4 Divided and Finite Differences

These are expressions that are helpful in writing, computing and implementing various iterative numerical procedures.

4.1 Divided Differences

4.1.1 Definition and computation

Definition 4.1. *Let $f : [a, b] \to \mathbb{R}$ be a differentiable function on $[a, b]$, and $x_i \in [a, b]$, $i = \overline{0, n}$, be $n + 1$ distinct nodes. The quantity*

$$f[x_0, x_1] = \frac{f(x_1) - f(x_0)}{x_1 - x_0}. \tag{4.1}$$

is called the ***first-order divided difference*** *of f at the nodes x_0 and x_1.*

Remark 4.2.
1. An alternative notation is $[x_0, x_1; f]$.
2. The first-order divided difference of a function can be thought of as a *discrete* version of the derivative.
3. If we consider $f[x_0] = f(x_0)$ the *divided difference of order* 0 at x_0, then (4.1) can be written as

$$f[x_0, x_1] = \frac{f[x_1] - f[x_0]}{x_1 - x_0}. \tag{4.2}$$

We define higher-order divided differences recursively using lower-order ones.

Definition 4.3. *The* ***divided difference of order*** *n of f at the distinct nodes $x_0, x_1, \dots, x_n$ is the quantity*

$$f[x_0, x_1, \dots, x_n] = \frac{f[x_1, x_2, \dots, x_n] - f[x_0, x_1, \dots, x_{n-1}]}{x_n - x_0}. \tag{4.3}$$

Remark 4.4.
1. The denominator in (4.3) is the difference between the uncommon nodes of the differences at the numerator.
2. For easy computation (and implementation) of divided differences, we generate the *table of divided differences*, illustrated below for 4 nodes. The divided differences are obtained on the first row.

$$
\begin{array}{cccccccc}
x_0 & f[x_0] & \longrightarrow & f[x_0,x_1] & \longrightarrow & f[x_0,x_1,x_2] & \longrightarrow & f[x_0,x_1,x_2,x_3] \\
 & & \nearrow & & \nearrow & & \nearrow & \\
x_1 & f[x_1] & \longrightarrow & f[x_1,x_2] & \longrightarrow & f[x_1,x_2,x_3] & & \\
 & & \nearrow & & \nearrow & & & \\
x_2 & f[x_2] & \longrightarrow & f[x_2,x_3] & & & & \\
 & & \nearrow & & & & & \\
x_3 & f[x_3] & & & & & &
\end{array}
$$

Example 4.5. Let $f(x) = \sin \pi x$, and the nodes $x_0 = 0, x_1 = \dfrac{1}{6}, x_2 = \dfrac{1}{2}$. Let us construct the divided difference table.

Solution

$$
\begin{array}{cccccc}
x_0 = 0 & f[x_0] = 0 & \longrightarrow & f[x_0,x_1] = \dfrac{1/2-0}{1/6-0} = 3 & \longrightarrow & f[x_0,x_1,x_2] = \dfrac{3/2-3}{1/2-0} = -3 \\
 & & \nearrow & & \nearrow & \\
x_1 = 1/6 & f[x_1] = 1/2 & \longrightarrow & f[x_1,x_2] = \dfrac{1-1/2}{1/2-1/6} = 3/2 & & \\
 & & \nearrow & & & \\
x_2 = 1/2 & f[x_2] = 1 & & & &
\end{array}
$$

■

Divided differences with multiple nodes

Divided differences with multiple nodes can be expressed in terms of the derivatives of the function f, as follows

$$f[x_0, x_0] = \lim_{x_1 \to x_0} f[x_0, x_1] = \lim_{x_1 \to x_0} \frac{f(x_1) - f(x_0)}{x_1 - x_0} = f'(x_0),$$

In general, the **divided difference of order** n at the node x_0, of multiplicity $n+1$, is defined as

$$f[x_0, x_0, \ldots, x_0] \quad = \quad \frac{f^{(n)}(x_0)}{n!}, \tag{4.4}$$

and further, for mixed nodes (some simple, some multiple), we use definition 4.3.

Example 4.6. Let us see the divided difference table for 3 double nodes.

Solution We have the nodes x_0, x_1, x_2 and the values $f(x_i), f'(x_i),\ i = 0, 1, 2$. We define the sequence of nodes $z_0, z_1, \ldots, z_5$ by

$$z_{2i} = z_{2i+1} = x_i, i = 0, 1, 2.$$

We build the divided difference table relative to the nodes $z_i, i = \overline{0, 5}$.
Since $z_{2i} = z_{2i+1} = x_i$ for every $i = 0, 1, 2$, $f[z_{2i}, z_{2i+1}] = f[x_i, x_i]$ is a divided difference with a double node and it is equal to $f'(x_i)$; therefore we will use $f'(x_0), f'(x_1), f'(x_2)$ instead of first order divided differences $f[z_0, z_1], f[z_2, z_3], f[z_4, z_5]$.

