Advances in Nonlinear Variational Inequalities
ISSN: 1092-910X
Vol 28 No. 3s (2025)

Characterization and Construction of Best Approximation Using Classical Tchebychev Alternation Theorem and Remez Algorithm: A Revisit

Raghava M[a], Jyothirmai Joshi[b], Naveen Nekuri[c], Ch Ram Mohan[d]

[a]CVR College of Engineering, Raghava.m@cvr.ac.in,

[b]VNR Vignan Jyothi Institute of Engineering and Technology, jyothirmai_j@vnrvjiet.in,

[c]University of Hyderabad, naveencis@uohyd.ac.in

[d]CVR College of Engineering, chramohan@cvr.ac.in

Article History:

Received: *14-10-2024*

Revised: *29-11-2024*

Accepted: *10-12-2024*

Abstract:

In this work, we study the Classical Tchebychev Alternation Theorem in Approximation Theory and an important application of this theorem. Using this characterization of the nearest polynomial given in the alternation theorem, many algorithms have been derived to construct the nearest polynomial in practical applications from the given data. One of the most important and often-used algorithms in this respect is the Remez Algorithm. Modern data-driven AI essentially needs deeper insights into the characterization of algorithms so that it behaves robustly. As a review, we give a detailed proof of the Tchebychev Alternation Theorem for the space C[a, b]. Using a generalized version of the theorem, we derive the Remez Algorithm giving iterative steps toward the construction of the nearest polynomial and prove the convergence theorem for the iterative process specified in this algorithm. Additionally, we explore how the Tchebychev Alternation Theorem and its associated algorithms, particularly the Remez Algorithm, can be leveraged in image inpainting. By treating missing pixel values as regions requiring approximation within a specified domain, the nearest polynomial formulation provides a mathematically sound framework for reconstructing high-fidelity images. This method ensures robustness against non-stationary artifacts and enables the preservation of curvilinear structures, edges, and other critical features in digital image restoration tasks. This insight highlights the significance of classical approximation theory in solving modern computational challenges such as image inpainting.

Keywords: Banach Spaces, Tchebychev Theorem, Remez Algorithm, Characterization, De- La Vallee Poussin Lemma, Image Inpainting.

Introduction

This article initially starts with few necessary definitions and formal theorems that become drivers for the review and develops all the necessary algorithms to formalize the approximation procedure.

Definition 1: Let X be normed linear space and Y is a closed subset of X. Given $x \in X$ we say that $y \in Y$ is a nearest element to x from Y if

$$\| x \| = d(x, Y) = \inf\{\| x - z \|; z \in Y\}.$$

Definition 2: Let Y is a closed subset of X; we say Y is said to be a minimizing sequence for $x \in X$ if $\lim_{n\to\infty} \| x - y_n \| = d(x, Y)$.

Definition 3: A sequence y_N contained in Y is said to be a minimizing sequence for $x \in X$ if

$$\lim_{n\to\infty} \| x - y_n \| = d(x, Y).$$

Lemma 1: Let Y be a finite-dimensional subspace of X, then Y is approximal in X.

Proof: Let $x \in X$ and $y_n \subseteq Y$ be a minimizing sequence for x, then $\| x - y_n \| \leq d(x, y) + 1$ eventually, and y_n is bounded. Since $dim\{Y\} < \infty$, y_n has a convergent subsequence, say y_{n_k} that converges to $y \in X$. But Y is closed, being a finite-dimensional subspace, and so $y \in Y$. Since

$$\| x - y \| = \lim_{n\to\infty} \| x - y_{n_k} \| = d(x, Y)$$

y is the nearest element to x. Q.E.D.

We note that there are infinite-dimensional subspaces that are not proximal.

Interpolation Theorem:

Let $x_0, x_1, \ldots, x_n$ be $n + 1$ distinct (real or complex) numbers and $w_0, w_1, \ldots, w_n$ be $n + 1$ given (real or complex) numbers. Then there exists a unique polynomial P of degree $\leq n$ such that

$$P(x_i) = w_i \text{ for } 0 \leq i \leq n.$$

In approximation problems, the polynomials $1, x, x,^2 \ldots, x^n$ are replaced by other fixed functions $g_0, g_1, \ldots, g_n$ and the polynomials of the form $\sum_{i=0}^{n} c_i\, g_i$ are called generalized polynomials.

If $span\{g_1, g_2, \ldots, g_n\}$ denotes the subspace generated by g then the elements of $span\{g_1, g_2, \ldots, g_n\}$ are called the generalized polynomials.

Weierstrass Theorem [1]:

Let $f \in C[a, b]$ then given $\epsilon > 0$, $\exists$ a polynomial p such that

$$\| f - p \| = \sup_{x\in[a,b]} |(f - p)(x)| < \epsilon.$$

Practically for small values of ϵ it is challenging to construct a polynomial to satisfy the above condition.

Haar Conditions [2]:

An n-dimensional subspace Y of $C[a, b]$ is said to fulfill the Haar condition or Y is called a Haar subspace if it possesses the property that every function in Y, which is not identically zero vanishes at no more than $n - 1$ points of $[a, b]$. Equivalently, Y is a Haar subspace if for every set of n distinct points $\{x_i\}$ $1 \leq i \leq n$ in $[a, b]$ and every prescribed vector $y_1, y_2, \ldots, y_n$ $\exists$ a unique $g(x) \in Y$ such that $g(x_i) = y_i$; $1 \leq i \leq n$ i.e., the interpolation problem is uniquely solvable.

Example: For $n \geq 0$, P_n the class of all polynomials of degree $\leq n$ form a $n + 1$ dimensional Haar subspace of $C[a, b]$.

