwicz [1960]), and for other decentralization schemes for solving (5.28) see (Geoffrion [1970b]).

Control of water quality in a stream.

The discussion in this section is mainly based on (Kendrick, *et al.*, [1971]). For an informal discussion of schemes of pollution control which derive their effectiveness from duality theory see (Solow [1971]). See (Dorfman and Jacoby [1970].)

Figure 5.7 is a schematic diagram of a part of a stream into which n sources (industries and municipalities) discharge polluting effluents. The pollutants consist of various materials, but for simplicity of exposition we assume that their impact on the quality of the stream is measured in terms of a single quantity, namely the biochemical oxygen demand (BOD) which they place on the dissolved oxygen (DO) in the stream. Since the DO in the stream is used to breakdown chemically the pollutants into harmless substances, the quality of the stream improves with the amount of DO and decreases with increasing BOD. It is a well-advertized fact that if the DO drops below a certain concentration, then life in the stream is seriously threatened; indeed, the stream can "die." Therefore, it is important to treat the effluents before they enter the stream in order to reduce the BOD to concentration levels which can be safely absorbed by the DO in the stream. In this example we are concerned with finding the optimal balance between costs of waste treatment and costs of high BOD in the stream.

We first derive the equations which govern the evolution in time of BOD and DO in the n areas of the streams. The fluctuations of BOD and DO will be cyclical with a period of 24 hours. Hence, it is enough to study the problem over a 24-hour period. We divide this period into T intervals, $t = 1, \ldots, T$. During interval t and in area i let
$z_i(t) =$ concentration of BOD measured in mg/liter,
$q_i(t) =$ concentration of DO measured in mg/liter,
$s_i(t) =$ concentration of BOD of effluent discharge in mg/liter, and
$m_i(t) =$ amount of effluent discharge in liters.

The principle of conservation of mass gives us equations (5.31) and (5.32):

$$z_i(t+1) - z_i(t) = -\alpha_i z_i(t) + \frac{\psi_{i-1} z_{i-1}(t)}{v_i} - \frac{\psi_i z_i(t)}{v_i} + \frac{s_i(t) m_i(t)}{v_i}\ , \tag{5.31}$$

$$\begin{aligned} q_i(t+1) - q_i(t) = \beta_i(q_i^s - q_i(t)) + \frac{\psi_{i-1} q_{i-1}(t)}{v_i} - \frac{\psi_i q_i(t)}{v_i} \\ +\alpha_i z_i(t) - \eta_i v_i,\ t = 1, \ldots, T \text{ and } i = 1, \ldots, N. \end{aligned} \tag{5.32}$$

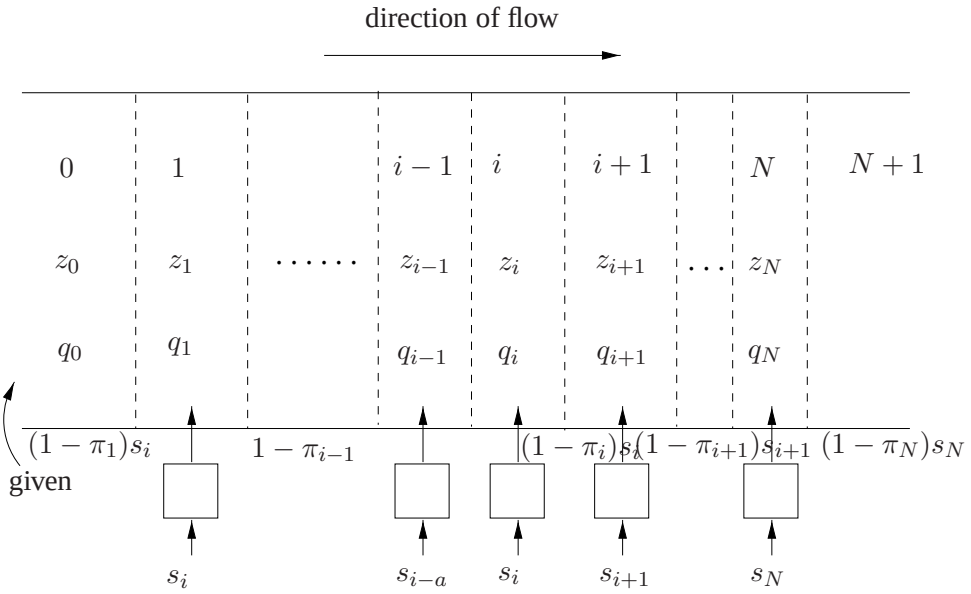


Figure 5.7: Schematic of stream with effluent discharges.

Here, v_i = volume of water in area i measured in liters, ψ_i = volume of water which flows from area i to are $i+1$ in each period measured in liters. α_i is the rate of decay of BOD per interval. This decay occurs by combination of BOD and DO. β_i is the rate of generation of DO. The increase in DO is due to various natural oxygen-producing biochemical reactions in the stream and the increase is proportional to $(q^s - q_i)$ where q^s is the saturation level of DO in the stream. Finally, η_i is the DO requirement in the bottom sludge. The $v_i, \psi_i, \alpha_i, \eta_i, q^s$ are parameters of the stream and are assumed known. They may vary with the time interval t. Also $z_0(t), q_0(t)$ which are the concentrations immediately upstream from area 1 are assumed known. Finally, the initial concentrations $z_i(1), q_i(1),\ i = 1, \ldots, N$ are assumed known.

Now suppose that the waste treatment facility in area i removes in interval t a fraction $\pi_i(t)$ of the concentration $s_i(t)$ of BOD. Then (5.31) is replaced by

$$z_i(t+1) - z_i(t) = -\alpha_i z_i(t) + \tfrac{\psi_i z_{i-1}}{v_i} - \tfrac{\psi_i z_i(t)}{v_i} + \tfrac{(1-\pi_i(t))s_i(t)m_i(t)}{v_i}\ . \tag{5.33}$$

We now turn to the costs associated with waste treatment and pollution. The cost of waste treatment can be readily identified. In period t the ith facility treats $m_i(t)$ liters of effluent with a BOD concentration $s_i(t)$ mg/liter of which the facility removes a fraction $\pi_i(t)$. Hence, the cost in period t will be $f_i(\pi_i(t), s_i(t), m_i(t))$ where the function must be monotonically increasing in all of its arguments. We further assume that f is convex.

The costs associated with increased amounts of BOD and reduced amounts of DO are much more difficult to quantify since the stream is used by many institutions for a variety of purposes (*e.g.*, agricultural, industrial, municipal, recreational), and the disutility caused by a decrease in the water quality varies with the user. Therefore, instead of attempting to quantify these costs let us suppose that some minimum water quality standards are set. Let $\underline{q}$ be the minimum acceptable DO concentration and let $\bar{z}$ be the maximum permissible BOD concentration. Then we face the

following NP:

$$\begin{aligned}&\text{Maximize } -\sum_{i=1}^{N}\sum_{t=1}^{T} f_i(\pi_i(t), s_i(t), m_i(t))\\ &\text{subject to (5.32), (5.33), and}\\ &\qquad -q_i(t) \le -\underline{q}\ ,\ i=1,\ldots,N;\ t=1,\ldots,T,\\ &\qquad z_i(t) \le \bar{z}\ ,\ i=1,\ldots,N;\ t=1,\ldots,T,\\ &\qquad 0 \le \pi_i(t) \le 1\ ,\ i=1,\ldots,N;\ t=1,\ldots,T.\end{aligned} \tag{5.34}$$

Suppose that all the treatment facilities are in the control of a single public agency. Then assuming that the agency is required to maintain the standards $(\underline{q}, \bar{z})$ and it does this at a minimum cost it will solve the NP (5.34) and arrive at an optimal solution. Let the minimum cost be $m(\underline{q}, \bar{z})$. But if there is no such centralized agency, then the individual polluters may not (and usually do not) have any incentive to cooperate among themselves to achieve these standards. Furthermore, it does not make sense to enforce legally a minimum standard $q_i(t) \ge \underline{q},\ z_i(t) \le \bar{z}$ on every polluter since the pollution levels in the *ith* area depend upon the pollution levels on all the other areas lying upstream. On the other hand, it may be economically and politically acceptable to tax individual polluters in proportion to the amount of pollutants discharged by the individual. The question we now pose is whether there exist tax rates such that if each individual polluter minimizes its own total cost (*i.e.,* cost of waste treatment + tax on remaining pollutants), then the resulting water quality will be acceptable and, furthermore, the resulting amount of waste treatment is carried out at the minimum expenditure of resources (*i.e.,* will be an optimal solution of (5.34)).

It should be clear from the duality theory that the answer is in the affirmative. To see this let $w_i(t) = (z_i(t), -q_i(t))'$, let $w(t) = (w_1(t), \ldots, w_N(t))$, and let $w = (w(1), \ldots, w(t))$. Then we can solve (5.32) and (5.33) for w and obtain

$$w = b + Ar\ \ , \tag{5.35}$$

where the matrix A and the vector b depend upon the known parameters and initial conditions, and r is the NT-dimensional vector with components $(1 - \pi_i(t))s_i(t)m_i(t)$. Note that the coefficients of the matrix must be non-negative because an increase in any component of r cannot decrease the BOD levels and cannot increase the DO levels. Using (5.35) we can rewrite (5.34) as follows:

$$\begin{aligned}&\text{Maximize } -\sum_{i}\sum_{t} f_i(\pi_i(t), s_i(t), m_i(t))\\ &\text{subject to } b + Ar \le \bar{w}\ ,\\ &\qquad 0 \le \pi_i(t) \le 1\ ,\ i=1,\ldots,N;\ t=1,\ldots,T,\end{aligned} \tag{5.36}$$

where the $2NT$-dimensional vector $\bar{w}$ has its components equal to $-\underline{q}$ or $\bar{z}$ in the obvious manner. By the duality theorem there exists a $2NT$-dimensional vector $\lambda^* \ge 0$, and an optimal solution $\pi_i^*(t),\ i=1,\ldots,N,\ t=1,\ldots,T$, of the problem:

$$\begin{aligned}&\text{Maximize } -\sum_{i}\sum_{t} f_i(\pi_i(t), s_i(t), m_i(t)) - \lambda^{*\prime}(b + Ar - w)\\ &\text{subject to } 0 \le \pi_i(t) \le 1,\ i=1,\ldots,N;\ t=1,\ldots,T\ ,\end{aligned} \tag{5.37}$$

such that $\{\pi_i^*(t)\}$ is also an optimal solution of (5.36) and, furthermore, the optimal values of (5.36) and (5.37) are equal. If we let $p^* = A'\lambda^* \ge 0$, and we write the components of p^* as $p_i^*(t)$ to match

with the components $(1-\pi_i(t))s_i(t)m_i(t)$ of r we can see that (5.37) is equivalent to the set of NT problems:

$$\begin{aligned} \text{Maximize } & -f_i(\pi_i(t), s_i(t), m_i(t)) - p_i^*(t)(1-\pi_i(t))s_i(t)m_i(t) \\ & 0 \le \pi_i(t) \le 1 \ , \\ & i = 1, \ldots, N; t = 1, \ \ldots, T \ . \end{aligned} \tag{5.38}$$

Thus, $p_i^*(t)$ is optimum tax per mg of BOD in area i during period t.

Before we leave this example let us note that the optimum dual variable or shadow price λ^* plays an important role in a larger framework. We noted earlier that the quality standard $(\underline{q}, \bar{z})$ was somewhat arbitrary. Now suppose it is proposed to change the standard in the ith area during period t to $\underline{q}+\Delta q_i(t)$ and $\bar{z}+\Delta z_i(t)$. If the corresponding components of λ^* are $\lambda_i^{q*}(t)$ and $\lambda_i^{z*}(t)$, then the change in the minimum cost necessary to achieve the new standard will be approximately $\lambda_i^{q*}(t)\Delta q_i(t) + \lambda_i^{z*}(t)\Delta z_i(t)$. This estimate can now serve as a basis in making a benefits/cost analysis of the proposed new standard.

5.3 *Quadratic Programming*

An important special case of NP is the quadratic programming (QP) problem:

$$\begin{aligned} &\text{Maximize } c'x - \tfrac{1}{2}x'Px \\ &\text{subject to } Ax \le b, \ x \ge 0 \ , \end{aligned} \tag{5.39}$$

where $x \in R^n$ is the decision variable , $c \in R^n$, $b \in R^m$ are fixed, A is a fixed $m \times n$ matrix and $P = P'$ is a fixed positive semi-definite matrix.

Theorem 1: A vector $x^* \in R^n$ is optimal for (5.39) iff there exist $\lambda^* \in R^m$, $\mu^* \in R^n$, such that

$$\begin{aligned} & Ax^* \le b, \ x^* \ge 0 \\ & c - Px^* = A'\lambda^* - \mu^*, \ \lambda^* \ge 0, \ \mu^* \ge 0 \ , \\ & (\lambda^*)'(Ax^* - b) = 0 \ , \ (\mu^*)'x^* = 0 \ . \end{aligned} \tag{5.40}$$

Proof: By Lemma 3 of 1.3, CQ is satisfied, hence the necessity of these conditions follows from Theorem 2 of 1.2. On the other hand, since P is positive semi-definite it follows from Exercise 6 of Section 1.2 that $f_0 : \ x \mapsto c'x - 1/2 \ x'Px$ is a concave function, so that the sufficiency of these conditions follows from Theorem 4 of 1.2. ♢

From (5.40) we can see that x^* is optimal for (5.39) iff there is a solution $(x^*, y^*, \lambda^*, \mu^*)$ to (5.41), (5.42), and (5.43):

$$\begin{aligned} & Ax + I_m Y = b \\ & -Px - A'\lambda + I_n\mu = -c \ , \end{aligned} \tag{5.41}$$

$$x \ge 0 \ y \ge 0, \ \lambda \ge 0, \ \mu \ge 0 \ , \tag{5.42}$$

$$\mu'x = 0 \ , \lambda'y = 0 \ . \tag{5.43}$$

Suppose we try to solve (5.41) and (5.42) by Phase I of the Simplex algorithm (see 4.3.2). Then we must apply Phase II to the LP:

$$\text{Maximize } \ -\sum_{i=1}^{m} z_i - \sum_{j=1}^{n} \xi_j$$

$$\begin{array}{l}\text{subject to}\\ \begin{array}{llll} Ax + I_m y & & + z & = b \\ -Px & - A'\lambda + I_n \mu & + \xi & = -c \\ x \geq 0, y \geq 0, \lambda \geq 0, \mu \geq 0, z \geq 0, \xi \geq 0, & & & \end{array}\end{array} \tag{5.44}$$

starting with a basic feasible solution $z = b$, $\xi = -c$. (We have assumed, without loss of generality, that $b \geq 0$ and $-c \geq 0$.) If (5.41) and (5.42) have a solution then the maximum value in (5.44) is 0. We have the following result.

Lemma 1: If (5.41), (5.42), and (5.43) have a solution, then there is an optimal basic feasible solution of (5.44) which is also a solution f (5.41), (5.42), and (5.43).

Proof: Let $\hat{x}, \hat{y}, \hat{\lambda}, \hat{\mu}$ be a solution of (5.41), (5.42), and (5.43). Then $\hat{x}, \hat{y}, \hat{\lambda}, \hat{\mu}, \hat{z} = 0, \hat{\xi} = 0$ is an optimal solution of (5.44). Furthermore, from (5.42) and (5.43) we see that at most $(n + m)$ components of $(\hat{x}, \hat{y}, \hat{\lambda}, \hat{\mu})$ are non-zero. But then a repetition of the proof of Lemma 1 of 4.3.1 will also prove this lemma. ♢

This lemma suggests that we can apply the Simplex algorithm of 4.3.2 to solve (5.44), starting with the basic feasible solution $z = b$, $\xi = -c$, in order to obtain a solution of (5.41), (5.42), and (5.43). However, Step 2 of the Simplex algorithm must be modified as follows to satisfy (5.43):

If a variable x_j is currently in the basis, do not consider μ_j as a candidate for entry into the basis; if a variable y_i is currently in the basis, do not consider λ_i as a candidate for entry into the basis. If it not possible to remove the z_i and ξ_j from the basis, stop.

The above algorithm is due to Wolfe [1959]. The behavior of the algorithm is summarized below.

Theorem 2: Suppose P is positive definite. The algorithm will stop in a finite number of steps at an optimal basic feasible solution $(\hat{x}, \hat{y}, \hat{\lambda}, \hat{\mu}, \hat{z}, \hat{\xi})$ of (5.44). If $\hat{z} = 0$ and $\hat{\xi} = 0$ then $(\hat{x}, \hat{y}, \hat{\lambda}, \hat{\mu})$ solve (5.41), (5.42), and (5.43) and $\hat{x}$ is an optimal solution of (5.39). If $\hat{z} \neq 0$ or $\hat{\xi} \neq 0$, then there is no solution to (5.41), (5.42), (5.43), and there is no feasible solution of (5.39).

For a proof of this result as well as for a generalization of the algorithm which permits positive *semi*-definite P see (Cannon, Cullum, and Polak [1970], p. 159 ff).

5.4 *Computational Method*

We return to the general NP (5.45),

$$\begin{array}{l}\text{Maximize } f_0(x)\\ \text{subject to } f_i(x) \leq 0,\ i = 1, \ldots, m\ ,\end{array} \tag{5.45}$$

where $x \in R^n$, $f_i : R^n \to R$, $0 \leq i \leq m$, are differentiable. Let $\Omega \subset R^n$ denote the set of feasible solutions. For $\hat{x} \in \Omega$ define the function $\psi(\hat{x}) : R^n \to R$ by

$$\psi(\hat{x})(h) = \max\{-f_{0x}(\hat{x})h, f_1(\hat{x}) + f_{1x}(\hat{x})h, \ldots, f_m(\hat{x}) + f_{mx}(\hat{x})h\}.$$

Consider the problem:

$$\begin{array}{l}\text{Minimize } \psi(\hat{x})(h)\\ \text{subject to } \begin{array}{ll} -\psi(\hat{x})(h) - f_{0x}(\hat{x})h & \leq 0\ , \\ -\psi(\hat{x})(h) + f_i(\hat{x}) f_{ix} h & \leq 0\ , \\ 1 \leq i \leq m\ ,\ -1 \leq h_j \leq 1\ ,\ 1 \leq j \leq n & \ . \end{array}\end{array} \tag{5.46}$$

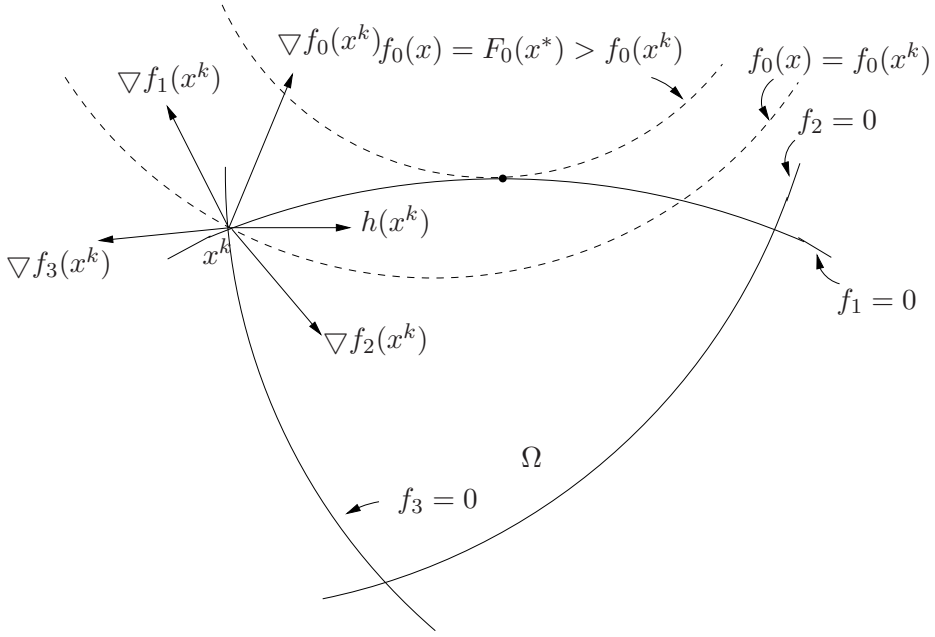


Figure 5.8: $h(x^k)$ is a feasible direction.

Call $h(\hat{x})$ an optimum solution of (5.46) and let $h_0(\hat{x}) = \psi(\hat{x})(h(\hat{x}))$ be the minimum value attained. (Note that by Exercise 1 of 4.5.1 (5.46) can be solved as an LP.)

The following algorithm is due to Topkis and Veinott [1967].

Step 1. Find $x^0 \in \Omega$, set $k = 0$, and go to Step 2.

Step 2. Solve (5.46) for $\hat{x} = x^k$ and obtain $h_0(x^k), h(x^k)$. If $h_0(x^k) = 0$, stop, otherwise go to Step 3.

Step 3. Compute an optimum solution $\mu(x^k)$ to the one-dimensional problem,

$$\begin{aligned} &\text{Maximize } f_0(x^k + \mu h(x^k)) \ , \\ &\text{subject to } (x^k + \mu h(x^k)) \in \Omega, \ \mu \geq 0 \ \ , \end{aligned}$$

and go to Step 4.

Step 4. Set $x^{k+1} = x^k + \mu(x^k)h(x^k)$, set $k = k + 1$ and return to Step 2.

The performance of the algorithm is summarized below.

Theorem 1: Suppose that the set

$$\Omega(x^0) = \{x | x \in \Omega, \ f_0(x) \geq f_0(x^0)\}$$

is compact, and has a non-empty interior, which is dense in $\Omega(x^0)$. Let x^* be any limit point of the sequence $x^0, x^1, \ldots, x^k, \ldots$, generated by the algorithm. Then the Kuhn-Tucker conditions are satisfied at x^*.

For a proof of this result and for more efficient algorithms the reader is referred to (Polak [1971]).

Remark: If $h_0(x^k) < 0$ in Step 2, then the direction $h(x^k)$ satisfies $f_{0x}(x^k)h(x^k) > 0$, and $f_i(x^k) + f_{ix}(x^K)h(x^k) < 0, \ 1 \leq i \leq m$. For this reason $h(x^k)$ is called a (desirable) *feasible direction*. (See Figure 5.8.)

