Set 3: Polynomial Interpolation – Part 3

Kyle A. Gallivan

Department of Mathematics

Florida State University

Foundations of Computational Math 2
Spring 2013

Conditioning, Stability, Error

- conditioning a polynomial with respect to representation
- conditioning of the interpolating polynomial with respect to function values
- stability and practical limitations
- interpolation error
- convergence of interpolation strategies

References

In addition to the text, the following are useful references for this topic.

- Isaacson and Keller, Analysis of Numerical Methods, Wiley Press, 1966.
- Higham, Accuracy and Stability of Numerical Algorithms, SIAM, Second Edition, 2002.
- W. Gautschi, Questions of numerical conditions related to polynomials, Studies in Numerical Analysis, Volume 24 of MAA Studies in Mathematics Series, G. H. Golub, Ed., pp. 140–177, 1984
- J. H. Wilkinson, The perfidious polynomial, Studies in Numerical Analysis, Volume 24 of MAA Studies in Mathematics Series, G. H. Golub, Ed., pp. 1–28, 1984

Conditioning of Representation

- Various representations of polynomials have different condition relative to perturbation of their parameters.
- Successive basis functions that are “close” to the span of the previous basis functions yield ill-conditioned representations.
- The monomial (power) and Newton representations have “nearly colinear” basis functions as n gets large and grow increasingly ill-conditioned relative to perturbations.
- Condition numbers that are exponential in n are possible.
- This theory will be very important later when discussing orthogonal polynomials and their uses.

Basics

Definition 3.1. If $f(x) \in \mathcal{C}^{(0)}[a, b]$ then its maximum or ∞ norm is

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

If $v \in \mathbb{R}^n$ then

$$\|v\|_\infty = \max_{1 \le i \le n} |e_i^T v|$$

Monomial (Power) Basis

Definition 3.2. The linear mapping $M_n : \mathbb{R}^n \to \mathbb{P}_{n-1}$ is defined by

$$M_n(a) = \alpha_0 + \alpha_1 x + \cdots + \alpha_{n-1} x^{n-1}$$

and the condition number κ_n is such that

$$\|p_n(x) - \tilde{p}_n(x)\|_\infty \leq \kappa_n \|a - \tilde{a}\|_\infty$$

$$a^T = \begin{pmatrix} \alpha_0 & \alpha_1 & \cdots & \alpha_n \end{pmatrix}$$

$$\tilde{a}^T = \begin{pmatrix} \tilde{\alpha}_0 & \tilde{\alpha}_1 & \cdots & \tilde{\alpha}_n \end{pmatrix}$$

Monomial (Power) Basis

Theorem 3.1 (Gautschi, 1984)**.** *For the linear mapping* $M_n : \mathbb{R}^n \to \mathbb{P}_{n-1}$ *defined by*

$$M_n(a) = \alpha_0 + \alpha_1 x + \cdots + \alpha_{n-1} x^{n-1}$$

and for the interval $[-\omega, \omega]$*, where* $\omega > 0$*, we have as* $n \to \infty$

$$\kappa(M_n) \approx \begin{cases} \left(1 + \sqrt{1 + \omega^2}\right)^n & \textit{for } \omega \geq 1 \\ \left(\frac{1+\sqrt{1+\omega^2}}{\omega}\right)^n & \textit{for } \omega < 1 \end{cases}$$

whose minimum is at $\omega = 1$.

Orthogonal Basis

- $\kappa(M_n)$ is a worst case perturbation result
- In practice, especially for moderate n, the power basis or the related Newton form may not be that sensitive.
- As in $\mathbb{R}^n$, an orthogonal basis is better conditioned.
- Families of polynomials that are orthogonal with respect to some inner product on $\mathbb{P}_n$ exist and will be considered later in detail.

Orthogonal Basis

Theorem 3.2 (Gautschi, 1984)**.** *The condition number for the representation on* $-1 \le x \le 1$

$$p(x) = \alpha_0 \pi_0(x) + \cdots + \alpha_{n-1}\pi_{n-1}(x)$$

where the $\pi_k(x)$ *are orthogonal polynomials is bounded:*

$$\kappa(M_n) \le \begin{cases} n \ \sqrt{2} & \textit{for Chebyshev polynomials} \\ n \ \sqrt{2n-1} & \textit{for Legendre polynomials} \end{cases}$$

