LINEAR APPROXIMATION

Arthur Sard

MATHEMATICAL SURVEYS · *Number 9*

LINEAR APPROXIMATION

BY
ARTHUR SARD

1963
AMERICAN MATHEMATICAL SOCIETY
190 Hope Street, Providence, Rhode Island

The writing of this book was supported in part by the Aeronautical Research Laboratories and the Office of Scientific Research, both of the United States Air Force.

Library of Congress Catalog Number: 63-11988

Preface

The form in which I have written "Linear approximation" seems to me to be suited to the subject, and to the use of mathematicians, scientists, and engineers.

Readers interested in the applications may wish to start with the illustrative examples and the statements of theorems. All readers are urged to skip boldly and to sample where they will.

I am grateful to my wife and to many men and women for help and teaching. In the words of David, Psalm 16, 6,

חבלים נפלו לי בנעימים אף נחלת שפרה עלי·

The lines are fallen unto me in pleasant places;
yea, I have a goodly heritage.

ARTHUR SARD

July 21, 1962

QUEENS COLLEGE
THE CITY UNIVERSITY OF NEW YORK
FLUSHING, NEW YORK

Table of Contents

Introduction

Many approximations of integrals, derivatives, functions, or sums are linear, and continuous relative to a suitable norm. The linearity and continuity provide a means to establish known properties of the approximations and to discover other properties.

Design, analysis, and appraisal all are illuminated and advanced by the theory in this book.

§ 1. **Functionals.** If a quantity e is a linear function on a three-dimensional linear space X, that fact alone permits us to gain a complete knowledge of e. We may introduce coordinates

$$x = (x^1, x^2, x^3), \qquad x \in X,$$

in X. Then the value of e at x is

$$e(x) = \alpha^1 x^1 + \alpha^2 x^2 + \alpha^3 x^3, \qquad x \in X, \tag{2}$$

where α^1, α^2, α^3 are appropriate numbers which characterize e and which can be calculated by putting x equal to

$$(1, 0, 0), \quad (0, 1, 0), \quad (0, 0, 1) \tag{3}$$

successively in (2).

For an infinite dimensional space the generalization of the above remark is interesting.

Suppose that X is a normed† linear space. By a functional F on X we mean a map of X into the numbers N, where N will be either the reals R or the complex numbers C. The set X^* of linear continuous functionals on X is known as the adjoint space. If

$$F \in X^*,$$

that is, if F is a linear continuous functional on X, then this fact alone tells us much about F.

If X is finite dimensional, the dimension of X characterizes X^*. This case is like the three-dimensional one alluded to earlier. If X is infinite dimensional, the dimension of X does not of itself characterize X^*.

For the most part we deal with spaces X of dimension $\aleph_0$ (aleph$_0$). For some such spaces, the adjoint X^* is well known and our knowledge includes an explicit direct procedure for obtaining a standard formula for an arbitrary element of X^*. This is important in the theory of approximation because

† Precise definitions of technical terms are in later chapters; exact references thereto are in the index.

remainders are often elements of X^* for suitable X. If the standard formula leads to sharp appraisals, it is particularly useful.

Consider $C_0(I)$, for example, the space of continuous functions on $I = \{\alpha \leqq s \leqq \tilde{\alpha}\}$ to R, with norm

(4) $$\|x\|_{C_0(I)} = \sup_{s \in I} |x(s)|, \qquad x \in C_0(I).$$

An arbitrary element F of $C_0(I)^*$, that is, an arbitrary linear continuous functional F on $C_0(I)$, may be written

(5) $$Fx = \int_I x(s)\, df(s), \qquad x \in X,$$

where f is a suitable function of bounded variation on I. Furthermore f may be calculated from the formula

(6) $$f(t) = \begin{cases} \lim_{\nu \to \infty} F\theta^\nu(t, s) & \text{if } t > \alpha, \\ 0 & \text{if } t \leqq \alpha, \end{cases}$$

where $\{\theta^\nu\}$, $\nu = 1, 2, \cdots$, is a standard sequence of continuous functions which approach the Heaviside step function θ monotonely (Theorem 3 : 34).†

To apply this theorem of F. Riesz's to a particular functional F we need only recognize F as an element of $C_0(I)^*$, which is easily done by § 3 : 8 or otherwise, calculate f by (6) or its equivalent, and conclude that (5) is valid. Note that (5) is the analogue of (2) and the use of (6) is the analogue of the substitution of (3) in (2).

