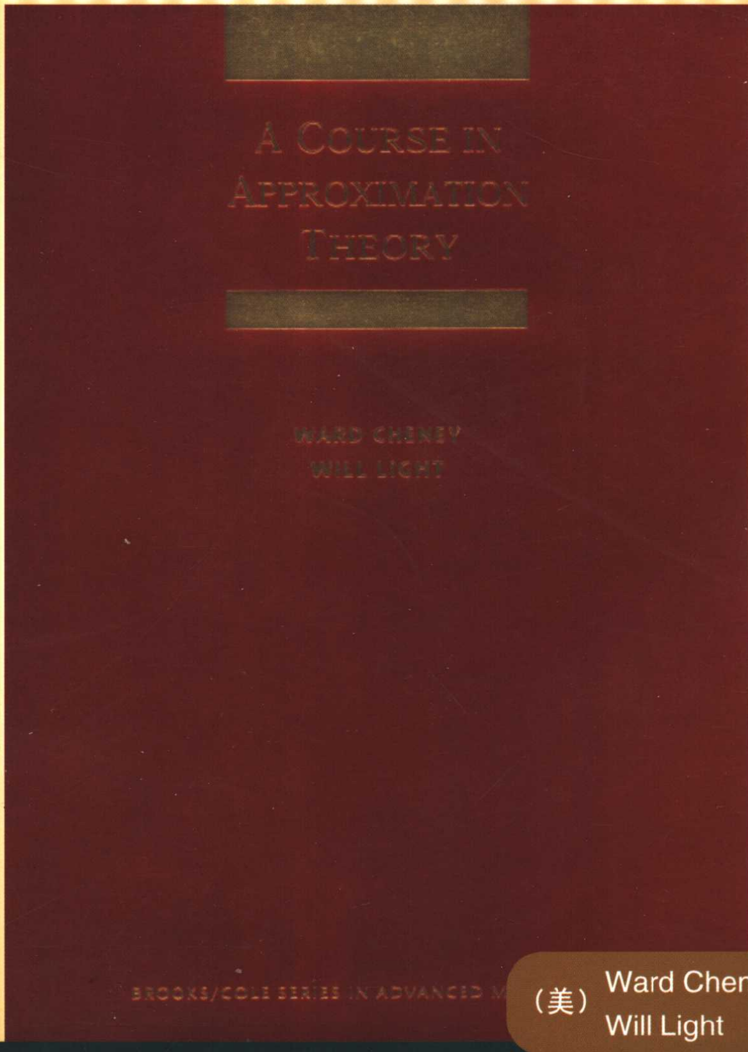
HZ BOOKS

经典原版书库

逼近论教程

(英文版)



(美) Ward Cheney
Will Light 著

机械工业出版社
China Machine Press

经典原版书库

逼近论教程

(英文版)

A Course in Approximation Theory

(美) Ward Cheney
Will Light 著

机械工业出版社
China Machine Press

Ward Cheney and Will Light

A Course in Approximation Theory

ISBN: 0-534-36224-9

981-254-190-X

本书版权登记号：图字：01-2003-8369

图书在版编目（CIP）数据

逼近论教程（英文版）/（美）谢尼（Cheney, W.）等著. －北京：机械工业出版社，2004.1

（经典原版书库）

书名原文：A Course in Approximation Theory

ISBN 7-111-13512-1

Ⅰ. 逼… Ⅱ. 谢… Ⅲ. 函数逼近论－教材－英文 Ⅳ. O174.41

中国版本图书馆CIP数据核字（2003）第112681号

机械工业出版社（北京市西城区百万庄大街22号 邮政编码 100037）

责任编辑：迟振春

北京昌平奔腾印刷厂印刷・新华书店北京发行所发行

2004年1月第1版第1次印刷

787mm × 1092mm 1/16 ・23印张

印数：0 001－3 000册

定价：39.00元

凡购本书，如有倒页、脱页、缺页，由本社发行部调换

本社购书热线电话：（010）68326294

Preface

This book offers a graduate-level exposition of selected topics in modern approximation theory. A large portion of the book focuses on multivariate approximation theory, where much recent research is concentrated. Although our own interests have influenced the choice of topics, the text cuts a wide swath through modern approximation theory, as can be seen from the table of contents. We believe the book will be found suitable as a text for courses, seminars, and even solo study. Although the book is at the graduate level, it does not presuppose that the reader already has taken a course in approximation theory.

Topics of This Book

A central theme of the book is the problem of interpolating data by smooth multivariable functions. Several chapters investigate interesting families of functions that can be employed in this task; among them are the polynomials, the positive definite functions, and the radial basis functions. Whether these same families can be used, in general, for approximating functions to arbitrary precision is a natural question that follows; it is addressed in further chapters.

