M. J. D. Powell

Approximation Theory and Methods

逼近理论和方法

CAMBRIDGE

世界图书出版公司
www.wpcbj.com.cn

M. J. D. POWELL
John Humphrey Plummer Professor of Applied Numerical Analysis
University of Cambridge

Approximation theory and methods

CAMBRIDGE
UNIVERSITY PRESS

图书在版编目（CIP）数据

逼近理论和方法 = Approximation theory and methods：英文/（英）鲍威尔（Powell，M. J. D.）著. —影印本. —北京：世界图书出版公司北京公司，2014.9
ISBN 978－7－5100－8625－0

Ⅰ.①逼…　Ⅱ.①鲍…　Ⅲ.①逼近论—教材—英文　Ⅳ.①O174.41

中国版本图书馆 CIP 数据核字（2014）第 211071 号

书　　名：	Approximation Theory and Methods
作　　者：	M. J. D. Powell
中 译 名：	逼近理论和方法
责任编辑：	高蓉　刘慧
出 版 者：	世界图书出版公司北京公司
印 刷 者：	三河市国英印务有限公司
发　　行：	世界图书出版公司北京公司（北京朝内大街 137 号 100010）
联系电话：	010－64021602，010－64015659
电子信箱：	kjb@wpcbj.com.cn
开　　本：	24 开
印　　张：	15
版　　次：	2015 年 1 月
版权登记：	图字：01－2013－8236
书　　号：	978－7－5100－8625－0　　定　　价：49.00 元

Approximation theory and methods

Approximation Theory and Methods (978-0-521-29514-7) by M. J. D. Powell, first published by Cambridge University Press 2001

PREFACE

There are several reasons for studying approximation theory and methods, ranging from a need to represent functions in computer calculations to an interest in the mathematics of the subject. Although approximation algorithms are used throughout the sciences and in many industrial and commercial fields, some of the theory has become highly specialized and abstract. Work in numerical analysis and in mathematical software is one of the main links between these two extremes, for its purpose is to provide computer users with efficient programs for general approximation calculations, in order that useful advances in the subject can be applied. This book presents the view of a numerical analyst, who enjoys the theory, and who is keenly interested in its importance to practical computer calculations. It is based on a course of twenty-four lectures, given to third-year mathematics undergraduates at the University of Cambridge. There is really far too much material for such a course, but it is possible to speak coherently on each chapter for about one hour, and to include proofs of most of the main theorems. The prerequisites are an introduction to linear spaces and operators and an intermediate course on analysis, but complex variable theory is not required.

Spline functions have transformed approximation techniques and theory during the last fifteen years. Not only are they convenient and suitable for computer calculations, but also they provide optimal theoretical solutions to the estimation of functions from limited data. Therefore seven chapters are given to spline approximations. The classical theory of best approximations from linear spaces with respect to the minimax, least squares and L_1-norms is also studied, and algorithms are described and analysed for the calculation of these approximations. Interpolation is considered also, and the accuracy of interpolation and other linear operators is related to the accuracy of optimal algorithms. Special attention is given to polynomial functions, and there is one chapter on rational functions, but, due to the constraints of twenty-four lectures, the approximation of functions of several variables is not included. Also there are no computer listings, and little attention is given to the consequences of the rounding errors of computer arithmetic. All theorems are proved, and the reader will find that the subject provides a wide range of techniques of proof. Some material is included in order to demonstrate these techniques, for example the analysis of the convergence of the exchange algorithm for calculating the best minimax approximation to a continuous function. Several of the proofs are new. In particular, the uniform boundedness theorem is established in a way that does not require any ideas that are more advanced than Cauchy sequences and completeness. Less functional analysis is used than in other books on approximation theory, and normally functions are assumed to be continuous, in order to simplify the presentation. Exercises are included with each chapter which support and extend the text. All references to related work are given in an appendix.

It is a pleasure to acknowledge the excellent opportunities I have received for research and study in the Department of Applied Mathematics and Theoretical Physics at the University of Cambridge since 1976, and before that at the Atomic Energy Research Establishment, Harwell. My interest in approximation theory began at Harwell, stimulated by the enthusiasm of Alan Curtis, and strengthened by Pat Gaffney, who developed some of the theory that is reported in Chapter 24. I began to write this book in the summer of 1978 at the University of Victoria, Canada, and I am grateful for the facilities of their Department of Mathematics, for the encouragement of Ian Barrodale and Frank Roberts, and for financial support from grants A5251 and A7143 of the National Research Council of Canada. At Cambridge David Carter of King's College kindly studied drafts of the chapters and offered helpful comments. The manuscript was typed most expertly by Judy Roberts, Hazel Felton, Margaret Harrison and Paula Lister. I wish to express special thanks to Hazel for her assistance and patience when I was redrafting the text. My wife, Caroline, not only showed sympathetic understanding at home during the time when I worked long hours to complete the manuscript, but also she assisted with the figures. This work is dedicated to Caroline.