We have concerned with the approximation of a continuous function from a Haar subspace or from P_n in the supnorm.

De La Vallee Poussin Lemma [3]:

Let $f \in C[a,b]$ and p is a generalized polynomial $p = \sum_{i=0}^{n} c_i\, g_i$ where $(g_1, \dots, g_n)$ is a system of continuous functions which satisfy Haar's condition. Assume that $f - p$ takes alternatives positive and negative values at $n + 1$ consecutive points $x_i; 0 \le i \le n$ of $[a,b]$ then

$$E(f) \ge \min_{0 \le i \le n} |f(x_i - p(x_i)|$$

where $E(f) = \inf\{\| p - f \|; p \in sp(g_1, \dots, g_n)\}$.
Proof: Let us assume that the conclusion is false. Then $\exists$ a generalized polynomial p_0 such that

$$\| f - p_0 \| < \min_{0 \le i \le n} |f(x_i) - p(x_i)|. \qquad \text{Eq (1)}$$

Then the generalized polynomial $p_0 - p = (f - p) - (f - p_0)$ is alternatively positive and negative at $n + 1$ points $\{x_i\}; 0 \le i \le n$. This implies $p_0 - p$ vanishes at n distinct points, which by the Haar condition implies $p_0 - p = 0$. But this contradicts equation 1. Q.E.D.
Notations: Let P_n be the set of all polynomials of degree not exceeding n. For any $f \in C[a,b]$ we define

$$E_n(f) = inf \ \| f - p \|.$$

Definition 4: Let $f \in C[a,b]$ and $p \in P_n$. Let $g = p - f$, as this is continuous on $[a,b] \exists$ at least one point $\tilde{x} \in [a,b]$ such that

$$|g(\tilde{x})| = \| f - p \| . \qquad \text{Eq (2)}$$

Any $x \in [a,b]$ satisfying equation 2 is called a e-point. Define according to sign of $g(\tilde{x})$, x is a positive e-point if $g(\tilde{x}) = \| f - p \|$ and x is a negative e-point if $g(\tilde{x}) = -\| f - p \|$.
We now state and prove the Classical Tchebychev Alternation Theorem that characterizes nearest polynomials of best approximation from P_n to a given continuous function.

Tchebychev Alternation Theorem [4]:

For a polynomial $p_0(x) \in P_n$ be to be a polynomial of best approximation to a continuous function $f \in C[a,b]$ it is necessary and sufficient that there exists on $[a,b]$ at-least $m = n + 2$ e-points $x_0 < x_1 < \dots < x_{n+1}$ which are alternatively positive and negative e-points of the error function $f(x) - p_0(x)$.

Proof:

Necessary: Suppose the given polynomial $p_0(x)$ is the best approximation to the function $f(x)$ on $[a,b]$. Let $g = p_0(x) - f(x)$. Since g is continuous on $[a,b]$, it is uniformly continuous on $[a,b]$ for $\delta > 0$ such that

$$|g(x_1) - g(x_2)| < \frac{E_n}{2} \ \text{ if only } |x_1 - x_2| < \delta.$$

Let us break the interval $[a,b]$ by division points $a = a_0 < a_1, \dots < a_k = b$ into segments $\Delta_i, i = 1,2, \dots, k$. The length of each segment is less than δ, and choose those segments which contain at least one e-point. The segments containing at least one positive e-point is called a positive segment, and those containing at-least one negative point is called a negative segment. By condition 3, one segment cannot contain positive and negative points simultaneously. Otherwise, if x_1 is a positive point and x_2 is a negative point such that

$$|x_1 - x_2| < \delta$$
$$\Rightarrow |g(x_1) - g(x_2)| = \| (P_0 - f) \| + \| (P_0 - f) \|$$
$$= \| E_n + E_n \| < E_n/2.$$

This gives a contradiction.

Let M_1 is the set of all positive and negative segments. M_2 is the set formed from the remaining segments. Choose from the set M_1 a segment Δ_{m_0} having the least number and assume for the sake of definiteness that it contains a positive point. From the portion of the set M_1 situated to the right of Δ_{m_0}, choose a negative segment Δ_{m_1} having the least number. From the portion of the set M_1 situated to the right of Δ_{m_1}, choose a positive segment Δ_{m_2} with the least number. Since M_1 consists of a finite number of segments, continuing this process, we can separate a finite sequence of segments $\Delta_{m_0}, \Delta_{m_1}, \ldots, \Delta_{m_l}$, which are alternatively positive and negative segments. We claim to show that $l > n$.

On contrary take, $l \leq n$.

By choice of δ positive and negative segments that cannot be adjacent. This, along with the construction of segments $\Delta_{m_i}; 0 \leq i \leq l$, implies that the segments $\Delta_{m_{1_1}}, \ldots, \Delta_{m_{l_1}}$ belong to the set M_2.

Let $x^{(j)} \in \Delta_{j_1}; j = 1,2,\ldots,l$.

The method of construction implies that the portion of the set M_1 to the left of the point $x^{(1)}$ contains only positive segments and the portion of the set M_1 between $X^{(1)}$ and $x^{(2)}$ encloses only negative segments. Therefore, the polynomial of degree $l \leq n$ is given by

$$R(x) = (-1)^l \left(x - x^{(1)}\right) \ldots \left(x - x^{(l)}\right) \qquad Eq\ (3)$$

will be positive on any positive segment and negative on any negative segment.