Conditioning of Interpolation

Given $x_0, x_1, \ldots, x_n$ consider two polynomials

- $p_n(x)$ that interpolates $y_0, y_1, \ldots, y_n$
- $\tilde{p}_n(x)$ that interpolates $\tilde{y}_0, \tilde{y}_1, \ldots, \tilde{y}_n$

We want a condition number κ_n

$$\|p_n(x) - \tilde{p}_n(x)\|_\infty \leq \kappa_n \|y - \tilde{y}\|_\infty$$

$$y = \begin{pmatrix} y_0 \\ y_1 \\ \vdots \\ y_n \end{pmatrix} \qquad \tilde{y} = \begin{pmatrix} \tilde{y}_0 \\ \tilde{y}_1 \\ \vdots \\ \tilde{y}_n \end{pmatrix}$$

Conditioning of Interpolation

The Lagrange basis relates function values to the interpolating polynomials:

$$\|p_n(x) - \tilde{p}_n(x)\|_\infty = \max_{x\in[a,b]} |\sum_{i=0}^{n} (y_i - \tilde{y}_i)\ell_i^{(n)}(x)|$$

$$\leq \max_{0\leq i\leq n} |(y_i - \tilde{y}_i)| \max_{x\in[a,b]} \sum_{i=0}^{n} |\ell_i(x)| = \Lambda_n \|y - \tilde{y}\|_\infty$$

$$\Lambda_n = \|\sum_{j=0}^{n} |\ell_j^{(n)}(x)|\|_\infty$$

Conditioning of Interpolation

- $\therefore \Lambda_n$, the Lebesgue constant, can be viewed as a condition number with respect to the ∞ norm of polynomial interpolation relative to changes in function values.
- It is also a condition number of the Lagrange representation of a polynomial.
- The choice of interpolation points can significantly affect the conditioning.

Conditioning of Interpolation

Examples of point selection effects:

- (Natonson, 1965) For equally spaced nodes

$$\Lambda_n(X) \approx \frac{2^{n+1}}{en\log n}$$

- (Gautschi, 1984) For the Chebyshev points, $0 \le j \le n$

$$x_j = \cos\frac{(2j+1)\pi}{2n+2}, \qquad \Lambda_n(X) \approx \frac{2}{\pi}\log n$$

- Among all Lagrange bases, best value, i.e., slowest growth, is

$$\Lambda_n(X) = O(\log n)$$

Interpolation Stability

- Two parts of process:
  1. evaluation of the parameters, e.g., divided differences
  2. evaluation of the polynomial given the computed differences
- Many analyses in the literature.
- Horner's rule has a backward error, i.e., the computed value is the exact value of a perturbed polynomial.
- It can be adapted to Newton and orthogonal bases (any basis with a definition based on a recurrence)

Newton Form Interpolation

- See Higham 2002 for a nice summary.
- It is possible to have significant errors in the difference table and still reproduce the original data accurately.
- $x_0 < x_1 < \cdots < x_n$ or $x_0 > x_1 > \cdots > x_n$ are "optimal" ordering for keeping $|fl(p_n(x_i)) - p_n(x_i)|$ acceptably small, i.e., error in interpolation conditions.

Newton Form Interpolation

- If keeping $|fl(p_n(x)) - p_n(x)|$ small for $x \neq x_i$ is the goal then Leja ordering (Reichel, BIT30:332–346, 1990) is useful.
- Ordered points satisfy

$$x_0 = \max_i |x_i|$$

$$\prod_{k=0}^{j-1} |x_j - x_k| = \max_{i \geq j} \prod_{k=0}^{j-1} |x_i - x_k|$$

- Two orderings:

$$-1,\ -0.5,\ 0,\ 0.5,\ 1, \quad \text{small reconstruction error}$$

$$1,\ -1,\ 0,\ 0.5,\ -0.5, \quad \text{Leja ordered}$$

- A Leja ordering can be computed in $O(n^2)$ operations.

Newton Form Interpolation

- If n is small or moderately sized the entire divided difference table may be kept.
- A dynamic point ordering heuristic can be used to choose a path through the table based on x
- Start at the x_i nearest x, then add to the polynomial the nearest in turn until desired degree achieved.
- Easy to locate initial x_{i_0}, then need only compare two values for each additional x_{i_j} included.

Newton Form

i	0		1		2		3
x_i	0		1		3		4
y_i	-5		1		25^*		55
$y[-,-]$		6		12		30^*	
$y[-,-,-]$			2		6^*		
$y[-,-,-,-]$				1^*			

- Table generated with $x_0 < x_1 < \cdots < x_n$.
- $p_3(x) = -5 + 7x - 2x^2 + x^3$
- $p_3(2.55) \approx 16.42638$ would use differences marked with $*$.