$$\begin{array}{llclcll}
z_0 = x_0 & f[z_0] & \longrightarrow & f[z_0, z_1] = f'(x_0) & \longrightarrow & f[z_0, z_1, z_2] & \ldots \\
 & & \nearrow & & \nearrow & & \\
z_1 = x_0 & f[z_1] & \longrightarrow & f[z_1, z_2] = \dfrac{f(z_2) - f(z_1)}{z_2 - z_1} & \longrightarrow & f[z_1, z_2, z_3] & \ldots \\
 & & \nearrow & & \nearrow & & \\
z_2 = x_1 & f[z_2] & \longrightarrow & f[z_2, z_3] = f'(x_1) & \longrightarrow & f[z_2, z_3, z_4] & \ldots \\
 & & \nearrow & & \nearrow & & \\
z_3 = x_1 & f[z_3] & \longrightarrow & f[z_3, z_4] = \dfrac{f(z_4) - f(z_3)}{z_4 - z_3} & \longrightarrow & f[z_3, z_4, z_5] & \ldots \\
 & & \nearrow & & \nearrow & & \\
z_4 = x_2 & f[z_4] & \longrightarrow & f[z_4, z_5] = f'(x_2) & & & \\
 & & \nearrow & & & & \\
z_5 = x_2 & f[z_5] & & & & &
\end{array}$$

■

4.1.2 Properties of divided differences

Theorem 4.7. *Divided differences have a number of special properties that can simplify work with them:*
a)

$$f[x_0, x_1, \ldots, x_n] \;=\; \sum_{i=0}^{n} \frac{f(x_i)}{u'(x_i)} = \sum_{i=0}^{n} \frac{f(x_i)}{u_i(x_i)}, \tag{4.5}$$

where $u(x) = (x - x_0)(x - x_1)\dots(x - x_n)$ *and* $u_i(x) = \dfrac{u(x)}{x - x_i}$.

b) *For any permutation* $\{i_0, i_1, \dots, i_n\}$ *of the integers* $\{0, 1, \dots, n\}$,

$$f[x_{i_0}, x_{i_1}, \dots, x_{i_n}] \;\; = \;\; f[x_0, x_1, \dots, x_n]. \tag{4.6}$$

c)

$$f[x_0, x_1, \dots, x_n] \;\; = \;\; \frac{(Wf)(x_0, x_1, \dots, x_n)}{V(x_0, x_1, \dots, x_n)}, \tag{4.7}$$

where

$$Wf(x_0, x_1, \dots, x_n) \;\; = \;\; \begin{vmatrix} 1 & x_0 & x_0^2 & \dots & x_0^{n-1} & f(x_0) \\ 1 & x_1 & x_1^2 & \dots & x_1^{n-1} & f(x_1) \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ 1 & x_n & x_n^2 & \dots & x_n^{n-1} & f(x_n) \end{vmatrix} \tag{4.8}$$

and $V(x_0, x_1, \dots, x_n)$ *is the Vandermonde determinant.*

d) *Let* $e_k(x) = x^k, k \geq 0$. *Then,*

$$e_k[x_0, x_1, \dots, x_n] \;\; = \;\; \begin{cases} 0, & k < n \\ 1, & k = n \end{cases}.$$

For a polynomial of degree k, $P_k = a_0 + a_1 x + \dots + a_k x^k$,

$$P_k[x_0, x_1, \dots, x_n] \;\; = \;\; \begin{cases} 0, & k < n \\ a_n, & k = n \end{cases}. \tag{4.9}$$

e) *If* $f \in C^n[a, b]$, *where* $[a, b]$ *is the smallest interval containing the distinct nodes* $\{x_0, \dots, x_n\}$, *then there exists* $\xi_n \in (a, b)$ *such that*

$$f[x_0, x_1, \dots, x_n] \;\; = \;\; \frac{1}{n!} f^{(n)}(\xi_n), \tag{4.10}$$

(the mean-value formula for divided differences).

Proof. (Selected)

a) First off, let us notice that

$$u'(x) \;=\; \sum_{i=0}^{n} u_i(x)$$

and $u_i(x_j) = 0, i \neq j$, from where it follows that

$$u'(x_i) \;=\; u_i(x_i),\; i = 0, 1, \ldots, n.$$

We prove (4.5) by induction on n. Let us look at several cases.

Although it's a convention, let us start with $\boldsymbol{n = 0}$. In this case,

$$\begin{aligned} u(x) &= x - x_0, \\ u_0(x) &= 1,\; u_0(x_0) \;=\; 1, \\ u'(x) &= 1,\; u'(x_0) \;= 1 \end{aligned}$$

and

$$f[x_0] \;= f(x_0) \;=\; \frac{f(x_0)}{u_0(x_0)} \;= \sum_{i=0}^{0} \frac{f(x_i)}{u_i(x_i)},$$

so (4.5) is trivially true.