That is, the signs of difference $p_0(x) - f(x)$ and polynomial $R(x)$ coincide on M_1. Therefore we choose $\lambda_1 > 0$ small enough to satisfy

$$\sup_{x \in M_1} |p_0 x - f(x)| - \lambda_1 |R(x)| < E_n(f). \qquad \text{Eq (4)}$$

Let us introduce the numbers

$$L = \max_{x \in M_2} |p_0(x) - f(x)|$$
$$N = \max_{x \in M_2} |R(x)|.$$

The set M_2 does not contain a single negative or a positive point. Hence $L < E_n(f)$. Therefore it is possible to choose a positive number λ_2 small enough to fulfill the inequality

$$L + \lambda_2 N < E_n(f).$$

Then on M_2 we shall have

$$\begin{aligned} |p_0(x) - f(x) - \lambda_2 R(x)| &\leq |p_0(x) - f(x)| + |\lambda R(x)| \\ &\leq L + \lambda N < E_n(f). \end{aligned} \qquad \text{Eq (5)}$$

choosing $\lambda = min\{\lambda_1, \lambda_2\}$,
and by combining equations 5 and 6, we get

$$|p_0(x) - f(x) - \lambda R(x)| < E_n(f)\ \forall x \in [a,b]. \qquad \text{Eq (6)}$$

But for $l \leq n$ we have $\left(p_0(x) - \lambda R(x)\right) \in P_n$, and the inequality 7 contradicts the fact that p_0 is the polynomial of the best approximation. Therefore $l \geq n + 1$.

Sufficient Condition:

Let the points $x_0 < x_1 <, \ldots, x_{n+1}$ be the alternatively positive and negative points of difference $p_0(x) - f(x)$. Then by setting $\alpha = sgn[p_0(x) - f(x)]$ we can write

$$(-1)^k \alpha[p_0(x_k) - f(x_k)] = \| p_0 - f \|; k = 0,1, \ldots, n.$$

Hence by de la Valle´e-Poussin lemma [1] $E_n(f) \geq \| p_0 - f \|$. By the definition of best approximation $E_n(f) = \inf_{p \in P_n}(\| p - f \|) \leq \| p_0 - f \|$. From the two inequalities, we can write $\| p_0 - f \| = E_n(f)$ $i.e.$, p_0 is the polynomial of the best approximation to $f(x) \in C[a,b]$.
Let Y be the n-dimensional Haar subspace of $C[X]$ where X is a non-empty closed subset of $[a,b]$ then the following generalized version of Tchebychev alternation hold, and we state the result without the proof and the theorem is needed in the later sections.

Generalized Version of Tchebychev Alternation Theorem [5]:

Let Y be the n-dimensional Haar subspace of $C[X]$ where X is the compact subset of $[a,b]$, then $h_0 \in Y$ is the nearest element to $f \in C[X]$ from Y iff $\exists\ n + 1$ points $x_0, x_1, \ldots, x_n$ in X such that

$$(f - h_0)(x_i) = \xi(-1)^i \| f - h_0 \|; 0 \leq i \leq n \text{ where } \xi \in \{+1, -1\}.$$

Corollary: Let Y be Haar Subspace and f be as in the above theorem, then $\exists$ a unique nearest element f from Y.

Unicity Theorem:

If functions $h_1, \ldots, h_n$ are continuous on $[a,b]$ and satisfy the Haar Condition, then the best approximation of each continous function by a generalized polynomial $\sum_{i=0}^{n} c_i h_i$ is unique. **Proof**: If possible, let p_0 and Q_0 be distinct generalized polynomials of best approximation to a given function f
$p_0(x_i) - f(x_i) = (-1)^i \| f - P_0 \|$ and $x_i \in X; 1 \leq i \leq n + 1$

$$Q_0(x_i) - f(x_i) = (-1)^i \| f - Q_0 \|.$$

By the triangular inequality, $1/2(P_0 + Q_0)$ is also the best approximation. By alternation theorem exists a system of points $x_1, x_2, \ldots, x_{n+1}$ in X such that

$$f(x_i)(1/2)[P_0 + Q_0](x_i) = (-1)^i \| f - P_0 \|.$$

Hence $(1/2)[f(x_i) - P_0(x_i)] + (1/2)[P_0(x_i) - Q_0(x_i)] = (-1)^i \parallel f - P_0 \parallel$.
Since $(1/2)[f(x_i) - P_0(x_i)] \leq [E_n(f)/2]$ and $(1/2)[f(x_i) - Q_0(x_i)] \leq [E_n(f)/2]$. ⇒
$(f - P_0)(x_i) = (f - Q_0)(x_i) = (-1)^i \parallel f - P \parallel$; $1 \leq i \leq n + 1$. But then P_0 and Q_0 are equal at $n + 1$ points $\{x_i\}$; $1 \leq i \leq n + 1$. This contradicts Haar's condition. Q.E.D.
However, in practical applications, it is difficult to determine the $n + 1$ consecutive points of X where the error function $f - p$ alternates. The method of Remez we describe in the next section helps to locate the points of alternation and to evaluate the $E_n(f)$ distance of f from Y using an iterative process.

