Set 7: Polynomial Interpolation – Part 3

Kyle A. Gallivan

Department of Mathematics

Florida State University

Foundations of Computational Math 2
Fall 2017

Conditioning, Stability and Error

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

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 7.1. 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 7.2. 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 7.1 (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 7.2 (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:*

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

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 $\tilde{y}_0, \tilde{y}_1, \ldots, \tilde{y}_n$

A condition number κ_n with respect to perturbations in the interpolated values, y_i, is desired and sastisfies

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

$$y^T = \begin{pmatrix} y_0 & y_1 & \ldots & y_n \end{pmatrix} \quad \tilde{y}^T = \begin{pmatrix} \tilde{y}_0 & \tilde{y}_1 & \ldots & \tilde{y}_n \end{pmatrix}$$

Note the x_i are not perturbed. Such a condition number can be used in concert with a backward error analysis that casts the effects of the finite precision back on the y_i.

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 and shows that the choice of interpolation points can significantly affect the conditioning.
- As given it is an absolute condition number. The relative form is

$$\Lambda^{(rel)} = \frac{\Lambda_n \|y_i\|_\infty}{\|p_n(x)\|_\infty}$$

Conditioning of Interpolation

Examples of point selection effects:

- (Natonson, Constructive Function Theory, VIII, Unger, 1965) For equally spaced nodes

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

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

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

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

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

Conditioning of Interpolation

Higham (IMA Jour. Num. Analysis 24, 2004) gives the following relative conditioning statement for the value of the interpolating polynomial at a point x

$$\kappa(x,n,y) = \frac{\sum_{i=0}^{n} |\ell_i(x) y_i|}{|p_n(x)|} \geq 1, \quad 1 \leq \kappa(x,n,1) \leq \Lambda_n$$

$$\forall \Delta y \in \mathbb{R}^{n+1} \text{ with } |\Delta y_i| \leq \epsilon |y_i|$$

$$\frac{|p_n(x) - \tilde{p}_n(x)|}{|p_n(x)|} \leq \kappa(x,n,y)\epsilon$$

Interpolation Stability

- Two parts of process:
  1. evaluation of the parameters, e.g., divided differences
  2. evaluation of the polynomial given the computed parameters, e.g., Horner's rule
- 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. (see Higham 2002)
- The algorithm can be adapted to Newton and orthogonal bases (any basis with a definition based on a recurrence)

Lagrange Form Interpolation

Higham (IMA Jour. Num. Analysis 24, 2004) gives stability results for the Barycentric forms 1 and 2 discussed by Berrut and Trefethen (Siam Review Vol. 46 No. 3)

- Barycentric form 1 (modified Lagrange) is backward stable with respect to perturbations to the y_i.
- Barycentric form 2 not proven to be backward stable.
- Barycentric form 2 forward error bound given.
- Barycentric form 2 forward error can be much larger than the Barycentric form 1 forward error.
- For well-conditioned problems, i.e., Λ_n acceptably small, Barycentric form 2 is forward stable and performs similarly to Barycentric form 1 in terms of forward error.

Higham's Basic Definitions and Facts

Let u denote unit roundoff

$$\mu_k = \frac{ku}{1 - ku}$$

$$\langle k \rangle = \prod_{i=1}^{k} (1 + \delta_i)^{\rho_i}, \quad \rho_i = \pm 1, \;\; |\delta_i| \leq u$$

$$|\langle k \rangle - 1| \leq \mu_k$$

In $\langle k \rangle_j$, j indicates that the δ_i and ρ_i depend on some iteration j.