Pointwise Error

Pointwise error is defined for all x in a region of interest via:

$$E_n(x) = f(x) - p_n(x)$$

Theorem 3.3. *Let* $x_i \in [a, b]$ *for* $0 \le i \le n$ *be distinct points,* $f(x) \in \mathcal{C}^{(n+1)}[a, b]$, *and* $p_n(x) \in \mathbb{P}_n$ *be the interpolating polynomial of degree* n. *For each* $x \in [a, b]$ $\exists \xi(x) \in [a, b] \ni$

$$E_n(x) = f(x) - p_n(x) = \frac{f^{(n+1)}(\xi)}{(n+1)!}\omega_{n+1}(x)$$

Proof. See text p. 335. $\square$

Pointwise Error

Remark.

- $E_n(x_i) = 0$
- $E_n(x) = 0$ if $f(x) \in \mathbb{P}_n$
- $E_n(x)$ tends to be very oscillatory.
- If $f^{(n+1)}(x)$ is nicely bounded on $[a, b]$ error can be estimated.
- Problems if $f^{(n+1)}(x)$ grows faster than $(n+1)!$ or $\omega_{n+1}(x)$ is large.
- This is not a general result for all continuous $f(x)$.

Error and Derivatives

Given nodes $x_0, \ldots, x_n$, an associated interpolant $p_n(t)$, an arbitrary x, and an associated interpolant $p_{n+1}(t)$, let I be the smallest interval containing all of the points.

$$p_{n+1}(t) = p_n(t) + f[x_0, \ldots, x_n, x]\omega_{n+1}(t)$$

$$\therefore E_n(x) = f(x) - p_n(x) = f[x_0, \ldots, x_n, x]\omega_{n+1}(x)$$

$$\therefore \text{ if, additionally, } f \in \mathcal{C}^{(n+1)}[I]$$

$$E_n(x) = \frac{f^{(n+1)}(\xi)}{(n+1)!}\omega_{n+1}(x)$$

$$f[x_0, \ldots, x_n, x] = \frac{f^{(n+1)}(\xi)}{(n+1)!}$$

The last one can be used as the basis for an error estimate.

Some Useful Identities

$$\omega'_{n+1}(x_i) = \prod_{k=0, k\neq i}^{n} (x_i - x_k)$$

$$f[x_0, \ldots, x_n] = \sum_{i=0}^{n} \frac{f(x_i)}{\omega'_{n+1}(x_i)}$$

$f[x_0, \ldots, x_n] = f[x_{i_0}, \ldots, x_{i_n}]$ for any permutation of indices

$$f[x_0, \ldots, x_n, x] = \sum_{i=0}^{n} \frac{f[x, x_i]}{\omega'_{n+1}(x_i)}$$

Many others, including nondistinct points, see Isaason and Keller (1966) and text.

Convergence on Interval

Approximation by polynomials is motivated by the following theorem:

Theorem 3.4. *(Weierstrass Approximation Theorem) If* $f(x) \in \mathcal{C}^{(0)}[a,b]$ *then* $\forall \epsilon > 0\ \exists n \in \mathbb{Z}$ *and polynomial* $p_n(x)$ *with degree at most* n *such that*

$$\|f(x) - p_n(x)\|_\infty < \epsilon.$$

This is uniform convergence, i.e., pointwise error at all points in interval is bounded and the bound is going to 0.

Convergence on Interval

- Theorem 3.4 gives no insight into how to choose $p_n(x)$ and does not relate necessarily to an interpolation strategy.
- The result can be derived as a corollary to a constructive theorem due to Bernstein.
- A sequence of polynomials is defined and shown to converge uniformly.

