Approximation of Functions

Orthogonal Polynomials

Numerical Methods

2023-2024

[DΣIM]

Orthogonal Polynomials

- The approximation of functions becomes simpler when we use an orthogonal set as polynomial basis.
- Unfortunately, this is not the case for the set of monic polynomials $\{1, x, x^2, \ldots, x^n\}$. In fact, using the Euclidean norm in the interval [0,1] they constitute a nearly dependent set of basis functions. Their scalar product gives:

$$\left\langle x^{i-1}, x^{j-1} \right\rangle = \int_0^1 x^{i-1} x^{j-1} dx = \frac{1}{i+j-1}$$

Orthogonal Polynomials

- This gives precisely the coefficients of the known Hilbert matrix:

$$\mathbf{A}_n = (a_{ij}) = \frac{1}{i+j-1} \qquad \mathbf{A}_4 = \begin{pmatrix} 1 & \frac{1}{2} & \frac{1}{3} & \frac{1}{4} \\ \frac{1}{2} & \frac{1}{3} & \frac{1}{4} & \frac{1}{5} \\ \frac{1}{3} & \frac{1}{4} & \frac{1}{5} & \frac{1}{6} \\ \frac{1}{4} & \frac{1}{5} & \frac{1}{6} & \frac{1}{7} \end{pmatrix}$$

- Which has a non-zero determinant, but has also a very large condition number, given very unstable linear systems which, eventually, need to be solved to obtain the polynomial coefficients.

Weierstrass Theorem

- Let f be a given continuous function in the interval $[a, b]$. The lower bound (infimum) of all possible values of:

$$\|f - p_n\|_\infty = \max_{x \in [a,b]} |f(x) - p_n(x)|$$

- Is denoted by $E_n(f)$.
- ***Theorem***. *Weierstrass approximation Theorem*.
- For every continuous function f defind on a closed bounded interval $[a, b]$, it holds that

$$\lim_{n \to \infty} E_n(f) = 0$$

Weierstrass Theorem

- In many cases $E_n(f)$ decreases so slowly toward zero (as n grows) that it is impractical to approximate f with only one polynomial in the entire interval $[a, b]$ as its degree will be too high for practical numerical purposes.
- The approximation methods give polynomials whose maximal errors will be significantly larger than $E_n(f)$, even if f is quite smooth.
- Even equidistant interpolation can give rise to convergence problems when the number of nodes become large (Ruge phenomenon)

Orthogonal Polynomials

- By a family of orthogonal polynomials, we mean a family of polynomials which is polynomials which is orthogonal in a broad sense:

$$\langle \varphi_n , \varphi_m \rangle = \int_a^b w(x)\varphi_n(x)\varphi_m(x)dx = 0, \quad n \neq m$$

- Where $w(x)$ is a given weight function, specified for the interval $[a, b]$.

Orthogonal Polynomials

- Expansions of functions in terms of orthogonal polynomials are very useful. They are easy to manipulate and can be implemented in any computer. They have good convergence properties, and they give a well-conditioned representation of a function.
- The theory of orthogonal polynomials also constitutes the background for numerical problems such as numerical integration, or the algebraic eigenvalue problem.

Gram-Schmidt Polynomials

- ***Theorem*** **(*Gram-Schmidt*)** There exists a sequence of polynomials $\{\varphi_n(x), n \geq 0\}$ with $degree\big(\varphi_n(x)\big) = n$, for all n and

$$\left\langle \varphi_n, \varphi_m \right\rangle = 0, \qquad n \neq m, \qquad n, m \geq 0$$

- Moreover, we can construct the sequence with the additional properties: (1) $\langle \varphi_n, \varphi_m \rangle = 1$ for all n. (2) The coefficient of x^n in $\varphi_n(x)$ is positive. (3) The sequence is unique.

Gram-Schmidt Polynomials

- We show a constructive and recursive method of obtaining the members of the sequence called the Gram-Schmidt process.
- Let

$$\varphi_0(x) = c$$

- A constant polynomial in $[a, b]$. Pick it such that $\|\varphi_0\|_2 = 1$ and $c > 0$. Then

$$\langle \varphi_0, \varphi_0 \rangle = c^2 \int_a^b w(x)dx = 1$$

$$c = \left[\int_a^b w(x)dx \right]^{-1/2}$$

Gram-Schmidt Polynomials

- For constructing $\varphi_1(x)$, begin with:

$$\psi_1(x) = x + a_{1,0}\varphi_0(x)$$

- Then

$$\langle \psi_1, \varphi_0 \rangle = 0 \Rightarrow 0 = \langle x, \varphi_0 \rangle + a_{1,0} \langle \varphi_0, \varphi_0 \rangle$$

$$a_{1,0} = -\langle x, \varphi_0 \rangle = \frac{-\int_a^b x w(x) dx}{\left[\int_a^b w(x) dx\right]^{1/2}}$$

Gram-Schmidt Polynomials

- Now define

$$\varphi_1(x) = \frac{\psi_1(x)}{\left\|\psi_1\right\|_2}$$

- And we have:

$$\left\|\varphi_1\right\|_2 = 1, \qquad \left\langle \varphi_1, \varphi_0 \right\rangle = 0$$

- While the coefficient of x is positive.
- From now on, the process will continue recursively.