The book then moves on to the consideration of methods for concocting approximations, such as by convolutions, by neural nets, or by interpolation at more and more points. Here there are questions of limiting behavior of sequences of operators, just as there are questions about interpolating on larger and larger sets of nodes.

A major departure from our theme of multivariate approximation is found in the two chapters on univariate wavelets, which comprise a significant fraction of the book. In our opinion wavelet theory is so important a development in recent times—and is so mathematically appealing—that we had to devote some space to expounding its basic principles.

The Style of This Book

In style, we have tried to make the exposition as simple and clear as possible, electing to furnish proofs that are complete and relatively easy to read without the reader needing to resort to pencil and paper. Any reader who finds this style too prolix can proceed quickly over arguments and calculations that are routine. To paraphrase Shaw: We have done our best to avoid conciseness! We have also made considerable efforts to find simple ways to introduce and explain each topic. We hope that in doing so, we encourage readers to delve deeper into some areas. It should be borne in mind that further exploration of some topics may require more mathematical sophistication than is demanded by our treatment.

Organization of the Book

A word about the general plan of the book: we start with relatively elementary matters in a series of about ten short chapters that do not, in general, require more of the reader than undergraduate mathematics (in the American university system). From that point on, the gradient gradually increases and the text becomes more demanding, although still largely self-contained. Perhaps the most significant demands made on the technical knowledge of the reader fall in the areas of measure theory and the Fourier transform. We have freely made use of the Lebesgue function spaces, which bring into play such measure-theoretic results as the Fubini Theorem. Other results such as the Riesz Representation Theorem for bounded linear functionals on a space of continuous functions and the Plancherel Theorem for Fourier transforms also are employed without compunction; but we have been careful to indicate explicitly how these ideas come into play. Consequently, the reader can simply accept the claims about such matters as they arise. Since these theorems form a vital part of the equipment of any applied analyst, we are confident that readers will want to understand for themselves the essentials of these areas of mathematics. We recommend Rudin's *Real and Complex Analysis* (McGraw-Hill, 1974) as a suitable source for acquiring the necessary measure theoretic ideas, and the book *Functional Analysis* (McGraw-Hill, 1973) by the same author as a good introduction to the circle of ideas connected with the Fourier transform.

Additional Reading

We call the reader's attention to the list of books on approximation theory that immediately precedes the main section of references in the bibliography. These books, in general, are concerned with what we may term the "classical" portion of approximation theory—understood to mean the parts of the subject that already were in place when the authors were students. As there are very few textbooks covering recent theory, our book should help to fill that "much needed gap," as some wag phrased it years ago. This list of books emphasizes only the systematic textbooks for the subject as a whole, not the specialized texts and monographs.

Acknowledgments

It is a pleasure to have this opportunity of thanking three agencies that supported our research over the years when this book was being written: the Division of Scientific Affairs of NATO, the Deutsche Forschungsgemeinschaft, and the Science and Engineering Research Council of Great Britain. For their helpful reviews of our manuscript, we thank Robert Schaback, University of Göttingen, and Larry Schumaker, Vanderbilt University. We acknowledge also the contribution made by many students, who patiently listened to us expound the material contained in this book and who raised incisive questions. Students and colleagues in Austin, Leicester, Würzburg, Singapore, and Canterbury (NZ) all deserve our thanks. Professor S. L. Lee was especially helpful.

A special word of thanks goes to Ms. Margaret Combs of the University of Texas Mathematics Department. She is a superb technical typesetter in the modern sense of the word, that is, an expert in TEX. She patiently created the TEX files for lecture notes, starting about six years ago, and kept up with the constant editing of these notes, which were to become the backbone of the book.

The staff of Brooks/Cole Publishing has been most helpful and professional in guiding this book to its publication. In particular, we thank Gary Ostedt, sponsoring editor; Ragu Raghavan, marketing representative; and Janet Hill, production editor, for their personal contact with us during this project.

How to Reach Us

Readers are encouraged to bring errors and suggestions to our attention. E-mail is excellent for this purpose: our addresses are cheney@math.utexas.edu and pwl@mcs.le.ac.uk. A web site for the book is maintained at http:www.math.utexas.edu/user/cheney/ATBOOK.

