APPROXIMATION OF FUNCTIONS

BY

G. G. LORENTZ

Professor of Mathematics
University of Texas
Austin

CHELSEA PUBLISHING COMPANY
NEW YORK, N. Y.

SECOND EDITION

Library of Congress Catalog Card No. 85- 72465
International Standard Book No. 0-8284-0322-8

Printed on 'long-life' acid-free paper

Printed in the United States of America

Preface to the Second Edition

My book, Approximation of Functions (Holt, Rinehart and Winston), has been out of print for a number of years. But a considerable demand for it still remains which, I hope, is in some measure justified. A Chinese translation of the book has appeared recently.

I have therefore been happy to accept an offer by Chelsea Publishing Company to publish a second edition of the book. The changes are minimal. Some errors have been corrected and a faulty result about rational approximation has been replaced by a beautiful theorem of my friend, the late Geza Freud. I am grateful to D. D. Stancu, Cluj, Rumania for many corrections.

G. G. Lorentz

Preface to the First Edition

My purpose has been to write an easily accessible book on the approximation of functions that is simple and without unnecessary details, and is also complete enough to include the main results of the theory, including some recent ones. In some cases (for example, Chapter 7, saturation classes), this has been made possible by restricting discussions to a few representative theorems from a field. The leitmotiv of the book is that of the degree of approximation. Only Chapter 2 (Chebyshev's theorem and related results), Chapter 3 (auxiliary results and notions), and Chapter 11 do not depend on this idea. The justification of the latter chapter lies in its coverage of some applications of entropy, which are significant because of Kolmogorov's theorem.

Except for a few sections, only functions of a real variable have been treated. The beautiful results of Runge, Bernstein, Walsh, Mergeljan, Dzjadyk, and others in the complex domain remain outside the scope of this book.

The book can be used as a textbook for a graduate or an advanced undergraduate course, or for self-study. Notes at the end of each chapter give information about important topics not treated in the main text. Problems serve as illustrations; some of them are not easy. It was felt that it is more useful to solve one difficult problem than several easy ones. The Bibliography is not all-inclusive. It has been limited to works that can be expected to be particularly useful for the reader, and to others of utmost historical interest.

I owe thanks to my friends, colleagues, and students, who have helped me with my work. Above all, I must thank Professor E. Hewitt, the editor of this series, Professors G. T. Cargo, G. F. Clements, H. S. Shapiro, Messrs. J. Case and J. T. Scheick, all students in my class on approximation theory, and the OSR of the U.S. Air Force, whose grant supported my work. I would be grateful for any suggestions that readers may send to me.

G. G. Lorentz

Contents

[1]

Possibility of Approximation

1. Basic Notions

The problem of linear approximation can be described in the following way: Let Φ be a set of functions, defined on a fixed space A. If a function f on A is given, can one find a linear combination $P = a_1\phi_1 + \cdots + a_n\phi_n$, $\phi_i \in \Phi$, which is close to the function f? Two preliminary problems arise: We must select the set Φ, and also decide how the deviation of P from f should be measured.

We begin with the second question. Let A be a compact Hausdorff topological space,[1] and let $C = C[A]$ be the set of all continuous real functions on A. The set C is a linear space over the reals: sums $f + g$ and products af with real a and $f, g \in C$ belong to C and satisfy the axioms of a linear space. The supremum

$$\| f \| = \sup_{x \in A} | f(x) | \tag{1}$$

is attained for all functions $f \in C$; thus, $\| f \| = \max_{x \in A} | f(x) |$. This supremum has the following properties, which define a norm on $C[A]$:

$$\| f \| \geqslant 0; \qquad \| f \| = 0, \qquad \text{if and only if} \qquad f = 0; \tag{2}$$

$$\| af \| = | a | \cdot \| f \|; \tag{3}$$

$$\| f + g \| \leqslant \| f \| + \| g \| . \tag{4}$$

Thus, C is a *normed linear space*. Similarly, the space of all continuous complex functions f on A with norm (1), which is also denoted by $C[A]$, is a normed linear space over the complex number field.