Bernstein Polynomials

Definition 3.3. Let $f(x)$ be a real function defined on $[0,1]$. The n-th Bernstein polynomial for f is

$$\begin{aligned}
B_n(x) = B_n(x;f) &= \sum_{k=0}^{n} f(\frac{k}{n})\binom{n}{k} x^k (1-x)^{n-k} \\
&= \sum_{k=0}^{n} f(\frac{k}{n})\phi_{n,k}(x) \\
&= \sum_{k=0}^{n} f(x_k)\phi_{n,k}(x) \\
x_k &= k/n
\end{aligned}$$

Bernstein Polynomials

- Sum of $f(x)$ at uniformly-spaced points.
- The weight $\phi_{n,k}(x)$ is non-negative on $[0, 1]$ and $\sum_{k=0}^{n} \phi_k(x) = 1$.
- The weight $\phi_{n,k}(x)$ can be very small for k where x is far from k/n.
- The weight $\phi_{n,k}(x)$ achieves its maximum on $[0, 1]$ at $x = k/n$.
- The construction is not interpolatory, i.e., $B_n(x_k)$ is not necessarily equal to $f(x_k)$.
- $B_n(x)$ usually interpolates $f(x)$ but where and how often it does is not controlled.

Bernstein Approximation

Theorem 3.5. *If* $f(x) \in \mathcal{C}^{(0)}[0,1]$ *then* $B_n(x)$ *converges uniformly to* $f(x)$ *on* $[0,1]$, *i.e.,*

$$\lim_{n \to \infty} \|f(x) - B_n(x)\|_\infty = 0$$

Proof. See Bartle, Elements of Real Analysis (1976) □

Corollary 3.6. *If, in addition, on* $[0,1]$, $f(x)$ *satisfies the Lipschitz condition* $|f(x) - f(\hat{x})| < \lambda |x - \hat{x}|$ *then*

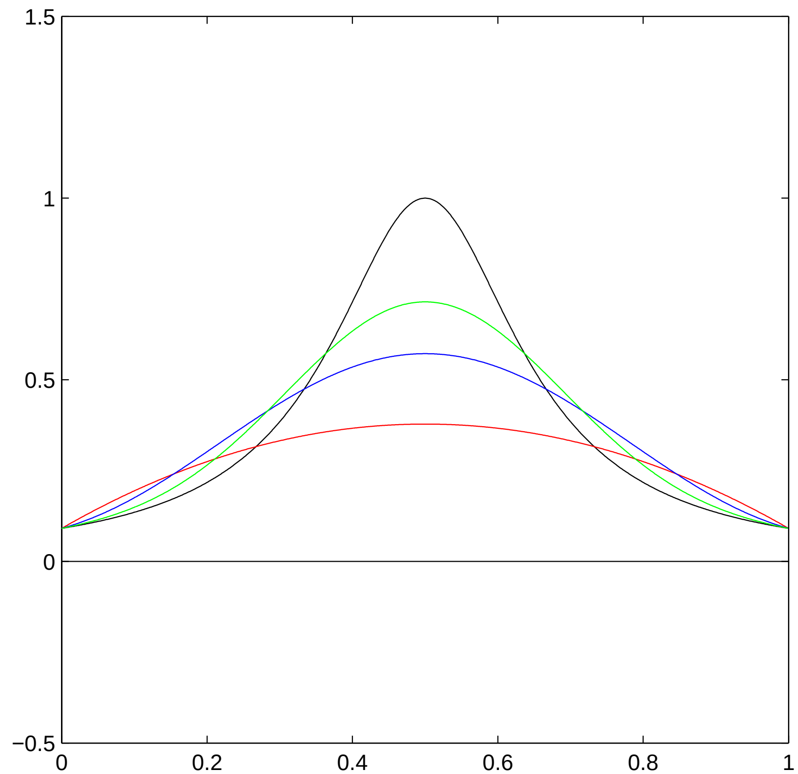
$$\|f(x) - B_n(x)\|_\infty < \frac{9}{4} \lambda n^{-1/2}$$

Proof. See Isaacson and Keller (1966) □

Bernstein Approximation

- Easily updated to apply to $[a, b]$.
- Convergence is much slower than other approximation methods.
- Even if $f(x) \in \mathcal{C}^{(p)}[0, 1]$ with $p \geq 2$ convergence remains relatively slow.
- Useful theoretical result but Bernstein polynomials are not used in practice for this type of approximation.
- Bernstein polynomials are used when “shape” is important.
- This shows that polynomials **can converge uniformly** to a continuous f.

Bernstein Convergence



$f(x) = 1/(1+10x^2)\ \ -1 \leq x \leq 1$ shifted to $[0,1]$ – black, $B_3(x)$ – red, $B_6(x)$ – blue, $B_{15}(x)$ – green

Convergence of Interpolating Polynomials

Definition 3.4. An interpolating strategy is defined by a sequence, X, of sets of nodes $X_n = \{x_0^{(n)}, \ldots, x_n^{(n)}\}$.

