Optimization Techniques
with Applications to Aerospace Systems

edited by

George Leitmann

Optimization Techniques

With Applications to Aerospace Systems

Edited by

GEORGE LEITMANN

Applied Mechanics Group, Mechanical Engineering Department
University of California, Berkeley, California

New York ACADEMIC PRESS London

ACADEMIC PRESS INC.
111 Fifth Avenue
New York 3, N. Y.

United Kingdom Edition
Published by
ACADEMIC PRESS INC. (London) Ltd.
Berkeley Square House, London, W.1

Library of Congress Catalog Card Number 62-13108

PRINTED IN THE UNITED STATES OF AMERICA

Contributors

RICHARD BELLMAN, *The RAND Corporation, Santa Monica, California*

JOHN BREAKWELL, *Lockheed Missiles and Space Company, Sunnyvale, California*

R. W. BUSSARD,* *Los Alamos Scientific Laboratory of the University of California, Los Alamos, New Mexico*

THEODORE N. EDELBAUM, *Research Laboratories, United Aircraft Corporation, East Hartford, Connecticut*

FRANK D. FAULKNER, *U. S. Naval Postgraduate School, Monterey, California*

ROBERT KALABA, *The RAND Corporation, Santa Monica, California*

C. M. KASHMAR, *General Electric Company, Santa Barbara, California*

HENRY J. KELLEY, *Grumman Aircraft Engineering Corporation, Bethpage, New York*

RICHARD E. KOPP, *Grumman Aircraft Engineering Corporation, Bethpage, New York*

DEREK F. LAWDEN, *University of Canterbury, Christchurch, New Zealand*

G. LEITMANN, *University of California, Berkeley, California*

ANGELO MIELE, *Boeing Scientific Research Laboratories, Seattle, Washington*

E. L. PETERSON, *General Electric Company, Santa Barbara, California*

---

* *Present address*: Space Technology Laboratories, Redondo Beach, California

Foreword

Whatever man does he strives to do in the "best" possible manner. In attempting to reach a desired goal in an "optimum" fashion, he is faced immediately with two problems. The first of these is the decision of choosing the goal itself—the "payoff." For what is one man's optimum may well be another man's pessimum. Here he may decide to ignore the desires of the other interested parties and to choose a payoff solely on the basis of his own interests, whether or not these are in conflict with the interests of others. Thus, the scientist who wishes to carry out experiments using satellite-borne apparatus may desire an optimum ascent trajectory which results in maximum payload in orbit, even if such an "optimum" trajectory involves excessively large accelerations which are not tolerable from the point of view of the structural engineer whose job is the design of the carrier vehicle. On the other hand, he may temper his choice of payoff by considering the interests of others, that is, by imposing restrictions on the optimal policy so as not to violate the requirements of other interested parties. Consequently, the scientist may be forced to accept a somewhat smaller payload in orbit by placing bounds on the accelerations, resulting in loads which are tolerable and hence acceptable to the structural engineer. But even if a payoff and constraints agreeable to all parties involved can be decided upon, there still remains the choice of technique to be used for arriving at the optimum. It is primarily to this latter question that this book addresses itself.

During the past decade there has been a remarkable growth of interest in problems of systems optimization and of optimal control. And with this interest has come an increasing need for methods useful for rendering systems optimum. Rising to meet this challenge there have sprung up various "schools," often championing one method and regarding it superior to all others. Long experience has shown that life is not so simple, that the picture is not all white and black. In short, one may expect that a particular method is superior to others for the solution of some problems—rarely for all problems. Furthermore, since the basic mathematical formulation of optimization problems is often essentially the same in many approaches, it is not unreasonable to expect that there may be a great deal of similarity among various methods, a similarity—often, indeed, an identity—which is obscured by dissimilarities in language and notation. To help the uncommitted in his search for and choice of the optimum optimization technique is the fundamental aim of this volume.