$\boldsymbol{n = 1}$. We have

$$\begin{aligned} u(x) &= (x - x_0)(x - x_1), \\ u_0(x) &= x - x_1,\; u_0(x_0) \;=\; x_0 - x_1, \\ u_1(x) &= x - x_0,\; u_1(x_1) \;=\; x_1 - x_0, \\ u'(x) &= (x - x_1) + (x - x_0), \\ u'(x_0) &= x_0 - x_1 \;=\; u_0(x_0), \\ u'(x_1) &= x_1 - x_0 \;=\; u_1(x_1). \end{aligned}$$

The first-order divided difference is

$$\begin{aligned} f[x_0, x_1] &= \frac{f[x_1] - f[x_0]}{x_1 - x_0} \;=\; \frac{f(x_1) - f(x_0)}{x_1 - x_0} \\ &= \frac{f(x_0)}{x_0 - x_1} + \frac{f(x_1)}{x_1 - x_0} \;=\; \frac{f(x_0)}{u_0(x_0)} + \frac{f(x_1)}{u_1(x_1)}. \end{aligned}$$

To better understand the induction step, let us consider one more case, $n = 2$.

$\boldsymbol{n = 2}$. Now, $u(x) = (x - x_0)(x - x_1)(x - x_2)$. The second-order divided difference can be written as

$$\begin{aligned}
f[x_0, x_1, x_2] &= \frac{f[x_1, x_2] - f[x_0, x_1]}{x_2 - x_0} \\
&= \frac{1}{x_2 - x_0}\left[\frac{f(x_1)}{x_1 - x_2} + \frac{f(x_2)}{x_2 - x_1} - \frac{f(x_0)}{x_0 - x_1} - \frac{f(x_1)}{x_1 - x_0}\right] \\
&= \frac{f(x_0)}{u_0(x_0)} + \frac{f(x_2)}{u_2(x_2)} + \frac{f(x_1)}{x_2 - x_0}\left[\frac{1}{x_1 - x_2} - \frac{1}{x_1 - x_0}\right] \\
&= \frac{f(x_0)}{u_0(x_0)} + \frac{f(x_2)}{u_2(x_2)} + \frac{f(x_1)}{x_2 - x_0} \cdot \frac{x_1 - x_0 - x_1 + x_2}{u_1(x_1)} \\
&= \sum_{i=0}^{2} \frac{f(x_i)}{u_i(x_i)}.
\end{aligned}$$

Assume (4.5) is true for $\boldsymbol{n - 1}$. For the case $\boldsymbol{n}$, we have

$$u(x) = (x - x_0)(x - x_1)\dots(x - x_{n-1})(x - x_n),$$

and to simplify the writing, for $u_i(x)$ we use the notation

$$u_i(x) = (x - x_0)\dots \cdot / \cdot \dots (x - x_n).$$

Then,

$$\begin{aligned}
f[x_0, \dots, x_n] &= \frac{f[x_1, \dots, x_n] - f[x_0, \dots, x_{n-1}]}{x_n - x_0} \\
&= \frac{1}{x_n - x_0}\left[\sum_{i=1}^{n-1} \frac{f(x_i)}{(x_i - x_1)\dots \cdot / \cdot \dots (x_i - x_n)} + \frac{f(x_n)}{(x_n - x_1)\dots(x_n - x_{n-1})}\right. \\
&\quad \left. - \frac{f(x_0)}{(x_0 - x_1)\dots(x_0 - x_{n-1})} - \sum_{i=1}^{n-1} \frac{f(x_i)}{(x_i - x_0)\dots \cdot / \cdot \dots (x_i - x_{n-1})}\right] \\
&= \frac{f(x_0)}{u_0(x_0)} + \frac{f(x_n)}{u_n(x_n)} + \frac{1}{x_n - x_0} \sum_{i=1}^{n-1} \frac{f(x_i)}{u_i(x_i)} \cdot (x_i - x_0 - x_i + x_n),
\end{aligned}$$

hence,

$$f[x_0,\ldots,x_n] \;=\; \sum_{i=0}^{n} \frac{f(x_i)}{u_i(x_i)}.$$

b) This follows directly from part a), since the right-hand-side of (4.5) is symmetric with respect to the indexes $\{1,2,\ldots,n\}$.

c) Let us recall that for Vandermonde determinants, the following are true:

$$V(x_0,x_1,\ldots,x_n) \;=\; \begin{vmatrix} 1 & x_0 & x_0^2 & \ldots & x_0^n \\ 1 & x_1 & x_1^2 & \ldots & x_1^n \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & x_n & x_n^2 & \ldots & x_n^n \end{vmatrix} = \prod_{0\le j<i\le n} (x_i - x_j)$$

and

$$V(x_0,\ldots,/,\ldots,x_n) \;=\; \frac{V(x_0,x_1,\ldots,x_n)}{(x_i-x_0)\ldots/\ldots(x_i-x_n)} \;=\; (-1)^{n+i}\frac{V(x_0,x_1,\ldots,x_n)}{u_i(x_i)}.$$

The idea for the proof of (4.7) is to expand (Wf) over the elements of the last column, to get

$$\begin{aligned}(Wf)(x_0,x_1,\ldots,x_n) &= V(x_0,x_1,\ldots,x_n)\sum_{i=0}^{n}\frac{f(x_i)}{u_i(x_i)} \\ &= V(x_0,x_1,\ldots,x_n) f[x_0,\ldots,x_n],\end{aligned}$$

by part a).