5.5 *Appendix*

The proofs of Lemmas 4,7 of Section 2 are based on the following extremely important theorem (see Rockafeller [1970]).
Separation theorem for convex sets. Let F, G be convex subsets of R^n such that the relative interiors of F, G are disjoint. Then there exists $\lambda \in R^n,\ \lambda \neq 0$, and $\theta \in R$ such that

$$\begin{aligned} &\lambda' g \leq \theta \text{ for all } \ g \in G \\ &\lambda' f \geq \theta \text{ for all } f \in F\ . \end{aligned}$$

Proof of Lemma 4: Since M is stable at $\hat{b}$ there exists K such that

$$M(b) - M(\hat{b}) \leq K|b - \hat{b}| \text{ for all } b \in B\ . \tag{5.47}$$

In R^{1+m} consider the sets

$$\begin{aligned} F &= \{(r,b)|b \in R^m,\ r > K|b - \hat{b}|\}\ , \\ G &= \{(r,b)|b \in B,\ r \leq M(b) - M(\hat{b})\}\ . \end{aligned}$$

It is easy to check that F, G are convex, and (5.47) implies that $F \cap G = \phi$. Hence, there exist $(\lambda_0, \ldots, \lambda_m) \neq 0$, and θ such that

$$\begin{aligned} &\lambda_0 r + \sum_{i=1}^{m} \lambda_i b_i \leq \theta \ \text{ for } \ (r,b) \in G\ , \\ &\lambda_0 r + \sum_{i=1}^{m} \lambda_i b_i \geq \theta \ \text{ for } (r,b) \in F\ . \end{aligned} \tag{5.48}$$

From the definition of F, and the fact that $(\lambda_0, \ldots, \lambda_m) \neq 0$, it can be verified that (5.49) can hold only if $\lambda_0 > 0$. Also from (5.49) we can see that $\sum_{i=1}^{m} \lambda_i \hat{b}_i \geq \theta$, whereas from (5.48) $\sum_{i=1}^{m} \lambda_i \hat{b}_i \leq \theta$, so that $\sum_{i=1}^{m} \lambda_i \hat{b}_i = \theta$. But then from (5.48) we get

$$M(b) - M(\hat{b}) \leq \tfrac{1}{\lambda_0}[\theta - \sum_{i=1}^{m} \lambda_i b_i] = \sum_{i=1}^{m} (-\frac{\lambda_i}{\lambda_0})(b_i - \hat{b}). \qquad \diamondsuit$$

Proof of Lemma 7: Since $\hat{b}$ is in the interior of B, there exists $\varepsilon > 0$ such that

$$b \in B \text{ whenever } |b - \hat{b}| < \varepsilon\ . \tag{5.49}$$

In R^{1+m} consider the sets

$$\begin{aligned} F &= \{(r,\hat{b})|r > M(\hat{b}\} \\ G &= \{(r,b)|b \in B,\ r \leq M(b)\}\ . \end{aligned}$$

Evidently, F, G are convex and $F \cap G = \phi$, so that there exist $(\lambda_0, \ldots, \lambda_m) \neq 0$, and θ such that

$$\lambda_0 r + \sum_{i=1}^{m} \lambda_i \hat{b}_i \geq \theta\ , \text{ for } r > M(\hat{b})\ \ , \tag{5.50}$$

$$\lambda_0 r + \sum_{i=1}^{m} \lambda_i \hat{b}_i \leq \theta \ , \text{ for } (r, b) \in G \ . \tag{5.51}$$

From (5.49), and the fact that $(\lambda_0, \ldots, \lambda_m) \neq 0$ we can see that (5.50) and (5.51) imply $\lambda_0 > 0$. From (5.50),(5.51) we get

$$\lambda_0 M(\hat{b}) + \sum_{i=1}^{m} \lambda_i \hat{b}_i = \theta \ ,$$

so that (5.52) implies

$$M(b) \leq (\hat{b}) + \sum_{i=1}^{m} (-\frac{\lambda_i}{\lambda_0})(b_i - \hat{b}_i) \ . \qquad \diamondsuit$$

Chapter 6

SEQUENTIAL DECISION PROBLEMS: DISCRETE-TIME OPTIMAL CONTROL

In this chapter we apply the results of the last two chapters to situations where decisions have to be made sequentially over time. A very important class of problems where such situations arise is in the control of dynamical systems. In the first section we give two examples, and in Section 2 we derive the main result.

6.1 *Examples*

The trajectory of a vertical sounding rocket is controlled by adjusting the rate of fuel ejection which generates the thrust force. Specifically suppose that the equations of motion are given by (6.1).

$$\begin{aligned} \dot{x}_1(t) &= x_2(t) \\ \dot{x}_2(t) &= -\tfrac{C_D}{x_3(t)}\rho(x_1(t))x_2^2(t) - g + \tfrac{C_T}{x_3(t)}u(t) \\ \dot{x}_3(t) &= -u(t) \ , \end{aligned} \tag{6.1}$$

where $x_1(t)$ is the height of the rocket from the ground at time t, $x_2(t)$ is the (vertical) speed at time t, $x_3(t)$ is the weight of the rocket (= weight of remaining fuel) at time t. The "dot" denotes differentiation with respect to t. These equations can be derived from the force equations under the assumption that there are four forces acting on the rocket, namely: inertia = $x_3\ddot{x}_1 = x_3\dot{x}_2$; drag force $= C_D\rho(x_1)x_2^2$ where C_D is constant, $\rho(x_1)$ is a friction coefficient depending on atmospheric density which is a function of x_1; gravitational force $= gx_3$ with g assumed constant; and thrust force $C_T\dot{x}_3$, assumed proportional to rate of fuel ejection. See Figure 6.1. The decision variable at time t is $u(t)$, the rate of fuel ejection. At time 0 we assume that $(x_1(0), x_2(0), x_3(0)) = (0, 0, M)$; that is, the rocket is on the ground, at rest, with initial fuel of weight M. At a prescribed final time t_f, it is desired that the rocket be at a position as high above the ground as possible. Thus, the

decision problem can be formalized as (6.2).

$$\begin{array}{l}\text{Maximize } x_1(t_f)\\ \text{subject to } \dot{x}(t) = f(x(t), u(t)),\ 0 \le t \le t_f\\ \qquad\qquad x(0) = (0, 0, M)\\ \qquad\qquad u(t) \ge 0,\ x_3(t) \ge 0,\ 0 \le t \le t_f\ \ ,\end{array} \tag{6.2}$$

where $x = (x_1, x_2, x_3)',\ f:\ R^{3+1} \to R^3$ is the right-hand side of (6.1). The constraint inequalities $u(t) \ge 0$ and $x_3(t) \ge 0$ are obvious physical constraints.

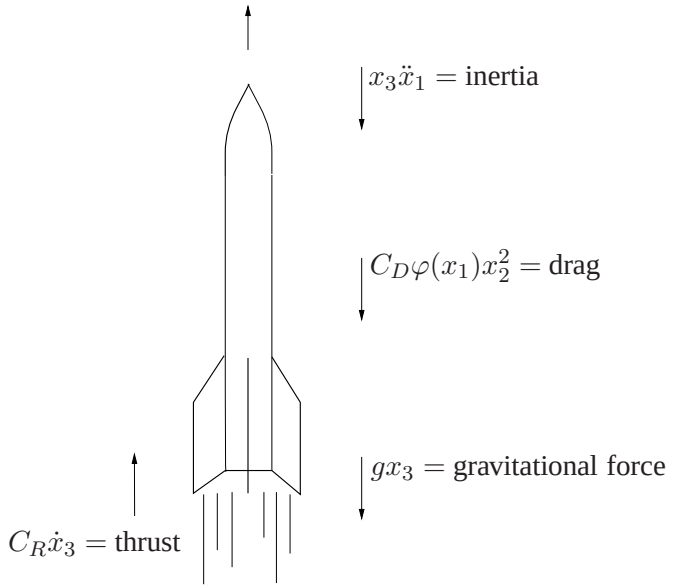


Figure 6.1: Forces acting on the rocket.

The decision problem (6.2) differs from those considered so far in that the decision variables, which are functions $u:\ [0, t_f] \to R$, *cannot* be represented as vectors in a *finite*-dimensional space. We shall treat such problems in great generality in the succeeding chapters. For the moment we assume that for computational or practical reasons it is necessary to approximate or restrict the permissible function $u(\cdot)$ to be constant over the intervals $[0, t_1),\ [t_1, t_2), \ldots, [t_{N-1}, t_f)$, where $t_1, t_2, \ldots, t_{N-1}$ are fixed *a priori*. But then if we let $u(i)$ be the constant value of $u(\cdot)$ over $[t_i, t_{i+1})$, we can reformulate (6.2) as (6.3):

$$\begin{array}{l}\text{Maximize } x_1(t_N)(t_N = t_f)\\ \text{subject to } x(t_{i+1}) = g(i, x(t_i), u(i)),\ i = 0, 1, \ldots, N-1\\ \qquad\qquad x(t_0) = x(0) = (0, 0, M)\\ \qquad\qquad u(i) \ge 0,\ x_3(t_i) \ge 0,\ i = 0, 1, \ldots, N\ \ .\end{array} \tag{6.3}$$

In (6.3) $g(i, x(t_1), u(i))$ is the state of the rocket at time t_{i+1} when it is in state $x(t_i)$ at time t_i and $u(t) \equiv u(i)$ for $t_i \le t < t_{i+1}$.

As another example consider a simple inventory problem where time enters discretely in a natural fashion. The Squeezme Toothpaste Company wants to plan its production and inventory schedule for the coming month. It is assumed that the demand on the *ith* day, $0 \le i \le 30$, is $d_1(i)$ for

their orange brand and $d_2(i)$ for their green brand. To meet unexpected demand it is necessary that the inventory stock of either brand should not fall below $\underline{s} > 0$. If we let $s(i) = (s_1(i), s_2(i))'$ denote the stock at the beginning of the *ith* day, and $m(i) = (m_1(i), m_2(i))'$ denote the amounts manufactured on the *ith* day, then clearly

$$s(i+1) + s(i) + m(i) - d(i) \ ,$$

where $d(i) = (d_1(i), d_2(i))'$. Suppose that the initial stock is $\hat{s}$, and the cost of storing inventory s for one day is $c(s)$ whereas the cost of manufacturing amount m is $b(m)$. The the cost-minimization decision problem can be formalized as (6.4):

$$\begin{aligned} &\text{Maximize } \sum_{i=0}^{30}(c(s(i)) + b(m(i))) \\ &\text{subject to } s(i+1) = s(i) + m(i) - d(i),\ 0 \le i \le 29 \\ &\qquad\qquad s(0) = \hat{s} \\ &\qquad\qquad s(i) \ge (\underline{s}, \underline{s})',\ m(i) \ge 0,\ 0 \le i \le 30 \ . \end{aligned} \tag{6.4}$$

Before we formulate the general problem let us note that (6.3) and (6.4) are in the form of non-linear programming problems. The reason for treating these problems separately is because of their practical importance, and because the conditions of optimality take on a special form.

6.2 *Main Result*

The general problem we consider is of the form (6.5).

$$\begin{aligned} &\text{Maximize } \sum_{i=0}^{N-1} f_0(i, x(i), u(i)) \\ &\text{subject to} \\ &\textit{dynamics}: \quad x(i+1) - x(i) = f(i, x(i), u(i)),\ i = 0, \ldots, N-1 \ , \\ &\textit{initial condition:}\ \ q_0(x(0) \le 0,\ g_0(x(0)) = 0 \ , \\ &\textit{final condition:}\ \ q_N(x(N)) \le 0, g_N(x(N)) = 0 \ , \\ &\textit{state-space constraint:}\ \ q_i(x(i)) \le 0, \quad i = 1, \ldots, N-1 \ , \\ &\textit{control constraint:}\ \ h_i(u(i)) \le 0, \qquad i = 0, \ldots, N-1 \ . \end{aligned} \tag{6.5}$$

Here $x(i) \in R^n,\ u(i) \in R^p,\ f_0(i, \cdot, \cdot) : R^{n+p} \to R,\ f(i, \cdot, \cdot) : R^{n+p} \to R^n,\ q_i : R^n \to R^{m_i},\ g_i : R^n \to R^{\ell_i},\ h_i : R^p \to R^{s_i}$ are given differentiable functions. We follow the control theory terminology, and refer to $x(i)$ as the *state* of the system at time i, and $u(i)$ as the *control* or *input* at time i.

We use the formulation mentioned in the Remark following Theorem 3 of V.1.2, and construct the Lagrangian function L by

$$\begin{gathered} L(x(0), \ldots, x(N); u(0), \ldots, u(N-1); p(1), \ldots, p(N); \\ \lambda^0, \ldots, \lambda^N; \alpha^0, \alpha^N; \gamma^0, \ldots, \gamma^{N-1}) \end{gathered}$$

$$= \sum_{i=0}^{N-1} f_0(i, x(i), u(i)) - \left\{ \sum_{i=0}^{N-1} (p(i+1))'(x(i+1) - x(i) - f(i, x(i), u(i))) + \right.$$
$$\left. \sum_{i=0}^{N} (\lambda^i)' q_i(x(i)) + (\alpha^0)' g_0(x(0)) + (\alpha^N)' g_N(x(N)) + \sum_{i=0}^{N-1} (\gamma^i)' h_i(u(i)) \right\} \quad .$$

Suppose that CQ is satisfied for (6.5), and $x^*(0), \ldots, x^*(N);\ u^*(0), \ldots, u^*(N-1)$, is an optimal solution. Then by Theorem 2 of 5.1.2, there exist $p^*(i)$ in R^n for $1 \leq i \leq N,\ \lambda^{i*} \geq 0$ in R^{m_i} for $0 \leq i \leq N,\ \alpha^{i*}$ in R^{ℓ_i} for $i = 0, N$, and $\gamma^{i*} \geq 0$ in R^{s_i} for $0 \leq i \leq N-1$, such that
(A) the derivative of L evaluated at these points vanishes,
and
(B) $\lambda^{i*} q_i(x^*(i)) = 0$ for $0 \leq i \leq N,\ \gamma^{i*} h_i(u^*(i)) = 0$ for $0 \leq i \leq N-1$.
We explore condition (A) by taking various partial derivatives.

Differentiating L with respect to $x(0)$ gives

$$f_{0x}(0, x^*(0), u^*(0)) - \{-(p^*(1))' - (p^*(1))'[f_x(0, x^*(0), u^*(0))]$$
$$+(\lambda^{0*})'[q_{0x}(x^*(0))] + (\alpha^{0*})'[g_{0x}(x^*(0))]\} = 0 \ ,$$

or

$$p^*(0) - p^*(1) = [f_x(0, x^*(0), u^*(x))]' p^*(1)$$
$$+[f_{0x}(0, x^*(0), u^*(0))]' - [q_{0x}(x^*(0))]' \lambda^{0*}, \tag{6.6}$$

where we have defined

$$p^*(0) = [g_{0x}(x^*(x))]' \alpha^{0*} \quad . \tag{6.7}$$

Differentiating L with respect to $x(i),\ 1 \leq i \leq N-1$, and re-arranging terms gives

$$p^*(i) - p^*(i+1) = [f_x(i, x^*(i), u^*(i))]' p^*(i+1)$$
$$+[f_{0x}(i, x^*(i), u^*(i))]' - [q_{ix}(x^*(i))]' \lambda^{i*} \ . \tag{6.8}$$

Differentiating L with respect to $x(N)$ gives,

$$p^*(N) = -[g_{Nx}(x^*(N))]' \alpha^{N*} - [q_{Nx}(x^*(N))]' \lambda^{N*} \ .$$

It is convenient to replace α^{N*} by $-\alpha^{N*}$ so that the equation above becomes (6.9)

$$p^*(N) = [g_{Nx}(x^*(N))]' \alpha^{N*} - [q_{Nx}(x^*(N))]' \lambda^{N*} \ . \tag{6.9}$$

Differentiating L with respect to $u(i),\ 0 \leq i \leq N-1$ gives

$$[f_{0u}(i, x^*(i), u^*(i))]' + [f_u(i, x^*(i), u^*(i))]' p^*(i+l) - [h_{iu}(u^*(i))]' \gamma^{i*} = 0 \ . \tag{6.10}$$

We summarize our results in a convenient form in
Table 6.1

Remark 1: Considerable elegance and mnemonic simplification is achieved if we define the *Hamiltonian function* H by

Suppose $x^*(0), \ldots, x^*(N)$; $u^*(0), \ldots, u^*(N-1)$ maximizes $\sum_{i=0}^{N-1} f_0(i, x(i), u(i))$ subject to the constraints below	then there exist $p^*(N); \lambda^{0_*}, \ldots, \lambda N_*; \alpha^{0_*}, \alpha^{N_*};$ $\gamma^{0_*}, \ldots, \gamma^{N-1_*}$, such that	
dynamics: $i = 0, \ldots, N-1$ $x(i+1) - x(i) = f(i, x(i), u(i))$	*adjoint equations:* $i = 0, \ldots, N-1$ $p^*(i) - p^*(i+1) = [f_x(i, x^*(i), u^*(i)]' p^*(i+1)$ $+[f_{0x}(i, x^*(i), u^*(i)]' - [q_{ix}(x^*(i)]' \gamma^{i_*}$	
initial condition: $q_0(x^*(0)) \leq 0, g_0(x^*(0)) = 0$	*transversality conditions:* $p^*(0) = [g_{0x}(x^*(0))]' \alpha^{0_*}$	$\lambda^{0_*} \geq 0,$ $(\lambda^{0_*})' q_0(x^*(0)) = 0$
final conditions: $q_N(x^*(N)) \leq 0, g_N(x^*(N)) = 0$	$p^*(N) = [g_{Nx}(x^*(N))]' \alpha^{N_*} - [q_{Nx}(x^*(N))]' \lambda^{N_*}$	$\lambda^{N_*} \geq 0,$ $(\lambda^{N_*})' q_N(x^*(N)) = 0$
state space constraint: $i = 1, \ldots, N-1$ $q_i(x^*(i)) \leq 0$		$\lambda^{i_*} \geq 0,$ $(\lambda^{i_*})' q_i(x^*(i)) = 0$
control constraint: $i = 0, \ldots, N-1$ $h_i(u^*(i)) \leq 0$	$[f_{0u}(i, x^*(i), u^*(i))]' + [f_u(i, x^*(i) u^*(i))]'.$ $p^*(i_1) = [h_{iu}(u^*(i))]' \gamma^{i_*}$	$\gamma^{i_*} \geq 0$ $(\gamma^{i_*})' h_i(u^*(i) = 0$

Table 6.1:

$$H(i,x,u,p) = f_0(i,x,u) + p'f(i,x,u) \ .$$

The dynamic equations then become

$$\begin{array}{c} x^*(i+1) - x^*(i) = [H_p(i,x^*(i),u^*(i),p^*(i+1))]' \ , \\ 0 \le i \le N-1 \ . \end{array} \tag{6.11}$$

and the adjoint equations (6.6) and (6.8) become

$$\begin{array}{c} p^*(i) - p^*(i+1) = [H_x(i,x^*(i),u^*(i),u^*(i),p^*(i+1))]' - [q_{ix}(x^*(i))]'\lambda^{i*}, \\ 0 \le i \le N-1 \ \ , \end{array}$$

whereas (6.10) becomes

$$[h_{iu}(u^*(i))]'\gamma^{i*} = [H_u(i,x^*(i),u^*(i),p^*(i+1))]', \ \ 0 \le i \le N-1 \ . \tag{6.12}$$

Remark 2: If we linearize the dynamic equations about the optimal solution we obtain

$$\delta x(i+1) - \delta x(i) = [f_x(i,x^*(i),u^*(i))]\delta x(i) + [f_u(i,x^*(i,x^*,(i),u^*(i))]\delta u(i) \ ,$$

whose homogeneous part is

$$z(i+1) - z(i) = [f_x(i,x^*(i),u^*(i))]z(i) \ ,$$

which has for it adjoint the system

$$r(i) - r(i+1) = [f_x(i,x^*(i),u^*(i))]'r(i+1) \ . \tag{6.13}$$

Since the homogeneous part of the linear difference equations (6.6), (6.8) is (6.13), we call (6.6), (6.8) the *adjoint* equations, and the $p^*(i)$ are called *adjoint variables*.

Remark 3: If the $f_0(i,\cdot,\cdot)$ are concave and the remaining function in (6.5) are linear, then CQ is satisfied, and the necessary conditions of Table 6.1 are also sufficient. Furthermore, in this case we see from (6.13) that $u^*(i)$ is an optimal solution of

$$\begin{array}{l} \text{Maximize } H(i,x^*(i),u,p^*(i+1)), \\ \text{subject to } h_i(u) \le 0 \ \ . \end{array}$$

For this reason the result is sometimes called the *maximum principle*.

Remark 4: The conditions (6.7), (6.9) are called *transversality* conditions for the following reason. Suppose $q_0 \equiv 0, \ q_N \equiv 0$, so that the initial and final conditions read $g_0(x(0)) = 0, \ g_N(x(N)) = 0$, which describe surfaces in R^n. Conditions (6.7), (6.9) become respectively $p^*(0) = [g_{0x}(x^*(0))]'\alpha^{0*}, \ p^*(N) = [g_{Nx}(x(N))]'\alpha^{N*}$ which means that $p^*(0)$ and $p^*(N)$ are respectively orthogonal or transversal to the initial and final surfaces. Furthermore, we note that in this case the initial and final conditions specify $(\ell_0 + \ell_n)$ conditions whereas the transversality conditions specify $(n-\ell_0)+(n-\ell_n)$ conditions. Thus, we have a total of $2n$ boundary conditions for the $2n$-dimensional system of difference equations (6.5), (6.12); but note that these $2n$ boundary conditions are *mixed, i.e.*, some of them refer to the initial time 0 and the rest refer to the final time.

Exercise 1: For the regulator problem,

$$\begin{aligned}
&\text{Maximize } \tfrac{1}{2}\sum_{i=0}^{N-1} x(i)'Qx(i) + \frac{1}{2}\sum_{i=0}^{N-1} u(i)'Pu(i) \\
&\text{subject to } x(i+1) - x(i) = Ax(i) + Bu(i),\ 0 \le i \le N-1 \\
&\qquad\qquad x(0) = \hat{x}(0), \\
&\qquad\qquad u(i) \in R^p\ ,\ 0 \le i \le N-1\ ,
\end{aligned}$$

where $x(i) \in R^n,\ A$ and B are constant matrices, $\hat{x}(0)$ is fixed, $Q = Q'$ is positive semi-definite, and $P = P'$ is positive definite, show that the optimal solution is unique and can be obtained by solving a $2n$-dimensional linear difference equation with mixed boundary conditions.

Exercise 2: Show that the minimal fuel problem,

$$\begin{aligned}
&\text{Minimize } \sum_{i=0}^{N-1}\left(\sum_{j=1}^{P} |(u(i))_j|\right)\ , \\
&\text{subject to } x(i+1) - x(i) = Ax(i) + Bu(i),\ 0 \le i \le N-1 \\
&\qquad\qquad x(0) = \hat{x}(0),\ x(N) = \hat{x}(N)\ , \\
&\qquad\qquad u(i) \in R^p\ ,\ |u(i))_j| \le 1,\ 1 \le j \le p,\ 0 \le i \le N-1
\end{aligned}$$

can be transformed into a linear programming problem. Here $\hat{x}(0), \hat{x}(N)$ are fixed, A and B are as in Exercise 1.

Chapter 7

SEQUENTIAL DECISION PROBLEMS: CONTINUOUS-TIME OPTIMAL CONTROL OF LINEAR SYSTEMS

We will investigate decision problems similar to those studied in the last chapter with one (mathematically) crucial difference. A choice of control has to be made at each instant of time t where t varies continuously over a finite interval. The evolution in time of the state of the systems to be controlled is governed by a differential equation of the form:

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

where $x(t) \in R^n$ and $u(t) \in R^p$ are respectively the state and control of the system at time t.

To understand the main ideas and techniques of analysis it will prove profitable to study the linear case first. The general nonlinear case is deferred to the next chapter. In Section 1 we present the general linear problem and study the case where the initial and final conditions are particularly simple. In Section 2 we study more general boundary conditions.

7.1 *The Linear Optimal Control Problem*

We consider a dynamical system governed by the linear differential equation (7.1):

$$\dot{x}(t) = A(t)x(t) + B(t)u(t), \ t \geq t_0 \ . \tag{7.1}$$

Here $A(\cdot)$ and $B(\cdot)$ are $n \times n$- and $n \times p$-matrix valued functions of time; we assume that they are piecewise continuous functions. The control $u(\cdot)$ is constrained to take values in a fixed set $\Omega \subset R^p$, and to be piecewise continuous.

Definition: A piecewise continuous function $u : [t_0, \infty) \to \Omega$ will be called an *admissible control*. $\mathcal{U}$ denotes the set of all admissible controls.