One may deduce from Riesz's theorem the corresponding result for the space $C_n(I)$ of functions on I n-fold continuously differentiable, with norm

$$|||x|||_{C_n(I)} = \max_{i=0,1,\cdots,n} \sup_{s \in I} |x_i(s)|, \qquad x \in C_n(I),$$

where x_i denotes the ith derivative of x (Theorem 3 : 39). The standard form of an element F of $C_n(I)^*$ is

(7) $$Fx = \sum_{i<n} c^i x_i(a) + \int_I x_n(s)\, df(s), \qquad x \in C_n(I),$$

† In our cross-references a number before a colon is a chapter number and a number after a colon is the number of the object referred to within the designated chapter. If no colon is present, the reference is to the numbered object in the current chapter.

Thus in Chapter 2, for example, (3: 4) would refer to item (4) of Chapter 3; (3: 4, 5) would refer to items (4) and (5) of Chapter 3; and (6) would refer to item (6) of Chapter 2 itself. Theorem 7: 8 would refer to Theorem 8 of Chapter 7. Theorem 7: 8 is not necessarily the 8th theorem in that chapter, but the 8th numbered object, which happens to be a theorem. Similarly § 1 : 2 is the section labeled 2 in Chapter 1. Chapter numbers will appear in the captions of left-hand pages.

References to the bibliography are in square brackets, near the author's name or including his name. Thus [Newton 3] or . . . Newton . . . [3] would refer to the 3rd work listed under Newton.

where the constants c^i and the function f of bounded variation are given in (3 : 40, 41) and a is an arbitrary fixed element of I. The relation (7) leads to sharp appraisals of Fx, $x \in C_n(I)$, because the elements

$$x_0(a), x_1(a), \cdots, x_{n-1}(a); \qquad x_n \in C_0(I),$$

that describe x are independent of one another: For any set $b^0, b^1, \cdots, b^{n-1}$ of numbers and any function $y \in C_0(I)$, there is a function $x \in C_n(I)$ such that

$$x_i(a) = b^i, \qquad i < n;$$

$$x_n(s) = y(s), \qquad s \in I.$$

The function f in (7) is determined by the functional F. If f is absolutely continuous, then

$$df(s) = f_1(s)\, ds;$$

and the integral in (7) becomes an ordinary integral. Ordinary integrals have the advantage of being appraisable by all the Hölder inequalities (§ 1 : 21).

In actual practice the functional F is often given as a linear combination of integrals. Then relations like (7) may be established by a transformation of formulas. For functions of one variable, the powerful and elementary Theorem 1 : 8 states this fact. The theorem is due to Peano and dominates the subject. The theorem provides a standard form (7) in which f is absolutely continuous and a direct procedure for calculating the density f_1, for certain functionals F.

Every formula for a remainder in Steffensen's useful book [1], for example, can be deduced from Theorem 1 : 8. The theorem provides alternative formulas also, which are useful and have been neglected.

There are illustrations in Chapter 1 and extensive applications in Chapter 2.

In the approximation of a given functional G, we often consider a family $\mathscr{A}$ of functionals and put to ourselves the problem of choosing one functional $A \in \mathscr{A}$ to serve as the approximation of G. For each $A \in \mathscr{A}$, there is a remainder

$$Rx = Gx - Ax, \qquad x \in X.$$

A procedure often followed is to choose $A \in \mathscr{A}$ so that R vanish for polynomials of as high degree, say $n - 1$, as possible. Such a criterion of choice seems to me to be indirect. What it achieves is that Rx is expressible in terms of the nth derivative x_n. What we want is that Rx be as small as possible in some sense for the set of functions x on which we will operate, or that the appraisal of Rx that we use be as small as possible.