The Remez Algorithm [6]:

The following notations and remarks are needed to gain self-contained nature and to demonstrate the Remez algorithm. Throughout the discussion, Y denotes an n-dimensional Haar space of $C[a, b]$. Let $\{x_i\}$; $1 \leq i \leq n + 1$ be $n + 1$ distinct points of the compact interval $[a, b]$.
Let $X = \{x_i; 1 \leq i \leq n + 1\}$ be $n + 1$ distinct points of compact interval $[a, b]$ and consider $n + 1$ dimensional space $C[X]$ of continuous and real-valued functions on X with supnorm. If $h \in C[a, b]$, we denote h the restriction of the function h on $C[X]$ to X. Also, set $Y = h$; $h \in Y$, then we have the following observations.
Property: Let Y be a Haar subspace of $C[a, b]$ of dimension n and $X = x_i;1 < i < n + 1$ are $n + 1$ distinct points of $[a, b]$. If $(h_1, \ldots, h_n)$ is a basis of Y, then $\left(h_1, \ldots, h_n\right)$ is a basis of Y .
Proof: Assume $\sum_{i=1}^{n} \alpha_i h_i = 0$ where α_i's are scalars. Then $\sum_{i=1}^{n} \alpha_i h_i \left(x_j\right) = 0$ for $1 \leq j \leq n + 1$. This implies $\sum_{i=1}^{n} \alpha_i h_i \in Y$ vanishes at $n + 1$ distinct points $X = x_i$; $i = 1ton + 1$. By Haar condition $\sum_{i=1}^{n} \alpha_i h_i = 0$ on $[a, b]$. Since $(h_1, \ldots, h_n)$ is a basis of Y we have $\alpha_i = 0$ for $i = 1ton + 1$. Hence $\left(h_1, \ldots, h_n\right)$ is the basis of Y .
Remark: Note $C[X]$ is of dimension $n + 1$. For $f \in C[a, b]$ and $f \notin Y$, $\{h_1, \ldots, h_n, f\}$ is a linearly independent set of $C[X]$ and so is a basis for $C[X]$.

Model Building:

Let $f \in C[a, b]$ such that $f \notin Y$ then $h_1, \ldots, h_n, f$ form a basis for $n + 1$-dimensional subspace of $C[X]$ and $d(f, Y) > 0$.
For each i; $1 \leq i \leq n + 1$, let δ_{x_i} represent the evaluation functional at x_i, that is $\delta_{x_i}(g) = g(x_i)$ $\forall g \in C[X]$. If $L \in C[X]^* and \parallel L \parallel < 1$, it is well known that there exist scalars λ_i; $1 \leq i \leq n + 1$ such that

$$L = \sum_{i=1}^{n+1} \lambda_i \, \delta_{x_i} \text{ and } \sum_{i=1}^{n} |\lambda_i| \leq 1. \qquad \text{Eq (7)}$$

We now recall the following consequence of the Hahn-Banach theorem that holds for general normed linear spaces.
Theorem 2: If X is a normed linear space, Y is a closed subspace of X with $x \in \{X \ \ Y\}$ then ∃ a linear functional $\Phi \in X^*$ such that $\Phi = 0$ on Y, $\parallel \Phi \parallel = 1$ and $\Phi(x) = d(x, Y)$.
Now for any $g \in C[X]$, we have

$L(g) = \sum_{i=1}^{n+1} \lambda_i \, g(x_i)$. Since $(h_1, \dots, h_n, f)$ is a basis of $C[X]$ ∃, a unique linear functional say $L \in C[X^*]$ such that

$$L(h_i) = 0; 1 \leq i \leq n \text{ and } L(f) = d(f, Y) > 0. \qquad Eq\ (8)$$

By the theorem 2, $\| L \| = 1$.
Hence we get a unique set of scalars $\lambda_i; 1 \leq i \leq n+1$ such that

$$\sum_{i=1}^{n+1} |\lambda_i| = 1 \text{ and } L = \sum_{i=1}^{n+1} \lambda_i \, \delta_{x_i} \qquad Eq\ (9)$$

then L satisfies equation 8.
Note that

$$\lambda_j = 0$$

for any $j; 1 \leq j \leq n+1$, then

$$\sum_{i=1, i\neq j}^{n+1} \lambda_i \, h_i(x_i) = 0; 1 \leq k \leq n$$

that is $\{\lambda_i; 1 \leq i \leq n+1, i \neq j\}$ is a unique solution to

$$\Rightarrow \begin{bmatrix} h_1(x_1) & \dots & h_1(x_{(j-1)}) & h_1(x_{(j+1)}) & \dots & h_1(x_{(n+1)}) \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ h_n(x_1) & \dots & h_n(x_{(j-1)}) & h_n(x_{(j+1)}) & \dots & h_n(x_{(n+1)}) \end{bmatrix} \begin{bmatrix} \lambda_1 \\ \lambda_2 \\ \vdots \\ \lambda_{j-1} \\ \lambda_{j+1} \\ \vdots \\ \lambda_{n+1} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ \vdots \\ \vdots \\ 0 \\ 0 \end{bmatrix}$$

Here h_i belongs to Haar subspace Y of $C[a, b]$.
As $(h_1, \dots, h_n)$satisfy Haar condition we must have $\{\lambda_i\}_{1 \leq i \leq n+1} = 0$.
Thus if $L(f) = d(f, Y) > 0$ then λ_i 's are nonzero $\forall i; 1 \leq i \leq n+1$.
Otherwise $L \equiv 0$ on $C[x]$ and $\lambda_i = 0, \forall i; 1 \leq i \leq n+1$.
Let $f \in C[a, b]$ and assume that $d(f, Y) > 0$.
Consider the set of equations given by

$$h(x_i) + (-1)^i \eta = f(x_i), \forall i; 1 \leq i \leq n+1 \qquad Eq\ (10)$$

where $h \in$ Haar subspace Y of C[a, b]. Since Y is a Haar subspace, we can uniquely solve for h and η. That is, we can get suitable coefficients $\alpha_1, \dots, \alpha_n$ and η such that $h = \sum_{j=1}^{n} \alpha_j \, h_j$ satisfies equation 11.
By the Generalized Tchebychev Alternation Theorem applied to $C[X]$ where $X = (x_i)_{i=1}^{n+1}$ and the Haar subspace Y, h is the nearest element to f from Y .
Let $\Phi = f - h$. Since $L(h) = 0$, we have

$$L(\Phi) = L(f) = |\eta| = d(f, h)$$

Thus if $d(f, h) > 0$ then

$$|\eta| = L(f) = L(\Phi) = \sum_{i=1}^{n+1} \lambda_i \, \Phi(x_i) = \sum_{i=0}^{n+1} \lambda_i \, \eta(-1)^i.$$