Unless something to the contrary is said, all our functions and scalars will be real. The convergence $f_n \to f$ in the norm of C, that is, $\| f_n - f \| \to 0$ as $n \to \infty$, is equivalent to the *uniform convergence* of $f_n(x)$ to $f(x)$ for all $x \in A$. It follows from this interpretation that the space C is complete: If f_n is a Cauchy sequence (that is, $\| f_n - f_m \| \to 0$ for $n, m \to \infty$), then f_n converges to some element f of C:

$$\| f_n - f \| \to 0. \tag{5}$$

[1] Without essential loss, the reader can substitute for this, here and in the remainder of the book, a compact metric space, or even a compact subset of a euclidean space.

Complete normed linear spaces are called *Banach spaces.* Many types of Banach spaces are important in the theory of approximation; for example, the spaces $L^p = L^p[a, b]$, $p \geqslant 1$, with the norm

$$\| f \| = \left\{ \int_a^b | f(x) |^p \, dx \right\}^{1/p}.$$

However, approximation in the spaces C remains both the most interesting and the most important special case (if one excludes the theory of orthogonal polynomials in the space L^2), and this book is devoted almost entirely to it.

The following definitions apply to any Banach space X with elements f and a distinguished subset Φ. We call f *approximable* by linear combinations

$$P = a_1\phi_1 + a_2\phi_2 + \cdots + a_n\phi_n , \qquad \phi_i \in \Phi, \qquad a_i \text{ real}, \tag{6}$$

if for each $\epsilon > 0$ there is a P with $\| f - P \| < \epsilon$. Often, Φ is a sequence: $\phi_1 , \phi_2 , \cdots, \phi_n , \cdots$. Then

$$E_n(f) = E_n^{\Phi}(f) = \inf_{a_1,\ldots,a_n} \| f - (a_1\phi_1 + \cdots + a_n\phi_n) \| \tag{7}$$

is the nth *degree of approximation* of f by the ϕ_i . If the infimum in (7) is attained for some P, this P is called *a linear combination of best approximation.* There is an exception to this notation: If the P are algebraic or trigonometric polynomials of a given degree, then n in (7) will refer to the degree of the polynomials rather than to the number of functions ϕ_i .

For the space $C[a, b]$ of continuous real functions on $[a, b]$, a natural sequence Φ is given by the powers $1, x, \cdots, x^n, \cdots$. In this case, the linear combinations of the first $n + 1$ functions are the *algebraic polynomials of degree n*: $P_n(x) = a_0 + a_1x + \cdots + a_nx^n$. In this definition we do *not* require that $a_n \neq 0$. A similar remark applies to some later definitions.

Another important compact set K is the additive group of real numbers $R = (-\infty, +\infty)$, taken modulo 2π; the distance $| x - x' |$ between $x, x' \in K$ is the minimal distance between the representations of x, x' in R. For obvious reasons, we can call K the *unit circle.* This K is a metric space with the distance $| x - x' |$. We shall follow the practice of identifying functions $f \in C^* = C[K]$ with the continuous 2π-periodic functions on R. For such functions, $\int_K f dx$ is the integral of f over any interval of R of length 2π. A function $f \in C^*$ does not necessarily have an indefinite integral F in C^*, for $\int_0^x f(t)\, dt$ is not necessarily periodic. Clearly, an $F \in C^*$ exists if and only if f has *mean-value zero*; that is, if $(1/2\pi) \int_K f dt = 0$. In this case, $F(x) = \text{const} + \int_0^x f dt$. Among these F there is *exactly one* with mean-value zero. Iterating this, we see that for each $p = 1, 2, \cdots$, a function $f \in C^*$ with mean-value zero has a family of pth indefinite integrals, which depends upon one additive constant. We shall call each of these integrals *a* pth *indefinite integral of f.*