Ward Cheney

Will Light

Contents

1 Introductory Discussion of Interpolation

We shall be concerned with real-valued functions defined on a domain X, which need not be specified at this moment. (It will often be a subset of $\mathbb{R}$, $\mathbb{R}^2, \ldots$, but can be more general.) In the domain X a set of n distinct points is given:

$$\mathcal{N} = \{x_1, x_2, \ldots, x_n\}$$

These points are called **nodes**, and $\mathcal{N}$ is the **node set**. For each node x_i an **ordinate** $\lambda_i \in \mathbb{R}$ is given. (Each λ_i is a real number.) The problem of **interpolation** is to find a suitable function $F : X \to \mathbb{R}$ that takes these prescribed n values. That is, we want

$$F(x_i) = \lambda_i \qquad (1 \leq i \leq n)$$

When this occurs, we say that F **interpolates** the given data $\{(x_i, \lambda_i)\}_{i=1}^n$. Usually F must be chosen from a preassigned family of functions on X.

A wide variety of functions F may be suitable. Figures 1.1 and 1.2 show 12 different interpolation functions for a single data set. The nodes are 5 real numbers. They and the specified ordinates are given in this table:

x	1.2	3.6	4.3	6.1	7.8
y	1.2	3.1	−1.3	2.7	1.4

In Figure 1.1a, the raw data points are shown. In Figures 1.1b to 1.1f, F has the form $F(x) = \sum_1^5 c_j u(x - x_j)$, in which u is a function of our choosing. First we took a B-spline of degree 0. To avoid the discontinuous nature of this example, we then took u to be a B-spline of degree 1, as shown in Figure 1.1c. To avoid discontinuities in the first two derivatives, we then let u be a cubic B-spline, as in Figure 1.1d. In Figure 1.1e we show the interpolant when $u(x) = |x|$, and in Figure 1.1f we used $u(x) = |x|^{1/2}$.

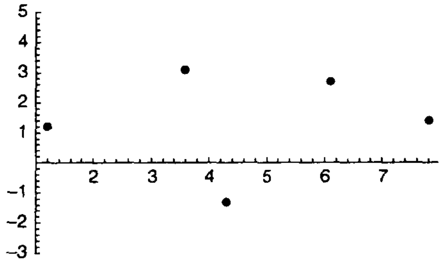


Figure 1.1a

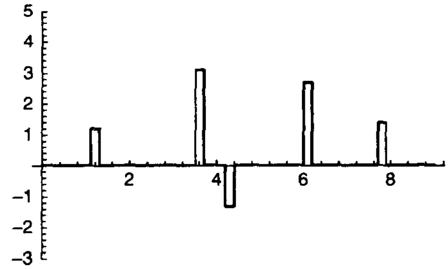


Figure 1.1b

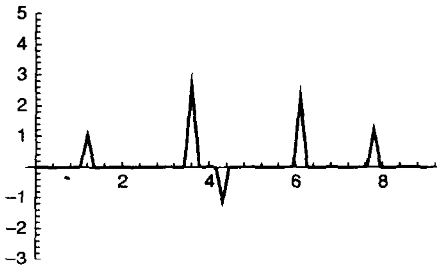


Figure 1.1c

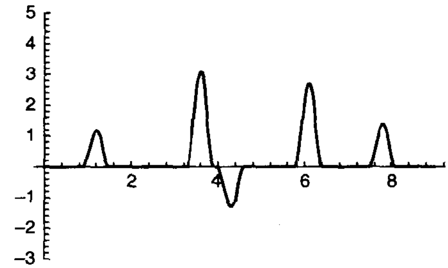


Figure 1.1d

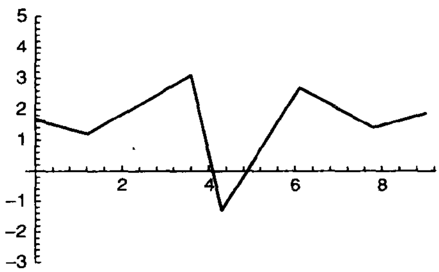


Figure 1.1e

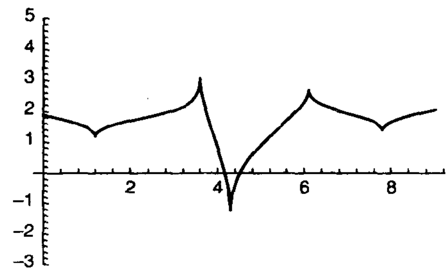


Figure 1.1f

Further examples are shown in Figure 1.2. Here we have used the same data as in Figure 1.1, but a different choice of interpolating functions. Specifically, 1.2a employs a fourth-degree polynomial; 1.2b employs a natural cubic spline; 1.2c is given by the Interpolation command in Mathematica and is also a cubic spline. In 1.2d, we used a cubic B-spline, B^3, determined by integer knots, and interpolated with $\sum_1^5 c_i B^3(x - x_i)$. In 1.2e, we used $\sum c_i e^{-(x-x_i)^2}$, and in 1.2f we used, in the same manner, a 0-degree B-spline. Some variations in scaling are noticeable in the figures.