d) For $f(x) = e_k(x) = x^k$, we have

$$(We_k)(x_0,x_1,\ldots,x_n) \;=\; \begin{vmatrix} 1 & x_0 & x_0^2 & \ldots & x_0^{n-1} & x_0^k \\ 1 & x_1 & x_1^2 & \ldots & x_1^{n-1} & x_1^k \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ 1 & x_n & x_n^2 & \ldots & x_n^{n-1} & x_n^k \end{vmatrix} = \begin{cases} 0, & k<n \\ V(x_0,x_1,\ldots,x_n), & k=n \end{cases},$$

so, by part c),

$$e_k[x_0,x_1,\ldots,x_n] \;=\; \begin{cases} 0, & k<n \\ 1, & k=n \end{cases}.$$

Then, for $P_k = a_0 + a_1 x + \cdots + a_k x^k$,

$$P_k[x_0, x_1, \ldots, x_n] \;=\; (a_0 + a_1 x + \cdots + a_k x^k)[x_0, x_1, \ldots, x_n] \;=\; \begin{cases} 0, & k < n \\ a_n, & k = n \end{cases} .$$

□

Remark 4.8. As a consequence of part e), if $f \in C^n[a, b]$ and $\alpha \in [a, b]$, then

$$\lim_{x_0, \ldots, x_n \to \alpha} f[x_0, x_1, \ldots, x_n] \;=\; \lim_{\xi_n \to \alpha} \frac{f^{(n)}(\xi_n)}{n!} \;=\; \frac{1}{n!} f^{(n)}(\alpha),$$

the computational formula for divided differences with multiple nodes.

4.2 Finite Differences

Definition 4.9. *Consider the equidistant nodes* $x_i = x_0 + ih, i = 0, 1, \ldots, n,\ h > 0$*.*
The quantity

$$\Delta^1 f(x_i) \;=\; f(x_{i+1}) - f(x_i) \;=\; f_{i+1} - f_i \tag{4.11}$$

is called the ***first-order forward difference*** *of* f *with step* h *at* x_i*, and*

$$\Delta^k f(x_i) \;=\; \Delta^{k-1} f(x_{i+1}) - \Delta^{k-1} f(x_i) \;=\; \Delta^{k-1} f_{i+1} - \Delta^{k-1} f_i \tag{4.12}$$

is the k***th-order forward difference*** *of* f *with step* h*, at* x_i*.*

Remark 4.10.
1. As a convention, we use $\Delta^0 f(x_i) = f(x_i) = f_i$.
2. For easy computation (and implementation) of forward differences, we construct a *table of forward differences*, similar to the one used for divided differences, illustrated below for 4 nodes.

$$
\begin{array}{c|ccccccc}
x_0 & f_0 & \longrightarrow & \Delta f_0 & \longrightarrow & \Delta^2 f_0 & \longrightarrow & \Delta^3 f_0 \\
 & & \nearrow & & \nearrow & & \nearrow & \\
x_1 & f_1 & \longrightarrow & \Delta f_1 & \longrightarrow & \Delta^2 f_1 & & \\
 & & \nearrow & & \nearrow & & & \\
x_2 & f_2 & \longrightarrow & \Delta f_2 & & & & \\
 & & \nearrow & & & & & \\
x_3 & f_3 & & & & & &
\end{array}
$$

In a similar way, we define the **backward difference** ∇ by

$$
\begin{aligned}
\nabla^0 f_i &= f_i, \\
\nabla^1 f_i &= f_i - f_{i-1}, \\
\nabla^k f_i &= \nabla^{k-1} f_i - \nabla^{k-1} f_{i-1},
\end{aligned}
\tag{4.13}
$$

and they can also be easily computed in a table.

$$
\begin{array}{c|ccccccc}
x_0 & f_0 & & & & & & \\
 & & \searrow & & & & & \\
x_1 & f_1 & \longrightarrow & \nabla f_1 & & & & \\
 & & \searrow & & \searrow & & & \\
x_2 & f_2 & \longrightarrow & \nabla f_2 & \longrightarrow & \nabla^2 f_2 & & \\
 & & \searrow & & \searrow & & \searrow & \\
x_3 & f_3 & \longrightarrow & \nabla f_3 & \longrightarrow & \nabla^2 f_3 & \longrightarrow & \nabla^3 f_3
\end{array}
$$

Remark 4.11. These differences are referred to collectively as *finite differences*. Usually, if nothing is specified, by "finite" differences we mean "forward" differences.

Denote by

$$X = \{x_i \mid x_i = x_0 + ih,\ i = \overline{0,n}, x_0, h \in \mathbb{R}\}$$

and for $f : X \to \mathbb{R}$, by $f_i = f(x_i)$.

Let us write a few finite differences and notice a pattern.

$$\begin{aligned}
\Delta^1 f(x_i) &= f_{i+1} - f_i, \\
\Delta^2 f(x_i) &= \Delta^1 f_{i+1} - \Delta^1 f_i \;=\; f_{i+2} - f_{i+1} - \big(f_{i+1} - f_i\big) \;=\; f_{i+2} - 2f_{i+1} + f_i, \\
\Delta^3 f(x_i) &= \Delta^2 f_{i+1} - \Delta^2 f_i \;=\; f_{i+3} - 2f_{i+2} + f_{i+1} - \big(f_{i+2} - 2f_{i+1} + f_i\big) \\
&= f_{i+3} - 3f_{i+2} + 3f_{i+1} - f_i.
\end{aligned}$$

It can easily be proved (by induction) that

$$\Delta^n f(x_i) \;=\; \sum_{k=0}^{n} (-1)^k \binom{n}{k} f_{n-k+i},$$

or, equivalently, by the symmetry of combinations,

$$\Delta^n f(x_i) \;=\; \sum_{k=0}^{n} (-1)^{n-k} \binom{n}{k} f_{k+i}.$$

In particular, we have

$$\Delta^n f(x_0) \;=\; \sum_{k=0}^{n} (-1)^{n-k} \binom{n}{k} f_k. \tag{4.14}$$

Finite and divided differences for equally spaced nodes are closely related.

