Math 340 • Fall 2026 • Victor Matveev

Polynomial Interpolation Error and Chebyshev Nodes

Let's consider again the error of polynomial interpolation of some smooth function:

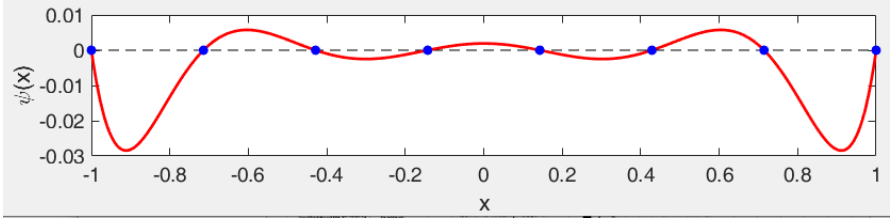
$$f(x) = P_{n-1}(x) + \underbrace{f[x_1,\ldots,x_n,x](x-x_1)\ldots(x-x_n)}_{\text{Remainder / Error}} = P_{n-1}(x) + \underbrace{\frac{f^{(n)}(c)}{n!}\psi(x)}_{\text{Remainder / Error}}$$

Until now, we chose an equidistant partition of an interval of interest as our nodes x_n. However, it turns out that this choice is never optimal in terms of error behavior. To see why, let's plot the polynomial factor in the error term:

$$\psi(x) \equiv (x-x_1)\ldots(x-x_n)$$

Below, we will focus on this so-called "nodal polynomial", even though this is only one of two "parts" of the error.

If points x_n represent an equidistant partition of interval [-1, 1], this contribution to the error looks like this (here $n = 8$):



Note that the absolute magnitude of the nodal polynomial is largest near the edges: this is the reason why interpolation error is usually the greatest near the edges. Moreover, this plot also makes it clear that the largest size of this factor near the edges can be reduced by moving the interpolation nodes near those edges closer to each other.

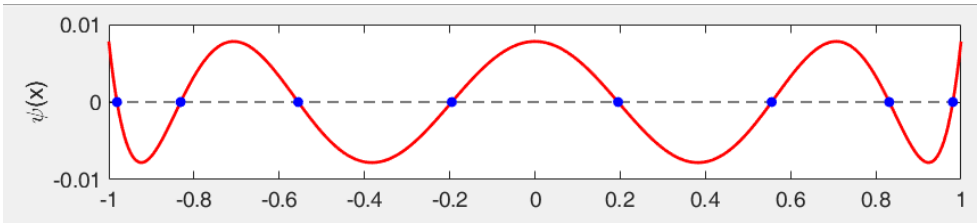
In fact, a very deep and very elegant mathematical result (the Chebyshev equioscillation theorem) says that:

The best mini-max error is always equioscillatory

In other words, all local extreme values of lowest mini-max interpolation error should have the same size.

The reason for this result is simple: if the error is **not** equioscillatory, but achieves the highest value on some of the intervals (like in the figure above), then we can move some of the nodes to decrease those largest errors at the expense of increasing the smaller local extreme values, until these local extreme values all have the same size. When they are indeed all equal, any further adjustment of the node positions will increase at least *some* of the local error extrema.

This is what the **equioscillatory** polynomial factor looks like, for the same number of nodes ($n = 8$) as in the figure above:



The maximal value of $|\psi(x)|$ is smaller by a factor of about 3 compared to the 1st figure above! Note also that this time we didn't pick the endpoints as our nodes, and the values at the endpoints has the same size as its local extreme values.

But what are these optimal node locations in this figure? They are zeros of the so-called Chebyshev polynomial, given by

$$T_n(x) = \cos(n\cos^{-1}x)$$

This polynomial is obviously **equioscillatory** since it is expressible in terms of a **cosine** function, therefore it automatically represents an optimal polynomial $\psi(x)$ of a given order *n* that has the lowest magnitude over the entire interval!