The examples in Figures 1.1 and 1.2 suggest the great diversity among different types of interpolating functions. The selection of an appropriate type of interpolant must be made according to further criteria, above and beyond the basic requirement of inter-

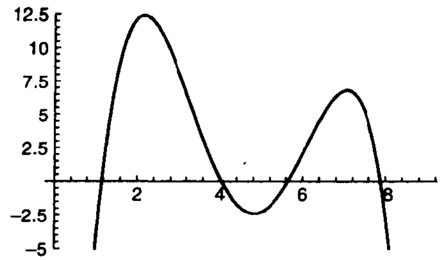


Figure 1.2a

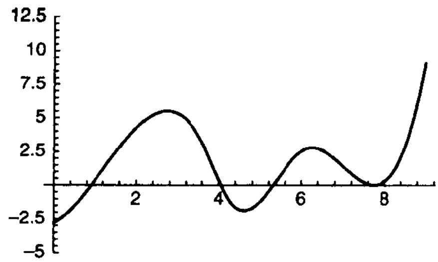


Figure 1.2b

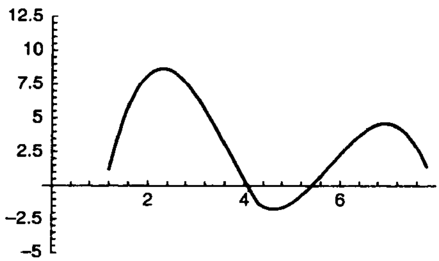


Figure 1.2c

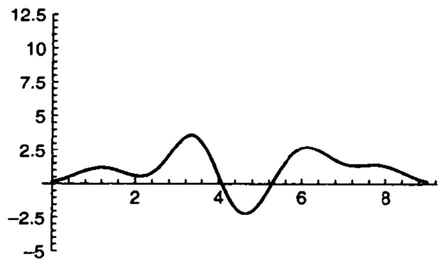


Figure 1.2d

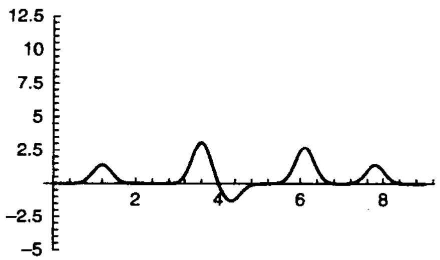


Figure 1.2e

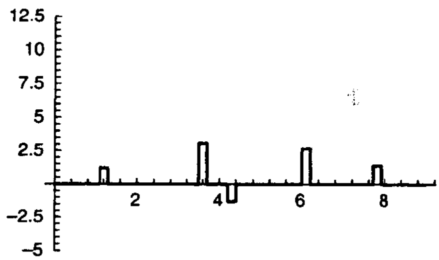


Figure 1.2f

polation. For example, in a specific application we may want the interpolating function to have a continuous first derivative. (That requirement would disqualify most of the functions in Figure 1.1.)

The **linear interpolation problem** is a special case that arises when F is to be chosen from a prescribed n-dimensional vector space of functions on X. Suppose that U is this vector space and that a basis for U is $\{u_1, u_2, \ldots, u_n\}$. The function F that we seek must have the form

$$F = \sum_{j=1}^{n} c_j u_j$$

When the interpolation conditions are imposed on F, we obtain

$$\lambda_i = F(x_i) = \sum_{j=1}^{n} c_j u_j(x_i) \qquad (1 \le i \le n)$$

This is a system of n linear equations in n unknowns. It can be written in matrix form as $Ac = y$, or in detail as

$$\begin{bmatrix} u_1(x_1) & u_2(x_1) & \cdots & u_n(x_1) \\ u_1(x_2) & u_2(x_2) & \cdots & u_n(x_2) \\ \cdots & \cdots & \cdots & \cdots \\ u_1(x_n) & u_2(x_n) & \cdots & u_n(x_n) \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \\ \cdots \\ c_n \end{bmatrix} = \begin{bmatrix} \lambda_1 \\ \lambda_2 \\ \cdots \\ \lambda_n \end{bmatrix}$$

The $n \times n$ matrix A appearing here is called the **interpolation matrix.** In order that our problem be solvable for any choice of ordinates λ_i it is necessary and sufficient that the interpolation matrix be nonsingular. The ideal situation is that this matrix be nonsingular for all choices of n distinct nodes.