To accomplish this aim there are assembled in one book ten chapters dealing with the various methods currently espoused for the solution of problems in systems optimization and optimal control. The choice of authors has been dictated solely by a consideration of an author's interest and expertness in a particular method. With the advantages of such an eclectic approach and the ensuing multiple authorship there comes some loss of smoothness of over-all presentation, for which the Editor must take the sole blame. On the one hand, correlation between the various chapters has been achieved by cross-referencing; on the other hand, each chapter can be read as a separate entity setting forth the technique championed by a particular "school."

While each of the ten chapters dealing with methods includes simple examples, primarily for didactic purposes, it has been thought useful to present four additional chapters dealing with applications alone. Of these, the first three, Chapters 11–13, cover specific optimization problems, and the final chapter contains a discussion of problems in the optimization of a complete system, in this case a nuclear propulsion system.

A word concerning coverage is in order. Whenever a method or peculiarities in its applications are not treated in standard works of reference, these points are covered in detail; such is the case especially in Chapters 1–7. When a technique is fully exposed in readily available sources or when applications to aerospace systems are as yet sparse, the method is presented in outline only, together with appropriate remarks concerning its application to the systems under discussion; this is largely so in Chapters 8–10. Niceties in notation for their own sake have been avoided in order to make the subject matter accessible to the widest possible audience which may include engineers, scientists, and applied mathematicians whose training in mathematics need not have progressed past the first graduate year of a standard engineering curriculum.

GEORGE LEITMANN

Berkeley, California

Contents

4 *The Calculus of Variations in Applied Aerodynamics and Flight Mechanics*

ANGELO MIELE

5 *Variational Problems with Bounded Control Variables*

G. LEITMANN

6 *Methods of Gradients*

HENRY J. KELLEY

7 *Pontryagin Maximum Principle*

RICHARD E. KOPP

8 *On the Determination of Optimal Trajectories Via Dynamic Programming*

RICHARD BELLMAN

9 *Computational Considerations for Some Deterministic and Adaptive Control Processes*

ROBERT KALABA

10 *General Imbedding Theory*

C. M. KASHMAR AND E. L. PETERSON

14 *Some Topics in Nuclear Rocket Optimization*

R. W. BUSSARD

—1—

Theory of Maxima and Minima

THEODORE N. EDELBAUM

Research Laboratories, United Aircraft Corporation,
East Hartford, Connecticut

1.1 Necessary Conditions for Maxima or Minima

1.11 Introduction—Variational Terminology

The theory of ordinary maxima and minima is concerned with the problem of finding the values of each of n independent variables $x_1, x_2, \cdots, x_n$ at which some function of the n variables $f(x_1, x_2, \cdots, x_n)$ reaches either a maximum or a minimum (an extremum). This problem may be interpreted geometrically as the problem of finding a point in an n-dimensional space at which the desired function has an extremum. This geometrical interpretation can be helpful in understanding this type of problem, particularly when there are only two independent variables. A representation of such a problem is shown as a contour map in Fig. 1.

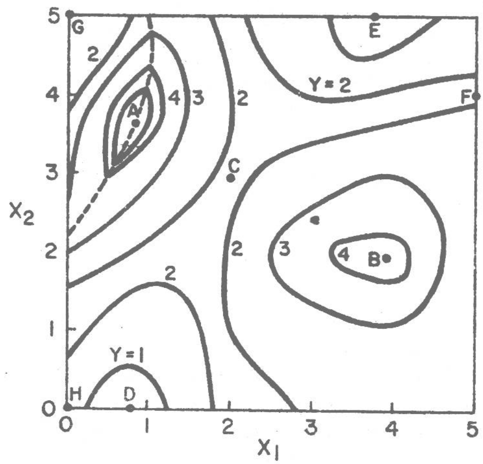

FIG. 1. Extrema and stationary points.

The independent variables are x_1 and x_2 while the dependent variable $y = f(x_1, x_2)$ is represented by the contour lines. The maximum of the function is located at point A at the top of a sharp ridge where the derivative of y with respect to both x_1 and x_2 is discontinuous. A second but lower maximum is located at point B which is "higher" than all points in its immediate vicinity. The highest of all the points in a suitably defined region, such as point A for the region shown, is called an absolute maximum while a point, such as B, that is higher than all the points in a suitably defined small neighborhood is called a local maximum.

