

Lecture 12 - Calculating Derivatives

Dinesh Krishnamoorthy

Department of Engineering Cybernetics, NTNU

Spring 2026

Outline

1. Finite Difference
2. Automatic Differentiation

Calculating derivatives

Search direction requires calculating gradients and Hessians

- ▶ Gradient descent: $s_k = -\nabla f(x_k)$
- ▶ Newton method: $s_k = -H^{-1}(x_k)\nabla f(x_k)$

Calculating derivatives

- ▶ Analytically (by hand)
- ▶ Symbolically (e.g., Mathematica, Maple)
- ▶ Automatic differentiation (e.g., CasADi¹)
- ▶ Finite differencing (e.g., MATLAB `fmincon()`)

¹<https://web.casadi.org/>

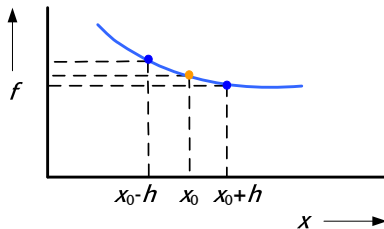
Finite differencing

- Taylor forward

$$f(x_0 + h) = f(x_0) + hf'(x_0) + \frac{1}{2}h^2f''(x_0) + O(h^3)$$

- Forward finite difference

$$f'(x) = \frac{f(x_0+h)-f(x)}{h} + O(h)$$



- Taylor backward

$$f(x_0 - h) = f(x_0) - hf'(x_0) + \frac{1}{2}h^2f''(x_0) + O(h^3)$$

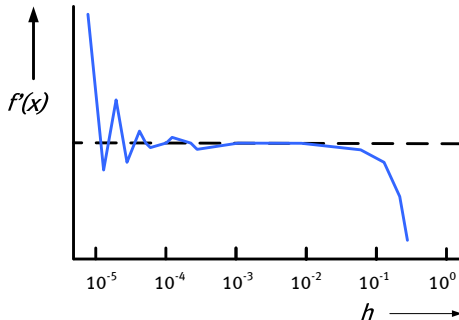
- Central difference

$$f'(x) = \frac{f(x_0+h)-f(x_0-h)}{2h} + O(h^2)$$

Finite differencing

Computational error

- ▶ Truncation error (neglecting higher order terms)
- ▶ Cancellation error (depends on # significant digits)



Automatic differentiation

Key idea

- ▶ Any function can be evaluated by a sequence of elementary operations involving one or two arguments for which the derivatives are already known.

$$a + b, \quad a * b, \quad a/b, \quad a^b$$

$$\log(a), \quad \sin(a), \quad e^a$$

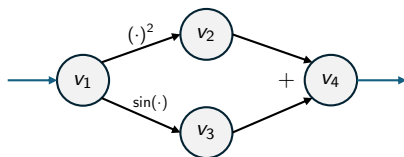
- ▶ systematically apply the chain rule to these elementary steps to calculate the exact numerical derivative of the function.

Example

$$f(x) = x^2 + \sin(x)$$

Decompose into elementary operations

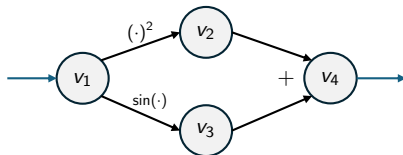
- ▶ $v_1 = x$
- ▶ $v_2 = v_1^2$
- ▶ $v_3 = \sin(v_1)$
- ▶ $v_4 = v_2 + v_3$



Forward mode AD

Propagate derivatives forward using chain rule to compute $\dot{v}_i = \frac{\partial v_i}{\partial x}$

- ▶ $\dot{v}_1 = 1$
- ▶ $\dot{v}_2 = 2v_1 \dot{v}_1 = 2v_1 = 2x$
- ▶ $\dot{v}_3 = \dot{v}_1 \cos(v_1) = \cos(x)$
- ▶ $\dot{v}_4 = \dot{v}_2 + \dot{v}_3 = 2x + \cos(x)$



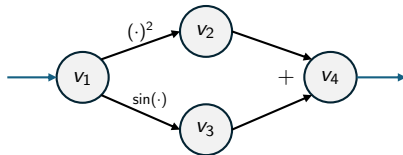
One forward sweep gives us the derivative $\nabla f(x) = 2x + \cos(x)$.

Reverse mode AD

Propagate sensitivities (a.k.a. adjoints) $\bar{v}_i = \frac{\partial f}{\partial v_i}$ backwards, with the chain rule,

$$\frac{\partial f}{\partial v_i} = \sum_{j \text{ child of } i} \frac{\partial f}{\partial v_j} \frac{\partial v_j}{\partial v_i} \Rightarrow \bar{v}_i = \sum_{j \text{ child of } i} \bar{v}_j \frac{\partial v_j}{\partial v_i}$$

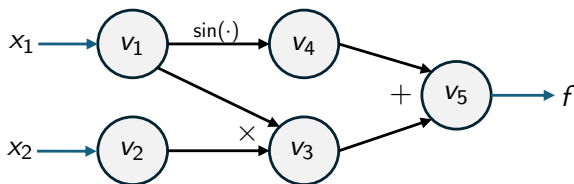
- ▶ $\bar{v}_4 = \frac{\partial f}{\partial v_4} = 1$
- ▶ $\bar{v}_3 = \frac{\partial f}{\partial v_3} = \frac{\partial f}{\partial v_4} \frac{\partial v_4}{\partial v_3} = \bar{v}_4 \frac{\partial (v_2 + v_3)}{\partial v_3} = 1$
- ▶ $\bar{v}_2 = \frac{\partial f}{\partial v_2} = \frac{\partial f}{\partial v_4} \frac{\partial v_4}{\partial v_2} = \bar{v}_4 \frac{\partial (v_2 + v_3)}{\partial v_2} = 1$
- ▶ $\bar{v}_1 = \frac{\partial f}{\partial v_1} = \frac{\partial f}{\partial v_2} \frac{\partial v_2}{\partial v_1} + \frac{\partial f}{\partial v_3} \frac{\partial v_3}{\partial v_1} = 2\bar{v}_2 v_1 + \bar{v}_3 \cos(v_1) = 2x + \cos(x)$



One backward sweep gives us the derivative $\nabla f(x) = 2x + \cos(x)$.

Example 2

$$f(x) = x_1 x_2 + \sin(x_1)$$



See Handwritten notes for solution!

Automatic differentiation

For a general function with $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$, obtaining its derivative $\nabla f(x)$ requires

- ▶ n forward passes using the forward mode AD
- ▶ m backward passes using the reverse mode AD