THEOREM 1. *Let U be an n-dimensional linear space of functions on X. Let $x_1, x_2, \ldots, x_n$ be n distinct nodes in X. In order that U be capable of interpolating arbitrary data at the nodes it is necessary and sufficient that zero data be interpolated only by the zero-element in U.*

Proof. The space U can furnish an interpolant for arbitrary data if and only if the interpolation matrix A is nonsingular. An equivalent condition on the matrix A is that the equation $Ac = 0$ can be true only if $c = 0$. ■

Example. Let $X = \mathbb{R}$ and let $u_j(x) = x^{j-1}$, for $j = 1, 2, \ldots, n$. An $n \times n$ interpolation matrix in this special case is called a **Vandermonde matrix.** It looks like this:

$$V = \begin{bmatrix} 1 & x_1 & x_1^2 & \cdots & x_1^{n-1} \\ 1 & x_2 & x_2^2 & \cdots & x_2^{n-1} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 1 & x_n & x_n^2 & \cdots & x_n^{n-1} \end{bmatrix}$$

The determinant of V is given by the formula

$$\det V = \prod_{1 \le j < i \le n} (x_i - x_j)$$

This is obviously nonzero if and only if the nodes are distinct. Hence the interpolation problem has a unique solution for any choice of distinct nodes. We can also use Theorem 1 to see that V is nonsingular. Thus, we consider the "homogeneous" linear problem, in which we attempt to interpolate zero data. The solution will be a polynomial of degree at most $n - 1$ that takes the value 0 at each of the n nodes. Since a nonzero polynomial of degree at most $n-1$ can have at most $n-1$ zeros, we conclude that the zero polynomial is the only possible solution. ■

The Vandermonde matrix occurs often in mathematics. Refer to Rushanan [Rush], Grosof and Taiani [GT], Cheney [C1], for example. It is ill-conditioned for numerical work. See Gautschi, [Gau1, Gau2].

An n-dimensional vector space U of functions on a domain X is said to be a **Haar space** if the only element of U which has more than $n-1$ roots in X is the zero element. The next theorem provides some properties equivalent to the Haar property. In the theorem, we refer to **point-evaluation functionals.** If V is a vector space of functions on a set X, and if x is a point of X, then the point-evaluation functional corresponding to x is denoted by x^* and is defined on V by

$$x^*(f) = f(x) \qquad (f \in V)$$

Obviously x^* is linear, because

$$x^*(\alpha f + \beta g) = (\alpha f + \beta g)(x) = \alpha f(x) + \beta g(x) = \alpha x^*(f) + \beta x^*(g)$$

THEOREM 2. *Let U have the basis $\{u_1, u_2, \ldots, u_n\}$. These properties are equivalent:*

- **a.** *U is a Haar space*
- **b.** $\det\left(u_j(x_i)\right) \neq 0$ *for any set of distinct points $x_1, x_2, \ldots, x_n$ in X*
- **c.** *For any distinct points $x_1, x_2, \ldots, x_n$ in X, the set of point-evaluation functionals $x_1^*, x_2^*, \ldots, x_n^*$ spans the algebraic dual space U^**
- **d.** *If $x_1, x_2, \ldots, x_m$ are distinct in X and if $\sum_{i=1}^m \lambda_i u_j(x_i) = 0$ for $j = 1, 2, \ldots, n$ then either at least $n+1$ of the coefficients λ_i are nonzero, or $\sum_{i=1}^m |\lambda_i| = 0$*

Proof. To show that **a** implies **b**, suppose **b** false. Since the determinant of $(u_j(x_i))$ is zero, the matrix is singular, and there exists a nonzero vector $(c_1, c_2, \ldots, c_n)$ such that $\sum_{j=1}^n c_j u_j(x_i) = 0$, $(1 \leq i \leq n)$. Put $u = \sum_{j=1}^n c_j u_j$. Since $\{u_1, u_2, \ldots, u_n\}$ is linearly independent, $u \neq 0$. But $u(x_i) = 0$ for $1 \leq i \leq n$. Hence **a** is false.

To show that **b** implies **c**, suppose **b** true. Then the set $\{x_1^*, x_2^*, \ldots, x_n^*\}$ is linearly independent when these functionals are restricted to U. Indeed, if $\sum_{i=1}^n a_i x_i^* | U = 0$, then $\sum_{i=1}^n a_i x_i^*(u_j) = 0$ for $1 \leq j \leq n$, and by **b**, $\sum_{i=1}^n |a_i| = 0$. Since U^* is of dimension n, the functionals span U^*.