Let $c \in R^n, \ x^0 \in R^n$ be fixed and let $t_f \geq t_0$ be a fixed time. We are concerned with the

decision problem (7.2).

$$\begin{array}{l}\text{Maximize } c'x(t_f),\\ \text{subject to}\\ \quad \textit{dynamics: } \dot{x}(t) = A(t)x(t) + B(t)u(t)\ ,\ t_0 \le t \le t_f\ ,\\ \quad \textit{initial condition: } x(t_0) = x^0\ ,\\ \quad \textit{final condition: } x(t_f) \in R^n\ ,\\ \quad \textit{control constraint: } u(\cdot) \in \mathcal{U}\ .\end{array} \tag{7.2}$$

Definition: (i) For any piecewise continuous function $u(\cdot):\ [t_0, t_f] \to R^p$, for any $z \in R^n$, and any $t_0 \le t_1 \le t_2 \le t_f$ let

$$\phi(t_2, t_1, z, u)$$

denote the state of (7.1) at time t_2, if a time t_1 it is in state z, and the control $u(\cdot)$ is applied.
(ii) Let

$$K(t_2, t_1, z) = \{\phi(t_2, t_1, z, u) | u \in \mathcal{U}\}\ .$$

Thus, $K(t_2, t_1, z)$ is the set of states reachable at time t_2 starting at time t_1 in state z and using admissible controls. We call K *the reachable* set.
Definition: Let $\Phi(t, \tau),\ t_0 \le \tau \le t \le t_f$, be the *transition-matrix* function of the homogeneous part of (7.1), *i.e.*, Φ satisfies the differential equation

$$\frac{\partial \Phi}{\partial t}(t, \tau) = A(t)\Phi(t, \tau)\ ,$$

and the boundary condition

$$\Phi(t, t) \equiv I_n\ .$$

The next result is well-known. (See Desoer [1970].)
Lemma 1: $\phi(t_2, t_1, z, u) = \Phi(t_2, t_1)z + \displaystyle\int_{t_1}^{t_2} \Phi(t_2, \tau)B(\tau)u(\tau)d\tau.$

Exercise 1: (i) Assuming that Ω is convex, show that $\mathcal{U}$ is a convex set. (ii) Assuming that $\mathcal{U}$ is convex show that $K(t_2, t_1, z)$ is a convex set. (It is a deep result that $K(t_2, t_1, z)$ is convex even if Ω is not convex (see Neustadt [1963]), provided we include in $\mathcal{U}$ any measurable function $u:\ [t_0, \infty) \to \Omega$.)

Definition: Let $K \subset R^n$, and let $x^* \in K$. We say that c is the *outward normal to a hyperplane supporting K at* x^* if $c \ne 0$, and

$$c'x^* \ge c'x \ \text{ for all } \ x \in K\ .$$

The next result gives a geometric characterization of the optimal solutions of (2).
Lemma 2: Suppose $c \ne 0$. Let $u^*(\cdot) \in \mathcal{U}$ and let $x^*(t) = \phi(t, t_0, x^0, u^*)$. Then u^* is an optimal solution of (2) iff
(i) $x^*(t_f)$ is on the boundary of $K = K(t_f, t_0, x^0)$, and
(ii) c is the outward normal to a hyperplane supporting K at x^*. (See Figure 7.1.)

Proof: Clearly (i) is implied by (ii) because if $x^*(t_f)$ is in the interior of K there is $\delta > 0$ such that $(x^*(t_f) + \delta c) \in K$; but then

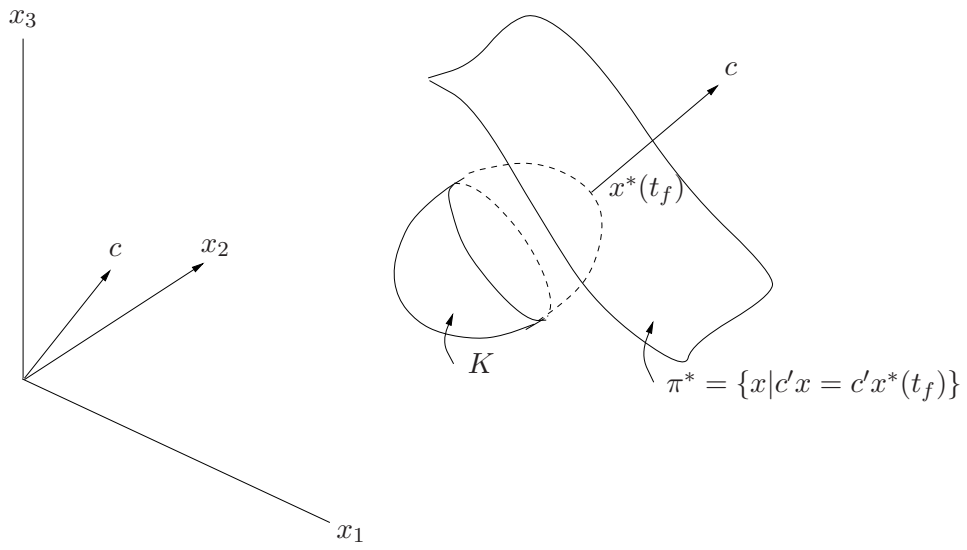


Figure 7.1: c is the outward normal to π^* supporting K at $x^*(t_f)$

.

$$c'(x^*(t_f) + \delta c) = c'x^*(t_f) + \delta|c|^2 > c'x^*(t_f) \ .$$

Finally, from the definition of K it follows immediately that u^* is optimal iff $c'x^*(t_f) \geq c'x$ for all $x \in K$. ♢

The result above characterizes the optimal control u^* in terms of the final state $x^*(t_f)$. The beauty and utility of the theory lies in the following result which translates this characterization directly in terms of u^*.

Theorem 1: Let $u^*(\cdot) \in \mathcal{U}$ and let $x^*(t) = \phi(t, t_0, x^0, u^*),\ t_0 \leq t \leq t_f$. Let $p^*(t)$ be the solution of (7.3) and (7.4):

$$\textit{adjoint equation: } \dot{p}^*(t) = -A'(t)p^*(t) \ , \ t_0 \leq t \leq t_f. \tag{7.3}$$

$$\textit{final condition: } p^*(t_f) = c \ . \tag{7.4}$$

Then $u^*(\cdot)$ is optimal iff

$$(p^*(t))'B(t)u^*(t) = \sup\{(p^*(t))'B(t)v | v \in \Omega\} \ , \tag{7.5}$$

for all $t \in [t_0, t_f]$, except possibly for a finite set.

Proof: $u^*(\cdot)$ is optimal iff for every $u(\cdot) \in \mathcal{U}$

$$\begin{array}{c}(p^*(t_f))'[\Phi(t_f, t_0)x^0 + \int_{t_0}^{t_f} \Phi(t_f, \tau)B(\tau)u^*(\tau)d\tau] \\ \geq (p^*(t_f))'[\Phi(t_f, t_0)x^0 + \int_{t_0}^{t_f} \Phi(t_f, \tau)B(\tau)u(\tau)d\tau] \ ,\end{array}$$

which is equivalent to (7.6).

$$\begin{array}{c}\int_{t_0}^{t_f} (p^*(t_f))'\Phi(t_f, \tau)B(\tau)u^*(\tau)d\tau \\ \geq \int_{t_0}^{t_f} (p^*(t_f))'\Phi(t_f, \tau)B(\tau)u(\tau)d\tau\end{array} \tag{7.6}$$

Now by properties of the adjoint equation we know that $p^*(t))' = (p^*(t_f))'\Phi(t_f, t)$ so that (7.6) is equivalent to (7.7),

$$\int_{t_0}^{t_f} (p^*(\tau))' B(\tau) u^*(\tau) d\tau \geq \int_{t_0}^{t_f} (p^*(\tau))' B(\tau) u(\tau) d\tau, \tag{7.7}$$

and the sufficiency of (7.5) is immediate.

To prove the necessity let D be the finite set of points where the function $B(\cdot)$ or $u^*(\cdot)$ is discontinuous. We shall show that if $u^*(\cdot)$ is optimal then (7.5) is satisfied for $t \notin D$. Indeed if this is not the case, then there exists $t^* \in [t_0, t_f],\ t^* \notin D$, and $v \in \Omega$ such that

$$(p^*(t^*))' B(t^*) u^*(t^*) < (p^*(t^*))' B(t^*) v \ ,$$

and since t^* is a point of continuity of $B(\cdot)$ and $u^*(\cdot)$, it follows that there exists $\delta > 0$ such that

$$(p^*(t))' B(t) u^*(t) < (p^*(t))' B(t) v, \text{ for } |t - t^*| < \delta \ . \tag{7.8}$$

Define $\tilde{u}(\cdot) \in \mathcal{U}$ by

$$\tilde{u}(t) = \begin{cases} v & |t - t^*| < \delta,\ t \in [t_0, t_f] \\ u^*(t) & \text{otherwise} \ . \end{cases}$$

Then (7.8) implies that

$$\int_{t_0}^{t_f} (p^*(t))' B(t) \tilde{u}(t) dt > \int_{t_0}^{t_f} (p^*(t))' B(t) u^*(t) dt \ .$$

But then from (7.7) we see that $u^*(\cdot)$ cannot be optimal, giving a contradiction. ♢

Corollary 1: For $t_0 \leq t_1 \leq t_2 \leq t_f$,

$$(p^*(t_2)) x^*(t_2) \geq (p^*(t_2))' x \ \text{ for all } \ x \in K(t_2, t_1, x^*(t_1)). \tag{7.9}$$

Exercise 2: Prove Corollary 1.

Remark 1: The geometric meaning of (7.9) is the following. Taking $t_1 = t_0$ in (7.9), we see that if $u^*(\cdot)$ is optimal, *i.e.*, if $c = p^*(t_f)$ is the outward normal to a hyperplane supporting $K(t_f, t_0, x^0)$ at $x^*(t_f)$, then $x^*(t)$ is on the boundary of $K(t, t_0, x^0)$ and $p^*(t)$ is the normal to a hyperplane supporting $K(t, t_0, x^0)$ at $x^*(t)$. This normal is obtained by transporting backwards in time, via the adjoint differential equation, the outward normal $p^*(t_f)$ at time t_f. The situation is illustrated in Figure 7.2.

Remark 2: If we define the *Hamiltonian* function H by

$$H(t, x, u, p) = p'(A(t)x + B(t)u) \ ,$$

and we define M by

$$M(t, x, p) = \sup\{H(t, x, u, p) | u \in \Omega\},$$

then (7.5) can be rewritten as

$$H(t, x^*(t), u^*(t), p^*(t)) = M(t, x^*(t), p^*(t)) \ . \tag{7.10}$$

This condition is known as the *maximum principle.*

Exercise 3: (i) Show that $m(t) = M(t, x^*(t), p^*(t))$ is a Lipschitz function of t. (ii) If $A(t), B(t)$ are constant, show that $m(t)$ is constant. (Hint: Show that $(dm/dt) \equiv 0$.)

The next two exercises show how we can obtain important qualitative properties of an optimal control.

Exercise 4: Suppose that Ω is bounded and closed. Show that there exists an optimal control $u^*(\cdot)$ such that $u^*(t)$ belongs to the boundary of Ω for all t.
Exercise 5: Suppose $\Omega = [\alpha, \beta]$, so that $B(t)$ is an $n \times 1$ matrix. Suppose that $A(t) \equiv A$ and $B(t) \equiv B$ are constant matrices and A has n real eigenvalues. Show that there is an optimal control $u^*(\cdot)$ and $t_0 \leq t_1 \leq t_2 \leq \ldots \leq t_n \leq t_f$ such that $u^*(t) \equiv \alpha$ or β on $[t_i, t_{i+1}), 0 \leq i \leq n$. (Hint: first show that $(p^*(t))'B = \gamma_1 \exp(\delta_1 t) + \ldots + \gamma_n \exp(\delta_n(t))$ for some γ_i, δ_i in R.)
Exercise 6: Assume that $K(t_f, t_0, x^0)$ is convex (see remark in Exercise 1 above). Let $f_0 : R^n \to R$ be a differentiable function and suppose that the objective function in (7.2) is $f_0(x(t_f))$ instead of $c'x(t_f)$. Suppose $u^*(\cdot)$ is an optimal control. Show that $u^*(\cdot)$ satisfies the maximum principle (7.10) where $p^*(\cdot)$ is the solution of the adjoint equation (7.3) with the final condition

$$p^*(t_f) = \bigtriangledown f_0(x^*(t_f)) \ .$$

Also show that this condition is sufficient for optimality if f_0 is concave. (Hint: Use Lemma 1 of 5.1.1 to show that if $u^*(\cdot)$ is optimal, then $f_{0x}(x^*(t_f)(x^*(t_f) - x) \leq$ for all $x \in K(t_f, t_0, x^0)$.)

7.2 *More General Boundary Conditions*

We consider the following generalization of (7.2). The notion of the previous section is retained.

$$\begin{array}{l} \text{Maximize } c'x(t_f) \\ \text{subject to} \\ \textit{dynamics: } \dot{x}(t) = A(t)x(t) + B(t)u(t),\ t_0 \leq t \leq t_f \ , \\ \textit{initial condition: } \ G^0 x(t_0) = b^0 \ , \\ \textit{final condition: } \ G^f x(t_f) = b^f \ , \\ \textit{control constraint: } \ u(\cdot) \in \mathcal{U},\ \textit{i.e.},\ \sqcap :\ [\sqcup\!\prime, \sqcup\!\{] \to \otimes \text{ and} \\ \qquad\qquad u(\cdot) \text{piecewise continuous}. \end{array} \tag{7.11}$$

In (7.11) G^0 and G^f are fixed matrices of dimensions $\ell^0 \text{x} n$ and $\ell^f \times n$ respectively, while $b^0 \in R^{\ell_0}, b^f \in R^{\ell_f}$ are fixed vectors.

We will analyze the problem in the same way as before. That is, we first characterize optimality in terms of the state at the final time, and then translate these conditions in terms of the control. For convenience let

$$T^0 = \{z \in R^n | G^0 z = b^0\} \ ,$$
$$T^f = \{z \in R^n | G^f z = b^f\} \ .$$

Definition: Let $p \in R^n$. Let $z^* \in T^0$. We say that p is *orthogonal to* T^0 *at* z^* and we write $p \perp T^0(z^*)$ if

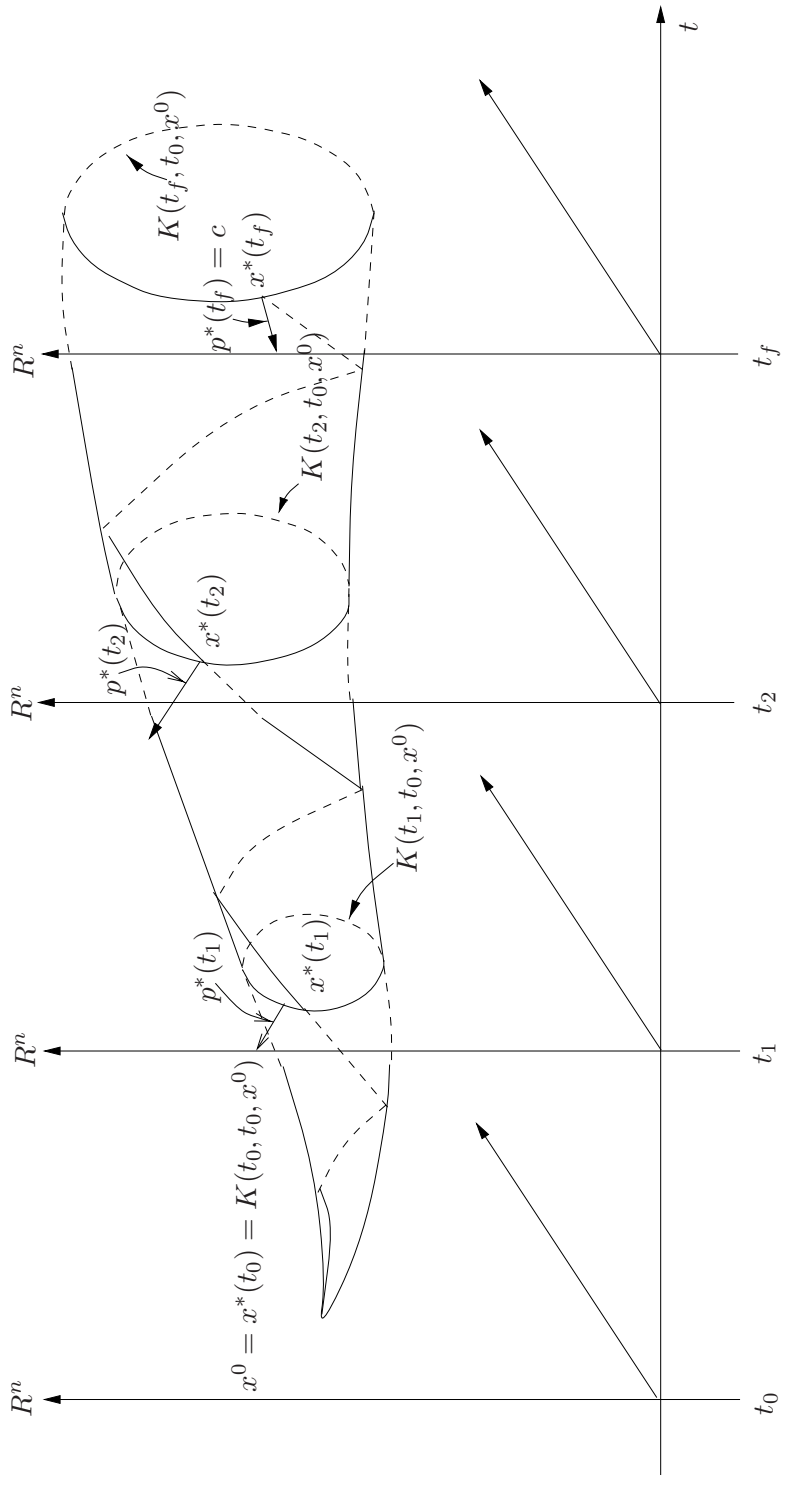


Figure 7.2: Illustration of (7.9) for $t_1 = t_0$.

$$p'(z - z^*) = 0 \quad \text{for all} \quad z \in T^0 \ .$$

Similarly if $z^* \in T^f,\ p \perp T^f(z^*)$ if

$$p'(z - z^*) = 0 \quad \text{for all} \quad z \in T^f \ .$$

Definition: Let $X(t_f) = \{\Phi(t_f, t_0)z + w | z \in T^0,\ w \in K(t_f, t_0, 0)\}$.

Exercise 1: $X(t_f) = \{\Phi(t_f, t_0, z, u) | z \in T^0,\ u(\cdot) \in \mathcal{U}\}$.

Lemma 1: Let $x^*(t_0) \in T^0$ and $u^*(\cdot) \in \mathcal{U}$. Let $x^*(t) = \phi(t, t_0, x^*(t_0), u^*)$, and suppose that $x^*(t_f) \in T^f$.
(i) Suppose the Ω is convex. If $u^*(\cdot)$ is optimal, there exist $\hat{p}_0 \in R, \hat{p}_0 \geq 0$ and $\hat{p} \in R^n$, not both zero, such that

$$(\hat{p}_0 c + \hat{p})' x^*(t_f) \geq (\hat{p}_0 c + \hat{p})' x \ \text{ for all } \ x \in X(t_f) \ , \tag{7.12}$$

$$\hat{p} \perp T^f(x^*(t_f)) \ , \tag{7.13}$$

$$[\Phi(t_f, t_0)]'(\hat{p}_0 c + \hat{p}) \perp T^0(x^*(t_0)) \ . \tag{7.14}$$

(ii) Conversely if there exist $\hat{p}_0 > 0$, and $\hat{p}$ such that (7.12) and (7.13) are satisfied, then $u^*(\cdot)$ is optimal and (7.14) is also satisfied.
Proof: Clearly $u^*(\cdot)$ is optimal iff

$$c' x^*(t_f) \geq c' x \ \text{ for all } \ x \in X(t_f) \cap T^f \ . \tag{7.15}$$

(i) Suppose that $u^*(\cdot)$ is optimal. In R^{1+m} define sets S^1, S^2 by

$$S^1 = \{(r, x) | r > c' x^*(t_f), x \in T^f\} \ , \tag{7.16}$$

$$S^2 = \{(r, x) | r = c' x \ , \ x \in X(t_f)\} \ . \tag{7.17}$$

First of all $S^1 \cap S^2 = \phi$ because otherwise there exists $x \in X(t^f) \cap T^f$ such that $c'x > c'x^*(t_f)$ contradicting optimality of $u^*(\cdot)$ by (7.15).

Secondly, S^1 is convex since T^f is convex. Since Ω is convex by hypothesis it follows by Exercise 1 of Section 1 that S^2 is convex.

But then by the separation theorem for convex sets (see 5.5) there exists $\hat{p}_0 \in R, \hat{p} \in R^n$, not both zero, such that

$$\hat{p}_0 r^1 + \hat{p}' x^1 \geq \hat{p}_0 r^2 + \hat{p}' x^2 \ \text{ for all } \ (r^i, x^i) \in S^i \ , \ i = 1, 2. \tag{7.18}$$

In particular (7.18) implies that

$$\hat{p}_0 r + \hat{p}' x^*(t_f) \geq \hat{p}_0 c' x + \hat{p}' x \ \text{ for all } \ x \in X(t_f), r > c' x^*(t_f). \tag{7.19}$$

Letting $r \to \infty$ we conclude that (7.19) can hold only if $\hat{p}_0 \geq 0$. On the other hand letting $r \to c' x^*(t_f)$ we see that (7.19) can hold only if

$$\hat{p}_0 c' x^*(t_f) + \hat{p}' x^*(t_f) \geq \hat{p}_0 c' x + \hat{p}' x \quad \text{for all } \ x \in X(t_f) \ , \tag{7.20}$$

which is the same as (7.12). Also from (7.18) we get

$$\hat{p}_0 r + \hat{p}\prime x \geq \hat{p}_0 c' x^*(t_f) + \hat{p}' x^*(t_f) \text{ for all } r > c' x^*(t_f), x \in T^f \ ,$$

which can hold only if

$$\hat{p}_1 c' x^*(t_f) + \hat{p}' x \geq \hat{p}_0 c' x^*(t_f) + \hat{p}' x^*(t_f) \text{ for all } x \in T^f \ ,$$

or

$$\hat{p}'(x - x^*(t_f)) \geq 0 \text{ for all } x \in T^f \tag{7.21}$$

But $\{x - x^*(t_f) | x \in T^f\} = \{z | G^f z = 0\}$ is a subspace of R^n, so that (7.21) can hold only if

$$\hat{p}'(x - x^*(t_f)) = 0 \text{ for all } x \in T^f \ ,$$

which is the same as (7.13). Finally (7.12) always implies (7.14), because by the definition of $X(t_f)$ and Exercise 1, $\{\Phi(t_f, t_0)(z - x^*(t_0)) + x^*(t_f)\} \in X(t_f)$ for all $z \in T^0$, so that from (7.12) we get

$$0 \geq (\hat{p}_0 c + \hat{p})' \Phi(t_f, t_0)(z - x^*(t_0)) \text{ for all } z \in T^0 \ ,$$

which can hold only if (7.14) holds.
(ii) Now suppose that $\hat{p}_0 > 0$ and $\hat{p}$ are such that (7.12), (7.13) are satisfied. Let $\tilde{x} \in X(t_f) \cap T^f$. Then from (7.13) we conclude that

$$\hat{p}' x^*(t_f) = \hat{p}' \tilde{x} \ \ ,$$

so that from (7.12) we get

$$\hat{p}_0 c' x^*(t_f) \geq \hat{p}_0 c' \tilde{x} \ ;$$

but then by (7.15) $u^*(\cdot)$ is optimal. $\diamondsuit$

Remark 1: If it is possible to choose $\hat{p}_0 > 0$ then $\hat{p}_0 = 1, \ \hat{p} = (\hat{p}/\hat{p}_0)$ will also satisfy (7.12), (7.13), and (7.14). In particular, in part (ii) of the Lemma we may assume $\hat{p}_0 = 1$.
Remark 2: it would be natural to conjecture that in part (i) $\hat{p}_0$ may be chosen > 0. But in Figure 7.3 below, we illustrate a 2-dimensional situation where $T^0 = \{x^0\}, \ T^f$ is the vertical line, and $T^f \cap X(t_f)$ consists of just one vector. It follows that the control $u^*(\cdot) \in \mathcal{U}$ for which $x^*(t_f) = \phi(t_f, t_0, x^0, u^*) \in T^f$ is *optimal for any c*. Clearly then for some c (in particular for the c in Figure 7.3) we are forced to set $\hat{p}_0 = 0$. In higher dimensions the reasons may be more complicated, but basically if T^f is "tangent" to $X(t_f)$ we may be forced to set $\hat{p}_0 = 0$ (see Exercise 2 below). Finally, we note that part (i) is not too useful if $\hat{p}_0 = 0$, since then (7.12), (7.13), and (7.14) hold for any vector c whatsoever. Intuitively $\hat{p}_0 = 0$ means that it is so difficult to satisfy the initial and final boundary conditions in (7.11) that optimization becomes a secondary matter.
Remark 3: In (i) the convexity of Ω is only used to guarantee that $K(t_f, t_0, 0)$ is convex. But it is known that $K(t_f, t_0, 0)$ is convex even if Ω is not (see Neustadt [1963]).
Exercise 2: Suppose there exists z in the *interior* of $X(t_f)$ such that $z \in T^f$. Then in part (i) we must have $\hat{p}_0 > 0$.