Proposition 4.12. *Let* $f : X \to \mathbb{R}$. *Then*

$$f[a, a+h, \ldots, a+nh] \;=\; \frac{1}{n!h^n} \Delta^n f(a). \tag{4.15}$$

Chapter 2. Function Approximation

Approximation of functions is one of the most important tasks in Numerical Analysis.

Most functions encountered in mathematical problems and applications cannot be evaluated exactly, even though we usually handle them as if they where completely known quantities. The simplest and most important of these are $\sqrt{x}, e^x, \log x$, and the trigonometric functions; and there are many other functions that occur commonly in physics, engineering, and other disciplines. In evaluating functions, by hand or using a computer, we are essentially limited to the elementary arithmetic operations $+, -, \times$ and $\div$. Combining these operations means that we can evaluate polynomials and rational functions, which are polynomials divided by polynomials. All other functions must be evaluated by using approximations based on polynomials or rational functions, including piecewise variants of them (e.g., spline functions). Although rational functions generally give slightly more efficient approximations, polynomials are adequate for most problems and their theory is much easier to work with.

Interpolation is the process of finding and evaluating a function whose graph goes through a set of given points. The points may arise as measurements in a physical problem, or they may be obtained from a known function. The interpolating function is usually chosen from a restricted class of functions and *polynomials* are the most commonly used class.

Interpolation is an important tool in producing computable approximations to commonly used functions. Moreover, to numerically integrate or differentiate a function, we often replace the function with a simpler approximating expression, and it is then integrated or differentiated. These simpler expressions are almost always obtained by interpolation. Also, some of the most widely used numerical methods for solving differential equations are obtained from interpolating approximations. Finally, interpolation is widely used in computer graphics, to produce smooth curves and surfaces when the geometric object of interest is given at only a discrete set of data points.

1 Polynomial Interpolation

Interpolation problem. Given $n+1$ distinct points – called *nodes (or knots)* – $x_i \in [a, b], i = \overline{0, n}$ and the values $f(x_i) = y_i$ of an unknown function $f : [a, b] \to \mathbb{R}$, find a polynomial $P(x)$ of minimum degree, satisfying

$$P(x_i) \quad = \quad f(x_i),\ i = \overline{0, n}, \tag{1.1}$$

called *interpolation conditions*. This polynomial approximates function f.

1.1 Lagrange Interpolation

Linear interpolation

We start with a simple case: consider two interpolation nodes, $(x_0, y_0), (x_1, y_1), x_0 \neq x_1$.
We know that there is a unique *line* passing through these points. That means we can find a polynomial of degree 1 that interpolates the data. Let us find it.

The slope of the line is

$$m = \frac{y_1 - y_0}{x_1 - x_0}$$

and its equation is

$$y - y_0 = \frac{y_1 - y_0}{x_1 - x_0}(x - x_0).$$

We find the linear interpolation polynomial as

$$\begin{aligned} P_1(x) &= y_0 + \frac{y_1 - y_0}{x_1 - x_0}(x - x_0) \\ &= \left(1 - \frac{x - x_0}{x_1 - x_0}\right) y_0 + \frac{x - x_0}{x_1 - x_0} y_1 \\ &= \frac{x - x_1}{x_0 - x_1} y_0 + \frac{x - x_0}{x_1 - x_0} y_1. \end{aligned}$$

Example 1.1. Consider the function $f(x) = \sqrt{x}$ and the nodes $x_0 = 1, x_1 = 4$, i.e. the data $(1, 1), (4, 2)$.

Solution The linear interpolation polynomial is

$$P_1(x) = \frac{1}{3}x + \frac{2}{3}.$$

The graphs of f and P_1 on the interval $[0, 15]$ are shown in Figure 1.