A tool of approximation for functions $f \in C^*$ is the following set of *trigonometric polynomials*:

$$T_n(x) = \frac{a_0}{2} + a_1 \cos x + b_1 \sin x + \cdots + a_n \cos nx + b_n \sin nx. \quad (8)$$

The polynomial (8) is said to have *degree n*. It is *even* (or *odd*) if only cosines $\cos kx$, $k = 0, \cdots, n$ (or sines) appear in the representation. Simple trigonometric formulas imply that the product of two trigonometric polynomials of degrees n and m is equal to a trigonometric polynomial of degree $n + m$.

In analogy with these two cases, the linear combinations (6) will also sometimes be called polynomials.

In the present chapter, we shall discuss the possibility of approximation of functions by polynomials. In Chapter 2, properties of polynomials of best approximation will be treated.

2. Linear Operators

The following theorem (which will be proved in Sec. 3) is due to Weierstrass [103]:

THEOREM 1. Each continuous real function f on $[a, b]$ is approximable by algebraic polynomials: For each $\epsilon > 0$ there is a polynomial $P_n(x) = \sum_0^n a_k x^k$ with

$$|f(x) - P_n(x)| < \epsilon, \qquad a \leqslant x \leqslant b. \quad (1)$$

The most natural way to prove a theorem of this type is to give an explicit formula for the polynomial $P_n(x)$. In terms of f, this formula is usually linear.

A function $g = L(f)$ from a Banach space X into a Banach space Y is called a *linear operator* if it satisfies the conditions

$$L(f + f') = L(f) + L(f'); \qquad L(af) = aL(f) \quad (2)$$

for all $f, f' \in X$ and all real a. If Y is the real line R, L is called a *linear functional*. A linear operator L is called *bounded* if

$$\| L(f) \| \leqslant M \| f \|, \qquad f \in X \quad (3)$$

for some positive constant M. In this case, the infimum of all M for which (3) is true is still an admissible M. This minimal M is called the *norm* of L, and is denoted by $\| L \|$. From this definition it follows that

$$\| L \| = \sup_{f \neq 0} \frac{\| L(f) \|}{\| f \|} = \sup_{f \neq 0} \left\| L\left(\frac{f}{\| f \|}\right) \right\| = \sup_{\| f \| = 1} \| L(f) \|. \quad (4)$$

A bounded linear operator is continuous: From $f_n \to f$ (in the norm of X), it follows that $L(f_n) \to L(f)$ (in the norm of Y), since

$$\| L(f_n) - L(f) \| \leqslant \| L \| \cdot \| f_n - f \|.$$

In the case $X = C[A]$, we can distinguish *positive* elements of C: We write $f \geqslant 0$ if $f \in C$ and $f(x) \geqslant 0$ for all $x \in A$. An operator L that maps C into itself is called a *positive operator* if it transforms each positive element f into a positive element g. For a positive linear operator, we have $L(f) \leqslant L(g)$ if $f \leqslant g$ (that is, if $g - f \geqslant 0$); also, $|L(f)| \leqslant L(|f|)$, where $|f|$ is the function with values $|f(x)|$. An operator of this type is always bounded: From

$$|L(f)| \leqslant L(|f|) \leqslant L(\|f\|\, e) = \|f\|\, L(e)$$

(here e is the constant function $e(x) = 1$), it follows that

$$\|L(f)\| \leqslant \|L(e)\| \cdot \|f\|,$$

so that

$$\|L\| = \|L(e)\|.$$

For the value of $L(f)$ at $x \in A$, we write $L(f, x)$.