To show that **c** implies **d**, assume **c**. Let $x_1, \ldots, x_m$ be distinct points that satisfy $\sum_{i=1}^m \lambda_i u_j(x_i) = 0$ for $1 \leq j \leq n$. If $m \leq n$, then by **c** we can take additional points and obtain a basis $\{x_1^*, \ldots, x_n^*\}$ for U^*. Then the subset $\{x_1^*, \ldots, x_m^*\}$ is linearly independent on U and all λ_i are zero. If $m > n$ and $\sum_{i=1}^m |\lambda_i| \neq 0$ then at least $n+1$ of the λ_i are nonzero, for otherwise we will have a nontrivial linear combination of n (or fewer) x_i^* that vanishes on U, contrary to **c**.

To prove that **d** implies **a**, assume **d** and take $m = n$. Then the equation $\sum_{i=1}^n \lambda_i u_j(x_i) = 0$ for $1 \leq j \leq n$ implies $\sum_{i=1}^n |\lambda_i| = 0$. Hence the matrix $(u_j(x_i))$ is nonsingular. Thus if $(c_1, c_2, \ldots, c_n) \neq 0$, we cannot have $\sum_{j=1}^n c_j u_j(x_i) = 0$. In other words, a nonzero member of U cannot have n zeros. ■

Any basis for a Haar space is called a **Chebyshev system.** Here are some examples of Chebyshev systems on $\mathbb{R}$:

1. $1, x, x^2, \ldots, x^n$
2. $e^{\lambda_1 x}, e^{\lambda_2 x}, \ldots, e^{\lambda_n x} \qquad (\lambda_1 < \lambda_2 < \cdots < \lambda_n)$
3. $1, \cosh x, \sinh x, \ldots, \cosh nx, \sinh nx$

Here are some Chebyshev systems on $(0, \infty)$:

4. $x^{\lambda_1}, x^{\lambda_2}, \ldots, x^{\lambda_n} \qquad (\lambda_1 < \lambda_2 < \cdots < \lambda_n)$
5. $(x + \lambda_1)^{-1}, \ldots, (x + \lambda_n)^{-1} \qquad (0 \le \lambda_1 < \lambda_2 < \cdots < \lambda_n)$

Here is a Chebyshev system on the circle $\mathbb{R}/2\pi$:

6. $1, \cos\theta, \sin\theta, \ldots, \cos n\theta, \sin n\theta$

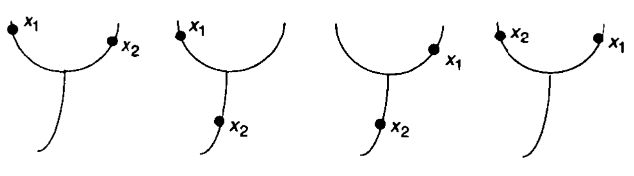
The notation $\mathbb{R}/2\pi$ denotes the set of reals with an equivalence relation: $x \equiv y$ if $x - y$ is an integer multiple of 2π.

Are there any Chebyshev systems of continuous functions on $\mathbb{R}^2$ and on the higher-dimensional Euclidean spaces? No, there is an immediate and absolute barrier:

THEOREM 3. *On $\mathbb{R}^2, \mathbb{R}^3, \ldots$ there are no Haar subspaces of continuous functions except one-dimensional ones.*

Proof. Suppose that $\{u_1, u_2, \ldots, u_n\}$ is a Chebyshev system of continuous functions on $\mathbb{R}^s$, where $s \ge 2$ and $n \ge 2$. By Theorem 2, $\det(u_j(x_i)) \ne 0$ for any set of distinct nodes $x_1, x_2, \ldots, x_n$ in $\mathbb{R}^s$. Select a closed path in $\mathbb{R}^s$ containing x_1 and x_2 but no other nodes. By moving x_1 and x_2 in the same direction continuously along this path, we can made x_1 and x_2 exchange positions without allowing them to coincide at any stage in the process. In the determinant above, rows 1 and 2 will exchange positions, and the determinant will change sign. Since the determinant is a continuous function of x_1 and x_2, it will assume the value 0 during this process, contrary to Theorem 2. ■

Even on domains X that are subsets of $\mathbb{R}^s (s \ge 2)$ it may be impossible to have Haar subspaces (of continuous functions) with dimension 2 or higher. Suppose that X is, or contains, a subset homeomorphic to the letter Y. (For example, in $\mathbb{R}^2$ consider the case when X is the union of two nonparallel lines.) Then there can exist no continuous Haar space of dimension 2 or more on X. The argument is as before. By a continuous movement of nodes x_1 and x_2 along the Y-shaped figure, their positions can interchange without being coincident at any stage. See the diagram.



The general theorem along these lines is due to Mairhuber [Mai]. Later elucidations and extensions are by Curtis [Cu], Sieklucki [Si], Lutts [Lut], McCullough and Wulbert [McW], Schoenberg [S10] and Schoenberg and Yang [SY]. The result is as follows.