■

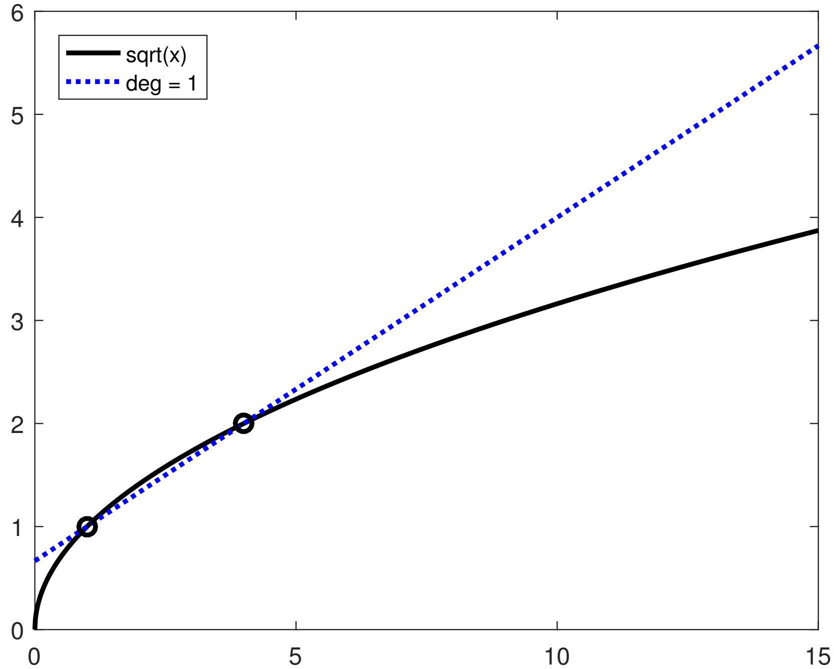

Fig. 1: Linear interpolation of function $\sqrt{x}$

Quadratic interpolation

We go further and consider 3 distinct nodes $(x_0, y_0), (x_1, y_1), (x_2, y_2)$. It can easily be checked that the quadratic polynomial

$$P_2(x) \quad = \quad \frac{(x-x_1)(x-x_2)}{(x_0-x_1)(x_0-x_2)}y_0 + \frac{(x-x_0)(x-x_2)}{(x_1-x_0)(x_1-x_2)}y_1 + \frac{(x-x_0)(x-x_1)}{(x_2-x_0)(x_2-x_1)}y_2$$

interpolates these data.

Example 1.2. In Example 1.1 we add the node $(9, 3)$. The graphs of f and the two interpolation polynomials P_1 and P_2 are plotted in Figure 2.

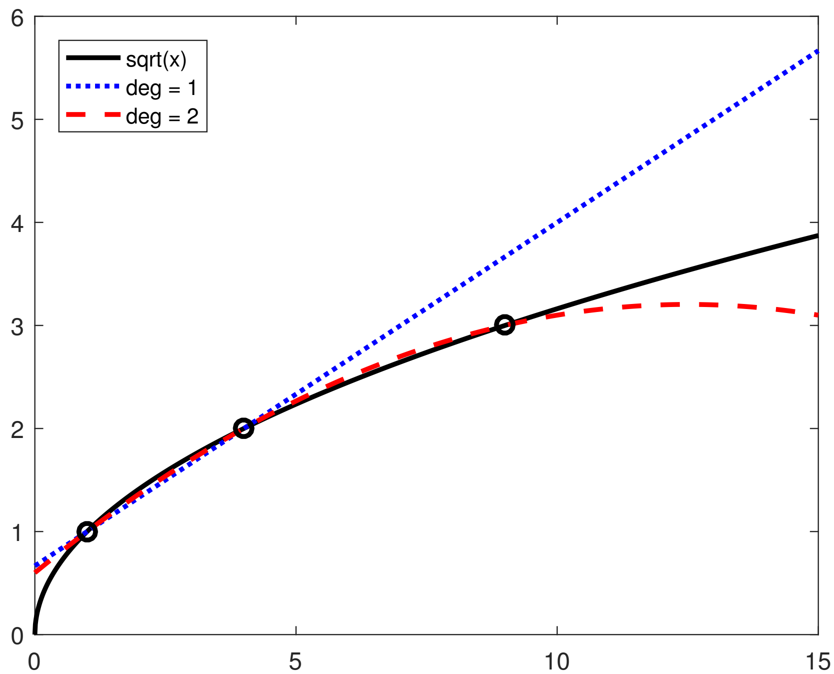

Fig. 2: Linear and quadratic interpolation of function $\sqrt{x}$

General case

Consider the interval $[a, b] \subset \mathbb{R}$, a function $f : [a, b] \to \mathbb{R}$ and a set of $n + 1$ distinct nodes $\{x_0, x_1, \ldots, x_n\} \subset [a, b]$.
Recall the notations

$$\begin{aligned} u(x) &= \prod_{j=0}^{n}(x - x_j), \\ u_j(x) &= \frac{u(x)}{x - x_j},\ j = 0, 1, \ldots, n. \end{aligned} \tag{1.2}$$

Theorem 1.3. *There is a unique polynomial $L_n f$ of degree at most n, satisfying the interpolation conditions* (1.1)*. This polynomial can be written as*

$$L_n f(x) = \sum_{i=0}^{n} l_i(x) f(x_i), \tag{1.3}$$

where

$$l_i(x) \;=\; \prod_{\substack{j=0\\ j\neq i}}^{n} \frac{x-x_j}{x_i-x_j} \;=\; \frac{u_i(x)}{u_i(x_i)} \;=\; \frac{u_i(x)}{u'(x_i)}. \tag{1.4}$$

*$L_n f$ is called the **Lagrange interpolation polynomial** of f at the nodes $x_0, x_1, \ldots, x_n$. The functions $l_i(x), i = \overline{0,n}$ are called **Lagrange fundamental (basis) polynomials** associated with these points.*

Proof. It can easily be checked that l_i is a polynomial of degree at most n and that

$$l_i(x_j) \;= \delta_{ij} = \begin{cases} 0, & i \neq j \\ 1, & i = j \end{cases}.$$

Hence, the polynomial $L_n f$ defined in(1.3) is also a polynomial of degree at most n and it satisfies conditions (1.1).

