

ATILIM UNIVERSITY
DEPARTMENT of MATHEMATICS

MATH 411 Seminar Studies
Term Project Report

Submitted by: Fazil AMIRLI
Instructor: Prof. Dr. Mehmet TURAN

January 8, 2024

1 Introduction

In this paper, we will state **Weierstrass approximation theorem** and provide some historical information about the theorem. The theorem firstly stated by Karl Theodor Wilhelm Weierstrass, who was a German mathematician often cited as the "father of modern analysis" [1]. Despite leaving university without a degree, he studied mathematics and trained as a school teacher, eventually teaching mathematics, physics, botany, and gymnastics. He later received an honorary doctorate and became professor of mathematics in Berlin. Among many other contributions, Weierstrass formalized the definition of the continuity of a function and complex analysis, proved the intermediate value theorem and the Bolzano-Weierstrass theorem, and used the later to study the properties of continuous functions on closed bounded intervals. The theorem states that every continuous function defined on a closed interval $[a, b]$ can be uniformly approximated as closely as desired by a polynomial function. Because polynomials are among the simplest functions, and because computers can directly evaluate polynomials, this theorem has both practical and theoretical relevance, especially in polynomial interpolation. The original version of this result was established by Karl Weierstrass himself in 1885, while he was 70, using the Weierstrass transform.

1.1 Uniform Convergence

Definition 1.1 [5] We say that a sequence of functions $\{f_n\}$, $n = 1, 2, 3, \dots$, converges uniformly on E to a function f if for every $\epsilon > 0$ there is an integer N such that $n \geq N$ implies

$$|f_n(x) - f(x)| < \epsilon \quad (1)$$

for all $x \in E$.

Example 1.2 Let $0 < b < 1$ and consider the sequence of functions $\{f_n\}$ defined on $[0, b]$ by $f_n(x) = x^n$. Then, $\{f_n\}$ converges to 0 uniformly on $[0, b]$. To see this, take any integer $N > \frac{\ln \epsilon}{\ln b}$, then for every $n \geq N$, we have $b^n < \epsilon$, which implies $|x^n - 0| \leq b^n < \epsilon$.

Example 1.3 Now consider, the sequence of functions f_n defined on \mathbb{R} such that $f_n(x) = \frac{x^2 + nx}{n}$. This sequence of functions converges to x , but not uniformly. Since for each fixed x , we have

$$|f_n(x) - x| = \left| \frac{x^2}{n} + x - x \right| = \left| \frac{x^2}{n} \right| \rightarrow 0 \text{ as } n \rightarrow \infty.$$

We have $\{f_n(x)\} \rightarrow x$. Now, since we cannot find any n such that, $\frac{x^2}{n} < \epsilon$, for every $x \in \mathbb{R}$, we do not have a uniform convergence.

2 Bernstein Polynomials

Sergei Natanovich Bernstein, also known as Serguei Bernshtein, was a distinguished Russian and Soviet mathematician, whose contributions to the field of mathematics are highly valued. Born on March 5, 1880, in Odessa, Bernstein exhibited a strong inclination towards mathematics from an early age. He pursued higher education in mathematics at the University of Paris (Sorbonne), where he was influenced by the work of the renowned mathematician Henri Poincaré.

Upon returning to Russia, Bernstein embarked on a prolific academic career. He held several positions, including a professorship at the University of Kharkov and at the Institute of Mathematics of the Ukrainian Academy of Sciences. Bernstein's research interests were vast, encompassing areas such as probability, statistics, and function theory. He is best known for Bernstein polynomials, which are foundational in the field of approximation theory and were key in his proof of the Stone-Weierstrass theorem.

Bernstein made seminal contributions to the theory of partial differential equations and constructive function theory. His work in probability also laid the groundwork for the development of modern control and filtering theory.

Despite facing political challenges during the Soviet era, Bernstein maintained his academic pursuits and contributed significantly to the mathematical community until his passing on October 26, 1968, in Moscow. His legacy endures, with Bernstein polynomials continuing to be a fundamental tool in numerical analysis and approximation theory which will be discussed later in the report.