This implies $|\sum_{i=0}^{n+1} \lambda_i (-1)^i| = 1 = \sum_{i=0}^{n+1} |\lambda_i|$ which in-turn implies the λ_i's alternate inside. Recall that $\lambda \neq 0$ for $1 \leq i \leq n+1$ if $f \notin Y$ that is $sgn(\lambda_i) = (-1)^i \xi$ where $\xi \in \{-1, 1\}$.

Numerical Method for Linear Tchebychev Approximation

Let Y be a linear Haar Subspace of $C[a, b]$ of dimension n. Then the best approximation given function $f \in C[a, b]$ with respect to Y can constructed iteratively. In the following we consider a sequence (X^K) in R^{n+1} where $X^k = ({x_1}^k, \ldots, {x^k}_{n+1}), k \geq 0$ and ${x_1}^k < \ldots < {x^k}_{n+1}$ are distinct points of $[a, b]$. In the iterative process we replace X by X^k to obtain unique $h_k \in Y$ satisfying

$$h({x_i}^k) + (-1)^k \eta_k = f({x_i}^k), 1 \leq i \leq n+1.$$

As no confusion is likely to arise, we denote by f the restriction of f to x^k for any $f \in C[a, b]$ and $Y_k = (f ; f \in Y)$. Then we have $|\eta_k| = |d(f, Y_k)|$.
Also given$f \in C[a, b]$ $\exists$ a unique linear functional L^k in $(C[X^k])^*$ such that

$$L_k = \sum_{i=0}^{n+1} {\lambda_i}^k \, \delta_{{x_i}^k} \qquad Eq\ (11)$$

satisfying the following four conditions:

i.$L \equiv 0$ on Y_k

ii.$L_k(f) = d(f, Y_k)$

iii. $\| L_k \| = \sum_{i=0}^{n+1} |{\lambda_i}^k| = 1$ and

iv. ${\lambda_k}^k$ alternate in sign.

Finally, we set $\Phi_k = f - h_k$ for $k \geq 0$.
We now proceed to describe how the sequence $(X^k), k \geq 0$ of $n+1$ tuples are selected or prescribed so that the convergence theorem for the resulting algorithm can be worked out.

Construction of sequence:

To begin with we select any $n+1$ distinct points $\{{x_i}^0\}; 1 < i < n+1$ with $a \leq {x_1}^0 < {x_2}^0 \ldots < {x_{n+1}}^0 \leq b$ we construct $h^0 \in Y$ and linear functional $L_0 \in C[x^0]^*$ satisfying * and conditions 1 to 5, respectively. For $k = 0$ we have $|\eta_0| = d(f, Y_0) = L_0(f)$. If $\| f - h^0 \| = |\eta_0|$, then clearly, h^0 is the nearest element to f from Y, and X^0 is the required set of points of an alternation.
We assume that

$$\| f - h^0 \| = \| \Phi_0 \| > |\eta_0|.$$

We then construct a new set X^1 from X^0, which consists of $n+1$ points but for which the corresponding linear function $L_1(f)$ has a larger magnitude than $L_0(f)$. Let $x_i^1; 1 < i < n+1$ be $n+1$ distinct points of $[a,b]$ such that the following hold

1. The error function $\Phi_0 = f - h_0$ satisfies

$$\left|\Phi_0(x_i^1)\right| \geq L_0(f); i = 1, \dots, n+1. \qquad \text{Eq (12)}$$

2. For at least one integer $i = i_0$

$$\left|\Phi_0(x_i^1)\right| > L_0(f) \qquad Eq\ (13)$$

3. For $\xi \in \{-1, +1\}$

$$sgn\left(\Phi_0(x_i^1)\right) = \xi sgn\left(\Phi_0(x_i^0)\right); i = 1, \dots, n+1. \qquad \text{Eq (14)}$$

Now if $L_1(h) = \sum_{i=1}^{n+1} \lambda_i^1 h_i(x_i^1)$ is a linear functional given by 12 corresponding to the set M_1 then $L_1(f) = |\eta_1|$. Note that since 13 holds $|\eta_1| = L_1(f) > 0$.
Hence $\lambda_i^1 \neq 0$ for $i = 1,2, \dots, n+1$.
Then $L_1(f) = \sum_{i=1}^{n+1} \lambda_i^1 \Phi_i(x_i^1) = -\xi \sum_{i=1}^{n+1} \lambda_i^1 |\Phi_i(x_i^1)| sgn\Phi_0(x_i^0)$; $\xi \in$ {-1, +1}.
As (λ_i^1)'s alternate in sign

$$L_1(f) = \pm \sum_{i=1}^{n+1} | \lambda_i^1 |\Phi_0(x_i^1)| \xi sgn(L_0(f)$$

and hence

$$L_1(f) = L_0(f) + \sum_{i=1}^{n+1} | \lambda_i^1 | \{\Phi_0(x_i^1) - |L_0(f)|\}$$

$\Rightarrow |L_1(f)| > |L_0(f)|$ that is $|\eta_1| > |\eta_0|$.
Similarly, now starting with the set X^1 $\exists$ a function $h_1(x) \in Y$, which is the best approximation of $f(x)$ on the set X^1.
Again if $\| f - h^1 \| = |\eta_1|$, the algorithm ends at this step. Otherwise, we start with the set X^1 and go on to construct X^2. Finally, the algorithm stops after a finite number of steps or yields a sequence of steps X^k and the corresponding linear functional L_K with the property that the quantities $L_k(f)$ are monotonically increasing.

