Homework 3

Math 76 Summer 2026

Automatic Differentiation with Dual Numbers

Name: ______________________________

Throughout, the *dual numbers* are $\mathbb{D} = \{\, a + b\varepsilon : a, b \in \mathbb{R} \,\}$, where $\varepsilon \neq 0$ but $\varepsilon^2 = 0$; addition and multiplication are those of polynomials in ε, reduced using $\varepsilon^2 = 0$. The single fact driving everything below is that evaluating a differentiable function on a dual number returns its value and derivative together:

$$f(x+\varepsilon) = f(x) + f'(x)\,\varepsilon, \qquad \text{and more generally} \qquad f(a+b\varepsilon) = f(a) + b\,f'(a)\,\varepsilon.$$

Problem 1. Fluency with the infinitesimal.

For each function, evaluate on $x+\varepsilon$, expand using $\varepsilon^2 = 0$, and read the derivative off the dual part. Then verify with ordinary differentiation.

(a) $f(x) = x^4 - 3x^2 + 7$.

(b) $f(x) = \dfrac{x-1}{x+1}$. (Find the reciprocal $\big((x+1)+\varepsilon\big)^{-1} = c+d\varepsilon$ by solving $\big((x+1)+\varepsilon\big)(c+d\varepsilon) = 1$, then multiply by $(x-1)+\varepsilon$.)

Problem 2. Inverses in $\mathbb{D}$.

Prove that $a + b\varepsilon$ has a multiplicative inverse in $\mathbb{D}$ *if and only if* $a \neq 0$, and when $a \neq 0$ derive the formula

$$(a + b\varepsilon)^{-1} = \frac{1}{a} - \frac{b}{a^2}\,\varepsilon.$$

(Equivalently, the non-invertible elements are exactly the multiples $b\varepsilon$ of ε — the sense in which $\mathbb{D}$ differs from a field, though you need not discuss that here.)

Problem 3. Deriving the differentiation rules algebraically.

(a) Use the binomial theorem to show directly that $(x + \varepsilon)^n = x^n + n\,x^{n-1}\varepsilon$ for every integer $n \geq 1$, recovering the power rule.

(b) Derive the quotient rule for dual numbers. For $u = f + f'\varepsilon$ and $v = g + g'\varepsilon$ with $g \neq 0$, first compute v^{-1} explicitly, then multiply by u and collect the ε-term to obtain the dual part $\big(f/g\big)' = (f'g - fg')/g^2$.

Problem 4. The chain rule, step by step.
Let $f(x) = \exp\big(\sin(x^2)\big)$, with intermediates $r = x^2$, $s = \sin r$, $f = e^s$.

(a) Seed the input as $x + \varepsilon$ and push it through the three stages in turn, using the dual rules for squaring, sin, and exp. Read $f'(x)$ off the dual part.

(b) Confirm your answer with the ordinary chain rule.

(c) Evaluate $f'(0)$ and $f'(\sqrt{\pi})$.

Problem 5. Reverse-mode AD: differentiating a graph backward.

Take the function

$$f(x,y) = \sin(xy) + x^2, \qquad p = xy, \quad q = \sin p, \quad r = x^2, \quad f = q + r,$$

so the input x is used twice: it feeds both p and r. Reverse-mode differentiation computes, for *every* variable v in this list, the sensitivity $\partial f/\partial v$ of the output to v, via the chain-rule recurrence

$$\frac{\partial f}{\partial v} = \sum_{w\colon v \to w} \frac{\partial f}{\partial w}\,\frac{\partial w}{\partial v}, \qquad \frac{\partial f}{\partial f} = 1,$$

where the sum runs over every variable w computed directly from v, and each local factor $\partial w/\partial v$ is obtained by evaluating that *single* elementary operation on a dual number and reading off the dual part — for instance $\sin(p+\varepsilon) = \sin p + \cos p\,\varepsilon$ gives $\partial q/\partial p = \cos p$.

(a) **Forward sweep.** At a general point $(x,y) = (a,b)$, evaluate and store p, q, r, f, and record each edge's local derivative $\partial w/\partial v$ from a dual evaluation.

(b) **Backward sweep.** Starting from $\partial f/\partial f = 1$ and moving from the output down to the inputs, compute in order $\partial f/\partial q$, $\partial f/\partial r$, $\partial f/\partial p$, $\partial f/\partial y$, and finally $\partial f/\partial x$. Show explicitly that $\partial f/\partial x$ picks up one term from each of the two variables computed directly from x (namely p and r).

(c) Read off the gradient ∇f and verify it by differentiating f directly.

Problem 6. Beyond first derivatives *(optional)*.

Because $\varepsilon^2 = 0$, evaluating $f(x+\varepsilon)$ yields only f'. Introduce instead a new infinitesimal δ with $\delta^3 = 0$ but $\delta^2 \neq 0$.

(a) Show that $f(x+\delta) = f(x) + f'(x)\,\delta + \frac{1}{2}f''(x)\,\delta^2$, so the δ^2 coefficient delivers $\frac{1}{2}f''(x)$.

(b) Verify this on $f(x) = x^3$.