2.1 Bernstein Polynomials

Given a function f on $[0, 1]$, the Bernstein polynomial of f of degree n is

$$B_n(f; x) = \sum_{r=0}^n f\left(\frac{r}{n}\right) \binom{n}{r} x^r (1-x)^{n-r} \quad (2)$$

for each positive integer n . It is clear from (2) that for all $n \geq 1$,

$$B_n(f; 0) = f(0) \quad \text{and} \quad B_n(f; 1) = f(1), \quad (3)$$

so that a Bernstein polynomial for f interpolates f at both endpoints of the interval $[0, 1]$. It follows from the binomial expansion that

$$B_n(1; x) = \sum_{r=0}^n \binom{n}{r} x^r (1-x)^{n-r} = (x + (1-x))^n = 1 \quad (4)$$

so that the Bernstein polynomial for the constant function 1 is also 1. Since $\frac{r}{n} \binom{n}{r} = \binom{n-1}{r-1}$ for $1 \leq r \leq n$, the Bernstein polynomial for the function x is

$$B_n(x; x) = \sum_{r=0}^n \frac{r}{n} \binom{n}{r} x^r (1-x)^{n-r} = x \sum_{r=1}^n \binom{n-1}{r-1} x^{r-1} (1-x)^{n-r}. \quad (5)$$

Note that the term corresponding to $r = 0$ in the equation is zero. On putting $s = r - 1$ in the second summation, we obtain

$$B_n(x; x) = x \sum_{s=0}^{n-1} \binom{n-1}{s} x^s (1-x)^{n-1-s} = x B_{n-1}(1; x) = x.$$

Thus the Bernstein polynomial for the function x is also x .

Example 2.1

$$B_n(x^2; x) = x^2 + \frac{x(1-x)}{n}. \quad (6)$$

We have

$$\begin{aligned}
B_n(x^2; x) &= \sum_{k=0}^n \frac{k^2}{n^2} \binom{n}{k} x^k (1-x)^{n-k} \\
&= x \sum_{k=1}^n \frac{k}{n} \binom{n-1}{k-1} x^{k-1} (1-x)^{n-k} \\
&= \frac{x}{n} \sum_{k=1}^n k \binom{n-1}{k-1} x^{k-1} (1-x)^{n-k} \\
&= \frac{x}{n} \sum_{s=0}^{n-1} (s+1) \binom{n-1}{s} x^s (1-x)^{n-s-1} \\
&= \frac{x}{n} \sum_{s=0}^{n-1} s \binom{n-1}{s} x^s (1-x)^{n-s-1} + \frac{x}{n} \sum_{s=0}^{n-1} \binom{n-1}{s} x^s (1-x)^{n-s-1} \\
&= \frac{x}{n} \sum_{s=0}^{n-1} \frac{s}{n-1} (n-1) \binom{n-1}{s} x^s (1-x)^{n-s-1} + \frac{x}{n} \\
&= \frac{x^2(n-1)}{n} + \frac{x}{n} \\
&= x^2 + \frac{x(1-x)}{n}.
\end{aligned}$$

We call B_n the *Bernstein operator*; it maps a function f , defined on $[0, 1]$ to $B_n f$, where the function $B_n f$ evaluated at x is denoted by $B_n(f; x)$. The Bernstein operator is obviously linear, since it follows from (2) that

$$B_n(\lambda f + \mu g) = \lambda B_n f + \mu B_n g \quad (7)$$

for all functions f and g defined on $[0, 1]$, and all real λ and μ .

Definition 2.2 Let L denote a linear operator that maps a function f defined on $[a, b]$ to a function Lf defined on $[c, d]$. Then L is said to be a *monotone operator* or, equivalently, a *positive operator* if

$$f(x) \geq g(x), \quad x \in [a, b] \Rightarrow (Lf)(x) \geq (Lg)(x), \quad Lf(x) \in [c, d], \quad (8)$$

where we write $(Lf)(x)$ to denote the value of the function Lf at the point $x \in [a, b]$.

