

Jim Lambers
MAT 460/560
Fall Semester 2009-10
Lecture 23 Notes

These notes correspond to Section 4.1 in the text.

Numerical Differentiation

We now discuss the other fundamental problem from calculus that frequently arises in scientific applications, the problem of computing the derivative of a given function $f(x)$.

Finite Difference Approximations

Recall that the derivative of $f(x)$ at a point x_0 , denoted $f'(x_0)$, is defined by

$$f'(x_0) = \lim_{h \rightarrow 0} \frac{f(x_0 + h) - f(x_0)}{h}.$$

This definition suggests a method for approximating $f'(x_0)$. If we choose h to be a small positive constant, then

$$f'(x_0) \approx \frac{f(x_0 + h) - f(x_0)}{h}.$$

This approximation is called the *forward difference formula*.

To estimate the accuracy of this approximation, we note that if $f''(x)$ exists on $[x_0, x_0 + h]$, then, by Taylor's Theorem, $f(x_0 + h) = f(x_0) + f'(x_0)h + f''(\xi)h^2/2$, where $\xi \in [x_0, x_0 + h]$. Solving for $f'(x_0)$, we obtain

$$f'(x_0) = \frac{f(x_0 + h) - f(x_0)}{h} + \frac{f''(\xi)}{2}h,$$

so the error in the forward difference formula is $O(h)$. We say that this formula is *first-order accurate*.

The forward-difference formula is called a *finite difference approximation* to $f'(x_0)$, because it approximates $f'(x)$ using values of $f(x)$ at points that have a small, but finite, distance between them, as opposed to the definition of the derivative, that takes a limit and therefore computes the derivative using an "infinitely small" value of h . The forward-difference formula, however, is just one example of a finite difference approximation. If we replace h by $-h$ in the forward-difference formula, where h is still positive, we obtain the *backward-difference formula*

$$f'(x_0) \approx \frac{f(x_0) - f(x_0 - h)}{h}.$$

Like the forward-difference formula, the backward difference formula is first-order accurate.

If we average these two approximations, we obtain the *centered-difference formula*

$$f'(x_0) \approx \frac{f(x_0 + h) - f(x_0 - h)}{2h}.$$

To determine the accuracy of this approximation, we assume that $f'''(x)$ exists on the interval $[x_0 - h, x_0 + h]$, and then apply Taylor's Theorem again to obtain

$$f(x_0 + h) = f(x_0) + f'(x_0)h + \frac{f''(x_0)}{2}h^2 + \frac{f'''(\xi_+)}{6}h^3,$$

$$f(x_0 - h) = f(x_0) - f'(x_0)h + \frac{f''(x_0)}{2}h^2 - \frac{f'''(\xi_-)}{6}h^3,$$

where $\xi_+ \in [x_0, x_0 + h]$ and $\xi_- \in [x_0 - h, x_0]$. Subtracting the second equation from the first and solving for $f'(x_0)$ yields

$$f'(x_0) = \frac{f(x_0 + h) - f(x_0 - h)}{2h} - \frac{f'''(\xi_+) + f'''(\xi_-)}{12}h^2.$$

By the Intermediate Value Theorem, $f'''(x)$ must assume every value between $f'''(\xi_-)$ and $f'''(\xi_+)$ on the interval (ξ_-, ξ_+) , including the average of these two values. Therefore, we can simplify this equation to

$$f'(x_0) = \frac{f(x_0 + h) - f(x_0 - h)}{2h} - \frac{f'''(\xi)}{6}h^2,$$

where $\xi \in [x_0 - h, x_0 + h]$. We conclude that the centered-difference formula is *second-order accurate*. This is due to the cancellation of the terms involving $f''(x_0)$.

Example Consider the function

$$f(x) = \frac{\sin^2\left(\frac{\sqrt{x^2+x}}{\cos x-x}\right)}{\sin\left(\frac{\sqrt{x-1}}{\sqrt{x^2+1}}\right)}.$$

Our goal is to compute $f'(0.25)$. Differentiating, using the Quotient Rule and the Chain Rule, we obtain

$$f'(x) = \frac{2 \sin\left(\frac{\sqrt{x^2+x}}{\cos x-x}\right) \cos\left(\frac{\sqrt{x^2+x}}{\cos x-x}\right) \left[\frac{2x+1}{2\sqrt{x^2+1}(\cos x-x)} + \frac{\sqrt{x^2+1}(\sin x+1)}{(\cos x-x)^2} \right]}{\sin\left(\frac{\sqrt{x-1}}{\sqrt{x^2+1}}\right)} - \frac{\sin\left(\frac{\sqrt{x^2+x}}{\cos x-x}\right) \cos\left(\frac{\sqrt{x-1}}{\sqrt{x^2+1}}\right) \left[\frac{1}{2\sqrt{x}\sqrt{x^2+1}} - \frac{x(\sqrt{x-1})}{(x^2+1)^{3/2}} \right]}{\sin^2\left(\frac{\sqrt{x-1}}{\sqrt{x^2+1}}\right)}.$$

Evaluating this monstrous function at $x = 0.25$ yields $f'(0.25) = -9.066698770$.