Pembroke College, Cambridge M. J. D. POWELL
January 1980

CONTENTS

1

The approximation problem and existence of best approximations

1.1 Examples of approximation problems

A simple example of an approximation problem is to draw a straight line that fits the curve shown in Figure 1.1. Alternatively we may require a straight line fit to the data shown in Figure 1.2. Three possible fits to the discrete data are shown in Figure 1.3, and it seems that lines B and C are better than line A. Whether B or C is preferable depends on our confidence in the highest data point, and to choose between the two straight lines we require a measure of the quality of the trial approximations. These examples show the three main ingredients of an approximation calculation, which are as follows: (1) A function, or some data, or

Figure 1.1. A function to be approximated.

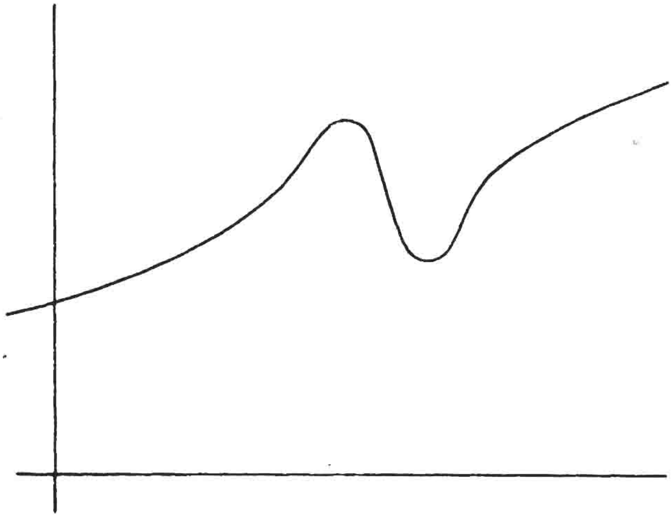


Figure 1.2. Some data to be approximated.

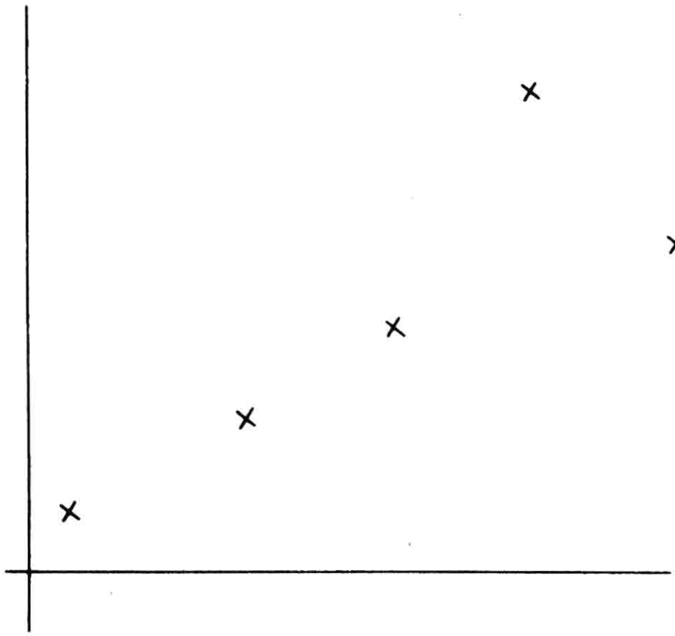
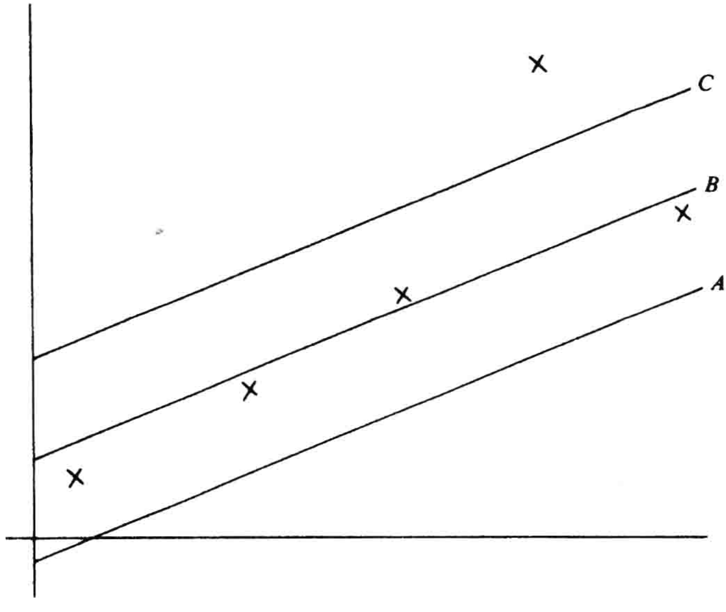


Figure 1.3. Three straight-line fits to the data of Figure 1.2.



more generally a member of a set, that is to be approximated. We call it f. (2) A set, $\mathscr{A}$ say, of approximations, which in the case of the given examples is the set of all straight lines. (3) A means of selecting an approximation from $\mathscr{A}$.