We can see from (2) that B_n is a monotone operator. It then follows from the monotonicity of B_n and (4) that

$$m \leq f(x) \leq M, x \in [0, 1] \Rightarrow m \leq B_n(f; x) \leq M, x \in [0, 1]. \quad (9)$$

In particular, if we choose $m = 0$ in (9), we obtain

$$f(x) \geq 0, x \in [0, 1] \Rightarrow B_n(f; x) \geq 0, x \in [0, 1]. \quad (10)$$

It follows from (4), (5), and the linear property (7) that

$$B_n(ax + b; x) = ax + b, \quad (11)$$

for all real a and b . We therefore say that the Bernstein operator *reproduces* linear polynomials. The Bernstein operator does not reproduce any polynomial of degree greater than one[4].

2.2 Proof of Weierstrass Approximation Theorem with Bernstein Polynomials

Theorem 2.3 (Weierstrass Approximation Theorem) *Let $f \in C[a, b]$. Given $\epsilon > 0$, one can find a polynomial $p_n(x)$ for which*

$$|f(x) - p_n(x)| \leq \epsilon, \quad a \leq x \leq b. \quad (12)$$

Proof. Notice that the linear transformation $y = (x - a)/(b - a)$ converts the interval $[a, b]$ into $[0, 1]$. For this reason, it will be enough to consider $[0, 1]$ instead of $[a, b]$.

We will give the proof of theorem presented in [4]. Begin with the identity

$$\left(\frac{r}{n} - x\right)^2 = \left(\frac{r}{n}\right)^2 - 2\left(\frac{r}{n}\right)x + x^2,$$

multiply each term by $\binom{n}{r}x^r(1-x)^{n-r}$, and sum from $r = 0$ to n , to give

$$\sum_{r=0}^n \left(\frac{r}{n} - x\right)^2 \binom{n}{r} x^r (1-x)^{n-r} = B_n(x^2; x) - 2xB_n(x; x) + x^2 B_n(1; x) = \frac{1}{n}x(1-x). \quad (13)$$

For any fixed $x \in [0, 1]$, let us estimate the sum of the polynomials $p_{n,r}(x) = \binom{n}{r}x^r(1-x)^{n-r}$ over all values of r for which $\frac{r}{n}$ is not close to x . To make this notion precise, we choose a number $\delta > 0$ and let S_δ denote the set of all values of r satisfying $|\frac{r}{n} - x| \geq \delta$. We now consider the sum of the polynomials $p_{n,r}(x)$ over all $r \in S_\delta$. Note that $|\frac{r}{n} - x| \geq \delta$ implies that

$$\frac{1}{\delta^2} \left(\frac{r}{n} - x\right)^2 \geq 1. \quad (14)$$

Then, using (14), we have

$$\sum_{r \in S_\delta} \binom{n}{r} x^r (1-x)^{n-r} \leq \frac{1}{\delta^2} \sum_{r \in S_\delta} \left(\frac{r}{n} - x\right)^2 \binom{n}{r} x^r (1-x)^{n-r}.$$

The latter sum is not greater than the sum of the same expression over all r , and using (13), we have

$$\frac{1}{\delta^2} \sum_{r=0}^n \left(\frac{r}{n} - x \right)^2 \binom{n}{r} x^r (1-x)^{n-r} = \frac{x(1-x)}{n\delta^2}.$$

Since $0 \leq x(1-x) \leq \frac{1}{4}$ on $[0, 1]$, we have

$$\sum_{r \in S_\delta} \binom{n}{r} x^r (1-x)^{n-r} \leq \frac{1}{4n\delta^2}. \quad (15)$$

Let us write

$$\sum_{r=0}^n = \sum_{r \in S_\delta} + \sum_{r \notin S_\delta},$$

where the latter sum is therefore over all r such that $|\frac{r}{n} - x| < \delta$. Having split the summation into these two parts, which depend on a choice of δ that we still have to make, we are now ready to estimate the difference between $f(x)$ and its *Bernstein polynomial*. Using (4), we have

$$f(x) - B_n(f; x) = \sum_{r=0}^n \left(f(x) - f\left(\frac{r}{n}\right) \right) \binom{n}{r} x^r (1-x)^{n-r},$$

and hence

$$\begin{aligned} f(x) - B_n(f; x) &= \sum_{r \in S_\delta} \left(f(x) - f\left(\frac{r}{n}\right) \right) \binom{n}{r} x^r (1-x)^{n-r} \\ &\quad + \sum_{r \notin S_\delta} \left(f(x) - f\left(\frac{r}{n}\right) \right) \binom{n}{r} x^r (1-x)^{n-r}. \end{aligned}$$