Various criteria of choice of $A \in \mathscr{A}$ are discussed in § 2 : 1 and in Chapters 9, 10.

§ 8. **Several variables.** For spaces X of functions of a single real variable, our knowledge of the adjoint space X^* is usually adequate. This is not the case for spaces of functions of several variables.

Consider a normed space X of functions on a closed bounded region D of the s, t-plane to the reals R. For some such spaces a standard form of elements in X^* and a direct procedure for finding sharp appraisals are not known.

An example is the space $C_1^2(D)$ of functions which have continuous first partial derivatives on D, with norm

$$\|x\|_{C_1^2(D)} = \max[\sup|x(s,t)|, \sup|x_{1,0}(s,t)|, \sup|x_{0,1}(s,t)|], \qquad x \in C_1^2(D),$$

where the suprema are taken for $(s, t) \in D$. There is no known procedure for calculating a standard form of an element F of $C_1^2(D)^*$, even when D is an interval. We know that Fx may be written as a sum of Stieltjes integrals on x, $x_{1,0}$, $x_{0,1}$; but we do not know a direct way of obtaining such a formula for an arbitrary $F \in C_1^2(D)^*$. Nor, if we have such a formula, do we know how to obtain effectively a sharp appraisal of Fx, $x \in C_1^2(D)$. The difficulties here are related to the fact that the derivatives $x_{1,0}$ and $x_{0,1}$ are not independent on D.

Thus it is important to discover particular spaces X for which we can find direct procedures of obtaining standard forms and sharp appraisals of arbitrary elements of X^*. Our spaces B, K, Z of Chapters 4, 5, 6, 7 are of this sort.

The spaces B, K and the related spaces $\mathscr{B}^*$, $\mathscr{K}^*$ have the following properties:

(i) Elements of B^*, K^*, $\mathscr{B}^*$, or $\mathscr{K}^*$ are readily recognizable (§§ 4: 80; 6: 13, 21, 47).

(ii) Remainders in approximation are often elements of suitable B^*, K^*, $\mathscr{B}^*$, or $\mathscr{K}^*$.

(iii) Each element of B^*, K^*, $\mathscr{B}^*$, or $\mathscr{K}^*$ may be written in a standard form. There is an explicit direct procedure for obtaining that form and sharp appraisals thereof.

Chapter 4 provides the elementary part of the theory. It is the counterpart for functions of two variables of Chapter 1.

Chapter 5 discusses applications to integration, substitution (interpolation or smoothing), and differentiation.

Chapter 6 provides the complete theory of the spaces B, K and their adjoints. The functionals which are elements of B^*, K^* are considered intrinsically; it is not required that such elements be given as sums of Stieltjes integrals. The culmination of the chapter is Theorem 6: 58 which gives a standard form for Fx in terms of ordinary integrals, when $F \in K^*$ and $x \in B$. It is Theorem 6: 58 which motivates the definition of the space K and which shows that K is the proper companion of B.

The spaces B, K are defined in terms of a compact interval

$$I = I_s \times I_t, \qquad I_s = \{\alpha \leqq s \leqq \tilde{\alpha}\}, \qquad I_t = \{\beta \leqq t \leqq \tilde{\beta}\},$$

of the s, t-plane; and an arbitrary fixed point (a, b) of I. The spaces B, K consist of functions x on I for which certain specified partial derivatives $x_{i,j}$ are continuous on I, and other specified partial derivatives $x_{i,j}$, with $s = a$, are continuous on I_t, and other specified partial derivatives $x_{i,j}$, with $t = b$, are continuous on I_s. There is a great variety of spaces B and K (§ 4 : 49; § 6 : 38).

The spaces B and K have a rectangular character which may appear to give our theory a limited scope but which on the contrary is a source of strength. Each condition that restricts a space X at the same time broadens the adjoint space X^*. Now our principal hypothesis in the first chapters is that

$$F \in K^* \quad \text{and} \quad x \in B.$$

The condition $x \in B$ is easily carried by functions that we encounter, as a rule. Indeed x usually has more than enough continuous differentiability. And the condition $F \in K^*$, because of the narrowness of K, is relatively broad. The functional F need not have a rectangular character but may depend entirely on an arbitrary subset of the interval I that enters in the definition of B and K. Cf. the illustrations in §§ 5 : 2, 18, 26. Thus the weight of our hypothesis falls on x rather than on F.