- The sets X_n are chosen independently of any particular $f(x)$.
- Each X_n defines an interpolatory polynomial, $p_n(x)$, of degree n such that given an $f(x)$, $p_n(x_i^{(n)}) = f(x_i^{(n)})$ for $0 \leq i \leq n$.

Uniform interpolation:

$$X_n = \{x_i^{(n)} = x_0 + ih, \quad h = (b-a)/n\}$$

Chebyshev interpolation:

$$X_n = \{x_j^{(n)} = \cos(\frac{2j+1}{n+1}\frac{\pi}{2})\}$$

Convergence of Interpolating Polynomials

The convergence of

$$\|f(x) - p_n(x)\|_\infty$$

on a closed interval $[a, b]$ for $f(x) \in \mathcal{C}^{(0)}[a, b]$ is complicated.

The result depends on

- the choice of X,
- the class of functions $f(x)$ that may be more constrained than $\mathcal{C}^{(0)}[a, b]$

Runge's Phenomenon

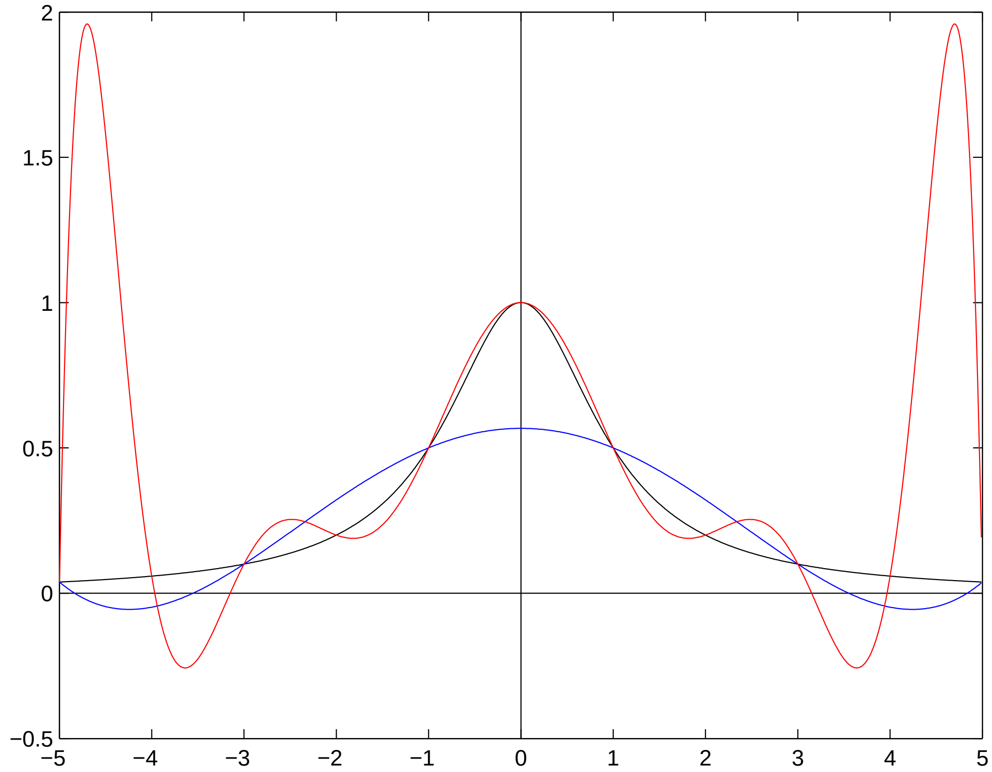
Let $I = [-5, 5]$ and define $x_j^{(n)} = -5 + jh_n$ with $h_n = 10/n$ and $0 \le j \le n$. The sets $X_n = \{x_0^{(n)}, \ldots, x_n^{(n)}\}$ define a sequence, X, of sets of nodes each of which define an interpolatory polynomial, $p_n(x)$, of degree n. It can be shown that

$$\lim_{n\to\infty} \|f(x) - p_n(x)\|_\infty$$

does not converge on I for $f(x) = 1/(1 + x^2)$.

Proof. See Isaacson and Keller (1966) $\square$

Runge's Phenomenon



$f(x) = 1/(1+x^2)$ – black, $p_5(x)$ – blue, $p_{10}(x)$ – red

Runge's Phenomenon

- The divergence occurs near the endpoints of the interval.
- This is typical behavior so keep order low to be effective with uniformly spaced points.
- Non-uniform points more dense near endpoints are needed for better interpolation strategies, e.g., Chebyshev.