Cosine achieves its zeros at half-integer multiples of π, therefore the zeros of $T_n(x)$ (the nodes in above figure) are given by cosines of equidistant angles breaking the half-circle $0 < \theta < \pi$ into equal "wedges" of size $\frac{\pi}{n}$, as follows:

$$x_k = \cos\left(\left(k - \tfrac{1}{2}\right)\tfrac{\pi}{n}\right), \quad k = 1..n.$$

It may not be obvious that $T_n(x)$ in this trigonometric formula is a polynomial, so let's consider several values of n:

$$T_0(x) = \cos(0) = 1$$
$$T_1(x) = \cos\left(\cos^{-1} x\right) = x$$
$$T_2(x) = \cos\left(2\cos^{-1} x\right) = 2\cos^2\left(\cos^{-1} x\right) - 1 = 2x^2 - 1$$

In fact, simple trigonometric identity for the sum of cosines can be used to prove a recursive relationship:

$$T_{n+1}(x) = 2xT_n(x) - T_{n-1}(x)$$

Note: this relationship is easy to prove by recognizing $T_{n+1}(x) + T_{n-1}(x) = 2xT_n(x)$ as an example of the identity $\cos(\alpha + \beta) + \cos(\alpha - \beta) = 2\cos\alpha\cos\beta$. Therefore, for instance, we have

$$T_3(x) = 2xT_2(x) - T_1(x) = 2x(2x^2 - 1) - x = 4x^3 - 3x\,.$$

It is easy to check that for the case $n = 2$ and $n = 3$, the zeros agree with the general formula on the previous page:

$$n = 2\text{: } \{x_k\} = \left\{\cos\left(\left(k - \tfrac{1}{2}\right)\tfrac{\pi}{2}\right)\right\} = \left\{\cos\left(\tfrac{\pi}{4}\right), \cos\left(\tfrac{3\pi}{4}\right)\right\} = \left\{\pm\tfrac{1}{\sqrt{2}}\right\}, \text{ therefore } 2x_k^2 = 1$$

$$n = 3\text{: } \{x_k\} = \left\{\cos\left(\left(k - \frac{1}{2}\right)\frac{\pi}{3}\right)\right\} = \left\{\cos\left(\frac{\pi}{6}\right), \cos\left(\frac{\pi}{2}\right), \cos\left(\frac{5\pi}{6}\right)\right\} = \left\{\pm\frac{\sqrt{3}}{2}, 0\right\}, \text{ therefore } 4x_k^3 = 3x_k$$

It is important to note that the l**eading coefficient of $T_n(x)$ equals 2^{n-1}**.

On the other hand, the leading coefficient of the nodal function $\psi(x) = (x - x_1)\ldots(x - x_n) = x^n + C_{n-1}x^{n-1} + \cdots$ is by definition equal to 1 (it's a "monic polynomial"), so we should divide $T_n(x)$ by its leading coefficient to obtain this improved new choice for $\psi(x)$:

$$\psi(x) = (x - x_1)(x - x_2)\ldots(x - x_n) = \frac{T_n(x)}{2^{n-1}}$$

Because $|T_n(x)| = |\cos(n\cos^{-1} x)| \le 1$ by simple trigonometry, not only do we obtain the near-optimal node locations, but we simultaneously obtain the **maximal absolute value** of $|\psi(x)|$:

$$\|\psi\|_\infty \equiv \max_{x\in[-1,1]}|\psi| = \frac{\max_{x\in[-1,1]}|T_n|}{2^{n-1}} = \frac{1}{2^{n-1}}$$

Note: the maximal value over an interval is known as the *sup-norm*, or *infinity norm*. This is what we minimize when dealing with numerical error, while the Euclidean L^2 norm (sum of squares) appears when dealing with statistical errors.

To explain the above result in more detail, recall the factorization for any polynomial in terms of its highest coefficient c_n and its zeros x_n:

$$T_n(x) = c_n x^n + c_{n-1}x^{n-1} + \cdots + c_1 x + c_0 = c_n(x - x_1)(x - x_2)\ldots(x - x_n)$$

From this immediately follows that

$$(x - x_1)(x - x_2)\ldots(x - x_n) = \frac{T_n(x)}{c_n}$$

Applying this to the Chebyshev polynomial proves the needed result:

$$|\psi(x)| \equiv |(x - x_1)(x - x_2)\ldots(x - x_n)| = \frac{|T_n(x)|}{c_n} = \frac{|\cos(n\cos^{-1} x)|}{2^{n-1}} \le \frac{1}{2^{n-1}}$$

where the final important step comes from the simple fact that $|\cos\alpha| \le 1$ for *any* angle/argument.