Furthermore, and this point is compelling, the hypotheses of our theorems are necessary as well as sufficient for the conclusions.

In Chapter 7, the theory is extended to spaces of functions of m variables.

Chapters 1–7 consider functions on compact intervals. Compactness makes the formulas for masses and kernels simpler than they otherwise would be (cf. Theorem 6 : 9). This gain in simplicity is important in the theory of approximation where the task to be done includes the calculation of masses and kernels as means to obtain explicit formulas and sharp appraisals.

§ 9. **General linear formulas.** In the first part of the book functionals are evaluated in terms of derivatives. It is natural to ask what objects, if any, other than derivatives may be used. We give a complete answer to this question for linear continuous operators, in† Chapter 8.

By an operator F we mean a map of a function space X into a space Y. In many cases Y is itself a function space. If Y is the space N of numbers, then the operator F is a functional.

The spaces that we consider are normed. Different norms and different sorts of norms are useful in the theory of approximation. There are norms based on suprema such as those in the spaces C_n, B, K; and norms based on averages (§ 9 : 1). There are norms based on the function itself, such as the norm in C_0; and norms based on certain derivatives of the function, such as the norm in C_2. The norms in X and Y determine the open sets and the meaning of continuity of operators.

† Chapter 8 does not depend on its predecessors.

In approximation we often start with a preproblem and then construct a precise problem which is an acceptable instance of the preproblem. If we are interested in an operator (F) vaguely defined on a vague space (X) to a vague space (Y), we replace these vague objects by precise ones F, X, Y. Whenever possible we arrange our choice of F, X, Y so that F be continuous. The wide choice of norms and of normed function spaces is very helpful here. It is natural to require continuity of F because otherwise a small change in input might induce a large change in output. But the meaning of smallness is to an extent at our disposal.

What we seek is a problem which fits the preproblem and which allows strong conclusions from relatively weak hypotheses.

Continuity is desirable and often attainable. Linearity is less universal. We say that an operator F on X is linear if

$$F(ax + by) = aFx + bFy$$

whenever

$$x, y \in X; \qquad a, b \in \mathsf{N}.$$

In this book we consider only linear operators. The hypothesis of linearity is very rewarding, as will be seen. Furthermore the preproblem often permits hypotheses of linearity, because derivatives, integrals, sums, and values of a function all are linear.

Minimax approximation is an example of a nonlinear process. Nonlinear theories tend to involve greater computational difficulties than do linear theories.

In the study of linear continuous operators we often seek to describe one operator in terms of another. Suppose that R is a linear continuous operator on X to Y and that U is a linear continuous operator on X to all of $\tilde{X}$, where X, $\tilde{X}$, Y are normed linear complete spaces. In order that there exist a linear continuous operator Q on $\tilde{X}$ to Y such that

$$(10) \qquad Rx = QUx, \qquad x \in X;$$

it is necessary and sufficient that

$$(11) \qquad Rx = 0 \quad \text{whenever} \quad Ux = 0, \qquad x \in X.$$

This is the quotient theorem of Chapter 8.

That the operator Q is continuous is important here. The relation (10) implies the sharp appraisal

$$(12) \qquad \|Rx\|_Y \leqq \|Q\| \, \|Ux\|_{\tilde{X}}, \qquad x \in X,$$

where $\|Q\|$ is a finite number determined by Q.

It is striking that the simple condition (11) and the preliminary hypotheses permit one to deduce the representation (10) and the appraisal (12).

The representations of elements of C_n^*, B^*, K^*, and Z^* of the earlier chapters all are instances of (10). In Theorem 3 : 39, for example, U is the operator which assigns to $x \in C_n$ the ordered set

$$x(a), x_1(a), \cdots, x_{n-1}(a), \quad \text{and} \quad x_n \in C_0(I).$$

In this case Ux is a set of n numbers and one function.