Barycentric Form 1

Assuming, x_i and y_i are floating point numbers, the computed $\hat{p}_n(x)$ using the Barycentric interpolation formula form 1

$$p_n(x) = \omega_{n+1}(x) \sum_{i=0}^{n} y_i \frac{\gamma_i}{(x - x_i)}$$

$$\gamma_i^{-1} = \omega'_{n+1}(x_i) = \prod_{j=0, i \neq j}^{n} (x_i - x_j) \quad \text{and} \quad \omega_{n+1}(x) = \prod_{i=0}^{n} (x - x_i)$$

satisfies

$$\tilde{p}_n(x) = \omega_{n+1}(x) \sum_{i=0}^{n} \frac{\gamma_i}{(x - x_i)} \, y_i \, \langle 5n + 5 \rangle_i$$

using the notation of Higham 2002.

Barycentric Form 1

The computed $\hat{p}_n(x)$ using the Barycentric interpolation formula form 1 satisfies the forward error bound:

$$\frac{|p_n(x) - \hat{p}_n(x)|}{|p_n(x)|} \le \mu_{5n+5}\kappa(x, n, y)$$

Barycentric Form 2

Recall the Barycentric interpolation formula form 2

$$p_n(x) = \frac{\sum_{i=0}^{n} y_i \frac{\gamma_i}{(x-x_i)}}{\sum_{i=0}^{n} \frac{\gamma_i}{(x-x_i)}}$$

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

Barycentric Form 2

Assuming, x_i and y_i are floating point numbers, the computed $\hat{p}_n(x)$ using the Barycentric form 2 satisfies the forward error bound

$$\frac{|p_n(x) - \hat{p}_n(x)|}{|p_n(x)|} \leq (3n+4)\kappa(x,n,y)u + (3n+2)\kappa(x,n,1)u + O(u^2)$$

$$\leq (3n+4)\kappa(x,n,y)u + (3n+2)\Lambda_n u + O(u^2)$$

Barycentric Form 2

- Weakly stable and acceptable unless $\kappa(x, n, 1) \gg \kappa(x, n, y)$
- $\kappa(x, n, 1) \leq \Lambda_n$ implies that stability is acceptable except for poorly chosen point sets as measured by Λ_n
- Such problems are ill-conditioned and therefore difficult to solve for any algorithm.
- Computational complexity advantages of Barycentric form 2 and weak stability in practice indicates a preference compared to form 1.

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.
- The computation of the coefficients of the Newton form can be shown to be the multiplication of a vector containing $y_0, \ldots y_n$ by n structured and sparse lower triangular matrices.
- The structure of the stability of analysis of these matrix vector products indicates how the inaccurate computed divided differences can still produce accurate reconstruction i.e., small $|fl(p_n(x_i)) - p_n(x_i)|$
- It also follows that $x_0 < x_1 < \cdots < x_n$ or $x_0 > x_1 > \cdots > x_n$ are “optimal”orderings to keep reconstruction error small.

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

- Surprisingly, the numerical properties of computing the divided differences by the usual recurrence and the evaluating of the Newton form of $p_n(x)$ via Horner's rule are still not completely understood.
- The algorithm of Smoktunowicz et al. (Computing V79, pp. 33-52, 2007) for polynomial interpolation and evaluating the divided differences has an improved stability analysis.
- Their analysis shows backward stability with respect to perturbations to the specified function values.
- For evaluation of the interpolating polynomial it has the $O(n^2)$ computational complexity of Aitken.

Newton Form Interpolation

- The backward stability of the algorithm is independent of the ordering of the points, unlike the standard recurrence form.
- Once the divided differences are evaluated, the polynomial can be evaluated via Horner's rule in $O(n)$ operations but the combination of these two backward stable algorithms is not known to be backward stable for an arbitrary evaluation point x.

Pointwise Error

Let $x_i \in [a, b]$ for $0 \leq i \leq n$ be distinct points, and $p_n(x)$ be the interpolating polynomial of degree n for the function $f(x)$ on $[a, b]$.

The pointwise error is defined for all $x \in [a, b]$ as

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

- $E_n(x_i) = 0$
- $E_n(x) = 0$ if $f(x) \in \mathbb{P}_n$
- $E_n(x)$ tends to be very oscillatory.