An alternative approach is to use a *centered difference* approximation,

$$f'(x) \approx \frac{f(x+h) - f(x-h)}{2h}.$$

Using this formula with $x = 0.25$ and $h = 0.005$, we obtain the approximation

$$f'(0.25) \approx \frac{f(0.255) - f(0.245)}{0.01} = -9.067464295,$$

which has absolute error 7.7×10^{-4} . While this complicated function must be evaluated twice to obtain this approximation, that is still much less work than using differentiation rules to compute $f'(x)$, and then evaluating $f'(x)$, which is much more complicated than $f(x)$. \square

While Taylor's Theorem can be used to derive formulas with higher-order accuracy simply by evaluating $f(x)$ at more points, this process can be tedious. An alternative approach is to compute the derivative of the interpolating polynomial that fits $f(x)$ at these points. Specifically, suppose we want to compute the derivative at a point x_0 using the data

$$(x_{-j}, y_{-j}), \dots, (x_{-1}, y_{-1}), (x_0, y_0), (x_1, y_1), \dots, (x_k, y_k),$$

where j and k are known nonnegative integers, $x_{-j} < x_{-j+1} < \dots < x_{k-1} < x_k$, and $y_i = f(x_i)$ for $i = -j, \dots, k$. Then, a finite difference formula for $f'(x_0)$ can be obtained by analytically computing the derivatives of the Lagrange polynomials $\{\mathcal{L}_{n,i}(x)\}_{i=-j}^k$ for these points, where $n = j + k + 1$, and the values of these derivatives at x_0 are the proper weights for the function values y_{-j}, \dots, y_k . If $f(x)$ is $n + 1$ times continuously differentiable on $[x_{-j}, x_k]$, then we obtain an approximation of the form

$$f'(x_0) = \sum_{i=-j}^k y_i \mathcal{L}'_{n,i}(x_0) + \frac{f^{(n+1)}(\xi)}{(n+1)!} \prod_{i=-j, i \neq 0}^k (x_0 - x_i),$$

where $\xi \in [x_{-j}, x_k]$.

Among the best-known finite difference formulas that can be derived using this approach is the second-order-accurate three-point formula

$$f'(x_0) = \frac{-3f(x_0) + 4f(x_0+h) - f(x_0+2h)}{2h} + \frac{f'''(\xi)}{3}h^2, \quad \xi \in [x_0, x_0+2h],$$

which is useful when there is no information available about $f(x)$ for $x < x_0$. If there is no information available about $f(x)$ for $x > x_0$, then we can replace h by $-h$ in the above formula to obtain a second-order-accurate three-point formula that uses the values of $f(x)$ at x_0 , $x_0 - h$ and $x_0 - 2h$.

Another formula is the five-point formula

$$f'(x_0) = \frac{f(x_0-2h) - 8f(x_0-h) + 8f(x_0+h) - f(x_0+2h)}{12h} + \frac{f^{(5)}(\xi)}{30}h^4, \quad \xi \in [x_0-2h, x_0+2h],$$

which is fourth-order accurate. The reason it is called a five-point formula, even though it uses the value of $f(x)$ at four points, is that it is derived from the Lagrange polynomials for the five points $x_0 - 2h, x_0 - h, x_0, x_0 + h,$ and $x_0 + 2h$. However, $f(x_0)$ is not used in the formula because $\mathcal{L}'_{4,0}(x_0) = 0$, where $\mathcal{L}_{4,0}(x)$ is the Lagrange polynomial that is equal to one at x_0 and zero at the other four points.

If we do not have any information about $f(x)$ for $x < x_0$, then we can use the following five-point formula that actually uses the values of $f(x)$ at five points,

$$f'(x_0) = \frac{-25f(x_0) + 48f(x_0 + h) - 36f(x_0 + 2h) + 16f(x_0 + 3h) - 3f(x_0 + 4h)}{12h} + \frac{f^{(5)}(\xi)}{5}h^4,$$

where $\xi \in [x_0, x_0 + 4h]$. As before, we can replace h by $-h$ to obtain a similar formula that approximates $f'(x_0)$ using the values of $f(x)$ at $x_0, x_0 - h, x_0 - 2h, x_0 - 3h,$ and $x_0 - 4h$.

The strategy of differentiating Lagrange polynomials to approximate derivatives can be used to approximate higher-order derivatives. For example, the second derivative can be approximated using a centered difference formula,

$$f''(x_0) \approx \frac{f(x_0 + h) - 2f(x_0) + f(x_0 - h)}{h^2},$$

which is second-order accurate.

Example We will construct a formula for approximating $f'(x)$ at a given point x_0 by interpolating $f(x)$ at the points $x_0, x_0 + h,$ and $x_0 + 2h$ using a second-degree polynomial $p_2(x)$, and then approximating $f'(x_0)$ by $p_2'(x_0)$. Since $p_2(x)$ should be a good approximation of $f(x)$ near x_0 , especially when h is small, its derivative should be a good approximation to $f'(x)$ near this point.