The derivative of y with respect to x_1 and x_2 at point B is equal to zero. A point at which a function has all its partial derivatives with respect to the independent variables equal to zero is called a stationary point. The fact that a stationary point need not represent a local extremum is illustrated by point C. Point C represents the highest point in a "pass" between the "mountains" on either side. There are both lower points (along the pass) and higher points (towards the mountains) in the immediate neighborhood of C. A stationary point of this type is called a saddle point.

The minimum of the function y does not occur in the interior of the region illustrated but occurs on the boundary of this region defined by $x_1 = 0, 5$ and $x_2 = 0, 5$. Along the boundary $x_2 = 0$, a local minimum occurs at point D, while it occurs at point E along $x_2 = 5$ and at point F along $x_1 = 5$. The boundary $x_2 = 0$ has two local minima, one at each end of the interval in which the function is defined. It should be noted here that the "suitable neighborhood" for the definition of a local minimum does not include points outside of the region of definition of the function. The absolute minimum must be found by comparing the values of the local minima D, E, F, G, and H.

The basic problem of the theory of ordinary maxima and minima is to determine the location of local extrema and then to compare these so as to determine which is the absolute extremum. The example of Fig. 1 illustrates that a place to look for local extrema is along discontinuities in the first derivative and also along boundaries (another type of discontinuity). Where the function and its derivatives are continuous the local extema will always occur at stationary points although, as point C illustrates, stationary points are not always local extrema.

1.12 Necessary Conditions for Maxima or Minima

The existence of a solution to an ordinary minimum problem is guaranteed by the theorem of Weierstrass as long as the function is continuous. This theorem states[1]:

> Every function which is continuous in a closed domain possesses a largest and a smallest value either in the interior or on the boundary of the domain.

There is no corresponding general existence theorem for the solutions of problems in the calculus of variations, a circumstance which sometimes leads to difficulties. It should be noted that this theorem does not require the derivatives to be continuous so that the theorem applies to problems such as the example of Fig. 1.

The location of extrema in the interior of the region may be determined from the following theorem[2]:

> A continuous function $f(x_1, x_2, \cdots, x_n)$ of n independent variables $x_1, x_2, \cdots, x_n$ attains a maximum or a minimum in the interior of a

region R only at those values of the variables x_i for which the n partial derivatives $f_{x_1}, f_{x_2}, \cdots, f_{x_n}$ either vanish simultaneously (a stationary point), or at which one or more of these derivatives cease to exist (are discontinuous).

The location of the stationary points may be found by simultaneous solution of the n algebraic equations obtained by setting the n partial derivatives equal to zero. The question as to whether these stationary points constitute extrema will be considered in the next section.

The theorem of Weierstrass indicates that the extrema may occur on the boundary of the region. If the problem is an n-dimensional one, the search for an extremum on the boundary will generally lead to one or more ordinary minimum problems in $n-1$, $n-2$, $\cdots$, 1 dimensions. A good illustrative example of this is given by Cicala.[3] He considers a 3-dimensional problem where the function is defined in a cubic region. The complete solution of this problem requires the determination of the extrema on each of the 6 sides and on each of the 12 edges and comparison of the values of these extrema with the value of the function at the 8 vertices and at the extrema in the interior of the cube. A spherical region, on the other hand, requires only the determination of the extrema on the unbounded spherical surface and in the interior. The determination of the extrema on boundaries which are not coordinate surfaces may be considered as a problem with a subsidiary condition, to be considered in Section 1.3.

The existence of discontinuous first derivatives along lines, surfaces, etc., also requires the solution of extremal problems in 1, 2, etc., dimensions. The methods of Section 1.3 are not as satisfactory here because most of them require the existence of all partial derivatives. The substitution method of Section 1.31 is applicable in many cases.

1.2 Sufficient Conditions for Maxima or Minima

1.21 Introduction

The theorem of the preceding section states that an extremum in the interior of a region must occur at either a stationary point or at a point where one or more first partial derivatives are discontinuous. However, neither stationary points nor discontinuities have to be extrema. When the location of stationary points or points of discontinuity has been determined, the question as to whether or not they constitute extrema still has to be answered.