EXAMPLES

1. **Bernstein Polynomials.** For a function f defined on $[0, 1]$, let

$$B_n(f, x) = \sum_{k=0}^{n} \binom{n}{k} x^k(1-x)^{n-k} f\left(\frac{k}{n}\right), \qquad n = 0, 1, \cdots. \tag{5}$$

Clearly, this is a positive linear operator which maps $C[0, 1]$ into itself. By the binomial formula,

$$B_n(e, x) = \sum_{k=0}^{n} p_{nk}(x) = 1, \qquad p_{nk}(x) = \binom{n}{k} x^k(1-x)^{n-k}, \tag{6}$$

and hence $\| B_n \| = 1$, for $n = 0, 1, \cdots$.

2. **Fourier Series.** Let f be a 2π-periodic, integrable function. The coefficients of its *Fourier series*,

$$\frac{a_0}{2} + \sum_{k=1}^{\infty} (a_k \cos kx + b_k \sin kx), \tag{7}$$

are given by the formulas

$$a_k = \frac{1}{\pi} \int_{-\pi}^{\pi} f(t) \cos ktdt, \qquad b_k = \frac{1}{\pi} \int_{-\pi}^{\pi} f(t) \sin ktdt. \tag{8}$$

For example, the finite sum 1(8), augmented by zero terms, is the Fourier series of the trigonometric polynomial T_n. We consider the nth partial sum s_n of the series (7). A standard computation gives

$$s_n = s_n(f, x) = \frac{1}{\pi} \int_{-\pi}^{\pi} f(t) \frac{\sin (2n+1) \dfrac{t-x}{2}}{2 \sin \dfrac{t-x}{2}}\, dt. \tag{9}$$

To obtain this, one writes $s_n = u_0 + u_1 + \cdots + u_n$, where

$$u_0 = \frac{1}{\pi}\int_{-\pi}^{\pi} \tfrac{1}{2} f(t)\,dt, \qquad u_k = \frac{1}{\pi}\int_{-\pi}^{\pi} f(t)\cos k(t-x)\,dt, \quad k = 1, 2, \ldots,$$

and applies the formula

$$\frac{1}{2} + \cos\alpha + \cdots + \cos n\alpha = \frac{\sin(2n+1)(\alpha/2)}{2\sin(\alpha/2)} = D_n(\alpha), \tag{10}$$

which is obtained by multiplying both sides with $2\sin(\alpha/2)$. In the same way, by means of the formula

$$\sin\frac{\alpha}{2} + \sin\frac{3}{2}\alpha + \cdots + \sin(2n-1)\frac{\alpha}{2} = \frac{\sin^2(n\alpha/2)}{\sin(\alpha/2)}, \tag{11}$$

we obtain a representation of the arithmetic mean σ_n of the s_n:

$$\sigma_n = \sigma_n(f, x) = \frac{s_0 + \cdots + s_{n-1}}{n} = \frac{1}{2\pi n}\int_{-\pi}^{\pi} f(t)\left(\frac{\sin\dfrac{n(t-x)}{2}}{\sin\dfrac{t-x}{2}}\right)^2 dt. \tag{12}$$

Thus, $\sigma_n(f)$ is a sequence of positive linear operators, mapping C^* into itself. We have $\| \sigma_n \| = 1$, since $\sigma_n(e) = e$ (for the function $e(x) \equiv 1$, all $s_n(x) \equiv 1$). The operators $s_n(f), f \in C^*$ are also linear, but not positive. It follows from the definition of s_n and σ_n that for each f, both $s_n(f, x)$ and $\sigma_n(f, x)$ are trigonometric polynomials of degrees n and $n-1$, respectively. The norm of $s_n(f)$ is given by the following theorem.[2]

THEOREM 2 (Fejér). The norm $\| s_n \|$ of the operator $s_n(f)$ is equal to

$$A_n = \frac{1}{\pi}\int_{-\pi}^{\pi} \left| \frac{\sin(2n+1)(t/2)}{2\sin(t/2)} \right| dt = \frac{4}{\pi^2}\log n + O(1); \tag{13}$$

also, the norm of $s_n(f, x)$, for each fixed x, considered as a linear functional from C^* to R, is equal to (13).