Using Lagrange interpolation, we obtain

$$p_2(x) = f(x_0)\mathcal{L}_{2,0}(x) + f(x_0 + h)\mathcal{L}_{2,1}(x) + f(x_0 + 2h)\mathcal{L}_{2,2}(x),$$

where $\{\mathcal{L}_{2,j}(x)\}_{j=0}^2$ are the Lagrange polynomials for the points $x_0, x_1 = x_0 + h$ and $x_2 = x_0 + 2h$. Recall that these polynomials satisfy

$$\mathcal{L}_{2,j}(x_k) = \delta_{jk} = \begin{cases} 1 & \text{if } j = k \\ 0 & \text{otherwise} \end{cases}.$$

Using the formula for the Lagrange polynomials,

$$\mathcal{L}_{2,j}(x) = \prod_{i=0, i \neq j}^2 \frac{(x - x_i)}{(x_j - x_i)},$$

we obtain

$$\mathcal{L}_{2,0}(x) = \frac{(x - (x_0 + h))(x - (x_0 + 2h))}{(x_0 - (x_0 + h))(x_0 - (x_0 + 2h))}$$

$$\begin{aligned}
&= \frac{x^2 - (2x_0 + 3h)x + (x_0 + h)(x_0 + 2h)}{2h^2}, \\
\mathcal{L}_{2,1}(x) &= \frac{(x - x_0)(x - (x_0 + 2h))}{(x_0 + h - x_0)(x_0 + h - (x_0 + 2h))} \\
&= \frac{x^2 - (2x_0 + 2h)x + x_0(x_0 + 2h)}{-h^2}, \\
\mathcal{L}_{2,2}(x) &= \frac{(x - x_0)(x - (x_0 + h))}{(x_0 + 2h - x_0)(x_0 + 2h - (x_0 + h))} \\
&= \frac{x^2 - (2x_0 + h)x + x_0(x_0 + h)}{2h^2}.
\end{aligned}$$

It follows that

$$\begin{aligned}
\mathcal{L}'_{2,0}(x) &= \frac{2x - (2x_0 + 3h)}{2h^2} \\
\mathcal{L}'_{2,1}(x) &= -\frac{2x - (2x_0 + 2h)}{h^2} \\
\mathcal{L}'_{2,2}(x) &= \frac{2x - (2x_0 + h)}{2h^2}
\end{aligned}$$

We conclude that $f'(x_0) \approx p'_2(x_0)$, where

$$\begin{aligned}
p'_2(x_0) &= f(x_0)\mathcal{L}'_{2,0}(x_0) + f(x_0 + h)\mathcal{L}'_{2,1}(x_0) + f(x_0 + 2h)\mathcal{L}'_{2,2}(x_0) \\
&\approx f(x_0)\frac{-3}{2h} + f(x_0 + h)\frac{2}{h} + f(x_0 + 2h)\frac{-1}{2h} \\
&\approx \frac{3f(x_0) + 4f(x_0 + h) - f(x_0 + 2h)}{2h}.
\end{aligned}$$

Using Taylor's Theorem to write $f(x_0 + h)$ and $f(x_0 + 2h)$ in terms of Taylor polynomials centered at x_0 , it can be shown that the error in this approximation is $O(h^2)$, and that this formula is exact when $f(x)$ is a polynomial of degree 2 or less. \square

In a practical implementation of finite difference formulas, it is essential to note that roundoff error in evaluating $f(x)$ is bounded independently of the spacing h between points at which $f(x)$ is evaluated. It follows that the roundoff error in the approximation of $f'(x)$ actually *increases* as h decreases, because the errors incurred by evaluating $f(x)$ are divided by h . Therefore, one must choose h sufficiently small so that the finite difference formula can produce an accurate approximation, and sufficiently large so that this approximation is not too contaminated by roundoff error.

Automatic Differentiation

If the function $f(x)$ can be represented as a sequence of computations, each consisting of the evaluation of some differentiable function with differentiable arguments, then a similar sequence of

computations representing $f'(x)$ can be obtained using *automatic differentiation*.

Suppose that the evaluation of $f(x)$ consists of the evaluation of simpler differentiable functions t_1, \dots, t_n , where t_1 only depends on x , and t_j , for $j > 1$, depends on x and t_1, \dots, t_{j-1} , with $t_n(x, t_1, \dots, t_{n-1})$ representing the desired value $f(x)$. Automatic differentiation proceeds by computing the derivatives of each function t_j , $j = 1, \dots, n$, to obtain a new sequence of functions t'_j , $j = 1, \dots, n$, where t'_j depends on t_1, \dots, t_j and t'_1, \dots, t'_{j-1} , and $f'(x) = t'_n(x, t_1, \dots, t_n, t'_1, \dots, t'_n)$.

This process, while conceptually simple, can be very difficult to implement since much optimization of the generated derivative is necessary. Fortunately, techniques used in compiler design are very useful in building a practical automatic differentiation procedure. An additional benefit of automatic differentiation is that its analysis of $f(x)$ can be used to estimate its sensitivity to perturbations.