Gram-Schmidt Polynomials

- To construct $\varphi_n(x)$, first define:

$$\psi_n(x) = x^n + a_{n,n-1}\varphi_{n-1}(x) + \cdots + a_{n,0}\varphi_0(x)$$

- And choose the constants as to make ψ_n orthogonal to φ_j for $j = 0, \dots, n-1$. Then:

$$\left\langle \psi_n, \varphi_j \right\rangle = 0 \Rightarrow a_{n,j} = -\left\langle x^n, \varphi_j \right\rangle, \quad j = 0, 1, \dots, n-1$$

- And the desired $\varphi_n(x)$ will be:

$$\varphi_n(x) = \frac{\psi_n(x)}{\left\| \psi_n \right\|_2}$$

Orthogonal Polynomials

- *Theorem.* Let $\{\varphi_n, n \geq 0\}$ be an orthogonal family of polynomials on (a, b) with weight function $w(x)$. If $f(x)$ is a polynomial of degree m, then:

$$f(x) = \sum_{n=0}^{m} \frac{\langle f, \varphi_n \rangle}{\langle \varphi_n, \varphi_n \rangle} \varphi_n(x)$$

- Every polynomial can be written as a combination of orthogonal polynomials of no greater degree, since we have $\varphi_0 = c$ and

$$1 = \frac{1}{c}\varphi_0(x)$$

Orthogonal Polynomials

- The Gram-Schmidt process gives:

$$\varphi_1(x) = c_{1,1}x + c_{1,0}\varphi_0(x), \qquad c_{1,1} \neq 0.$$

- And:

$$x = \frac{1}{c_{1,1}}\left[\varphi_1(x) - c_{1,0}\varphi_0(x)\right]$$

- Finally, using induction

$$\varphi_r(x) = c_{r,r}x^r + c_{r,r-1}\varphi_{r-1}(x) + \cdots + c_{r,0}\varphi_0(x), \qquad c_{r,r} \neq 0$$

Orthogonal Polynomials

- And we have

$$x^r = \frac{1}{c_{r,r}}\left[\varphi_r(x) - c_{r,r-1}\varphi_{r-1}(x) - \cdots - c_{r,0}\varphi_0(x)\right]$$

- Thus, every monomial can be rewritten as a combination of orthogonal polynomials of no greater degree. For an arbitrary polynomial of degree m we can write:

$$f(x) = b_m\varphi_m(x) + \cdots + b_0\varphi_0(x)$$

- For some choice of $b_0, \dots, b_m$

Orthogonal Polynomials

- To calculate each b_i, we multiply both sides by $w(x)$ and $\varphi_i(x)$ and integrate over (a, b). Then:

$$\langle f, \varphi_i \rangle = \sum_{j=0}^{m} b_j \langle \varphi_j, \varphi_i \rangle = b_i \langle \varphi_i, \varphi_i \rangle$$

$$b_i = \frac{\langle f, \varphi_i \rangle}{\langle \varphi_i, \varphi_i \rangle}$$

- ***Corollary:*** If $f(x)$ is a polynomials of degree $\leq m-1$, then

$$\langle f, \varphi_m \rangle = 0$$

- And $\varphi_m(x)$ is orthogonal to $f(x)$

Legendre Polynomials

- This family of polynomials is defined in the interval $[-1,1]$ as follows:

$$P_0(x) = 1$$

$$P_n(x) = \frac{1}{2^n n!}\frac{d^n}{dx^n}\left[\left(x^2-1\right)^n\right], \qquad n = 1, 2, \ldots$$

- Since $(x^2-1)^n$ is a polynomial of degree $2n$, $P_n(x)$ is a polynomial of degree n

Legendre Polynomials

- Another common definition is based on a formal expansion in powers of its generating function:

$$\frac{1}{\sqrt{1-2xt+t^2}} = \sum_{n=0}^{\infty} P_n(x)t^n$$

- The coefficient of t^n in this series is a polynomial of degree n in the variable x.