Proof. Since $D_n(t)$ has period 2π,

$$| s_n(f, x) | \leqslant \|f\| \frac{1}{\pi}\int_{-\pi}^{\pi} | D_n(t-x) |\,dt = A_n \|f\|.$$

[2] We use the following notation: If u_n and $v_n > 0$ are functions of n, we write (a) $u_n = O(v_n)$ if $| u_n | \leqslant M v_n$ for some constant M; (b) $u_n = o(v_n)$ if $u_n/v_n \to 0$ as $n \to \infty$; (c) $u_n \sim v_n$ if $u_n/v_n \to 1$ as $n \to \infty$, and (d) $u_n \approx v_n$ if u_n/v_n is contained between two constants m, M, where $0 < m < M$.

This inequality shows that the functionals as well as the operator s_n have norms not exceeding A_n. To show that these norms are actually equal to A_n, it is sufficient to find, for given $x \in K$ and $\epsilon > 0$, a continuous function g for which $\| g \| = 1$ and

$$s_n(g, x) = \frac{1}{\pi} \int_{-\pi}^{\pi} g(t)\, D_n(t - x)\, dt > A_n - \epsilon. \tag{14}$$

For the function $g_0(t) = \operatorname{sign} D_n(t - x)$, we have

$$\frac{1}{\pi} \int_{-\pi}^{\pi} g_0(t)\, D_n(t - x)\, dt = A_n .$$

But g_0 is not continuous; it has jump discontinuities at the finitely many points t_ν of $[-\pi, \pi]$, where $D_n(t - x)$ changes sign. We surround each t_ν by a small interval $I_\nu = (t_\nu - \delta, t_\nu + \delta)$ and change g_0 on each I_ν so as to obtain a continuous function g, which has values between -1 and $+1$ everywhere and coincides with g_0 outside the I_ν. The difference between the integrals $\int_{-\pi}^{\pi} g D_n dt$ and $\int_{-\pi}^{\pi} g_0 D_n dt$ does not exceed $2 \int_E | D_n(x - t) |\, dt$, where E is the union of the intervals I_ν. If δ is sufficiently small, we have (14).

To obtain an asymptotic formula for A_n, we write $A_n = \pi^{-1} \int_0^{\pi} 2 \,| D_n(t) |\, dt$, since $D_n(t)$ is even. The function under the integral sign is equal to

$$\left| \cot \frac{t}{2} \sin nt + \cos nt \right| = \left| \frac{2}{t} \sin nt + \left(\cot \frac{t}{2} - \frac{2}{t}\right) \sin nt + \cos nt \right| .$$

Since $\cot u - u^{-1}$ is bounded in $(0, \pi/2)$, we have

$$A_n = \frac{2}{\pi} \int_0^{\pi} \frac{| \sin nt |}{t}\, dt + O(1).$$

The integral of $t^{-1} | \sin nt |$ over $(0, \pi/n)$ is bounded, since $| \sin nt | \leqslant nt$. Thus,

$$A_n = \frac{2}{\pi} \sum_{k=1}^{n-1} \int_{k\pi/n}^{(k+1)\pi/n} \frac{| \sin nt |}{t}\, dt + O(1)$$

$$= \frac{2}{\pi} \int_0^{\pi/n} \sin nt \sum_{k=1}^{n-1} \frac{1}{t + n^{-1} k\pi}\, dt + O(1).$$

Let $S(t)$ denote the last sum. For $0 \leqslant t \leqslant \pi/n$, $S(t)$ lies between

$$S(0) = n\pi^{-1} \left(1 + \frac{1}{2} + \cdots + \frac{1}{n-1}\right) \quad \text{and} \quad S(\pi/n) = S(0) + O(n).$$

If we use the facts that

$$1 + \frac{1}{2} + \cdots + \frac{1}{n-1} = \log n + O(1),$$

and that

$$\int_0^{\pi/n} \sin ntdt = \frac{2}{n},$$

we obtain:

$$A_n = \frac{4}{\pi^2} \log n + O(1). \qquad \blacksquare \qquad (15)$$

3. Approximation Theorems

The operator s_n is not suitable for the uniform approximation of arbitrary continuous functions, as we shall see in Chapter 6. However, both B_n and σ_n can be used for this purpose.