We thus obtain the inequality

$$|f(x) - B_n(f; x)| \leq \sum_{r \in S_\delta} \left| f(x) - f\left(\frac{r}{n}\right) \right| \binom{n}{r} x^r (1-x)^{n-r} + \sum_{r \notin S_\delta} \left| f(x) - f\left(\frac{r}{n}\right) \right| \binom{n}{r} x^r (1-x)^{n-r}.$$

Since $f \in C[0, 1]$, it is bounded on $[0, 1]$, and we have $|f(x)| \leq M$, for some $M > 0$. We can therefore write

$$\left| f(x) - f\left(\frac{r}{n}\right) \right| \leq 2M$$

for all r and all $x \in [0, 1]$, and so

$$\sum_{r \in S_\delta} \left| f(x) - f\left(\frac{r}{n}\right) \right| \binom{n}{r} x^r (1-x)^{n-r} \leq 2M \sum_{r \in S_\delta} \binom{n}{r} x^r (1-x)^{n-r}.$$

On using (15) we obtain

$$\sum_{r \in S_\delta} \left| f(x) - f\left(\frac{r}{n}\right) \right| \binom{n}{r} x^r (1-x)^{n-r} \leq \frac{M}{2n\delta^2} \quad (16)$$

Since f is continuous, it is also uniformly continuous, on $[0, 1]$. Thus, corresponding to any choice of $\epsilon > 0$ there is a number $\delta > 0$, depending on ϵ and f , such that

$$|x - x_0| < \delta \quad \Rightarrow \quad |f(x) - f(x_0)| < \frac{\epsilon}{2},$$

for all $x, x_0 \in [0, 1]$. Thus, for the sum over $r \notin S_\delta$, we have

$$\begin{aligned} \sum_{r \notin S_\delta} \left| f(x) - f\left(\frac{r}{n}\right) \right| \binom{n}{r} x^r (1-x)^{n-r} &< \frac{\epsilon}{2} \sum_{r \notin S_\delta} \binom{n}{r} x^r (1-x)^{n-r} \\ &< \frac{\epsilon}{2} \sum_{r=0}^n \binom{n}{r} x^r (1-x)^{n-r}, \end{aligned}$$

and hence, again using (4), we find that

$$\sum_{r \notin S_\delta} \left| f(x) - f\left(\frac{r}{n}\right) \right| \binom{n}{r} x^r (1-x)^{n-r} < \frac{\epsilon}{2}. \quad (17)$$

On combining (15) and (17), we obtain

$$|f(x) - B_n(f; x)| < \frac{M}{2n\delta^2} + \frac{\epsilon}{2}.$$

It follows from the line above that if we choose $N > M/(e\delta^2)$, then

$$|f(x) - B_n(f; x)| < \epsilon$$

for all $n \geq N$, and this completes the proof. ■

References

- [1] E.T. Bell. *Men of Mathematics*. A Fireside book. Touchstone, 1986. ISBN: 9780671628185. URL: <https://books.google.com.tr/books?id=BLFL3coT5i4C>.
- [2] S. N. Bernstein. “Démonstration du théorème de Weierstrass fondée sur le calcul des probabilités”. In: *Communications de la Société Mathématique de Kharkov* 13.1 (1912), pp. 1–2.
- [3] Gale. *Bernstein, Sergey Natanovich*. Encyclopedia.com. Available online: <https://www.encyclopedia.com/science/dictionaries-thesauruses-pictures-and-press-releases/bernstein-sergey-natanovich> (Accessed on January 7, 2024). 2023. (Visited on 01/07/2024).
- [4] G.M. Phillips. *Interpolation and Approximation by Polynomials*. CMS Books in Mathematics. Springer, 2003. ISBN: 9780387002156. URL: <https://books.google.com.tr/books?id=87vciTxMcF8C>.
- [5] W. Rudin. *Principles of Mathematical Analysis*. International series in pure and applied mathematics. McGraw-Hill, 1976. ISBN: 9780070856134. URL: <https://books.google.com.tr/books?id=kwqzPAAACAAJ>.