We now translate the conditions obtained in Lemma 1 in terms of the control u^*.

Theorem 1: Let $x^*(t_0) \in T^0$ and $u^*(\cdot) \in \mathcal{U}$. Let $x^*(t) = \phi(t, t_0, x^*(t_0), u^*)$ and suppose that $x^*(t_f) \in T^f$.
(i) Suppose that Ω is convex. If $u^*(\cdot)$ is optimal for (7.11), then there exist a number $p_0^* \geq 0$, and a function $p^* : \ [t_0, t_f] \to R^n$, not both identically zero, satisfying

$$\textit{adjoint equation:} \ \ \dot{p}^*(t) = -A'(t)p^*(t) \ , \ t_0 \leq t \leq t_f \tag{7.22}$$

$$\textit{initial condition:} \ \ p^*(t_0) \perp T^0(x^*(t_0)) \tag{7.23}$$

$$\textit{final condition:} \ \ (p^*(t_f) - p_0^* c) \perp T^f(x^*(t_f)) \ , \tag{7.24}$$

and the *maximum principle*

$$H(t, x^*(t), u^*(t), p^*(t)) = M(t, x^*(t), p^*(t)) \ , \tag{7.25}$$

holds for all $t \in [t_0, t_f]$ except possibly for a finite set.
(ii) Conversely suppose there exist $p_0^* > 0$ and $p^*(\cdot)$ satisfying (7.22), (7.23), (7.24), and (7.25). Then $u^*(\cdot)$ is optimal.

[Here

$$H(t, x, u, p) = p'(A(t)x + B(t)u), \ \ M(t, x, p) = \sup\{H(t, x, v, p) | v \in \Omega\}.]$$

Proof: A repetition of a part of the argument in the proof of Theorem 1 of Section 1 show that if p^* satisfies (7.22), then (7.25) is equivalent to (7.26):

$$(p^*(t_f))' x^*(t_f) \geq (p^*(t_f))' x \ \text{ for all } \ x \in K(t_f, t_0, x^*(t_0)) \ . \tag{7.26}$$

(i) Suppose $u^*(\cdot)$ is optimal and Ω is convex. Then by Lemma 1 there exist $\hat{p} \geq 0, \ \hat{p} \in R^n$, not both zero, such that (7.12), (7.13) and (7.14) are satisfied. Let $p_0^* = \hat{p}_0$ and let $p^*(\cdot)$ be the solution of (7.22) with the final condition

$$p^*(t_f) = p_0^* c + \hat{p} = \hat{p}_0 c + \hat{p} \ .$$

Then (7.14) and (7.13) are respectively equivalent to (7.23) and (7.24), whereas since $K(t_f, t_0, x^*(t_0)) \subset X(t_f)$, (7.26) is implied by (7.12).
(ii) Suppose $p_0^* > 0$ and (7.22), (7.23), (7.24), and (7.26) are satisfied. Let $\hat{p}_0 = p_0^*$ and $\hat{p} = p^*(t_f) - p_0^* c$, so that (7.24) becomes equivalent to (7.13). Next if $x \in X(t_f)$ we have

$$\begin{aligned} (\hat{p}_0 c + \hat{p})' x &= (p^*(t_f))' x \\ &= (p^*(t_f))' (\Phi(t_f, t_0) z + w) \ , \end{aligned}$$

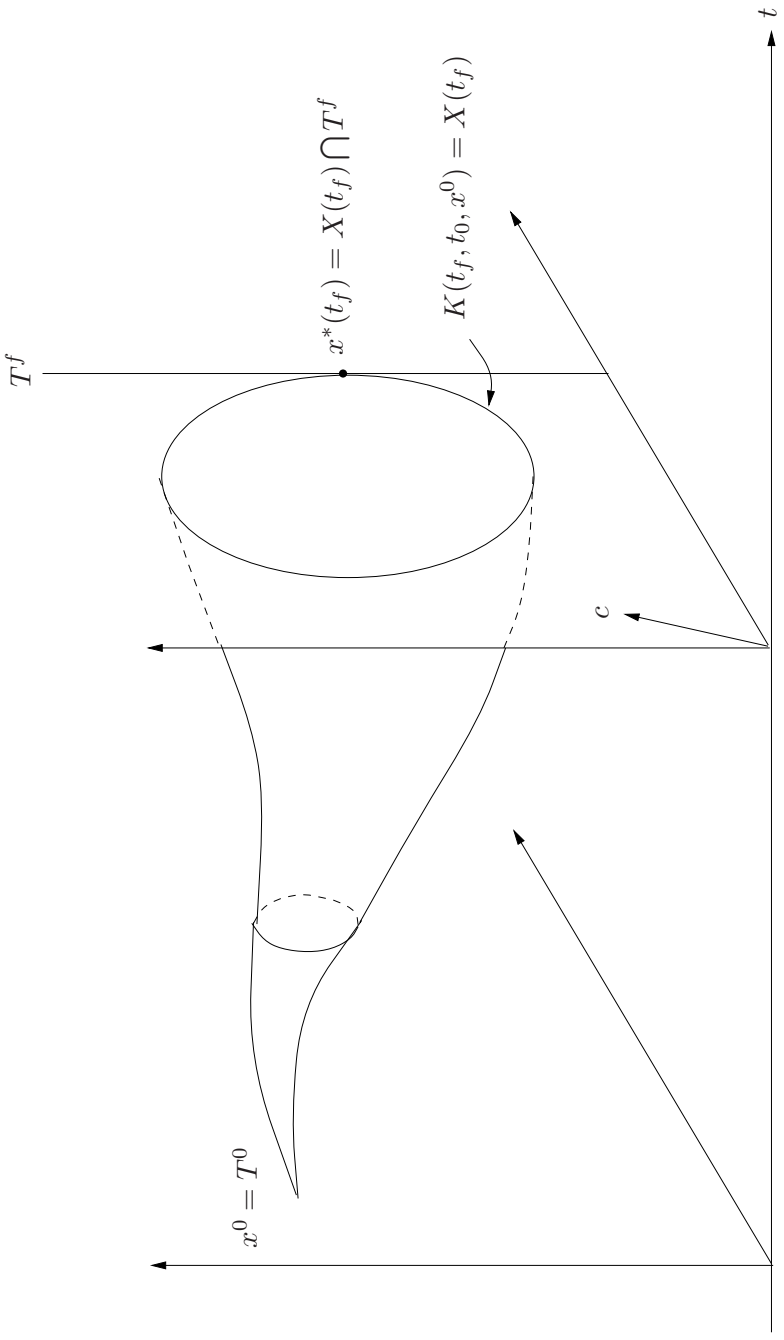


Figure 7.3: Situation where $\hat{p}_0 = 0$

for some $z \in T^0$ and some $w \in K(t_f, t_0, 0)$. Hence

$$\begin{aligned}(\hat{p}_0 c + \hat{p})' x &= (p^*(_f))' \Phi(t_f, t_0)(z - x^*(t_0)) \\ &\quad + (p^*(t_f))'(w + \phi(t_f, t_0) x^*(t_0)) \\ &= (p^*(t_0))'(z - x^*(t_0)) \\ &\quad + (p^*(t_f))'(w + \Phi(t_f, t_0) x^*(t_0)) \ .\end{aligned}$$

But by (7.23) the first term on the right vanishes, and since $(w+\phi(t_f, t_0)x^*(t_0)) \in K(t_f, t_0, x^*(t_0))$, it follows from (7.26) that the second term is bounded by $(p^*(t_f))'x^*(t_f)$. Thus

$$(\hat{p}_0 c + \hat{p})' x^*(t_f) \geq (\hat{p}_0 c + \hat{p})' x \ \text{ for all } \ x \in X(t_f) \ ,$$

and so $u^*(\cdot)$ is optimal by Lemma 1. $\diamondsuit$

Exercise 3: Suppose that the control constraint set is $\Omega(t)$ which varies continuously with t, and we require that $u(t) \in \Omega(t)$ for all t. Show that Theorem 1 also holds for this case where, in (7.25), $M(t, x, p) = \sup\{H(t, x, v, p) | v \in \Omega(t)\}$.

Exercise 4: How would you use Exercise 3 to solve Example 3 of Chapter 1?

Chapter 8

SEQUENTIAL DECISION PROBLEMS: CONTINUOUS-TIME OPTIMAL CONTROL OF NONLINEAR SYSTEMS

We now present a sweeping generalization of the problem studied in the last chapter. Unfortunately we are forced to omit the proofs of the results since they require a level of mathematical sophistication beyond the scope of these *Notes*. However, it is possible to convey the main ideas of the proofs at an intuitive level and we shall do so. (For complete proofs see (Lee and Markus [1967] or Pontryagin, *et al.*, [1962].) The principal result, which is a direct generalization of Theorem 1 of 7.2 is presented in Section 1. An alternative form of the objective function is discussed in Section 2. Section 3 deals with the minimum-time problem and Section 4 considers the important special case of linear systems with quadratic cost. Finally, in Section 5 we discuss the so-called singular case and also analyze Example 4 of Chapter 1.

8.1 *Main Results*

8.1.1 *Preliminary results based on differential equation theory.*

We are interested in the optimal control of a system whose dynamics are governed by the nonlinear differential equation

$$\dot{x}(t) = f(t, x, (t), u(t)) \quad , \quad t_0 \leq t \leq t_f \ , \tag{8.1}$$

where $x(t) \in R^n$ is the state and $u(t) \in R^p$ is the control. Suppose $u^*(\cdot)$ is an optimal control and $x^*(\cdot)$ is the corresponding trajectory. In the case of linear systems we obtained the necessary conditions for optimality by comparing $x^*(\cdot)$ with trajectories $x(\cdot)$ corresponding to other admissible controls $u(\cdot)$. This comparison was possible because we had an explicitly characterization of $x(\cdot)$ in terms of $u(\cdot)$. Unfortunately when f is nonlinear such a characterization is not available. Instead we shall settle for a comparison between the trajectory $x^*(\cdot)$ and trajectories $x(\cdot)$ obtained by perturbing the control $u^*(\cdot)$ and the initial condition $x^*(t_0)$. We can then estimate the difference between $x(\cdot)$ and $x^*(\cdot)$ by the solution to a linear differential equation as shown in Lemma 1 below. But first we need to impose some regularity conditions on the differential equation (8.1). We assume throughout that the function $f : \ [t_0, t_f] \times R^n \times R^p \rightarrow R^n$ satisfies the following conditions:

1. for each fixed $t \in [t_0, t_f]$, $f(t,\cdot,\cdot) : \ R^n \text{x} R^p \rightarrow R^n$ is continuously differentiable in the remaining variables (x,u),

2. except for a finite subset $D \subset [t_0, t_f]$, the functions f, f_x, f_u are continuous on $[t_0, t_f] \times R^n \times R^p$, and

3. for every finite α, there exist finite number β and γ such that

$$|f(t,x,u)| \leq \beta + \gamma|x| \text{ for all } t \in [t_0, t_f],\ x \in R^n,\ u \in R^p \text{ with } |u| \leq \alpha \ .$$

The following result is proved in every standard treatise on differential equations.

Theorem 1: For every $z \in R^n$, for every $t_1 \in [t_0, t_f]$, and every piecewise continuous function $u(\cdot) : \ [t_0, t_f] \rightarrow R^p$, there exists a unique solution

$$x(t) = \phi(t, t_1, z, u(\cdot)) \ , \ t_1 \leq t \leq t_f \ ,$$

of the differential equation

$$\dot{x}(t) = f(t, x(t), u(t)) \ , \ t_1 \leq t \leq t_f \ ,$$

satisfying the initial condition

$$x(t_1) = z \ .$$

Furthermore, for fixed $t_1 \leq t_2$ in $[t_0, t_f]$ and fixed $u(\cdot)$, the function $\phi(t_2, t_1, \cdot, u(\cdot)) : \ R^n \rightarrow R^n$ is differentiable. Moreover, the $n \times n$ matrix-valued function Φ defined by

$$\Phi(t_2, t_1, z, u(\cdot)) = \frac{\partial \phi}{\partial z}(t_2, t_1, z, u(\cdot))$$

is the solution of the linear homogeneous differential equation

$$\frac{\partial \Phi}{\partial t}(t, t_1, z, u, (\cdot)) = [\frac{\partial f}{\partial x}(t, x, (t), u(t))]\Phi(t, t_1, z, u(\cdot)),\ t_1 \leq t \leq t_f \ ,$$

and the initial condition

$$\Phi(t_1, t_1, z, u(\cdot)) = I_n \ \ .$$

Now let $\Omega \subset R^p$ be a fixed set and let $\mathcal{U}$ be set of all piecewise continuous functions $u(\cdot) : [t_0, t_f] \rightarrow \Omega$. Let $u^*(\cdot) \in \mathcal{U}$ be fixed and let D^* be the set of discontinuity points of $u^*(\cdot)$. Let $x_0^* \in R^n$ be a fixed initial condition.

Definition: $\pi = (t_1, \ldots, t_m;\ \ell_1, \ldots, \ell_m;\ u_1, \ldots, u_m)$ is said to be a *perturbation data* for $u^*(\cdot)$ if

1. m is a nonnegative integer,

2. $t_0 < t_1 < t_2 < \ldots t_m < t_f$, and $t_i \notin D^* \bigcup D,\ i = 1, \ldots, m$ (recall that D is the set of discontinuity points of f),

3. $\ell_i \geq 0,\ i = 1, \ldots, m,$ and

4. $u_i \in \Omega,\ i = 1, \ldots, m.$

Let $\varepsilon(\pi) > 0$ be such that for $0 \leq \varepsilon \leq \varepsilon(\pi)$ we have $[t_i - \varepsilon\ell_i, t_i] \subset [t_0, t_f]$ for all i, and $[t_i - \varepsilon\ell_i, t_i] \bigcap [t_j - \varepsilon\ell_j, t_j] = \phi$ for $i \neq j$. Then for $0 \leq \varepsilon \leq \varepsilon(\pi)$,*the perturbed control* $u_{(\pi,\varepsilon)}(\cdot) \in \mathcal{U}$ *corresponding to* π is defined by

$$u_{(\pi,\varepsilon)}(t) = \begin{cases} u_i \text{ for all } t \in [t_i - \varepsilon\ell_i, t_i] \ , \ i = 1, \ldots, m \\ u^*(t) \text{ otherwise } . \end{cases}$$

Definition: Any vector $\xi \in R^n$ is said to be a *perturbation* for x_0^*, and a function $x_{(\xi,\varepsilon)}$ defined for $\varepsilon > 0$ is said to be a *perturbed initial condition* if

$$\lim_{\varepsilon \to 0} \quad x_{(\xi,\varepsilon)} = x_0^* \ ,$$

and

$$\lim_{\varepsilon \to 0} \quad \tfrac{1}{\varepsilon}(x_{(\xi,\varepsilon)} - x_0^*) = \xi \ .$$

Now let $x^*(t) = \phi(t, t_0, x_0^*, u^*(\cdot))$ and let $x_\varepsilon(t) = \phi(t, t_0, x_{(\xi,\varepsilon)}, u_{(\pi,\varepsilon)}(\cdot))$. Let $\Phi(t_2, t_1) = \Phi(t_2, t_1, x^*(t_1), \ u^*(\cdot))$. The following lemma gives an estimate of $x^*(t) - x_\varepsilon(t)$. The proof of the lemma is a straightforward exercise in estimating differences of solutions to differential equations, and it is omitted (see for example (Lee and Markus [1967])).

Lemma 1: $\lim_{\varepsilon \to 0} |x_\varepsilon(t) - x^*(t) - \varepsilon h_{(\pi,\varepsilon)}(t)| = 0$ for $t \in [t_0, t_1]$, where $h_{(\pi,\varepsilon)}(\cdot)$ is given by

$$\begin{aligned} h_{(\pi,\varepsilon)}(t) &= \Phi(t, t_0)\xi \qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad , \ t \in [t_0, t_1) \\ &= \Phi(t, t_0)\xi + \Phi(t, t_1)[f(t_1, x^*(t_1), u_1) - f(t_1, x^*(t_1), u^*(t_1))]\ell_1 \ , \ t \in [t_1, t_2) \\ &= \Phi(t, t_0)\xi + \sum_{j=1}^{i} \Phi(t, t_j)[f(t_j, x^*(t_j), u_j) - f(t_j, x^*(t_j), u^*(t_j))]\ell_j \ , \ t \in [t_i, t_{i+1}) \\ &= \Phi(t, t_0)\xi + \sum_{j=1}^{m} \Phi(t, t_m)[f(t_j, x^*(t_j), u_j) - f(t_j, x^*(t_j), u^*t_j))]\ell_j \ , \ t \in [t_m, t_f] \ . \end{aligned}$$

(See Figure 8.1.)

We call $h_{(\pi,\xi)}(\cdot)$ the *linearized (trajectory) perturbation corresponding to* (π, ξ).
Definition: For $z \in R^n, \ t \in [t_0, t_f]$ let

$$K(t, t_0, z) = \{\phi(t, t_0, z, u(\cdot)) | u(\cdot) \in \mathcal{U}\}$$

be the set of states reachable at time t, starting at time t_0 in state z, and using controls $u(\cdot) \in \mathcal{U}$.
Definition: For each $t \in [t_0, t_f]$, let

$$\begin{aligned} Q(t) = \{h_{(\pi,0)}(t) | \pi &\text{is a perturbation data for } u^*(\cdot), \text{and} \\ &h_{(\pi,0)}(\cdot) \text{is the linearized perturbation} \\ &\text{corresponding to} (\pi, 0)\} \ . \end{aligned}$$

Remark: By Lemma 1 $(x^*(t) + \varepsilon h_{(\pi,\xi)})$ belongs to the set $K(t, t_0, x_{(\xi,\varepsilon)})$ up to an error of order $o(\varepsilon)$. In particular for $\xi = 0$, the set $x^*(t) + Q(t)$ can serve as an approximation to the set $K(t, t_0, x_0^*)$. More precisely we have the following result which we leave as an exercise.

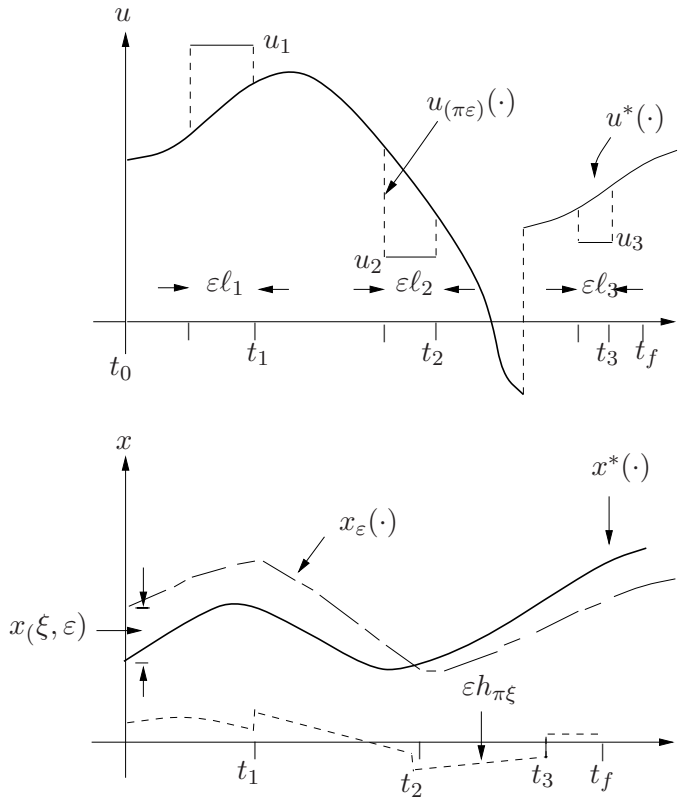


Figure 8.1: Illustration for Lemma 1.

Exercise 1: (Recall the definition of the tangent cone in 5.1.1.) Show that

$$Q(t) \subset C(K(t, t_0, x_0^*), x^*(t)) \ . \tag{8.2}$$

We can now prove a generalization of Theorem 1 of 7.1.
Theorem 2: Consider the optimal control problem (8.3):

$$\begin{array}{l} \text{Maximize } \psi(x(t_f)) \\ \text{subject to} \\ \textit{dynamics: } \dot{x}(t) = f(t, x(t), u(t)) \ , \ t_0 \le t \le t_f \ , \\ \textit{initial condition: } x(t_0) = x_0^* \ , \\ \textit{final condition: } x(t_f) \in R^n \ , \\ \textit{control constraint: } u(\cdot) \in \mathcal{U}, \textit{ i.e., } u: \ [t_0, t_f] \to \Omega \text{ and} \\ \qquad\qquad u(\cdot) \text{ piecewise continuous} \ , \end{array} \tag{8.3}$$

where $\psi: \ R^n \to R$ is differentiable and f satisfies the conditions listed earlier.

Let $u^*(\cdot) \in \mathcal{U}$ be an optimal control and let $x^*(t) = \phi(t, t_0, x_0^*, u^*(\cdot)), \ t_0 \le t \le t_f$, be the corresponding trajectory. Let $p^*(t), \ t_0 \le t \le t_f$, be the solution of (8.4) and (8.5):

$$\textit{adjoint equation: } \dot{p}^*(t) = -[\tfrac{\partial f}{\partial x}(t, x^*(t), u^*(t))]' p^*(t), \ t_0 \le t \le t_f \ , \tag{8.4}$$

$$\textit{final condition: } p^*(t_f) = \bigtriangledown\psi(x^*(t_f)) \ . \tag{8.5}$$

Then $u^*(\cdot)$ satisfies the *maximum principle*

$$H(t, x^*(t), u^*(t), p^*(t)) = M(t, x^*(t), p^*(t)) \tag{8.6}$$

for all $t \in [t_0, t_f]$ except possibly for a finite set. [Here $H(t, x, u, p) = p'f(t, x, u,),\ M(t, x, p) = \sup\{H(t, x, v, p) | v \in \Omega\}$].
Proof: Since $u^*(\cdot)$ is optimal we must have

$$\psi(x^*(t_f)) \geq \psi(z) \text{ for all } z \in K(t_f, t_0, x_0^*) \ ,$$

and so by Lemma 1 of 5.1.1

$$\psi(x^*(t_f))h \leq 0 \text{ for all } h \in C(K(t_f, t_0, x_0^*), x^*(t_f)) \ ,$$

and in particular from (8.2)

$$\psi_x(x^*(t_f))h \leq 0 \text{ for all } h \in Q(t_f) \ . \tag{8.7}$$

Now suppose that (8.6) does not hold from some $t^* \notin D^* \cup D$. Then there exists $v \in \Omega$ such that

$$p^*(t^*)'[f(t^*, x(t^*), v) - f(t^*, x(t^*), u^*(t^*))] > 0 \ . \tag{8.8}$$

If we consider the perturbation data $\pi = (t^*; 1; v)$, then (8.8) is equivalent to

$$p^*(t^*)'h_{(\pi,0)}(t^*) > 0 \ . \tag{8.9}$$

Now from (8.4) we can see that $p^*(t^*)' = p^*(t_f)'\Phi(t_f, t^*)$. Also $h_{(\pi,0)}(t_f) = \Phi(t_f, t^*)h_{(\pi,0)}(t^*)$ so that (8.9) is equivalent to

$$p^*(t_f)'h_{(\pi,0)}(t_f) > 0$$

which contradicts (8.7). ♢

8.1.2 *More general boundary conditions.*

In Theorem 2 the initial condition is fixed and the final condition is free. The problem involving more general boundary conditions is much more complicated and requires more refined analysis. Specifically, Lemma 1 needs to be extended to Lemma 2 below. But first we need some simple properties of the sets $Q(t)$ which we leave as exercises.

