Lecture Notes in Mathematics

A collection of informal reports and seminars

Edited by A. Dold, Heidelberg and B. Eckmann, Zürich

293

Ronald A. DeVore

The Approximation of Continuous Functions by Positive Linear Operators

Lecture Notes in Mathematics

A collection of informal reports and seminars
Edited by A. Dold, Heidelberg and B. Eckmann, Zürich

293

Ronald A. DeVore
Oakland University, Rochester, MI/USA

The Approximation of Continuous Functions by Positive Linear Operators

Springer-Verlag
Berlin · Heidelberg · New York 1972

AMS Subject Classifications (1970): 41-02, 41 A 10, 41 A 25, 41 A 35, 42 A 04, 42 A 08

ISBN 3-540-06038-3 Springer-Verlag Berlin · Heidelberg · New York
ISBN 0-387-06038-3 Springer-Verlag New York · Heidelberg · Berlin

Offsetdruck: Julius Beltz, Hemsbach/Bergstr.

PREFACE

These notes study linear methods of approximation which are given by a sequence (L_n) of positive linear operators. The essential ingredient being that of positivity. The main theme is to relate the smoothness of the function f being approximated with the rate of decrease of $||f - L_n(f)||$. This is accomplished in the usual setting of direct theorems, inverse theorems, saturation and approximation of classes of functions.

The fundamental ideas involved in direct estimates can be found in the pioneering book of P.P. Korovkin[4] and several of the more recent textbooks on approximation. However, most of the material appears in "book form" here for the first time. The main exception being the results on approximation by positive convolution operators, which have considerable overlap with the recent book of P.L. Butzer and R.J. Nessel[1].

I have written the notes at a level which presupposes a knowledge of the fundamental aspects of approximation theory, especially as pertains to the degree of approximation. Most of the necessary background material can be found in the now classic book of G.G. Lorentz[2]. For a good understanding of the material developed here for convolution operators, I expect that the reader will have to make several excursions into Butzer and Nessel[1].

The notes concentrate solely on spaces of continuous function (periodic and non-periodic) on a finite interval. I have not developed the theory for L_p - spaces since I know of little in these spaces that goes beyond what is already contained in Butzer and Nessel[1]. The only examples considered are those which fall comfortably in the grasps of the general theory. However, I believe that the reader will find that most of the better known examples are covered.

ACKNOWLEDGEMENT. I would like to thank Professor P.L. Butzer and his group at Aachen, most notably Professor R.J. Nessel and Dr.'s E. Görlich, E. Stark, K. Scherer, who were generous enough to comment on a rough first draft of these notes which I sent them in the spring of 1971. During the academic year 1971-1972, I had the opportunity to lecture on this material in a seminar on approximation theory at the University of Alberta. This effected considerably the final version of the notes, and I would like to thank all those who participated in this seminar, especially Professors A. Meir and A. Sharma who invited me to spend the year in Alberta. Finally, I would like to thank Mrs. Dorothy Farren for typing the first draft and Miss Olwyn Buckland for typing the final version of the manuscript.

Edmonton, June, 1972

CONTENTS

CHAPTER 1

PRELIMINARIES

1.1. Introduction. This first chapter will be devoted to stating, generally without proof, some fundamental results from approximation theory and Fourier analysis. Proofs can usually be found in the classical books on the appropriate subject. The material on approximation of functions can be found in the book of G.G. Lorentz$_2$, while the treatise of A. Zygmund$_1$ should be referred to for results from Fourier analysis. There is some middle ground material between approximation and Fourier analysis, which is best found in the work of P.L. Butzer and R.J. Nessel$_1$. We also quote some results on Chebyshev systems from S. Karlin and W.J. Studden$_1$, and orthogonal polynomials from G. Szegő$_1$. In the rare instance, that there is no reference for the exact result we need, a proof will be supplied.

In some cases, proofs of results stated in this chapter can actually be found in subsequent chapters. For example, Jackson's Theorem is proved in Chapter 2 and the Bernstein Inverse Theorem follows easily from the material in Chapter 8. However, we do not strive for this kind of completeness.

Most of our notation and terminology will be introduced in this chapter. Later in the text, when there may be some question as to notation, we will generally give reference back to the original usage.