Construction Methods:

We now present two methods to prescribe the set X^{k+1}, assuming X^i i=0 to K are constructed.

Simultaneous Exchange Method [7]:

We know that $\Phi_0(x)$ possess at least n zeros Z_i^0 in $[a,b]$ and

$$x_i^0 < Z_i^0 < x_{i+1}^0; 1 \leq i \leq n$$

Set $Z_0{}^0 = a$, $Z_{n+1}^0 = b$.
Take the interval $J_i = [Z_i{}^0, Z_{i+1}{}^0]; 0 \leq i \leq n$
We determine a point $x_{i+1}{}^1$ such that

$$\Phi_0(x_{i+1}{}^1) > \Phi_0(x_i{}^1) \text{for} x \in J_i \text{if} sgn(\Phi_0(x_{i+1}{}^0) = 1 \quad \text{and}$$
$$\Phi_0(x_{i+1}{}^1) \leq \Phi_0(x_i{}^1) \text{for} x \in J_i \text{if} sgn(\Phi_0(x_{i+1}{}^0) = -1.$$

Single Exchange Method [7]:

In this method one of the points of X^k is replaced by a new point which satisfies Eq 7. To make sure that Eq 8 a special rune in exchange, let ξ be a point such that $|\Phi_k(\xi)| = \rho > |L_k(f)|$. Then the substitution is given by the following table.

Case	Condition	ξ Replaces
$a \leq \xi \leq x_1{}^k$	$sgn\Phi_k(\xi) = sgn\Phi_k x_1{}^k$	$x_1{}^k$
$a \leq \xi \leq x_1{}^k$	$sgn\Phi_k(\xi) = -= -sgn\Phi_k x_1{}^k$	$x^k{}_{n+1}$
$1 \leq i \leq n$		
$x_i{}^k \leq \xi \leq x_{i+1}{}^k$	$sgn\Phi_k(\xi) = sgn\Phi_k x_i{}^k$	$x_i{}^k$
$x_i{}^k \leq \xi \leq x_{i+1}{}^k$	$sgn\Phi_k(\xi) = -sgn\Phi_k x_i{}^k$	$x_i{}^k$
$x^k{}_{n+1} \leq \xi \leq b$	$sgn\Phi_k(\xi) = sgn\Phi_k x_{n+1}{}^k$	$x^k{}_{n+1}$
$x^k{}_{n+1} \leq \xi \leq b$	$sgn\Phi_k(\xi) = -sgn\Phi_k x_{n+1}{}^k$	$x_1{}^k$

This rule is well defined except in case $L_k(f) = 0$ and in which case any point can be replaced byξ. By renumbering, according to the magnitude, we can get the set X^{k+1}. We can now prove the convergence theorem of the Remez algorithm.

Convergence Theorem:

The successive functions that are a sequence of functions $h_k \in Y$ generated in the previous algorithm converges to the best approximation of f from Y in the space of $C[a, b]$.
Proof: We have $X^k = \{x_1{}^k, x_2{}^k, \ldots, x_{n+1}{}^k\}$ and $a < x_1{}^k, x_2{}^k, \ldots, x_{n+1}{}^k < b$.
We first show that $\exists \epsilon > 0$ such that

$$|x^k{}_{i+1} - x^k{}_i| \geq \epsilon \quad \forall \ 1 \leq i \leq n, \text{and } k \geq 0. \qquad \text{Eq (15)}$$

Suppose not, then by taking a subsequence if necessary; we can assume that $\exists$ an index $1 \leq i \leq n + 1$ such that

$$\lim_{k \to \infty} |x^k{}_{i+1} - x^k{}_i| = 0$$

Let $x^-{}_1, \ldots, x^-{}_{n+1}$ be limit points of $(x^-{}_1, \ldots, x^-{}_{n+1})_k > 0$. Let $h \in Y$ be the function that interpolates to f at $x^-, \ldots, x^-{}_{n+1}$. We know that $(f - h)$ is uniformly continuous on $[a, b]$. Then select δ such that

$$w(\delta) = \max_{|x-f|<\delta} |(f - h)(x) - (f - h)(y)| < |\eta_1|$$

Then choose $k > 1$ large such that

$$|x_i{}^k - x^-{}_i| < \delta \text{ for } 1 \le i \le n+1.$$

and

$$\eta_k \le \max_{1\le i\le n+1} |(f-h)(x_i{}^k)|$$

Since $h \in Y$, $|\eta_k| = d(f^-, Y^-{}_k) \le \max_{1\le i\le n+1} |(f-h)(x_i{}^k)|$.
And this implies

$$\eta_k \le \max_{1\le i\le n+1} |(f-h)(x_i{}^k)|$$

$$\le \max|(f-h)(x_i{}^-)| + \max|(f-h)(x_i{}^k) - (f-h)(x_i{}^-)|$$

$$= w(\delta) < |\eta_1| (\text{ as } h \text{ interpolates } f \text{ at } x_i{}^-) \qquad \text{Eq (16)}$$