To prove uniqueness, assume there exists another polynomial P_n^* (of degree at most n) satisfying conditions (1.1) and consider

$$Q_n \;=\; L_n - P_n^*.$$

By (1.1), $Q_n(x_i) = 0, i = 0, \ldots, n$, which means Q_n, a polynomial of degree at most n, has $n+1$ distinct roots. By the Fundamental Theorem of Algebra, Q_n must be identically zero, thus proving the uniqueness of L_n.

□

Error and convergence

We want to use the approximation

$$f(x) \;\approx\; L_n f(x),\; x \in [a,b].$$

To this end, we must assess (bound) the error (the remainder)

$$(R_n f)(x) \;=\; f(x) - (L_n f)(x),\; x \in [a,b]. \tag{1.5}$$

Theorem 1.4. *Let $[a,b] \subset \mathbb{R}$, $f : [a,b] \to \mathbb{R}$ a function of class $C^{n+1}[a,b]$ and consider the distinct nodes $\{x_0, x_1, \ldots, x_n\} \subset [a,b]$. Then there exists $\xi \in (a,b)$ such that*

$$(R_n f)(x) \;=\; \frac{u(x)}{(n+1)!} f^{(n+1)}(\xi). \tag{1.6}$$

Example 1.5. For linear and quadratic interpolation, the remainders are given by

$$\begin{aligned}(R_1f)(x) &= \frac{(x-x_0)(x-x_1)}{2}f''(\xi),\\ (R_2f)(x) &= \frac{(x-x_0)(x-x_1)(x-x_2)}{6}f'''(\xi).\end{aligned}$$

For $f(x)=\sqrt{x}=x^{1/2}$, the derivatives are

$$\begin{aligned}f'(x) &= \frac{1}{2}x^{-1/2} = \frac{1}{2}\cdot\frac{1}{\sqrt{x}},\\ f''(x) &= \frac{1}{2}\left(-\frac{1}{2}\right)x^{-3/2} = -\frac{1}{4}\cdot\frac{1}{x\sqrt{x}},\\ f'''(x) &= -\frac{1}{4}\left(-\frac{3}{2}\right)x^{-5/2} = \frac{3}{8}\cdot\frac{1}{x^2\sqrt{x}}.\end{aligned}$$

So, for the remainders, we have

$$\begin{aligned}|(R_1f)(x)| &= \frac{|(x-x_0)(x-x_1)|}{8}\cdot\frac{1}{\xi\sqrt{\xi}},\\ |(R_2f)(x)| &= \frac{3}{8}\frac{|(x-x_0)(x-x_1)(x-x_2)|}{6}\cdot\frac{1}{\xi^2\sqrt{\xi}} = \frac{|(x-x_0)(x-x_1)(x-x_2)|}{16}\cdot\frac{1}{\xi^2\sqrt{\xi}}.\end{aligned}$$

Remark 1.6. In general, an upper bound of the interpolation error is given by

$$|(R_nf)(x)| \le \frac{|u(x)|}{(n+1)!}M_{n+1}(f), \tag{1.7}$$

where

$$M_{n+1}(f) = \sup_{t\in[a,b]}|f^{(n+1)}(t)|.$$

The term $|u(x)|$ can be minimized by a suitable choice of the nodes.
Assume the interval is $[-1,1]$ (then, for a general interval $[a,b]$, we use the linear change of variables $x=\frac{b-a}{2}t+\frac{b+a}{2},\ t\in[-1,1],\ x\in[a,b]$).

An optimal choice of nodes are the roots of the **Chebyshev polynomial of the first kind**:

$$T_m(x) = \cos(m\arccos x). \tag{1.8}$$

With the change of variables $x = \cos t, t \in [0, \pi]$, we get

$$\begin{aligned} T_m(x) &= \cos(mt) = \frac{1}{2}\left(e^{imt} + e^{-imt}\right) \\ &= \frac{1}{2}\left[(\cos t + i \sin t)^m + (\cos t - i \sin t)^m\right] \\ &= \frac{1}{2}\left[\left(x + i\sqrt{1-x^2}\right)^m + \left(x - i\sqrt{1-x^2}\right)^m\right]. \end{aligned}$$

The odd powers of the radical will be canceled, resulting in a polynomial of degree m in x, with leading coefficient 2^{m-1}.