THEOREM 4. *Let X be a compact Hausdorff space, and suppose that $C(X)$ contains a Haar subspace of dimension 2 or more. Then X is homeomorphic to a subset of the circumference of a circle.*

Linear interpolation is closely connected to the notion of linear independence of a set of functions. Suppose that $\{u_1, u_2, \ldots, u_m\}$ is a set of m real-valued functions defined on a set X. The set of functions is said to be **linearly independent on** X if this implication is valid:

(1) $$\sum_{j=1}^{m} c_j u_j(x) = 0 \text{ for all } x \in X \implies \sum_{j=1}^{m} |c_j| = 0$$

If D is a subset of X, we will say that $\{u_1, u_2, \ldots, u_m\}$ is **linearly independent on** D if the following implication is valid:

(2) $$\sum_{j=1}^{m} c_j u_j(x) = 0 \text{ for all } x \in D \implies \sum_{j=1}^{m} |c_j| = 0$$

Observe that the property in Implication (2) is stronger than the one in (1).

To connect these notions to the interpolation problem, let us consider a finite set of nodes in X:

(3) $$\mathcal{N} = \{x_1, x_2, \ldots, x_n\}$$

Interpolating arbitrary data on $\mathcal{N}$ by the functions u_j requires the solution of the system

(4) $$\sum_{j=1}^{m} c_j u_j(x_i) = \lambda_i \qquad (1 \le i \le n)$$

THEOREM 5. *Equation (4) is solvable for arbitrary $\lambda_1, \ldots, \lambda_n$ if and only if the $n \times m$ matrix $A_{ij} = u_j(x_i)$ has rank n.*

Proof. Write Equation (4) in the expanded form

(5) $$c_1 \begin{bmatrix} u_1(x_1) \\ \vdots \\ u_1(x_n) \end{bmatrix} + \cdots + c_m \begin{bmatrix} u_m(x_1) \\ \vdots \\ u_m(x_n) \end{bmatrix} = \begin{bmatrix} \lambda_1 \\ \vdots \\ \lambda_n \end{bmatrix}$$

This system is solvable if and only if the vector $\boldsymbol{\lambda} = (\lambda_1, \lambda_2, \ldots, \lambda_n)^T$ is in the column space of A. Equation (5) is solvable for all $\boldsymbol{\lambda}$ if and only if the column space of A contains $\mathbb{R}^n$. This occurs if and only if the column rank (and the rank) of A is n. (The rank cannot exceed n.) ∎

THEOREM 6. *If $m \le n$ and if Equation (4) is solvable for arbitrary $\lambda_1, \ldots, \lambda_n$, then the set of functions $u_1, \ldots, u_m$ is linearly independent on the set of nodes $x_1, \ldots, x_n$.*

Proof. If $\{u_1, \ldots, u_m\}$ is linearly dependent on the node set, then for suitable c, not zero, we have $\sum_{j=1}^{m} c_j u_j(x_i) = 0$ for $1 \le i \le n$. The set of columns displayed in Equation (5) is linearly dependent and has at most n elements. Hence, the column space of A is a proper subspace of $\mathbb{R}^n$, and System (5) fails to be solvable for some vectors λ. ■

Problems

One must learn by doing the thing;
for though you think you know it,
you have no certainty until you try.

—Sophocles

1. Prove the formula given for a Vandermonde determinant.
2. Prove that the functions $e^{\lambda_j x} (1 \leq j \leq n)$ form a Chebyshev system on $\mathbb{R}$.
3. Under what conditions will the functions $\cosh \lambda_j x$ $(1 \leq j \leq n)$ form a Chebyshev system on $\mathbb{R}$? What about $(0, \infty)$?
4. Consider the space $C(X)$, where X is a compact Hausdorff space. Use the norm $\|f\| = \sup_{x \in X} |f(x)|$ in this space. Prove that each point-evaluation functional has norm 1.
5. In the space $C[0, 1]$ use the norm $\|f\| = \int_0^1 |f(x)|\, dx$. What is the norm of a point-evaluation functional?
6. Prove that if X is a compact Hausdorff space and if $C(X)$ contains an n-dimensional Haar subspace with $n \geq 2$, then X is homeomorphic to a subset of $\mathbb{R}^n$. Hint: Let $\{u_1, u_2, \ldots, u_n\}$ be a basis for the Haar subspace, and define a map $f : X \to \mathbb{R}^n$ by writing $f(x) = [u_1(x), u_2(x), \ldots, u_n(x)]$. What is the correct theorem when $n = 1$?
7. Prove that Examples 4 and 5 given in the text are indeed Chebyshev systems on $(0, \infty)$.
8. Let $\{u_1, u_2, \ldots, u_n\}$ be a set of real-valued functions on a set X. Prove that the set of functions is linearly independent if and only if there exist n distinct points $x_1, x_2, \ldots, x_n$ in X such that $\det(u_j(x_i)) \neq 0$.
9. Let $\{u_1, u_2, \ldots, u_n\}$ be a linearly independent set of functions from a set X to the reals. Prove that there is a subset Y of X on which $\{u_1, u_2, \ldots, u_n\}$ is a Chebyshev system. Does there necessarily exist a maximal such Y? Illustrate with $n = 2$, $u_1(x) = x$, $u_2(x) = x^2$ on $\mathbb{R}$.
10. In the space Π_n consisting of all polynomials of degree at most n, let $\|p\|_\infty = \max_{-1 \leq x \leq 1} |p(x)|$. Find the norm of the functional x^* when $x = 2$. That is, compute