Exercise 2: Show that
(i) $Q(t)$ is a cone, *i.e.*, if $h \in Q(t)$ and $\lambda \geq 0$, then $\lambda h \in Q(t)$,
(ii) for $t_0 \leq t_1 \leq t_2 \leq t_f$, $\Phi(t_2, t_1)Q(t_1) \subset Q(t_2)$.
Definition: Let $C(t)$ denote the closure of $Q(t)$.
Exercise 3: Show that
(i) $C(t)$ is a convex cone,
(ii) for $t_0 \leq t_1 \leq t_2 \leq t_f$, $\Phi(t_2, t_1)C(t_1) \subset C(t_2)$.
Remark: From Lemma 1 we know that if $h \in C(t)$ then $(x^*(t) + \varepsilon h)$ belongs to $K(t, t_0, x^*(t_0))$ up to an error of order $o(\varepsilon)$. Lemma 2, below, asserts further that if h is in the interior of $C(t)$ then in fact $(x^*(t) + \varepsilon h) \in K(t, t_0, x^*(t_0))$ for $\varepsilon > 0$ sufficiently small. The proof of the lemma depends upon some deep topological results and is omitted. Instead we offer a plausibility argument.
Lemma 2: Let h belong to the interior of the cone $C(t)$. Then for all $\varepsilon > 0$ sufficiently small,

$$(x^*(t) + \varepsilon h) \in K(t, t_0, x_0^*) \ . \tag{8.10}$$

Plausibility argument. (8.10) is equivalent to

$$\varepsilon h \in K(t, t_0, x^*(t_0)) - \{x^*(t)\} \ , \tag{8.11}$$

where we have moved the origin to $x^*(t)$. The situation is depicted in Figure 8.2.

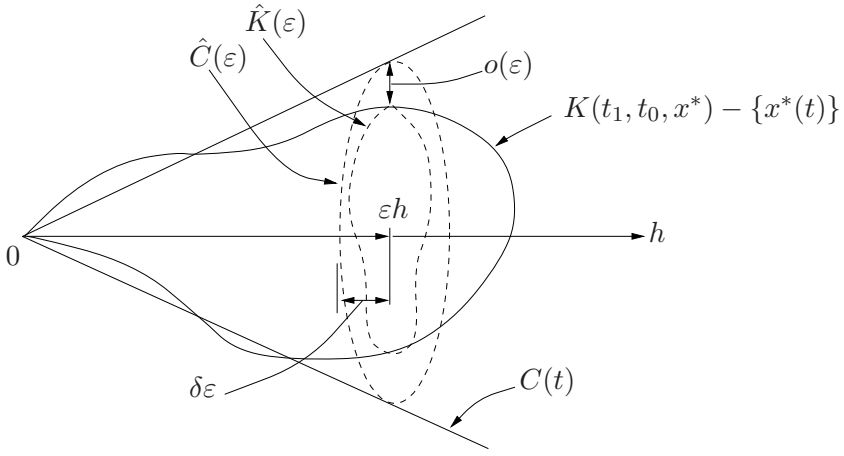


Figure 8.2: Illustration for Lemma 2.

Let $\hat{C}(\varepsilon)$ be the cross-section of $C(t)$ by a plane orthogonal to h and passing through εh. Let $\hat{K}(\varepsilon)$ be the cross-section of $K(t, t_0, x_0^*) - \{x^*(t_0)\}$ by the same plane. We note the following:
(i) by Lemma 1 the distance between $\hat{C}(\varepsilon)$ and $\hat{K}(\varepsilon)$ is of the order $o(\varepsilon)$;
(ii) since h is in the interior of $C(t)$, the minimum distance between εh and $\hat{C}(\varepsilon)$ is $\delta\varepsilon$ where $\delta > 0$ is independent of ε.
Hence for $\varepsilon > 0$ sufficiently small εh must be trapped inside the set $\hat{K}(\varepsilon)$.

(This would constitute a proof except that for the argument to work we need to show that there are no "holes" in $\hat{K}(\varepsilon)$ through which εh can "escape." The complications in a rigorous proof arise precisely from this drawback in our plausibility argument.) $\diamondsuit$

Lemmas 1 and 2 give us a characterization of $K(t, t_0, x_0^*)$ in a neighborhood of $x^*(t)$ when we perturb the control $u^*(\cdot)$ leaving the initial condition fixed. Lemma 3 extends Lemma 2 to the case when we also allow the initial condition to vary over a fixed surface in a neighborhood of x_0^*.

Let $g^0 : \ R^n \to R^{\ell_0}$ be a differentiable function such that the $\ell_0 \times n$ matrix $g_x^0(x)$ has rank ℓ_0 for all x. Let $b^0 \in R^n$ be fixed and let $T^0 = \{x | g^0(x) - b^0\}$. Suppose that $x_0^* \in T^0$ and let $T^0(x_0^*) = \{\xi | g_x^0(x_0^*)\xi = 0\}$. Thus, $T^0(x_0^*) + \{x_0^*\}$ is the plane through x_0^* tangent to the surface T^0. The proof of Lemma 3 is similar to that of Lemma 2 and is omitted also.
Lemma 3: Let h belong to the interior of the cone $\{C(t) + \Phi(t, t_0)T^0(x_0^*)\}$. For $\varepsilon \geq 0$ let $h(\varepsilon) \in R^n$ be such that $\lim h(\varepsilon) = 0$, and $\lim_{\varepsilon \to 0}(\frac{1}{\varepsilon})h(\varepsilon) = h$. Then for $\varepsilon > 0$ sufficiently small there exists $x_0(\varepsilon) \in T^0$ such that

$$(x^*(t) + h(\varepsilon)) \in K(t, t_0, x_0(\varepsilon)) \ .$$

We can now prove the main result of this chapter. We keep all the notation introduced above. Further, let $g^f : \quad R^n \to R^{\ell_f}$ be a differentiable function such that $g^f_x(x)$ has rank ℓ_f for all x. Let $b^f \in R^n$ be fixed and let $T^f = \{x | g^f(x) - b^f\}$. Finally, if $x^*(t_f) \in T^f$ let $T^f(x^*(t_f)) = \{\xi | g^f_x(x^*(t_f))\xi = 0\}$.

Theorem 3: Consider the optimal control problem (8.12):

$$\begin{array}{l} \text{Maximize } \psi(x(t_f)) \\ \text{subject to} \\ \textit{dynamics: } \dot{x}(t) = f(t, x(t), u(t)) \ , \ t_0 \le t \le t_f \ \ , \\ \textit{initial conditions: } \ g^0(x(t_0)) = b^0 \ \ , \\ \textit{final conditions: } \ g^f(x(t_f)) = b^f \ \ , \\ \textit{control constraint: } \ u(\cdot) \in \mathcal{U}, \textit{i.e.}, \ u : \ [t_0, t_f] \to \Omega \text{ and} \\ \qquad\qquad u(\cdot) \text{ piecewise continuous} \ \ . \end{array} \tag{8.12}$$

Let $u^*(\cdot) \in \mathcal{U}$, let $x_0^* \in T^0$ and let $x^*(t) = \phi(t, t_0, x_0^*, u^*(\cdot))$ be the corresponding trajectory. Suppose that $x^*(t_f) \in T^f$, and suppose that $(u^*(\cdot), x_0^*)$ is optimal. Then there exist a number $p_0^* \ge 0$, and a function $p^* : \ [t_0, t_f] \to R^n$, not both identically zero, satisfying

$$\textit{adjoint equation: } \ \dot{p}^*(t) = -[\tfrac{\partial f}{\partial x}(t, x^*(t), u^*(t))]' p^*(t), \ t_0 \le t \le t_f \ , \tag{8.13}$$

$$\textit{initial condition: } \ p^*(t_0) \perp T^0(x_0^*) \ \ , \tag{8.14}$$

$$\textit{final condition: } \ (p^*(t_f) - p_0^* \nabla \psi(x^*(t_f))) \perp T^f(x^*(t_f)) \ . \tag{8.15}$$

Furthermore, the *maximum principle*

$$H(t, x^*(t), u^*(t), p^*(t)) = M(t, x^*(t), p^*(t)) \tag{8.16}$$

holds for all $t \in [t_0, t_f]$ except possibly for a finite set. [Here $H(t, x, p, u) = p' f(t, x, u,), \ M(t, x, p) = \sup\{H(t, x, v, p) | v \in \Omega\}$].

Proof: We break the proof up into a series of steps.
Step 1. By repeating the argument presented in the proof of Theorem 2 we can see that (8.15) is equivalent to

$$p^*(t_f)' h \le 0 \ \text{ for all } \ h \in C(t_f) \ . \tag{8.17}$$

Step 2. Define two convex sets S_1, S_2 in R^{1+m} as follows:
$S_1 = \{(y, h) | y > 0, \ h \in T^f(x^*(t_f))\}$,
$S_2 = \{(y, h) | y = \psi_x(x^*(t_f))h, \ h \in \{C(t_f) + \Phi(t_f, t_0) T^0(x_0^*)\}\}$.
We claim that the optimality of $(u^*(\cdot), x_0^*)$ implies that $S_1 \cap$ Relative Interior $(S_2) = \phi$. Suppose this is not the case. Then there exists $h \in T^f(x^*(t_f))$ such that

$$\psi_x(x^*(t_f))h > 0 \ \ , \tag{8.18}$$

$$h \in \text{Interior}\{C(t_f) + \Phi(t_f, t_0)T^0(x_0^*)\} \ . \tag{8.19}$$

Now by assumption $g_x^f(x^*(t_f)$ has maximum rank. Since $g_x^f(x^*(t_f))h = 0$ it follows that the Implicit Function Theorem that for $\varepsilon > 0$ sufficiently small there exists $h(\varepsilon) \in R^n$ such that

$$g^f(x^*(t_f) + h(\varepsilon)) = b^f \ \ , \tag{8.20}$$

and, moreover, $h(\varepsilon) \to 0,\ (1/\varepsilon)h(\varepsilon) \to h$ as $\varepsilon \to 0$. From (8.18) and Lemma 3 it follows that for $\varepsilon > 0$ sufficiently small there exists $x_0(\varepsilon) \in T^0$ and $u_\varepsilon(\cdot) \in \mathcal{U}$ such that

$$x^*(t_f) + h(\varepsilon) = \phi(t_f, t_0, x_0(\varepsilon), u_\varepsilon(\cdot)) \ \ .$$

Hence we can conclude from (8.20) that the pair $(x_0(\varepsilon), u_\varepsilon(\cdot))$ satisfies the initial and final conditions, and the corresponding value of the objective function is

$$\psi(x^*(t_f) + h(\varepsilon)) = \psi(x^*(t_f)) + \psi_x(x^*(t_f))h(\varepsilon) + o(|h(\varepsilon)|) \ ,$$

and since $h(\varepsilon) = \varepsilon h + o(\varepsilon)$ we get

$$\psi(x^*(t_f) + h(\varepsilon)) = \psi(x^*(t_f)) + \varepsilon)\psi_x(x^*(t_f))h + o(\varepsilon) \ \ ;$$

but then from (8.18)

$$\psi(x^*(t_f) + h(\varepsilon)) > \psi(x^*(t_f))$$

for $\varepsilon > 0$ sufficiently small, thereby contradicting the optimality of $(u^*(\cdot), x_0^*)$.
Step 3. By the separation theorem for convex sets there exist $\hat{p}_0 \in R, \hat{p}_1 \in R^n$, not both zero, such that

$$\hat{p}_0 y^1 + \hat{p}_1' h^1 \geq \hat{p}_0 y^2 + \hat{p}_1' h^2 \text{ for all } (y^i, h^i) \in S_1 \ ,\ i = 1, 2 \ \ . \tag{8.21}$$

Arguing in exactly the same fashion as in the proof of Lemma 1 of 7.2 we can conclude that (8.21) is equivalent to the following conditions:

$$\begin{array}{l} \hat{p}_0 \geq 0 \ \ , \\ \hat{p}_1 \perp T^f(x^*(t_f)) \ \ , \end{array} \tag{8.22}$$

$$\Phi(t_f, t_0)'(\hat{p}_0 \nabla\psi(x^*(t_f)) + \hat{p}_1) \perp T^0(x_0^*) \ \ , \tag{8.23}$$

and

$$(\hat{p}_0\psi_x(x^*(t_f)) + \hat{p}_1')h \leq 0 \text{ for all } h \in C(t_f) \ \ . \tag{8.24}$$

If we let $\hat{p}_0^* = \hat{p}_0$ and $p^*(t_f) = \hat{p}_0 \nabla\psi(x^*(t_f)) + \hat{p}_1$ then (8.22), (8.23), and (8.24) translate respectively into (8.15), (8.14), and (8.17). $\diamondsuit$

8.2 *Integral Objective Function*

In many control problems the objective function is not given as a function $\psi(x(t_f))$ of the final state, but rather as an integral of the form

$$\int_{t_0}^{t_f} f_0(t, x(t), u(t))dt \quad . \tag{8.25}$$

The dynamics of the state, the boundary conditions, and control constraints are the same as before. We proceed to show how such objective functions can be treated as a special case of the problems of the last section. To this end we defined the *augmented system* with state variable $\tilde{x} = (x_0, x) \in R^{1+m}$ as follows:

$$\dot{\tilde{x}} = \begin{bmatrix} \dot{x}_0(t) \\ \dot{x}(t) \end{bmatrix} = \tilde{f}(t, \tilde{x}(t), u(t)) = \begin{bmatrix} f_0(t, x(t), u(t)) \\ f(t, x(t), u(t)) \end{bmatrix} \quad .$$

The initial and final conditions which are of the form

$$g^0(x) = b^0, \; g^f(x) = b^f \text{ are augmented } \tilde{g}^0(\tilde{x}) = \begin{bmatrix} x_0 \\ g^0(x) \end{bmatrix} = \tilde{b}^0 = \begin{bmatrix} 0 \\ b^0 \end{bmatrix}$$

and $\tilde{g}^f(\tilde{x}) = g^f(x) = b^f$. Evidently then the problem of maximizing (8.25) is equivalent to the problem of maximizing

$$\psi(\tilde{x}(t_f)) = x_0(t_f) \quad ,$$

subject to the augmented dynamics and constraints which is of the form treated in Theorem 3 of Section 1, and we get the following result.

Theorem 1: Consider the optimal control problem (8.26):

$$\begin{aligned} &\text{Maximize } \int_{t_0}^{t_f} f_0(t, x(t), u(t))dt \\ &\text{subject to} \\ &\textit{dynamics: } \dot{x}(t) = f(t, x(t), u(t)), \; t_0 \le t \le t_f \quad , \\ &\textit{initial conditions: } g^0(x(t_0)) = b^0 \quad , \\ &\textit{final conditions: } g^f(x(t_f)) = b^f \quad , \\ &\textit{control constraint: } u(\cdot) \in \mathcal{U} \quad . \end{aligned} \tag{8.26}$$

Let $u^*(\cdot) \in \mathcal{U}$, let $x_0^* \in T^o$ and let $x^*(t) = \phi(t, t_0, x_0^*, u^*(\cdot))$, and suppose that $x^*(t_f) \in T^f$. If $(u^*(\cdot), x_0^*)$ is optimal, then there exists a function $\tilde{p}^* = (p_0^*, p^*) : \; [t_0, t_f] \to R^{1+m}$, not identically zero, and with $p_0^*(t) \equiv$ constant and $p_0^*(t) \ge 0$, satisfying

(augmented) adjoint equation: $\dot{\tilde{p}}^*(t) = -[\frac{\partial \tilde{f}}{\partial \tilde{x}}(t, x^*(t), u^*(t))]' \tilde{p}^*(t)$,
initial condition: $p^*(t_0) \perp T^0(x_0^*)$,
final condition: $p^*(t_f) \perp T^f(x^*(t_f))$.
Futhermore, the *maximum principle*

$$\tilde{H}(t, x^*(t), \tilde{p}^*(t), u^*(t)) = \tilde{M}(t, x^*(t), \tilde{p}^*(t))$$

holds for all $t \in [t_0, t_f]$ except possibly for a finite set. [Here $\tilde{H}(t, x, \tilde{p}, u) = \tilde{p}' \tilde{f}(t, x, u) = p_0 f_0(t, x, u) + p' f(t, x, u)$, and $\tilde{M}(t, x, \tilde{p}) = \sup\{\tilde{H}(t, x, \tilde{p}, v) | v \in \Omega\}$.]

Finally, if f_0 and f do not explicitly depend on t, then $\tilde{M}(t, x^*(t), \tilde{p}^*(t)) \equiv$ constant.

Exercise 1: Prove Theorem 1. (Hint: For the final part show that $(d/dt)\; \tilde{M}(t, x^*(t), \tilde{p}^*(t)) \equiv 0$.)

8.3 *Variable Final Time*

8.3.1 *Main result.*

In the problem considered up to now the final time t_f is assumed to be fixed. In many important cases the final time is itself a decision variable. One such case is the minimum-time problem where we want to transfer the state of the system from a given initial state to a specified final state in minimum time. More generally, consider the optimal control problem (8.27).

$$\begin{array}{l} \text{Maximize } \int_{t_0}^{t_f} f_0(t,x(t),u(t))dt \\ \text{subject to} \\ \textit{dynamics: } \dot{x}(t) = f(t,x,(t),u(t)),\ ,\ t_0 \le t \le t_f\ , \\ \textit{initial condition: } g^0(x(t_0)) = b^0\ , \\ \textit{final condition: } g^f(x(t)f)) = b^f\ , \\ \textit{control constraint: } u(\cdot) \in \mathcal{U}\ , \\ \textit{final-time constraint: } t_f \in (t_0,\infty)\ . \end{array} \tag{8.27}$$

We analyze (8.27) by converting the variable time interval $[t_0, t_f]$ into a fixed-time interval $[0, 1]$. This change of time-scale is achieved by regarding t as a new state variable and selecting a new time variable s which ranges over $[0, 1]$. The equation for t is

$$\tfrac{dt(s)}{ds} = \alpha(s)\ ,\ 0 \le s \le 1\ ,$$

with initial condition

$$t(0) = t_0\ .$$

Here $\alpha(s)$ is a new control variable constrained by $\alpha(s) \in (0,\infty)$. Now if $x(\cdot)$ is the solution of

$$\dot{x}(t) = f(t,x(t),u(t))\ ,\ t_0 \le t \le t_f\ ,\ x(t_0) = x_0 \tag{8.28}$$

and if we define

$$z(s) = x(t(s)),\ v(s) = u(t(s))\ ,\ 0 \le s \le 1\ ,$$

then it is easy to see that $z(\cdot)$ is the solution of

$$\tfrac{dz}{ds}(s) = \alpha(s)f(s,z(s),v(s))\ ,\ 0 \le s \le 1\ \ z(0) = x_0\ . \tag{8.29}$$

Conversely from the solution $z(\cdot)$ of (8.29) we can obtain the solution $x(\cdot)$ of (8.28) by

$$x(t) = z(s(t))\ \ ,\ \ t_0 \le t \le t_f\ \ ,$$

where $s(\cdot)\ :\ \ [t_0,t_f] \to [0,1]$ is the functional inverse of $s(t)$; in fact, $s(\cdot)$ is the solution of the differential equation $\dot{s}(t) = 1/\alpha(s(t)),\ s(t_0) = 0$.

With these ideas in mind it is natural to consider the fixed-final-time optimal control problem (8.30), where the state vector $(t, z) \in R^{1+m}$, and the control $(\alpha, v) \in R^{1+p}$:

$$\begin{aligned} &\text{Maximize} \quad \int_0^1 f_0(t(s), z(s), v(s))\alpha(s)ds \\ &\text{subject to} \\ &\textit{dynamics:} \ \ (\dot{z}(s), \dot{t}(s)) = (f(t(s), z(s), v(s))\alpha(s), \alpha(s)), \\ &\textit{initial constraint:} \ \ g^0(z(0)) = b^0, \ t(0) = t_0 \ , \\ &\textit{final constraint:} \ \ g^f(z(1)) = b^f, t(1) \in R \ , \\ &\textit{control constraint:} \ \ (v(s), \alpha(s)) \in \Omega \times (0, \infty) \\ &\qquad \text{for } 0 \leq s \leq 1 \text{ and } v(\cdot), \alpha(\cdot) \text{ piecewise continuous.} \end{aligned} \tag{8.30}$$

The relation between problems (8.27) and (8.30) is established in the following result.

Lemma 1: (i) Let $x_0^* \in T^0, u^*(\cdot) \in \mathcal{U}, \ t_f^* \in (t_0, \infty)$ and let $x^*(t) = \phi(t, t_0, x_0^*, u^*(\cdot))$ be the corresponding trajectory. Suppose that $x^*(t_f^*) \in T^f$, and suppose that $(u^*(\cdot), x_0^*, t_f^*)$ is optimal for (8.27). Define $z_0^*, v^*(\cdot)$, and $\alpha^*(\cdot)$ by

$$\begin{aligned} z_0^* \quad &= x_0^* \\ v^*(s) &= u^*(t_0 + s(t_f^* - t_0)) \quad , \ 0 \leq s \leq 1 \ , \\ \alpha^*(s) &= (t_f^* - t_0) \qquad\qquad , \ 0 \leq s \leq 1 \ . \end{aligned}$$

Then $((v^*(\cdot), \alpha^*(\cdot)), z_0^*)$ is optimal for (8.30).
(ii) Let $z_0^* \in T^0$, and let $(v^*(\cdot), \alpha^*(\cdot))$ be an admissible control for (8.30) such that the corresponding trajectory $(t^*(\cdot), z^*(\cdot))$ satisfies the final conditions of (8.30). Suppose that $((v^*(\cdot), \alpha^*(\cdot)), z_0^*)$ is optimal for (8.30). Define $x_0^*, u^*(\cdot) \in \mathcal{U}$, and t_f^* by

$$\begin{aligned} x_0^* \quad &= z_0^* \ , \\ u^*(t) &= v^*(s^*(t)) \ , \ \ t_0 \leq t \leq t_f^* \ , \\ t_f^* \quad &= t^*(1) \ , \end{aligned}$$

where $s^*(\cdot)$ is functional inverse of $t^*(\cdot)$. Then $(u^*(\cdot), z_0^*, t_f^*)$ is optimal for (8.27).

Exercise 1: Prove Lemma 1.
Theorem 1: Let $u^*(\cdot) \in \mathcal{U}$, let $x_0^* \in T^0$, let $t_f^* \in (0, \infty)$, and let
$x^*(t) = \phi(t, t_0, x_0^*, u^*(\cdot)), \ t_0 \leq t \leq t_f$, and suppose that $x^*(t_f^*) \in T^f$. If $(u^*(\cdot), x_0^*, t_f^*)$ is optimal for (8.27), then there exists a function $\tilde{p}^* = (p_0^*, p^*) : \ [t_0, t_f^*] \to R^{1+m}$, not identically zero, and with $p_0^*(t) \equiv$ constant and $p_0^*(t) \geq 0$, satisfying

$$\begin{aligned} &\textit{(augmented) adjoint equation:} \\ &\qquad \dot{\tilde{p}}^* \ (t) = -[\tfrac{\partial \tilde{f}}{\partial \tilde{x}}(t, x^*(t), u^*(t))]' \tilde{p}^*(t) \ , \end{aligned} \tag{8.31}$$

$$\textit{initial condition:} \ \ p^*(t_0) \perp T^0(x_0^*) \ , \tag{8.32}$$

$$\textit{final condition:}\ \ p^*(t_f^*)\perp T^f(x^*(t_f^*))\ \ . \tag{8.33}$$

Also the *maximum principle*

$$\tilde{H}(t,x^*(t),\tilde{p}^*(t),u^*(t))=\tilde{M}(t,x^*(t),\tilde{p}^*(t))\ \ , \tag{8.34}$$

holds for all $t\in[t_0,t_f]$ except possibly for a finite set. Furthermore, t_f^* must be such that

$$\hat{H}(t_f^*,x^*(t_f^*),\tilde{p}^*(t_f^*),u^*(t_f^*))=0\ \ . \tag{8.35}$$

Finally, if f_0 and f do not explicitly depend on t, then $\hat{M}(t,x^*(t),\tilde{p}^*(t))\equiv 0$.