Approximation problems of this type arise frequently. For instance we may estimate the solution of a differential equation by a function of a certain simple form that depends on adjustable parameters, where the measure of goodness of the approximation is a scalar quantity that is derived from the residual that occurs when the approximating function is substituted into the differential equation. Another example comes from the choice of components in electrical circuits. The function f may be the required response from the circuit, and the range of available components gives a set $\mathscr{A}$ of attainable responses. We have to approximate f by a member of $\mathscr{A}$, and we require a criterion that selects suitable components. Moreover, in computer calculations of mathematical functions, the mathematical function is usually approximated by one that is easy to compute.

Many closely related questions are of interest also. Given f and $\mathscr{A}$, we may wish to know whether any member of $\mathscr{A}$ satisfies a fixed tolerance condition, and, if suitable approximations exist, we may be willing to accept any one. It is often useful to develop methods for selecting a member of $\mathscr{A}$ such that the error of the chosen approximation is always within a certain factor of the least error that can be achieved. It may be possible to increase the size of $\mathscr{A}$ if necessary, for example $\mathscr{A}$ may be a linear space of polynomials of any fixed degree, and we may wish to predict the improvement in the best approximation that comes from enlarging $\mathscr{A}$ by increasing the degree. At the planning stage of a numerical method we may know only that f will be a member of a set $\mathscr{B}$, in which case it is relevant to discover how well any member of $\mathscr{B}$ can be approximated from $\mathscr{A}$. Further, given $\mathscr{B}$, it may be valuable to compare the suitability of two different sets of approximating functions, $\mathscr{A}_0$ and $\mathscr{A}_1$. Numerical methods for the calculation of approximating functions are required. This book presents much of the basic theory and algorithms that are relevant to these questions, and the material is selected and described in a way that is intended to help the reader to develop suitable techniques for himself.

1.2 Approximation in a metric space

The framework of metric spaces provides a general way of measuring the goodness of an approximation, because one of the basic properties of a metric space is that it has a distance function. Specifically, the distance function $d(x, y)$ of a metric space $\mathscr{B}$ is a real-valued function, that is defined for all pairs of points (x, y) in $\mathscr{B}$, and that has the following properties. If $x \neq y$, then $d(x, y)$ is positive and is equal to $d(y, x)$. If $x = y$, then the value of $d(x, y)$ is zero. The triangle inequality

$$d(x, y) \leqslant d(x, z) + d(z, y) \tag{1.1}$$

must hold, where x, y and z are any three points in $\mathscr{B}$.

In most approximation problems there exists a suitable metric space that contains both f and the set of approximations $\mathscr{A}$. Then it is natural to decide that $a_0 \in \mathscr{A}$ is a better approximation than $a_1 \in \mathscr{A}$ if the inequality

$$d(a_0, f) < d(a_1, f) \tag{1.2}$$

is satisfied. We define $a^* \in \mathscr{A}$ to be a best approximation if the condition

$$d(a^*, f) \leqslant d(a, f) \tag{1.3}$$

holds for all $a \in \mathscr{A}$.

The metric space should be chosen so that it provides a measure of the error of each trial approximation. For example, in the problem of fitting the data of Figure 1.2 by a straight line, we approximate a set of points $\{(x_i, y_i); i = 1, 2, 3, 4, 5\}$ by a function of the form

$$p(x) = c_0 + c_1 x, \tag{1.4}$$

where c_0 and c_1 are scalar coefficients. Because we are interested in only five values of x, the most convenient space is $\mathscr{R}^5$. The fact that $p(x)$ depends on two parameters is not relevant to the choice of metric space. We measure the goodness of the approximation (1.4) as the distance, according to the metric we have chosen, from the vector of function values $\{p(x_i); i = 1, 2, 3, 4, 5\}$ to the data values $\{y_i; i = 1, 2, 3, 4, 5\}$.

It may be important to know whether or not a best approximation exists. One reason is that many methods of calculation are derived from properties that are obtained by a best approximation. The following theorem shows existence in the case when $\mathscr{A}$ is compact.

Theorem 1.1

If $\mathscr{A}$ is a compact set in a metric space $\mathscr{B}$, then, for every f in $\mathscr{B}$, there exists an element $a^* \in \mathscr{A}$, such that condition (1.3) holds for all $a \in \mathscr{A}$.