1.2. Chebyshev Systems. $C[a,b]$ denotes the space of continuous functions on $[a,b]$ and $C^* = C^*[-\pi,\pi]$, the space of 2π-periodic and continuous functions on the line. Both of these spaces are equipped with the supremum norm over the appropriate interval. For example, for $f \in C[a,b]$

$$||f|| = \sup_{a \le x \le b} |f(x)|$$

We will use the notation $||\cdot||[a,b]$ to indicate that the supremum is taken over $[a,b]$ whenever it is necessary to make it clear which interval the norm is taken over.

A set of functions $\{u_0, u_1, \ldots, u_n\} \subseteq C[a,b]$ is called a Chebyshev system on $[a,b]$ if each function $u \in sp(u_0, \ldots, u_n)$ has at most n zeros unless it is identically zero on $[a,b]$. For the periodic case, $\{u_0, \ldots, u_n\} \subseteq C^*$ and any $u \in sp(u_0, \ldots, u_n)$ has at most n zeros on $[-\pi, \pi)$, unless it is identically zero. The space $\mathcal{U}_n = sp(u_0, \ldots, u_n)$ is called a <u>Chebyshev</u> <u>space</u>. $\mathcal{P}_n$, the space of algebraic polynomials of degree $\leq n$ is a Chebyshev space for any $[a,b]$ and $\mathcal{T}_n$ the space of trigonometric polynomials of degree $\leq n$ is a Chebyshev subspace of C^*.

A Chebyshev space on $[a,b]$ can also be characterized by the following interpolation property.

1.2.1. *If $a \leq x_0 < x_1 < \ldots < x_n \leq b$ and $(y_i)_{i=0}^n$ are real numbers then there is a unique $u \in sp(u_0, \ldots, u_n)$ such that*

$$u(x_i) = y_i \qquad i = 0,1, \ldots n.$$

If $\mathcal{U}_n$ is a Chebyshev subspace of $C[a,b]$ with each non-zero $u \in \mathcal{U}_n$ having n continuous derivatives on $[a,b]$ and at most n zeros on $[a,b]$ <u>counting multiplicity</u> then we say $\mathcal{U}_n$ is an <u>extended</u> <u>Chebyshev space</u>. If for each $k = 0,\ldots,n$, $sp(u_o,\ldots,u_k)$ is an extended Chebyshev system, then we say $\mathcal{U}_n$ is an <u>extended complete Chebyshev space.</u> For an extended complete Chebyshev space there is a canoncial basis $\{u_o,\ldots,u_n\}$ described by

1.2.2.
$$u_o(t) = w_o(t)$$

$$u_1(t) = w_o(t)\int_a^t w_1(x_1)dx_1$$

$$\vdots$$

$$u_n(t) = w_o(t) \int_a^t w_1(x_1) \int_a^{x_1} w_2(x_2) \dots \int_a^{x_{n-1}} w_n(x_n) dx_n \dots dx_1$$

where each w_i is a continuous strictly positive function on $[a,b]$, $w_i \varepsilon C^{(n-i)}[a,b]$.

Let $f \in C[a,b]$ (C^*) and U_n be a Chebyshev subspace of $C[a,b]$ (C^*). A function $u^* \in U_n$ is called a best approximation to f from U_n if

$$\| f-u^* \| = \inf_{u \in U_n} \| f-u \|$$

The following theorem of Chebyshev gives the existence, uniqueness and characterization of u^*.

<u>Theorem 1.1.</u> <u>Let</u> $f \varepsilon C[a,b]$ $(C^*[-\pi,\pi])$ <u>and</u> $\{u_0, \dots, u_n\}$ <u>be a Chebyshev system on</u> $[a,b]([-\pi,\pi])$. <u>Then the best approximation</u> u^* <u>to</u> f <u>exists and is unique. If</u> u <u>is any function in</u> $sp(u_o, \dots, u_n)$ <u>then</u> u <u>is the best approximation to</u> f <u>if and only if there exist points</u>

$a \le x_0 < x_1 < \dots < x_{n+1} \le b$ $(-\pi \le x_0 < x_1 < \dots < x_{n+1} < \pi)$ <u>such that</u>

$f-u$ <u>alternately takes on the values</u> $\pm \| f-u \|$ <u>at the points</u> x_i, $i = 0, \dots, n+1$.