$$\sup_{\|p\|_\infty \leq 1} |p(2)|$$

You may need the theory of Chebyshev polynomials.

11. Let $S = \{u_1, u_2, \ldots, u_n\}$, where each u_i is a continuous function on a domain X. Let $Y \subset X$. Prove these assertions: (a) If S is linearly independent on Y then it is linearly

independent on X. (b) If S is a Chebyshev system on X then it is a Chebyshev system on Y, provided that Y has at least n elements. Show by examples that in assertions (a) and (b) we cannot interchange X and Y.

12. Is a subset of a Chebyshev system necessarily a Chebyshev system? For each $n = 1, 2, 3, \ldots$ give an example of a Chebyshev system having n elements such that each subset is also a Chebyshev system.

13. Let U be an n-dimensional subspace in $C^{(n)}[a, b]$. Let $D = d/dt$. Thus $D^k(U) = \{u^{(k)} : u \in U\}$. Prove that if $\dim D^k(U) = n - k$ then $\dim D^i(U) = n - i$ for all $i \in \{0, 1, \ldots, k\}$. Under the same hypothesis, prove that $\Pi_{k-1} \subset U$. Prove that if $D^k(U)$ is an $(n-k)$-dimensional Haar space, then $D^i(U)$ is an $(n-i)$-dimensional Haar space for each $i \in \{0, 1, \ldots, k\}$.

14. In the discussion of Mairhuber's Theorem, sets homeomorphic to the letter Y were employed. These sets are called "triods." Prove that any set of disjoint triods in $\mathbb{R}^2$ must be countable. *References:* Moore [RLM] and Problem 6598 proposed by W. Rudin in *American Math. Monthly 98* (1991), 70–71.

15. Let U be an m-dimensional space of functions defined on a set X. Let $\mathcal{N}$ be a set of n points ("nodes") in X. Prove the equivalence of these properties:
  a. for any function f on $\mathcal{N}$ there is a unique element u in U such that $f(x) = u(x)$ on $\mathcal{N}$.
  b. $m = n$, and no element $u \in U$ (other than $u = 0$) vanishes on $\mathcal{N}$.

16. There exist discontinuous Haar spaces on all the spaces $\mathbb{R}^s$. Prove that if X is any infinite set of cardinality at most c (the cardinal number of $\mathbb{R}$), then for each n there is an n-dimensional Haar space of functions on X. (A suitable reference is Zielke's book [Zi].)

17. This chapter emphasizes linear interpolation problems. Investigate the question of whether the function $F(x) = a(1 + bx)^{-1}$ can be used to interpolate arbitrary data at two points.

18. (Continuation of Problem 17.) Investigate whether the function $F(x) = ae^{bx} + ce^{dx}$ can be used to interpolate arbitrary data at four points.

19. Find a set of three functions (defined on $\mathbb{R}^2$) such that interpolation of arbitrary data at any two points in $\mathbb{R}^2$ is possible by a linear combination of the three functions. Explain why this does not contradict Theorem 3. References: Wulbert [Wul1] and Shekhtman [Shek1].

20. Explain why point-evaulation functionals cannot be defined on $L^p[0, 1]$, $(1 \le p \le \infty)$.

21. Prove or disprove the converse of Theorem 6.

22. A concept stronger than that of a Chebyshev system is that of a Markov system. A sequence of continuous functions $u_0, u_1, u_2, \ldots$ defined on $\mathbb{R}$ is a **Markov system** if, for each n, $\{u_0, u_1, \ldots, u_n\}$ is a Chebyshev system. For example, the functions $u_n(x) = x^n$ form a Markov system. Notice that if the order of this sequence is changed, the Markov property is lost. If possible, make Markov systems from the examples 1–6 given on page 6.