Convergence of Interpolating Polynomials

For each degree n we can define the "best" polynomial approximation:

Definition 3.5. Let $p_n^*(x) \in \mathbb{P}_n$ be such that
$E_n^* = \|f(x) - p_n^*(x)\|_\infty \leq \|f(x) - q_n(x)\|_\infty \ \ \forall q_n(x) \in \mathbb{P}_n$.

This approximation will be discussed in much more detail later.

Convergence of Interpolating Polynomials

Lemma. *Let the sequence X define an interpolating strategy, and let the Lebesgue constant be*

$$\Lambda_n(X) = \| \sum_{j=0}^{n} |\ell_j^{(n)}(x)| \|_\infty$$

for the set of nodes $X_n = \{x_0^{(n)}, \ldots, x_n^{(n)}\}$ where $\ell_j^{(n)}(x)$ are the Lagrange characteristic functions associated with X_n.

If $f(x) \in \mathcal{C}^{(0)}[a, b]$ then

$$E_n^* \leq \|f(x) - p_n(x)\|_\infty \leq (1 + \Lambda_n(X))E_n^*$$

for $n = 0, 1, \ldots$.

Convergence of Interpolating Polynomials

- A small Lebesgue constant $\Lambda_n(X)$ guarantees good ∞ norm approximation of $f(x)$ for the associated $p_n(x)$.
- Bounding the Lebesgue constant $\Lambda_n(X)$ is a key task when analyzing an interpolating strategy.
- Erdos (1961) showed $\forall X \;\; \exists C > 0$ such that

$$\Lambda_n(X) > \frac{2}{\pi}\log(n+1) - C \;\; n = 0, 1, \ldots$$

  so $\Lambda_n(X) \to \infty$.

- Natonson (1965) showed for equally spaced nodes

$$\Lambda_n(X) \approx \frac{2^{n+1}}{en\log n}$$

Convergence of Interpolating Polynomials

- The error bound predicted by the Lebesgue constant is not achieved for all $f(x) \in \mathcal{C}^{(0)}[a,b]$.
- A particular strategy may work well with a particular f or some particular class of f
- Unfortunately, no interpolating strategy, X, converges for all $f(x) \in \mathcal{C}^{(0)}[a,b]$.

Convergence of Interpolating Polynomials

Theorem 3.7. *(Faber 1914) Given an interpolating strategy defined by any sequence of node sets* X *on* $[a, b]$, $\exists\ f(x) \in \mathcal{C}^{(0)}[a, b]$ *such that* $\|f(x) - p_n(x)\|_\infty$ *does not converge.*

Summary

- (Bernstein) $B_n(x)$ converge uniformly for all $f(x) \in \mathcal{C}^{(0)}[a,b]$ but not an interpolating strategy since the number and position of points where they agree with $f(x)$ depend on $f(x)$.
- (Faber) No $p_n(x)$ defined by an X converges for all $f(x) \in \mathcal{C}^{(0)}[a,b]$.
- (Bernstein) and (Brutman, Passow) interpolant for $|x|$ on $[-1,1]$ diverges almost everywhere for a variety of well-known node sets.
- For an interpolating strategy to converge uniformly:
  - the class of $f(x)$ is more restrictive than $\mathcal{C}^{(0)}[a,b]$,
  - the nodes in $X_n = \{x_0^{(n)}, \ldots, x_n^{(n)}\}$ are chosen carefully

Uniform Convergence of Interpolating Polynomials

Theorem 3.8. *Let $I = [-1, 1]$ and let the interpolating strategy be defined by the sets $X_n = \{x_0^{(n)}, \ldots, x_n^{(n)}\}$ given by the Chebyshev zeros*

$$x_j^{(n)} = \cos(\frac{2j+1}{n+1}\frac{\pi}{2}) \;\; 0 \leq j \leq n.$$

- *If $f(x) \in \mathcal{C}^{(2)}[I]$ then $\|f(x) - p_n(x)\|_\infty$ converges uniformly on I.*
- *If $f(x) \in \mathcal{C}^{(0)}[I]$ satisfies the Lipschitz condition $|f(x) - f(\hat{x})| < \lambda|x - \hat{x}|$ then $\|f(x) - p_n(x)\|_\infty$ converges uniformly on I.*

Proof. See Isaacson and Keller (1966), Ueberhuber (1995) $\square$

We will discuss this interpolation strategy in more detail later.