<u>1.3. Classes of Functions.</u> There are various methods for measuring the smoothness of functions. The divided difference operator $\Delta_t(f,x)$ is defined for $t \ge 0$ by

$$\Delta_t(f,x) = f(x+t) - f(x)$$

If $f \in C[a,b]$, then the modulus of continuity of f is defined by

$$\omega(f,h) = \sup_{0 \le t \le h} \| \Delta_t(f,x) \| [a,b-t] \quad 0 \le h \le b-a$$

where the norm is taken with respect to the variable x. A similar definition holds for $f \in C^*$ except now the norm can be taken over the whole line. The modulus of continuity has the following fundamental properties:

1.3.1. $\omega(f,t)$ <u>is non-decreasing</u>, <u>continuous</u>, <u>and</u> $\omega(f,0) = 0$.

1.3.2. <u>If</u> $t_1, t_2 > 0$, <u>then</u> $\omega(f,t_1+t_2) \le \omega(f,t_1) + \omega(f,t_2)$.

1.3.3. <u>If</u> $\lambda > 0$ <u>and</u> $t > 0$, <u>then</u> $\omega(f,\lambda t) \le (\lambda+1)\omega(f,t)$.

1.3.4. <u>If</u> $t_1 < t_2$ <u>then</u> $t_2^{-1}\omega(f,t_2) \le 2t_1^{-1}\omega(f,t_1)$.

The conditions (1.3.1.) and (1.3.2.) characterize the modulus of continuity in the sense that if ω is any function which has these two properties, then ω is a modulus of continuity of a function in $C[a,b]$ (namely $\omega(t-a)$).

When ω is a modulus of continuity, we denote by $C_\omega(M)$ the set of all functions f for which

$$\omega(f,t) \le M\omega(t) \qquad 0 \le t \le b-a$$

The notation C_ω stands for the union of all the $C_\omega(M)$, $M > 0$. Sometimes, we will need to have a compact set and so we will restrict the norms of the functions by $\| f \| \le M_o$ and denote this new class by $C_\omega(M,M_o)$.

The most important case here is when $\omega(t) = t^\alpha$, with $0 < \alpha \le 1$, which are called the Lipschitz α classes. The common notation is $\text{Lip}\ \alpha = C_{t^\alpha}$. Similar notations, $\text{Lip}(\alpha,M)$ and $\text{Lip}(\alpha,M,M_o)$ are used for $C_{t^\alpha}(M)$ and $C_{t^\alpha}(M,M_o)$.

For higher smoothness, we let $W^{(r)}(\alpha,M,M_o, \ldots,M_r)$ denote the set of functions for which

$$\|f^{(i)}\| \leq M_i \qquad i = 0,1, \ldots,r$$

$$\omega(f^{(r)},t) \leq Mt^\alpha \qquad 0 \leq t \leq b-a$$

When we do not restrict the norms of the $f^{(i)}$ then we denote this class by $W^{(r)}(\alpha,M)$. The class $W^{(r)}(\alpha)$ is the union of all the classes $W^{(r)}(\alpha,M), M > 0$.

We denote by $L_\infty^{(r)}(M)$, the class of all functions f whose r^{th} derivative is $\leq M$ a.e.. $L_\infty^{(r)}$ is the union of all the $L_\infty^{(r)}(M)$. The classes $L_\infty^{(r)}$ and $W^{(r-1)}(1)$ are the same.

We can also get at higher smoothness by using higher order divided differences. The r^{th} order divided difference Δ_t^r is defined as the r-fold composition of Δ_t with itself. The r^{th} order modulus of continuity is then given by

$$\omega_r(f,h) = \sup_{0 \leq t \leq h} ||\Delta_t^r(f,x)|| \; [a,b-rt] \qquad 0 \leq h \leq \frac{b-a}{r}$$

Particularly important is the second order modulus of continuity since its behaviour cannot be characterized in terms of the first order modulus of continuity.

If $0 < \alpha \leq 2$, we define the class $Lip^*(\alpha,M)$ as the collection of all functions $f \in C[a,b]$ $(C^*[-\pi,\pi])$ such that

$$\omega_2(f,t) \leq 2M\, t^\alpha .$$

$\text{Lip}^*\alpha$ is the union of all $\text{Lip}^*(\alpha,M)$ over all M's. For $0 < \alpha < 1$, the classes $\text{Lip}^*\alpha$ and Lip α are the same. In fact, there are constants M_1, M_2 and M_3 depending on α, such that

$$\text{Lip}(\alpha,M_1) \subseteq \text{Lip}^*(\alpha,M_2) \subseteq \text{Lip}(\alpha,M_3) \qquad (1.3.5)$$