How does this compare with the **equidistant node** case? As it turns out, the polynomial factor does decrease with *n* for that case as well, but at a much slower exponential rate or $(2/e)^n$ instead of $1/2^n$:

$$\left|\left|\psi(x)\right|\right|_\infty \equiv \left|\left|(x-x_1)(x-x_2)\dots(x-x_n)\right|\right|_\infty \approx \frac{C}{\log n}\left(\frac{2}{e}\right)^n, \quad \text{where } C = const$$

Comparing with the result above shows that Chebyshev nodes are clearly superior, since $\frac{2}{e} \approx 0.74 > \frac{1}{2}$.

Relationship between Newton's divided differences and function derivatives

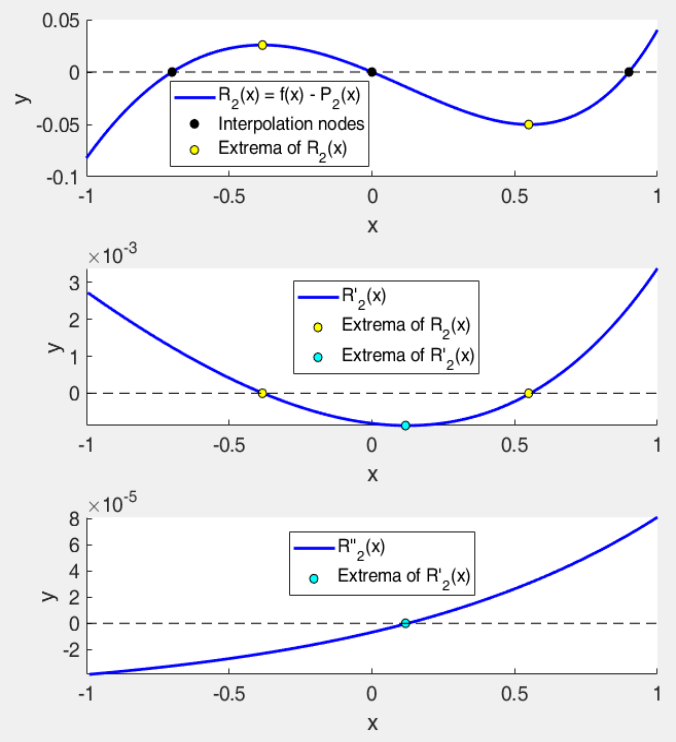
Let us finally prove the relationsip between Newton's divided differences and function derivatives:

$$f[x_1, \dots, x_n] = \frac{f^n(c)}{n!}, \qquad where\ c \in [\min(x_k), \max(x_k)]$$

The proof is very elegant. Without loss of generality, let's consider the case $n = 3$ for the sake of simplicity, and let's focus on the remainder (error) of the interpolant (where the sign is changed for convenience):

$$R_2(x) \equiv P_2(x) - f(x) = f[x_1] + f[x_1, x_2](x - x_1) + f[x_1, x_2, x_3](x - x_1)(x - x_2) - f(x)$$

Let's visualize this interpolation remainder (error), and its two derivatives:



The remainder $R_2(x)$ has 2 local extreme points (yellow circles)

This extreme point are zeros of derivative $R_2'(x)$ plotted in this panel. This derivative therefore has one extreme point, labeled by a cyan circle

This extreme point of $R_2'(x)$ is a zero of the 2nd derivative, $R_2''(x)$

The graphs above show that there exists some point *c* somewhere between interpolation points where

$$R_2''(c) = f''(c) - P_2''(x) = 0$$

Note that 2nd derivative of the 1st two terms in $P_2(x)$ is zero, while in the 3rd term $[(x - x_1)(x - x_2)]'' = 2$, therefore:

$$R_2''(c) = f''(c) - 2f[x_1, x_2, x_3] = 0$$

This proves that $f[x_1, x_2, x_3] = \frac{f''(c)}{2}$ ☺

The proof is analogous for any $n > 2$: differentiate $R_n(x)$ *n* times to obtain $R^{(n)}(c) = 0 \rightarrow f^{(n)}(c) = n!\, f[x_1, \dots, x_n]$