Error and Divided Differences

Given nodes $x_0, \ldots, x_n$, an associated interpolant $p_n(t)$, an arbitrary but fixed x all in $[a, b]$, let $p_{n+1}(t)$ be the interpolating polynomial of $f(x)$ on at those $n + 2$ points.

From the Newton form incremental construction, we have

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

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

Note that values of $f[x_0, \ldots, x_n, x]$ for multiple x points other than x_i can be used to estimate the error.

Error and Derivatives

- When $f(x)$ has $n+1$ continuous derivatives and $f^{(n+1)}(x)$ is nicely bounded on $[a,b]$ the error can be bounded.
- Problems if $f^{(n+1)}(x)$ grows faster than $(n+1)!$ or $\omega_{n+1}(x)$ is large.
- This requires relating $f[x_0,\ldots,x_n,t]$ to $f^{(n+1)}(t)$ on $[a,b]$
- In fact, the continuity and differentiability of the divided differences under various assumptions on $f(x)$ is very interesting and useful for approximation, numerical differentiation, numerical quadrature and numerical integration. See for example Chapters 5 and 6 of Isaacson and Keller.

Pointwise Error

Theorem 7.3. *Let* $x_i \in [a,b]$ *for* $0 \le i \le n$ *be distinct points. If* $f(t)$ *is defined on* $[a,b]$ *and* $p_n(t) \in \mathbb{P}_n$ *is the interpolating polynomial of degree* n *defined at the points* x_i *then for any* $x \in [a,b]$

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

If, additionally, $f \in \mathcal{C}^{(n+1)}[a,b]$ *then*

$$\exists \xi(x) \in [a,b] \;\; \textit{such that} \;\; f[x_0, \ldots, x_n, x] = \frac{f^{(n+1)}(\xi(x))}{(n+1)!}$$

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

Pointwise Error and Derivatives

The proof of this is instructive and fairly standard, e.g., see textbook p. 335, or Chapters 5 and 6 of Isaacson and Keller.

Assuming $x_0, \ldots, x_n, x$ are all distinct we have the two interpolating polynomials on $[a, b]$ of degrees n and $n+1$ with the relationship

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

Consider the function on $[a, b]$

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

Pointwise Error and Derivatives

$G(t)$ has at least $n+2$ roots in $[a, b]$ by construction at the x_i and x and since sufficient continuous differentiability is assumed $G'(t)$ has at least $n+1$ roots in $[a, b]$ by Rolle's Theorem. This can be repeated to see that $G^{(n+1)}(t)$ has at least 1 root in $[a, b]$ which we denote ξ. This is clearly dependent on x.

We have

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

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

$$0 = f^{(n+1)}(\xi) - f[x_0, \ldots, x_n, x](n+1)!$$

and the result follows.

Divided Difference Differentiability

- Since $E_n(x_i) = 0$ the result of Theorem 7.3 holds even when $x = x_i$, i.e., ξ is arbitrary then.
- The proof of Theorem 7.3 does not require divided differences, e.g., neither reference proof uses them, but it is a convenient merger of two important results.
- Distinct points are not needed for the correspondence of divided differences and derivatives when they exist.
- This leads to a set of beautiful results that can be found, e.g., in Chapter 6 of Isaacson and Keller.

Nondistinct Points and Differences

Corollary 7.4. *If* $f^{(n)}(x)$ *is continuous in* $[a,b]$ *and* $x_0,\ldots,x_n$ *are in* $[a,b]$ *(not necessarily distinct) then*

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

with $\min(x_0,\ldots,x_n)\le\xi\le\max(x_0,\ldots,x_n)$.

Nondistinct Points and Differences

Corollary 7.5. *If $f^{(n)}(x)$ is continuous in a neighborhood of x then*

$$f[x, \ldots, x] = \frac{f^{(n)}(x)}{n!}$$

with x appearing $n + 1$ times in the argument of the divided difference.