It has been observed by Bohman and Korovkin [7] that for a sequence L_n of positive linear operators, convergence often can be established quite simply by checking it for certain finite sets of functions f. Let A be a compact Hausdorff topological space (with at least two points). Let $f_1, \cdots, f_m$ be continuous real functions on A that have the following property[3]:

$$\left.\begin{array}{l}\text{there exist continuous real functions } a_i(y), y \in A, i = 1, \cdots, m \text{ such that} \\[1ex] \qquad\qquad P_y(x) = \sum_{i=1}^{m} a_i(y) f_i(x) \\[1ex] \text{is positive, and equal to zero if and only if } x = y.\end{array}\right\} \qquad (1)$$

THEOREM 3. If the functions $f_1, \cdots, f_m$ satisfy (1) and if L_n is a sequence of positive linear operators that map $C[A]$ into itself and satisfy

$$L_n(f_i, x) \to f_i(x) \quad \text{uniformly for} \quad x \in A, \qquad i = 1, \cdots, m, \qquad (2)$$

then

$$L_n(f, x) \to f(x) \quad \text{uniformly in } x \text{ for each} \quad f \in C[A]. \qquad (3)$$

Proof. We begin with some properties of the functions $P(x) = \sum_1^m a_i f_i(x)$. There exists a $\bar{P}$ with $\bar{P}(x) > 0$ for all $x \in A$: if $y_1 \neq y_2$ are two points of A, we can take $\bar{P} = P_{y_1} + P_{y_2}$. From (2) we have $L_n(P, x) \to P(x)$ uniformly in x for each P with constant coefficients. We also have

$$L_n(P_y, y) = \sum_{i=1}^{m} a_i(y) L_n(f_i, y) \to \sum_{i=1}^{m} a_i(y) f_i(y) = 0,$$

[3] The assumption that $a_i(y)$ are continuous could be omitted, but it simplifies the proof of Theorem 3.

and the convergence is uniform in y because the $a_i(y)$ are bounded. Finally, for some constant $M_0 > 0$, $\| L_n(e) \| \leqslant M_0$. This follows from

$$L_n(e, x) \leqslant aL_n(\bar{P}, x) \to a\bar{P}(x),$$

where $a > 0$ is taken so that $1 = e(x) \leqslant a\bar{P}(x)$, $x \in A$.

LEMMA. Let $f_y \in C[A]$, $y \in A$, be a family of functions for which $f_y(x)$ is a continuous function of the point $(x, y) \in A \times A$ and $f_y(y) = 0$ for all $y \in A$. Then

$$L_n(f_y, y) \to 0 \qquad \text{uniformly in } y. \tag{4}$$

Proof. Consider the "diagonal" set $B = \{(y, y)\}$ in $A \times A$ and some $\epsilon > 0$. Each point of B has a neighborhood U in $A \times A$ for which $|f_y(x)| < \epsilon$ if $(x, y) \in U$. The union G of all these U is an open set; its complement F is compact. Let

$$m = \min_{(x,y)\in F} P_y(x) > 0, \qquad M = \max_{(x,y)\in F} |f_y(x)|.$$

For all x, y we have

$$|f_y(x)| < \epsilon + \frac{M}{m} P_y(x). \tag{5}$$

In fact, $|f_y(x)|$ does not exceed the first term on the right if $(x, y) \in G$, nor the second term if $(x, y) \in F$. From (5) we derive