Proof: By Lemma 1, $z_0^*=x_0^*,\ v^*(s)=u^*(t_0+s(t_f^*-t_0))$ and $\alpha^*(s)=(t_f^*-t_0)$ for $0\le s\le 1$ constitute an optimal solution for (8.30). The resulting trajectory is
$z^*(s)=x^*(t_0+s(t_f^*-t_0)), t^*(s)=t_0+s(t_f^*-t_0),\ 0\le s\le 1$, so that in particular
$z^*(1)=x^*(t_f^*)$.
By Theorem 1 of Section 2, there exists a function $\tilde{\lambda}^*=(\lambda_0^*,\lambda^*,\lambda_{n+1}^*):\ \ [0,1]\to R^{1+n+1}$, not identically zero, and with $\lambda_0^*(s)\equiv$ constant and $\lambda_0^*(s)\ge 0$, satisfying

$$\textit{adjoint equation:}\ \begin{bmatrix}\dot{\lambda}_0^*(t)\\ \dot{\lambda}^*(t)\\ \\ \dot{\lambda}_{n+1}^*(t)\end{bmatrix}=-\begin{bmatrix}0\\ \{[\frac{\partial f_0}{\partial z}(t^*(s),z^*(s),v^*(s))]'\lambda_0^*(s)\\ +[\frac{\partial f}{\partial z}(t^*(s),z^*(s),v^*(s))]'\lambda^*(s)\}\alpha^*(s)\\ \{[\frac{\partial f_0}{\partial t}(t^*(s),z^*(s),v^*(s))]'\lambda_0^*(s)\\ +[\frac{\partial f}{\partial t}(t^*(s),z^*(s),v^*(s))]'\lambda^*(s)\}\alpha^*(s)\end{bmatrix} \tag{8.36}$$

$$\textit{initial condition:}\ \ \lambda^*(0)\perp T^0(z_0^*) \tag{8.37}$$

$$\textit{final condition:}\ \ \lambda^*(1)\perp T^f(z^*(1))\ ,\ \lambda_{n+1}^*(1)=0\ . \tag{8.38}$$

Furthermore, the maximum principle

$$\begin{aligned}&\lambda_0^*(s)f_0(t^*(s),z^*(s),v^*(s))\alpha^*(s)\\ &\quad+\lambda^*(s)'f(t^*(s),z^*(s),v^*(s))\alpha^*(s)+\lambda_{n+1}^*(s)\alpha^*(s)\\ &=\sup\{[\lambda_0^*(s)f_0(t^*(s),z^*(s),w)\beta\\ &\quad+\lambda^*(s)'f(t^*(s),z^*(s),w)\beta+\lambda_{n+1}^*(s)\beta]|w\in\Omega,\ \beta\in(0,\infty)\}\end{aligned} \tag{8.39}$$

holds for all $s\in[0,1]$ except possibly for a finite set.

Let $s^*(t)=(t-t_0)/(t_f^*-t_0),\ t_0\le t\le t_f^*$, and define $\tilde{p}^*=(p_0^*,p^*):\ \ [t_0,t_f^*]\to R^{1+n}$ by

$$p_0^*(t)=\lambda_0^*(s^*(t)),\ p^*(t)=\lambda^*(s^*(t)),\ t_0\le t\le t_f^*\ \ . \tag{8.40}$$

First of all, $\tilde{p}^*$ is not identically zero. Because if $\tilde{p}^*\equiv 0$, then from (8.40) we have $(\lambda_0^*,\lambda^*)\equiv 0$ and then from (8.36), $\lambda_{n+1}^*\equiv$ constant, but from (8.38), $\lambda_{n+1}^*(1)=0$ so that we would have $\tilde{\lambda}^*\equiv 0$

which is a contradiction. It is trivial to verify that $\tilde{p}^*(\cdot)$ satisfies (8.31), and, on the other hand (8.37) and (8.38) respectively imply (8.32) and (8.33). Next, (8.39) is equivalent to

$$\begin{aligned}&\lambda_0^*(s)f_0(t^*(s),z^*(s),v^*(s))\\&\quad+\lambda^*(s)'f(t^*(s),z^*(s),v^*(s))+\lambda_{n+1}^*(s)=0\end{aligned}\tag{8.41}$$

and

$$\begin{aligned}&\lambda_0^*(s)f_0(t^*(s),z^*(s),v^*(s))+\lambda^*(s)'f(t^*(s),z^*(s),v^*(s))\\&=\text{Sup }\{[\lambda_0^*(s)f_0(t^*(s),z^*(s),w)+\lambda^*(s)'f(t^*(s),z^*(s),w)]|w\in\Omega\}.\end{aligned}\tag{8.42}$$

Evidently (8.42) is equivalent to (8.34) and (8.35) follows from (8.41) and the fact that $\lambda_{n+1}^*(1)=0$. Finally, the last assertion of the Theorem follows from (8.35) and the fact that $\tilde{M}(t,x^*(t),\tilde{p}^*(t))\equiv$ constant if f_0, f are not explicitly dependent on t. ♢

8.3.2 *Minimum-time problems*

.

We consider the following special case of (8.27):

$$\begin{aligned}&\text{Maximize }\int_{t_0}^{t_f}(-1)dt\\&\text{subject to}\\&\textit{dynamics: }\dot{x}(t)=f(t,x(t),u(t)),\ t_0\le t\le t_f\\&\textit{initial condition: }x(t_0)=x_0\ ,\\&\textit{final condition: }x(t_f)=x_f\ ,\\&\textit{control constraint: }u(\cdot)\in\mathcal{U}\ ,\\&\textit{final-time constraint: }t_f\in(t_0,\infty)\ .\end{aligned}\tag{8.43}$$

In (8.43), x_0, x_f are fixed, so that the optimal control problem consists of finding a control which transfers the system from state x_0 at time t_0 to state x_f in minimum time. Applying Theorem 1 to this problem gives Theorem 2.

Theorem 2: Let $t_f^*\in(t_0,\infty)$ and let $u^*:\ [t_0,t_f^*]\to\Omega$ be optimal. Let $x^*(\cdot)$ be the corresponding trajectory. Then there exists a function $p^*:\ [t_0,t_f^*]\to R^n$, not identically zero, satisfying

adjoint equation: $\dot{p}^*(t)=-[\frac{\partial f}{\partial x}(t,x^*(t),u^*(t))]'p^*(t),\ t_0\le t\le t_f^*\ ,$
initial condition: $p^*(t_0)\in R^n\ ,$
final condition: $p^*(t_f^*)\in R^n\ .$
Also the *maximum principle*

$$H(t,x^*(t),p^*(t),u^*(t))=M(t,x^*(t),p^*(t))\tag{8.44}$$

holds for all $t\in[t_0,t_f^*]$ except possibly for a finite set.
Finally,

$$M(t_f^*,x^*(t_f),p^*(t_f))\ge 0\tag{8.45}$$

and if f does not depend explicitly on t then

$$M(t,x^*(t),p^*,(t))\equiv\ \text{constant}\ .\tag{8.46}$$

Exercise 2: Prove Theorem 2.
We now study a simple example illustrating Theorem 2. *Example 1:* The motion of a particle is described by

$$m\ddot{x}(t) + \sigma\dot{x}(t) = u(t) \ ,$$

where m = mass, σ = coefficient of friction, u = applied force, and x = position of the particle. For simplicity we suppose that $x \in R,\ u \in R$ and $u(t)$ constrained by $|u(t)| \leq 1$. Starting with an initial condition $x(0) = x_{01}, \dot{x}(0) = x_{02}$ we wish to find an admissible control which brings the particle to the state $x = 0, \dot{x} = 0$ in minimum time.

Solution: Taking $x_1 = x,\ x_2 = \dot{x}$ we rewrite the particle dynamics as

$$\begin{bmatrix} \dot{x}_1(t) \\ \dot{x}_2(t) \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 0 & -\alpha \end{bmatrix} \begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix} + \begin{bmatrix} 0 \\ b \end{bmatrix} u(t) \ , \tag{8.47}$$

where $\alpha = (\sigma/m) > 0$ and $b = (1/m) > 0$. The control constraint set is $\Omega = [-1, 1]$.

Suppose that $u^*(\cdot)$ is optimal and $x^*(\cdot)$ is the corresponding trajectory. By Theorem 2 there exists a non-zero solution $p^*(\cdot)$ of

$$\begin{bmatrix} \dot{p}_1^*(t) \\ \dot{p}_2^*(t) \end{bmatrix} = -\begin{bmatrix} 0 & 0 \\ 1 & -\alpha \end{bmatrix} \begin{bmatrix} p_1^*(t) \\ p_2^*(t) \end{bmatrix} \tag{8.48}$$

such that (8.44), (8.45), and (8.46) hold. Now the transition matrix function of the homogeneous part of (8.47) is

$$\Phi(t, \tau) = \begin{bmatrix} 1 & \frac{1}{\alpha}(1 - e^{-\alpha(t-\tau)}) \\ 0 & e^{-\alpha(t-\tau)} \end{bmatrix} \ ,$$

so that the solution of (8.48) is

$$\begin{bmatrix} p_1^*(t) \\ p_2^*(t) \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ \frac{1}{\alpha}(1 - e^{\alpha t}) & e^{\alpha t} \end{bmatrix} \begin{bmatrix} p_1^*(0) \\ p_2^*(0) \end{bmatrix} \ ,$$

or

$$p_1^*(t) \equiv p_1^*(0) \ ,$$

and

$$p_2^*(t) = \tfrac{1}{\alpha}p_1^*(0) + e^{\alpha t}(-\tfrac{1}{\alpha}p_1^*(0) + p_2^*(0)) \ . \tag{8.49}$$

The Hamiltonian H is given by

$$\begin{aligned} H(x^*(t), p^*(t), v) &= (p_1^*(t) - \alpha p_2^*(t))x_2^*(t) + bp_2^*(t)v \\ &= e^{\alpha t}(p_1^*(0) - \alpha p_2^*(0))x_2^*(t) + pb_2^*(t)v \ , \end{aligned}$$

so that from the maximum principle we can immediately conclude that

$$u^*(t) = \begin{cases} +1 \text{ if } p_2^*(t) > 0, \\ -1 \text{ if } p_2^*(t) < 0, \\ ? \quad \text{if } p_2^*(t) = 0 \ . \end{cases} \tag{8.50}$$

Furthermore, since the right-hand side of (8.47) does not depend on t explicitly we must also have

$$e^{\alpha t}(p_1^*(0) - \alpha p_2^*(0))x_2^*(t) + b p_2^*(t)u^*(t) \equiv \text{constant}. \tag{8.51}$$

We now proceed to analyze the consequences of (8.49) and (8.50). First of all since $p_1^*(t) \equiv p_1^*(0), p_2^*(\cdot)$ can have three qualitatively different forms.
Case 1. $-p_1^*(0) + \alpha p_2^*(0) > 0$: Evidently then, from (8.49) we see that $p_2^*(t)$ must be a strictly monotonically increasing function so that from (8.50) $u^*(\cdot)$ can behave in one of two ways:

either

$$u^*(t) = \begin{cases} -1 \text{ for } t < \hat{t} \text{ and } p_2^*(t) < 0 \text{ for } t < \hat{t}, \\ +1 \text{ for } t > \hat{t} \text{ and } p_2^*(t) > 0 \text{ for } t > \hat{t}, \end{cases}$$

or

$$u^*(t) \equiv +1 \quad \text{and} \quad p_2^*(t) > 0 \quad \text{for all } t.$$

Case 2. $-p_1^*(0) + \alpha p_2^*(0) < 0$: Evidently $u^*(\cdot)$ can behave in one of two ways:

either

$$u^*(t) = \begin{cases} +1 \text{ for } t < \hat{t} \text{ and } p_2^*(t) > 0 \text{ for } t < \hat{t}, \\ -1 \text{ for } t > \hat{t} \text{ and } p_2^*(t) < 0 \text{ for } t > \hat{t}, \end{cases}$$

or

$$u^*(t) \equiv -1 \quad \text{and} \quad p^*(t) < 0 \text{ for all } t.$$

Case 3. $-p_1^*(0) + \alpha p_2^*(0) = 0$: In this case $p_2^*(t) \equiv (1/\alpha)p_1^*(0)$. Also since $p^*(t) \not\equiv 0$, we must have in this case $p_1^*(0) \neq 0$. Hence $u^*(\cdot)$ we can behave in one of two ways:

either

$$u^*(t) \equiv +1 \quad \text{and} \quad p_2^*(t) \equiv \tfrac{1}{\alpha} p_1^*(0) > 0 \ ,$$

or

$$u^*(t) \equiv -1 \quad \text{and} \quad p_2^*(t) \equiv \tfrac{1}{\alpha} p_1^*(0) < 0 \ ,$$

Thus, the optimal control u^* is always equal to +1 or -1 and it can switch at most once between these two values. The optimal control is given by

$$\begin{aligned} u^*(t) &= \text{sgn } \ p_2^*(t) \\ &= \text{sgn } [\tfrac{1}{\alpha}p_1^*(0) + e^{\alpha t}(-\tfrac{1}{\alpha}p_1^*(0) + p_2^*(0))] \ . \end{aligned}$$

Thus the search for the optimal control reduces to finding $p_1^*(0), p_2^*(0)$ such that the solution of the differential equation

$$\begin{aligned} \dot{x} &= x_2 \\ \dot{x}_2 &= -\alpha x_2 + b \text{ sgn}[\tfrac{1}{\alpha}p_1^*(0) + e^{\alpha t}(-\tfrac{1}{\alpha}p_1^*(0) + p_2^*(0))] \ , \end{aligned} \tag{8.52}$$

with initial condition

$$x_1(0) = x_{10}, x_{20} = x_{20} \tag{8.53}$$

also satisfies the final condition

$$x_1(t_f^*) = 0, \ \ x_2(t_f^*) = 0 \ , \tag{8.54}$$

for some $t_f^* > 0$; and then t_f^* is the minimum time.

There are at least two ways of solving the two-point boundary value problem (8.52), (8.53), and (8.54). One way is to guess at the value of $p^*(0)$ and then integrate (8.52) and (8.53) *forward* in time and check if (8.54) is satisfied. If (8.54) is not satisfied then modify $p^*(0)$ and repeat. An alternative is to guess at the value of $p^*(0)$ and then integrate (8.52) and (8.54) *backward* in time and check of (8.53) is satisfied. The latter approach is more advantageous because we know that any trajectory obtained by this procedure is optimal for initial conditions which lie on the trajectory. Let us follow this procedure.

Suppose we choose $p^*(0)$ such that $-p_1^*(0) = \alpha p_2^*(0) = 0$ and $p_2^*(0) > 0$. Then we must have $u^*(t) \equiv 1$. Integrating (8.52) and (8.54) backward in time give us a trajectory $\xi(t)$ where

$$\begin{aligned} \dot{\xi}_1(t) &= -\dot{\xi}_2(t) \\ \dot{\xi}_2(t) &= \alpha\xi_2(t) - b \ , \end{aligned}$$

with

$$\xi_1(0) - \xi_2(0) = 0 \ .$$

This gives

$$\xi_1(t) = \tfrac{b}{\alpha}(-t + \tfrac{e^{\alpha t}-1}{\alpha}) \ , \ \xi_2(t) = \tfrac{b}{\alpha}(1 - e^{\alpha t}) \ ,$$

which is the curve OA in Figure 8.3.

On the other hand, if $p^*(0)$ is such that $-p_1^*(0) + \alpha p_2^*(0) = 0$ and $p_2^*(0) < 0$, then $u^*(t) \equiv -1$ and we get

$$\xi_1(t) = -\tfrac{b}{\alpha}(-t + \tfrac{e^{\alpha t}-1}{\alpha}) \ , \ \xi_2(t) = -\tfrac{b}{\alpha}(1 - e^{\alpha t}) \ ,$$

which is the curve OB.

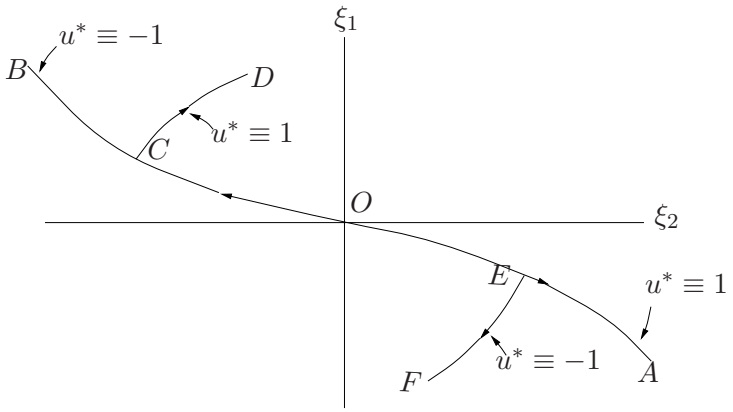


Figure 8.3: Backward integration of (8.52) and (8.54).

Next suppose $p^*(0)$ is such that $-p_1^*(0) + \alpha p_2^*(0) > 0$, and $p_2^*(0) < 0$. Then $[(1/\alpha)p_1^*(0) + e^{\alpha t}(-(1/\alpha)p_1^*(0) + p_2^*(0))]$ will have a negative value for $t \in (0, \hat{t})$ and a positive value for $t \in (\hat{t}, \infty)$. Hence, if we integrate (8.52), (8.54) backwards in time we get trajectory $\xi(t)$ where

$$\begin{array}{l} \dot{\xi}(t) = -\xi_2(t) \\ \dot{\xi}_2(t) = \alpha\xi_2(t) + \end{array} \begin{cases} -b & \text{for} \quad t < \hat{t} \\ b & \text{for} \quad t > \hat{t} \end{cases},$$

with $\xi_1(0) = 0, \xi_2(0) = 0$. This give us the curve OCD. Finally if $p^*(0)$ is such that $-p_1^*(0) + \alpha p_2^*(0) < 0$, and $p_2^*(0) < 0$, then $u^*(t) = 1$ for $t < \hat{t}$ and $u^*(t) = -1$ for $t > \hat{t}$, and we get the curve OEF.

We see then that the optimal control $u^*(\cdot)$ has the following characterizing properties:

$$u^*(t) = \begin{cases} 1 & \text{if} \quad x^*(t) \quad \text{is above } \textit{BOA} \text{ or on } \textit{OA} \\ -1 & \text{if} \quad x^*(t) \quad \text{is below } \textit{BOA} \text{ or on } \textit{OB} \end{cases}.$$

Hence we can synthesize the optimal control in feedback from: $u^*(t) = \psi(x^*(t))$ where the

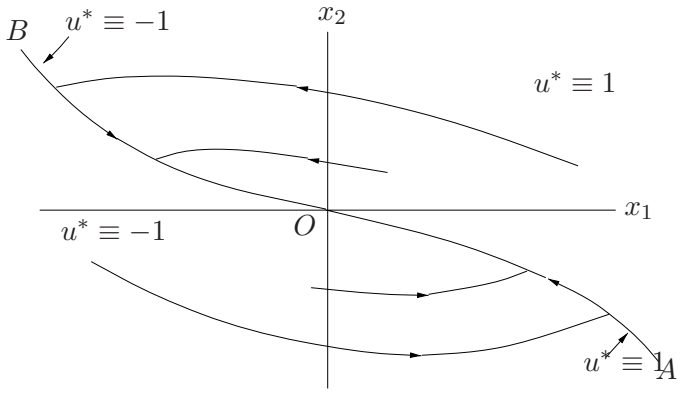


Figure 8.4: Optimal trajectories of Example 1.

function $\psi: \ R^2 \to \{1, -1\}$ is given by (see Figure 8.4)

$$\psi(x_1, x_2) = \begin{cases} 1 & \text{if } \ (x_1, x_2) \text{ is above } \textit{BOA} \text{ or on } \textit{OA} \\ -1 & \text{if } \ (x_1, x_2) \text{ is below } \textit{BOA} \text{ or on } \textit{OB} \ . \end{cases}$$

8.4 *Linear System, Quadratic Cost*

An important class of problems which arise in practice is the case when the dynamics are linear and the objective function is quadratic. Specifically, consider the optimal control problem (8.55):

$$\begin{array}{l} \text{Minimize } \ \displaystyle\int_0^T \frac{1}{2}[x'(t)P(t)x(t) + u'(t)Q(t)u(t)]dt \\ \text{subject to} \\ \textit{dynamics: } \ \dot{x}(t) = A(t)x(t) + B(t)u(t), \ \ 0 \le t \le T \ , \\ \textit{initial condition: } \ x(0) = x_0 \ , \\ \textit{final condition: } \ G^f x(t) = b^f \ , \\ \textit{control constraint: } \ u(t) \in R^p, \ u(\cdot) \text{ piecewise continuous.} \end{array} \tag{8.55}$$

In (8.56) we assume that $P(t)$ is an $n \times n$ symmetric, positive semi-definite matrix whereas $Q(t)$ is a $p \times p$ symmetric, positive definite matrix. G^f is a given $\ell_f \times n$ matrix, and $x_0 \in R^n, \ b^f \in R^{\ell_f}$ are given vectors. T is a fixed final time.

We apply Theorem 1 of Section 2, so that we must search for a number $p_0^* \ge 0$ and a function $p^*: \ [0, T] \to R^n$, not both zero, such that

$$\dot{p}^*(t) = -p_0^*(-P(t)x^*(t)) - A'(t)p^*(t) \ \ , \tag{8.56}$$

and

$$p^*(t) \perp T^f(x^*(t)) = \{\xi | G^f \xi = 0\} \ . \tag{8.57}$$

The Hamiltonian function is

$$\begin{array}{rcl} H(t, x^*(t), \tilde{p}^*(t), v) & = & -\frac{1}{2}p_0^*[x^*(t)'P(t)x^*(t) + v'Q(t)v] \\ & & +p^*(t)'[A(t)x^*(t) + B(t)v] \end{array}$$

so that the optimal control $u^*(t)$ must maximize

$$-\tfrac{1}{2}p_0^* v'Q(t)v + p^*(t)'B(t)v \ \text{ for } \ v \in R^p \ \ . \tag{8.58}$$

If $p_0^* > 0$, this will imply

$$u^*(t) = \tfrac{1}{p_0^*}Q^{-1}(t)B'(t)p^*(t) \ \ , \tag{8.59}$$

whereas if $p_0^* = 0$, then we must have

$$p^*(t)'B(t) \equiv 0 \tag{8.60}$$

because otherwise (8.58) cannot have a maximum.