Remark 1.7. Chebyshev polynomials of the first kind have some remarkable properties:
1. Polynomials of degree $0, 1, 2$ and 3 are easily computable using trigonometric identities. They are

$$\begin{aligned} T_0(x) &= 1, \\ T_1(x) &= x, \\ T_2(x) &= 2x^2 - 1, \\ T_3(x) &= 4x^3 - 3x. \end{aligned}$$

2. Higher degree polynomials can be obtained from the recurrence relation

$$T_{k+1}(x) = 2xT_k(x) - T_{k-1}(x), \; k = 1, 2, \ldots.$$

For example, the next polynomial is

$$T_4(x) = 8x^4 - 8x^2 + 1.$$

3. $\{T_n(x)\}_{n\in\mathbb{N}}$ is a sequence of *orthogonal* polynomials on $(-1, 1)$ with respect to the weight function $w(x) = \dfrac{1}{\sqrt{1-x^2}}$, i.e.

$$\int_{-1}^{1} \frac{T_n(x)T_m(x)}{\sqrt{1-x^2}}\, dx = 0, \; n \neq m.$$

4. To minimize the term $|u(x)|$ in (1.7) on the interval $[-1,1]$, we choose

$$u(x) = \widetilde{T}_{n+1}(x) = \frac{1}{2^n} T_{n+1}(x),$$

i.e. the nodes

$$x_k = \cos \frac{2k-1}{2n}\pi, \ k = 0, 1, \ldots, n. \tag{1.9}$$

In this case, we have

$$||R_n f|| \leq \frac{1}{2^n (n+1)!} ||f^{(n+1)}||.$$

For the general case, on the interval $[a,b]$, we take

$$u(x) = \widetilde{T}_{n+1}(x; a, b) = \frac{(b-a)^{n+1}}{2^{2n+1}} \cos\left((n+1) \arccos \frac{2x-a-b}{b-a}\right)$$

and for the remainder, we have

$$||R_n f|| \leq \frac{(b-a)^{n+1}}{2^{2n+1}(n+1)!} ||f^{(n+1)}||.$$

Regarding the convergence of the Lagrange polynomial $L_n f$ to f, this *does not* always happen. The polynomial does converge, if, for instance, $f \in C^\infty[a,b]$, with $|f^k(x)| \leq M_k, \ \forall x \in [a,b], k = 0, 1, 2, \ldots$ and satisfies

$$\lim_{k\to\infty} \frac{(b-a)^k}{k!} M_k = 0.$$

Example 1.8 (**Runge's Example**). Consider the function

$$f(x) = \frac{1}{1+x^2}, \ x \in [-5, 5]$$

and the equally spaced nodes

$$x_k^{(n)} = -5 + 10\,\frac{k}{n}, \ k = \overline{0, n}.$$

It can be shown that

$$\lim_{n\to\infty} |f(x) - L_n f(x)| = \begin{cases} 0, & \text{if} \quad |x| < 3.633\ldots \\ \infty, & \text{if} \quad |x| > 3.633\ldots \end{cases}$$

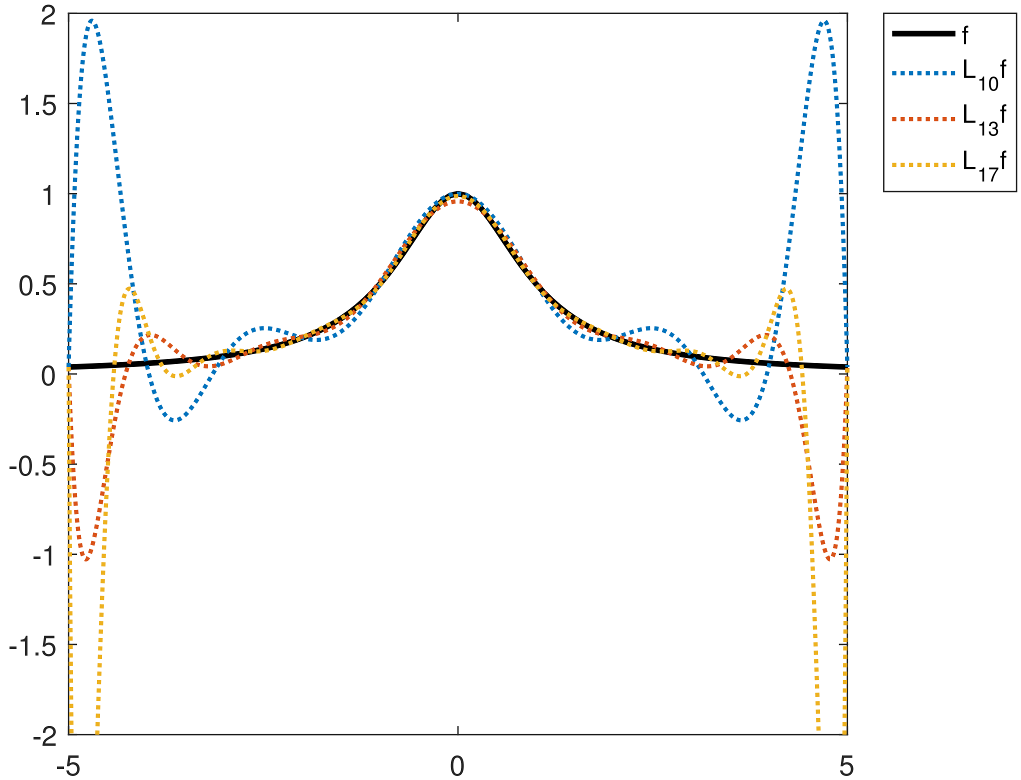

Fig. 3: Runge's Example, $n = 10, 13, 17$