This is a contradiction as $|\eta_1|, |\eta_k|$ for $k > 1$.
Fix j, $1 \le j \le n+1$. Using the fact that Y is a Haar subspace, we can get a unique $g_k \in Y$ such that

$$g_k(x^k{}_i) = f(x^k{}_i) \text{for} 1 \le i \le n+1, i \ne j$$

We know that since 16 holds $\exists M > 0$ such that $\| g_k \| \le M, \forall k > 0$.
For any positive integer k let $g_k = \sum_{i=1}^n \beta_i{}^k h_i$ for suitable scalars $\beta_i{}^k$. These scalars $\beta_i{}^k$ are uniquely determined as the solution of the matrix equation.

$$\begin{bmatrix} h_1(x_1{}^k) & \dots & h_1(x^k{}_{(j-1)}) & h_1(x^k{}_{(j+1)}) & \dots & h_1(x^k{}_{(n+1)}) \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ h_n(x_1{}^k) & \dots & h_n(x^k{}_{(j-1)}) & h_n(x^k{}_{(j+1)}) & \dots & h_n(x^k{}_{(n+1)}) \end{bmatrix} \begin{bmatrix} \beta_1{}^k \\ \vdots \\ \beta^k{}_{j-1} \\ \beta^k{}_{j+1} \\ \vdots \\ \beta^k{}_{n+1} \end{bmatrix} = \begin{bmatrix} f(x^k{}_1) \\ \vdots \\ f(x^k{}_{j-1}) \\ f(x^k{}_{j+1}) \\ \vdots \\ f(x^k{}_{n+1}) \end{bmatrix}.$$

Now observe that
$A = \{y_1, \dots, y_n\}$ with $a \le y_1 < y_2, \dots, y_n \le b$ and $|y_i - y_{i+1}| \ge \epsilon$ is a closed subset of the compact set $[a,b]$ x $[a,b]$ x …x $[a,b]$.

Define $D(y_1, \dots, y_n) = \begin{bmatrix} h_1(y_1) & \dots & h_1(y_n) \\ \vdots & \vdots & \vdots \\ h_n(y_1) & \dots & h_n(y_n) \end{bmatrix}$.

Then $D(y_1, \dots, y_n) \ne 0$ for each $(y_1, \dots, y_n) \in A$ as $h_1, \dots, h_n$ satisfy the Haar condition. Also, D is a continuous map from the compact set A into R.
Hence

$$\inf\{D(y_1, \dots, y_n)\} > 0. \qquad \text{Eq (17)}$$

Now

$$\| h_i \| = \sup_{t\in[a,b]} \| h_i(t) \| < \infty$$

Using Crammer's rule [8] and , it is now easy to see that the set of scalars $\{|\beta_i{}^k|; 1 < i \le n \text{ and } k > 0\}$ is bounded above and the upper bound is independent of the points $\{x_i\}^k$ and $k > 0; 1 < i \le n$.

This clearly implies that $\exists M > 0$ such that $\| g_k \| \leq M, \forall k \geq 0$.
Now

$$|\eta_1| \leq |\eta_k|$$
$$= L_k(f) = L_k(f - g_k)$$
$$= \sum_{i=1}^{n+1} \lambda_i{}^k (f - g_k)(\{x_i\}^k)$$
$$= \lambda_j{}^k |f - g_k(\{x_i\}^k)|$$
$$\leq \lambda_i{}^k (\| f \| + M)$$

Thus $\lambda_j{}^k \geq \frac{|\eta_1|}{\{\|f\|+M\}} \forall k > 0$ and $j \leq n+1$
Convergence: [8]

By equation 10, the absolute values of $\eta_k's$ in the iteration process have a positive lower bound except for some case $L_0(f)$. By construction $\exists x \in X_{k+1}$ such that $\| \Phi_{k(x)} \| = \| \Phi_k \|$ and by Equation 14 we prove that $\| \Phi_k \| > L_k(f)$.
If this process does not terminate, then there exists an integer i_0 such that

$$|L_{k+1}(f) - L_k(f)| \geq |\lambda_{i_0}{}^{k+1}| \{\| \Phi_k \| - |L_k(f)|\} \qquad Eq\ (18)$$

The sequence $|L_k(f)|$ is monotone increasing and bounded above as $|L_k(f)| \leq d(f, y)$.
Now $|L_{k+1}(f) - L_k(f)| = \epsilon_k \geq 0$ where $\lim \epsilon_k = 0$.
By the equation 18 for suitable value of $\epsilon_{k,}$ $d(f, y) \leq \| \Phi_k \| < d(f, y) + \epsilon_k$ and

$$d(f, y) \leq \| f - h_k \| \leq d(f, y) + \epsilon_k \text{ for } h_k \in Y. \qquad \text{Eq (19)}$$

Consider $h \in Y$ satisfying

$$\| f - h \| \leq d(f, Y) + \epsilon. \qquad \text{Eq (20)}$$

Assume h_1, h_2 are two real functions in Y satisfying 16, then $\| h_1 - h_2 \| \leq \delta(\epsilon)$ where $\lim_{\epsilon \to 0} \delta(\epsilon) = 0$. This is because of the uniqueness of the best approximation.
If the assumption is not true, then there exist two sequences $\{h_1{}^m\}, \{h_2{}^m\} \in Y$ such that
$\lim \| f - h_1{}^m \| = d(f, Y)$ and $\lim \| f - h_2{}^m \| = d(f, Y)$ but $\| h_1{}^m - h_2{}^m \| > k > 0$ $m = 1,2, \ldots$.
Because Y is a finite dimension Bolzanov Weierstrass [7] theorem implies that we can get convergence subsequences from each of these sequences.
Let h_1 and h_2 be the limits they satisfy the relation $\| f - h_1 \| = \| f - h_2 \| = d(f, Y)$ and $h_1 \neq h_2$.
This contradicts the uniqueness of best approximation. The convergence theorem is proved.