We make the following assumption about the system dynamics.
Assumption: The control system $\dot{x}(t) = A(t)x(t) + B(t)u(t)$ is controllable over the interval $[0, T]$. (See (Desoer [1970]) for a definition of controllability and for the properties we use below.) Let $\Phi(t, \tau)$ be the transition matrix function of the homogeneous linear differential equation $\dot{x}(t) = A(t)x(t)$. Then the controllability assumption is equivalent to the statement that for any $\xi \in R^n$

$$\xi' \Phi(t, \tau) B(\tau) = 0 \ , \ 0 \leq \tau \leq T \ , \text{ implies } \xi = 0 \ . \tag{8.61}$$

Next we claim that if the system is controllable then $p_0^* \neq 0$, because if $p_0^* = 0$ then from (8.56) we can see that

$$p^*(t) = (\Phi(T, t))' p^*(T)$$

and hence from (8.60)

$$(p^*(t))' \Phi(T, t) B(t) = 0 \ , \ 0 \leq t \leq T \ ,$$

but then from (8.61) we get $p^*(T) = 0$. Hence if $p_0^* = 0$, then we must have $\tilde{p}^*(t) \equiv 0$ which is a contradiction. Thus, under the controllability assumption, $p_0^* > 0$, and hence the optimal control is given by (8.59). Now if $p_0^* > 0$ it is trivial that $\hat{p}^*(t) = (1, (p^*(t)/p_0^*))$ will satisfy all the necessary conditions so that we can assume that $p_0^* = 1$. The optimal trajectory and the optimal control is obtained by solving the following two-point boundary value problem:

$$\begin{aligned} &\dot{x}^*(t) = A(t)x^*(t) + B(t)Q^{-1}(t)B'(t)p^*(t) \\ &\dot{p}(t) = P(t)x^*(t) - A'(t)p^*(t) \\ &x^*(0) = x_0, G^f x^*(T) = b^f \ , \ p^*(T) \perp T^f(x^*(T)) \ . \end{aligned}$$

For further details regarding the solution of this boundary value problem and for related topics see (See and Markus [1967]).

8.5 *The Singular Case*

In applying the necessary conditions derived in this chapter it sometimes happens that $H(t, x^*(t), p^*(t), v)$ is independent of v for values of t lying in a non-zero interval. In such cases the maximum principle does not help in selecting the optimal value of the control. We are faced with the so-called singular case (because we are in trouble–not because the situation is rare). We illustrate this by analyzing Example 4 of Chapter 1.

The problem can be summarized as follows:

$$\begin{aligned} &\text{Maximize } \int_0^T c(t)dt = \int_0^T (1 - s(t)) f(k(t)) dt \\ &\text{subject to} \\ &\textit{dynamics:} \ \dot{k}(t) = s(t) f(k(t)) - \mu k(t) \ , \ 0 \leq t \leq T \\ &\textit{initial constraint:} \ k(0) = k_0 \ , \\ &\textit{final constraint:} \ k(t) \in R \ , \\ &\textit{control constraint:} \ s(t) \in [0, 1], \ s(\cdot) \text{ piecewise continuous.} \end{aligned}$$

We make the following assumptions regarding the production function f:

$$f_k(k) > 0, \quad f_{kk}(K) < 0 \quad \text{for all} \quad k \ , \tag{8.62}$$

$$\lim_{k \to 0} f_k(k) = \infty \ . \tag{8.63}$$

Assumption (8.62) says that the marginal product of capital is positive and this marginal product decreases with increasing capital. Assumption (8.63) is mainly for technical convenience and can be dispensed with without difficulty.

Now suppose that $s^* : \ [0,T] \to [0,1]$ is an optimal savings policy and let $k^*(t), \ 0 \le t \le T$, be the corresponding trajectory of the capital-to-labor ratio. Then by Theorem 1 of Section 2, there exist a number $p_0^* \ge 0$, and a function $p^* : \ [0,T] \to R$, not both identically zero, such that

$$\dot{p}^*(t) = -p_0^*(1 - s^*(t))f_k(k^*(t)) - p^*(t)[s^*(t)f_k(k^*(t)) - \mu] \tag{8.64}$$

with the final condition

$$p^*(T) = 0 \ , \tag{8.65}$$

and the maximum principle holds. First of all, if $p_0^* = 0$ then from (8.64) and (8.65) we must also have $p^*(t) \equiv 0$. Hence we must have $p_0^* > 0$ and then by replacing (p_0^*, p^*) by $(1/p_0^*)(p_0^*, p^*)$ we can assume without losing generality that $p_0^* = 1$, so that (8.64) simplifies to

$$\dot{p}^*(t) = -1(1 - s^*(t))f_k(k^*(t)) - p^*(t)[s^*(t)f_k(k^*(t)) - \mu] \ . \tag{8.66}$$

The maximum principle says that

$$H(t, k^*(t), p^*(t), s) = (1 - s)f(k^*(t)) + p^*(t)[sf(k^*(t)) - \mu k^*(t)]$$

is maximized over $s \in [0,1]$ at $s^*(t)$, which immediately implies that

$$s^*(t) = \begin{cases} 1 & \text{if} \quad p^*(t) > 1 \\ 0 & \text{if} \quad p^*(t) < 1 \\ ? & \text{if} \quad p^*(t) = 1 \end{cases} \tag{8.67}$$

We analyze separately the three cases above.

Case 1. $p^*(t) > 1, s^*(t) = 1$: Then the dynamic equations become

$$\begin{aligned} \dot{k}^*(t) &= f(k^*(t)) - \mu k^*(t) \ , \\ \dot{p}^*(t) &= -p^*(t)[f_k(k^*(t)) - \mu] \ . \end{aligned} \tag{8.68}$$

The behavior of the solutions of (8.68) is depicted in the $(k,p)-, (k,t)-$ and $(p,t)-$planes in Figure 8.5. Here k_G, k_H are the solutions of $f_k(k_G) - \mu = 0$ and $f(k_M) - \mu k = 0$. Such solutions exist and are unique by virtue of the assumptions (8.62) and (8.63). Futhermore, we note from (8.62) that $k_G < k_M$, and $f_k(k) - \mu \overset{<}{>} 0$ according as $k \overset{<}{>} k_G$ whereas $f(k) - \mu k \overset{>}{<} 0$ according as $k \overset{<}{>} k_M$. (See Figure 8.6.)

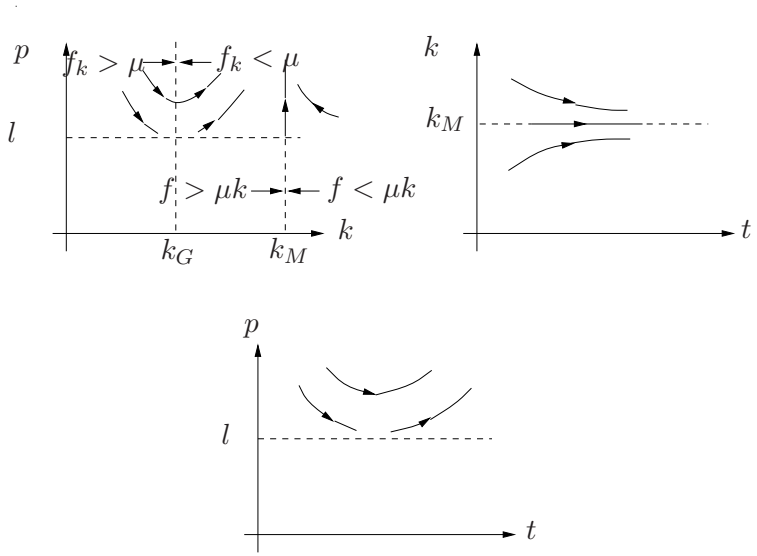


Figure 8.5: Illustration for Case 1.

Case 2. $p^*(t) < 1, s^*(t) = 0$: Then the dynamic equations are

$$\begin{aligned}\dot{k}^*(t) &= -\mu k^*(t) \ , \\ \dot{p}^*(t) &= -f_k(k^*(t)) + \mu p^*(t) \ ,\end{aligned}$$

giving rise to the behavior illustrated in Figure 8.7.
Case 3. $p^*(t) = 1, s^*(t) =$?: (Possibly singular case.) Evidently if $p^*(t) = 1$ only for a finite set of times t then we do not have to worry about this case. We face the singular case only if $p^*(t) = 1$ for $t \in I$, where I is a non-zero interval. But then we have $\dot{p}^*(t) = 0$ for $t \in I$ so that from (8.66) we get

$$-(1 - s^*(t))f_k(k^*(t)) - [s^*(t)f_k(k^*(t)) - \mu] = 0 \text{ for } t \in I \ ,$$

so

$$-f_k(k^*(t)) + \mu = 0 \quad \text{for} \quad t \in I \ ,$$

or

$$k^*(t) = k_G \ \text{ for } \ t \in I \ . \tag{8.69}$$

In turn then we must have $\dot{k}^*(t) = 0$ for $t \in I$ so that

$$s^*(t)f(k_G) - \mu K_G = 0 \text{ for } t \in I \ ,$$

and hence,

$$s^*(t) = \mu \tfrac{k_G}{f(k_G)} \quad \text{for} \quad t \in I \ . \tag{8.70}$$

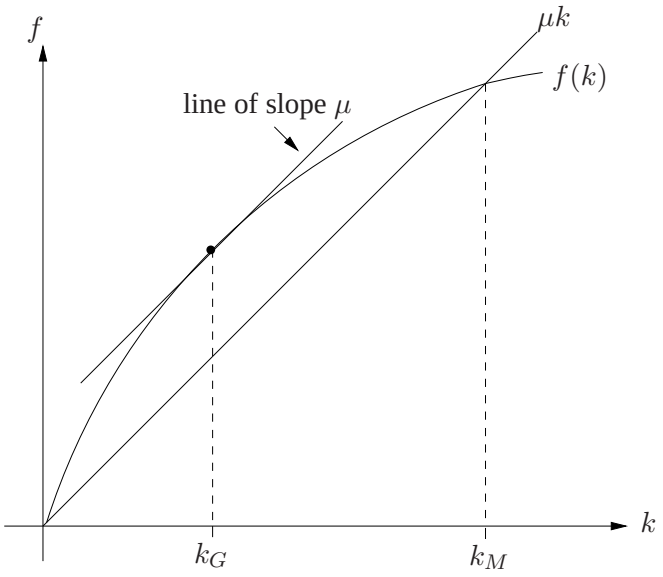


Figure 8.6: Illustration for assumptions (8.62), (8.63).

Thus in the singular case the optimal solution is characterized by (8.69) and (8.70), as in Figure 8.8.

We can now assemble separate cases to obtain the optimal control. First of all, from the final condition (8.65) we know that for t close to T, $p^*(t) < 1$ so that we are in Case 2. We face two possibilities: Either (A)

$$p^*(t) < 1 \quad \text{for all} \quad t < [0, T]$$

and then $s^*(t) = 0, k^*(t) = k_0 e^{-\mu t}$, for $0 \leq t \leq T$, or (B)

$$\text{there exists } t_2 \in (0, T) \text{ such that } p^*(t_2) = 1 \text{ and } p^*(t) < 1 \text{ for } t_2 < t \leq T \ .$$

We then have three possibilities depending on the value of $k^*(t_2)$:
(Bi) $k^*(t_2) < k_G$: then $\dot{p}^*(t_2) < 0$ so that $p^*(t) > 1$ for $t < t_2$ and we are in Case 1 so that $s^*(t) = 1$ for $t < t_2$. In particular we must have $k_0 < k_G$.
(Bii) $k^*(t_2) > k_G$: then $\dot{p}^*(_2) > 0$ but then $p^*(t_2 + \varepsilon) > 1$ for $\varepsilon > 0$ sufficiently small and since $p^*(T) = 0$ there must exist $t_3 \in (t_2, T)$ such that $p^*(t_3) = 1$. This contradicts the definition of t_2 so that this possibility cannot arise.
(Biii) $k^*(t_2) - k_G$: then we can have a singular arc in some interval (t_1, t_2) so that $p^*(t) = 1, k^*(t) = k_G$, and $s^*(t) = \mu(k_G/f(k_G))$ for $t \in (t_1, t_2)$. For $t < t_1$ we either have $p^*(t) > 1, s^*(t) > 1$ if $k_0 < k_G$, or we have $p^*(t) < 1,\ s^*(t) = 0$ if $k > k_G$.

The various possibilities are illustrated in Figure 8.9.

The capital-to-labor ratio k_G is called the *golden mean* and the singular solution is called the *golden path.* The reason for this term is contained in the following exercise.

Exercise 1: A capital-to-labor ratio $\hat{k}$ is said to be *sustainable* if there exists $\hat{s} \in [0, 1]$ such that $\hat{s} f(\hat{k}) - \mu \hat{k} = 0$. Show that k_G is the unique sustainable capital-to-labor ratio which maximizes sustainable consumption $(1 - s) f(k)$.

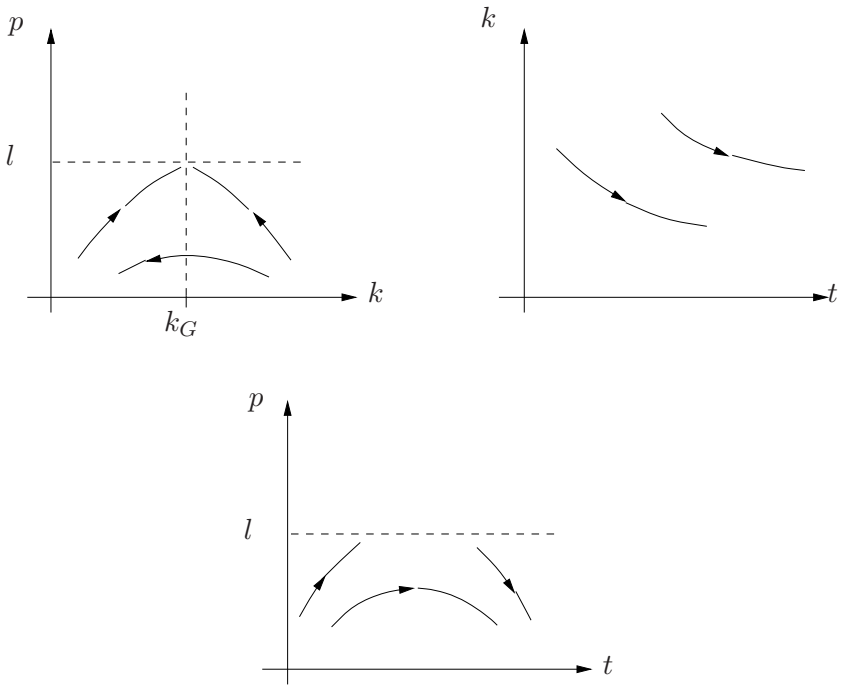


Figure 8.7: Illustration for Case 2.

8.6 *Bibliographical Remarks*

The results presented in this chapter appeared in English in full detail for the first time in 1962 in the book by Pontryagin, *et al.*, cited earlier. That book contains many extensions and many examples and it is still an important source. However, the derivation of the maximum principle given in the book by Lee and Markus is more satisfactory. Several important generalizations of the maximum principle have appeared. On the one hand these include extensions to infinite-dimensional state spaces and on the other hand they allow for constraints on the state more general than merely initial and final constraints. For a unified, but mathematically difficult, treatment see (Neustadt [1969]). For a less rigorous treatment of state-space constraints see (Jacobson, *et al*, [1971]), whereas for a discussion of the singular case consult (Kelley, *et al.* [1968]).

For an applications-oriented treatment of this subject the reader is referred to (Athans and Falb [1966]) and (Bryson and Ho [1969]). For applications of the maximum principle to optimal economic growth see (Shell [1967]). There is no single source of computational methods for optimal control problems. Among the many useful techniques which have been proposed see (Lasdon, *et al.*, [1967]), (Kelley [1962]), (McReynolds [1966]), and (Balakrishnan and Neustadt [1964]); also consult (Jacobson and Mayne [1970]), and (Polak [1971]).

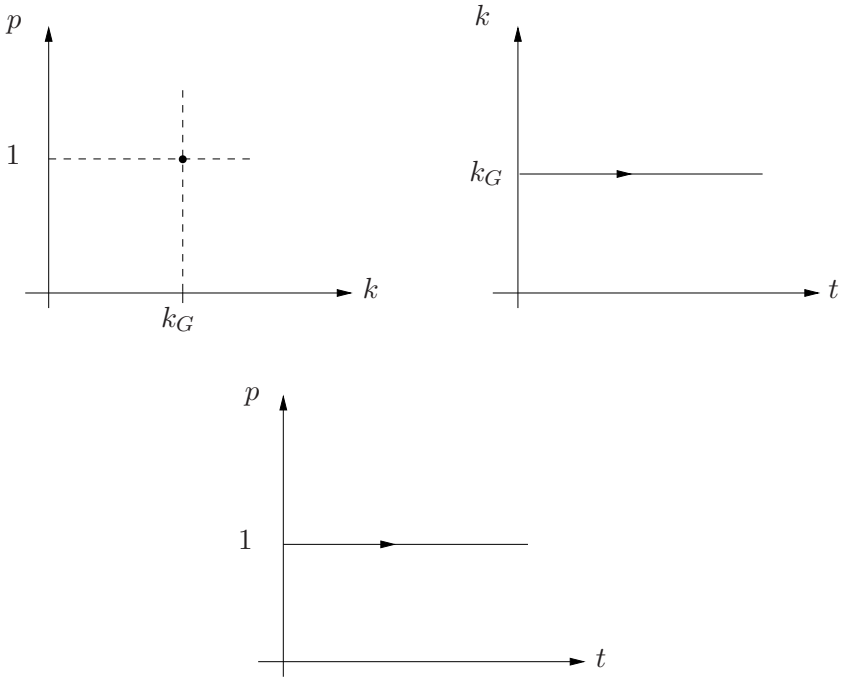


Figure 8.8: Case 3. The singular case.

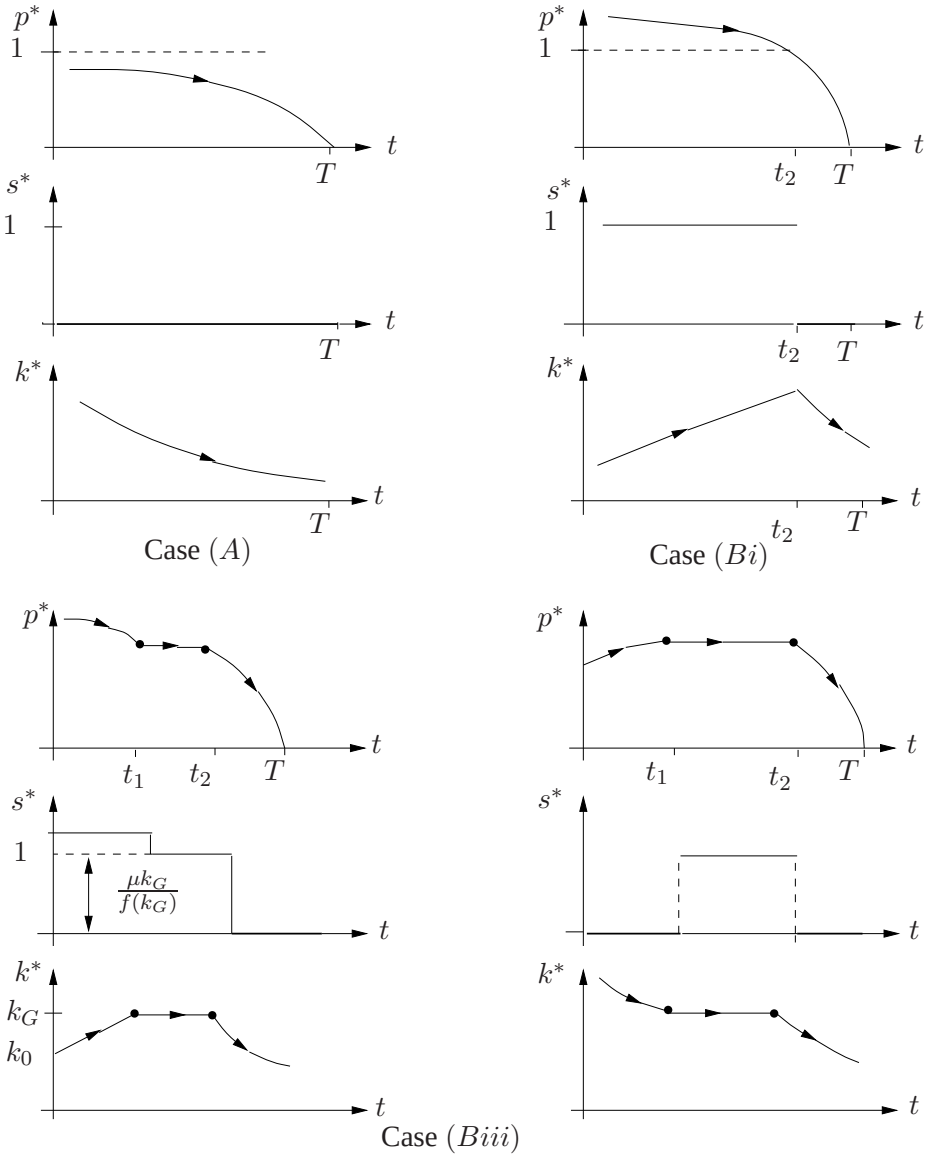


Figure 8.9: The optimal solution of example.

Chapter 9

Dynamic programing

SEQUENTIAL DECISION PROBLEMS: DYNAMIC PROGRAMMING FORMULATION

The sequential decision problems discussed in the last three Chapters were analyzed by variational methods, *i.e.*, the necessary conditions for optimality were obtained by comparing the optimal decision with decisions in a small neighborhood of the optimum. Dynamic programming (DP is a technique which compares the optimal decision with *all* the other decisions. This global comparison, therefore, leads to optimality conditions which are *sufficient*. The main advantage of DP, besides the fact that it give sufficiency conditions, is that DP permits very general problem formulations which do not require differentiability or convexity conditions or even the restriction to a finite-dimensional state space. The only disadvantage (which unfortunately often rules out its use) of DP is that it can easily give rise to enormous computational requirements.

In the first section we develop the main recursion equation of DP for discrete-time problems. The second section deals with the continuous-time problem. Some general remarks and bibliographical references are collected in the final section.

9.1 *Discrete-time DP*

We consider a problem formulation similar to that of Chapter VI. However, for notational convenience we neglect final conditions and state-space constraints.

$$\begin{aligned} &\text{Maximize} \quad \sum_{i=0}^{N-1} f_0(i, x(i), u(i)) + \Phi(x(N)) \\ &\text{subject to} \\ &\textit{dynamics:} \;\; x(i+1) = f(i, x(i), u(i)) \;,\; i = 0, 1, \ldots, N-1 \;, \\ &\textit{initial condition:} \;\; x(0) = x_0 \;\;, \\ &\textit{control constraint:} \;\; u(i) \in \Omega_i \;,\; i = 0, 1, \ldots, N-1 \;\;. \end{aligned} \tag{9.1}$$

In (9.1), the state $x(i)$ and the control $u(i)$ belong to arbitrary sets X and U respectively. X and U may be finite sets, or finite-dimensional vector spaces (as in the previous chapters), or even infinite-dimensional spaces. $x_0 \in X$ is fixed. The Ω_i are fixed subsets of U. Finally $f_0(i, \cdot, \cdot) : \;\; X \times U \to R, \Phi : \;\; X \to R, \;\; f(i, \cdot, \cdot) : \;\; X \times U \to X$ are fixed functions.

The main idea underlying DP involves embedding the optimal control problem (9.1), in which the system starts in state x_0 at time 0, into a family of optimal control problems with the same dynamics, objective function, and control constraint as in (9.1) but with different initial states and initial times. More precisely, for each $x \in X$ and k between) and $N-1$, consider the following problem:

$$\begin{array}{l} \text{Maximize} \quad \displaystyle\sum_{i=k}^{N-1} f_0(i, x(i), u(i)) + \Phi(x(N)) \ , \\ \text{subject to} \\ \textit{dynamics:} \ \ x(i+1) = f(i, x(i), u(i)), \ i = k, k+1, \ldots, N-1, \\ \text{initial condition:} \ \ x(k) = x, \\ \text{control constraint:} \ \ u(i) \in \Omega_i, \ i = k, k+1, \cdot, N-1 \ . \end{array} \tag{9.2}$$

Since the initial time k and initial state x are the only parameters in the problem above, we will sometimes use the index $(9.2)_{k,x}$ to distinguish between different problems. We begin with an elementary but crucial observation.

Lemma 1: Suppose $u^*(k), \ldots, u^*(N-1)$ is an optimal control for $(9.2)_{k,x}$, and let $x^*(k) = x, \ x^*(k+1), \ldots, x^*(N)$ be the corresponding optimal trajectory. Then for any $\ell, k \le \ell \le N-1, u^*(\ell), \ldots, u^*(N-1)$ is an optimal control for $(9.2)_{\ell, x^*(\ell)}$.