Legendre Polynomials

- These polynomials verify the orthogonality property. If we use the scalar product

$$\langle f, g \rangle = \int_{-1}^{1} f(x) g(x) dx$$

- Then

$$\langle P_i, P_j \rangle = \begin{cases} 0 & \text{if } n \neq j \\ \dfrac{2}{2i+1} & \text{if } n = j \end{cases}$$

Legendre Polynomials

- This family satisfies also the recurrence relation:

$$P_{n+1}(x) = \frac{2n+1}{n+1} x P_n(x) - \frac{n}{n+1} P_{n-1}, \quad n \geq 1$$

- And can be used in any interval $[a, b]$ with the transformation:

$$t = \frac{2}{b-a}\left(x - \frac{a+b}{2}\right) \in [-1,1]$$

- Obtaining the polynomials

$$\varphi_n(x) = P_n\left(\frac{2}{b-a}\left(x - \frac{a+b}{2}\right)\right)$$

Legendre Polynomials

- The first Legendre Polynomials are as follows:

$$P_0(t) = 1$$

$$P_1(t) = t$$

$$P_2(t) = \frac{1}{2}(3t^2 - 1)$$

$$P_3(t) = \frac{1}{2}\left(5t^3 - 3t\right)$$

$$P_4(t) = \frac{1}{8}\left(35t^4 - 30t^2 + 3\right)$$

$$P_5(t) = \frac{1}{8}\left(63t^5 - 70t^3 + 15t\right)$$

Tchebycheff Polynomials

- The polynomials of Tchebycheff is also a family of orthogonal polynomials on the interval $[-1,1]$. They are defined by the expression

$$T_n(x) = \cos\left(n \arccos x\right)$$

- Thus, if $n = 0$, we have $T_n(x) = 1$ while if $n = 1$, we obtain $T_1(x) = \cos(\arccos(x)) = x$. In the case $n = 2$

$$T_2(x) = \cos(2\arccos x) = 2\cos^2\left(\arccos x\right) - 1 = 2x^2 - 1$$

Tchebycheff Polynomials

- The rest of the polynomials can be obtained recursively. If we make $\theta = \arccos(x)$, we have the relation:

$$T_{n-1}(x) + T_{n+1}(x) = \cos\left[(n-1)\theta\right] + \cos\left[(n+1)\theta\right]$$
$$= 2\cos\theta\cos(n\theta) = 2nT_n(x)$$

- Frow which we deduce the recursive relation between the Tchebycheff polynomials:

$$\boxed{T_{n+1}(x) = 2nT_n(x) - T_{n-1}(x)}$$

Tchebycheff Polynomials

- The first Tchebycheff polynomials are the following

$$
\begin{aligned}
T_0(x) &= 1 \\
T_1(x) &= x \\
T_2(x) &= 2x^2 - 1 \\
T_3(x) &= 4x^3 - 3x \\
T_4(x) &= 8x^4 - 8x^2 + 1 \\
T_5(x) &= 16x^5 - 20x^3 + 5x \\
T_6(x) &= 32x^6 - 48x^4 + 18x^2 - 1
\end{aligned}
$$

Tchebycheff Polynomials

- In the continuous case, they are a family of orthogonal polynomials with the weight function $w(x) = (1 - x^2)^{-1/2}$. If we set

$$\langle f, g \rangle = \int_{-1}^{1} f(x) g(x) (1 - x^2)^{-1/2} \, dx$$

- We have

$$\langle T_i, T_j \rangle = \begin{cases} 0 & \text{if } i \neq j \\ \frac{1}{2}\pi & \text{if } i = j \neq 0 \\ \pi & \text{if } i = j = 0 \end{cases}$$

Tchebycheff Polynomials

- Using the appropriate coordinates, this family is also orthogonal in the discrete case. If we select the nodes as the zeros of the polynomial $T_{m+1}(x)$ and we use the scalar product:

$$\langle f, g \rangle = \sum_{k=0}^{m} f(x_k) g(x_k)$$

- Then for $0 \leq i \leq m, 0 \leq j \leq m,$ we have

$$\langle T_i, T_j \rangle = \begin{cases} 0 & \text{if } \; i \neq j \\ \frac{1}{2}(m+1) & \text{if } \; i = j \neq 0 \\ \text{m+1} & \text{if } \; i = i = 0 \end{cases}$$