Application of the Tchebychev Alternation Theorem and Remez Algorithm in Image Inpainting

In modern computational tasks such as **image inpainting**, restoring missing or corrupted parts of an image is crucial. This involves reconstructing pixel values based on available neighboring data while preserving important features like edges, curvilinear structures, and textures. The Tchebychev Alternation Theorem and the Remez Algorithm provide a robust mathematical foundation for achieving high-fidelity reconstructions.

Image Inpainting as an Approximation Problem

Image inpainting can be framed as an approximation problem [9], where the goal is to find a smooth and accurate function that approximates the underlying structure of the image in the missing regions. This corresponds to finding a **nearest polynomial approximation** that fits the known pixel values while adhering to the constraints imposed by the geometric features of the image.

Utilizing the Tchebychev Alternation Theorem

1. **Characterizing the Problem**:
    - The alternation theorem ensures that the reconstructed polynomial alternates in error across predefined points, minimizing the overall error uniformly.
    - By treating pixel values as discrete data points in a compact domain, the theorem guarantees a globally optimal polynomial approximation.
2. **Ensuring Structural Integrity**:
    - The alternation theorem's emphasis on uniform error minimization is particularly advantageous for preserving curvilinear structures, such as edges and boundaries in images.

Remez Algorithm for Iterative Inpainting

1. **Initialization**: The initial step involves identifying the known pixel values and defining a set of alternating error points along the boundary of the missing region.
2. **Iterative Refinement**: The Remez Algorithm iteratively updates the polynomial coefficients to converge to the best approximation, ensuring minimal deviation from the true values.
3. **Convergence and Robustness**: The convergence of the Remez Algorithm guarantees that the iterative process stabilizes, leading to robust reconstructions even in non-stationary regions of the image.

Advantages in Image Inpainting

- **Robust to Artifacts**: The mathematical rigor of the alternation theorem minimizes non-stationary artifacts, which are common in conventional interpolation methods.
- **Preservation of Key Features**: The polynomial approximation inherently respects the global and local characteristics of the image, preserving critical features like edges and textures.
- **Computational Efficiency**: The iterative nature of the Remez Algorithm ensures computational efficiency, making it suitable for large-scale image data.

Case Study: Curvilinear Structures

Consider an image with missing data along a curvilinear structure, such as roads in satellite imagery or vessels in medical imaging. Using the Tchebychev Alternation Theorem and the Remez

Advances in Nonlinear Variational Inequalities
ISSN: 1092-910X
Vol 28 No. 3s (2025)

Algorithm:

- A nearest polynomial approximation is constructed to align with the known data points.
- The alternation theorem ensures that the error between the reconstructed and actual data alternates uniformly, minimizing deviation.
- The Remez Algorithm iteratively refines the polynomial, resulting in a seamless reconstruction that blends naturally with the original image.

Results:

This subsection demonstrates the capabilities of the Remez algorithm to solve image inpainting problems. Initially a benchmark is taken, and cracks are added programmatically. The regions of cracks are informed to the algorithm as a mask. The challenges to solve the inpainting problem are related to singularity aspects that arise due to the system of linear equations constructed during the Remez Algorithm is not solvable. This usually happens due to insufficient or poorly distributed known points in the inpainting process. To address this, safeguards are added to the algorithm by ensuring that there are at least n+2 known points and switching over to simple linear interpolation if there are fewer points. And as a preprocessing step the x-coordinates are normalized to a smaller range (e.g., [0, 1]) to avoid numerical instability in the Vandermonde matrix. Figure 1 presents the results for simple crack removal experiments.

Figure 1 Crack removal example

Conclusions

This article presented an end-to-end discussion on the polynomial approximation of a continuous function. The Tchebychev characterization and algorithmic construction made the discussion more elegant and implementable. The proof of convergence of the Remez algorithm gave completeness to the article, and new researchers find it very convenient to access all aspects of approximation theory in one place. This maiden effort of compiling all involved theorems and proofs is proposed to develop visual analytics in the future. And, finally, the application of these mathematical ideas is presented by solving an instance of image inpainting problem.

Conflicts of Interest Statement: Manuscript title: Characterization and Construction of Best Approximation Using Classical Tchebychev Alternation Theorem and Remez Algorithm: Inpaining

Case Study. The authors certify that they have NO affiliations with or involvement in any organization or entity with any financial interest.

References:

[1] Jan Brinkhuis., Vladimir Tikhomirov, Optimization: Insights and Applica- tions, Princeton University Press, (2005).
[2] E.W. Cheney, ”Introduction to approximation theory” , McGraw-Hill, pp. Chapt. 3, (1966).
[3] C. Dellacherie., P. A. Meyer. Probabilit ⁀ es et potentiel, chapter I-IV, pages 85–115. Hermann, Paris, (1975).
[4] G.G., Lorentz., Holt, Rinehart., Winston, ”Approximation of functions” , pp. Chapt. 2, (1966).
[5] Crouzeix, JP., Sukhorukova, N., Ugon, J. Finite Alternation Theorems and a Constructive Approach to Piecewise Polynomial Approximation in Cheby- shev Norm. Set-Valued Var. Anal 28, 123–147 (2020).
[6] Remez, E.: General computational methods of Chebyshev approximation. Atom. Energy Transl., 4491 (1957).
[7] G. Meinardus, Approximation of Functions: Theory and Numerical Meth- ods, (1967).
[8] Dunford.s, and Schwartz, Linear Operators Vol1, General Theory, Wiley, (1988).
[9] Raghava M, Arun Agarwal, C Raghavendra Rao, Spatial Anisotropic Interpolation Approach for Text Removal from an Image, 153—164, LNAI, MIWAI 2013.