$$\begin{aligned} |L_n(f_y, y)| &\leqslant \epsilon L_n(e, y) + \frac{M}{m} L_n(P_y, y) \\ &\leqslant M_0\epsilon + \frac{M}{m} L_n(P_y, y) \leqslant (M_0 + 1)\,\epsilon, \end{aligned}$$

for all large n. ∎

Now the proof of the theorem can be completed easily. If $f \in C[A]$ is given, we put

$$f_y(x) = f(x) - \frac{f(y)}{\bar{P}(y)} \bar{P}(x).$$

By the lemma,

$$L_n(f, y) - \frac{f(y)}{\bar{P}(y)} L_n(\bar{P}, y) \to 0,$$

and since $L_n(\bar{P}, y) \to \bar{P}(y)$, we obtain (3). ∎

For example, if $A = [a, b]$, the system $f_1 = 1$, $f_2 = x$, $f_3 = x^2$ satisfies the condition (1). We can take

$$P_y(x) = (y - x)^2 = y^2 f_1 - 2y f_2 + f_3.$$

Theorem 3 allows us to check the convergence of certain operators with a minimum of computations. We begin by proving Theorem 1. The linear substitution $t = (x - a)/(b - a)$ reduces the interval $a \leqslant x \leqslant b$ to the interval $0 \leqslant t \leqslant 1$. Thus, Theorem 1 follows from

THEOREM 4. If the function f is continuous on [0, 1], then

$$\lim_{n\to\infty} B_n(f, x) = f(x) \qquad \text{uniformly for} \qquad 0 \leqslant x \leqslant 1. \tag{6}$$

Proof. For the polynomials p_{nk} of 2(6), we have

$$\sum_{k=0}^{n} k p_{nk}(x) = \sum_{k=1}^{n} k \binom{n}{k} x^k (1 - x)^{n-k}$$

$$= nx \sum_{l=0}^{n-1} \binom{n-1}{l} x^l (1 - x)^{n-1-l} = nx; \tag{7}$$

$$\sum_{k=0}^{n} k(k-1)\, p_{nk}(x) = n(n-1)\, x^2 \sum_{l=0}^{n-2} \binom{n-2}{l} x^l (1-x)^{n-2-l} = n(n-1)\, x^2,$$

so that

$$\sum_{k=0}^{n} k^2 p_{nk}(x) = n^2 x^2 + nx(1 - x). \tag{8}$$

Formulas 2(6), (7), and (8) mean that the functions $f_1 = 1, f_2 = x$, and $f_3 = x^2$ have as their Bernstein polynomials, respectively, 1, x, and $x^2 + n^{-1}x(1 - x)$, for $n \geqslant 2$. Conditions (1) and (2) hold, and (6) follows from Theorem 3. ∎

Useful in connection with Bernstein polynomials are the sums

$$T_{nr}(x) = \sum_{k=0}^{n} (k - nx)^r\, p_{nk}(x), \qquad r = 0, 1, \cdots. \tag{9}$$

Using formulas (7) and (8), we see that $T_{n0} = 1$, $T_{n1} = 0$, $T_{n2} = nx(1 - x)$. In order to compute the T_{nr} for $r \geqslant 2$, it is convenient to use the recurrence relation

$$T_{n,r+1} = x(1 - x)\,(T'_{nr} + nrT_{n,r-1}), \qquad r \geqslant 1, \tag{10}$$

which follows from (9) by differentiation, if one notices that

$$\frac{d}{dx} p_{nk}(x) = \frac{k - nx}{x(1 - x)}\, p_{nk}(x). \tag{11}$$

From (9) we obtain

$$T_{n0} = 1, \qquad T_{n1} = 0, \qquad T_{n2} = nX, \qquad T_{n3} = n(1 - 2x)\, X,$$
$$T_{n4} = 3n^2X^2 - 2nX^2 + nX(1 - 2x)^2, \qquad X = x(1 - x). \tag{12}$$