There are several methods of answering this question. Probably the most widely used method is the obvious one of direct comparison of the values of the function at stationary points, discontinuities, and at the boundaries. In spite of its simplicity, this method is the only rigorous method of determining absolute optima and is the only common method of treating discontinuities. The methods of the last section provide a means of determining all of the possible locations for both local extrema and the absolute extremum. A comparison of all these points is all that is needed for a rigorous determination of the absolute extremum even though it will not determine which of the other points is a local extremum.

In practice, the physical interpretation of the mathematical model will often result in an "obvious" determination of the extremal character of some stationary point. For many practical problems this intuitive result will be sufficient, although few of us are blessed with infallible intuition. The determination of the value of the function at a few neighboring points will often help to reinforce intuition. However, it should be realized that the investigation of many neighboring points cannot provide a rigorous proof of even a local extremum. Failure to realize this has led to serious errors in the past in calculus of variations problems and can lead to serious errors in ordinary minimum problems involving many independent variables.

If it is impractical to determine the values of the function at all possible stationary points, discontinuities, and boundaries, all that can be determined is whether or not a given point constitutes a local extremum. In many problems the determination of local extrema may be all that is desired. Sections 1.21 and 1.22 will consider the necessary conditions for a stationary point, having continuous second derivatives, to constitute a local extremum. Necessary conditions for points where the first derivatives are discontinuous are treated by Hancock.[4]

Excellent examples of the detailed proof that a solution is an absolute extremum may be found in Horner[5] and Munick *et al.*[6]

1.22 Two Independent Variables

The behavior of a function $f(x_1, x_2)$ in the vicinity of a point (a, b) may be determined by means of a Taylor series expansion:

$$f(x_1, x_2) = f(a, b) + f_{x_1}(a, b)(x_1 - a) + f_{x_2}(a, b)(x_2 - b)$$
$$+ \frac{1}{2!}[f_{x_1x_1}(a,b)(x_1 - a)^2 + 2f_{x_1x_2}(a,b)(x_1 - a)(x_2 - b) + f_{x_2x_2}(a,b)(x_2 - b)^2]$$
$$+ \cdots \qquad (1.1)$$

If the point a, b is a stationary point of $f(x_1, x_2)$ the two first order terms will be zero. It is necessary to examine the three second order terms in order to determine whether a, b is a maximum, a minimum, a saddle point, etc.

The sum of the three second order terms will always be positive, so that a, b will be a local minimum if

$$\begin{aligned} f_{x_1x_1} &> 0 \\ f_{x_1x_1}f_{x_2x_2} - (f_{x_1x_2})^2 &> 0 \end{aligned} \tag{1.2}$$

the function will be a local maximum if

$$\begin{aligned} f_{x_1x_1} &< 0 \\ f_{x_1x_1}f_{x_2x_2} - (f_{x_1x_2})^2 &> 0 \end{aligned} \tag{1.3}$$

In exceptional cases extrema may occur when the inequality in the second of Eqs. (1.2) and (1.3) becomes an equality. These cases are discussed by Hancock,[4] pages 20–69.

1.23 *n* Independent Variables

The corresponding sufficiency conditions for n variables may be expressed concisely by using a notation similar to that of Leitmann.[7] The necessary condition for a stationary point to be a local minimum is that

$$D_i > 0, \qquad i = 1, 2, \cdots, n \tag{1.4}$$

The necessary condition for a stationary point to be a local maximum is that

$$\begin{aligned} D_i &> 0, \qquad i = 2, 4, 6, \cdots \\ D_i &< 0, \qquad i = 1, 3, 5, \cdots \end{aligned} \tag{1.5}$$

where

$$D_i \equiv \begin{vmatrix} f_{x_1x_1} & f_{x_1x_2} & \cdots & f_{x_1x_i} \\ f_{x_2x_1} & f_{x_2x_2} & \cdots & f_{x_2x_i} \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ f_{x_ix_1} & f_{x_ix_2} & \cdots & f_{x_ix_i} \end{vmatrix}$$

A special case of this equation is the well-known sufficiency condition for a maximum or minimum of a function of one independent variable:

$$f_{x_1x_1} \lesseqgtr 0$$