When studying an approximation in which the remainder is Rx, $x \in X$, a mathematician may review the variety of ways to construct operators U and spaces $\tilde{X}$ for which the hypotheses of the quotient theorem would be valid. Thus Ux, instead of being a derivative or set of derivatives as in earlier chapters, may be a specified linear combination of derivatives or sets thereof or a difference or a specified linear combination of differences or other things.

§ 13. **The effect of error in input.** We often wish to approximate

$$Gx, \qquad x \in X,$$

by

$$A(x + \delta x), \qquad x + \delta x \in X,$$

where

- X is a space of functions (or equivalence sets of functions) on a space S;
- Y is a space;
- G is a given operator on X to Y;
- A is an operator on X to Y.

We may say that x is the ideal input, $x + \delta x$ the actual input, δx the error in the input, Gx the desired output, and $A(x + \delta x)$ the actual output. The error in the approximation is

$$e = A(x + \delta x) - Gx \tag{14}$$

and is an element of Y.

The theory of approximation is concerned with the calculation of $A(x + \delta x)$, the appraisal of e, and the choice of $A \in \mathscr{A}$ when a family $\mathscr{A}$ of admissible approximations is given.

It is sometimes advantageous to write

$$e = e_A + e_{\delta x}, \tag{15}$$

where

$$e_A = Ax - Gx,$$
$$e_{\delta x} = A(x + \delta x) - Ax.$$

We may call e_A the truncation error or error due to A and $e_{\delta x}$ the error due to δx. We may put

$$R = G - A,$$
$$Rx = -e_A,$$

and study the operator R by the quotient theorem.

If A is linear, then

$$e_{\delta x} = A\,\delta x$$

and $e_{\delta x}$ is independent of x.

The above decomposition of e into e_A and $e_{\delta x}$ may be useful when our knowledge or hypotheses about x and δx are of different sorts. For example, we may know that δx is small in one sense and that Ux is small in another sense, for suitable U.

§ 16. **The use of probability.** In the problem of the preceding section more than one x and one δx enter. Either the operator A is to be used repeatedly on different inputs or, if A is to be used only once, we nonetheless do not know precisely what x and δx will be and we must provide for a number of possibilities.

There are two ways of describing the multiplicity of inputs. For brevity let us consider the input x. Our comments will apply similarly to the input δx.

We may say that x is an arbitrary element of a specified set $M \subset X$ and we may treat all elements of M as equally important. Alternatively we may introduce a probability to indicate the anticipated importance of different $x \in X$.

Each element of X is a function on S. We may introduce a probability as follows. Let Ω be a space and let p be a probability defined on Ω. Assume that there is given a function x on $S \times \Omega$ and that, for each fixed $\omega \in \Omega$ with null exceptions,

$$x_\omega \in X$$

and x_ω is a possible ideal input of § 13 above, where x_ω is the function on S obtained by fixing ω in $x(s, \omega)$, $s \in S$, $\omega \in \Omega$. Assume further that

$$p(\Omega_1)$$

is the probability that the ideal input of § 13 will be x_ω with $\omega \in \Omega_1$, where Ω_1 is an arbitrary measurable subset of Ω (§§ 9 : 174, 175). Thus $p(\Omega_1)$ is a measure of the anticipated frequency of occurrence of inputs x_ω with $\omega \in \Omega_1$. Alternatively $p(\Omega_1)$ is a measure of the anticipated importance of such inputs. Since x is a function on $S \times \Omega$, x is now a stochastic process.

Similarly δx may be a stochastic process and the same number $p(\Omega_1)$ may be the probability that the input error of § 13 will be δx_ω with $\omega \in \Omega_1$.

The space Ω and the probability p may be simple or not. In many pre-problems it is natural to suppose that an appropriate Ω and p exist. If p is not known we may attempt to estimate p by some statistical technique. Known and future statistical theories of stochastic processes may afford effective methods of estimating p and x or δx (§§ 9 : 175, 304). Even without a technique of statistical estimation we may assume that p and x or δx are of