Similarly, for $1 < \alpha < 2$, the classes $\text{Lip}^*\alpha$ and $W^{(1)}(\alpha)$ are the same. Again a relation like (1.3.5.) holds. When $\alpha = 2$, we can say more

$$\text{Lip}^*(2,M) = W^{(1)}(1,2M) \qquad (1.3.6).$$

When $\alpha = 1$, the class Lip 1 is properly contained in $\text{Lip}^* 1$. The class $\text{Lip}^* 1$ is commonly referred to as the class of quasi-smooth functions, also called the Zygmund class. The class of smooth functions S consists of all functions f such that

$$\omega_2(f,t) = o(t) .$$

Given a function in C[a,b], it is sometimes necessary to define f outside [a,b] in such a way that as much smoothness as possible can be retained. It is always possible to extend f to get a new function g defined on the entire line (see Timan_1, p. 121) with

$$\omega_2(g,t) \leq 5\omega_2(f,t) \qquad 0 \leq t \leq \tfrac{1}{2}(b-a) \qquad (1.3.7.)$$

where now in the definition of $\omega_2(g,t)$ the norm is taken over all $-\infty < x < \infty$.

1.4. Fourier Series. If $f \in C^*$, the complex Fourier coefficients of f are defined by

$$\hat{f}(k) = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(t)e^{-ikt}dt \quad k = 0, \pm 1, \ldots$$

and the real Fourier coefficients by

$$a_k(f) = \frac{1}{\pi}\int_{-\pi}^{\pi} f(t)\cos kt\,dt\ , \ b_k(f) = \frac{1}{\pi}\int_{-\pi}^{\pi} f(t)\sin kt, \quad k = 0,1,2, \ldots .$$

Thus, the Fourier series of f is

$$f \sim \sum_{k-\infty}^{\infty} \hat{f}(k)e^{ikx} = \frac{a_o(f)}{2} + \sum_{1}^{\infty} a_k(f)\cos kx + b_k(f)\sin kx = \sum_{o}^{\infty} A_k(f,x)$$

Related to this is the conjugate series

$$\sum_{1}^{\infty}(-b_k(f)\cos kx + a_k(f)\sin kx) \qquad (1.4.1.)$$

If $f \in C^*$ then (1.4.1.) is the Fourier series of a function $g \in L_1[-\pi,\pi]$. We call g the conjugate function of f and write $g = \tilde{f}$.

The n^{th} partial sum of the Fourier series of f is denoted by

$$S_n(f,x) = \sum_{o}^{n} A_k(f,x) .$$

Similarly, for a Borel measure $d\mu$ defined on $[-\pi,\pi)$, we define the complex Fourier coefficients of $d\mu$ by

$$\check{\mu}(k) = \frac{1}{2\pi}\int_{-\pi}^{\pi} e^{-ikt}d\mu(t) , \quad k = 0, \pm 1, \ldots$$

In the case that $d\mu$ is an even measure, then the sine terms drop and

$$d\mu \sim \sum_{-\infty}^{\infty} \overset{\vee}{\mu}(k)e^{ikx} = \frac{\rho_o}{2} + \sum_{1}^{\infty} \rho_k \cos kx$$

where

$$\rho_k = \frac{1}{\pi} \int_{-\pi}^{\pi} \cos kt \; d\mu(t)$$

We call the ρ_k's the real Fourier coefficients of $d\mu$.

If f and g are in $C^*[-\pi,\pi]$ then the convolution of f with g is

$$(f*g)(x) = \frac{1}{\pi} \int_{-\pi}^{\pi} f(t)\; g(x-t)dt \qquad (1.4.2.)$$

The function $f*g$ is also in $C^*[-\pi,\pi]$ and has the Fourier series

$$f*g \sim 2 \sum_{-\infty}^{\infty} \hat{f}(k)\; \hat{g}(k)e^{ikx} \qquad (1.4.3.)$$

More generally we can define the convolution of f with a measure $d\mu$ by replacing $g(x-t)dt$ by $d\mu(x-t)$ in (1.4.2.) . Then $f*d\mu$ is in $C^*[-\pi,\pi]$ and (1.4.3.) also holds with $d\mu$ in place of g . In particular if $d\mu$ is even then

$$f*d\mu \sim \sum_{o}^{\infty} \rho_k \; A_k(f,x) \qquad (1.4.4.)$$

It is important to point out that the factor 2 that appears in (1.4.3.) arises since we used $\frac{1}{\pi}$ rather than $\frac{1}{2\pi}$ in (1.4.2.) . The only reason for doing this is that the formula (1.4.4.) comes out in a more convenient