PHOENICIA UNIVERSITY

College of Arts and Sciences

Numerical Differentiation

Chapter 4 — Approximating Derivatives from Function Values

Math 213 – Numerical Methods

1 Lesson Roadmap

This chapter develops methods for estimating the derivative of a function when only sampled values (not a formula) are available, or when differentiating directly is impractical.

1 Forward & backward difference formulas

The simplest first-order approximations, $O(h)$

2 Central difference formulas

More accurate, second-order approximations, $O(h^2)$

3 Higher-order formulas

Three- and five-point stencils for greater accuracy

4 Error analysis

Balancing truncation error against round-off error

5 Optimal step size

Choosing h to minimize total error

6 Richardson extrapolation

Combining estimates to cancel leading error terms

2 Why Numerical Differentiation?

The definition of a derivative from calculus is a limit:

$$f'(x) = \lim_{h \to 0} \; [\, f(x+h) - f(x) \,] \,/\, h$$

In practice, we often cannot take this limit directly: the function may only be known through a table of measured data, or its formula may be too complex to differentiate by hand. Numerical differentiation replaces the limit with a small, finite step size h, giving an approximation instead of an exact value.

Tabulated data

Sensor readings, experimental measurements, or survey data given only at discrete points.

Costly evaluations

The function comes from a simulation that is expensive to evaluate at many points.

Building block

Derivative estimates feed into root-finding, optimization, and differential-equation solvers.

3 Forward & Backward Difference Formulas

Both formulas come from a first-order Taylor expansion of f around x, and both use only two function values.

Forward Difference

$$f'(x) \approx [\, f(x+h) - f(x) \,] \,/\, h$$

Error: $O(h)$

Uses the current point and one step ahead. Simple to apply, but the leading error term is proportional to h, so accuracy improves slowly as h shrinks.

Backward Difference

$$f'(x) \approx [\, f(x) - f(x-h) \,] \,/\, h$$

Error: $O(h)$

Uses the current point and one step behind. Useful when a value ahead of x is unavailable — for example, at the last point of a data table.

4 Central Difference Formulas

Averaging the forward and backward Taylor expansions cancels the h-term entirely, leaving a formula that is far more accurate for the same step size.

First Derivative

$$f'(x) \approx [\ f(x+h) - f(x-h)\ ] \,/\, (2h)$$

Error: $O(h^2)$

Second Derivative

$$f''(x) \approx [\ f(x+h) - 2f(x) + f(x-h)\ ] \,/\, h^2$$

Error: $O(h^2)$

Because the error shrinks with h^2 rather than h, halving the step size roughly quarters the error — central formulas are the default choice whenever both neighboring points are available.

5 Higher-Order Formulas

Combining more Taylor expansions cancels additional error terms, producing formulas that stay accurate even with a larger h.

Three-Point Formulas — $O(h^2)$

$$f'(x) \approx [-3f(x) + 4f(x+h) - f(x+2h)] / (2h)$$

Forward-leaning, $O(h^2)$

$$f'(x) \approx [3f(x) - 4f(x-h) + f(x-2h)] / (2h)$$

Backward-leaning, $O(h^2)$

Five-Point Central Formula — $O(h^4)$

$$f'(x) \approx [-f(x+2h) + 8f(x+h) - 8f(x-h) + f(x-2h)] / (12h)$$

Error: $O(h^4)$

The three-point forward and backward formulas are useful near the edges of a data table, where a full central stencil is not available. The five-point central formula gives very high accuracy when four neighboring values exist.

6 Worked Example: Comparing Formulas

Let $f(x) = e^x$. Approximate $f'(1)$ using the central difference formula $O(h^2)$, for decreasing step sizes h. The exact value is $f'(1) = e \approx 2.718281828$.

h	Approximation	Absolute Error
0.1	2.72281456	4.53×10^{-3}
0.01	2.71832713	4.53×10^{-5}
0.001	2.71828228	4.53×10^{-7}
0.0001	2.71828183	4.53×10^{-9}

As h shrinks by a factor of 10, the error shrinks by a factor of about 100 — confirming the $O(h^2)$ behavior of the central difference formula.

7 Understanding the Error: Two Sources

Every numerical derivative carries two competing sources of error, and they move in opposite directions as h changes.

Truncation Error

Comes from stopping the Taylor series after a few terms. It shrinks as h gets smaller — for a central formula, it is proportional to h^2 times a bound on a higher derivative of f.

Round-off Error

Comes from the limited precision used to store and evaluate f(x). Because it is divided by h in the formula, it grows as h gets smaller — tiny steps amplify tiny measurement or rounding noise.

Total error(h) ≈ (round-off term, grows as h ↓) + (truncation term, grows as h ↑) — so the smallest h is not always the best choice.

8 Finding the Optimal Step Size

For the central difference formula $f'(x) \approx [f(x+h) - f(x-h)] / (2h)$, the total error can be bounded by a sum of a round-off term and a truncation term:

$$\text{Error}(h) \le \varepsilon / h + (h^2 / 6)\, M$$

ε = round-off bound on f, M = bound on $|f'''|$

Treating Error(h) as a function of h and setting its derivative to zero gives the step size that balances the two effects:

$$h\_opt = (3\varepsilon / M)\text{^}(1/3)$$

This is why decreasing h indefinitely does not keep improving accuracy: past a certain point, round-off error — not the formula itself — dominates the result. The worked example on the earlier slide shows this pattern begin to emerge at very small h.

9 Richardson Extrapolation

Richardson extrapolation combines two lower-order estimates to cancel the leading error term, producing a higher-order result without deriving a new formula from scratch.

If $D(h)$ is a central-difference estimate whose error is dominated by a term in h^2, then $D(h)$ and $D(h/2)$ can be combined as follows:

$$D\_better = [\, 4 \cdot D(h/2) - D(h) \,] / 3$$

New error: $O(h^4)$

Why it works

$D(h) = D + c_1h^2 + c_2h^4 + \ldots$ and $D(h/2) = D + c_1h^2/4 + c_2h^4/16 + \ldots$

Multiplying $D(h/2)$ by 4 and subtracting $D(h)$ eliminates the h^2 term exactly, leaving an error that starts at h^4.

The same idea extends to combine two $O(h^4)$ estimates into an $O(h^6)$ estimate, and so on — this general pattern is called Richardson's method.

10 Worked Example: Richardson Extrapolation

Continuing with $f(x) = e^{xx}$ at $x = 1$ (true value $f'(1) = e \approx 2.718281828$), apply Richardson extrapolation to the central-difference estimates at $h = 0.1$ and $h = 0.05$.

Estimate	Value	Absolute Error
$D(0.1)$ — $O(h^2)$	2.72281456	4.53×10^{-3}
$D(0.05)$ — $O(h^2)$	2.71941459	1.13×10^{-3}
Richardson — $O(h^4)$	**2.71828126**	**5.66×10^{-7}**

Combining two modest $O(h^2)$ estimates produced a result almost four orders of magnitude more accurate than either one alone — without evaluating f at any new points beyond $x \pm h$ and $x \pm h/2$.

Key Takeaways

- Forward and backward differences are simple, two-point formulas with error $O(h)$.
- Central differences use symmetry to reach $O(h^2)$ accuracy with the same two neighboring points.
- Higher-order stencils (three-point edge formulas, five-point central formula) trade more function evaluations for greater accuracy.
- Total error balances truncation error (grows with h) against round-off error (grows as h shrinks) — there is an optimal h, not an infinitely small one.
- Richardson extrapolation combines two lower-order estimates to cancel the leading error term, boosting accuracy without a new derivation.