Proof: Suppose not. Then there exists a control $\hat{u}(\ell), \hat{u}(\ell+1), \ldots, \hat{u}(N-1)$, with corresponding trajectory $\hat{x}(\ell) = x^*(\ell), \hat{x}(\ell+1), \ldots, \hat{x}(N)$, such that

$$\begin{array}{l} \displaystyle\sum_{i=\ell}^{N-1} f_0(i, \hat{x}(i), \hat{u}(i)) + \Phi(\hat{x}(N)) \\ > \displaystyle\sum_{i=\ell}^{N-1} f_0(i, x^*(i), u^*(i)) + \Phi(x^*(N)) \ . \end{array} \tag{9.3}$$

But then consider the control $\tilde{u}(k), \ldots, \tilde{u}(N-1)$ with

$$\tilde{u}(i) \begin{cases} u^*(i) \ , & i = k, \ldots, \ell - 1 \\ \hat{u}(i) \ , & i = \ell, \ldots, N-1 \ , \end{cases}$$

and the corresponding trajectory, starting in state x at time k, is $\tilde{x}(k), \ldots, \tilde{x}(N)$ where

$$\tilde{x}(i) = \begin{cases} x^*(i) \ , & i = k, \ldots, \ell \\ \hat{x}(i) \ , & i = \ell + 1, \ldots, N \ . \end{cases}$$

The value of the objective function corresponding to this control for the problem $(9.2)_{k,x}$ is

$$\begin{array}{l} \displaystyle\sum_{i=k}^{N-1} f_0(i, \tilde{x}(i), \tilde{u}(i)) + \Phi(\tilde{x}(n)) \\ = \displaystyle\sum_{i=k}^{\ell-1} f_0(i, x^*(i), u^*(i)) + \sum_{i=\ell}^{N-1} f_0(i, \hat{x}(i), \hat{u}(i)) + \Phi(\hat{x}(N)) \\ > \displaystyle\sum_{i=k}^{N-1} f_0(i, x^*(i), u^*(i)) + \Phi(x^*(N)) \ , \end{array}$$

by (9.3), so that $u^*(k), \ldots, u^*(N-1)$ cannot be optimal for $9.2)_k, x$, contradicting the hypothesis. (end theorem)

From now on we assume that an optimal solution to $(9.2)_{k,x}$ exists for all $0 \leq k \leq N-1$, and all $x \in X$. Let $V(k,x)$ be the maximum value of $(9.2)_{k,x}$. We call V the *(maximum) value function*.
Theorem 1: Define $V(N,\cdot)$ by $(V(N,x) = \Phi(x).\ \ V(k,x)$ satisfies the backward recursion equation

$$V(k,x) = \text{ Max}\{f_0,(k,x,u) + V(k_1, f(k,x,u,))|u \in \Omega_k\},\ 0 \leq k \leq N-1\ \ . \tag{9.4}$$

Proof: Let $x \in X$, let $u^*(k), \ldots, u^*(N-1)$ be an optimal control for $(9.2)_{k,x}$, and let $x^*(k) = x, \ldots, x^*(N)$ be the corresponding trajectory be $x(k) = x, \ldots, x(N)$. We have

$$\begin{aligned} &\sum_{i=k}^{N-1} f_0(i, x^*(i), u^*(i)) + \Phi(x^*(N)) \\ &\geq \sum_{i=k}^{N-1} f_0(i, x(i), u(i)) + \Phi(x(N))\ \ . \end{aligned} \tag{9.5}$$

By Lemma 1 the left-hand side of (9.5) is equal to

$$f_0(k, x, u^*(k)) + V(k+1, f(k, x^*, u^*(k)))\ \ .$$

On the other hand, by the definition of V we have

$$\begin{aligned} &\sum_{i=k}^{N-1} f_0(i, x(i), u(i)) + \Phi(x(N)) = f_0(k, x, u(k)) \\ &+\{\sum_{i=k+1}^{N} f_0(i, x(i), u(i)) + \Phi(x(N)) \leq f_0(k, x, u, (k)) + V(k+1, f(k, x, u(k)))\}\ \ , \end{aligned}$$

with equality if and only if $u(k+1), \ldots, u(N-1)$ is optimal for $(9.2)_{k+1,x(k+1)}$. Combining these two facts we get

$$\begin{aligned} &f_0(k, xu^*(k)) + V(k+1, f(k, x, u^*(k))) \\ &\geq f_0(k, x, u(k)) + V(k+1, f(x, k, u(k)))\ \ , \end{aligned}$$

for all $u(k) \in \Omega_k$, which is equivalent to (9.4).(end theorem)
Corollary 1: Let $u(k), \ldots, u(N-1)$ be any control for the problem $(9.2)_{k,x}$ and let $x(k) = x, \ldots, x(N)$ be the corresponding trajectory. Then

$$V(\ell, x(\ell)) \leq f_0(\ell, x(\ell), u(\ell)) + V(\ell+1, f(\ell, x(\ell), u(\ell)),\ k \leq \ell \leq N-1\ \ ,$$

and equality holds for all $k \leq \ell \leq N-1$ if and only if the control is optimal for $(9.2)_{k,x}$.
Corollary 2: For $k = 0, 1, \ldots, N-1$, let $\psi(k,\cdot):\ \ X \to \Omega_k$ be such that

$$\begin{aligned} &f_0(k, x, \psi(k,x)) + V(k+1, f(k, x, \psi(k,x)) \\ &= \text{Max}\{f_0(k, x, u) + V(k+1, f(k, x, u))|u \in \Omega_k\}\ \ . \end{aligned}$$

Then $\psi(k,\cdot),\ \ k = 0, \ldots, N-1$ is an *optimal feedback control, i.e.,* for any k, x the control $u^*(k), \ldots, u^*(N-1)$ defined by $u^*(\ell) = \psi(\ell, x^*(\ell)),\ k \leq \ell \leq N-1$, where

$$x^*(\ell+1) = f(\ell, x^*(\ell), \psi(\ell, x^*(\ell)),\ k \le \ell \le N-1\ ,\ x^*(k) = x\ \ ,$$

is optimal for $(\alpha)_{k,x}$.

Remark: Theorem 1 and Corollary 2 are the main results of DP. The recursion equation (9.4) allows us to compute the value function, and in evaluating the maximum in (9.4) we also obtain the optimum feedback control. Note that this feedback control is optimum for *all* initial conditions. However, unless we can find a "closed-form" analytic solution to (9.4), the DP formulation may necessitate a prohibitive amount of computation since we would have to compute and store the values of V and ψ for all k and x. For instance, suppose $n = 10$ and the state-space X is a finite set with 20 elements. Then we have to compute and store 10×20 values of V, which is a reasonable amount. But now suppose $X = R^n$ and we approximate each dimension of x by 20 values. Then for $N = 10$, we have to compute and store $10x(20)^n$ values of V. For $n = 3$ this number is 80,000, and for $n = 5$ it is 32,000,000, which is quite impractical for existing computers. This "curse of dimensionality" seriously limits the applicability of DP to problems where we cannot solve (9.4) analytically.

- *Exercise 1:* An instructor is preparing to lead his class for a long hike. He assumes that each person can take up to W pounds in his knapsack. There are N possible items to choose from. Each unit of item i weighs w_i pounds. The instructor assigns a number $U_i > 0$ for each unit of item i. These numbers represent the relative utility of that item during the hike. How many units of each item should be placed in each knapsack so as to maximize total utility? Formulate this problem by DP.

9.2 *Continuous-time DP*

We consider a continuous-time version of (9.2):

$$\begin{array}{l} \text{Maximize}\ \ \int_0^{t_f} f_0(t, x(t), u(t))dt + \Phi(x(t_f)) \\ \text{subject to} \\ \textit{dynamics:}\ \ \dot{x}(t) = f(t, x(t), u(t))\ \ ,\ \ t_0 \le t \le t_f \\ \textit{initial condition:}\ \ x(0) = x_0\ \ , \\ \textit{control constraint:}\ \ u:\ \ [t_0, t_f] \to \Omega\ \text{ and }\ u(\cdot)\ \text{ piecewise continuous.} \end{array} \tag{9.6}$$

In (9.6), $x \in R^n,\ u \in R^p,\ \Omega \subset R^p$. $\Phi:\ \ R^n \to R$ is assumed differentiable and f_0, f are assumed to satisfy the conditions stated in VIII.1.1.

As before, for $t_0 \le t \le t_f$ and $x \in R^n$, let $V(t,x)$ be the maximum value of the objective function over the interval $[t, t_f]$ starting in state x at time t. Then it is easy to see that V must satisfy

$$\begin{array}{r} V(t,x) = \text{Max}\{\int_t^{t+\Delta} f_0(\tau, x(\tau), u(\tau))d\tau \\ +V(t+\Delta, x(t+\Delta))|u:\ \ [t, t+\Delta] \to \Omega\}, \Delta \ge 0\ \ , \end{array} \tag{9.7}$$

and

$$V(t_f, x) = \Phi(x)\ \ . \tag{9.8}$$

In (9.7), $x(\tau)$ is the solution of

$$\begin{aligned} &\dot{x}(\tau) = f(\tau, x(\tau), u(\tau)) \ , \ \ t \le \tau \le t + \Delta \ , \\ &x(t) = x \ . \end{aligned}$$

Let us suppose that V is differentiable in t and x. Then from (9.7) we get

$$\begin{aligned} V(t,x) = \ \ \text{Max}\{ & f_0(t,x,u)\Delta + V(t,x) + \tfrac{\partial V}{\partial x} f(t,x,u)\Delta \\ & + \tfrac{\partial V}{\partial t}(t,x)\Delta + o(\Delta) | u \in \Omega\}, \ \Delta > 0 \ . \end{aligned}$$

Dividing by $\Delta > 0$ and letting Δ approach zero we get the *Hamilton-Jacobi- Bellman* partial differentiable equation for the value function:

$$\tfrac{\partial V}{\partial t}(t,x) + \ \text{Max}\{ f_0(t,x,u) + \tfrac{\partial V}{\partial x}(t,x) f(t,x,u) | u \in \Omega\} = 0. \tag{9.9}$$

Theorem 1: Suppose there exists a differentiable function $V : \ \ [t_0, t_f] \times R^n \to R$ which satisfies (9.9) and the boundary condition (9.8). Suppose there exists a function $\psi : \ \ [t_0, t_f] \times R^n \to \Omega$ with ψ piecewise continuous in t and Lipschitz in x, satisfying

$$\begin{aligned} & f_0(t,x,\psi(t,x)) + \tfrac{\partial V}{\partial x} f(t,x,\psi(t,x)) \\ & \quad = \text{Max}\{ f_0(t,x,u) + \tfrac{\partial V}{\partial x} f(t,x,u) | u \in \Omega\} \ . \end{aligned} \tag{9.10}$$

Then ψ is an optimal feedback control for the problem (9.6), and V is the value function.
Proof: Let $t \in [t_0, t_f]$ and $x \in R^n$. Let $\hat{u} : \ \ [t, t_f] \to \Omega$ be any piecewise continuous control and let $\hat{x}(\tau)$ be the solution of

$$\begin{aligned} &\dot{\hat{x}}\ (\tau) = f(\tau, \hat{x}(\tau), \hat{u}(\tau)) \ , \ t \le \tau \le t_f \ , \\ &\hat{x}(t) = x \ . \end{aligned} \tag{9.11}$$

Let $x^*(\tau)$ be the solution of

$$\begin{aligned} &\dot{x}^*(\tau) = f(\tau, x^*(\tau), \psi(\tau, x^*(\tau))) \ , \ t \le \tau \le t_f \ , \\ &x^*(\tau) = x \ . \end{aligned} \tag{9.12}$$

Note that the hypothesis concerning ψ guarantees a solution of (9.12). Let $u^*(\tau) = \psi(\tau, x^*, (\tau)), \ t \le \tau \le t_f$. To show that ψ is an optimal feedback control we must show that

$$\begin{aligned} & \int_t^{t_f} f_0(t\tau, x^*(\tau), u^*(\tau)) d\tau + \Phi(x^*(\tau)) \\ \le & \int_t^{t_f} f_0(\tau, x^*(\tau), \hat{u}(\tau)) d\tau + \Phi(\hat{x}(t_f)) \ . \end{aligned} \tag{9.13}$$

To this end we note that

$$\begin{aligned} V(t_f, x^*(t_f)) - V(t, x^*(t)) &= \int_f^{t_f} \frac{dV}{d\tau}(\tau, x^*(\tau)) d\tau \\ &= \in_t^{t_f} \{ \frac{\partial V}{\partial \tau}(\tau, x^*(\tau) + \frac{\partial V}{\partial x} \dot{x}^*(\tau) \} d\tau \\ &= - \int_t^{t_f} F - 0(\tau, x^*(\tau), u^*(\tau)) d\tau \ \ , \end{aligned} \tag{9.14}$$

using (9.9), (9.10). On the other hand,

$$\begin{aligned} V(t_f, \hat{x}(t_f)) - V(t, \hat{x}, (t)) &= \int_t^{t_f} \{\frac{\partial V}{\partial \tau}(\tau, \hat{x}(\tau)) + \frac{\partial V}{\partial x}\ \dot{\tilde{x}}\ (\tau)\} d\tau \\ &\leq -\int_t^{t_f} f_0(\tau, \hat{x}(\tau), \hat{u}^*(\tau)) d\tau \ , \end{aligned} \tag{9.15}$$

using (9.9). From (9.14), (9.15), (9.8) and the fact that $x^*(t) = \hat{x}(t) = x$ we conclude that

$$\begin{aligned} V(t,x) &= \Phi(x^*(t_f)) + \int_t^{t_f} f_0(\tau, x^*(\tau), u^*(\tau)) \\ &\geq \Phi(\hat{x}(t_f)) + \int_t^{t_f} f_0(\tau, \hat{x}(\tau), \hat{u}(\tau)) d\tau \end{aligned}$$

so that (9.13) is proved. It also follows that V is the maximum value function. $\diamondsuit$

- *Exercise 1:* Obtain the value function and the optimal feedback control for the linear regulatory problem:

$$\begin{aligned} &\text{Minimize } \tfrac{1}{2}x'(T)P(T)x(t) + \tfrac{1}{2}\int T_0\{x'(t)P(t)x(t) \\ &\qquad\qquad +u'(t)Q(t)u(t)\}dt \\ &\text{subject to} \\ &\textit{dynamics: } \dot{x}(t) = A(t)x(t) + B(t)u(t)\ ,\ 0 \leq t \leq T\ , \\ &\textit{initial condition: } x(0) = x_0\ , \\ &\textit{control constraint: } u(t) \in R^p\ , \end{aligned}$$

  where $P(t) = P'(t)$ is positive semi-definite, and $Q(t) = Q'(t)$ is positive definite. [Hint: Obtain the partial differential equation satisfied by $V(t,x)$ and try a solution of the form $V(t,x) = x'R(t)x$ where R is unknown.]

9.3 *Miscellaneous Remarks*

There is vast literature dealing with the theory and applications of DP. The most elegant applications of DP are to various problems in operations research where one can obtain "closed-form" analytic solutions to be recursion equation for the value function. See (Bellman and Dreyfus [1952]) and (Wagner [1969]). In the case of sequential decision-making under uncertainties DP is about the only available general method. For an excellent introduction to this area of application see (Howard [1960]). For an important application of DP to computational considerations for optimal control problems see (Jacobson and Mayne [1970]). Larson [1968] has developed computational techniques which greatly increase the range of applicability of DP where closed-form solutions are not available. Finally, the book of Bellman [1957] is still excellent reading. []

Bibliography

[1] J.J. Arrow and L. Hurwicz. *Decentralization and Computation in Resource in resource Allocation*. Essays in Economics and Econometrics. University of North Carolina Press, 1960. in Pfouts R.W. (ed.).

[2] M. Athans and P.L. Falb. *Optimal Control*. McGraw-Hill, 1966.

[3] A.V. Balakrishnan and L.W. Neustadt. *Computing Methods in Optimization Problems*. Academic Press, 1964.

[4] K. Banerjee. *Generalized Lagrange Multipliers in Dynamic Programming*. PhD thesis, College of Engineering, University of California, Berkeley, 1971.

[5] R.E. Bellman. *Dynamic Programming*. Princeton University Press, 1957.

[6] R.E. Bellman and S.E. Dreyfus. *Applied Dynamic Programming*. Princeton University Press, 1962.

[7] D. Blackwell and M.A. Girshick. *Theory of Games and Statistical Decisions*. John Wiley, 1954.

[8] J.E. Bruns. The function of operations research specilists in large urban schools. *IEEE Trans. on Systems Science and Cybernetics*, SSC-6(4), 1970.

[9] A.E. Bryson and Y.C. Ho. *Applied Optimal Control*. Blaisdell, 1969.

[10] J.D. Cannon, C.D. Cullum, and E. Polak. *Theory of Optimal Control and Mathematical Programming*. McGraw-Hill, 1970.

[11] G. Dantzig. *Linear Programming and Extensions*. Princeton University Press, 1963.

[12] C.A. Desoer. *Notes for a Second Course on Linear Systems*. Van Nostrand Reinhold, 1970.

[13] S.W. Director and R.A. Rohrer. On the design of resistance n-port networks by digital computer. *IEEE Trans. on Circuit Theory*, CT-16(3), 1969a.

[14] S.W. Director and R.A. Rohrer. The generalized adjoint network and network sensitivities. *IEEE Trans. on Circuit Theory*, CT-16(3), 1969b.

[15] S.W. Director and R.A. Rohrer. Automated network design–the frequency-domain case. *IEEE Trans. on Circuit Theory*, CT-16(3), 1969c.

[16] R. Dorfman and H.D. Jacoby. *A Model of Public Decisions Illustrated by a Water Pollution Policy Problem*. Public Expenditures and Policy Analysis. Markham Publishing Co, 1970. in Haveman, R.H. and Margolis, J. (eds.).

[17] R. Dorfman, P.A. Samuelson, and R.M.Solow. *Linear Programming and Economic Analysis*. McGraw-Hill, 1958.

[18] R.I. Dowell and R.A. Rohrer. Automated design of biasing circuits. *IEEE Trans. on Circuit Theory*, CT-18(1), 1971.

[19] Dowles Foundation Monograph. *The Economic Theory of Teams*. John Wiley, 1971. to appear.

[20] W.H. Fleming. *Functions of Several Variables*. Addison-Wesley, 1965.

[21] C.R. Frank. *Production Theory and Indivisible Commodities*. Princeton University Press, 1969.

[22] D. Gale. A geometric duality theorem with economic applications. *Review of Economic Studies*, XXXIV(1), 1967.

[23] A.M. Geoffrion. *Duality in Nonlinear Programming: a Simplified Application-Oriented Treatment*. The Rand Corporation, 1970a. Memo RM-6134-PR.

[24] A.M. Geoffrion. Primal resource directive approaches for optimizing nonlinear decomposable programs. *Operations Research*, 18, 1970b.

[25] F.J. Gould. Extensions of lagrange multipliers in nonlinear programming. *SIAM J. Apl. Math*, 17, 1969.

[26] J.J. Greenberg and W.P. Pierskalla. Surrogate mathematical programming. *Operations Research*, 18, 1970.

[27] H.J.Kushner. *Introduction to Stochastic Control*. Holt, Rinehart, and Winston, 1971.

[28] R.A. Howard. *Dynamic Programming and Markov Processes*. MIT Press, 1960.

[29] R. Isaacs. *Differential Games*. John Wiley, 1965.

[30] D.H. Jacobson, M.M. Lele, and J.L. Speyes. New necessary condtions of optimality for problems with state-variable inequality constraints. *J. Math. Analysis and Applications*, 1971. to appear.

[31] D.H. Jacobson and D.Q. Mayne. *Differential Dynamic Programming*. American Elsevier Publishing Co., 1970.

[32] S. Karlin. *Mathematical Methods and Theory in Games, Programming, and Economics*, volume 1. Addison-Wesley, 1959.

[33] H.J. Kelley. *Method of Gradients*. Optimization Techniques. Academic Press, 1962. in Leitmann, G.(ed.).

[34] J.H. Kelley, R.E. Kopp, and H.G. Mayer. *Singular Extremals*. Topics in Optimization. Academic Press, 1970. in Leitman, G. (ed.).

[35] D.A. Kendrick, H.S. Rao, and C.H. Wells. Water quality regulation with multiple polluters. In *Proc. 1971 Jt. Autom. Control Conf.*, Washington U., St. Louis, August 11-13 1971.

[36] T.C. Koopmans. Objectives, constraints, and outcomes in optimal growth models. *Econometrica*, 35(1), 1967.

[37] H.W. Kuhn and A.W. Tucker. Nonlinear programming. In *Proc. Second Berkeley Symp. on Math. Statistics and Probability*. University of California Press, Berkeley, 1951.

[38] R.E. Larson. *State Increment Dynamic Programming*. American Elsevier Publishing Co., 1968.

[39] L.S. Lasdon, S.K. Mitter, and A.D. Waren. The conjugate gradient method for optimal control problems. *IEEE Trans. on Automatic Control*, AC-12(1), 1967.

[40] E.B. Lee and L. Markus. *Foundation of Optimal Control Theory*. John Wiley, 1967.

[41] R. Luce and H. Raiffa. *Games and Decisions*. John Wiley, 1957.

[42] D.G. Luenberger. Quasi-convex programming. *Siam J. Applied Math*, 16, 1968.

[43] O.L. Mangasarian. *Nonlinear Programming*. McGraw-Hill, 1969.

[44] S.R. McReynolds. The successive sweep method and dynamic programming. *J. Math. Analysis and Applications*, 19, 1967.

[45] J.S. Meditch. *Stochastic Optimal Linear Estimation and Control*. McGraw-Hill, 1969.

[46] M.D. Mesarovic, D. Macho, and Y. Takahara. *Theory of Hierarchical, Multi-level Systems*. Academic Press, 1970.

[47] C.E. Miller. *The Simplex Method for Local Separable Programming*. Recent Advance Programming. McGraw-Hill, 1963. in Graves, R.L. and Wolfe, P. (eds.).

[48] L.W. Neustadt. The existence of optimal controls in the absence of convexity conditions. *J. Math. Analysis and Applications*, 7, 1963.

[49] L.W. Neustadt. A general theory of extremals. *J. Computer and System Sciences*, 3(1), 1969.

[50] H. Nikaido. *Convex Structures and Economic Theory*. Academic Press, 1968.

[51] G. Owen. *Game Theory*. W.B. Saunders & Co., 1968.

[52] E. Polak. *Computational Methods in Optimization: A Unified Approach*. Academic Press, 1971.

[53] L.S. Pontryagin, R.V. Boltyanski, R.V. Gamkrelidze, and E.F.Mischenko. *The Mathematical Theory of Optimal Processes*. Interscience, 1962.

[54] R.T. Rockafeller. *Convex Analysis*. Princeton University Press, 1970.

[55] M. Sakarovitch. *Notes on Linear Programming*. Van Nostrand Reinhold, 1971.

[56] L.J. Savage. *The Foundation of Statistics*. John Wiley, 1954.

[57] K. Shell. *Essays in the Theory of Optimal Economic Growth*. MIT Press, 1967.

[58] R.M. Solow. The economist's approach to pollution and its control. *Science*, 173(3996), 1971.

[59] D.M. Topkis and A. Veinnott Jr. On the convergence of some feasible directions algorithms for nonlinear programming. *SIAM J. on Control*, 5(2), 1967.

[60] H.M. Wagner. *Principles of Operations Research*. Prentice-Hall, 1969.

[61] P. Wolfe. The simplex method for quadratic programming. *Econometrica*, 27, 1959.

[62] W.M. Wonham. On the seperation theorem of optimal control. *SIAM J. on Control*, 6(2), 1968.

[63] E. Zangwill. *Nonlinear Programming: A Unified Approach*. Prentice-Hall, 1969.

Index