Proof. Let d^* be the quantity

$$d^* = \inf_{a \in \mathscr{A}} d(a, f). \tag{1.5}$$

If there exists a^* in $\mathscr{A}$ such that this bound on the distance is achieved, then there is nothing to prove. Otherwise there is a sequence $\{a_i; i = 1, 2, \ldots\}$ of points in $\mathscr{A}$ which gives the limit

$$\lim_{i \to \infty} d(a_i, f) = d^*. \tag{1.6}$$

By compactness the sequence has at least one limit point in $\mathscr{A}$, a^+ say. Expression (1.6) and the definition of a^+ imply that, for any $\varepsilon > 0$, there exists an integer k such that the inequalities

$$d(a_k, f) < d^* + \tfrac{1}{2}\varepsilon \tag{1.7}$$

and

$$d(a_k, a^+) < \tfrac{1}{2}\varepsilon \tag{1.8}$$

are obtained. Hence the triangle inequality (1.1) provides the bound

$$\begin{aligned} d(a^+, f) &\leq d(a^+, a_k) + d(a_k, f) \\ &< d^* + \varepsilon. \end{aligned} \tag{1.9}$$

Because ε can be arbitrarily small, the distance $d(a^+, f)$ is not greater than d^*. Therefore a^+ is a best approximation. □

When $\mathscr{A}$ is not compact it is easy to find examples to show that best approximations may not exist. For instance, let $\mathscr{B}$ be the Euclidean space $\mathscr{R}^2$ and let $\mathscr{A}$ be the set of points that are strictly inside the unit circle. There is no best approximation to any point of $\mathscr{B}$ that is outside or on the unit circle.

1.3 Approximation in a normed linear space

The properties of metric spaces are not sufficiently strong for most of our work, so it is assumed that $\mathscr{A}$ and f are contained in a normed linear space, which we call $\mathscr{B}$ also when we want to refer to it. The norm is a real-valued function $\|x\|$ that is defined for all $x \in \mathscr{B}$. Its properties are such that the function

$$d(x, y) = \|x - y\| \tag{1.10}$$

is suitable as a distance function. Therefore, by letting z be zero in expression (1.1) and by reversing the sign of y, we may deduce the triangle inequality

$$\|x + y\| \leq \|x\| + \|y\|. \tag{1.11}$$

Moreover, the norm must satisfy the homogeneity condition

$$\|\lambda x\| = |\lambda| \, \|x\| \tag{1.12}$$

for all $x \in \mathscr{B}$ and for all scalars λ.

The specialization from metric spaces to normed linear spaces does not exclude any of the approximation problems that we will consider. Therefore mostly we use the distance function (1.10). It occurs naturally in the approximation calculations that are of practical interest, and it allows the existence of a best approximation to be proved when $\mathscr{A}$ is a linear space.

Theorem 1.2

If $\mathscr{A}$ is a finite-dimensional linear space in a normed linear space $\mathscr{B}$, then, for every $f \in \mathscr{B}$, there exists an element of $\mathscr{A}$ that is a best approximation from $\mathscr{A}$ to f.

Proof. Let the subset $\mathscr{A}_0$ contain the elements of $\mathscr{A}$ that satisfy the condition

$$\|a\| \leqslant 2\|f\|. \tag{1.13}$$

It is compact because it is a closed and bounded subset of a finite-dimensional space. It is not empty: for example it contains the zero element. Therefore, by Theorem 1.1, there is a best approximation from $\mathscr{A}_0$ to f which we call a_0^*. By definition the inequality

$$\|a - f\| \geqslant \|a_0^* - f\|, \qquad a \in \mathscr{A}_0, \tag{1.14}$$

holds. Alternatively, if the element a is in $\mathscr{A}$ but is not in $\mathscr{A}_0$ then, because condition (1.13) is not obtained we have the bound

$$\begin{aligned} \|a - f\| &\geqslant \|a\| - \|f\| \\ &> \|f\| \\ &\geqslant \|a_0^* - f\|, \end{aligned} \tag{1.15}$$

where the last line makes further use of the fact that the zero element is in $\mathscr{A}_0$. Hence expression (1.14) is satisfied for all a in $\mathscr{A}$, which proves that a_0^* is a best approximation. □

1.4 The L_p-norms

In most of the approximation problems that we consider, f and $\mathscr{A}$ are in the space $\mathscr{C}[a\ b]$, which is the set of continuous real-valued functions that are defined on the interval $[a, b]$ of the real line. Occasionally we turn to discrete problems, where f and $\mathscr{A}$ are in $\mathscr{R}^m$, which is the set of real m-component vectors. Both of these spaces are linear and we have a choice of norms.

We study the three norms that are used most frequently, namely the L_p-norms in the cases when $p = 1$, 2 and ∞. For finite p the L_p-norm in