Tchebycheff Polynomials

- The Tchebycheff polynomials are used for the approximation of transcendental functions like $f(x) = \sin(x)$. As this family forms a basis for the linear space of polynomials, all the powers of the variable x^n can be expressed as a linear combination of this basis. Then

$$1 = T_0(x) \qquad x^4 = \frac{1}{8}\left(3T_0(x) + 4T_2(x) + T_4(x)\right)$$

$$x = T_1(x) \qquad x^5 = \frac{1}{16}\left(10T_1(x) + 5T_3(x) + 5T_5(x)\right)$$

$$x^2 = \frac{1}{2}\left(T_0(x) + T_2(x)\right) \qquad x^6 = \frac{1}{32}\left(10T_0(x) + 15T_2(x) + 6T_4(x) + T_6(x)\right)$$

$$x^3 = \frac{1}{4}\left(3T_1(x) + T_3(x)\right) \qquad \vdots$$

Lanczos Economization

- If we want to approximate a function $f(x)$ using a power (Taylor) series around a point $x = c$, $c \in [a, b]$ we will have an expression of the form

$$f(x) = c_0 + c_1(x-c) + c_2(x-c)^2 + \cdots + c_n(x-c)^n + R_n(x)$$

- Where $R_n(x)$ is the residual of the approximation. If we introduce the new variable:

$$t = 2\frac{x-a}{b-a} - 1, \qquad x \in [a,b] \Leftrightarrow t \in [-1,1]$$

Lanczos Economization

- We can rewrite the power series as:

$$f(t) = a_0 + a_1 t + a_0 t^2 + \cdots + a_n t^n + S_n(t) = P_n(t) + S_n(t)$$

- We can reduce the number of operations if we change the polynomial $P_n(t)$ by a lower order polynomial $P_{n-1}(t)$ in such a way that

$$\max_{x \in [-1,1]} \left| P_n(t) - P_{n-1}(t) \right|$$

- Is as small as possible

Lanczos Economization

- Note that $(P_n(t) - P_{n-1}(t))/a_n$ is a monic polynomial of degree n. Then, for all monic polynomials:

$$\max_{x\in[-1,1]}\left|\frac{1}{a_n}\left(P_n(t)-P_{n-1}(t)\right)\right| \geq \frac{1}{2^{n-1}} = \max_{x\in[-1,1]}\left|T_n(t)\right|$$

- The best that we can do is then

$$P_n(t) = P_{n-1}(t) + a_n T_n(t)$$

Lanczos Economization

- Then we have:

$$f(t) = P_{n-1}(t) + \frac{a_n}{2^{n-1}} T_n(t) + S_n(t)$$

- If we now make the approximation $f(t) \approx P_{n-1}(t)$, the error will depend on the term:

$$R_{n-1}(t) = \frac{a_n}{2^{n-1}} T_n(t) + S_n(t)$$

- But the error of the first term is bounded by $\frac{a_n}{2^{n-1}}$, so if the combined error is lower than the desired precision, we obtain an approximation of lower order.

Lanczos Economization

- Consider the approximation of $f(x) = e^x$ by the power series:

$$e^x = 1 + x + \frac{x^2}{2} + \frac{x^3}{6} + \frac{x^4}{24} + R_4(x) = P_4(x) + R_4(x)$$

- The error of this approximation is bounded by

$$\left|R_4(x)\right| \le \frac{\left|f^{(5)}(\xi)\right|}{120} \le \frac{e}{120} \simeq 0.023, \quad \forall x \in [-1,1]$$

Lanczos Economization

- If we substitute x^4 by the expression on the Tchebycheff polynomials we have:

$$\begin{aligned} P_4(x) &= 1 + x + \frac{1}{2}x^2 + \frac{1}{6}x^3 + \frac{1}{24}\left[\frac{3}{8}T_0(x) + \frac{1}{2}T_2(x) + \frac{1}{8}T_4(x)\right] \\ &= 1 + x + \frac{1}{2}x^2 + \frac{1}{6}x^3 + \frac{1}{64} + \frac{1}{48}\left(2x^2 - 1\right) + \frac{1}{192}T_4(x) \\ &= \frac{191}{192} + x + \frac{13}{24}x^2 + \frac{1}{6}x^3 + \frac{1}{192}T_4(x) \end{aligned}$$

- If we now discard the last term, the total error is:

$$|R_4(x)| + \frac{1}{192}|T_4(x)| \le 0.023 + \frac{1}{192} = 0.0283$$

Lanczos Economization

- Thus, keeping the approximation:

$$P_4(x) = \frac{191}{192} + x + \frac{13}{24}x^2 + \frac{1}{6}x^3$$

- We obtain an approximation of the function $f(x)$ with an error of similar magnitude and with lower order. This new approximation has the additional advantage that will keep the error bounded for the whole interval, while the Taylor polynomial will increase the error as we go apart from the initial point.